POINCARÉ-TRESCHEV MECHANISM IN MULTI-SCALE, NEARLY INTEGRABLE HAMILTONIAN SYSTEMS

LU XU, YONG LI, AND YINGFEI YI

ABSTRACT. This paper is a continuation to our work [20] concerning the persistence of lower dimensional tori on resonant surfaces of a multi-scale, nearly integrable Hamiltonian system. This type of systems, being properly degenerate, arise naturally in planar and spatial lunar problems of celestial mechanics for which the persistence problem ties closely to the stability of the systems. For such a system, under certain non-degenerate conditions of Rüssmann type, the majority persistence of non-resonant tori and the existence of a nearly full measure set of Poincaré non-degenerate, lower dimensional, quasi-periodic invariant tori on a resonant surface corresponding to the highest order of scale is proved in [6] and [20], respectively. In this work, we consider a resonant surface corresponding to any intermediate order of scale and show the existence of a nearly full measure set of Poincaré non-degenerate, lower dimensional, quasi-periodic invariant tori on the resonant surface. The proof is based on a normal form reduction which consists of a finite step of KAM iterations in pushing the non-integrable perturbation to a sufficiently high order and the splitting of resonant tori on the resonant surface according to the Poincaré-Treshchev mechanism.

1. INTRODUCTION

With respect to the symplectic structure $dx \wedge dy$ on $\mathbb{T}^n \times \mathbb{R}^n$, we consider a multi-scale, real analytic, nearly integrable Hamiltonian system of the form

$$H(x, y, \varepsilon) = H_0(y^{n_0}) + \varepsilon^{\bar{m}_1} H_1(y^{n_1}) + \cdots + \varepsilon^{\bar{m}_\alpha} H_\alpha(y^{n_\alpha}) + \varepsilon^{\bar{m}_{\alpha+1}} P(x, y, \varepsilon), \tag{1.1}$$

where $x = (x_1, \cdots, x_n)^\top \in \mathbb{T}^n$, $y = (y_1, \cdots, y_n)^\top \in G$ with $G \subset \mathbb{R}^n$ being a bounded closed region, $\varepsilon > 0$ is a small parameter, $\alpha,\ n_i,\ \bar{m}_j,\ i = 0, 1, \cdots, \alpha,\ j = 1, 2, \cdots, \alpha + 1$, are positive integers such that $n_0 \le \cdots \le n_\alpha := n,\ \bar{m}_1 < \bar{m}_2 < \cdots < \bar{m}_\alpha < \bar{m}_{\alpha+1},\ y^{n_i} = (y_1, \cdots, y_{n_i})^\top$, $i = 0, 1, \cdots, \alpha$, and the perturbation P depends on ε smoothly. We note that $y^{n_\alpha} = y$.

Multi-scale, nearly integrable Hamiltonian systems of the form (1.1) are rooted in many problems of celestial mechanics, for instance, the restricted three-body problem coupling two massive bodies with a body of very small mass. The multiple scales are due to the significant differences in masses and distances between the bodies. Averaging and normalization techniques lead to a Hamiltonian of the form (1.1) which is nearly integrable but admits properly degeneracy in the integrable part (see e.g. [2, 13, 14, 15, 18]). Indeed, all terms H_i, $i = 0, 1, \cdots, \alpha - 1$, in (1.1) only depend on parts of the action variables.

2000 *Mathematics Subject Classification.* Primary 37J40, 70H08.

Key words and phrases. Multi-scale Hamiltonian systems, high order proper degeneracy, resonant tori, lower dimensional tori, KAM theory.

The first author was partially supported by NSFC grant 11401251 and a visiting scholar programm from Jilin University and CSC. The second author was partially supported by National Basic Research Program of China grant 2013CB834100 and NSFC grant 11171132. The third author was partially supported by NSERC discovery grant 1257749, a faculty development grant from the University of Alberta, and a Scholarship from Jilin University.

The majority persistence of quasi-periodic invariant tori in a multi-scale, nearly integrable system like (1.1) is an important problem concerning the (metric) stability of the system. The two-scale case was first studied by Arnold in [1] under a degeneracy-removing condition that $H_0 + \varepsilon H_1$ satisfies either Kolmogorov or iso-energetic non-degenerate condition. The case of general scales was treated in a recent work of Han, Li, and Yi [6] under the following degeneracy-removing condition of Bruno-Rüssman type:

A*) There is a positive integer N such that

$$\mathrm{Rank}\{\partial_y^l \Omega^*(y): \quad 0 \le |l| \le N\} = n, \ \forall\, y \in G,$$

where

$$\Omega^*(y) =: (\nabla_{\hat{y}^{n_0}} H_0(y^{n_0}), \nabla_{\hat{y}^{n_1}} H_1(y^{n_1}), \cdots, \nabla_{\hat{y}^{n_\alpha}} H_\alpha(y^{n_\alpha}))^\top, \tag{1.2}$$

$\hat{y}^{n_0} = y^{n_0}$, $\hat{y}^{n_i} = (y_{n_{i-1}+1}, \cdots, y_{n_i})^\top$, $\nabla_{\hat{y}^{n_i}}$ denotes the gradient with respect to $\hat{y}^{n_i}$, for each $i = 1, 2, \cdots, \alpha$ respectively, and the matrix $\{\partial_y^l \Omega^*(y): \quad 0 \le |l| \le N\}$ consists of coefficients of simultaneous Taylor expansions of components of $\Omega^*(\cdot)$ at y up to order N.

We refer the reader to [15] and references therein for applications of the main result in [6] to the stability problem of certain spatial three body problems.

Like in the case of standard nearly integrable Hamiltonian systems, an important approach in studying the existence of quasi-periodic motions in the resonant zone of a multi-scale, nearly integrable Hamiltonian system is to show the persistence of lower dimensional, quasi-periodic invariant tori split from resonant ones according to the Poincaré-Treshchev mechanism (see [9, 10, 19]). More precisely, consider the integrable part of (1.1):

$$N_\varepsilon(y) = H_0(y^{n_0}) + \varepsilon^{\bar{m}_1} H_1(y^{n_1}) + \cdots + \varepsilon^{\bar{m}_\alpha} H_\alpha(y^{n_\alpha}),$$

and set

$$\omega_\varepsilon^*(y) =: \nabla N_\varepsilon(y) = (\omega_\varepsilon^{*,n_0}(y), \varepsilon^{\bar{m}_1} \hat{\omega}_\varepsilon^{*,n_1}(y), \cdots, \varepsilon^{\bar{m}_\alpha}, \hat{\omega}_\varepsilon^{*,n_\alpha}(y))^\top. \tag{1.3}$$

To work with a fixed resonant type, we let g be a subgroup of $\mathbb{Z}^n$, called *resonant group*, and consider the *g-resonant surface*

$$\mathcal{O}_0(g, G) = \{y \in G: \ \langle \hat{k}^{n_j}, \ \nabla_{\hat{y}^{n_j}} H_i(y) \rangle = 0, \ 0 \le j \le i \le \alpha, \ k \in g\}, \tag{1.4}$$

where for each $k = (k_1, \cdots, k_n)^\top \in g$, $\hat{k}^{n_0} = (k_1, \cdots, k_{n_0})^\top$ and $\hat{k}^{n_j} = (k_{n_{j-1}+1}, \cdots, k_{n_j})^\top$, $j = 1, \cdots, \alpha$. We note that if $y \in \mathcal{O}_0(g, G)$, then $\langle k, \omega_\varepsilon^*(y) \rangle = 0$, $k \in g$, and thus, for any $\varepsilon > 0$, $\omega_\varepsilon^*(y)$ is a resonant frequency vector of resonant type characterized by g.

In [20], we have treated the case that resonance occurs at the highest ε-order term H_α of the integrable part. Under a non-degenerate condition on $\mathcal{O}_0(g, G)$ resembling **A*****)**, we showed the majority persistence of Poincaré non-degenerate sub-tori on the resonant surface $\mathcal{O}(g, G)$, where $g = \{0\} \oplus \hat{g}^\alpha$ with $\hat{g}^\alpha$ being a subgroup of $\mathbb{Z}^{n_\alpha - n_{\alpha-1}}$.

In this work, we pay attention to the more general case that resonance occurs among some lower ε-order terms $H_{I+1}, \cdots, H_\alpha$, where $0 \le I < \alpha$ is a fixed integer such that

A1) $n_I < n_{I+1}$ in (1.1).

To characterize resonances among the last $(n_\alpha - n_I)$-components of the frequency map $\omega_\varepsilon^*(y)$, we restrict the resonant group g to the form

$$g = \{0\} \oplus \hat{g}^{I+1} \oplus \cdots \oplus \hat{g}^\alpha,$$

where for each $j = 1, \cdots, \alpha - I$, $\hat{g}^{I+j} \subset \mathbb{Z}^{n_{I+j}-n_{I+j-1}}$ is a subgroup of $\mathbb{Z}^{n_{I+j}-n_{I+j-1}}$. Then each $k \in g$ can be expressed as $k = (0, \hat{k}^{n_{I+1}}, \cdots, \hat{k}^{n_\alpha})^\top$, where $\hat{k}^{n_{I+j}} \in g^{I+j}$, $j = 1, \cdots, \alpha - I$, and the g-resonant surface $\mathcal{O}_0(g, G)$ becomes

$$\mathcal{O}(g, G) = \{y \in G : \langle \hat{k}^{n_{I+j}}, \nabla_{\hat{y}^{n_{I+j}}} H_i(y)\rangle = 0,\ \forall \hat{k}^{n_{I+j}} \in \hat{g}^{I+j},\ I < I + j \le i \le \alpha\}.$$

If there exists $1 \le j_0 < \alpha - 1$ such that $n_{I+j_0} = n_{I+j_0+1}$, then we take $\hat{g}^{I+j_0+1} = \emptyset$ and

$$g = \{0\} \oplus \hat{g}^{I+1} \oplus \cdots \oplus \hat{g}^{I+j_0} \oplus \hat{g}^{I+j_0+2} \oplus \cdots \oplus \hat{g}^{\alpha}.$$

In this case, the g-resonant surface becomes

$$\begin{aligned} \mathcal{O}_1(g, G) \quad = \quad & \{y \in G : \langle \hat{k}^{n_{I+j}}, \nabla_{\hat{y}^{n_{I+j}}} H_i(y)\rangle = 0, \\ & \forall \hat{k}^{n_{I+j}} \in \hat{g}^{I+j},\ i \ge I + j,\ 1 \le j \le \alpha - I,\ j \ne j_0 + 1\}. \end{aligned}$$

Let $\hat{K}_2^{I+1} = (\hat{\tau}^{n_I+1}, \cdots, \hat{\tau}^{n_I+d_1})$ be an $(n_{I+1} - n_I) \times d_1$ integral matrix whose columns form the basis of g^{I+1}, and let $\hat{K}_1^{I+1} = (\hat{\iota}^{n_I+1}, \cdots, \hat{\iota}^{n_I+m_1})$ be an $(n_{I+1} - n_I) \times m_1$ integral matrix such that $\det(\hat{K}_1^{I+1}, \hat{K}_2^{I+1}) = 1$, where $m_1 + d_1 = n_{I+1} - n_I$. For each $j = 2,\ 3, \cdots, \alpha - I$, also let $\hat{K}_2^{I+j} = (\hat{\tau}^{n_{I+j-1}+1} \ \cdots \ \hat{\tau}^{n_{I+j}})$ be an $(n_{I+j} - n_{I+j-1}) \times (n_{I+j} - n_{I+j-1})$ integral matrix whose columns consist of a basis of $\hat{g}^{I+j}$. Clearly, $\hat{K}_2^{I+j}$ is of the full rank $n_{I+j} - n_{I+j-1}$.

We make the following assumptions:

A2) In the case $n_{I+1} < n_{I+2}$, H_{I+1} is $\hat{g}^{I+1}$-non-degenerate on $\mathcal{O}(g, G)$, i.e., $\forall y \in \mathcal{O}(g, G)$,

$$\det\ (\hat{K}_2^{I+1})^\top \frac{\partial^2 H_{I+1}}{\partial(\hat{y}^{n_{I+1}})^2}(y) \hat{K}_2^{I+1} \ne 0,$$

$$\det \frac{\partial^2 H_{I+j}}{\partial(\hat{y}^{n_{I+j}})^2}(y) \ne 0, \qquad j = 2, \cdots, \alpha - I.$$

In the case $n_{I+1} = n_{I+2} < n_{I+3}$, either H_{I+1} or H_{I+2} is $\hat{g}^{I+1}$-non-degenerate on $\mathcal{O}(g, G)$ and for any $y \in \mathcal{O}(g, G)$,

$$\det \frac{\partial^2 H_{I+j}}{\partial(\hat{y}^{n_{I+j}})^2}(y) \ne 0, \tag{1.5}$$

$j = 3, \cdots, \alpha - I$.

The consideration of the second case above is motivated by the example in Section 5.

Remark 1.1. For each $j = 2, \cdots, \alpha - I$, if $\hat{g}^{I+j}$ is of the full rank $n_{I+j} - n_{I+j-1}$, then the condition (1.5) implies that

$$\det\ (\hat{K}_2^{I+j})^\top \frac{\partial^2 H_{I+j}}{\partial(\hat{y}^{n_{I+j}})^2}(y) \hat{K}_2^{I+j} \ne 0, \qquad \forall y \in \mathcal{O}(g, G),$$

i.e., each H_{I+j} is $\hat{g}^{I+j}$-non-degenerate.

Denote

$$\hat{K}_2 = \begin{pmatrix} \hat{K}_2^{I+1} & 0 & \cdots & 0 \\ 0 & \hat{K}_2^{I+2} & \cdots & 0 \\ \cdots & & \cdots & \\ 0 & 0 & \cdots & \hat{K}_2^{\alpha} \end{pmatrix}_{(n-n_I)\times d}, \quad \hat{K}_1 = \begin{pmatrix} \hat{K}_1^{I+1} \\ 0 \\ \vdots \\ 0 \end{pmatrix}_{(n-n_I)\times m_1},$$

and

$$K_1 = \begin{pmatrix} I & O \\ O & \hat{K}_1 \end{pmatrix}_{n\times m}, \ K_2 = \begin{pmatrix} O \\ \hat{K}_2 \end{pmatrix}_{n\times d}, \ K_0 = (K_1, K_2),$$

where $d =: n - n_{I+1} + d_1,\ m =: n_I + m_1$. Then the g-resonant surface can be expressed as

$$\begin{aligned}\mathcal{O}(g,G) &= \{y \in G:\ \langle \hat{k}^{n_{I+j}}, \nabla_{\hat{g}^{n_{I+j}}} H_i(y)\rangle = 0,\ \forall \hat{k}^{n_{I+j}} \in \hat{g}^{I+j},\ I < I+j \le i \le \alpha\} \\ &= \{y \in G:\ K_2^\top \Omega_*(y)\rangle = 0\}.\end{aligned}$$

We note that under conditions **A1)** and **A2)**, the map $K_2^\top \Omega : G \to \mathbb{R}^d$ is of maximal rank, and consequently, $\mathcal{O}(g,G)$, as the kernel of this map, is a m-dimensional, real analytic sub-manifold of G.

We also assume the following non-degenerate condition of Bruno-Rüssman type:

A3) There is a positive integer N such that

$$\mathrm{Rank}\{\partial_y^l K_1^\top \Omega^*(y);\quad 0 \le |l| \le N\} = m,\ \forall\ y \in \mathcal{O}(g,G).$$

The subgroup g induces a symplectic transformation which uniquely determines the following splitting of resonant tori T_y^ε:

$$y \to y,\ x \to \begin{pmatrix}\varphi \\ \psi\end{pmatrix} \in \mathbb{T}^m \times \mathbb{T}^d,$$

where

$$\begin{pmatrix}\varphi \\ \psi\end{pmatrix} = K_0^\top x.$$

Under the new coordinate, the unperturbed motion of (1.1) becomes

$$\begin{cases} \dot{\varphi} &= \omega_\varepsilon(y), \\ \dot{\psi} &= 0, \\ \dot{y} &= 0, \end{cases}$$

where

$$\omega_\varepsilon(y) = K_1^\top \omega_\varepsilon^*(y) = (\omega_\varepsilon^0(y), \cdots, \varepsilon^{\bar{m}_I}\hat{\omega}_\varepsilon^I(y), \varepsilon^{\bar{m}_{I+1}}\hat{\omega}_\varepsilon^{I+1}(y))^\top. \tag{1.6}$$

It follows that for each $y \in \mathcal{O}(g,G)$, the resonant torus T_y^ε is foliated into invariant m-tori

$$T_y^\varepsilon(\psi) = \mathbb{T}^m \times \{\psi\} \times \{y\}, \qquad \psi \in \mathbb{T}^d$$

with linear flows $\{\varphi_0 + \omega_\varepsilon(y)t\} \times \{\psi\} \times \{y\}$.

With respect to the multi-scale Hamiltonian (1.1), we now introduce degeneracy-removing conditions of Poincaré-Treshchev type, similarly to the case of standard nearly integrable Hamiltonian systems ([10, 19]). Consider $h_0 : \mathbb{T}^d \times \mathcal{O}(g,G) \to \mathbb{R}$:

$$h_0(\psi, y) = \int_{\mathbb{T}^m} \tilde{P}(\psi, \varphi, y)\mathrm{d}\varphi,$$

where

$$\tilde{P}(\psi, \varphi, y) = P((K_0^\top)^{-1}\begin{pmatrix}\varphi \\ \psi\end{pmatrix}, y, 0).$$

For each $y \in \mathcal{O}(g,G)$, a m-torus $T_y^\varepsilon(\psi)$ of the unperturbed system is said to be *Poincaré non-degenerate* if ψ is a non-degenerate critical point of $h_0(\cdot, y)$, i.e., $\frac{\partial h_0}{\partial \psi}(\psi, y) = 0$ and $\frac{\partial^2 h_0}{\partial \psi^2}(\psi, y)$ is non-singular. By the implicit function theorem, if there exists one Poincaré non-degenerate m-torus, then in its neighborhood there is an analytic family of them. Thus, instead of assuming the existence of one such torus, we assume the following condition:

A4) There is a real analytic function $\psi : \mathcal{O}(g,G) \to \mathbb{T}^d$ such that $T_y^\varepsilon =: T_y^\varepsilon(\psi(y))$ is a Poincaré non-degenerate m-torus for each $y \in \mathcal{O}(g,G)$.

We will show that Poincaré non-degeneracy is an important mechanism for the majority persistence of these m-tori on $\mathcal{O}(g,G)$. We refer this mechanism as the *Poincaré-Treshchev mechanism* because it was first discovered by Poincaré for the maximal resonance and generalized by Treshchev for general resonances, in standard nearly integrable Hamiltonian systems.

Let $b =: 24d^2(N+1)$, $d = \text{rank } g$, and N be as in **A3)**. Our main result states as follows.

Main Theorem. *Assume the condition* **A1)** *and let g be a resonant group satisfying conditions* **A2) - A4)**. *Then there exist an $\varepsilon_0 > 0$ sufficiently small and Cantor sets $\mathcal{O}_\varepsilon \subset \mathcal{O}(g,G)$, $0 < \varepsilon < \varepsilon_0$ with $|\mathcal{O}(g,G) \setminus \mathcal{O}_\varepsilon| = O(\varepsilon^{\frac{\iota}{2bN}})$ for some fixed constant $0 < \iota < \frac{1}{3}$ such that for each $0 < \varepsilon < \varepsilon_0$ the Hamiltonian*(1.1) *admits a C^{N-1} Whitney smooth family of quasi-periodic, invariant m-tori $\hat{T}_y^\varepsilon$, $y \in \mathcal{O}_\varepsilon$. Moreover, for each $y \in \mathcal{O}_\varepsilon$ and $0 < \varepsilon < \varepsilon_0$, $\hat{T}_y^\varepsilon$ and its frequency vector are only slightly deformed from those of the unperturbed Poincaré non-degenerate m-torus T_y^ε.*

Remark 1.2. (1) We note that the Main Theorem actually holds when the Hamiltonian (1.1) is of the class C^∞. The proof follows from our proof in Section 3 and Section 4 with more derivative estimates involved.

(2) The Main Theorem also holds on a sub-manifold M of G if the condition **A2)-A4)** are assumed to hold on M instead. In particular, if M is taken as an energy surface $\{H_0 = h_0,\ H_1 = h_1,\ \cdots,\ H_\alpha = h_\alpha\}$, then the Main Theorem will lead to a persistence result for lower-dimensional tori in the iso-energetic case. To prove such result on a sub-manifold M, one applies the same iterative scheme in this paper with M in place of G then uses the measure estimate on a sub-manifold contained in [4]. We note that the validity of assumptions **A2)-A4)** on M now depends on both choices of sub-manifold M and resonant group g, which can be a non-trivial matter in applications.

The proof of the above result uses the approaches of our early work [20] but involves more complicated technical treatments. In Section 2, we reduce (1.1) to a resonant-splitting normal form containing multi-scale tangential frequencies, i.e., for each $\xi \in \mathcal{O}(g,G)$,

$$\begin{aligned} & H(x,y,z,\xi,\varepsilon) \\ = \quad & e_\varepsilon(\xi) + \langle \omega_\varepsilon(\xi), y\rangle + \frac{\delta}{2}\langle \begin{pmatrix} y \\ z \end{pmatrix}, \varepsilon^{\bar{m}_{I+1}} M(\xi,\varepsilon) \begin{pmatrix} y \\ z \end{pmatrix}\rangle + \delta h(y,\xi,\varepsilon) + \delta\varepsilon^{\bar{m}_\alpha} P(x,y,z,\xi,\varepsilon), \end{aligned}$$

where $(x,y,z) \in \mathbb{T}^m \times \mathbb{R}^m \times \mathbb{R}^{2d}$, $\delta = \varepsilon^{\frac{\bar{m}_{\alpha+1} - \bar{m}_\alpha}{2}}$. Unlike the case of a standard nearly integrable Hamiltonian system considered in [9] - [11], the appearance of the multi-scale frequency vector ω_ε is a major obstacle for standard KAM iterations to apply because in order to carry out such iterations small divisor conditions on ω_ε require that the perturbation is of at least an order of $O(\varepsilon^{2(N+6)c_*})$, where

$$c_* = \sum_{i=1}^{\alpha} \bar{m}_i (n_i - n_{i-1}),$$

which needs not be the case in many applications (see the examples in [14] or Section 5 of this paper). Thus, in Section 3, we perform a finite step of iterations to the resonant-splitting normal form to obtain an improved Hamiltonian with the new perturbation in a desired order. This finite step of iterations is carried out using small divisor conditions relating to the ε-independent frequency map Ω^* in order to have concrete controls of sizes of both phase and frequency domains. We note that if the perturbation in the original Hamiltonian is already in the order of $O(\varepsilon^{2(N+6)c_*})$

(see e.g., [3]), then a finite step of quasi-linear iterations is not necessary and the excluding measure for the persistence of invariant, quasi-periodic, m-dimensional tori can be improved to an order of $O(\varepsilon^{\frac{12c_*}{N}})$.

The rest sections are organized as follows. In Section 4, we apply the standard KAM iterations to the improved normal form to complete the proof of Main Theorem. In Section 5, we adopt a multi-scale, nearly integrable celestial system of 3-degree of freedom from [14, Section 4.4] to show the validity of conditions **A1)-A3)** on certain resonant surfaces. We left the condition **A4)** unchecked because of the unavailability of an explicit perturbation. We remark that, as a generic condition, the condition **A4)** is satisfied by generic perturbations. It should also be largely satisfied by particular perturbations which are not highly degenerate, simply because of the freedom of choosing resonance types in verifying this condition.

Through the rest sections, we let g be the resonant group considered in the Main Theorem. We use the symbol $|\cdot|$ to denote absolutely value for constant, vector norm and matrix norm induced by vector norm and the Lebesgue measure of a set. Also, we use $|\cdot|_D$ to denote the sup-norm of vector or matrix functions on a domain D and for each $r, s > 0$,

$$D(r,s) = \{(x,y,u,v) \in \mathbb{T}^m \times \mathbb{C}^m \times \mathbb{C}^d \times \mathbb{C}^d : \ |\mathrm{Im}\ x| < r,\ |y| < s,\ |z| < s\}$$

denotes a complex neighborhood of $\mathbb{T}^m \times \{0\} \times \{0\} \times \{0\}$.

2. Normal form

In this section, we reduce the Hamiltonian (1.1) into a resonant-splitting normal form near the family of Poincaré non-degenerate m-tori T_ξ^ε, $\xi \in \mathcal{O}(g,G)$. Firstly, we expand the Hamiltonian (1.1) at each $\xi \in \mathcal{O}(g,G)$ and obtain the following parametrized Hamiltonian

$$\begin{aligned} & H(x,y,\varepsilon) \\ (2.1) \qquad = \ & \langle \omega_\varepsilon^*(\xi), y-\xi\rangle + \frac{1}{2}\langle \hat{\Gamma}(\xi,\varepsilon)(y-\xi), y-\xi\rangle + \varepsilon^{\bar{m}_{\alpha+1}} P(x,y,\varepsilon) + O(|y-\xi|^3) \end{aligned}$$

up to the omission of a constant, where $\omega_\varepsilon^*(\xi)$ is of the form (1.3). For each $j = 0, 1, \cdots, \alpha$, denote $\xi^{n_j} = (\xi_1, \cdots, \xi_{n_j})^\top$, $\hat{\xi}^{n_j} = (\xi_{n_{j-1}+1}, \cdots, \xi_{n_j})^\top$, where $\xi^{n_0} = \hat{\xi}^{n_0}$ and $\xi^{n_\alpha} = \xi$. Then the matrix $\hat{\Gamma}(\xi,\varepsilon)$ in (2.1) can be expressed as

$$\hat{\Gamma}(\xi,\varepsilon) = \hat{\Gamma}_0(\xi^{n_0}) + \cdots + \varepsilon^{\bar{m}_\alpha}\hat{\Gamma}_\alpha(\xi^{n_\alpha}),$$

where

$$\hat{\Gamma}_i(\xi^{n_i}) = \begin{pmatrix} \frac{\partial^2 H_i}{\partial {y^{n_i}}^2}(\xi^{n_i}) & 0 \\ 0 & 0 \end{pmatrix}_{n\times n}, \quad i = 0, 1, \cdots, \alpha.$$

Denote

$$\hat{\Gamma}_{i,j}(\xi^{n_i}) =: \frac{\partial^2 H_i}{\partial (\hat{y}^{n_i})^2}(\xi^{n_i}),\ j = 0, 1, \cdots, \alpha.$$

Then, for each $j = 0, 1, \cdots, \alpha$, the matrix $\hat{\Gamma}_i(\xi^{n_i})$ can be rewritten as

$$\hat{\Gamma}_i(\xi^{n_i}) = \begin{pmatrix} \hat{\Gamma}_{i,1}(\xi^{n_i}) & * & \cdots & * & \\ * & \hat{\Gamma}_{i,2}(\xi^{n_i}) & \cdots & * & O \\ \vdots & \cdots & \ddots & \vdots & \\ * & * & \cdots & \hat{\Gamma}_{i,i}(\xi^{n_i}) & \\ & & & & \\ & O & & & O \end{pmatrix}_{n\times n}, \ \forall\ 0 \ \le i \le \alpha,$$

where $*$ denotes constant matrices independent of ε and O denotes zero matrices of appropriate dimension.

Secondly, we consider the linear symplectic transformation

$$\begin{aligned} y-\xi &= K_0 p, \\ q &= {K_0}^{\top} x, \end{aligned} \tag{2.2}$$

under which the Hamiltonian (2.1) becomes

$$\begin{aligned} &\bar{H}(\bar{p},p,q,\xi,\varepsilon) \\ =\ & \langle \omega_\varepsilon(\xi),\bar{p}\rangle + \frac{1}{2}\langle \bar{\Gamma}(\xi,\varepsilon)p,p\rangle + \varepsilon^{\bar{m}_{\alpha+1}}\hat{P}(p,q,\xi,\varepsilon) + O(|p|^3), \end{aligned} \tag{2.3}$$

where $\bar{p}$ is first m-components of p, $\omega_\varepsilon(\xi) = K_1^{\top}\omega_\varepsilon^*(\xi)$ is as in (1.6), $\bar{\Gamma}(\xi,\varepsilon) = {K_0}^{\top}\hat{\Gamma}(\xi,\varepsilon)K_0$, and $\hat{P}(q,p,\xi,\varepsilon) = P(({K_0}^{\top})^{-1}q, \xi + K_0 p, \varepsilon)$. Rewrite $\bar{\Gamma}(\xi,\varepsilon)$ as

$$\bar{\Gamma}(\xi,\varepsilon) = \bar{\Gamma}_0(\xi^{n_0}) + \cdots + \varepsilon^{\bar{m}_\alpha}\bar{\Gamma}_0(\xi^{n_\alpha}),$$

where $\bar{\Gamma}_i(\xi^{n_i}) = K_0^{\top}\hat{\Gamma}_i(\xi^{n_i})K_0$, $i = 0,1,\cdots,\alpha$. For each $i = 0,1,\cdots,\alpha$, we further express $\bar{\Gamma}_i(\xi^{n_i})$ as

$$\bar{\Gamma}_i(\xi^{n_i}) = \begin{pmatrix} \bar{\Gamma}_i^{11} & \bar{\Gamma}_i^{12} \\ \bar{\Gamma}_i^{21} & \bar{\Gamma}_i^{22} \end{pmatrix},$$

where $\bar{\Gamma}_i^{11},\ \bar{\Gamma}_i^{12},\ \bar{\Gamma}_i^{21},\ \bar{\Gamma}_i^{22}$ are $m\times m,\ m\times d,\ d\times m,\ d\times d$ blocks, respectively. It is easy to see that $\bar{\Gamma}_i^{12},\ \bar{\Gamma}_i^{21}, \bar{\Gamma}_i^{22}$ are zero matrices for each $i = 0,1,\cdots,I$.

For each $i = I+1,\cdots,\alpha$, we have

$$\hat{\Gamma}_i^{22} = \begin{pmatrix} \bar{\Gamma}_{i,I+1} & * & \cdots & * & \\ * & \bar{\Gamma}_{i,I+2} & \cdots & * & O \\ \vdots & \vdots & \ddots & \vdots & \\ * & * & \cdots & \bar{\Gamma}_{i,i} & \\ & O & & & O \end{pmatrix}_{d\times d},$$

where $*$ denote constant matrices independent of ε and O denote zero matrices of appropriate dimension. More specifically, $\bar{\Gamma}_{i,I+j} = (\hat{K}_2^{I+j})^{\top}\hat{\Gamma}_{i,I+j}\hat{K}_2^{I+j}$, for $i = I+1,\cdots,\alpha,\ j = 1,\cdots,i$.

Let $p = (\bar{p},p^d)^{\top} \in \mathbb{R}^m \times \mathbb{R}^d$ and $q = (\varphi,\psi)^{\top}$. The Hamiltonian (2.3) then becomes

$$\begin{aligned} H(\varphi,\psi,\bar{p},p^d,\xi,\varepsilon) &= \langle \omega_\varepsilon(\xi),\bar{p}\rangle + \frac{1}{2}\langle \varepsilon^{\bar{m}_{I+1}}\Gamma(\xi,\varepsilon)p,p\rangle \\ &\quad + h(\bar{p},p^d,p,\xi,\varepsilon) + \varepsilon^{\bar{m}_{\alpha+1}}\bar{P}(\varphi,\psi,\bar{p},p^d,\xi,\varepsilon), \end{aligned} \tag{2.4}$$

where

$$\Gamma(\xi,\varepsilon) = \begin{pmatrix} O & \Gamma^{12}(\xi,\varepsilon) \\ \Gamma^{21}(\xi,\varepsilon) & \Gamma^{22}(\xi,\varepsilon) \end{pmatrix}_{n\times n}, \tag{2.5}$$

$$\Gamma^{\imath\jmath}(\xi,\varepsilon) = \bar{\Gamma}^{\imath\jmath}_{I+1} + \varepsilon^{\bar{m}_{I+2}-\bar{m}_{I+1}}\bar{\Gamma}^{\imath\jmath}_{I+2} + \cdots + \varepsilon^{\bar{m}_\alpha-\bar{m}_{I+1}}\bar{\Gamma}^{\imath\jmath}_{\alpha},\ (\imath\jmath) = (12),(21),(22), \tag{2.6}$$

$$\begin{aligned} h(p,\xi,\varepsilon) &= h_1(\bar{p},\xi,\varepsilon) + h_2(p,\xi,\varepsilon), \\ h_1(\bar{p},\xi,\varepsilon) &= O(|\bar{p}^0|^2 + \cdots + \varepsilon^{\bar{m}_I}|\bar{p}^I|^2 + \varepsilon^{\bar{m}_{I+1}}|\bar{p}|^2), \\ h_2(p,\xi,\varepsilon) &= O(\varepsilon^{\bar{m}_{I+1}}|(\bar{p},p^{d_{I+1}})|^3 + \cdots + \varepsilon^{\bar{m}_\alpha}|(\bar{p},p^{d_\alpha})|^3), \\ \bar{p}^j &= (p_1,\cdots,p_{n_j})^{\top},\ j = 0,1,\cdots,I, \\ p^{d_{I+j}} &= (p_{m+1},\cdots,p_{n_{I+j}}),\ j = I+1,\cdots,\alpha,\ p^{d_\alpha} = p^d, \\ \bar{P}(\varphi,\psi,\bar{p},p^d,\xi,\varepsilon) &= \hat{P}(\varphi,\psi+\psi(\xi),\bar{p},p^d,\xi,\varepsilon). \end{aligned} \tag{2.7}$$

As to be seen later, lower order terms $h(p,\xi,\varepsilon)$ will play an important role during the finite KAM iterations, since they simply cannot be included in the perturbation.

Rewrite the matrix $\Gamma^{22}(\xi,\varepsilon)$ as

$$\begin{pmatrix} \bar\Gamma_{I+1,I+1}+O(\varepsilon) & O(\varepsilon^{\bar m_{I+2}-\bar m_{I+1}}) & \cdots & O(\varepsilon^{\bar m_\alpha-\bar m_{I+1}}) \\ O(\varepsilon^{\bar m_{I+2}-\bar m_{I+1}}) & \varepsilon^{\bar m_{I+2}-\bar m_{I+1}}(\bar\Gamma_{I+2,I+2}+O(\varepsilon)) & \cdots & O(\varepsilon^{\bar m_\alpha-\bar m_{I+1}}) \\ \vdots & \vdots & \ddots & \vdots \\ O(\varepsilon^{\bar m_\alpha-\bar m_{I+1}}) & O(\varepsilon^{\bar m_\alpha-\bar m_{I+1}}) & \cdots & \varepsilon^{\bar m_\alpha-\bar m_{I+1}}\bar\Gamma_{\alpha,\alpha} \end{pmatrix}_{d\times d}.$$

Then it is easy to verify that

$$\det\Gamma^{22}(\xi,\varepsilon)=\varepsilon^{\sum_{j=1}^{\alpha}(\bar m_{I+j+1}-\bar m_{I+j})n_j}\Big(\prod_{I+1\le I+j\le\alpha}\det\bar\Gamma_{I+j,I+j}-o(\varepsilon)\Big),$$

where "O,o" are asymptotic orders involving constants depending on ε. Since by **A2)** and Remark 1.1, all sub-matrices $\{\bar\Gamma_{I+j,I+j}\}$ are non-singular, we deduce that $\Gamma^{22}(\xi,\varepsilon)$ is non-singular for all $\xi\in\mathcal{O}(g,G)$, as ε sufficiently small.

Denote

$$\Omega_\varepsilon(\xi)=(\omega_\varepsilon^0(\xi),\cdots,\hat\omega_\varepsilon^i(\xi),\cdots,\hat\omega_\varepsilon^{I+1}(\xi))^\top,\qquad \xi\in\mathcal{O}(g,G), \tag{2.8}$$

where ω_ε^i, $i=0,1,\cdots,I+1$, are as in (1.6). Then

$$\partial_\xi^l\Omega_\varepsilon(\xi)=\partial_\xi^l K_1^\top\Omega^*(\xi)+O(\varepsilon),\qquad l=0,1,\cdots,N,\qquad \xi\in\mathcal{O}(g,G), \tag{2.9}$$

where Ω^* is as in (1.2). For fixed positive constants $\gamma_0,\ \tau>\max\{m(m+1)-1,N(N+1)-1\}$, consider sets

$$\Lambda_\varepsilon=\{\xi\in\mathcal{O}(g,G):|\langle k,\Omega_\varepsilon(\xi)\rangle|>\frac{\gamma_0}{|k|^\tau},\ \forall 0\neq k\in\mathbb{Z}^m\},\qquad 0<\varepsilon\ll 1.$$

It follows from **A3)**, (2.9) and the standard measure estimate under Rüssmann conditions ([21, 22]) that

$$|\mathcal{O}(g,G)\setminus\Lambda_\varepsilon|=O(\gamma_0^{\frac{1}{N}}). \tag{2.10}$$

Next, we separate the first-order resonant terms from the perturbation $\bar P$. For each $\xi\in\Lambda_\varepsilon$, expand $\bar P$ as

$$\begin{aligned}\bar P(\varphi,\psi,p,\xi,\varepsilon) &= \sum_{k\in\mathbb{Z}^m}h_k(\psi,\xi)\mathrm{e}^{\sqrt{-1}\langle k,\varphi\rangle}+O(|p|^2)\\ &= h_0(0,\xi)+\frac12\langle\psi,\tilde\Gamma(\xi)\psi\rangle+\sum_{k\in\mathbb{Z}^m\setminus\{0\}}h_k(\psi,\xi)\mathrm{e}^{\sqrt{-1}\langle k,\varphi\rangle}\\ &\quad +O(|p|^2)+O(|\psi|^3),\end{aligned} \tag{2.11}$$

where $\tilde\Gamma(\xi)=\frac{\partial^2h_0}{\partial\psi^2}(\xi)$ and $h_k(\psi,\xi)=\int_{\mathbb{T}^d}\bar P(\varphi,\psi,0,\xi)\mathrm{e}^{\sqrt{-1}\langle k,\varphi\rangle}\mathrm{d}\varphi$, $k\in\mathbb{Z}^n$. Consider

$$S_\varepsilon^\xi(q,Y)=\langle Y,q\rangle+\varepsilon^{\bar m_{\alpha+1}}\sum_{k\in Z^m\setminus\{0\}}S_k\mathrm{e}^{\sqrt{-1}\langle k,\varphi\rangle},\quad \xi\in\Lambda_\varepsilon,$$

where $S_k=\frac{\sqrt{-1}h_k(\psi,\xi)}{\langle\omega_\varepsilon(\xi),k\rangle}$, $k\in\mathbb{Z}^n$. Then $\{S_\varepsilon^\xi\}$ generates the following Whitney smooth family of real analytic, symplectic transformations on $(\mathbb{T}^m\times\mathbb{R}^d)\times\mathbb{R}^n$:

$$(q,p)=(\varphi,\psi,Y):q=\frac{\partial S_\varepsilon^\xi(q,Y)}{\partial Y},\ p=\frac{\partial S_\varepsilon^\xi(q,Y)}{\partial q},\quad \xi\in\Lambda_\varepsilon.$$

More specifically, we have

$$\bar{p} = \bar{Y} + \sqrt{-1}\varepsilon^{\bar{m}_{\alpha+1}} \sum_{k\in\mathbb{Z}^m} kS_k \mathrm{e}^{\sqrt{-1}\langle k,\varphi\rangle} = \bar{Y} + O(\varepsilon^{\bar{m}_{\alpha+1}}),$$

and

$$p^d = Y^d + \sqrt{-1}\varepsilon^{\bar{m}_{\alpha+1}} \sum_{k\in\mathbb{Z}^m\setminus\{0\}} \frac{1}{\langle k,\omega_\varepsilon(\xi)\rangle}\frac{\partial h_k}{\partial\psi}\mathrm{e}^{\sqrt{-1}\langle k,\varphi\rangle} = Y^d + O(\varepsilon^{\bar{m}_{\alpha+1}}),$$

where $Y = (\bar{Y}, Y^d)^\top \in \mathbb{R}^m \times \mathbb{R}^d$. Under this family of symplectic transformations, the Hamiltonian (2.4) becomes

$$\begin{aligned} & H(\varphi,\psi,\bar{Y},Y^d,\xi,\varepsilon) \\ = \; & \langle\omega_\varepsilon(\xi),\bar{Y}\rangle + \frac{1}{2}\langle\varepsilon^{\bar{m}_{I+1}}\Gamma(\xi,\varepsilon)Y,Y\rangle + \frac{1}{2}\langle\varepsilon^{\bar{m}_{\alpha+1}}\tilde{\Gamma}\psi,\psi\rangle \\ & + h(Y,\xi,\varepsilon) + O(\varepsilon^{\bar{m}_{\alpha+1}}|Y|^2) + O(\varepsilon^{2\bar{m}_{\alpha+1}}) + O(\varepsilon^{\bar{m}_{\alpha+1}}|\psi|^3). \end{aligned}$$

Finally, the rescaling $Y \to \varepsilon^{\frac{\bar{m}_{\alpha+1}-\bar{m}_\alpha}{2}} Y$ yields that

$$\begin{aligned} \tilde{H} \; = \; & \frac{H(\varphi,\psi,\varepsilon^{\frac{\bar{m}_{\alpha+1}-\bar{m}_\alpha}{2}}\bar{Y},\varepsilon^{\frac{\bar{m}_{\alpha+1}-\bar{m}_\alpha}{2}}Y^d,\xi,\varepsilon)}{\varepsilon^{\frac{\bar{m}_{\alpha+1}-\bar{m}_\alpha}{2}}} \\ = \; & \langle\omega_\varepsilon(\xi),\bar{Y}\rangle + \frac{\varepsilon^{\frac{\bar{m}_{\alpha+1}-\bar{m}_\alpha}{2}}}{2}\langle\varepsilon^{\bar{m}_\alpha}\tilde{\Gamma}\psi,\psi\rangle + \varepsilon^{\frac{\bar{m}_{\alpha+1}-\bar{m}_\alpha}{2}}\langle\varepsilon^{\bar{m}_{I+1}}\Gamma(\xi,\varepsilon)Y,Y\rangle + \varepsilon^{\frac{\bar{m}_{\alpha+1}-\bar{m}_\alpha}{2}}h(Y,\xi,\varepsilon) \\ & + O(\varepsilon^{\bar{m}_{\alpha+1}+\frac{\bar{m}_{\alpha+1}-\bar{m}_\alpha}{2}}|Y|^2) + O(\varepsilon^{\frac{3\bar{m}_{\alpha+1}+\bar{m}_\alpha}{2}}) + O(\varepsilon^{\frac{\bar{m}_{\alpha+1}+\bar{m}_\alpha}{2}}|\psi|^3). \end{aligned}$$

By replacing $\varphi,\ \bar{Y},\ Y^d,\ \psi, \tilde{H}, \varepsilon^{\frac{\bar{m}_{\alpha+1}-\bar{m}_\alpha}{2}}$ with $x,\ y,\ u,\ v,\ H,\ \delta$ respectively, we obtain the following resonant-splitting normal form:

$$H(x,y,z,\xi,\varepsilon) \;=\; N + \delta\varepsilon^{\bar{m}_\alpha}P, \tag{2.12}$$

where

$$\begin{aligned} N \; &= \; \langle\omega_\varepsilon(\xi),y\rangle + \frac{\delta}{2}\langle\begin{pmatrix} y \\ z\end{pmatrix},\varepsilon^{\bar{m}_{I+1}}M(\xi,\varepsilon)\begin{pmatrix} y \\ z\end{pmatrix}\rangle + \delta h(y,z,\xi,\varepsilon), \\ M(\xi,\varepsilon) \; &= \; \begin{pmatrix} O & M_{12} \\ M_{21} & M_{22}\end{pmatrix}_{(m+2d)\times(m+2d)}, \qquad (2.13) \\ M_{12}(\xi,\varepsilon) \; &= \; \begin{pmatrix} \Gamma^{12} & O \end{pmatrix}_{m\times 2d},\ M_{21}(\xi,\varepsilon) = \begin{pmatrix} \Gamma^{21} \\ O \end{pmatrix}_{2d\times m}, \\ M_{22}(\xi,\varepsilon) \; &= \; \begin{pmatrix} \Gamma^{22} & O \\ O & \varepsilon^{\bar{m}_\alpha - \bar{m}_{I+1}}\tilde{\Gamma}\end{pmatrix}_{2d\times 2d}, \\ h(y,z,\xi,\varepsilon) \; &= \; h_1(y,\xi,\varepsilon) + h_2(y,z,\xi,\varepsilon), \qquad (2.14) \\ h_1(y,\xi,\varepsilon) \; &= \; O(|y^0|^2 + \cdots + \varepsilon^{\bar{m}_I}|y^I|^2 + \varepsilon^{\bar{m}_{I+1}}|y|^2), \\ h_2(y,z,\xi,\varepsilon) \; &= \; O(\varepsilon^{\bar{m}_{I+1}}|(y,u^{I+1})|^3 + \cdots + \varepsilon^{\bar{m}_\alpha}|(y,u)|^3 + \varepsilon^{\bar{m}_\alpha}|v|^3), \\ y^j \; &= \; (y_1,\cdots,y_{n_j})^\top,\ j = 0,1,\cdots,I,\ y = \bar{p}, \\ u^{I+j} \; &= \; (p_{m+1},\cdots,p_{n_{I+j}}),\ j = 1,\cdots,\alpha - I,\ u^\alpha = u = p^d, \\ P(x,y,z,\xi,\varepsilon) \; &= \; O(\varepsilon^{\bar{m}_{\alpha+1}-\bar{m}_\alpha}|(y,u)|) + O(\varepsilon^{\bar{m}_{\alpha+1}}) + O(|\psi|^3), \end{aligned}$$

and $\Gamma^{12},\ \Gamma^{21}$ are as (2.6). We note that M_{22} is non-singular and $M_{12}^\top = M_{21}$ for each $\xi \in \Lambda_\varepsilon$ and ε sufficiently small. Let $0 < r_0 \ll 1$ be fixed and $s_0 = \varepsilon^{\frac{\bar{m}_{\alpha+1}-\bar{m}_\alpha}{6}}$ for sufficiently small $\varepsilon > 0$. Then the normal form (2.12) is real analytic in $(x,y,z) \in D(r_0,s_0)$ and Whitney smooth in $\xi \in \Lambda_\varepsilon$.

Denote $\gamma_0^b = \varepsilon^{\frac{\iota(\bar{m}_{\alpha+1}-\bar{m}_\alpha)}{2}}$, $\mu_0 = \varepsilon^{\frac{(1-\iota)(\bar{m}_{\alpha+1}-\bar{m}_\alpha)}{2}}$, where $0 < \iota < \frac{1}{3}$ is fixed and $b = 24d^2(N+1)$. Then we have

$$|\partial_\xi^l P|_{D(r_0,s_0)\times\Lambda_\varepsilon} = \gamma_0^b s_0^2 \mu_0, \quad \forall l \le N. \tag{2.15}$$

3. Improving the Resonant Splitting Normal Form

In this section, we will perform a finite step of quasi-linear KAM iterations to the normal form (2.12) in order to push the perturbation to a desired high order. We start from the parametrized real analytic Hamiltonian (2.12) and re-write it as

$$\begin{aligned} &H^0(x,y,z,\xi,\varepsilon) \\ = \; & e_\varepsilon^0(\xi) + \langle \omega_\varepsilon^0(\xi), y\rangle + \frac{\delta}{2}\langle \begin{pmatrix} y \\ z \end{pmatrix}, \varepsilon^{\bar{m}_{I+1}} M^0(\xi,\varepsilon)\begin{pmatrix} y \\ z \end{pmatrix}\rangle + \delta h^0(y,\xi,\varepsilon) + \delta\varepsilon^{\bar{m}_\alpha} P^0(x,y,z,\xi,\varepsilon), \end{aligned} \tag{3.1}$$

where $(x,y,z) \in D(r_0,s_0)$, $\xi \in \Lambda^0 =: \Lambda_\varepsilon$, $e_\varepsilon^0 \equiv 0$, and $\omega_\varepsilon^0 =: \omega_\varepsilon$, $M^0 =: M$, $h^0 =: h$, and $P^0 =: P$ are as in (1.6), (2.13), (2.14), and (2.15) respectively . In the following, we show the quasi-linear iteration process from the ν-th step to the $\nu+1$-th step. For simplicity, we omit the subscript at the ν-th step and denote the $\nu+1$ step by $+$.

Assume that, after the ν-th iterative step, we obtain the following smooth family of real analytic Hamiltonians

$$\begin{aligned} &H(x,y,z,\xi,\varepsilon) \\ = \; & e_\varepsilon(\xi) + \langle \omega_\varepsilon(\xi), y\rangle + \frac{\delta}{2}\langle \begin{pmatrix} y \\ z \end{pmatrix}, \varepsilon^{\bar{m}_{I+1}} M(\xi,\varepsilon)\begin{pmatrix} y \\ z \end{pmatrix}\rangle + \delta h(y,\xi,\varepsilon) + \delta\varepsilon^{\bar{m}_\alpha} P(x,y,z,\xi,\varepsilon), \end{aligned} \tag{3.2}$$

where $(x,y,z) \in D(r,s)$ for some smaller $0 < r < r_0$, $0 < s < s_0$, $\xi \in \Lambda \subset \Lambda^0 \subset \mathcal{O}(g,G)$, ω_ε, M, h are of the same forms as in (1.6), (2.13), (2.14) respectively, and

$$|\partial_\xi^l P|_{D(r,s)\times\Lambda} < \gamma^b s^2 \mu, \ |l| \le N$$

for some constant $0 < \mu < \mu_0$. For each $\xi \in \Lambda$, we write

$$\omega_\varepsilon(\xi) = (\omega_\varepsilon^0(\xi), \cdots, \varepsilon^{\bar{m}_i}\hat{\omega}_\varepsilon^i(\xi), \cdots \varepsilon^{\bar{m}_{I+1}}\hat{\omega}_\varepsilon^{I+1}(\xi))^\top,$$

and denote

$$\Omega_\varepsilon(\xi) = (\omega_\varepsilon^0(\xi), \cdots, \hat{\omega}_\varepsilon^i(\xi), \cdots, \hat{\omega}_\varepsilon^{I+1}(\xi))^\top. \tag{3.3}$$

For the $\nu+1$-th step, we will find a smooth family of symplectic transformations Φ^+, to yield a new smooth family of real analytic Hamiltonians $H^+ = H \circ \Phi^+$ of the same form as (2.12) whose perturbations P^+ are dramatically smaller on new frequency and phase domains.

Define

$$
\begin{aligned}
r_+ &= \frac{r}{2} + \frac{7}{16} r_0, \\
\gamma_+ &= \frac{\gamma}{2} + \frac{\gamma_0}{4}, \\
s_+ &= \frac{1}{8}\alpha s, \ \alpha = \mu^{\frac{1}{3}}, \\
K_+ &= ([\log \frac{1}{\mu}] + 1)^3, \\
D_{i\alpha} &= D(r_+ + \frac{i-1}{8}(r - r_+), \frac{i}{8}\alpha s), \ i = 1, 2, \cdots, 8, \\
D_+ &= D_\alpha = D(r_+, s_+), \\
\hat{D}(s) &= D(r_+ + \frac{7}{8}(r - r_+), s), \\
\Lambda^+ &= \{\xi \in \Lambda : \ |\langle k, \Omega_\varepsilon(\xi)\rangle| > \frac{\gamma}{|k|^\tau}\}, \\
\Gamma(r - r_+) &= \mathrm{e}^{\frac{r_0}{108}} \sum_{0<|k|\le K_+} |k|^\chi \mathrm{e}^{-|k|\frac{r-r_+}{8}},
\end{aligned}
$$

where $\chi = 2(N+1)(4d^2+1)\tau$, $\tau > n-1$ is fixed.

In the rest of the paper, besides c, c_i, $i = 0, 1 \cdots, 9$ are also intermediate constants which are independent of ε and the iteration process.

3.1. **Truncation.** For each $\xi \in \Lambda$, we write the perturbation P into Taylor-Fourier series, i.e.,

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

Let

$$R = \sum_{|k|\le K_+} (p_{k00} + \langle p_{k10}, y\rangle + \langle p_{k01}, z\rangle + \langle y, p_{k20} y\rangle + \langle z, p_{k11} y\rangle + \langle z, p_{k02} z\rangle)\mathrm{e}^{\sqrt{-1}\langle k, x\rangle}$$

be the truncation of the Taylor-Fourier series of P up to the order K_+.

Lemma 3.1. *Assume that*

(H1) $\displaystyle\int_{K_+}^{\infty} t^{d+N+3} \mathrm{e}^{-t\frac{r-r_+}{16}} \mathrm{d}t \le \mu.$

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

$$|\partial_\xi^l (P - R)|_{D_{7\alpha}} \le c_1 \gamma^b s^2 \mu^2, \quad |\partial_\xi^l R|_{D_{7\alpha}} \le c_1 \gamma^b s^2 \mu.$$

Proof. See Lemma 3.1 in [11]. □

3.2. **Quasi-linear homological equation.** For each $\xi \in \Lambda^+$, we will construct a symplectic transformation ϕ_F^1 to the Hamiltonian H as the time-1 map of the flow ϕ_F^t generated by a Hamiltonian F of the following form:

$$\begin{aligned} F &= \sum_{0<|k|\le K_+} (f_{k00} + \langle f_{k10}, y\rangle + \langle f_{k01}, z\rangle + \langle y, f_{k20}y\rangle \\ &\quad + \langle z, f_{k11}y\rangle + \langle z, f_{k02}z\rangle)\mathrm{e}^{\sqrt{-1}\langle k,x\rangle} + \langle f_{001}, z\rangle, \end{aligned} \tag{3.4}$$

where $f_{k\imath\jmath}$ are (matrix valued) functions of (ξ, ε) and $f_{k20},\ f_{k02}$ are symmetric for each k. The transformed Hamiltonian then reads

$$\begin{aligned} H\circ\phi_F^1 &= N\circ\phi_F^1 + \delta\varepsilon^{\bar m_\alpha}(P-R)\circ\phi_F^1 \\ &= N + \{N,F\} + \delta\varepsilon^{\bar m_\alpha}R + \int_0^1 \{(1-t)\{N,F\},F\}\circ\phi_F^t \mathrm{d}t \\ &\quad + \delta\varepsilon^{\bar m_\alpha}\int_0^1 \{(1-t)R,F\}\circ\phi_F^t\mathrm{d}t + \delta\varepsilon^{\bar m_\alpha}(P-R)\circ\phi_F^1 \end{aligned}$$

$$= N + \delta\varepsilon^{\bar m_\alpha}([R] - \langle p_{001}, z\rangle) + (\{N,F\} + \delta\varepsilon^{\bar m_\alpha}R - \delta\varepsilon^{\bar m_\alpha}[R] - Q + \delta\varepsilon^{\bar m_\alpha}\langle p_{001}, z\rangle) \tag{3.5}$$

$$+ \int_0^1 \{(1-t)\{N,F\},F\}\circ\phi_F^t\mathrm{d}t + \delta\varepsilon^{\bar m_\alpha}\int_0^1 \{R,F\}\circ\phi_F^t\mathrm{d}t + \delta\varepsilon^{\bar m_\alpha}(P-R)\circ\phi_F^1 + Q, \tag{3.6}$$

where $[R] = \frac{1}{(2\pi)^d}\int_{\mathbb{T}^d} R(x,\cdot)\mathrm{d}x$.

To average out the resonant terms among R, we will determine the Hamiltonian F by solving the following quasi-linear homological equation

$$\{N,F\} + \delta\varepsilon^{\bar m_\alpha}(R - [R] + \langle p_{001}, z\rangle) - Q = 0, \tag{3.7}$$

where

$$\begin{aligned} Q &= -\sum_{0<|k|\le K_+} \langle \delta\varepsilon^{\bar m_{I+1}}\frac{\partial h_2}{\partial z}, J(f_{k01} + f_{k11}y + f_{k02}z + f_{k02}^\top z)\rangle e^{\sqrt{-1}\langle k,x\rangle} \\ &\quad - \langle \delta\varepsilon^{\bar m_{I+1}}\frac{\partial h_2}{\partial z}, Jf_{001}\rangle \end{aligned} \tag{3.8}$$

which contains all terms in $\{N,F\}$ of size $O(\delta\varepsilon^{\bar m_\alpha}s^3\mu)$ and thus can be included in the new perturbation term. We note that, by introducing the modified term Q, the homological equation (3.7) modifies the linear equations used in the standard KAM iterative scheme.

Now, if (3.7) is solvable on $\xi \in \Lambda^+$, then the transformed Hamiltonian $H\circ\phi_F^1$ is of the following form:

$$H\circ\phi_F^1 = N + \delta\varepsilon^{\bar m_\alpha}([R] - \langle p_{001}, z\rangle) + \bar P^+,$$

where

$$\begin{aligned} \bar P^+ &= \delta\varepsilon^{\bar m_\alpha}\int_0^1 \{(1-t)(R - [R] + \langle p_{001}, z\rangle) + R, F\}\circ\phi_F^t\mathrm{d}t + \delta\varepsilon^{\bar m_\alpha}(P-R)\circ\phi_F^1 \\ &\quad + \int_0^1 \{Q,F\}\circ\phi_F^t\mathrm{d}t + Q. \end{aligned} \tag{3.9}$$

This is the main idea of the quasi-linear KAM scheme.

3.3. **Solving the homological equation.** Write M into blocks

$$M=\left(\begin{array}{cc} O & M_{12}\\ M_{21} & M_{22}\end{array}\right), \tag{3.10}$$

where $M_{12},\ M_{21},\ M_{22}$ are $m\times 2d,\ 2d\times m, 2d\times 2d$ blocks of M respectively. Substituting (3.4), $[R]$, (3.8) into (3.7), we have

$$\begin{aligned}
& -\sum_{0<|k|\le K_+}\sqrt{-1}\langle k,\omega_\varepsilon(\xi)+\delta\partial_y h+\delta\varepsilon^{\bar m_{I+1}}M_{12}z\rangle(f_{k00}+\langle f_{k10},y\rangle\\
& +\langle f_{k01},z\rangle+\langle y,f_{k20}y\rangle+\langle z,f_{k02}z\rangle)\mathrm{e}^{\sqrt{-1}\langle k,x\rangle}\\
& +\sum_{0<|k|\le K_+}(\langle\delta\varepsilon^{\bar m_{I+1}}M_{21}y+\delta\varepsilon^{\bar m_{I+1}}M_{22}z,J(f_{k01}+f_{k11}y+f_{k02}z+f_{k02}^\top z)\mathrm{e}^{\sqrt{-1}\langle k,x\rangle}\\
& +\langle\delta\varepsilon^{\bar m_{I+1}}M_{21}y+\delta\varepsilon^{\bar m_{I+1}}M_{22}z,Jf_{001}\rangle\\
= & -\sum_{0<|k|\le K_+}\delta\varepsilon^{\bar m_\alpha}(p_{k00}+\langle p_{k10},y\rangle+\langle p_{k01},z\rangle+\langle y,p_{k20}y\rangle+\langle z,p_{k11}y\rangle+\langle z,p_{k02}z\rangle)\mathrm{e}^{\sqrt{-1}\langle k,x\rangle}\\
& -\delta\varepsilon^{\bar m_\alpha}\langle p_{001},z\rangle.
\end{aligned}$$

By comparing coefficients of the same order terms in the above, we obtain the following equations for all $\xi\in\Lambda^+,\ 0<|k|\le K_+$:

$$L_{0k}f_{k00} = \delta\varepsilon^{\bar m_\alpha}p_{k00}, \tag{3.11}$$

$$L_{1k}f_{k01} = \delta\varepsilon^{\bar m_\alpha}p_{k01}, \tag{3.12}$$

$$L_{2k}f_{k02} = \delta\varepsilon^{\bar m_\alpha}p_{k02}, \tag{3.13}$$

$$M_{22}Jf_{001} = -\varepsilon^{\bar m_\alpha-\bar m_{I+1}}p_{001}, \tag{3.14}$$

$$L_{0k}f_{k10} = \delta\varepsilon^{\bar m_\alpha}p_{k10}-\delta\varepsilon^{\bar m_{I+1}}M_{12}Jf_{001}, \tag{3.15}$$

$$L_{0k}f_{k20} = \delta\varepsilon^{\bar m_\alpha}p_{k20}-\frac{1}{2}\delta\varepsilon^{\bar m_{I+1}}(M_{12}Jf_{k11}-f_{k11}^\top JM_{21}), \tag{3.16}$$

$$L_{1k}f_{k11} = \delta\varepsilon^{\bar m_\alpha}p_{k11}-\delta\varepsilon^{\bar m_{I+1}}(f_{k02}^\top+f_{k02})JM_{21}, \tag{3.17}$$

where

$$\begin{aligned}
L_{0k} &= \sqrt{-1}\langle k,\omega_\varepsilon(\xi)+\delta\partial_y h+\delta\varepsilon^{\bar m_{I+1}}M_{12}z\rangle,\\
L_{1k} &= \sqrt{-1}\langle k,\omega_\varepsilon(\xi)+\delta\partial_y h+\delta\varepsilon^{\bar m_{I+1}}M_{12}z\rangle I_{2d}-\delta\varepsilon^{\bar m_{I+1}}M_{22}J,\\
L_{2k} &= \sqrt{-1}\langle k,\omega_\varepsilon(\xi)+\delta\partial_y h+\delta\varepsilon^{\bar m_{I+1}}M_{12}z\rangle I_{4d^2}-(\delta\varepsilon^{\bar m_{I+1}}M_{22}J)\otimes I_{2d}-I_{2d}\otimes(\delta\varepsilon^{\bar m_{I+1}}JM_{22}).
\end{aligned}$$

We first solve equations (3.11)-(3.14), then use the estimates of $f_{001},\ f_{k11},\ f_{k02}$ to solve equations (3.15)-(3.17).

For any $0<|k|\le K_+$, denote $k=(\hat k^0,\cdots,\hat k^j,\cdots,\hat k^{I+1})$, where $\hat k^{n_0}=(k_1,\cdots,k_{n_0}),\ \hat k^j=(k_{n_{j-1}+1},\cdots,k_{n_j}),\ j=1,\cdots,I$ and $\hat k^{I+1}=(k_{n_I+1},\cdots,k_{n_I+m})$. For some $j=0,\cdots,I+1$, let $\hat k^j$ be the first nonzero components of k. Then equation (3.11) becomes

$$\tilde L_{0k}=\delta\varepsilon^{\bar m_\alpha-\bar m_j}p_{k00},$$

where $\tilde L_{0k}=\langle\hat k^j,\hat\omega_\varepsilon^j(\xi)+O(\delta s)\rangle+\cdots+\varepsilon^{\bar m_{I+1}-\bar m_j}\langle\hat k^{I+1},\hat\omega_\varepsilon^{I+1}(\xi)+O(\delta s)\rangle$.

Assume that

(H2) $\max\{s,\delta s\}K_+^{\tau+1}\le o(\gamma_0)$.

Then

$$
\begin{aligned}
|\tilde{L}_{0k}| &\geq |\langle \hat{k}^j, \hat{\omega}_\varepsilon^j + O(\delta s)\rangle| - |O(\varepsilon^{\bar{m}_{j+1}-\bar{m}_j}))K_+| \\
&\geq |\langle k, \Omega_\varepsilon(\xi)\rangle| - |(O(\delta s)K_+| - |O(\varepsilon^{\bar{m}_{j+1}-\bar{m}_j}))K_+| \\
&\geq \frac{\gamma}{2|k|^\tau}.
\end{aligned}
$$

Combining the above with the estimate

$$|\partial_\xi^l L_{0k}| \leq \varepsilon^{\bar{m}_j} c|k|, \qquad |l| \leq N$$

and the inductive equations

$$\partial_\xi^l L_{qk}^{-1} = -\sum_{l'=1}^{l} C_l^{l'}(\partial_\xi^{l-l'} L_{qk}^{-1} \partial_\xi^{l'} L_{qk}) L_{qk}^{-1}, \qquad |l| \leq N,$$

we have

$$|\partial_\xi^l L_{0k}^{-1}| \leq c\frac{|k|^{(|l|+1)\tau+|l|}}{\varepsilon^{\bar{m}_{I+1}}\gamma^{|l|+1}}, \qquad |l| \leq N. \tag{3.18}$$

It follows that for each $0<|k|\leq K_+,\ \xi\in\Lambda^+$, equation (3.11) is uniquely solvable and the solution f_{k00} satisfies

$$|\partial_\xi^l f_{k00}| \leq c\delta\varepsilon^{\bar{m}_\alpha-\bar{m}_{I+1}}|k|^{|l|+(|l|+1)\tau} s^2\mu \mathrm{e}^{-|k|^\tau}. \tag{3.19}$$

To solve equation (3.12), we let $\hat{k}^j$, for some $j=0,\cdots,I+1$, contains the first non-zero components of k and rewrite L_{1k} as

$$L_{1k} = \varepsilon^{\bar{m}_j}\tilde{L}_{1k}, \tag{3.20}$$

where

$$
\begin{aligned}
\tilde{L}_{1k} &= \sqrt{-1}\langle \hat{k}^j, \hat{\omega}_\varepsilon^j(\xi) + O(\delta s)\rangle I_{2d} + \varepsilon^{\bar{m}_{j+1}-\bar{m}_j}\sqrt{-1}\langle \hat{k}^j, \hat{\omega}_\varepsilon^j(\xi) + O(\delta s)\rangle I_{2d} + \cdots \\
&\quad + \varepsilon^{\bar{m}_{I+1}-\bar{m}_j}\sqrt{-1}\langle \hat{k}^{I+1}, \hat{\omega}_\varepsilon^{I+1}(\xi) + O(\delta s)\rangle I_{2d} - \delta\varepsilon^{\bar{m}_{I+1}-\bar{m}_j} M_{22} J \\
&= \sqrt{-1}\langle k, \Omega_\varepsilon + O(\delta s)\rangle I_{2d} - \delta\varepsilon^{\bar{m}_{I+1}-\bar{m}_j} M_{22} J.
\end{aligned}
$$

Then equation (3.12) becomes

$$\tilde{L}_{1k} f_{k10} = \delta\varepsilon^{\bar{m}_\alpha-\bar{m}_j} p_{k01}.$$

For each $\xi\in\Lambda^+, 0<|k|\leq K_+$, it follows from **(H2)** that

$$|\langle k, \Omega_\varepsilon + O(\delta s)\rangle| \geq |\langle k, \Omega_\varepsilon\rangle| - |O(\delta s)K_+| \geq \frac{\gamma}{2|k|^\tau}.$$

Since

$$
\begin{aligned}
|\det\tilde{L}_{1k}| &\geq |\langle \hat{k}^j, \hat{\omega}_\varepsilon^j\rangle|^{2d}(1 - (\frac{2|k|^\tau\delta\varepsilon^{\bar{m}_{I+1}-\bar{m}_j}}{\gamma})^2 + \cdots + (\frac{2|k|^\tau\delta\varepsilon^{\bar{m}_{I+1}-\bar{m}_j}}{\gamma})^{2d}), \\
&\geq \frac{\gamma^{2d}}{2^{d+1}|k|^{2d\tau}},
\end{aligned}
$$

we see that for each $\xi\in\Lambda^+,\ 0<|k|\leq K_+$, $\tilde{L}_{1k}$ is invertible and

$$|\tilde{L}_{1k}^{-1}| = |\frac{\mathrm{adj}\bar{L}_{1k}}{\det\tilde{L}_{1k}}| \leq c\frac{|k|^{2d\tau+2d-1}}{\gamma^{2d}}. \tag{3.21}$$

This, together with the estimate

$$|\partial_\xi^l \tilde{L}_{1k}| \leq c|k|, \qquad |l| \leq N, \tag{3.22}$$

implies that

$$|\partial_\xi^l \tilde{L}_{1k}^{-1}| \leq c\frac{|k|^{|l|+(|l|+1)2d\tau}}{|\gamma|^{2d(|l|+1)}}, \qquad |l| \leq N. \tag{3.23}$$

Hence for each $\xi \in \Lambda^+,\ 0<|k|\le K_+$ equation (3.12) is uniquely solvable to yield solution

$$f_{k01} = \delta\varepsilon^{\bar m_\alpha - \bar m_j}\tilde L_{1k}^{-1}p_{k01}$$

which satisfies the estimate

$$|\partial_\xi^l f_{k01}| \le c\delta\varepsilon^{\bar m_\alpha-\bar m_{I+1}}|k|^{|l|+(|l|+1)2d\tau}s\mu \mathrm{e}^{-|k|^\tau},\quad |l|\le N. \tag{3.24}$$

Similarly, equation (3.13) is uniquely solvable for each $\xi\in\Lambda^+,\ 0<|k|\le K_+$ whose solution satisfies estimates

$$\begin{aligned}&|\partial_\xi^l f_{k02}| \le c\delta\varepsilon^{\bar m_\alpha-\bar m_I}|k|^{|l|+(|l|+1)4d^2\tau}\mu \mathrm{e}^{-|k|^\tau},\quad |l|\le N,\\ &|\delta\varepsilon^{\bar m_{I+1}}\partial_\xi^l(f_{k02}^\top+f_{k02})JM_{21}| \le c\delta\varepsilon^{\bar m_\alpha}|k|^{|l|+(|l|+1)4d^2\tau}\mu \mathrm{e}^{-|k|^\tau},\quad |l|\le N.\end{aligned} \tag{3.25}$$

By (3.23) and (3.25), we also see that equation (3.17) is uniquely solvable for each $\xi\in\Lambda^+,\ 0<|k|\le K_+$ whose solution satisfies the estimates

$$|\partial_\xi^l f_{k11}| \le c\delta\varepsilon^{\bar m_\alpha-\bar m_{I+1}}|k|^{|l|+(|l|+1)4d^2\tau}s\mu \mathrm{e}^{-|k|^\tau},\quad |l|\le N. \tag{3.26}$$

Similarly, it follows from (3.18) and (3.26) that the equation (3.16) is uniquely solvable for each $\xi\in\Lambda^+,\ 0<|k|\le K_+$ whose solution satisfies the estimates

$$|\partial_\xi^l f_{k20}| \le c\delta\varepsilon^{\bar m_\alpha-\bar m_{I+1}}|k|^{|l|+(|l|+1)4d^2\tau}\mu \mathrm{e}^{-|k|^\tau},\quad |l|\le N. \tag{3.27}$$

To solve equation (3.14), we assume that

(H3) $|\partial_\xi^l M_{12}-\partial_\xi^l M_{12}^0|,\ |\partial_\xi^l M_{22}-\partial_\xi^l M_{22}^0| \le \mu^{\frac14},\quad |l|\le N, \xi\in\Lambda^+.$

Then

$$|(M_{22})^{-1}|_{\Lambda^\nu} \le \frac{|(M_{22}^0)^{-1}|_{\Lambda^0}}{1-|M_{22}-M_{22}^0|_{\Lambda^\nu}|(M_{22}^0)^{-1}|_{\Lambda^0}} \le 2|(M_{22}^0)^{-1}|_{\Lambda^0}.$$

Recall that

$$M_{22}^0(\xi,\varepsilon) = \begin{pmatrix}\Gamma^{22} & O\\ O & \varepsilon^{\bar m_\alpha-\bar m_{I+1}}\tilde\Gamma\end{pmatrix},$$

where $\Gamma^{22},\ \tilde\Gamma$ are defined as in (2.6), (2.11) respectively. It is easy to see that

$$|(\Gamma^{22})^{-1}| = \frac{1}{|\det\Gamma^{22}|}|(\Gamma^{22})^*| \le O(\frac{1}{\varepsilon^{\bar m_\alpha-\bar m_{I+1}}}),$$

where $(\Gamma^{22})^*$ is the adjoint matrix of M. Since

$$(M_{22}^0)^{-1} = \begin{pmatrix}(\Gamma^{22})^{-1} & O\\ O & \frac{1}{\varepsilon^{\bar m_\alpha-\bar m_{I+1}}}\tilde\Gamma^{-1}\end{pmatrix},$$

where $|\tilde\Gamma^{-1}|=O(1)$, we have

$$|(M_{22}^0)^{-1}| \le O(\frac{1}{\varepsilon^{\bar m_\alpha-\bar m_{I+1}}}).$$

It follows that equation (3.14) is uniquely solvable for each $\xi\in\Lambda^+,\ 0<|k|\le K_+$ whose solution satisfies the estimates

$$\begin{aligned}&|\partial_\xi^l f_{001}| \le |M^{-1}||p_{001}| \le cs\mu,\quad |l|\le N,\\ &|\delta\varepsilon^{\bar m_{I+1}}M_{12}Jf_{001}| \le c\delta\varepsilon^{\bar m_{I+1}}|M_{12}M_{22}^{-1}||p_{001}| \le c\delta\varepsilon^{\bar m_\alpha}s\mu,\quad |l|\le N.\end{aligned} \tag{3.28}$$

Finally, it follows from (3.18) that equation (3.15) is solvable for each $\xi \in \Lambda^+,\ 0 < |k| \le K_+$ whose solution satisfies estimates

$$|\partial_\xi^l f_{k10}| \le c\delta\varepsilon^{\bar m_\alpha - \bar m_{I+1}} |k|^{|l|+(|l|+1)\tau} s\mu \mathrm{e}^{-|k|^\tau}, \qquad |l| \le N. \tag{3.29}$$

With the solvability of (3.11)-(3.17), the Hamiltonian F in (3.4) is now well defined.

Lemma 3.2. *Assume* **(H2)**, **(H3)** *and let* F *be as in* (3.4). *Then, there exists a positive constants* c_2 *such that on* $\hat D(s) \times \Lambda^+$,

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

and

$$|\partial_\xi^l \partial_x^i \partial_{(y,z)}^{(p,q)} F| \le c_2 \mu \Gamma(r - r_+)$$

for all $0 \le |l|, |i| \le N,\ 0 < |p| + |q| \le 2$.

Proof. It follows directly from (3.19) - (3.28). □

Lemma 3.3. *Assume that*

(H4) $c_2\mu\Gamma(r - r_+) + c_2\mu < \frac{1}{8}(r - r_+)$;
(H5) $c_2 s\mu\Gamma(r - r_+) + c_2 s\mu < s_+$.

Then for each $0 \le t \le 1$, *the transformation*

$$\phi_F^t : D_3 \to D_4$$

is well defined, real analytic, and depends on $\xi \in \Lambda^+$ *smoothly. Moreover, there is a constant* c_3 *such that*

$$|\partial_\xi^l (\phi_F^t - id)|_{D(s)\times\Lambda_+} \quad \le \quad c_3 s\mu\Gamma(r - r_+) \tag{3.30}$$

for all $|l| \le N,\ 0 \le t \le 1,\ i = 0, 1,$ *where* $D = \partial_{(x,y,z)}$.

Proof. It immediately follows from Lemma 3.2. □

3.4. **New Hamiltonian.** Denote $\Phi^+ := \phi_F^1$. Lemmas 3.2-3.3 imply that for each $\xi \in \Lambda^+$ the transformation $\Phi^+ :\ D(r_+, s_+) \to D(r, s)$ is well defined, real analytic and symplectic. The new Hamiltonian after transformation is defined on $D_+ \times \Lambda^+$ as follows:

$$\begin{aligned}
H^+ &=: H \circ \Phi^+ = N^+ + \delta\varepsilon^{\bar m_\alpha} P^+,\\
N^+ &= N + \delta\varepsilon^{\bar m_\alpha}([R] - \langle p_{001}, z\rangle)\\
&= e_\varepsilon^+(\xi) + \langle \omega_\varepsilon^+(\xi), y\rangle + \frac{\delta}{2}\langle \begin{pmatrix} y \\ z \end{pmatrix}, \varepsilon^{\bar m_{I+1}} M^+(\xi, \varepsilon)\begin{pmatrix} y, \\ z \end{pmatrix}\rangle + \delta h^+(y, \xi, \varepsilon),
\end{aligned}$$

$$\begin{aligned}
P^+ &= \frac{1}{\delta\varepsilon^{\bar m_\alpha}}\bar P^+ = \frac{1}{\delta\varepsilon^{\bar m_\alpha}}\int_0^1 \{Q, F\} \circ \phi_F^t \mathrm{d}t + \frac{1}{\delta\varepsilon^{\bar m_\alpha}} Q\\
&\quad + \int_0^1 \{(1-t)(R - [R] + \langle p_{001}, z\rangle) + R, F\} \circ \phi_F^t \mathrm{d}t + (P - R)\circ\phi_F^1,
\end{aligned} \tag{3.31}$$

where $\bar{P}^+$ is defined as (3.9), e_ε^+ is a smooth function on Λ^+ and

$$e_\varepsilon^+(\xi) = e_\xi(\xi) + \delta\varepsilon^{\bar{m}_\alpha} p_{000}, \tag{3.32}$$

$$\omega_\varepsilon^+(\xi) = \omega_\varepsilon(\xi) + \delta\varepsilon^{\bar{m}_\alpha} p_{010}, \tag{3.33}$$

$$M^+(\xi,\varepsilon) = \begin{pmatrix} O & M_{12} \\ M_{21} & M_{22} \end{pmatrix} + 2\varepsilon^{\bar{m}_\alpha - \bar{m}_{I+1}} \begin{pmatrix} O & p_{011} \\ p_{011}^\top & p_{002} \end{pmatrix}, \tag{3.34}$$

$$\begin{aligned} h^+(y,z,\xi,\varepsilon) &= h_1(y,\xi,\varepsilon) + \varepsilon^{\bar{m}_\alpha}\langle p_{020}y, y\rangle + h_2(y,z,\xi,\varepsilon) \\ &= O(|y^0|^2 + \cdots + \varepsilon^{\bar{m}_{I+1}}|y^{I+1}|^2) + h_2(y,z,\xi,\varepsilon). \end{aligned} \tag{3.35}$$

Rewrite ω_ε^+ as

$$\omega_\varepsilon^+ := (\hat{\omega}_\varepsilon^{+,0}, \cdots, \varepsilon^{\bar{m}_{I+1}}\hat{\omega}_\varepsilon^{+,I+1})^\top,$$

where $\hat{\omega}_\varepsilon^{+,j}$ is an $n_j - n_{j-1}$ dimensional vector for each $j = 0, \cdots, I$ and $\hat{\omega}_\varepsilon^{+,I+1}$ consists of the last $m - n_I$ components of ω_ε^+. Let

$$\Omega_\varepsilon^+(\xi) = (\hat{\omega}_\varepsilon^{+,0}(\xi), \cdots, \hat{\omega}_\varepsilon^{+,I+1}(\xi))^\top.$$

The following lemmas give estimates of the new Hamiltonian.

Lemma 3.4. *There exits a constant $c_4 > 0$ such that the followings hold for all $0 \le |l| \le N$:*

$$\begin{aligned} |\partial_\xi^l(e_\varepsilon^+ - e_\varepsilon)|_{\Lambda^+} &\le c_4\delta\varepsilon^{\bar{m}_\alpha}\gamma s^2\mu, \\ |\partial_\xi^l(\omega_\varepsilon^+ - \omega_\varepsilon)|_{\Lambda^+} &\le c_4\delta\varepsilon^{\bar{m}_\alpha}\gamma s\mu, \\ |\partial_\xi^l(\Omega_\varepsilon^+ - \Omega_\varepsilon)|_{\Lambda^+} &\le c_4\varepsilon^{\bar{m}_\alpha - \bar{m}_{I+1}}\gamma s\mu, \\ |\partial_\xi^l(M^+ - M)|_{\Lambda^+} &\le c_4\varepsilon^{\bar{m}_\alpha - \bar{m}_{I+1}}\gamma\mu, \\ |\partial_\xi^l(h^+ - h)|_{\Lambda^+} &\le c_4\varepsilon^{\bar{m}_\alpha}\gamma\mu. \end{aligned}$$

Proof. It follows from (3.32)-(3.35). □

Lemma 3.5. *Assume that*

(H6) $c_5\gamma s\mu K_+^{\tau+1} \le \gamma - \gamma_+$.

Then, for all $0 < |k| \le K_+$, $\xi \in \Lambda^+$,

$$|\langle k, \Omega_\varepsilon^+(\xi)\rangle| > \frac{\gamma_+}{|k|^\tau}. \tag{3.36}$$

Proof. It follows from (3.33) that there is a constant $c_4 > 0$ such that for all $|l| \le N$,

$$|\partial_\xi^l(\Omega_\varepsilon^+(\xi) - \Omega_\varepsilon(\xi))|_{\Lambda^+} \le c_4\varepsilon^{\bar{m}_\alpha - \bar{m}_{I+1}}\gamma s\mu.$$

This, together with **(H6)**, implies that

$$|\langle k, \Omega_\varepsilon^+(\xi)\rangle| \ge |\langle k, \Omega_\varepsilon(\xi)\rangle| - |\langle k, (\Omega_\varepsilon^+(\xi) - \Omega_\varepsilon(\xi))| \ge \frac{\gamma_+}{|k|^\tau},$$

for all $0 < |k| \le K_+,\ \xi \in \Lambda^+$. □

Lemma 3.6. *Let*

$$\Delta = \gamma^b(s_+^2 s\mu\Gamma(r - r_+) + s_+^2 s\mu + s^2\mu^2 + s^2\mu^2\Gamma^2(r - r_+))$$

and assume that

(H7) $c_0\Delta \le \gamma_+^b s_+^2\mu_+$, *where* $c_0 = \max\{1, c_1, \cdots, c_7\}$. *Then*

$$|\partial_\xi^l P^+|_{D_+\times\Lambda^+} \le \gamma_+^b s_+^2 \mu_+, \qquad |l| \le N.$$

Proof. Combing estimates (3.8), (3.24) - (3.26), (3.28) with Lemma 3.2, we have

$$\begin{aligned} |Q|_{D_+\times\Lambda^+} &\le \delta\varepsilon^{\bar m_{I+1}}|\frac{\partial h_2}{\partial z}||J(f_{k01} + f_{k11}y + f_{k02}z + f_{k02}^\top z)e^{\sqrt{-1}\langle k,x\rangle}| \\ &\quad +\delta\varepsilon^{\bar m_{I+1}}||M_{12}\frac{\partial h_2}{\partial z}||p_{001}| \\ &\le c\delta\varepsilon^{\bar m_\alpha}\gamma^b(s_+^2 s\mu\Gamma(r-r_+) + s_+^2 s\mu). \end{aligned}$$

It follows that there exists a positive constant c_6 such that the following estimates hold for any $|l| \le N$:

$$\frac{1}{\delta\varepsilon^{\bar m_\alpha}}|\partial_\xi^l Q|_{D_+\times\Lambda^+} \le c_6(\gamma^b s_+^2 s\mu\Gamma(r-r_+) + \gamma^b s_+^2 s\mu),$$

$$\frac{1}{\delta\varepsilon^{\bar m_\alpha}}|\int_0^1 \partial_\xi^l\{Q, F\}\circ\phi_F^t \mathrm{d}t|_{D_+\times\Lambda^+} \le c_6\gamma^b s^2\mu^2\Gamma^2(r-r_+).$$

By Lemmas 3.1-3.2, there exits a positive constant c_7 such that for any $|l| \le N$,

$$|\partial_\xi^l(P-R)\circ\phi_F^1|_{D_+\times\Lambda^+} \le c_7\gamma^b s^2\mu^2,$$

$$|\partial_\xi^l\int_0^1\{(1-t)(R-[R]+\langle p_{001}, z\rangle) + R, F\}\circ\phi_F^t \mathrm{d}t|_{D_+\times\Lambda^+} \le c_7\gamma^b s^2\mu^2\Gamma^2(r-r_+).$$

The lemma now follows from above estimates and the definition of P^+ in (3.32). □

3.5. **Improved Hamiltonian normal form.** In this subsection, we will improve the Hamiltonian (3.1) by performing the quasi-linear iterative scheme from $\nu = 0$ to some step ν^* in order to obtain a new Hamiltonian H^* with the perturbation P^* of a sufficiently high order. Tracing quantities defined at the beginning of Section 3, we have the following iterative sequences:

$$\begin{aligned} r_\nu &= r_0(1 - \frac{1}{4}\sum_{i=1}^{\nu}\frac{1}{2^{i+1}}), \\ s_\nu &= \frac{1}{8}\alpha_{\nu-1}s_{\nu-1},\ \alpha_\nu = \mu_\nu^{\frac{1}{3}}, \\ \mu_\nu &= \mu_{\nu-1}^{1+\hat\iota}, \quad \text{for some fixed } \hat\iota \in (0,\iota), \\ K_\nu &= ([\log(\frac{1}{\mu_{\nu-1}})] + 1)^3, \\ \Lambda^\nu &= \{\xi \in \Lambda^{\nu-1} : |\langle k, \Omega_\varepsilon^{\nu-1}(\xi)\rangle| > \frac{\gamma_{\nu-1}}{|k|^\tau},\ 0 < |k| \le K_\nu\} \end{aligned}$$

for $\nu = 1, 2, \cdots$.

It is easy to verify that

$$\mu_\nu = \mu_0^{(1+\hat\iota)^\nu} \le \varepsilon^{\frac{(1-\iota)(1+\hat\iota)^\nu}{6}}, \quad \nu = 1, 2, \cdots. \tag{3.37}$$

The assumptions **(H1)**, **(H4)-(H7)** and part of **(H2)** that

$$\delta s_{\nu-1}K_\nu^{\tau+1} \le s_{\nu-1}K_\nu^{\tau+1} = o(\gamma_0)$$

are easily seen to hold for all $\nu = 1, 2, \cdots$ as long as ε is sufficiently small. However, the other part of **(H2)**, i.e.,

$$\varepsilon K_\nu^{\tau+1} = o(\gamma_0) \tag{3.38}$$

can only hold for a finite number of ν's. More precisely, let

$$\nu_* = \left[\frac{\log(2(N+6)c_* + \log 8d^2(N+1)+1) - \log(1-\iota)}{\log(1+\hat{\iota})}\right] + 1, \tag{3.39}$$

where $[x]$ denotes the maximal integer less than x and $c_* = \sum_{i=1}^{\alpha} \bar{m}_i(n_i - n_{i-1})$. A simple computation shows that (3.38) holds as long as $(1+\hat{\iota})^{\nu_*}(1-\iota)\varepsilon^{1-2\iota}$ is bounded from the above by a constant that is independent of ν_*, i.e., **(H2)** holds for all $\nu = 1, 2, \cdots, \nu_*$ as long as ε is sufficiently small. Finally, **(H3)** holds for all ν because

$$\begin{aligned} & |\partial_\xi^l (M^\nu(\xi,\varepsilon) - M^0(\xi,\varepsilon))|_{\Lambda^\nu} \\ \le\ & |\partial_\xi^l (M^\nu(\xi,\varepsilon) - M^{\nu-1}(\xi,\varepsilon))|_{\Lambda^\nu} + \cdots + |\partial_\xi^l (M^1(\xi,\varepsilon) - M^0(\xi,\varepsilon))|_{\Lambda^1} \\ \le\ & \gamma_0^b \mu_0^{\frac{1}{4}} (\frac{1}{2^\mu} + \cdots + 1) \le \mu_0^{\frac{1}{4}}. \end{aligned}$$

With the validity of these assumptions, the quasi-linear iterative scheme can be performed inductively from step $\nu = 0$ to the step ν_*. When $\nu = \nu_*$, we obtain the following improved Hamiltonian normal form

$$\begin{aligned} H^* &=: H^{\nu_*} \\ &= e_\varepsilon^*(\xi) + \langle \omega_\varepsilon^*(\xi), y\rangle + \frac{\delta}{2}\langle \begin{pmatrix} y \\ z \end{pmatrix}, M^*(\xi,\varepsilon)\begin{pmatrix} y \\ z \end{pmatrix}\rangle + \delta h^*(y,\xi,\varepsilon) + \delta P^*(x,y,z,\xi,\varepsilon) \end{aligned} \tag{3.40}$$

defined on $D(r_*, s_*) \times \Lambda^*$, where $r_* = r_{\nu_*}$, $s_* = s_{\nu_*}$, $e_\varepsilon^* = e_\varepsilon^{\nu_*}$, $\omega_\varepsilon = \omega_\varepsilon^{\nu_*}$, $h^* = h^{\nu_*}$, $M^* = \varepsilon^{\bar{m}_{I+1}} M^{\nu_*}$, $P^* = \varepsilon^{\bar{m}_\alpha} P^{\nu_*}$, $\Lambda^* = \Lambda^{\nu_*}$.

By (3.39) and Lemma 3.6, we have

$$|\partial_\xi^l P^*|_{D(r_*,s_*)\times\Lambda^*} \le \varepsilon^{\bar{m}_\alpha+\iota} s_*^2 \varepsilon^{16d^2(N+1)(N+6)c_*} \le \gamma_*^{16d^2(N+1)(N+6)} s_*^2 \mu_*^2, \qquad |l| \le N, \tag{3.41}$$

where $\gamma_* = \varepsilon^{c_*}$, $\mu_* = \varepsilon^{\frac{\bar{m}_\alpha+\iota}{2}}$.

3.6. **Measure estimate.** By **A3)** and (2.9), we have

$$\mathrm{Rank}\{\partial_\xi^l \Omega_\varepsilon^0 :\ 0 \le |l| \le N\} = m, \quad \forall \xi \in \Lambda^0.$$

It follows from Lemma 3.5 that

$$\mathrm{Rank}\{\partial_\xi^l \Omega_\varepsilon^i :\ 0 \le |l| \le N\} = m, \quad \forall \xi \in \Lambda^i,\ i = 1, \cdots, \nu_*.$$

Using Lemma 3.5 and the standard measure estimate under Rüssmann condition (see [22] or Lemma 4.2 of the present paper), we have

$$|\Lambda^0 \setminus \Lambda^*| = \sum_{i=1}^{\nu_*} |\Lambda^{i-1} \setminus \Lambda^i| \le \sum_{i=1}^{\nu_*} \sum_{K_i \le |k| \le K_{i+1}} (\frac{\gamma_i}{|k|^\tau})^{\frac{1}{N}} = O(\gamma_0^{\frac{1}{N}}) = O(\varepsilon^{\frac{\iota(\bar{m}_{\alpha+1}-\bar{m}_\alpha)}{2bN}}).$$

This, together with (2.10), yields that

$$|\mathcal{O}(g,G) \setminus \Lambda^*| = O(\varepsilon^{\frac{\iota}{2bN}}). \tag{3.42}$$

4. Proof of the Main Theorem

In this section, we will perform an infinite steps of standard KAM iterations to the improved Hamiltonian normal form (3.40) to prove the Main Theorem. First, we consider the following rescalings

$$y \to \gamma_*^{8d^2(N+1)(N+6)} s_* \mu_* y, \quad z \to \gamma_*^{8d^2(N+1)(N+6)} s_* \mu_* z, \quad H^* \to \frac{H^*}{\gamma_*^{8d^2(N+1)(N+6)} s_* \mu_*}$$

to the normal form (3.40). The re-scaled Hamiltonian reads

$$H_0 =: \frac{H^*}{\gamma_*^{8d^2(N+1)(N+6)} \mu_*} =: e_0(\xi,\varepsilon) + \langle \omega_0(\xi,\varepsilon), y\rangle + \delta P_0(x,y,z,\xi,\varepsilon)$$

which is defined on region $D(r_0,s_0)\times\Lambda_0$, where $r_0 =: r_*,\ s_0 =: s_*,\ \Lambda_0 = \Lambda^*,\ e_0(\cdot,\varepsilon) = e^*_\varepsilon$, $\omega_0(\cdot,\varepsilon) = \omega^*_\varepsilon$, and

$$P_0 = \frac{\langle \binom{y}{z}, M^* \binom{y}{z}\rangle + h^*(y,\xi,\varepsilon) + P^*}{\gamma_*^{8d^2(N+1)(N+6)} s_* \mu_*}.$$

It follows from (3.41) that

$$\begin{aligned} |\partial_\xi^l P_0|_{D(r_0,s_0)\times\Lambda_0} &\le \frac{|\partial_\xi^l P^*| + |\langle \binom{y}{z}, \partial_\xi^l M^* \binom{y}{z}\rangle| + |\partial_\xi^l h^*(y,z,\xi,\varepsilon)|}{\gamma_*^{8d^2(N+1)(N+6)} s_* \mu_*} \\ &\le \gamma_*^{8d^2(N+1)(N+6)} s_* \mu_*, \qquad |l| \le N. \end{aligned}$$

Denote $\gamma_0 =: \gamma_*^{2(N+6)},\ \mu_0 =: \mu_*,\ a =: 4d^2(N+1)$. We then have

$$|\partial_\xi^l P_0|_{D(r_0,s_0)\times\Lambda_0} \le \gamma_0^a s_0 \mu_0, \qquad |l| \le N.$$

4.1. Iteration and convergence.

Consider the following sequences

$$\begin{aligned} & r_\nu = r_0(1 - \sum_{i=1}^{\nu} \frac{1}{2^{i+1}}), \\ & s_\nu = \frac{1}{8}\alpha_{\nu-1} s_{\nu-1}, \\ & \alpha_\nu = \mu_\nu^{\frac{1}{2}}, \\ & \mu_\nu = c_0 \mu_{\nu-1}^{\frac{6}{5}}, \\ & \gamma_\nu = \gamma_0(1 - \sum_{i=1}^{\nu} \frac{1}{2^{i+1}}), \\ & K_\nu = ([\log(\frac{1}{\mu_{\nu-1}})] + 1)^{3\eta}, \\ & \Lambda_\nu = \{\xi \in \Lambda_{\nu-1} : |\langle k, \omega_{\nu-1}\rangle| > \frac{\gamma_{\nu-1}}{|k|^\tau}, 0 < |k| \le K_\nu\}, \end{aligned}$$

$\nu = 1, 2, \cdots$, where $\eta \ge \frac{\log 2}{\log 6 - \log 5}$ is a fixed constant.

We have the following iteration lemma.

Lemma 4.1. *Let ε be sufficiently small. Then the followings hold for all $\nu = 1, 2, \cdots$.*

1) *There is a sequence of smooth families of symplectic, real analytic, near identity transformations*
$$\Phi^{\nu}_{\xi} : D(r_\nu, s_\nu) \to D(r_{\nu-1}, s_{\nu-1}), \qquad \xi \in \Lambda_\nu$$
such that
$$H_\nu \quad = \quad H_{\nu-1} \circ \Phi^{\nu}_{\xi} =: e_\nu(\xi,\varepsilon) + \langle \omega_\nu(\xi,\varepsilon), y\rangle + \delta P_\nu,$$
where
$$|\partial^l_\xi \omega_\nu - \partial^l_\xi \omega_0|_{\Lambda_\nu} \le \gamma_0^a \mu_0,$$
$$|\partial^l_\xi P_\nu|_{D_\nu \times \Lambda_\nu} \le \gamma_\nu^a s_\nu \mu_\nu$$
for all $|l| \le N$.
2) $\Lambda_\nu = \{\xi \in \Lambda_{\nu-1} : |\langle k, \omega_{\nu-1}\rangle| > \dfrac{\gamma_{\nu-1}}{|k|^\tau}, K_{\nu-1} < |k| \le K_\nu\}$.
3) *The Whitney extensions of*
$$\Psi^\nu =: \Phi^1_\xi \circ \Phi^2_\xi \circ \cdots \circ \Phi^\nu_\xi$$
converge C^N *uniformly to a smooth family of symplectic maps* Ψ^∞ *on* $D(\frac{r_0}{2}, 0) \times \Lambda_\infty$*, where*
$$\Lambda_\infty = \bigcap_{\nu \ge 0} \Lambda_\nu,$$
so that
$$H_\nu = H_0 \circ \Psi^{\nu-1} \to H_\infty =: H_0 \circ \Psi^\infty = e_\infty + \langle \omega_\infty, y\rangle + \delta P_\infty,$$
where $e_\infty = \lim_{\nu\to\infty} e_\nu$*,* $\omega_\infty = \lim_{\nu\to\infty} \omega_\nu$*, and* $P_\infty = \lim_{\nu\to\infty} P_\nu$ *satisfies*
$$\partial^j_{(y,z)} P_\infty|_{D(\frac{r_0}{2},0)\times\Lambda_\infty} = 0, \qquad |j| \le 2.$$

Proof. It follows from arguments similar to those in [6, Section 4]. □

The iteration lemma above shows that, for sufficiently small $\varepsilon > 0$ and each $\xi \in \Lambda_\infty$, $\mathbb{T}^m \times \{0\}$ is a real analytic, invariant, Diophantine torus of H_∞ of Diophantine type (γ_∞, τ), where $\gamma_\infty = \lim_{\nu\to\infty} \gamma_\nu$. Moreover, these m-tori form a Whitney smooth family.

4.2. **Measure estimate of Λ_∞.** Let $\mathcal{O}_\varepsilon = \Lambda_\infty$. We now estimate the measure $|\mathcal{O}(g,G) \setminus \mathcal{O}_\varepsilon|$.

Lemma 4.2. ([22, Lemma 2.1]) *Suppose that* $g(x)$ *is a* p*-times differentiable function on the closure* $\bar{I} \subset I$*, where* I *is a finite open interval. Let* $I_h = \{x :\ |g(x)| \le h,\ x \in I\},\ h > 0$*. If on* I*,* $|g^{(p)}(x)| \ge D > 0$*, where* D *is a constant, then* $|I_h| \le c_7 h^{\frac{1}{p}}$*, where* $c_7 = 2(2+3+\cdots+p+D^{-1})$.

For each $\nu = 0, 1, \cdots$ and $k \in \mathbb{Z}^m \setminus \{0\}$, denote
$$R^{\nu+1}_k \quad = \quad \{\xi \in \Lambda_\nu :\ |\sqrt{-1}\langle k, \omega_\nu\rangle| \le \frac{\gamma_\nu}{|k|^\tau}\}.$$

Then
$$\Lambda_0 \setminus \Lambda_\infty = \bigcup_{\nu=0}^{\infty} \bigcup_{K_\nu < |k| \le K_{\nu+1}} R^{\nu+1}_k(\xi). \tag{4.1}$$

Consider functions
$$g^\nu_{k,0}(\xi) \quad = \quad \langle \frac{k}{|k|}, \omega_\nu(\xi,\varepsilon)\rangle.$$

By Lemma 4.1 1) and Lemma 4.2, it is easy to see that there is a positive constant c_8 such that

$$|\frac{\partial^N g_{k,0}^\nu}{\partial \xi^N}|_{\Lambda^\nu} \geq c_8 \varepsilon^{c_*},$$

where $c_* = \sum_{i=1}^{\alpha} \bar{m}_i (m_i - m_{i-1})$. It follows from Lemma 4.2 that there exists a positive constant c_9 independent of ε such that

$$|R_k^{\nu+1}| \leq c(2+3+\cdots+\frac{1}{c_8\varepsilon^{c_*}})\frac{\varepsilon^{\frac{2(N+6)c_*}{N}}}{|k|^{\frac{\tau}{N}}} \leq c_9 \frac{\varepsilon^{\frac{12c_*}{N}}}{|k|^{\frac{\tau}{N}}}.$$

It follows that

$$|\Lambda_0 \setminus \Lambda_\infty| \leq \sum_{\nu=0}^{\infty} \sum_{K_\nu < |k| \leq K_{\nu+1}} |R_k^{\nu+1}| \leq c_9 \varepsilon^{\frac{12c_*}{N}} \sum_{\nu=0}^{\infty} \sum_{K_\nu < |k| \leq K_{\nu+1}} \frac{1}{|k|^{\frac{\tau}{N}}} = O(\varepsilon^{\frac{12c_*}{N}}).$$

Recall that $\Lambda_0 = \Lambda_*$. Combining the above with (3.42) yields that

$$|\mathcal{O}(g,G) \setminus \mathcal{O}_\varepsilon| \leq |\mathcal{O}(g,G) \setminus \Lambda^*| + |\Lambda_0 \setminus \Lambda_\infty| \leq O(\varepsilon^{\frac{12c_*}{N}}) + O(\varepsilon^{\frac{\iota}{2bN}}) = O(\varepsilon^{\frac{\iota}{2bN}}).$$

Now, tracing back all the symplectic transformations involved in Sections 2-4, proof of the Main Theorem is completed.

5. An Example

In [14, Section 4.4], the following normalized Hamiltonian of 3-degree-of-freedom is derived from the orbiting dust problem which describes the effect of radiation pressure on dust particles revolving around an idealized planet in a planar circular orbit around a star:

$$(5.1)\quad \begin{aligned} H_\varepsilon \;=\; & -Y_0 - \frac{\varepsilon^3}{2Y_0^6}(Y_1+Y_2)(5(Y_1+Y_2)^3 - 4Y_0(Y_1+Y_2)^2 + 3Y_0^2(Y_1+Y_2) - 2Y_0^3) \\ & - \frac{3\varepsilon^{13}\mu^2}{8Y_0^8}(29Y_1^2 + 20Y_1Y_2 + 4Y_2^2 - 2Y_0(3Y_1+Y_2)) + \varepsilon^{17}P(Y,\varphi,\varepsilon), \end{aligned}$$

where $Y = (Y_0, Y_1, Y_2)^\top \in \mathbb{R}^3$, $Y_0 > 0$, $\varphi \in \mathbb{T}^3$, and μ is a fixed constant. It is real analytic in a neighborhood of a relative equilibrium and depends on ε smoothly. The existence of a positive measure set of quasi-periodic invariant 3-tori of (5.1) is shown in [14, Section 4.4] based on the main result of [6].

We now examine the possible existence of lower dimensional tori of (5.1) in the resonance zone by showing the validity of assumptions **A2)**, **A3)** with respect to certain resonant type. Thus the Main Theorem is applicable to this example if the assumption **A4)** holds for certain particular perturbation term. As remarked at the end of Section 1, the condition **A4)** is satisfied by generic perturbations. It is also well-expected that it can be largely satisfied by particular perturbations which are not highly degenerate because of the freedom of choosing resonance types in verifying this condition.

Rewrite the Hamiltonian as

$$(5.2)\qquad H_\varepsilon = H_0(Y_0) + \varepsilon^3 H_1(Y) + \varepsilon^{13} H_2(Y) + \varepsilon^{17} P(Y,\varphi,\varepsilon),$$

where

$$\begin{aligned} H_0(Y_0) &= -Y_0, \\ H_1(Y) &= -\frac{1}{2Y_0^6}(Y_1+Y_2)(5(Y_1+Y_2)^3 - 4Y_0(Y_1+Y_2)^2 + 3Y_0^2(Y_1+Y_2) - 2Y_0^3), \\ H_2(Y) &= -\frac{3\mu^2}{8Y_0^8}(29Y_1^2 + 20Y_1Y_2 + 4Y_2^2 - 2Y_0(3Y_1+Y_2)). \end{aligned}$$

We restrict $Y \in G$ in (5.2), where G is a given bounded closed region in $(0,\infty)\times\mathbb{R}^2$. Then Hamiltonian (5.2) is a special case of (1.1) with with $n_0 = 1$, $n_1 = n_2 = 3$, $\bar{m}_1 = 3$, $\bar{m}_2 = 13$, $\bar{m}_3 = 17$. Since $n_0 < n_1$, **A1)** holds with $I = 0$.

Denote

$$\begin{aligned} \omega_\varepsilon(Y) &= (\frac{\partial H_0}{\partial Y_0} + \varepsilon^3\frac{\partial H_1}{\partial Y_0} + \varepsilon^7\frac{\partial H_2}{\partial Y_0}, \varepsilon^3\frac{\partial H_1}{\partial Y_1} + \varepsilon^7\frac{\partial H_2}{\partial Y_1}, \varepsilon^3\frac{\partial H_1}{\partial Y_2} + \varepsilon^7\frac{\partial H_2}{\partial Y_2})^\top, \\ \Omega^*(Y) &= (\frac{\partial H_0}{\partial Y_0}, \frac{\partial H_1}{\partial Y_1}, \frac{\partial H_1}{\partial Y_2}) =: (\Omega_0^*, \Omega_1^*, \Omega_2^*)^\top. \end{aligned}$$

Motivated by the fact that $\frac{\partial H_0}{\partial Y_0} \equiv -1$, the resonance of ω_ε cannot occur at the first integrable term H_0. Since $\frac{\partial H_1}{\partial Y_1} = \frac{\partial H_1}{\partial Y_2}$ and $n_1 = n_2 = 3$, we consider the resonant subgroup

$$g = \{0\} \oplus \hat{g}_1,$$

where $\hat{g}_1$ is a subgroup of $\mathbb{Z}^3$ spanned by $K_2 = (0, 1, -1)^\top$. Denote

$$K_1 = \begin{pmatrix} 1 & 0 \\ 0 & 0 \\ 0 & -1 \end{pmatrix}.$$

Then $\det(K_1, K_2) = 1$ and the g-resonant surface in G reads

$$\mathcal{O}(g, G) = \{Y \in G : K_2^\top \Omega_*(Y) = 0\} = \{I \in G : \ 38Y_1 + 12Y_2 - 4Y_0 = 0\}.$$

To verify the g-non-degenerate condition **A2)** for H_2 on $\mathcal{O}(g, G)$, we note that

$$\begin{aligned} &\frac{\partial^2 H_1}{\partial Y_1^2} = \frac{\partial^2 H_1}{\partial Y_1 \partial Y_2} = \frac{\partial^2 H_1}{\partial Y_2^2}, \\ &\frac{\partial^2 H_2}{\partial Y_1^2} = -\frac{174\mu^2}{8Y_0^8}, \ \frac{\partial^2 H_2}{\partial Y_1\partial Y_2} = -\frac{60\mu^2}{8Y_0^8}, \\ &\frac{\partial^2 H_2}{\partial Y_2 \partial Y_1} = -\frac{60\mu^2}{8Y_0^8}, \ \frac{\partial^2 H_2}{\partial Y_2^2} = -\frac{24\mu^2}{8Y_0^8}. \end{aligned} \tag{5.3}$$

It yields that

$$\det \hat{K}_2^\top \frac{\partial^2 H_2}{\partial(\hat{Y}^2)^2}\hat{K}_2 = \frac{\partial^2 H_2}{\partial Y_1^2} + \frac{\partial^2 H_2}{\partial I_2^2} - 2\frac{\partial^2 H_2}{\partial Y_1 Y_2} \neq 0, \quad \forall Y \in \mathcal{O}(g, G),$$

where $\hat{Y}^2 = (Y_1, Y_2)^\top$. Hence H_2 is g-non-degenerate on $\mathcal{O}(g, G)$, i.e., **A2)** holds.

To verify the degeneracy removing condition on $\mathcal{O}(g,G)$, we note that $K_1^\top\Omega^* = (\Omega_0^*, -\Omega_2^*)^\top =: (\Omega_0, \Omega_1)^\top$. By (5.3), it is easy to verify that

$$\text{Rank}\begin{pmatrix} \Omega_0 & \Omega_1 \\ \dfrac{\partial\Omega_0}{\partial Y_0} & \dfrac{\partial\Omega_1}{\partial Y_0} \\ \dfrac{\partial\Omega_0}{\partial Y_1} & \dfrac{\partial\Omega_1}{\partial Y_1} \\ \dfrac{\partial\Omega_0}{\partial Y_2} & \dfrac{\partial\Omega_1}{\partial Y_2} \end{pmatrix} \equiv 2$$

on $\mathcal{O}(g,G)$, i.e., **A3)** holds for $N=1$.

Summarizing up the above, we have the following result.

Corollary. *Consider the Hamiltonian* (5.1) *and the rank-*1 *subgroup g spanned by* $(0,1,-1)^\top$. *If* **A4)** *holds for the perturbation P on the resonant surface $\mathcal{O}(g,G)$, then the* Main Theorem *is applicable to yield a family of quasi-periodic, invariant* 2*-tori of the Hamiltonian* (5.1).

Remark 5.1. *Though a validation of* **A4)** *depends on the particular expression of $P(Y,\theta,0)$, one can give more specific validation of this condition in term of Fourier series*

$$P(I,\theta,0) = \sum_{k=(k_1,k_2,k_3)\in\mathbb{Z}^3} P_{k_1,k_2,k_3}(I)e^{\sqrt{-1}\langle k,\theta\rangle}.$$

Using the form of $K_1,\ K_2$, we have

$$\varphi = \begin{pmatrix}\varphi_1\\ \varphi_2\end{pmatrix} = K_1^\top\theta = \begin{pmatrix}\theta_0\\ -\theta_2\end{pmatrix} \in \mathbb{T}^2, \qquad \psi = K_2^\top\theta = \theta_1-\theta_2 \in \mathbb{T}.$$

Hence the function $h_0:\ \mathbb{T}\times\mathcal{O}(g,G)\to\mathbb{R}$ can be expressed as

$$\begin{aligned} h_0(\psi,Y) &= \int_{\mathbb{T}^2}\tilde{P}(\varphi,\psi,Y)\mathrm{d}\varphi = \int_{\mathbb{T}^2} P(\varphi_1,\psi-\varphi_2,-\varphi_2,Y,0)\mathrm{d}\varphi \\ &= \sum_{k\in\mathbb{Z}^3} P_{k_1,k_2,k_3}(Y)\int_{\mathbb{T}^2} e^{\sqrt{-1}(k_1\varphi_1-(k_2+k_3)\varphi_2+k_2\psi)}\mathrm{d}\varphi \\ &= \sum_{k_1=0,k_2+k_3=0} P_{k_1,k_2,k_3}(I)e^{\sqrt{-1}k_2\psi} = \sum_{j\in\mathbb{Z}} P_{0,j,-j}(Y)e^{\sqrt{-1}j\psi}. \end{aligned}$$

Thus, to verify **A4)***, one only needs to find, for a fixed $Y_0\in\mathcal{O}(g,G)$, a non-degenerate critical point φ_0 of $h_0(\cdot,Y_0)$, because the Implicit Function Theorem then implies the existence of a neighborhood U_{Y_0} of Y_0 and a real analytic family $\psi(Y)$ of non-degenerate critical points of $h_0(\cdot,Y)$ for $Y\in U_{Y_0}$.*

Remark 5.2. *After verifying* **A4)***, it then follows from the* Main Theorem *that there exists a Cantor subset $O_\varepsilon\subset\mathcal{O}_0(g,G) := \mathcal{O}(g,G)\cap U_{Y_0}$ of positive Lebesgue measure such that the unperturbed quasi-periodic* 2*-tori $T_Y^\varepsilon(\psi(Y)) = \mathbb{T}^2\times\{Y\}\times\{\psi(Y)\},\ Y\in O_\varepsilon$, of the Hamiltonian* (5.1)*, persist as ε sufficiently small. To be more precise, by taking the the constants $n_0=1, n_1=n_2=3,\ \bar{m}_1=3,\ \bar{m}_2=13,\ \bar{m}_3=17,\ N=1$, we have*

$$c_* = \sum_{i=1}^{2} m_i(n_i-n_{i-1}) = 6$$

and hence $2(N+6)c_* = 84 > 17$. This means that one needs to perform a finite step of iterations to push the order of the perturbation as higher as $O(\varepsilon^{84})$ in order to carry out the standard KAM iterations. It also follows from the Main Theorem *that the excluding measure has the estimate $|\mathcal{O}_0(g,G)\setminus O_\varepsilon| = O(\varepsilon^{\frac{1}{96}})$ by taking $\iota=\frac{1}{4},\ N=1,\ d=1$.*

References

[1] V. I. Arnold, Small denominators and problems of stability of motion in classical mechanics, *Usp. Math. Nauk.* **18** (6) (1963), 91-192.

[2] L. Biasco, L. Chierchia, and E. Valdinoci, N-dimensional elliptic invariant tori for the planar $(N+1)$-body problem, *SIAM J. Math. Anal.* **37** (2006), 2560-2588.

[3] H. W. Broer and H. Hannnβmann, On Jupiter and his Galilean satellites: Librations of De Sitters periodic motions, *Indagationes Math.* **27** (2016), 1305-1336.

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

[5] J. Féjoz, Quasiperiodic motions in the planar three-body problem, *J. Differential Equations* **183** (2002), 303-341.

[6] Y. Han, Y. Li, and Y. Yi, Invariant tori for Hamiltonian systems with high order proper degeneracy, *Ann. Henri Poincaré* **10** No. 8 (2010), 1419-1436.

[7] M. Kummer, On the regularization of the Kepler problem, *Comm. Math. Phys.* **84** (1982), 133-152.

[8] M. Inarrea, V. Lanchares, J. F. Palacián, A. I. Pascual, J. P. Salas, and P. Yanguas, Reduction of some perturbed Keplerian problems, *Chaos Solitons Fractals* **27** (2006), 527-536.

[9] Y. Li and Y. Yi, A quasi-periodic Poincare's Theorem, *Math. Annalen,* **326** (2003), 649-690.

[10] Y. Li and Y. Yi, On Poincaré-Treshchev tori in Hamiltonian systems, *Proc. Equadiff 2003*, Dumortier et al (Ed.), World Scientific, 2005, 136-151.

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

[12] M. Mazzocco, KAM theorem for generic analytic perturbations of the Euler system, *Z. Angew. Math. Phys.* **48** (1997), 193-219.

[13] K. R. Meyer, Scaling Hamiltonian systems, *SIAM J. Math. Anal.* **15** (1984), 877-889.

[14] K. R. Meyer, J. F. Palacián, and P. Yanguas, Geometric Averaging of Hamiltonian Systems: Periodic Solutions, Stability, and KAM Tori, *SIAM J. Appl. Dyn. Syst.,* **10(3)** (2011), 817C856.

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

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

[17] H. Rüssmann, Nondegeneracy in the perturbation theory of integrable dynamical systems, *Number theory and dynamical systems* (York, 1987), 5-18, London Math. Soc. Lecture Note Ser., 134, Cambridge Univ. Press, Cambridge, 1989, *Stochastics, algebra and analysis in classical and quantum dynamics* (Marseille, 1988), 211-223, Math. Appl., 59, Kluwer Acad. Publ., Dordrecht, 1990.

[18] B. Sommer, A KAM Theorem for the Spatial Lunar Problem, Ph.D thesis, 2003.

[19] D. V. Treshchev, The mechanism of destruciton of resonant tori of Hamiltonian systems, *Math. USSR Sb.* **68** (1991), 181-203.

[20] L. Xu, Y. Li and Y. Yi, Lower Dimensional Tori in multi-scale, nearly Integrabel Hamiltonian systems, *Ann. Henri Poincaré*, **18** (2017), 53-83.

[21] J. Xu and J. You, Corrigendum for the paper Invariant tori for nearly integrable Hamiltonian systems with degeneracy, *Math. Z.* **257** (2007), 939.

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

L. Xu: School of Mathematics, Jilin University, Changchun, 130012, P.R.China, and School of Mathematics, Georgia Institute of Technology, Atlanta, GA 30332, USA

E-mail address: xulu@jlu.edu.cn

Y. Li: School of Mathematics and Statistics, and Center of Mathematics and Interdisciplinary Sciences, Northeast Normal University, Changchun, 130024, P. R. China

E-mail address: liyong@jlu.edu.cn

Y. Yi (Corresponding author): Department of Mathematical & Statistical Sci, University of Alberta, Edmonton Alberta, Canada T6G 2G1 and School of Mathematics, Jilin University, Changchun, 130012, PRC

E-mail address: yingfei@ualberta.ca,yi@math.gatech.edu