LOWER DIMENSIONAL TORI IN MULTI-SCALE, NEARLY INTEGRABLE HAMILTONIAN SYSTEMS

LU XU, YONG LI, AND YINGFEI YI

ABSTRACT. We consider a multi-scale, nearly integrable Hamiltonian system. With proper degeneracy involved, such a Hamiltonian system arises naturally in problems of celestial mechanics such as Kepeler problems. Under suitable non-degenerate conditions of Bruno-Rüssmann type, the persistence of the majority of non-resonant, quasi-periodic invariant tori has been shown in [11]. This paper is devoted to the study of splitting of resonant invariant tori and the persistence of certain class of lower dimensional tori in the resonance zone. Similar to the case of standard nearly integrable Hamiltonian systems ([14, 15]), we show the persistence of the majority of Poincaré non-degenerate, lower dimensional invariant tori on a given resonant surface. The proof uses normal form reductions and KAM method in a non-standard way. More precisely, due to the involvement of multi-scales, finite steps of KAM iterations need to be firstly performed to the normal form in order to raise the non-integrable perturbation to a sufficiently high order for the standard KAM scheme to carry over.

1. INTRODUCTION

In this paper, we consider a multi-scale, real analytic, nearly integrable Hamiltonian system on $\mathbb{T}^n \times \mathbb{R}^n$, associated with the symplectic structure $dx \wedge dy$, whose Hamiltonian is 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 $\varepsilon > 0$ is a small parameter, α is a positive integer, $x = (x_1, \cdots, x_n)^\top \in \mathbb{T}^n, y = (y_1, \cdots, y_n)^\top \in G$ with $G \subset \mathbb{R}^n$ being a bounded closed region, $n_i, \bar{m}_j,\ i = 0, 1, \cdots, \alpha,\ j = 1, \cdots, \alpha + 1$ respectively, are positive integers such that $n_0 \le n_1 \le \cdots \le n_{\alpha-1} < n_\alpha := n,\ \bar{m}_1 \le \bar{m}_2 \le \cdots \le \bar{m}_\alpha < \bar{m}_{\alpha+1},\ y^{n_i} = (y_1, \cdots, y_{n_i})^\top,\ i = 0, 1, \cdots, \alpha$, and P depends on ε smoothly. We note with the above notation that $y^{n_\alpha} = y$.

Multi-scale, nearly integrable Hamiltonian systems of the form (1.1) arise naturally in many problems of celestial mechanics, for instance, the perturbed Kepler problems in which several bodies with very small masses are coupled with two massive bodies, resulting in different time scales (see e.g. [18, 20, 25]). After certain regularization and normalization ([2, 6, 12, 13, 17, 20, 21, 24]), they can be reduced from so-called nearly integrable, properly degenerate ones whose integrable parts only depend on part of the action variables. Indeed, if $\alpha = 1$, (1.1) can be reduced from a properly degenerate one in the classical sense as described in [1], while $\alpha > 1$ corresponds to the case with higher order proper degeneracy considered in [11].

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 supported by NSFC grant 11401251. 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 NSF grant DMS1109201, a Scholarship from Jilin University, and a faculty development fund from University of Alberta.

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 denote

$$\omega_\varepsilon^*(y) = \nabla N_\varepsilon(y) =: (\tilde{\omega}_\varepsilon(y^{n_{\alpha-1}}), \varepsilon^{\bar{m}_\alpha} \nabla_{\hat{y}^{n_\alpha}} H_\alpha(y^{n_\alpha}))^\top, \qquad y \in G.$$

Then the associated Hamiltonian system reads

$$\begin{cases} \dot{x} &= \omega_\varepsilon^*(y), \\ \dot{y} &= 0. \end{cases}$$

Hence $\mathbb{T}^n \times G$ is foliated into invariant tori

$$T_y^\varepsilon = \mathbb{T}^n \times \{y\}, \quad y \in G$$

with linear flows $\{x_0 + \omega_\varepsilon^*(y)t\} \times \{y\}$, $y \in G$, which can be non-resonant (quasi-periodic) or resonant depending on the resonant natures of the frequencies $\{\omega_\varepsilon^*(y)\}$.

Directly related to the stability of motions in applications arising in celestial mechanics, one important problem in studying the multi-scale Hamiltonian system (1.1) is to identify suitable conditions for the existence of quasi-periodic motions, or equivalently, the persistence, after perturbation by P, of some of the quasi-periodic, invariant tori or that of its lower dimensional sub-tori split from resonance. This problem was first studied by Arnold ([1]) who showed the existence of the majority of n-dimensional, quasi-periodic, invariant tori in the case $\alpha = \bar{m}_1 = 1$ under the degeneracy-removing condition that $H_0 + \varepsilon H_1$ satisfies either the Kolmogorov or iso-energetic non-degenerate condition. Motivated by the fact that the degeneracy-removing condition of Arnold fails in many cases of perturbed Kepler problems which admit more than two time scales, the existence of the majority of n-dimensional invariant tori of the Hamiltonian system (1.1) is proved in [11] for the case with higher order proper degeneracy under the following degeneracy-removing condition of Bruno-Rüssmann 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

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

and $\hat{y}^{n_0} = y^{n_0}$, $\hat{y}^{n_i} = (y_{n_{i-1}+1}, \cdots, y_{n_i})^\top$, $\nabla_{\hat{y}^{n_i}}$ denotes the gradient with respect to $\hat{y}^{n_i}$, for each $i = 1, 2, \cdots, \alpha$ respectively.

The persistence result in [11] has been shown to be useful in characterizing the stability of certain perturbed Kepler problems ([20, 24]). We note that the condition A*) implies that the set $\{y \in G : \omega_\varepsilon^*(y)$ is non-resonant for all ε sufficiently small$\}$ is of full Lebesgue measure in G. This is in fact the main reason for the desired persistence result holds for Hamiltonian (1.1) showed in [11].

For a standard nearly integrable Hamiltonian system

$$H(x, y, \varepsilon) = H(y) + \varepsilon P(x, y, \varepsilon), \qquad (y, x) \in G \times \mathbb{T}^n, \tag{1.3}$$

the stability problem also concerns the splitting of resonant tori and the persistence of lower dimensional sub-tori in the resonance zone. When $\varepsilon = 0$, the phase space $G \times \mathbb{T}^n$ of (1.3) is foliated into invariant tori

$$T_y = \mathbb{T}^n \times \{y\}, \quad y \in G$$

with linear flows $\{x_0 + \omega(y)t\} \times \{y\}$, where $\omega(y) = \frac{\partial H}{\partial y}(y)$, $y \in G$. While the persistence of the majority of non-resonant tori is guaranteed by KAM theorems under either Kolmogorov or Bruno-Rüssmann non-degenerate conditions on $\omega(y)$, it is well-known that the unperturbed resonant tori tend to be destroyed (under arbitrary generic perturbations) and give rise to a resonance zone containing both regular orbits and stochastic layers. The existence of regular orbits in the resonance zone is due to the fact that certain non-degenerate fractions of a resonant torus may still survive from the perturbation. A mechanism for splitting a resonant torus and the persistence of certain non-degenerate sub-tori first discovered by Poincaré ([22]) with respect to maximal resonance and later considered in [3, 7, 23] with respect to minimal resonance. The case of general resonance types was first studied by Treshchev ([26]) with respect to hyperbolic sub-tori under the Kolmogorov and a so-called g-non-degenerate condition, where g is a subgroup of $\mathbb{Z}^n$ of rank $0 < d < n$ which defines the g-resonant surface

$$\mathcal{O}(g, G) = \{y \in G : \ \langle k, \omega(y)\rangle = 0, \quad \forall k \in g\}$$

characterizing a unique class of resonant tori associated with resonant type g. More precisely, let K_2 be the integer matrix representing g and K_1 be the complementary integer matrix such that $K_0 = (K_1, K_2)$ is unimodular. Then K_0 defines a symplectic transformation

$$y \mapsto y, \qquad x \mapsto \begin{pmatrix} \varphi \\ \psi \end{pmatrix} = K_0^\top x,$$

where $\varphi \in \mathbb{T}^m$, $\psi \in \mathbb{T}^d$, such that for each $y \in \mathcal{O}(g, G)$, the resonant torus T_y is foliated into invariant m-tori

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

with linear flows $\{\varphi_0 + K_1^\top \omega(y)t\} \times \{\psi\} \times \{y\}$. Consider the function $h_0 : \mathbb{T}^d \times \mathcal{O}(g, G) \to \mathbb{R}$:

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

where

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

An $m-$torus $T_y(\psi)$ of (1.3) is said to be Poincaré-Treschev 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. It is shown in [26] that if $y_0 \in \mathcal{O}(g, G)$ is such that $\omega(y_0)$ is Kolmogorov non-degenerate, g-non-degenerate (i.e., $\det K_2^\top \frac{\partial^2 H}{\partial y^2}(y_0) K_2 \neq 0$), and $K_1^\top \omega(y_0)$ is Diophantine, then any hyperbolic, Poincaré-Treschev non-degenerate m-torus $T_{y_0}(\psi)$ will persist. This result is later generalized in [5] with respect to any Poincaré-Treschev non-degenerate m-torus on an almost full Lebesgue measure subset of $\mathcal{O}(g, G)$. In [14], two of the authors of this paper show that the result of [5] actually holds without the g-non-degenerate condition. A similar persistence result of Poincaré-Treschev tori is also shown by two of the authors of this paper in [15] under the Bruno-Rüssmann non-degenerate condition of $K_1^\top \omega(y)$ on $\mathcal{O}(g, G)$ in which case the g-non-degenerate condition of the m-tori turns out to be necessary.

In the present work, we would like to analyze similar Poincaré-Treschev mechanism for the multi-scale, nearly integrable Hamiltonian (1.1) whose frequency map $\omega_\varepsilon^*(y)$ depends on ε. Following the outline idea we introduced above, it seems necessary to consider resonance classes of ω_ε^* which are independent of ε. More precisely, corresponding to each subgroup g of $\mathbb{Z}^n$, let $k^{n_j} = (k_1, \cdots, k_{n_j})^\top$ be the first n_j-components of the vector $k \in g$, for all $j = 0, 1, \cdots, \alpha$, we will consider the following *g-resonant surface*:

$$\mathcal{O}(g, G) = \{y \in G : \ \langle k^{n_j}, \nabla_{y^{n_j}} H_j(y^{n_j})\rangle = 0, \, \forall k \in g, \, j = 0, 1, \cdots, \alpha\}.$$

The g-resonant surface characterizes a unique class of resonant tori among $\{T_y^\varepsilon : y \in G\}$. Indeed, for each $y \in \mathcal{O}(g, G)$, we have $\langle k, \omega_\varepsilon^*(y)\rangle = 0$, $\forall k \in g$ and ε sufficiently small.

However, for a multi-scale, nearly integrable Hamiltonian like (1.1), the persistence of regular fractions of resonant tori following the Poincaré-Treschev mechanism similarly to the case of standard nearly integrable Hamiltonian systems does not seem to be achievable with respect to resonance which occurs in a lower ε-order term of the frequency map, unless a special structure or lower degree of freedom is considered (see Theorem B). Technically speaking, if some resonance of the frequency map occurs in a lower order, then a resonance-splitting normal form (see Section 2) will have some components of tangential frequency in a higher order of ε than that of its normal frequency, resulting in an obstruction to the convergence of KAM iterations.

In this paper, we pay particular attention to unperturbed resonance tori of (1.1) whose resonance occurs at the highest order H_α of the Hamiltonian. This amounts to the characterizations of the resonance among the last $n_\alpha - n_{\alpha-1}$ components of $\Omega(y)$. To do so, we let g be a subgroup of $\{0\} \oplus \mathbb{Z}^{n_\alpha - n_{\alpha-1}} = \{(0, \hat{k}^{n_\alpha})^\top \in \mathbb{Z}^n : \hat{k}^{n_\alpha} \in \mathbb{Z}^{n_\alpha - n_{\alpha-1}}\}$ of rank d. Denote $\hat{g} = \{\hat{k}^{n_\alpha} \in \mathbb{Z}^{n_\alpha - n_{\alpha-1}} : (0, \hat{k}^{n_\alpha})^\top \in g\}$. Then $g = \{0\} \oplus \hat{g}$. Let $\hat{K}_2$ be an $(n_\alpha - n_{\alpha-1}) \times d$ integer matrix whose columns consist of a basis of $\hat{g}$ and let $\hat{K}_1$ be an $(n_\alpha - n_{\alpha-1}) \times (n_\alpha - n_{\alpha-1} - d)$ integer matrix such that $\det(\hat{K}_1, \hat{K}_2) = 1$. Denote

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

where O denotes a zero matrix and I denotes an identity matrix, of appropriate dimension. Then

$$\begin{aligned} \mathcal{O}(g, G) &= \{y \in G : \langle \hat{k}^{n_\alpha}, \nabla_{\hat{y}^{n_\alpha}} H_\alpha(y)\rangle = 0, \forall \hat{k}^{n_\alpha} \in \hat{g}\} = \{y \in G : \langle k, \Omega(y)\rangle = 0, \forall k \in g\} \\ &= \{y \in G : \hat{K}_2^\top \nabla_{\hat{y}^{n_\alpha}} H_\alpha(y) = 0\} = \{y \in G : K_2^\top \Omega(y) = 0\}. \end{aligned}$$

Following [15], we first assume the following condition for the multi-scale Hamiltonian (1.1):

A1) H_α is g-non-degenerate on $\mathcal{O}(g, G)$, i.e.,

$$\det \hat{K}_2^\top \frac{\partial^2 H_\alpha}{\partial (\hat{y}^{n_\alpha})^2}(y) \hat{K}_2 \neq 0, \qquad \forall y \in \mathcal{O}(g, G).$$

Under this condition, $\hat{K}_2^\top \nabla_{\hat{y}^{n_\alpha}} H_\alpha : G \to \mathbb{R}^d$ is of maximal rank, and consequently, $\mathcal{O}(g, G)$, being its kernel, is an m-dimensional, real analytic sub-manifold of G, called the *g-resonant surface*, where $m = n - d$.

Then, we also assume the following non-degenerate condition of Ω of Bruno-Rüssmann type on $\mathcal{O}(g, G)$:

A2) 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 resonant surface $\mathcal{O}(g, G)$ also determines a uniform splitting of the resonant tori $\{T_y^\varepsilon : y \in \mathcal{O}(g, G)\}$ as follows. If $x^{n_i} \in \mathbb{T}^{n_i}$ denotes the conjugate variable of y^{n_i} for each $i = 0, 1, \cdots, \alpha$, and let $\hat{x}^{n_\alpha}$ be the last $(n_\alpha - n_{\alpha-1})$-components of x, then

$$y \mapsto y, \qquad x \mapsto \begin{pmatrix} \varphi \\ \psi \end{pmatrix} \in \mathbb{T}^{n-d} \times \mathbb{T}^d$$

with

$$\begin{pmatrix} \varphi \\ \psi \end{pmatrix} = K_0^\top x = \begin{pmatrix} x^{n_{\alpha-1}} \\ \hat{K}_1^\top \hat{x}^{n_\alpha} \\ \hat{K}_2^\top \hat{x}^{n_\alpha} \end{pmatrix}$$

defines a symplectic transformation under which (1.2) becomes

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

where

$$\omega_\varepsilon(y) = \begin{pmatrix} \tilde\omega_\varepsilon(y^{n_{\alpha-1}}) \\ \varepsilon^{\bar m_\alpha} \hat K_1^\top \nabla_{\hat y^{n_\alpha}} H_\alpha(y^{n_\alpha}) \end{pmatrix}.$$

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

As for the perturbation, we consider the function $h_0 : \mathbb{T}^d \times \mathcal{O}(g,G) \to \mathbb{R}$:

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

where

$$\tilde P(\psi,\varphi,y) = P((K_0^\top)^{-1} \begin{pmatrix} \varphi \\ \psi \end{pmatrix}, y, 0) = P(x^{n_\alpha - 1}, (\hat K_1^\top, \hat K_2^\top)^{-1} \begin{pmatrix} \hat\varphi \\ \psi \end{pmatrix}, y, 0),$$

P is the perturbation in (1.1) and $\hat\varphi = \hat K_1^\top \hat x^{n_\alpha}$. For each $y \in \mathcal{O}(g,G)$, an $m-$torus $T_y^\varepsilon(\psi)$ of N_ε is said to be *Poincaré-Treschev 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. It follows from the Implicit Function Theorem that near each Poincaré-Treshchev non-degenerate m-torus, there is actually an analytic family of them. Thus, instead of assuming the existence of one such torus, without loss of generality we assume the following condition:

A3) 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 the following result:

Theorem A. *Assume the conditions* A1) - A3) *for a given subgroup g described as in the above. Then there exists a $\varepsilon_0 > 0$ sufficiently small and Cantor sets $\mathcal{O}_\varepsilon \subset \mathcal{O}(g,G)$, $0 < \varepsilon < \varepsilon_0$, with $|\mathcal{O}(g,G) \setminus \mathcal{O}_\varepsilon| = O(\varepsilon^{\frac{\iota}{2bN}})$, where $0 < \iota < \frac{1}{3}$ is a fixed constant, $b = 4d^2(N+1)$, d is the rank of a given subgroup g and N is as in* A2)*, such that for each $0 < \varepsilon < \varepsilon_0$ the Hamiltonian* (1.1) *admits a C^{N-1} Whitney smooth family of invariant, quasi-periodic 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 are only slightly deformed from the unperturbed Poincaré non-degenerate m-torus T_y^ε and the frequency of T_y^ε.*

In application of the theorem, we note that specifying the subgroup g is the same as specifying a matrix K_1 or K_2. A special case which can dramatically simplify the conditions A1), A2) is when $(\hat K_1, \hat K_2) =$ the $(n_\alpha - n_{\alpha-1}) \times (n_\alpha - n_{\alpha-1})$ identity matrix. In this case, $\hat\varphi$ consists of some components of $\hat x^{n_\alpha}$ and ψ consists of the remaining components.

As remarked before, though the persistence of lower dimensional tori in the resonance zone of N_ε does not seem to hold with respect to resonances occurred in lower order terms of ε in general, there are exceptions in the case of lower degree of freedom with respect to special type of resonances. Aiming at application to spatial three-body problems, we consider the following real analytic Hamiltonian on $\mathbb{T}^3 \times \mathbb{R}^3$ associated with the symplectic structure $dx \wedge dy$:

$$H(x,y,\varepsilon) = H_0(y_0) + \varepsilon^{\bar m_1} H_1(y) + \varepsilon^{\bar m_2} H_2(y) + \varepsilon^{\bar m_3} P(x,y,\varepsilon), \tag{1.4}$$

where ε is a small parameter, $x = (x_0, x_1, x_2)^\top \in \mathbb{T}^3$, $y = (y_0, y_1, y_2)^\top \in G$, G is a bounded closed region in $\mathbb{R}^3$, $\bar{m}_j$, $j = 1, 2, 3$, are positive integers satisfying $\bar{m}_1 < \bar{m}_2 < \bar{m}_3$, the perturbation P depends on ε smoothly. Comparing with the general case (1.1), we note that the condition $\bar{m}_1 < \bar{m}_2$ above is crucial for us to consider lower order resonances in (1.4).

It is clear that the resonance of the frequency map ω_ε^* will not occur in the first term H_0 unless $\partial_{y_0} H_0 \equiv 0$. Therefore, it is natural to consider resonances occurring in H_1, H_2. The case that H_1, H_2 are completely resonant corresponds to the resonant subgroup $g = \{0\} \times \mathbb{Z}^2$ and has already been considered in [10]. We now consider the remaining case.

Let g be any rank 1 subgroup of $\{0\} \times \mathbb{Z}^2$ and $\hat{K}_2 = \binom{\hat{\tau}_1}{\hat{\tau}_2}$ be an integer vector such that $(0, \hat{K}_2)^\top$ is a fixed basis of g. Then

$$\mathcal{O}(g, G) = \{y \in G : \hat{K}_2^\top \partial_{\hat{y}} H_1(y) = \hat{K}_2^\top \partial_{\hat{y}} H_2(y) = 0\},$$

where $\hat{y} = (y_1, y_2)^\top$. Let $\hat{K}_1 = \binom{\hat{\iota}_1}{\hat{\iota}_2}$ be an integer vector such that $\det(\hat{K}_1, \hat{K}_2) = 1$ and denote

$$K_1 = \begin{pmatrix} 1 & 0 \\ O & \hat{K}_1 \end{pmatrix}, \qquad K_2 = \begin{pmatrix} 0 \\ \hat{K}_2 \end{pmatrix}. \tag{1.5}$$

It is clear that such a choice of g does allow resonance in the lower order term H_1, for instance, when $\hat{K}_2 = \binom{1}{0}$ and $\hat{K}_1 = \binom{0}{-1}$.

Similar to A1), we assume the following condition:

A1)' Either H_1 or H_2 is g-non-degenerate on $\mathcal{O}(g, G)$, i.e.,

$$\text{either } \det \hat{K}_2^\top \frac{\partial^2 H_1}{\partial \hat{y}^2}(y) \hat{K}_2 \neq 0 \text{ or } \det \hat{K}_2^\top \frac{\partial^2 H_2}{\partial \hat{y}^2}(y) \hat{K}_2 \neq 0, \qquad \forall y \in \mathcal{O}(g, G).$$

.

We have the following result:

Theorem B. *Consider the Hamiltonian* (1.4) *and assume conditions A1)', A3) with respect to the rank* 1 *subgroup* g *above. Also assume condition* A2) *with respect to* K_1 *in* (1.5) *and*

$$\Omega(y) = (\partial_{y_0} H_0(y_0), \partial_{y_1} H_1(y), \partial_{y_2} H_1(y))^\top. \tag{1.6}$$

Then the conclusion of Theorem A holds to yield a family of quasi-periodic invariant 2*-tori of (1.4).*

Remark. (1) We remark that with more estimates involved in the proof both theorems actually hold when the Hamiltonians in (1.1), (1.4) are of the class C^∞.

(2) Theorem A also holds on a submanifold M of G if conditions A1)-A3) are assumed on M instead of G. In particular, if M is taken as an energy surface $\{H_0 = h_0, H_1 = h_1, \cdots, H_\alpha = h_\alpha\}$, then this will lead to an iso-energetic version of persistence result for lower dimensional tori. The proof of such a result more or less follows the arguments of [4] and the quasi-linear iterative scheme contained in this paper. But in application the verification of assumptions A1)-A3) on M will be a non-trivial matter, depending on a careful choice of M and g.

(3) Differing from the case of a standard nearly integrable Hamiltonian system considered in [14, 15], the excluding measure for the persistence of lower invariant tori on a resonant surface of the Hamiltonian (1.1) is of the order $O(\varepsilon^{\frac{\iota}{2bN}})$ for a pre-fixed small positive constant ι, as shown

in Theorem A above. This is due to the special nature of small divisor conditions required in finding quasi-periodic solutions in the multi-scaled system (1.1). More precisely, as to be shown in Section 2, for each $\xi \in \mathcal{O}(g,G)$, an initial normal form reduction of (1.1) yields

$$H(x,y,z,\xi,\varepsilon) = e_\varepsilon(\xi) + \langle \omega_\varepsilon(\xi), y\rangle + \delta h(y,\xi,\varepsilon) + \frac{\delta}{2}\langle z, \varepsilon^{\bar m_\alpha} M(\xi,\varepsilon)\ z\rangle + \delta\varepsilon^{\bar m_\alpha} P(x,y,z,\xi,\varepsilon), \tag{1.7}$$

$(x,y,z) \in \mathbb{T}^m \times \mathbb{R}^m \times \mathbb{R}^{2d}$, where $\delta = \varepsilon^{\frac{\bar m_{\alpha+1} - \bar m_\alpha}{2}}$. However, in viewing the multi-scales of ω_ε, the small divisor conditions in order to carry out the KAM iterations requires that the perturbation is of at least an order of $O(\varepsilon^{2(N+6)c_*})$, where

$$c_* = 4d^2 \sum_{i=1}^{\alpha} \bar m_i (n_i - n_{i-1}).$$

To overcome this obstacle to the application of the KAM method, we will adopt the approach from [11] to first perform a finite step of iterations to push the perturbation to the desired order. Another technical ingredient in proving our results is the derivative estimates, up to order N, of the frequency map in each KAM step which is necessary in verifying the Bruno-Rüssmann condition and proving measure estimate of the persistent invariant sub-tori. Unlike the case of a standard nearly integrable Hamiltonian, such derivative estimates cannot be done using Cauchy formula simply because for the ε-dependent frequency map its C^0 norm in a complex neighborhood of the domain cannot be explicitly estimated. We note that if the perturbation is already in an order of $O(\varepsilon^{2(N+6)c_*})$, then a finite number of KAM iterations will not be necessary, and the excluding measure for the persistence of invariant, quasi-periodic, m-dimensional tori can be improved to an order of $O(\varepsilon^{\frac{12c_*}{N}})$ (see the measure estimate in Section 4).

(4) Similar to the Melnikov persistence problem considered for standard Hamiltonian systems, it is meaningful to consider the persistence of non-resonant, lower dimensional tori for a multi-scale, partially nearly integrable Hamiltonian system. For such a system, one would typically work with the following Hamiltonian normal form

$$\begin{aligned} H(x,y,z,\xi,\varepsilon) &= e(\xi) + \langle \omega^0(\xi), y^{n_0}\rangle + \varepsilon^{\bar m_1}\langle \omega^1(\xi), y^{n_1}\rangle + \cdots + \varepsilon^{\bar m_\alpha}\langle \omega^\alpha(\xi), y^{n_\alpha}\rangle \\ &+ \frac{1}{2}\langle z^{2m_0}, M^0(\xi) z^{2m_0}\rangle + \frac{\varepsilon^{\bar m_1}}{2}\langle z^{2m_1}, M^1(\xi) z^{2m_1}\rangle + \cdots + \frac{\varepsilon^{\bar m_\alpha}}{2}\langle z^{2m_\alpha}, M^\alpha(\xi) z^{2m_\alpha}\rangle \\ &+ \varepsilon^{\bar m_{\alpha+1}} P(x,y,z,\xi), \end{aligned} \tag{1.8}$$

where $x \in \mathbb{T}^m$, $n_0 \le n_1 \le \cdots \le n_{\alpha-1} < n_\alpha := m$, $\bar m_1 \le \bar m_2 \le \cdots \le \bar m_\alpha < \bar m_{\alpha+1}$, $m_0 \le m_1 \le \cdots \le m_\alpha := d$ are positive integers, $y^{n_i} = (y_1, \cdots, y_{n_i})^\top$, $z^{2m_i} = (z_1,\ z_2,\ \cdots,\ z_{2m_i})^\top$, $i = 0,1,\cdots,\alpha$, $y = y^{n_\alpha}$, $z = z^{2m_\alpha}$, and $\xi \in \mathcal{O} \subset \mathbb{R}^m$ is a parameter. Denote

$$\Omega(\xi) = (\hat\omega^0(\xi), \hat\omega^1(\xi), \cdots, \hat\omega^\alpha(\xi))^\top,$$

where $\hat\omega^i(\xi) = (\omega^i_{n_{i-1}+1}(\xi),\ \omega^i_{n_{i-1}+2}(\xi),\ \cdots,\ \omega^i_{n_i}(\xi))^\top$, $i = 1,2,\ \cdots,\ \alpha$, $\hat\omega^0(\xi) = \omega^0(\xi)$. It is clear that the unperturbed part of (1.8) (i.e., when $P = 0$) admits a family of m-tori associated with the relative equilibrium $z = 0$ which are mostly quasi-periodic if Ω satisfies a Bruno-Rüssmann condition on $\mathcal{O}$. We note that unlike (1.7), the quadratic terms of (1.8) are of the same order of ε-scales as the linear terms. Inspired by the case of standard partially nearly integrable Hamiltonian systems, the persistence of the majority of unperturbed m-tori should require the non-degenerate and Melnikov non-resonant conditions that for each $i = 0,1,\cdots,\alpha$, the lower right $2(m_i - m_{i-1}) \times 2(m_i - m_{i-1})$-block $\hat M^i$ of M^i is non-singular on $\mathcal{O}$, and that the set

$$\{\xi \in \mathcal{O} : \sqrt{-1}\langle k,\ \hat\omega^i(\xi)\rangle - \lambda^i_j(\xi) - \lambda^i_l(\xi) \ne 0,\ \forall k \in \mathbb{Z}^{n_i - n_{i-1}} \setminus \{0\},\ 1 \le j,l \le 2(m_i - m_{i-1})\}$$

admits full Lebesgue measure relative to $\mathcal{O}$, where $\lambda^i_1(\xi),\ \cdots,\ \lambda^i_{2(m_i-m_{i-1})}(\xi)$ are eigenvalues of $J_i \hat M^i(\xi)$ with J_i being the standard $2(m_i - m_{i-1}) \times 2(m_i - m_{i-1})$ symplectic matrix. In the above, we take $n_{-1} = m_{-1} = 0$. We recall that the above non-resonant condition is the weak form of Melnikov's second non-resonant condition assumed in [16]. Under these conditions, it seems that the KAM scheme developed in [16] is applicable to the multi-scale Hamiltonian (1.8) but the convergence of the scheme should require to push the perturbation to a sufficiently high order first, similarly to the treatment contained in the present work. We leave the detail constructions in a subsequent work.

The rest sections are organized as following. In Section 2, we will rewrite (1.1) and (1.4) into a usual normal form. We will perform a finite step of KAM iterations in Section 3 to push the perturbation term of the normal form to a properly high order of ε. Theorems A, B will be proved in Section 4 by applying standard KAM method. We show in Section 5 an example contained in [20] of a properly degenerate Hamiltonian system reduced from spatial three-body problem for which our main results may be applicable to yield invariant lower dimensional tori in the resonance zone.

Through the rest of the paper, unless specified otherwise, we will use the same symbol $|\cdot|$ to denote an equivalent (finite dimensional) vector norm and its induced matrix norm, absolute value, and measure of sets etc., and use $|\cdot|_D$ to denote the sup-norm of functions on a domain D. For each $r, s > 0$, we denote

$$D(r,s) = \{(x,y,u,v) \in \mathbb{T}^m \times \mathbb{C}^m \times \mathbb{C}^d \times \mathbb{C}^d : \ |\mathrm{Im}\, x| < r, \ |y| < s^2, \ |u| + |v| < s\},$$

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

2. Normal form

In this section, with respect to a given resonant type, we convert Hamiltonians (1.1), (1.4) into a normal form near a family of Poincaré non-degenerate sub-tori in their resonance zones. First let g, K_1, K_2 be given as in Theorem A and $\psi(y)$ be as in A3).

For each $\xi \in \mathcal{O}(g, G)$, Taylor expansion of (1.1) at ξ reads

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

up to a constant. Write $\omega_\varepsilon^*(\xi)$ into the following form

$$\omega_\varepsilon^*(\xi) := (\omega_\varepsilon^{n_0}(\xi), \varepsilon^{\bar{m}_1}\hat{\omega}_\varepsilon^{n_1}(\xi), \cdots, \varepsilon^{\bar{m}_\alpha}\hat{\omega}_\varepsilon^{n_\alpha}(\xi))^\top,$$

where

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

and

$$\hat{\omega}_\varepsilon^{n_j}(\xi) = \nabla_{\hat{y}^{n_j}} H_j(\xi^{n_j}) + \varepsilon^{\bar{m}_{j+1}-\bar{m}_j}\nabla_{\hat{y}^{n_j}} H_{j+1}(\xi^{n_{j+1}}) + \cdots + \varepsilon^{\bar{m}_\alpha - \bar{m}_j}\nabla_{\hat{y}^{n_j}} H_\alpha(\xi^{n_\alpha}),$$

where $\xi^{n_j} = (\xi_1, \cdots \xi_{n_j})^\top$, $j = 0, 1, \cdots, \alpha$. Note that the matrix $\hat{\Gamma}(\xi,\varepsilon)$ in the second term of (2.1) can be expressed as

$$\begin{aligned} \hat{\Gamma}(\xi,\varepsilon) &:= \frac{\partial^2 N_\varepsilon}{\partial y^2}(\xi,\varepsilon) \\ &= \hat{\Gamma}_0(\xi^{n_0}) + \cdots + \varepsilon^{\bar{m}_\alpha}\hat{\Gamma}_\alpha(\xi^{n_\alpha}), \end{aligned} \tag{2.2}$$

where

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

Consider the linear symplectic transformation:

$$
\begin{aligned}
y-\xi &= K_0 p, \\
q &= {K_0}^{\top} x.
\end{aligned}
\tag{2.3}
$$

Then (2.1) can be re-written as

$$
\begin{aligned}
&\bar{H}(\bar{p},p,q,\xi,\varepsilon) \\
= \ &\langle \omega_\varepsilon(\xi),\bar{p}\rangle + \frac{1}{2}\langle \bar{\Gamma}(\xi,\varepsilon)p,p\rangle + \varepsilon^{\bar{m}_\alpha+1}\hat{P}(p,q,\xi,\varepsilon) + O(|p|^3),
\end{aligned}
\tag{2.4}
$$

where $\bar{p}$ is the first m-components of p, $\omega_\varepsilon(\xi) = K_1^\top \omega_\varepsilon^*(\xi)$, $\bar{\Gamma}(\xi,\varepsilon) = {K_0}^\top \hat{\Gamma}(\xi,\varepsilon)K_0$, and $\hat{P}(q,p,\xi,\varepsilon)$ $= P(({K_0}^\top)^{-1}q, \xi + K_0 p, \varepsilon)$. We rewrite ω_ε as follows:

$$
\begin{aligned}
\omega_\varepsilon(\xi) &= (\omega_\varepsilon^{n_0}(\xi),\cdots,\varepsilon^{\bar{m}_i}\hat{\omega}_\varepsilon^{n_i}(\xi),\cdots \varepsilon^{\bar{m}_{\alpha-1}}\hat{\omega}_\varepsilon^{n_{\alpha-1}}(\xi),\varepsilon^{\bar{m}_\alpha}\hat{K}_1^\top\hat{\omega}_\varepsilon^{n_\alpha}(\xi))^\top \\
&:= (\omega_\varepsilon^{0}(\xi),\cdots,\varepsilon^{\bar{m}_i}\hat{\omega}_\varepsilon^{i}(\xi),\cdots \varepsilon^{\bar{m}_{\alpha-1}}\hat{\omega}_\varepsilon^{\alpha-1}(\xi),\varepsilon^{\bar{m}_\alpha}\hat{\omega}_\varepsilon^{\alpha}(\xi))^\top.
\end{aligned}
\tag{2.5}
$$

To decompose the matrix $\bar{\Gamma}(\xi,\varepsilon)$, we have by (2.2) that

$$
\bar{\Gamma}(\xi,\varepsilon) = K_0^\top\hat{\Gamma}_0(\xi^{n_0})K_0 + \cdots + \varepsilon^{\bar{m}_\alpha}K_0^\top\hat{\Gamma}_\alpha(\xi^{n_\alpha})K_0.
$$

It is easy to verify that

$$
K_0^\top\hat{\Gamma}_j(\xi^{n_j})K_0 = \begin{pmatrix} \frac{\partial^2 H_i}{\partial {y^{n_i}}^2}(\xi^{n_i}) & O \\ O & O \end{pmatrix}_{n\times n}, \ \forall\, j = 0,1,\cdots,\alpha-1,
$$

where O denotes a zero matrix of appropriate dimension. As for the last term, write

$$
K_0^\top\hat{\Gamma}_\alpha(\xi^{n_\alpha})K_0 = \begin{pmatrix} \bar{\Gamma}^{11} & \bar{\Gamma}^{12} \\ \bar{\Gamma}^{21} & \bar{\Gamma}^{22} \end{pmatrix},
$$

$\bar{\Gamma}^{11},\ \bar{\Gamma}^{12},\ \bar{\Gamma}^{21},\ \bar{\Gamma}^{22}$ are $(n-d)\times(n-d),\ (n-d)\times d,\ d\times(n-d),\ d\times d$ matrices, respectively. We clearly have

$$
\bar{\Gamma}^{22} \ = \ K_2^\top\hat{\Gamma}_\alpha K_2 = \varepsilon^{\bar{m}_\alpha}K_2^\top\frac{\partial^2 H_\alpha}{\partial y^2}K_2 = \varepsilon^{\bar{m}_\alpha}\hat{K}_2^\top\frac{\partial^2 H_\alpha}{\partial^2 \hat{y}^{n_\alpha}}\hat{K}_2 := \varepsilon^{\bar{m}_\alpha}\Gamma^{22}.
$$

By A1), $\bar{\Gamma}^{22}$ is nonsingular for all $\xi \in \mathcal{O}(g,G)$.

With respect to the Hamiltonian (1.4), similar calculations based on the resonant type K_2 defined in (1.5) yields that

$$
\begin{aligned}
\omega_\varepsilon^*(\xi) &= (\partial_{y_0}H_0(\xi) + \varepsilon^{\bar{m}_1}\partial_{y_0}H_1(\xi) + \varepsilon^{\bar{m}_2}\partial_{y_0}H_2(\xi), \\
&\quad \varepsilon^{\bar{m}_1}\partial_{y_1}H_1(\xi) + \varepsilon^{\bar{m}_2}\partial_{y_1}H_2(\xi), \varepsilon^{\bar{m}_1}\partial_{y_2}H_1(\xi) + \varepsilon^{\bar{m}_2}\partial_{y_2}H_2(\xi))^\top \in \mathbb{R}^3
\end{aligned}
$$

and

$$
\hat{\Gamma}(\xi,\varepsilon) = \hat{\Gamma}_0 + \varepsilon^{\bar{m}_1}\hat{\Gamma}_1 + \varepsilon^{\bar{m}_2}\hat{\Gamma}_2,
$$

where

$$
\hat{\Gamma}_0 = \begin{pmatrix} \frac{\partial^2 H_0}{\partial^2 y^0}(\xi) & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \ \hat{\Gamma}_i = \frac{\partial^2 H_i}{\partial y^2}(\xi),\ i = 1,2.
$$

Under the same linear symplectic transformation (2.3), we then obtain (2.4) with $\bar{p} = (p_0,p_1)^\top$, $\bar{\Gamma}(\xi,\varepsilon) = {K_0}^\top\hat{\Gamma}(\xi,\varepsilon)K_0$, and

$$
\begin{aligned}
\omega_\varepsilon(\xi) &= (\partial_{\hat{y}^0}H_0(\xi) + \varepsilon^{\bar{m}_1}\partial_{\hat{y}^0}H_1(\xi) + \varepsilon^{\bar{m}_2}\partial_{\hat{y}^0}H_2(\xi), \varepsilon^{\bar{m}_1}\langle \hat{\imath}, \nabla_{\hat{y}^1}H_1\rangle + \varepsilon^{\bar{m}_2}\langle \hat{\imath}, \nabla_{\hat{y}^2}H_2\rangle)^\top \\
&:= (\omega_\varepsilon^0, \varepsilon^{\bar{m}_1}\omega_\varepsilon^1)^\top,
\end{aligned}
$$

where $\hat{y}^0 = y_0,\ \hat{y}^1 = \hat{y}^2 = (y_1, y_2)^\top,\ \hat{\iota} = (\hat{\iota}_1, \hat{\iota}_2)^\top$. Decomposing the matrix $\bar{\Gamma}$ as in the above, we have

$$\bar{\Gamma}^{22} = \varepsilon^{\bar{m}_1}\langle \frac{\partial^2 H_1}{\partial^2 \hat{y}^1}\hat{\tau}, \hat{\tau}\rangle + \varepsilon^{\bar{m}_2}\langle \frac{\partial^2 H_2}{\partial^2 \hat{y}^2}\hat{\tau}, \hat{\tau}\rangle,$$

where $\hat{\tau} = (\hat{\tau}_1, \hat{\tau}_2)^\top$. By A1)', we have $\langle \frac{\partial^2 H_1}{\partial^2 \hat{y}^1}\hat{\tau}, \hat{\tau}\rangle \neq 0$, or $\langle \frac{\partial^2 H_2}{\partial^2 \hat{y}^2}\hat{\tau}, \hat{\tau}\rangle \neq 0$ if $\langle \frac{\partial^2 H_1}{\partial^2 \hat{y}^1}\hat{\tau}, \hat{\tau}\rangle = 0$. In any case, $\bar{\Gamma}^{22} \neq 0$.

Let $p = (\bar{p}, p^d)^\top \in \mathbb{R}^m \times \mathbb{R}^d$ and $q = (\varphi, \psi + \psi(\xi))^\top$. The Hamiltonian (2.4) in both cases above becomes

$$\begin{aligned} H(\varphi, \psi, \bar{p}, p^d, \xi, \varepsilon) &= \langle \omega_\varepsilon(\xi), \bar{p}\rangle + \frac{1}{2}\langle \varepsilon^{\bar{m}_\alpha}\Gamma^{22}(\xi, \varepsilon)p^d, p^d\rangle + h(\bar{p}, \xi, \varepsilon) \\ &\quad + O(\varepsilon^{\bar{m}_\alpha}|\bar{p}||p^d| + \varepsilon^{\bar{m}_\alpha}|p|^3) + \varepsilon^{\bar{m}_\alpha+1}\bar{P}(\varphi, \psi, \bar{p}, p^d, \xi, \varepsilon), \end{aligned} \tag{2.6}$$

where

$$\begin{aligned} h(\bar{p}, \xi, \varepsilon) &= O(|\bar{p}^{n_0}|^2 + \cdots + \varepsilon^{\bar{m}_{\alpha-1}}|\bar{p}^{n_{\alpha-1}}|^2 + \varepsilon^{\bar{m}_\alpha}|\bar{p}|^2), \\ \bar{p}^{n_j} &= (\bar{p}_1, \cdots, \bar{p}_{n_j})^\top,\ j = 0, 1, \cdots, \alpha - 1, \\ \bar{P}(\varphi, \psi, \bar{p}, p^d, \xi, \varepsilon) &= \hat{P}(\varphi, \psi + \psi(\xi), \bar{p}, p^d, \xi, \varepsilon). \end{aligned}$$

As to be seen later, lower order terms $O(\varepsilon^{\bar{m}_j}|\bar{p}^{n_j}|^2),\ j = 0, \cdots, \alpha - 1$, in the above will play an important role during the iteration process in order to control small derivatives. Hence they cannot be simply included in the perturbation by rescaling.

Denote

$$\Omega_\varepsilon(\xi) = (\omega_\varepsilon^0(\xi), \cdots, \hat{\omega}_\varepsilon^i(\xi), \cdots \hat{\omega}_\varepsilon^{\alpha-1}(\xi), \hat{\omega}_\varepsilon^\alpha(\xi))^\top, \qquad \xi \in \mathcal{O}(g, G).$$

Then it is clear that

$$\partial_\xi^l \Omega_\varepsilon(\xi) = \partial_\xi^l K_1^\top \Omega(\xi) + O(\varepsilon), \qquad l = 0, 1, \cdots, N \tag{2.7}$$

uniformly for $\xi \in \mathcal{O}(g, G)$, where, $\Omega(\xi)$, is as in (1.2) in the case of (1.1) and equals in the case of (1.4).

For fixed positive constants $\gamma_0,\ \tau > \max\{m(m+1) - 1, N(N+1) - 1\}$, consider sets

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

It follows from A2), (2.7), and the measure estimate under Bruno-Rüssmann conditions ([27, 28], see also Section 3 of this paper) that

$$|\mathcal{O}(g, G) \setminus \Lambda_\varepsilon| = O(\gamma_0^{\frac{1}{N}}). \tag{2.8}$$

Moreover, it is easy to see that

$$\Lambda_\varepsilon \subset O_\varepsilon, \qquad 0 < \varepsilon \ll 1.$$

For each $\xi \in \Lambda_\varepsilon$, we can separate the first-order resonant terms from the perturbation $\bar{P}$ as follows. Using the Fourier expansion, we have

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

where $\tilde{\Gamma} := \frac{\partial^2 h_0}{\partial \psi^2}(\xi)$.

Consider the family $\{S_\xi : \xi \in O_\varepsilon\}$ of functions on $(\mathbb{T}^m \times \mathbb{R}^d) \times \mathbb{R}^n$ defined by

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

where $S_k = \frac{\sqrt{-1} h_k(\psi,\xi)}{\langle \omega_\varepsilon(\xi), k\rangle}$ and $h_k(\psi, \xi) = \int_{\mathbb{T}^d} \bar{P}(\varphi, \psi, 0, \xi) e^{\sqrt{-1}\langle k,\varphi\rangle} d\varphi$, for each $\xi \in O_\varepsilon$. We note that $\{S_\xi : \xi \in O_\varepsilon\}$ is a Whitney smooth family of real analytic functions. It follows that

$$(q, p) = (\varphi, \psi + \psi(\xi), p) \to (\varphi, \psi, Y) : q = \frac{\partial S(q, Y)}{\partial Y},\ p = \frac{\partial S(q, Y)}{\partial q}$$

defines a Whitney smooth family of real analytic, sympletic transformations on $(\mathbb{T}^m \times \mathbb{R}^d) \times \mathbb{R}^n$ such that

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

and

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

where $Y = (\bar{Y}, Y^d)^\top \in \mathbb{R}^m \times \mathbb{R}^d$. Under this family of transformations, the Hamiltonian (2.6) becomes

$$\begin{aligned}
& H(\varphi, \psi, \bar{Y}, Y^d, \xi, \varepsilon) \\
= \quad & \langle \omega_\varepsilon(\xi), \bar{Y}\rangle + \frac{1}{2}\langle \varepsilon^{\bar{m}_\alpha} \Gamma^{22}(\xi, \varepsilon) Y^d, Y^d\rangle + \frac{1}{2}\langle \varepsilon^{\bar{m}_{\alpha+1}} \tilde{\Gamma}\psi, \psi\rangle + h(\bar{Y}, \xi, \varepsilon) \\
& + O(\varepsilon^{\bar{m}_{\alpha+1}+\bar{m}_\alpha}|Y|) + O(\varepsilon^{2\bar{m}_{\alpha+1}}) + O(\varepsilon^{\bar{m}_{\alpha+1}}|\psi|^3) + O(\varepsilon^{\bar{m}_\alpha}|\bar{Y}||Y^d|) + O(\varepsilon^{\bar{m}_\alpha} Y^3).
\end{aligned}$$

Consider the rescaling $Y \to \varepsilon^{\frac{\bar{m}_{\alpha+1}-\bar{m}_\alpha}{2}} Y$. The re-scaled Hamiltonian reads

$$\begin{aligned}
\tilde{H} \quad = \quad & \frac{H(\varphi, \psi, \varepsilon^{\frac{\bar{m}_{\alpha+1}-\bar{m}_\alpha}{2}}\bar{Y}, \varepsilon^{\frac{\bar{m}_{\alpha+1}-\bar{m}_\alpha}{2}} Y^d, \xi, \varepsilon)}{\varepsilon^{\frac{\bar{m}_{\alpha+1}-\bar{m}_\alpha}{2}}} = \langle \omega_\varepsilon(\xi), \bar{Y}\rangle \\
+ \quad & \varepsilon^{\frac{\bar{m}_{\alpha+1}-\bar{m}_\alpha}{2}} h(\bar{Y}, \xi, \varepsilon) + \frac{\varepsilon^{\frac{\bar{m}_{\alpha+1}-\bar{m}_\alpha}{2}}}{2}\langle \varepsilon^{\bar{m}_\alpha}\Gamma^{22}(\xi, \varepsilon) Y^d, Y^d\rangle + \frac{\varepsilon^{\frac{\bar{m}_{\alpha+1}-\bar{m}_\alpha}{2}}}{2}\langle \varepsilon^{\bar{m}_\alpha}\tilde{\Gamma}\psi, \psi\rangle) \\
+ \quad & \frac{\varepsilon^{\frac{\bar{m}_{\alpha+1}-\bar{m}_\alpha}{2}}}{2}(O(\varepsilon^{\frac{\bar{m}_{\alpha+1}+3\bar{m}_\alpha}{2}}|Y|) + O(\varepsilon^{\frac{3\bar{m}_{\alpha+1}+\bar{m}_\alpha}{2}}) \\
+ \quad & O(\varepsilon^{\frac{\bar{m}_{\alpha+1}+\bar{m}_\alpha}{2}}|\psi|^3) + O(\varepsilon^{\bar{m}_\alpha}|\bar{Y}||Y^d|) + O(\varepsilon^{\bar{m}_\alpha}|Y|^3)).
\end{aligned}$$

With $\varphi,\ \bar{Y},\ Y^d,\ \psi, \tilde{H}, \varepsilon^{\frac{\bar{m}_{\alpha+1}-\bar{m}_\alpha}{2}}$ in place of $x,\ y,\ u,\ v,\ H,\ \delta$ respectively, we obtain the following normal form

$$H(x, y, z, \xi, \varepsilon) \quad = \quad N + \delta\varepsilon^{\bar{m}_\alpha} P, \tag{2.9}$$

where

$$\begin{aligned}
N \quad &= \quad \langle \omega_\varepsilon(\xi), y\rangle + \delta h(y, \xi, \varepsilon) + \frac{\delta}{2}\langle \varepsilon^{\bar{m}_\alpha} M(\xi, \varepsilon) z, z\rangle, \\
h(y, \xi, \varepsilon) \quad &= \quad O(|y^{n_0}|^2 + \cdots + \varepsilon^{\bar{m}_{\alpha-1}}|y^{n_{\alpha-1}}|^2 + \varepsilon^{\bar{m}_\alpha}|y|^2), && (2.10) \\
y^{n_j} \quad &= \quad (y_1, \cdots, y_{n_j})^\top,\ j = 0, 1, \cdots, \alpha - 1, \\
M(\xi, \varepsilon) \quad &= \quad \begin{pmatrix} \Gamma^{22} & 0 \\ 0 & \bar{\Gamma} \end{pmatrix},\ z = (u, v)^\top, && (2.11) \\
P(x, y, z, \xi, \varepsilon) \quad &= \quad O(\varepsilon^{\frac{\bar{m}_{\alpha+1}+\bar{m}_\alpha}{2}}|(y, u)|) + O(\varepsilon^{\frac{\bar{m}_{\alpha+1}+\bar{m}_\alpha}{2}}) \\
& \quad\ + O(\varepsilon^{\frac{\bar{m}_{\alpha+1}-\bar{m}_\alpha}{2}}|\psi|^3) + O(|y||u|) + O(|(y, u)|^3).
\end{aligned}$$

Let $0 < r_0 \ll 1$ be given and define $s_0 = \varepsilon^{\frac{\bar m_{\alpha+1} - \bar m_\alpha}{6}}$ for small $\varepsilon > 0$. Then the Hamiltonian (2.9) is real analytic in $(x, y, z) \in D(r_0, s_0)$ and Whitney smooth in $\xi \in O_\varepsilon$, and it is easy to see that

$$|\partial_\xi^l P|_{D(r_0,s_0)\times O_\varepsilon} = O(\varepsilon^{\frac{\bar m_{\alpha+1} - \bar m_\alpha}{2}}), \ \ \forall l \le N. \tag{2.12}$$

3. Improving the order of perturbation

In this section, we will use a finite step of quasi-linear KAM iterations to improve the normal form (2.9) by pushing the perturbation to a desired higher order. We re-label the real analytic Hamiltonian (2.9) as

$$H^0(x,y,z,\xi,\varepsilon) = e_\varepsilon^0(\xi) + \langle \omega_\varepsilon^0(\xi), y\rangle + \delta h^0(y,\xi,\varepsilon) + \frac{\delta}{2}\langle z, \varepsilon^{\bar m_\alpha} M^0(\xi,\varepsilon)\, z\rangle + \delta\varepsilon^{\bar m_\alpha} P^0(x,y,z,\xi,\varepsilon), \tag{3.1}$$

where $(x, y, z) \in D(r_0, s_0)$, $\xi \in \Lambda^0 := \Lambda_\varepsilon$, $e_\varepsilon^0 \equiv 0$, $\omega_\varepsilon^0 := \omega_\varepsilon$, $M^0 := M$, $h^0 := h$, and $P^0 := P$, as defined in (2.5), (2.11), (2.10), (2.12) respectively. We note that M^0 is invertible on its domain of definition and $|(M^0)^{-1}| = O(1)$. Denote $\gamma_0^b = \varepsilon^{\frac{\iota(\bar m_{\alpha+1} - \bar m_\alpha)}{2}}$, $\mu_0 = \varepsilon^{\frac{(1-3\iota)(\bar m_{\alpha+1} - \bar m_\alpha)}{6}}$, where $0 < \iota < \frac{1}{3}$ is fixed and $b = 4d^2(N + 1)$. It follows from (2.12) that

$$|\partial_\xi^l P^0|_{D(r_0,s_0)\times\Lambda^0} < \gamma_0^b s_0^2 \mu_0, \ \ |l| \le N. \tag{3.2}$$

We will use the quasi-linear iterative scheme introduced in [14] to iterate the Hamiltonian (3.1) so that after each iteration the term h^0 is kept essentially unchanged.

Assume that, after a ν-th iterative step, we have obtained the following Whitney smooth family of real analytic Hamiltonians

$$H(x,y,z,\xi,\varepsilon) = e_\varepsilon(\xi) + \langle \omega_\varepsilon(\xi), y\rangle + \delta h(y,\xi,\varepsilon) + \frac{\delta}{2}\langle z, \varepsilon^{\bar m_\alpha} M(\xi,\varepsilon)\, z\rangle + \delta\varepsilon^{\bar m_\alpha} P(x,y,z,\xi,\varepsilon), \tag{3.3}$$

where $(x, y, z) \in D(r, s)$ for some smaller $0 < r < r_0$, $0 < s < s_0$, $\xi \in \Lambda$ for a subset $\Lambda \subset \mathcal{O}(g, G)$, ω_ε, M, h are of the same forms as in (2.5), (2.11), (2.10) 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 $\xi \in \mathcal{O}(g, G)$, we write

$$\omega_\varepsilon(\xi) = (\omega_\varepsilon^0(\xi), \cdots, \varepsilon^{\bar m_i}\hat\omega_\varepsilon^i(\xi), \cdots \varepsilon^{\bar m_{\alpha-1}}\hat\omega_\varepsilon^{\alpha-1}(\xi), \varepsilon^{\bar m_\alpha}\hat\omega_\varepsilon^\alpha(\xi))^\top$$

and denote

$$\Omega_\varepsilon(\xi) = (\omega_\varepsilon^0(\xi), \cdots, \hat\omega_\varepsilon^i(\xi), \cdots \hat\omega_\varepsilon^{\alpha-1}(\xi), \hat\omega_\varepsilon^\alpha(\xi))^\top.$$

For $+ =: \nu + 1$, we will find a symplectic transformation Φ^+, such that, on a smaller phase domain $D(r_+, s_+)$, Hamiltonian (3.3) is transformed to a new Hamiltonian

$$H^+ = H\circ\Phi^+ = e_\varepsilon^+(\xi) + \langle \omega_\varepsilon^+(\xi), y\rangle + \delta h^+(y,\xi,\varepsilon) + \frac{\delta}{2}\langle z, \varepsilon^{\bar m_\alpha} M^+(\xi,\varepsilon)\, z\rangle + \delta\varepsilon^{\bar m_\alpha} P^+(x,y,z,\xi,\varepsilon)$$

of the same form as (2.9), in which M^+ is invertible on its domain of definition, $|(M^+)^{-1}| = O(1)$, and the perturbation P^+ is much smaller on $D(r_+, s_+)$.

Define

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

Through the rest of the paper, all constants $c_i, \ i = 0, 1 \cdots, 10$ are positive constants and independent of ε and the iteration process. We also use c to denote any intermediate constant that is independent of ε and the iteration process. For simplicity, we will also use the asymptotic notions "O, o" whose involving constants are understood to be independent of the iteration process as well.

3.1. **Truncation.** We first express P into the following Taylor-Fourier series

$$P = \sum_{k\in\mathbb{Z}^m, \imath\in\mathbb{Z}^d_+, \jmath\in\mathbb{Z}^{2d}_+} p_{k\imath\jmath} y^\imath z^\jmath \mathrm{e}^{\sqrt{-1}\langle k, x\rangle}$$

and let R be the truncation

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

Lemma 3.1. *Assume that*

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

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

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

Proof. The proof is similar to that of Lemma 3.1 in [14]. □

3.2. **Homological equations.** Next, we eliminate the resonant terms in R by constructing a symplectic transformation generated by a Hamiltonian F of the form

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

where f_{kij}'s are (matrix valued) functions of (ξ,ε) and f_{k02} is symmetric for each k. The main idea of the elimination scheme is to solve the following homological equation

$$
\{N, F\} + \delta\varepsilon^{\bar{m}_\alpha}(R - [R] + \langle p_{001}, z\rangle) = 0, \tag{3.5}
$$

where

$$
[R] \;=\; \frac{1}{(2\pi)^d}\int_{\mathbb{T}^d} R(x,\cdot)\mathrm{d}x. \tag{3.6}
$$

Substituting (3.4) and (3.6) into (3.5) yields

$$
\begin{aligned}
& - \sum_{0<|k|\le K_+} \sqrt{-1}\langle k, \omega_\varepsilon(\xi) + \delta\partial_y h\rangle (f_{k00} + \langle f_{k10}, y\rangle + \langle f_{k01}, z\rangle + \langle y, f_{k20} y\rangle + \langle z, f_{k02} z\rangle)\mathrm{e}^{\sqrt{-1}\langle k,x\rangle} \\
& + \delta \sum_{0<|k|\le K_+} (\langle \varepsilon^{\bar{m}_\alpha} M(\xi,\varepsilon) z, J f_{k01}\rangle + 2\langle \varepsilon^{\bar{m}_\alpha} M(\xi,\varepsilon) z, J f_{k02} z\rangle)\mathrm{e}^{\sqrt{-1}\langle k,x\rangle} + \delta\langle \varepsilon^{\bar{m}_\alpha} M(\xi,\varepsilon) z, J f_{001}\rangle \\
= & - \sum_{0<|k|\le K_+} \delta\varepsilon^{\bar{m}_\alpha}(p_{k00} + \langle p_{k10}, y\rangle + \langle p_{k01}, z\rangle + \langle y, p_{k20} y\rangle + \langle z, p_{k02} z\rangle)\mathrm{e}^{\sqrt{-1}\langle k,x\rangle} - \delta\varepsilon^{\bar{m}_\alpha}\langle p_{001}, z\rangle.
\end{aligned}
$$

Comparing the coefficients in the above, we deduce the following quasi-linear equations for all $0 < |k| \le K_+$, $\xi \in \Lambda^+$:

$$
\begin{aligned}
L_{0k} f_{k00} &= \delta\varepsilon^{\bar{m}_\alpha} p_{k00}, && (3.7)\\
L_{0k} f_{k10} &= \delta\varepsilon^{\bar{m}_\alpha} p_{k10}, && (3.8)\\
L_{0k} f_{k20} &= \delta\varepsilon^{\bar{m}_\alpha} p_{k20}, && (3.9)\\
L_{1k} f_{k01} &= \delta\varepsilon^{\bar{m}_\alpha} p_{k01}, && (3.10)\\
L_{2k} f_{k02} &= \delta\varepsilon^{\bar{m}_\alpha} p_{k02}, && (3.11)\\
M(\xi,\varepsilon) f_{001} &= -p_{001}, && (3.12)
\end{aligned}
$$

where

$$
\begin{aligned}
L_{0k} &= \sqrt{-1}\langle k, \omega_\varepsilon(\xi) + \delta\partial_y h\rangle, \\
L_{1k} &= \sqrt{-1}\langle k, \omega_\varepsilon(\xi) + \delta\partial_y h\rangle I_{2d} - \delta\varepsilon^{\bar{m}_\alpha} M(\xi,\varepsilon) J, \\
L_{2k} &= \sqrt{-1}\langle k, \omega_\varepsilon(\xi) + \delta\partial_y h\rangle I_{4d^2} - (\delta\varepsilon^{\bar{m}_\alpha} M(\xi,\varepsilon) J) \otimes I_{2d} + I_{2d} \otimes (\delta J \varepsilon^{\bar{m}_\alpha} M(\xi,\varepsilon)).
\end{aligned}
$$

In the above, $\otimes$ stands for the tensor product of matrices.

To solve equations (3.7)-(3.9), for each $k = (\hat{k}^0, \cdots, \hat{k}^i, \cdots, \hat{k}^\alpha)^\top$ with $0 < |k| \le K_+$, where $\hat{k}^i \in \mathbb{Z}^{n_i - n_{i-1}}$ for each $i = 0, 1, \cdots, \alpha - 1$ respectively and $\hat{k}^\alpha \in \mathbb{Z}^{m - n_{\alpha-1}}$. We let $\hat{k}^j$, for some $j = 0, 1, \cdots, \alpha$, be the first nonzero components of k. Then

$$
L_{0k} = \varepsilon^{\bar{m}_j}\langle \hat{k}^j, \hat{\omega}^j_\varepsilon(\xi) + O(\delta s)\rangle + \cdots + \varepsilon^{\bar{m}_\alpha}\langle \hat{k}^\alpha, \hat{\omega}^\alpha_\varepsilon(\xi) + O(\delta s)\rangle.
$$

We assume that

H2) $\max\{\varepsilon, \delta s\} K_+^{\tau+1} = o(\gamma_0)$, for all $\xi \in \Lambda^+$. Then

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

Thus L_{0k} is invertible. Combining the estimate

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

with the inductive equations

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

we have

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

It follows that for each $0 < |k| \leq K_+, \ \xi \in \Lambda^+$, equations (3.7)-(3.9) are uniquely solvable to yield solutions $f_{k00}, \ f_{k10}, \ f_{k20}$ which satisfy the following estimates

$$|\partial_\xi^l f_{k00}| \leq c\delta|k|^{|l|+(|l|+1)\tau} s^2\mu e^{-|k|^\tau}, \tag{3.13}$$

$$|\partial_\xi^l f_{k10}| \leq c\delta|k|^{|l|+(|l|+1)\tau} s\mu e^{-|k|^\tau}, \tag{3.14}$$

$$|\partial_\xi^l f_{k20}| \leq c\delta|k|^{|l|+(|l|+1)\tau} \mu e^{-|k|^\tau}, \qquad |l| \leq N. \tag{3.15}$$

To solve the equation (3.10) for each $k = (\hat{k}^0, \cdots, \hat{k}^j, \cdots, \hat{k}^\alpha)^\top$ with $0 < |k| \leq K_+$, we again let $\hat{k}^j$, for some $j = 0, 1, \cdots, \alpha$, be the first nonzero components of k. Then L_{1k} can be re-written into the form

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

where

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

Then (3.10) becomes

$$\tilde{L}_{1k} f_{k01} = \delta\varepsilon^{\bar{m}_\alpha-\bar{m}_j} p_{k01}. \tag{3.16}$$

For each $\xi \in \Lambda^+$, we note by H2) that

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

It follows from the definition of determinant that

$$\begin{aligned} |\det\tilde{L}_{1k}| &\geq \frac{\gamma^{2d}}{2^{2d}|k|^{2d\tau}}(1 - (\frac{2|k|^\tau\delta\varepsilon^{\bar{m}_\alpha-\bar{m}_j}}{\gamma})^2 + \cdots + (\frac{2|k|^\tau\delta\varepsilon^{\bar{m}_\alpha-\bar{m}_j}}{\gamma})^{2d}) \\ &\geq \frac{\gamma^{2d}}{2^{d+1}|k|^{2d\tau}}. \end{aligned}$$

Hence $\tilde{L}_{1k}$ is invertible and

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

The above together with the estimate

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

yields that

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

Hence (3.10) or (3.16) is uniquely solvable on Λ^+ for any $0 < |k| \le K_+$ and some calculations show that the solution f_{k01} satisfies

$$|\partial_\xi^l f_{k01}| \le c\delta|k|^{|l|+(|l|+1)2d\tau} s\mu \mathrm{e}^{-|k|^\tau}, \quad |l| \le N. \tag{3.17}$$

Similarly, (3.11) is uniquely solvable on Λ^+ for any $0 < |k| \le K_+$ and the solution f_{k02} satisfies

$$|\partial_\xi^l f_{k02}| \le c\delta|k|^{|l|+(|l|+1)4d^2\tau} \mu \mathrm{e}^{-|k|^\tau}, \qquad |l| \le N. \tag{3.18}$$

The unique solvability of (3.12) on Λ^+ is obvious and it is easy to see that the solution f_{001} satisfies

$$|\partial_\xi^l f_{001}| \le c\delta s\mu, \qquad |l| \le N. \tag{3.19}$$

Lemma 3.2. *Assume* H2) *and let F be as in* (3.4)*. Then there is an positive constant c_2 such that on $\hat{D}(s) \times \Lambda^+$,*

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

and

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

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

Proof. The lemma follows easily from (3.13)-(3.15) and (3.17)-(3.19). ◻

Lemma 3.3. *Let ϕ_F^t be the flow generated by F and assume*

H3) $c_2\delta\mu\Gamma(r - r_+) < \frac{1}{8}(r - r_+)$,
H4) $c_2\delta s\mu\Gamma(r - r_+)\mu < s_+$.

Then the following holds.

1) *For all $0 \le t \le 1$,*
$$\phi_F^t : D_3 \to D_4$$
are well defined, real analytic and depend smoothly on $\xi \in \Lambda^+$.
2) *Let $\Phi^+ = \phi_F^1$. Then for each $\xi \in \Lambda^+$,*
$$\Phi^+ : D_+ \to D.$$
3) *There is a constant c_3 such that*
$$\begin{aligned} |\partial_\xi^l(\phi_F^t - id)|_{D_+\times\Lambda^+} &\le c_3 s\mu\Gamma(r - r_+), \\ |\partial_\xi^l D^i(\Phi^+ - id)|_{D_+\times\Lambda^+} &\le c_3\mu\Gamma(r - r_+) \end{aligned}$$
for all $|l| \le N,\ 0 \le t \le 1,\ i = 0, 1,$ where $D = \partial_{(x,y,z)}$.

Proof. It easily follows from Lemma 3.2. ◻

3.3. **New Hamiltonian.** Lemma 3.3 shows that for each $\xi \in \Lambda^+$, $\Phi^+ : \ D(r_+, s_+) \to D(r, s)$ is a well defined, real analytic, and symplectic transformation. Applying this transformation to the Hamiltonian H, we obtain the new Hamiltonian

$$\begin{aligned} H^+ \ &=: \ H \circ \Phi^+ = H \circ \phi_F^1 = (N + \delta\varepsilon^{\bar{m}_\alpha} R) \circ \phi_F^1 + \delta\varepsilon^{\bar{m}_\alpha}(P - R) \circ \phi_F^1 \\ &= \ e_\varepsilon^+(\xi) + \langle \omega_\varepsilon^+(\xi), y \rangle + \delta h^+(y, \xi, \varepsilon) + \frac{\delta}{2}\langle \varepsilon^{\bar{m}_\alpha} M^+(\xi, \varepsilon) z, z \rangle + \delta\varepsilon^{\bar{m}_\alpha} P^+ \end{aligned}$$

on $D(r_+, s_+) \times \Lambda^+$, where e_ε^+ is a smooth function on Λ^+ and

$$\omega_\varepsilon^+(\xi) = \omega_\varepsilon(\xi) + \delta\varepsilon^{\bar{m}_\alpha} p_{010}, \tag{3.20}$$

$$M^+(\xi, \varepsilon) = M(\xi, \varepsilon) + p_{002}, \tag{3.21}$$

$$h^+(y, \xi, \varepsilon) = h(y, \xi, \varepsilon) + \varepsilon^{\bar{m}_\alpha} \sum_{|\jmath|=2} p_{0\jmath 0} y^\jmath, \tag{3.22}$$

$$P^+ = \int_0^1 \{(1-t)(R - [R] + \langle p_{001}, z \rangle) + R, F\} \mathrm{d}t + (P - R) \circ \phi_F^1. \tag{3.23}$$

We note that

$$h^+(y, \xi, \varepsilon) = h(y, \xi, \varepsilon) + O(\varepsilon^{\bar{m}_\alpha}|y|^2) = O(|y^{n_0}|^2 + \cdots + \varepsilon^{\bar{m}_j}|y^{n_j}|^2 + \cdots + \varepsilon^{\bar{m}_\alpha}|y|^2).$$

Denote $\omega_\varepsilon^+ := (\hat{\omega}_\varepsilon^{+,0}, \cdots, \varepsilon^{\bar{m}_\alpha}\hat{\omega}_\varepsilon^{+,\alpha})^\top$, where $\hat{\omega}_\varepsilon^{+,j}$ is an $n_j - n_{j-1}$ dimensional vector for each $j = 0, \cdots, \alpha - 1$, and $\hat{\omega}_\varepsilon^{+,\alpha}$ ia an $m - n_{\alpha-1}$ dimensional vector, let

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

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

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

Proof. It follows from (3.20)-(3.22). □

Lemma 3.5. *Assume that*

H5) $c_5\gamma s\mu K_+^{\tau+1} \le \gamma - \gamma_+$.

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

$$|\langle k, \Omega_\varepsilon^+(\xi) \rangle| > \frac{\gamma_+}{|k|^\tau}. \tag{3.24}$$

Proof. Let $0 < |k| \le K_+$, $\xi \in \Lambda^+$. By Lemma 3.4 and H5), we have

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

□

Lemma 3.6. *There exists a constant c_6 such that*

$$|\partial_\xi^l P^+|_{D_+ \times \Lambda^+} \le c_6\gamma^b(s^3\mu^2\Gamma(r - r_+) + s^3\mu^2 + s^2\mu^2), \ |l| \le N,$$

consequently, if

H6) $c_0\gamma^b(s^3\mu^2\Gamma(r-r_+)+s^3\mu^2)\le\gamma_+^b s_+^2\mu_+,$

where $c_0=\max\{1,c_1,\cdots,c_6\}$*, then*

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

Proof. We note that $[R]=O(\delta\varepsilon^{\bar m_\alpha})$. The lemma follows easily from the expression of P^+ as (3.23) and Lemma 3.2. □

This completes one cycle of KAM iterations.

3.4. **Finite steps of KAM iterations.** We now show that the general KAM iterative step outlined above can be carried over for a finite number of steps starting from $\nu=0$ to some $\nu=\nu_*$ to yield a new Hamiltonian normal form with a sufficiently high order of perturbation. Recursively applying the definitions of quantities right before Section 3.1 with $\nu+1$ in place of " + " for $\nu=0,1,\cdots$, we obtain the following iterative sequences:

$$\begin{aligned}
r_\nu &= r_0\Big(1-\sum_{i=1}^{\nu}\frac{1}{2^{i+1}}\Big),\\
s_\nu &= \frac{1}{8}\alpha_{\nu-1}s_{\nu-1},\ \ \alpha_\nu=\mu_\nu^{\frac13},\\
\mu_\nu &= \mu_{\nu-1}^{1+\hat\iota},\quad \text{for some fixed } \hat\iota\in(0,\iota),\\
K_\nu &= \Big(\Big[\log\Big(\frac{1}{\mu_{\nu-1}}\Big)\Big]+1\Big)^3,\\
\Lambda^\nu &= \Big\{\xi\in\Lambda^{\nu-1}:|\langle k,\Omega_\varepsilon^{\nu-1}(\xi)\rangle|>\frac{\gamma_{\nu-1}}{|k|^\tau},\ 0<|k|\le K_\nu\Big\},
\end{aligned}$$

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

It is easy to deduce that

$$\mu_\nu=\mu_0^{(1+\hat\iota)^\nu}=\varepsilon^{\frac{(1-3\iota)}{6}(\bar m_{\alpha+1}-\bar m_\alpha)(1+\hat\iota)^\nu}\le\varepsilon^{\frac{(1-3\iota)}{6}(1+\hat\iota)^\nu},\quad \nu=1,2,\cdots,\tag{3.25}$$

from which the hypotheses H1), H3), H4), H5), H6) can be verified for all $\nu=1,2,\cdots$ as ε is sufficiently small. One part of hypothesis H2) can be also verified form (3.25), i.e.,

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

for all $\nu=1,2,\cdots$. However, the other part of H2), i.e., $\varepsilon K_\nu^{\tau+1}=o(\gamma_0)$ only holds for a finite number of ν's. More precisely, define

$$\nu_*=\left[\frac{\log(2(N+6)c_*+\log 8d^2(N+1)+1)-\log\frac{(1-3\iota)}{6}}{\log(1+\hat\iota)}\right]+1,\tag{3.26}$$

where $c_*=4d^2\sum_{i=1}^{\alpha}\bar m_i(n_i-n_{i-1})$, $[x]$ denotes the maximum integer less than x. Some calculations show that H2) holds as long as $(1+\hat\iota)^{\nu_*}(1-3\iota)\varepsilon^{1-2\iota}$ bounds above by a constant independent of ν_*, consequently, H2) holds as $\varepsilon\ll1$ for all $\nu=1,2,\cdots,\nu_*$.

Hence the iteration scheme can be performed inductively to generate a sequence of Hamiltonians

$$H^\nu=H^{\nu-1}\circ\Phi^\nu=e_\varepsilon^\nu(\xi)+\langle\omega_\varepsilon^\nu(\xi),y\rangle+\delta h^\nu(y,\xi,\varepsilon)+\frac{\delta}{2}\langle\varepsilon^{\bar m_\alpha}M^\nu(\xi,\varepsilon)z,z\rangle+\delta\varepsilon^{\bar m_\alpha}P^\nu(x,y,z,\xi,\varepsilon)$$

defined on $D(r_\nu, s_\nu)\times\Lambda^\nu$, for all $\nu = 1, 2, \cdots, \nu_*$. When $\nu=\nu_*$, we obtain the new Hamiltonian normal form

$$(3.27)\quad H^* =: H^{\nu_*} = e^*_\varepsilon(\xi) + \langle \omega^*_\varepsilon(\xi), y\rangle + \delta h^*(y,\xi,\varepsilon) + \frac{\delta}{2}\langle M^*(\xi,\varepsilon)z, z\rangle + \delta P^*(x,y,z,\xi,\varepsilon)$$

defined on $D(r_*, s_*)\times\Lambda^*$, where $r_* = r_{\nu_*},\ s_* = s_{\nu_*},\ e^*_\varepsilon = e^{\nu_*}_\varepsilon,\ \omega^*_\varepsilon = \omega^{\nu_*}_\varepsilon,\ h^* = h^{\nu_*},\ M^* = \varepsilon^{\bar m_\alpha} M^{\nu_*},\ P^* = \varepsilon^{\bar m_\alpha}P^{\nu_*},\ \Lambda^* = \Lambda^{\nu_*}$.

Since, by (3.26),

$$\mu_{\nu_*} = \mu_0^{(1+\hat\iota)^{\nu_*}} \le \varepsilon^{\frac{(1-3\iota)}{6}(1+\hat\iota)^{\nu_*}} \le \varepsilon^{16d^2(N+1)(N+6)c_*+1},$$

we have

$$(3.28)\qquad |\partial^l_\xi P^*|_{D(r_*,s_*)\times\Lambda^*} \le \varepsilon^{\bar m_\alpha+\iota}s_*^2\varepsilon^{16d^2(N+1)(N+6)c_*} \le \gamma_*^{16d^2(N+1)(N+6)}s_*^2\mu_*^2, \qquad |l|\le N,$$

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

3.5. **Measure estimate.** By A2) and (2.7), we have

$$\mathrm{Rank}\{\partial^l_\xi\Omega^0_\varepsilon:\ 0\le|l|\le N\} = m,\quad \forall\xi\in\Lambda^0.$$

It follows from Lemma 3.4 that

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

Using Lemma 3.5 and the standard measure estimate under Bruno-Rüssmann condition (see [27] or see Lemma 4.2 in section 4) that

$$|\Lambda^0\setminus\Lambda^*| = \sum_{i=1}^{\nu_*}|\Lambda^{i-1}\setminus\Lambda^i| \le \sum_{i=1}^{\nu_*}\sum_{K_i\le|k|\le K_{i+1}}\left(\frac{\gamma_i}{|k|^\tau}\right)^{\frac1N} = O(\gamma_0^{\frac1N}) = O(\varepsilon^{\frac{\iota(\bar m_{\alpha+1}-\bar m_\alpha)}{2bN}}).$$

This, together with (2.8), yields that

$$(3.29)\qquad |\mathcal{O}(g,G)\setminus\Lambda^*| = O(\varepsilon^{\frac{\iota}{2bN}}).$$

4. Proof of main results

In this section, we will perform an infinite number of steps of standard KAM iterations to the normal form (3.27) to prove our main results Theorems A,B. To do so, we consider the following rescalings

$$y\to\gamma_*^{8d^2(N+1)(N+6)}\mu_* y,\quad z\to\sqrt{\mu_*}\gamma_*^{4d^2(N+1)(N+6)}z,\quad H^*\to\frac{H^*}{\gamma_*^{8d^2(N+1)(N+6)}\mu_*}$$

to the normal form (3.27). The re-scaled Hamiltonian reads

$$H_0 := \frac{H^*}{\gamma_*^{8d^2(N+1)(N+6)}\mu_*} := e_0(\xi,\varepsilon) + \langle\omega_0(\xi,\varepsilon), y\rangle + \frac{\delta}{2}\langle M_0(\xi,\varepsilon)z,z\rangle + \delta P_0(x,y,z,\xi,\varepsilon),$$

defined on new region $D(r_0,s_0)\times\Lambda_0$, where $r_0 =: r_*$, $s_0 =: s_*$, $\Lambda_0 = \Lambda^*$, $e_0(\cdot,\varepsilon) = e^*_\varepsilon$, $\omega_0(\cdot,\varepsilon) = \omega^*_\varepsilon$, $M_0 =: M^*$, and

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

Denote $b =: 4d^2(N+1),\ \gamma_0 =: \gamma_*^{2(N+6)},\ \mu_0 =: \mu_*$. Using (3.28) and the fact that

$$h^*(y,\xi,\varepsilon) = O(|y^{n_0}|^2 + \cdots + \varepsilon^{\bar m_j}|y^{n_j}|^2 + \cdots + \varepsilon^{\bar m_\alpha}|y|^2),$$

we have

$$|\partial^l_\xi P_0|_{D(r_0,s_0)\times\Lambda_0} \le \gamma_0^b s_0^2\mu_0,\qquad |l|\le N.$$

4.1. **Iteration and convergence.** We consider the following sequences

$$
\begin{aligned}
& r_\nu = r_0(1-\sum_{i=1}^{\nu}\frac{1}{2^{i+1}}),\\
& s_\nu = \frac{1}{8}\alpha_{\nu-1}s_{\nu-1},\\
& \alpha_\nu = \mu_\nu^{\frac{1}{3}},\\
& \mu_\nu = c_0\mu_{\nu-1}^{\frac{6}{5}},\\
& \gamma_\nu = \gamma_0(1-\sum_{i=1}^{\nu}\frac{1}{2^{i+1}}),\\
& K_\nu = ([\log(\frac{1}{\mu_{\nu-1}})]+1)^{3\eta},\\
& L_{1k,\nu-1} = \sqrt{-1}\langle k,\omega_{\nu-1}(\xi,\varepsilon)\rangle I_{2d} - \delta M_{\nu-1}(\xi,\varepsilon)J, \quad 0<|k|\le K_\nu,\\
& L_{2k,\nu-1} = \sqrt{-1}\langle k,\omega_{\nu-1}(\xi,\varepsilon)\rangle I_{4d^2} - (\delta M_{\nu-1}(\xi,\varepsilon)J)\otimes I_{2d} + I_{2d}\otimes(\delta J M_{\nu-1}(\xi,\varepsilon)), \quad 0<|k|\le K_\nu,\\
& \Lambda_\nu = \{\xi\in\Lambda_{\nu-1} : |\langle k,\omega_{\nu-1}\rangle| > \frac{\gamma_{\nu-1}}{|k|^\tau},\ |\det L_{1k,\nu-1}| > \frac{\gamma_{\nu-1}^{2d}}{|k|^{2d\tau}},\\
& |\det L_{2k,\nu-1}| > \frac{\gamma_{\nu-1}^{4d^2}}{|k|^{4d^2\tau}},\ 0<|k|\le K_\nu\},
\end{aligned}
$$

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

The following iteration lemma and convergence result are special cases of those contained in [14, Section 4].

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

1) *There is a sequence of smooth families of symplectic, real analytic, near identity transformations*
$$\Phi_\xi^\nu : D(r_\nu,s_\nu)\to D(r_{\nu-1},s_{\nu-1}),\ \xi\in\Lambda_\nu$$
such that
$$
\begin{aligned}
H_\nu &= H_{\nu-1}\circ\Phi_\xi^\nu =: N_\nu + \delta P_\nu,\\
N_\nu &= e_\nu(\xi,\varepsilon) + \langle\omega_\nu(\xi,\varepsilon),y\rangle + \frac{\delta}{2}\langle M_\nu(\xi,\varepsilon)z,z\rangle,
\end{aligned}
$$
where
$$
\begin{aligned}
&|\partial_\xi^l\omega_\nu - \partial_\xi^l\omega_0|_{\Lambda_\nu} \le \gamma_0^b\mu_0,\\
&|\partial_\xi^l M_\nu - \partial_\xi^l M_0|_{\Lambda_\nu} \le \gamma_0^b\mu_0,\\
&|\partial_\xi^l P_\nu|_{D_\nu\times\Lambda_\nu} \le \gamma_\nu^b s_\nu\mu_\nu
\end{aligned}
$$
for all $|l|\le N$.

2) $\Lambda_\nu = \{\xi\in\Lambda_{\nu-1} : |\langle k,\omega_{\nu-1}\rangle| > \frac{\gamma_{\nu-1}}{|k|^\tau},\ |\det L_{1k,\nu-1}| > \frac{\gamma_{\nu-1}^{2d}}{|k|^{2d\tau}},\ |\det L_{2k,\nu-1}| > \frac{\gamma_{\nu-1}^{4d^2}}{|k|^{4d^2\tau}},$ $K_{\nu-1} < |k| \le K_\nu\}$.

3) *The Whitney extensions of*
$$\Psi^\nu =: \Phi_\xi^1\circ\Phi_\xi^2\circ\cdots\circ\Phi_\xi^\nu$$

converge C^N uniformly to a Whitney smooth family of symplectic maps, say, Ψ_ε^∞, on $D(\frac{r_0}{2}, \frac{s_0}{2}) \times \Lambda_\infty$, where

$$\Lambda_\infty = \bigcap_{\nu \geq 0} \Lambda_\nu,$$

such that

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

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

$$\partial^j_{(y,z)} P_\infty|_{D(\frac{r_0}{2},0)\times\Lambda_\infty} = 0, \qquad |j| \leq 2.$$

The iteration lemma above shows that, for each $0 < |\varepsilon| \ll 1$ and $\xi \in \Lambda_\infty$, $\mathbb{T}^m \times \{0\} \times \{0\}$ is an 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. ([28, 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 $I_h = \{x : \ |g(x)| \leq h, \ x \in I\}$, $h > 0$. If on I, $|g^{(p)}(x)| \geq D > 0$, where D is a constant, then $|I_h| \leq c_7 h^{\frac{1}{p}}$, 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

$$R_k^{\nu+1}(\xi) = R_{k,0}^{\nu+1} \bigcup R_{k,1}^{\nu+1} \bigcup R_{k,2}^{\nu+1},$$

where

$$\begin{aligned}
R_{k,0}^{\nu+1} &= \{\xi \in \Lambda_\nu : \ |\sqrt{-1}\langle k, \omega_\nu\rangle| \leq \frac{\gamma_\nu}{|k|^\tau}\}, \\
R_{k,1}^{\nu+1} &= \{\xi \in \Lambda_\nu : \ |\det L_{1k,\nu}| \leq \frac{\gamma_{\nu-1}^{2d}}{|k|^{2d\tau}}\}, \\
R_{k,2}^{\nu+1} &= \{\xi \in \Lambda_\nu : \ |\det L_{2k,\nu}| \leq \frac{\gamma_{\nu-1}^{4d^2}}{|k|^{4d^2\tau}}\}.
\end{aligned}$$

Then

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

Consider functions

$$\begin{aligned}
g_{k,0}^\nu(\xi) &= \langle k, \omega_\nu(\xi, \varepsilon)\rangle, \\
g_{k,i}^\nu(\xi) &= \det L_{ik,\nu}, \quad i = 1, 2.
\end{aligned}$$

By Lemma 4.1 1) and Lemma 4.2, it is easy to see that there are positive constants c_8, c_9 such that

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

$$|\frac{\partial^{4d^2 N} g_{k,i}^\nu}{\partial \xi^{4d^2 N}}| \geq c_9 \varepsilon^{c_*},$$

where $a=\sum_{i=1}^{\alpha}\bar{m}_i(n_i-n_{i-1})$ and $c_*=4d^2\sum_{i=1}^{\alpha}\bar{m}_i(n_i-n_{i-1})$. It follows from Lemma 4.2 and Fubini's theorem that there exists a positive constant c_{10} depending only on d such that

$$|R_k^{\nu+1}|\leq c_{10}\frac{\varepsilon^{\frac{12c_*}{N}}}{|k|^{\frac{\tau}{N}}}.$$

Hence by (4.1),

$$|\Lambda_0\setminus\Lambda_\infty|\leq\sum_{\nu=0}^{\infty}\sum_{K_\nu<|k|\leq K_{\nu+1}}|R_k^{\nu+1}|\leq c_{10}\varepsilon^{\frac{12c_*}{N}}\sum_{\nu=0}^{\infty}\sum_{K_\nu<|k|\leq K_{\nu+1}}\frac{1}{|k|^{\frac{\tau}{N}}}=O(\varepsilon^{\frac{12c_*}{N}}).$$

Recall that $\Lambda_0=\Lambda_*$. Combining the above with (3.29) yields that

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

Now, tracing back all symplectic transformations involved in Sections 2,3 and this section, the proof of Theorems A,B are completed.

5. An Example

In this section, we give an example from celestial mechanics applications to show the validity of conditions of A1), (or A1)') A2) with respect to certain resonant type. The validity of the condition A3) in this example depends on the particular form of its perturbation.

Consider the following normalized Hamiltonian in [20] derived from a spatially restricted three-body problem:

$$\begin{aligned} H_\varepsilon &= -I_0-\frac{\varepsilon^3}{2I_0^6}(I_1+I_2)(5(I_1+I_2)^3\mp 4I_0(I_1+I_2)^2+3I_0^2(I_1+I_2)\mp 2I_0^3) \\ &\quad -\frac{3\varepsilon^7(1-\mu)\mu}{8I_0^8}(29I_1^2+20I_1I_2+4I_2^2\mp 2I_0(3I_1+I_2))+\varepsilon^{11}P(I,\theta,\varepsilon), \end{aligned} \tag{5.1}$$

where $I=(I_0,I_1,I_2)^\top\in\mathbb{R}^3$, $\theta=(\theta_0,\theta_1,\theta_2)^\top\in\mathbb{T}^3$, and μ is a fixed constant. The sigh " $\mp$ " appears in an expression involving I_0, the upper sigh applies for $I_0>0$ and the lower sigh applies for $I_0<0$. This Hamiltonian is valid and real analytic in the 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 [20] based on the main result of [11].

To examine the possible existence of lower dimensional tori in the resonance zone, we rewrite (5.1) as

$$H_\varepsilon=H_0(I_0)+\varepsilon^3H_1(I)+\varepsilon^7H_2(I)+\varepsilon^{11}P(I,\theta,\varepsilon),$$

where

$$\begin{aligned} H_0(I_0) &= -I_0, \\ H_1(I) &= -\frac{1}{2I_0^6}(I_1+I_2)(5(I_1+I_2)^3-4I_0(I_1+I_2)^2+3I_0^2(I_1+I_2)-2I_0^3), \\ H_2(I) &= -\frac{3(1-\mu)\mu}{8I_0^8}(29I_1^2+20I_1I_2+4I_2^2-2I_0(3I_1+I_2)). \end{aligned}$$

Here we take the upper sigh of "$\mp$" which implies $I\in G$ with G being a bounded region closed in $\mathbb{R}^+\times\mathbb{R}^2$. Hence (5.1) is a particular case of Hamiltonian (1.4) with $\bar{m}_1=3,\ \bar{m}_2=7,\ \bar{m}_3=11$.

Denote

$$\begin{aligned}
\omega_\varepsilon(I) &= (\frac{\partial H_0}{\partial I_0}+\varepsilon^3\frac{\partial H_1}{\partial I_0}+\varepsilon^7\frac{\partial H_2}{\partial I_0},\varepsilon^3\frac{\partial H_1}{\partial I_1}+\varepsilon^7\frac{\partial H_2}{\partial I_1},\varepsilon^3\frac{\partial H_1}{\partial I_2}+\varepsilon^7\frac{\partial H_2}{\partial I_2})^\top,\\
\Omega^*(I) &= (\frac{\partial H_0}{\partial I_0},\frac{\partial H_1}{\partial I_1},\frac{\partial H_1}{\partial I_2}) =: (\Omega_0^*,\Omega_1^*,\Omega_2^*)^\top.
\end{aligned}$$

We note that, since $\frac{\partial H_0}{\partial I_0}\equiv -1$, the resonance of ω_ε cannot occur at the first integrable term H_0.

We pay particular attention to the following resonant type occurring in H_1, H_2. Let g be a subgroup of $\mathbb{Z}^3$ spanned by $K_2=(0,\hat{K}_2)^\top$, where $\hat{K}_2=(1,-1)^\top$. Let

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

Then $\det(K_1,K_2)=1$.

Let G be an appropriate bounded open set in $\mathbb{R}$. Since

$$\begin{aligned}
\frac{\partial H_1}{\partial I_1} &= \frac{\partial H_1}{\partial I_2},\\
\frac{\partial H_2}{\partial I_1} &= \frac{3\mu(1-\mu)}{8I_0^8}(58I_1+20I_2-6I_0),\\
\frac{\partial H_2}{\partial I_2} &= \frac{3\mu(1-\mu)}{8I_0^8}(20I_1+8I_2-2I_0),
\end{aligned}$$

the g-resonant surface in G reads

$$\mathcal{O}(g,G)=\{I\in G: K_2^\top\Omega^*(I)=0\}=\{I\in G:\ 38I_1+12I_2-4I_0=0\}.$$

To verify the $g-$non-degeneracy of H_2 on $\mathcal{O}(g,G)$, we note that

$$\begin{aligned}
\frac{\partial^2 H_1}{\partial {I_1}^2} &= \frac{\partial^2 H_1}{\partial I_1\partial I_2}=\frac{\partial^2 H_1}{\partial {I_2}^2},\\
\frac{\partial^2 H_2}{\partial {I_1}^2} &= \frac{174\mu(1-\mu)}{8I_0^8},\ \frac{\partial^2 H_2}{\partial I_1\partial I_2}=\frac{60\mu(1-\mu)}{8I_0^8},\\
\frac{\partial^2 H_2}{\partial I_1\partial I_2} &= \frac{60\mu(1-\mu)}{8I_0^8},\ \frac{\partial^2 H_2}{\partial {I_2}^2}=\frac{24\mu(1-\mu)}{8I_0^8}.
\end{aligned}\tag{5.2}$$

It follows that

$$\det \hat{K}_2^\top\frac{\partial^2 H_2}{\partial(\hat{I}^2)^2}\hat{K}_2=\frac{\partial^2 H_2}{\partial I_1^2}+\frac{\partial^2 H_2}{\partial I_2^2}-2\frac{\partial^2 H_2}{\partial I_1 I_2}\neq 0,\qquad \forall I\in\mathcal{O}(g,G),$$

where $\hat{I}^2=(I_1,I_2)^\top$. Hence A1)' holds on $\mathcal{O}(g,G)$.

To verify the Bruno-Rüssmann condition on $\mathcal{O}(g,G)$, we let $(\Omega_0,\Omega_1)^\top := K_1^\top\Omega^*=(\Omega_0^*,-\Omega_2^*)^\top$. Since, by (5.2),

$$\text{Rank}\begin{pmatrix}\Omega_0 & \Omega_1\\ \frac{\partial\Omega_0}{\partial I_0} & \frac{\partial\Omega_1}{\partial I_0}\\ \frac{\partial\Omega_0}{\partial I_1} & \frac{\partial\Omega_1}{\partial I_1}\\ \frac{\partial\Omega_0}{\partial I_2} & \frac{\partial\Omega_1}{\partial I_2}\end{pmatrix}\equiv 2$$

on $\mathcal{O}(g,G)$, the condition A2) holds with $N=1$.

Though validation of assumption A3) depends on the particular expression of $P(I,\theta,0)$, one can be more specific about 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

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

To verify A3), one only needs to find, for a fixed $I_0\in\mathcal{O}(g,G)$, a non-degenerate critical point φ_0 of $h_0(\cdot,I_0)$, because the Implicit Function Theorem then implies the existence a neighborhood U_{I_0} of I_0 and a real analytic family $\psi(I)$ of non-degenerate critical points of $h_0(\cdot,I)$ for $I\in U_{I_0}$. Having done so, it then follows from Theorem B that for each ε sufficiently small there exists a Cantor subset $O_\varepsilon\subset\mathcal{O}_0(g,G):=\mathcal{O}(g,G)\cap U_{I_0}$ such that the unperturbed quasi-periodic 2-tori $T_I^\varepsilon(\psi(I))=\mathbb{T}^2\times\{I\}\times\{\psi(I)\}$, $I\in O_\varepsilon$, of the Hamiltonian (5.1) will persist. Applying Theorem A with $\iota=\frac{1}{4}$, $N=1$, $d=1,\bar{m}_2=7$, $\bar{m}_3=11.$, the excluding measure can be estimated as $|\mathcal{O}_0(g,G)\setminus O_\varepsilon|=O(\varepsilon^{\frac{1}{16}})$.

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] L. Chierchia, G.Gallavotti, Drift and diffusion phase space, *Ann. Inst. H. Poincaré Phy. Th.* **69** (1994), 1-144.

[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] F. Cong, T. Küpper, Y. Li, and J. You, KAM-type theorem on resonant surfaces for nearly integrable Hamiltonian systems, *J. Nonl. Sci.,* **10**(2000), 49-68.

[6] R. H. Cushman, A survey of normalization techniques applied to perturbed Keplerian systems, Dynamics Reported: Expositions in dynamical systems (Jones et al Ed.), Springer, Berlin, Heidelberg, 1992.

[7] L.H. Eliasson, Biasymptotic solutions of perturebed intergrable Hamiltonian systems, *Bol. Sco. Mat.,* **25** (1994) 57-76.

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

[9] J. Féjoz, Proof of 'Arnold's theorem' on the stability of a planetary system (following Herman), *Ergod. Th. & Dynam. Sys.* **24** (2004), 1521-1582.

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

[11] Y. C. 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.

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

[13] M. Iñarrea, 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.
[14] Y. Li and Y. Yi, A quasi-periodic Poincaré's Theorem, *Math. Annalen,* **326** (2003), 649-690.
[15] Y. Li and Y. Yi, On Poincaré-Treshchev tori in Hamiltonian systems, *Proc. Equadiff 2003*, Dumortier et al (Ed.), World Scientific, 2005, 136-151.
[16] Y. Li and Y. Yi, Persistence of lower dimensional tori of general types in Hamiltonian systems, *Trans. Amer. Math. Soc.* **357** (2005), 1565-1600.
[17] K. R. Meyer, Scaling Hamiltonian systems, *SIAM J. Math. Anal.* **15** (1984), 877-889.
[18] K. R. Meyer, Periodic Solutions of the N-Body Problem, Springer-Verlag, Berlin, Heidelberg, 1999.
[19] K. R. Meyer and D. S. Schmidt, From the restricted to the full three-body problem, *Trans. Amer. Math. Soc.* **352** (2000), 2283-2299.
[20] 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.
[21] 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.
[22] H. Poincaré, Les Méthodes Nouvelles de la Mécaniques Céleste,I-III, Gauthier-Villars, 1892, 1893, 1899. (The English translation: New Methods of Celestial Mechanics, AIP Press, Williston, 1992)
[23] M. Rudnev, S.Wiggins, KAM theory near multiplicity one resoant surfaces in perturbations of A-priori stable Hamiltonian systems, *J. Nonl. Sci.* **7** (1997), 177-209.
[24] F. Sayas, Averaging, reduction and reconstruction in the spatial three-body problem, Ph.D thesis, Universidad Pública de Navarra, 2015.
[25] B. Sommer, A KAM Theorem for the Spatial Lunar Problem, Ph.D thesis, 2003.
[26] D.V. Treshchev, The mechanism of destruciton of resonant tori of Hamiltonian systems, *Math. USSR Sb.* **68** (1991), 181-203.
[27] J. Xu and J. You, Corrigendum for the paper Invariant tori for nearly integrable Hamiltonian systems with degeneracy, *Math. Z.* **257** (2007), 939.
[28] 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, School of Mathematics, Jilin University, Changchun, 130012, PRC, and School of Mathematics, Georgia Institute of Technology, Atlanta, GA 30332, USA

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