

## PERSISTENCE OF LOWER DIMENSIONAL TORI OF GENERAL TYPES IN HAMILTONIAN SYSTEMS

YONG LI AND YINGFEI YI

ABSTRACT. The work is a generalization to [40] in which we study the persistence of lower dimensional tori of general type in Hamiltonian systems of general normal forms. By introducing a modified linear KAM iterative scheme to deal with small divisors, we shall prove a persistence result, under a Melnikov type of non-resonance condition, which particularly allows multiple and degenerate normal frequencies of the unperturbed lower dimensional tori.

### 1. INTRODUCTION

The present work concerns the study of the persistence of lower dimensional tori in perturbative, partially integrable Hamiltonian systems of the following normal form:

$$(1.1) \quad H = e(\omega) + \langle \omega, y \rangle + \frac{1}{2} \left\langle \begin{pmatrix} y \\ z \end{pmatrix}, M(\omega) \begin{pmatrix} y \\ z \end{pmatrix} \right\rangle + h(y, z, \omega) + P(x, y, z, \omega),$$

where  $(x, y, z) \in T^n \times R^n \times R^{2m}$ ,  $\omega$  is a parameter in a bounded closed region  $\mathcal{O} \subset R^n$ ,  $M(\omega)$  is a  $(n + 2m) \times (n + 2m)$  real symmetric for each  $\omega \in \mathcal{O}$ ,  $h = O(|(y, z)|^3)$ ,  $P$  is the perturbation whose smallness will be specified later, and both  $h$  and  $P$  are real analytic in  $(x, y, z)$  for  $(x, y, z)$  lying in a complex neighborhood  $D(r, s) = \{(x, y, z) : |\operatorname{Im} x| < r, |y| < s, |z| < s\}$  of  $T^n \times \{0\} \times \{0\}$ . In the above, all  $\omega$  dependences are of class  $C^{l_0}$  for some  $l_0 \geq 4m^2$ , and the derivatives of  $h$  and  $P$  up to order  $l_0$  are uniformly bounded on their domain of definitions.

With the symplectic form

$$\sum_{i=1}^n dx_i \wedge dy_i + \sum_{j=1}^m dz_j \wedge dz_{m+j},$$

the equation of motion associated to (1.1) reads

$$\begin{cases} \dot{x} &= \omega + M_{11}y + M_{12}z + \frac{\partial h}{\partial y} + \frac{\partial P}{\partial y}, \\ \dot{y} &= -\frac{\partial P}{\partial x}, \\ \dot{z} &= JM_{22}z + JM_{21}y + J\frac{\partial h}{\partial z} + J\frac{\partial P}{\partial z}, \end{cases}$$

---

2000 *Mathematics Subject Classification.* 37J40.

*Key words and phrases.* Hamiltonian systems, invariant tori, KAM theory, Melnikov problem, persistence.

The first author was partially supported by NSFC grant 19971042, National 973 Project of China: Nonlinearity, the outstanding young's project of Ministry of Education of China, and National outstanding young's award of China.

The second author was partially supported by NSF grants DMS9803581 and DMS-0204119.

where  $M_{11}, M_{12}, M_{21}, M_{22}$  denote the  $n \times n, n \times 2m, 2m \times n, 2m \times 2m$  blocks of  $M = M(\omega)$  respectively, and  $J$  is the standard  $2m \times 2m$  symplectic matrix. Clearly, as  $\omega$  varies in  $\mathcal{O}$ , the unperturbed system associated to (1.1) admits a family of invariant  $n$ -tori  $T_\omega = T^n \times \{0\} \times \{0\}$  which are parameterized by the toral frequency  $\omega \in \mathcal{O}$ .

Hamiltonian (1.1) can be viewed as a canonical form in studying the persistence of lower dimensional tori in Hamiltonian systems. Under suitable symplectic coordinates, typical Hamiltonian systems in the vicinity of a family of lower dimensional tori (e.g., those near an equilibrium point) may be reduced to a system of form (1.1) in which there is a natural dependence of  $M$  on the tori, or a natural coupling of  $M$  with the toral frequencies  $\omega$  (see Section 6 for more details).

Along the line of the classical KAM theorem (Kolmogorov [17], Arnold [1], Moser [28]), the persistence problem of lower dimensional tori for (1.1) is to find appropriate non-degenerate and non-resonance conditions on  $M$  and  $\omega$ , respectively, so that the majority (with respect to the Lebesgue measure on  $\mathcal{O}$ ) of these tori, associated to  $\omega \in \mathcal{O}$ , will persist after small perturbations.

Such persistence problem has been studied in various normally non-degenerate cases (i.e.  $M_{22}$  is non-singular over  $\mathcal{O}$ ). A persistence theorem of elliptic tori was formulated by Melnikov ([25, 26]) for Hamiltonian systems of the form (1.1) in which  $M_{11}(\omega), M_{12}(\omega), M_{21}(\omega) \equiv 0$  and  $M_{22}(\omega)$  is a diagonal matrix satisfying the so-called Melnikov's second non-resonance condition, i.e., the set

$$(1.2) \quad \{\omega \in \mathcal{O} : \langle k, \omega \rangle + \langle l, \Lambda \rangle \neq 0, \text{ for all } (k, l) \in (Z^n \times Z^{2m}) \setminus \{0\}, |l| \leq 2\}$$

admits full Lebesgue measure relative to  $\mathcal{O}$ , where  $\Lambda = \Lambda(\omega)$  is the column vector formed by the eigenvalues (diagonal) of  $M_{22}$ . The Melnikov's theorem was first proved by Eliasson in [11]. Refinements of the result with generalizations to infinite dimension were made by Kuksin ([19, 20]) and Pöschel ([32, 33]). We note that, with the Melnikov's second non-resonance condition, all normal frequencies have to be simple. Recently, Bourgain ([2, 3]) obtained a sharp persistence result (in both finite and infinite dimensions) of elliptic tori in the Melnikov's setting only under the Melnikov's first non-resonance condition (i.e.,  $|l| \leq 1$  in (1.2)), which therefore allows the multiplicity of normal frequencies.

Following a pioneer work of Moser ([27]), there were also studies on the persistence of normally non-degenerate, lower dimensional tori of general types. For example, Broer, Huitema and Takens ([5]), Broer, Huitema and Sevryuk ([6, 7]), Sevryuk ([34]) studied cases that  $M_{11}(\omega), M_{12}(\omega), M_{21}(\omega) \equiv 0$  and all eigenvalues of  $JM_{22}(\omega)$  are simple (hence  $JM_{22}(\omega)$  can be smoothly and symplectically diagonalized) and satisfy a Melnikov's second non-resonance condition. Jorba and Villanueva ([16]) considered a case with generic, non-degenerate perturbation that,  $M_{12}(\omega), M_{21}(\omega) \equiv 0$ ,  $M_{11}(\omega)$  is non-singular over  $\mathcal{O}$ , and  $JM_{22}(\omega)$  is diagonal with distinct eigenvalues  $\pm\lambda_i(\omega)$ ,  $i = 1, 2, \dots, m$ , satisfying a similar Melnikov second non-resonance condition. Recently, You ([40]) considered the same persistence problem for (1.1) with  $M_{11}(\omega), M_{12}(\omega), M_{21}(\omega) \equiv 0$ , but under the following weaker Melnikov second non-resonance condition on the eigenvalues  $\lambda_i$ ,  $i = 1, 2, \dots, 2m$ , of  $JM_{22}$ :

$$(1.3) \quad \sqrt{-1}\langle k, \omega \rangle - \lambda_i(\omega) - \lambda_j(\omega) \neq 0, \text{ for all } k \in Z^n \setminus \{0\}, 1 \leq i, j \leq 2m,$$

which in particular allows the multiplicity of these eigenvalues.

The persistence problem for hyperbolic tori was first considered in the work of Moser ([27]) and later studied by Graff ([12]) and Zehnder ([41]) for a fixed Diophantine toral frequency. Hyperbolic case with degeneracy and varying frequencies is recently considered by the authors in [22].

The aim of the present study is to generalize the result of You ([40]) to allow normal degeneracy (i.e., the singularity of  $M_{22}(\omega)$ ) of the unperturbed lower dimensional tori. We shall show under the same weaker Melnikov second non-resonance condition (1.3) that the majority of unperturbed lower dimensional tori in (1.1) will persist if either  $M_{22}(\omega)$  or  $M(\omega)$  is everywhere non-singular. In the latter case, the normal degeneracy of the lower dimensional tori can be made to an extreme. For example, one can simply take  $M_{22} \equiv 0$  if  $M(\omega)$  is everywhere non-singular.

The proof of our results uses the classical KAM procedure. However, to be able to handle a general normal form like (1.1) especially when normal degeneracy occurs, a modified KAM linear iterative scheme will be introduced in order to construct the desired symplectic transformations. An essential point of this scheme is to treat the tangential variable  $y$  and the normal variable  $z$  in a same scale rather than the traditional way of treating  $z$  as a much smaller variable than  $y$ . Having done so, the present problem is in spirit closely related to the KAM problem in Hamiltonian systems with more action than angle variables (we refer the readers to [24, 30] and references therein for KAM theory in Hamiltonian or Poisson-Hamilton systems with a distinct number of action and angle variables).

The persistence problem of lower dimensional tori can also be considered in the resonance zone of a nearly integrable Hamiltonian system in which the lower dimensional tori are naturally split from a family of unperturbed, full dimensional, resonant tori. However, the persistence problem of this nature depends on the Poincaré non-degenerate condition rather than conditions of Melnikov type. In this case the normal form near the lower dimensional tori, derived from splitting of the resonance, is similar to (1.1) but with  $M$  being in the same scale as the perturbation (see [9, 21, 23, 37] for details).

The paper is organized as follows. In Section 2, we describe our main result followed by some comments and remarks. We use the modified KAM linear scheme in Section 3 to construct the symplectic transformation for one KAM cycle and give estimates to the transformation and the transformed Hamiltonian. In Section 4, an iteration lemma will be proved to ensure the validity of all KAM steps. We shall complete the proof of our result in Section 5 by showing the convergence of iterative sequences and estimating the measure of the limiting frequency set. Some applications of our results will be discussed in Section 6 for quasi-periodic motions of Hamiltonian systems near equilibria, in the spirit of the Lyapunov center theorem.

Unless specified otherwise, we shall use the same symbol  $|\cdot|$  to denote an equivalent vector norm and its induced matrix norm, absolute value of functions, and measure of sets etc., and use  $|\cdot|_D$  to denote the sup-norm of functions on a domain  $D$ . Also, for any two complex column vectors  $\xi, \zeta$  of same dimension,  $\langle \xi, \zeta \rangle$  always stands for  $\xi^\top \zeta$ , i.e., the transpose of  $\xi$  times  $\zeta$ . For the sake of brevity, we shall not specify smoothness orders for functions having obvious orders of smoothness indicated by their derivatives taken.

We would like to thank the referee for valuable comments and suggestions which led to a significant improvement of the paper. This work was partially done when

the first author was visiting the Center of Dynamical Systems and Nonlinear Studies, Georgia Institute of Technology, and the second author was visiting the National University of Singapore, in 1998. We would also like to thank both Institutions for support.

## 2. STATEMENT OF MAIN RESULTS

Consider (1.1) and let  $\lambda_1(\omega), \dots, \lambda_{2m}(\omega)$  be eigenvalues of  $JM_{22}(\omega)$ . We assume the weak form of the Melnikov second non-resonance condition (1.3), i.e.,

**NR)** *The set*

$$\{\omega \in \mathcal{O} : \sqrt{-1}\langle k, \omega \rangle - \lambda_i(\omega) - \lambda_j(\omega) \neq 0, \text{ for all } k \in Z^n \setminus \{0\}, 1 \leq i, j \leq 2m\}$$

*admits full Lebesgue measure relative to  $\mathcal{O}$ .*

For  $\gamma > 0$ , let  $\hat{\mathcal{O}}_\gamma$  denote the Diophantine set

$$\{\omega \in \mathcal{O} : |\langle k, \omega \rangle| > \frac{\gamma}{|k|^\tau}, \text{ for all } k \in Z^n \setminus \{0\}\},$$

where  $|k| = \sum_{i=1}^n |k_i|$  and  $\tau > n - 1$  is fixed. Our main results are stated as follows.

**Theorem 1.** *Suppose NR) and that either  $M(\omega)$  or  $M_{22}(\omega)$  is non-singular on  $\mathcal{O}$ . For given  $r, s > 0$ , a positive integer  $l_0$  and  $\sigma_0 \in (0, 1)$ , if there is a sufficient small  $\mu = \mu(r, s, l_0, \sigma_0) > 0$  such that*

$$(2.1) \quad |\partial_\omega^l \partial_x^i \partial_{(y,z)}^j P| \leq \gamma^{3(1+\sigma_0)a} \mu$$

*for all  $(x, y, z) \in D(r, s)$ ,  $\omega \in \mathcal{O}$ , and  $(l, i, j) \in Z_+^n \times Z_+^n \times Z_+^{n+2m}$  with  $|l| + |i| + |j| \leq l_0$ , where  $a = 4m^2(l_0 + 1)$ , then there exist Cantor-like sets  $\mathcal{O}_\gamma$ , with  $|\mathcal{O} \setminus \mathcal{O}_\gamma| \rightarrow 0$  as  $\gamma \rightarrow 0$ , on which the following holds.*

- 1) *If  $M_{22}(\omega)$  is non-singular on  $\mathcal{O}$ , then there exists a  $C^{l_0-1}$  Whitney smooth family of real analytic, symplectic transformations*

$$\Psi_\omega : D\left(\frac{r}{2}, \frac{s}{2}\right) \rightarrow D(r, s), \quad \omega \in \mathcal{O}_\gamma,$$

*which is  $C^{l_0}$  uniformly close to the identity, such that*

$$(2.2) \quad H \circ \Psi_\omega = e_*(\omega) + \langle \Omega_*(\omega), y \rangle + \frac{1}{2} \left\langle \begin{pmatrix} y \\ z \end{pmatrix}, M_*(\omega) \begin{pmatrix} y \\ z \end{pmatrix} \right\rangle + h_*(y, z, \omega) + P_*(x, y, z, \omega),$$

*where  $e_*, \Omega_*, M_*$  are  $C^{l_0-1}$  Whitney smooth with  $|\partial_\omega^l e_* - \partial_\omega^l e|_{\mathcal{O}_\gamma}, |\partial_\omega^l (\Omega_* - id)|_{\mathcal{O}_\gamma}, |\partial_\omega^l M_* - \partial_\omega^l M|_{\mathcal{O}_\gamma} = O(\gamma^\alpha \mu^{\frac{1}{4}})$ ,  $|l| \leq l_0 - 1$ ,  $h_*(y, z, \omega) = O(|(y, z)|^3)$ ,  $h_*$  and  $P_*$  are real analytic in their associated phase variables and  $C^{l_0-1}$  Whitney smooth in the parameter  $\omega$ , and*

$$\partial_y^j \partial_z^k P_*|_{(y,z)=(0,0)} = 0$$

*for all  $x \in T^n$ ,  $\omega \in \mathcal{O}_\gamma$ ,  $j \in Z_+^n$ ,  $k \in Z_+^{2m}$  with  $|j| + |k| \leq 2$ . Moreover,  $\Omega_*(\mathcal{O}_\gamma) \subset \hat{\mathcal{O}}_\gamma$ .*

*Thus, for each  $\omega \in \mathcal{O}_\gamma$ , the unperturbed torus  $T_\omega$  persists and gives rise to a slightly deformed, analytic, Floquet, quasi-periodic, invariant torus of the perturbed system with the Diophantine toral frequency  $\Omega_*(\omega)$ . These perturbed tori also form a  $C^{l_0-1}$  Whitney smooth family.*

- 2) If  $M(\omega)$  is non-singular on  $\mathcal{O}$ , then 1) holds with  $\Omega_* = id$ , i.e., for each  $\omega \in \mathcal{O}_\gamma \subset \hat{\mathcal{O}}_\gamma$ , the perturbed torus also preserves the unperturbed (Diophantine) toral frequency  $\omega$ .

**Theorem 2.** *Suppose that the Hamiltonian (1.1) is also real analytic in parameter  $\omega$ . Then the conclusions of Theorem 1 hold for  $a = 4m^2$  and any  $l_0 > 1$ , provided that the condition (2.1) is replaced by*

$$(2.3) \quad |P| \leq \gamma^{3(1+\sigma_0)a} \mu, \quad (x, y, z) \in D(r, s), \quad \omega \in \mathcal{O}.$$

**Remark:**

- 1) Part 1) of Theorem 1 becomes the result of You ([40]) when  $M_{11}, M_{12}, M_{21}, h \equiv 0$ . There are certainly cases in application that some of these terms are non-zero and are not able to be eliminated via a smooth family of canonical transformations or via rescaling (see Section 6). In the case that  $M_{22}$  is everywhere non-singular, our results hold for arbitrary smooth matrices  $M_{11}, M_{12} = M_{21}^\top$  and apply to non-degenerate lower invariant tori of all types. Similar to [40], the results also allow multiple eigenvalues of  $JM_{22}$  (or  $M_{22}$ ), for example, when  $k \neq 0$ ,  $i$  can be equal to  $j$  in NR).
- 2) Part 2) of Theorem 1 is of a particular interest in the presence of normal degeneracy. As a simple example, we let  $M(\omega)$  be any smooth, non-singular, real symmetric,  $(n + 2m) \times (n + 2m)$  matrix on  $\mathcal{O}$  such that  $M_{22}(\omega) \equiv 0$  (in particular, all normal eigenvalues equal to 0), say,

$$M = \begin{pmatrix} 0 & 0 & \cdots & 0 & 1 \\ 0 & 0 & \cdots & 1 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 1 & \cdots & 0 & 0 \\ 1 & 0 & \cdots & 0 & 0 \end{pmatrix}.$$

Then the non-resonance condition NR) holds automatically, which therefore assures the persistence of  $n$ -tori when the perturbation is sufficiently small.

With respect to normal degeneracy, an essential point of the result in part 2) of Theorem 1 is that possible normal degeneracy can be compensated by the overall non-degeneracy of  $M(\omega)$ . Unless there are further non-degenerate constraints in the perturbation, such non-degeneracy is generally needed in the presence of normal degeneracy to eliminate first order resonant terms in the Hamiltonian, for otherwise, there would not be any persistent  $n$ -torus. For example, in (1.1) if  $M_{12} = M_{21}^\top \equiv 0$ ,  $M_{22} \equiv 0$ ,  $h \equiv 0$ ,  $P = \varepsilon \langle \xi, z \rangle$ , where  $\xi$  is a non-zero  $2m$  vector, then for any matrix  $M_{11}(\omega)$ , no invariant  $n$ -torus can exist when  $\varepsilon \neq 0$ .

- 3) Like the classical KAM case, Theorem 1 concludes that perturbed tori are actually Floquet (i.e., the quadratic terms in their normal forms (2.2) are independent of angle variables) - a useful structure characterizing the normal spectra of the perturbed tori which may be used in studying problems such as bifurcation from tori, stability and integrability etc. (see [7]).

The non-resonance condition NR) seems to be a best possible condition in general for the persistence of Floquet lower dimensional tori in Hamiltonians of form (1.1).

A parallel first non-resonance condition of Melnikov-Bourgain type with respect to the general setting (1.1) can be described as follows: on a frequency set of full measure,

$$(2.4) \quad \sqrt{-1}\langle k, \omega \rangle - \lambda_i(\omega) \neq 0, \quad \text{for all } k \in \mathbb{Z}^n \setminus \{0\}, 1 \leq i \leq 2m,$$

where  $\lambda_i$ 's are as in NR). Thus, the condition NR), when restricted to the elliptic case, is weaker than the Melnikov's second condition but is stronger than the first condition.

At this point, it is not clear to us whether it is possible to generalize the Bourgain's theorem and our results to obtain a persistence result for general type of non-degenerate, lower dimensional tori in Hamiltonian system of general form (1.1), by only imposing the first non-resonance condition (2.4). Even if it does, we suspect that the perturbed tori under the first non-resonance condition (2.4) would in general be non-Floquet. This is because, as far as KAM type of symplectic transformations are concerned, the first non-resonance condition (2.4) is not sufficient for the elimination of the angular dependence among quadratic order perturbative terms of a Hamiltonian system.

Some exciting developments of the quasi-periodic Floquet theory are made recently (e.g., [10, 13, 15, 18, 29, 38]) concerning the reducibility of certain linear systems with quasi-periodic coefficients including some cases of quasi-periodic linear Hamiltonian systems. With these developments, one might obtain the Floquet tori alternatively by first obtaining the invariant tori then reducing their normal parts to Floquet forms. However, since the quasi-periodic Floquet theory gives no particular information on the reducibility near a given invariant torus, we feel that the weak second Melnikov condition NR) is still needed for the reducibility near all invariant tori corresponding to the set  $\mathcal{O}_\gamma$ . Besides, the Hamiltonian system near each of these perturbed tori will not be quasi-periodically forced in general.

- 4) Part 2) of Theorem 1 also gives a general result on the preservations of toral frequencies, which has been known mainly in the classical KAM case and hyperbolic cases (see [12, 41]). However, it does not mean the persistence of any fixed lower dimensional torus satisfying the Melnikov condition NR).

Nevertheless, unless there are further non-degenerate constraints in the perturbation, the non-degeneracy of  $M$  assumed in part 2) of Theorem 1 is a sharp condition for the preservation of toral frequencies in Hamiltonian systems of form (1.1) in the sense stated in the theorem. Indeed consider (1.1) and let  $M = M(\omega)$  be any  $(n + 2m) \times (n + 2m)$  real symmetric matrix depending smoothly on  $\omega$  and satisfying NR). Take  $h \equiv 0, P = \varepsilon(\langle y_0, y \rangle + \langle z_0, z \rangle)$ , where  $\varepsilon \neq 0$  is sufficiently small and  $y_0 = y_0(\omega) \in \mathbb{R}^n, z_0 = z_0(\omega) \in \mathbb{R}^{2m}$ . Then the equation of motions associated to the Hamiltonian becomes

$$\begin{cases} \dot{x} &= \omega + M_{11}y + M_{12}z + \varepsilon y_0, \\ \dot{y} &= 0, \\ \dot{z} &= JM_{22}z + JM_{21}y + \varepsilon Jz_0. \end{cases}$$

Suppose that the Hamiltonian admits a family of invariant  $n$ -tori  $T_{\varepsilon, \omega}$  which preserves the toral frequencies  $\omega$  of the corresponding unperturbed tori. Let  $(x(t), y(t), z(t))^\top$  be an orbit on a perturbed torus  $T_{\varepsilon, \omega}$ . Then  $y$  is a constant

and  $z$  must be a quasi-periodic function with basic frequencies chosen from components of  $\omega$ . Since NR) implies the condition (2.4) (see Lemma 5.1 in Section 5), i.e., none of the eigenvalues of  $JM_{22}$  can rationally depend on the basic frequencies, it is easy to see that  $z$  must be a constant solution of the last equation above, i.e.,

$$(2.5) \quad M_{22}z + M_{21}y + \varepsilon z_0 = 0.$$

Substituting the constant solutions  $y, z$  into the  $x$ -equation above, we see that the frequencies of  $x$  will be drifted unless

$$(2.6) \quad M_{11}y + M_{12}z + \varepsilon y_0 = 0.$$

Now, if  $M$  is singular on a positive measure set  $\hat{\mathcal{O}} \subset \mathcal{O}$ , then by choosing  $(y_0, z_0)^\top$  as a basis vector of  $\ker M(\omega)$  (which varies smoothly in  $\omega$ ) we see that the system  $\{(2.5), (2.6)\}$  admits no solutions on  $\hat{\mathcal{O}}$  for any  $\varepsilon$ , which leads to a contradiction. In other word, we have shown that, given any  $(n + 2m) \times (n + 2m)$  real symmetric matrix  $M$  depending smoothly on  $\omega$  and satisfying NR), if  $M$  is singular on a positive measure set, then there exists an arbitrary small perturbation which preserves no toral frequency lying in a positive measure set.

### 3. KAM STEP

We shall only prove Theorem 1 in the sequel. The proof of Theorem 2 can be carried out by similar arguments along with the standard Cauchy estimate.

Due to the presence of small divisors, one cannot remove all  $x$ -dependent terms in a finite number of steps. An essential idea of the KAM theory is to construct a symplectic transformation, consisting of infinitely many successive steps (referred to as *KAM steps*) of iterations, so that the  $x$ -dependent terms are pushed into higher order perturbations after each step. As all KAM steps can be carried out inductively, below, we only show detailed constructions of the symplectic transformation for one KAM cycle (i.e., from one KAM step to the successive one).

Initially, we set  $N_0 = N$ ,  $e_0 = e$ ,  $\Omega_0 = id$ ,  $M^0 = M$ ,  $M_{22}^0 = M_{22}$ ,  $h_0 = h$ ,  $P_0 = P$ ,  $\mathcal{O}_0 = \mathcal{O}$ ,  $r_0 = r$ ,  $\beta_0 = s$ ,  $s_0 = \gamma^{(1+\sigma_0)a} \mu^{\frac{1}{4}}$ ,  $\mu_0 = \gamma^{\sigma_0} \mu^{\frac{1}{2}}$ ,  $\gamma_0 = \gamma$ . Without loss of generality, we assume that  $0 < r_0, \beta_0, \mu_0, \gamma_0 \leq 1$ ,  $s_0 \leq \beta_0$  and write  $\mu_* = \mu$ . For  $j \in \mathbb{Z}_+^{n+2m}$ , define

$$\begin{aligned} a_j &= a(1 - \operatorname{sgn}(|j|)\operatorname{sgn}(|j| - 1)\operatorname{sgn}(|j| - 2)) = \begin{cases} a, & |j| = 0, 1, 2, \\ 0, & |j| \geq 3, \end{cases} \\ \kappa_j &= 1 - \operatorname{sgn}(|j| - 1) = \begin{cases} 2, & |j| = 0, \\ 1, & |j| = 1, \\ 0, & |j| \geq 2, \end{cases} \\ d_j &= 1 - \lambda_0 \operatorname{sgn}(|j|)\operatorname{sgn}(|j| - 1)\operatorname{sgn}(|j| - 2) = \begin{cases} 1, & |j| = 0, 1, 2, \\ 1 - \lambda_0, & |j| \geq 3, \end{cases} \end{aligned}$$

where  $\lambda_0 \in (0, 1)$  is fixed. Then by (2.1),

$$|\partial_\omega^l \partial_x^i \partial_{(y,z)}^j P_0|_{D(r_0, s_0) \times \mathcal{O}_0} \leq \gamma_0^{a_j} s_0^{\kappa_j} \mu_0^{d_j}, \quad |l| + |i| + |j| \leq l_0.$$

Suppose that after a  $\nu$ th KAM step, we arrive at a real analytic, parameter-dependent Hamiltonian

$$(3.1) \quad \begin{aligned} H &= H_\nu = N + P, \\ N &= N_\nu = e + \langle \Omega(\omega), y \rangle + \frac{1}{2} \left\langle \begin{pmatrix} y \\ z \end{pmatrix}, M(\omega) \begin{pmatrix} y \\ z \end{pmatrix} \right\rangle + h(y, z, \omega), \end{aligned}$$

where  $(x, y, z) \in D = D_\nu = D(r, s)$ ,  $r = r_\nu \leq r_0$ ,  $s = s_\nu \leq s_0$ ,  $\omega \in \mathcal{O} = \mathcal{O}_\nu \subset \mathcal{O}_0$ ,  $e(\omega) = e_\nu(\omega)$ ,  $\Omega(\omega) = \Omega_\nu(\omega)$  are smooth on  $\mathcal{O}$ ,  $M(\omega) = M^\nu(\omega)$  is real symmetric and smooth on  $\mathcal{O}$ ,  $h = h_\nu(y, z, \omega) = O(|(y, z)|^3)$ ,  $h$  and  $P = P_\nu(x, y, z, \omega)$  are real analytic in  $(y, z) \in \mathcal{D} = \mathcal{D}(s) = \{|y| < s, |z| < s\}$  and in  $(x, y, z) \in D$  respectively and smooth in  $\omega \in \mathcal{O}$ , and moreover,

$$(3.2) \quad |\partial_\omega^l \partial_x^i \partial_{(y,z)}^j P|_{D \times \mathcal{O}} \leq \gamma^{a_j} s^{\kappa_j} \mu^{d_j}, \quad |l| + |i| + |j| \leq l_0,$$

for some  $\mu = \mu_\nu > 0$ ,  $\gamma = \gamma_\nu > 0$ .

We wish to construct a symplectic transformation  $\Phi_+ = \Phi_{\nu+1}$ , which, in smaller frequency and phase domains, should carry the above Hamiltonian into the next KAM cycle. More precisely, the transformed Hamiltonian should be of a form similar to (3.1) with a smaller perturbation term satisfying an inequality similar to (3.2). Thereafter, quantities (domains, normal form, perturbation, etc.) in the next KAM cycle will be simply indexed by  $+$  ( $=\nu+1$ ), and we shall not specify the dependence of  $M, M_+, P, P_+$  etc. on their arguments. Also, all constants  $c_1 - c_5$  below are positive and independent of the iteration process. For simplicity, we also use  $c$  to denote any intermediate positive constant which is independent of the iteration process.

Let  $b, \sigma, d, \delta$  be positive constants such that

$$\begin{aligned} \delta &= 1 - d, \quad \sigma - (b + \sigma)(4b + 5\sigma) > 0, \quad \delta(1 + b + \sigma) > 1, \\ \sigma_0 - \frac{\sigma}{b + \sigma}(1 + \sigma_0)a &> 0, \quad (1 + \sigma_0)(1 + b + \sigma)(1 - (b + 2\sigma)) > 1. \end{aligned}$$

Define

$$\begin{aligned} r_+ &= \delta r + d\left(1 - \frac{\delta^2}{2}\right)r_0, \\ s_+ &= s^{1+b+\sigma}, \\ \beta_+ &= \frac{\beta}{2} + \frac{\beta_0}{4}, \\ \gamma_+ &= \frac{\gamma}{2} + \frac{\gamma_0}{4}, \\ K_+ &= (\lceil \log \frac{1}{s} \rceil + 1)^3, \\ \mathcal{D}_+ &= \mathcal{D}(\beta_+), \\ D_+ &= D(r_+, s_+), \\ \tilde{D}_+ &= D\left(r_+ + \frac{5}{8}(r - r_+), \beta_+\right), \\ D_i &= D\left(r_+ + \frac{i-1}{8}(r - r_+), is_+\right), \quad i = 1, 2, \dots, 8, \end{aligned}$$



$$\begin{aligned}
D(\lambda) &= D\left(r_+ + \frac{7}{8}(r - r_+), \lambda\right), \\
\Gamma(\lambda) &= e^{\frac{r_0(1-\delta)\delta^2}{16}} \sum_{0 < |k| \leq K_+} |k|^\chi e^{-\frac{\lambda}{8}},
\end{aligned}$$

where  $\lambda > 0$ ,  $\chi = 2(l_0 + 1)(2m^2 + 1)\tau$ , and  $\tau > n - 1$  is fixed.

**3.1. Outline of the construction.** The transformation  $\Phi_+$  will be constructed in two steps. The first step is to average out terms in  $P$  up to quadratic order. First order resonant terms will then be removed in the second step by a translation of coordinate.

We first consider the averaging process.

Let  $R$  be the truncation of the Taylor-Fourier series of  $P$  up to quadratic order, i.e.,

$$\begin{aligned}
(3.3) \quad R &= \sum_{|k| \leq K_+, |i|+|j| < 3} p_{kij} y^i z^j e^{\sqrt{-1}\langle k, x \rangle} \\
&= \sum_{|k| \leq K_+} (P_{k00} + \langle P_{k10}, y \rangle + \langle P_{k01}, z \rangle + \langle y, P_{k20} y \rangle \\
&\quad + \langle z, P_{k11} y \rangle + \langle z, P_{k02} z \rangle) e^{\sqrt{-1}\langle k, x \rangle},
\end{aligned}$$

where  $K_+$  is the truncation order in  $x$  to be specified later.

We shall seek for an averaging transformation as the time 1-map  $\phi_F^1$  of the flow  $\phi_F^t$  generated by a Hamiltonian  $F$  of the form

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

where  $F_{kij}$ ,  $0 \leq i + j \leq 2$ , are  $\omega$ -dependent vectors or matrices of obvious dimensions.

We wish to determine  $F$  through the following linear homological equation:

$$(3.5) \quad \{N, F\} + R - [R] - Q = 0,$$

where  $[R] = \int_{T^n} R(x, \cdot) dx$  is the average of  $R$  with respect to  $x \in T^n$  and  $Q$  consists of all terms in  $\{N, F\}$  of size  $O(s^3\mu)$  including those of order  $O(y^i z^j)$  for  $|i| + |j| \geq 3$ . We note that, by taking the term  $Q$  into account, (3.5) modifies the usual linear homological equation adopted in standard KAM linear schemes.

At the moment, let us suppose that (3.5) is solvable on a suitable frequency domain. Then

$$\begin{aligned}
H \circ \phi_F^1 &= (N + R) \circ \phi_F^1 + (P - R) \circ \phi_F^1 \\
&= N + [R] + (\{N, F\} + R - [R] - Q) + \int_0^1 (1-t) \{\{N, F\}, F\} \circ \phi_F^t dt \\
&\quad + \int_0^1 \{R, F\} \circ \phi_F^t dt + (P - R) \circ \phi_F^1 + Q \\
&= N + [R] + \int_0^1 \{R_t, F\} \circ \phi_F^t dt + (P - R) \circ \phi_F^1 + Q,
\end{aligned}$$

where

$$(3.6) \quad R_t = (1-t)\{N, F\} + R = (1-t)(Q + [R] - R) + R.$$

Therefore, if we let

$$(3.7) \quad \begin{aligned} \bar{N}_+ &= N + [R], \\ \bar{P}_+ &= \int_0^1 \{R_t, F\} \circ \phi_F^t dt + (P - R) \circ \phi_F^1 + Q, \end{aligned}$$

then

$$H \circ \phi_F^1 = \bar{N}_+ + \bar{P}_+.$$

This completes the averaging process.

Write  $M$  into blocks:

$$M = \begin{pmatrix} M_{11} & M_{12} \\ M_{21} & M_{22} \end{pmatrix},$$

where  $M_{11}, M_{12}, M_{21}, M_{22}$  are  $n \times n, n \times 2m, 2m \times n, 2m \times 2m$  blocks of  $M$  respectively.

As for the second step, we need to eliminate the first order resonant terms in  $[R]$ . Assume the following condition:

**H1)** *If  $M_{22}^0$  ( $M^0$  resp.) is non-singular on  $\mathcal{O}_0$ , then so is  $M_{22}$  ( $M$  resp.), and moreover,*

$$|M_{22}^{-1}|_{\mathcal{O}} \leq 2|(M_{22}^0)^{-1}|_{\mathcal{O}_0} \quad (|M^{-1}|_{\mathcal{O}} \leq 2|(M^0)^{-1}|_{\mathcal{O}_0} \text{ resp.}).$$

Let  $(y_0, z_0)$  satisfy

$$(3.8) \quad (M(\omega) + \partial_{(y,z)}^2 h(y_0, z_0, \omega)) \begin{pmatrix} y_0 \\ z_0 \end{pmatrix} + \partial_{(y,z)} h(y_0, z_0, \omega) = - \begin{pmatrix} P_{010} \\ P_{001} \end{pmatrix}$$

if  $M$  is non-singular on  $\mathcal{O}$ , or

$$(3.9) \quad \begin{aligned} y_0 = 0, \quad & (\text{diag}(O, M_{22}(\omega)) + \partial_{(y,z)}^2 h(y_0, z_0, \omega)) \begin{pmatrix} y_0 \\ z_0 \end{pmatrix} \\ & + \partial_{(y,z)} h(y_0, z_0, \omega) = - \begin{pmatrix} 0 \\ P_{001} \end{pmatrix} \end{aligned}$$

if  $M_{22}$  is non-singular on  $\mathcal{O}$ . By the implicit function theorem,  $(y_0, z_0) = (y_0(\omega), z_0(\omega))$  exist and depend smoothly on  $\omega$ .

Let

$$(3.10) \quad \begin{aligned} e_+ &= e + P_{000} + h(y_0, z_0, \omega) + \langle \Omega_+(\omega), y_0 \rangle + \langle P_{001}, z_0 \rangle \\ &+ \frac{1}{2} \left\langle \begin{pmatrix} y_0 \\ z_0 \end{pmatrix}, \begin{pmatrix} P_{020} & P_{011} \\ P_{011}^\top & P_{002} \end{pmatrix} \begin{pmatrix} y_0 \\ z_0 \end{pmatrix} \right\rangle, \end{aligned}$$

$$(3.11) \quad \Omega_+ = \Omega + P^*,$$

$$M^+ = M + 2 \begin{pmatrix} P_{020} & P_{011} \\ P_{011}^\top & P_{002} \end{pmatrix},$$

$$(3.12) \quad \begin{aligned} h_+ &= h(y + y_0, z + z_0, \omega) - h(y_0, z_0, \omega) - \langle \partial_{(y,z)} h(y_0, z_0, \omega), \begin{pmatrix} y \\ z \end{pmatrix} \rangle \\ &- \frac{1}{2} \left\langle \begin{pmatrix} y \\ z \end{pmatrix}, \partial_{(y,z)}^2 h(y_0, z_0, \omega) \begin{pmatrix} y \\ z \end{pmatrix} \right\rangle, \end{aligned}$$

where

$$P^* = \begin{cases} P_{010}, & \text{if } M_{22} \text{ is non-singular on } \mathcal{O}, \\ 0, & \text{if } M \text{ is non-singular on } \mathcal{O}. \end{cases}$$

We consider the translation

$$\phi : x \rightarrow x, \begin{pmatrix} y \\ z \end{pmatrix} \rightarrow \begin{pmatrix} y \\ z \end{pmatrix} + \begin{pmatrix} y_0 \\ z_0 \end{pmatrix}.$$

Then  $\Phi_+ = \phi_F^1 \circ \phi$  will transform  $H$  into the Hamiltonian  $H_+$  in the next KAM cycle, i.e.,

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

where

$$\begin{aligned} (3.13) \quad N_+ &= \bar{N}_+ \circ \phi - 2 \left\langle \begin{pmatrix} y \\ z \end{pmatrix}, \begin{pmatrix} P_{020} & P_{011} \\ P_{011}^\top & P_{002} \end{pmatrix} \begin{pmatrix} y_0 \\ z_0 \end{pmatrix} \right\rangle \\ &= e_+ + \langle \Omega_+(\omega), y \rangle + \frac{1}{2} \left\langle \begin{pmatrix} y \\ z \end{pmatrix}, M^+ \begin{pmatrix} y \\ z \end{pmatrix} \right\rangle + h_+(y, z, \omega), \end{aligned}$$

$$(3.14) \quad P_+ = \bar{P}_+ \circ \phi + 2 \left\langle \begin{pmatrix} y \\ z \end{pmatrix}, \begin{pmatrix} P_{020} & P_{011} \\ P_{011}^\top & P_{002} \end{pmatrix} \begin{pmatrix} y_0 \\ z_0 \end{pmatrix} \right\rangle.$$

**3.2. Solving the linear homological equation.** In (3.5), we let

$$\begin{aligned} (3.15) \quad Q &= -\sqrt{-1} \sum_{0 < |k| \leq K_+} \langle k, M_{11}y + M_{12}z + \frac{\partial h}{\partial y} \rangle (F_{k00} + \langle F_{k10}, y \rangle + \langle F_{k01}, z \rangle \\ &\quad + \langle y, F_{k20}y \rangle + \langle z, F_{k11}y \rangle + \langle z, F_{k02}z \rangle) e^{\sqrt{-1}\langle k, x \rangle} \\ &\quad - \sum_{0 < |k| \leq K_+} \left\langle \frac{\partial h}{\partial z}, J(F_{k01} + F_{k11}y + F_{k02}z + F_{k02}^\top z) \right\rangle e^{\sqrt{-1}\langle k, x \rangle}. \end{aligned}$$

Then substitutions of (3.3), (3.4), (3.15) into (3.5) yield

$$\begin{aligned} &- \sum_{0 < |k| \leq K_+} \sqrt{-1} \langle k, \Omega \rangle (F_{k00} + \langle F_{k10}, y \rangle + \langle F_{k01}, z \rangle \\ &\quad + \langle y, F_{k20}y \rangle + \langle z, F_{k11}y \rangle + \langle z, F_{k02}z \rangle) e^{\sqrt{-1}\langle k, x \rangle} \\ &\quad + \sum_{0 < |k| \leq K_+} \langle M_{21}y + M_{22}z, J(F_{k01} + F_{k11}y + F_{k02}z + F_{k02}^\top z) \rangle e^{\sqrt{-1}\langle k, x \rangle} \\ &= - \sum_{0 < |k| \leq K_+} (P_{k00} + \langle P_{k10}, y \rangle + \langle P_{k01}, z \rangle \\ (3.16) \quad &+ \langle y, P_{k20}y \rangle + \langle z, P_{k11}y \rangle + \langle z, P_{k02}z \rangle) e^{\sqrt{-1}\langle k, x \rangle}. \end{aligned}$$

By comparing the coefficients of (3.16), we obtain the following linear equations for all  $0 < |k| \leq K_+$ :

$$(3.17) \quad \sqrt{-1}\langle k, \Omega \rangle F_{k00} = P_{k00},$$

$$(3.18) \quad \sqrt{-1}\langle k, \Omega \rangle F_{k10} + M_{12} J F_{k01} = P_{k10},$$

$$(3.19) \quad (\sqrt{-1}\langle k, \Omega \rangle I_{2m} - M_{22} J) F_{k01} = P_{k01},$$

$$(3.20) \quad \sqrt{-1}\langle k, \Omega \rangle F_{k20} + \frac{1}{2}(M_{12}^\top J F_{k11} - F_{k11}^\top J M_{21}) = P_{k20},$$

$$(3.21) \quad (\sqrt{-1}\langle k, \Omega \rangle - M_{22} J) F_{k11} + (F_{k02}^\top + F_{k02}) J M_{21} = P_{k11},$$

$$(3.22) \quad \sqrt{-1}\langle k, \Omega \rangle F_{k02} - M_{22} J F_{k02} + F_{k02} J M_{22} = P_{k02}.$$

For any given matrix  $A = (a_{ij})_{pq}$ , let  $\sigma(A)$  denote the  $pq$ -vector which lines up the row vectors of  $A$  from left to the right, i.e.,

$$\sigma(A) = (a_{11} \cdots a_{1q} \cdots a_{p1} \cdots a_{pq})^\top.$$

Then for any matrices  $A, B, C$  such that  $ABC$  is well defined, we have

$$(3.23) \quad \sigma(ABC) = (A \otimes C^\top) \sigma(B),$$

where  $\otimes$  denotes the usual tensor product of matrices, i.e.,

$$(3.24) \quad U \otimes V = (u_{ij} V), \quad U = (u_{ij})$$

for any two matrices  $U, V$ .

For  $0 < |k| \leq K_+$ , define

$$\begin{aligned} F_k &= \sigma(F_{k02}), \\ P_k &= \sigma(P_{k02}), \\ L_{0k} &= \sqrt{-1}\langle k, \Omega \rangle, \\ L_{1k} &= \sqrt{-1}\langle k, \Omega \rangle I_{2m} - M_{22} J, \\ L_{2k} &= \sqrt{-1}\langle k, \Omega \rangle I_{4m^2} - (M_{22} J) \otimes I_{2m} - I_{2m} \otimes (M_{22} J). \end{aligned}$$

Using (3.23), (3.24) and the above notions, the linear equations (3.17)-(3.22) can be rewritten equivalently as

$$(3.25) \quad L_{0k} F_{k00} = P_{k00},$$

$$(3.26) \quad L_{1k} F_{k01} = P_{k01},$$

$$(3.27) \quad L_{0k} F_{k10} = P_{k10} - M_{12} J F_{k01},$$

$$(3.28) \quad L_{0k} F_{k20} = P_{k20} + \frac{1}{2}(F_{k11}^\top J M_{21} - M_{12}^\top J F_{k11}),$$

$$(3.29) \quad L_{1k} F_{k11} = P_{k11} - (F_{k02}^\top + F_{k02}) J M_{21},$$

$$(3.30) \quad L_{2k} F_k = P_k,$$

$0 < |k| \leq K_+$ , which are clearly solvable as long as all  $L_{0k}, L_{1k}, L_{2k}$  are invertible.

Consider the set

$$(3.31) \quad \begin{aligned} \mathcal{O}_+ = \mathcal{O}(K_+) &= \{\omega \in \mathcal{O} : |L_{0k}| > \frac{\gamma}{|k|^\tau}, \quad |\det L_{1k}| > \frac{\gamma^{2m}}{|k|^{2m\tau}}, \\ &|\det L_{2k}| > \frac{\gamma^{4m^2}}{|k|^{4m^2\tau}}, \quad \text{for all } 0 < |k| \leq K_+\}. \end{aligned}$$

Then, on  $\mathcal{O}_+$ , equations (3.25)-(3.30) can be solved uniquely to obtain solutions  $F_{kij}$ , depending smoothly on  $\omega \in \mathcal{O}_+$  and satisfying  $\bar{F}_{kij} = F_{-kij}$ , for all  $0 < |k| \leq K_+$ ,  $0 \leq i + j \leq 2$ , which uniquely determine the Hamiltonian  $F$  in (3.4).

### 3.3. Estimate on the truncation $R$ .

**Lemma 3.1.** *Assume that*

$$\mathbf{H2)} \quad s_+ \leq \frac{s}{16};$$

$$\mathbf{H3)} \quad \int_{K_+}^{\infty} \lambda^{n+l_0} e^{-\lambda \frac{r-r_+}{16}} d\lambda \leq s.$$

Then there is a constant  $c_1$  such that for all  $|l| + |i| + |j| \leq l_0$ ,  $\omega \in \mathcal{O}$ ,

$$|\partial_\omega^l \partial_x^i \partial_{(y,z)}^j (P - R)|_{D_8} \leq c_1 \gamma^{a_j} (s^{\kappa_j+1} + s^{\kappa_j-3} s_+^3) \mu^{d_j}.$$

*Proof.* Denote

$$\begin{aligned} I &= \sum_{|k| > K_+} p_{kij} y^i z^j e^{\sqrt{-1}\langle k, x \rangle}, \\ II &= \sum_{|k| \leq K_+, |i| + |j| \geq 3} p_{kij} y^i z^j e^{\sqrt{-1}\langle k, x \rangle} \\ &= \int \frac{\partial^{(p,q)}}{\partial y^p \partial z^q} \sum_{|k| \leq K_+, |i| + |j| \geq 3} p_{kij} e^{\sqrt{-1}\langle k, x \rangle} y^i z^j dy dz, \end{aligned}$$

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

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

Since, by H2),  $D_8 \subset D(s)$ , and by Cauchy's estimate and (3.2),

$$\begin{aligned} |\partial_\omega^l \partial_{(y,z)}^j \sum_{i \in \mathbb{Z}_+^n, j \in \mathbb{Z}_+^{2m}} p_{kij} y^i z^j| &\leq |\partial_\omega^l \partial_{(y,z)}^j P|_{D(r,s)} e^{-|k|r} \\ &\leq \gamma^{a_j} s^{\kappa_j} \mu^{d_j} e^{-|k|r}, \quad |y|, |z| < s, \end{aligned}$$

for all  $k$  and  $|l| + |j| \leq l_0$ , H3) implies that

$$\begin{aligned} |\partial_\omega^l \partial_x^i \partial_{(y,z)}^j I|_{D(s)} &\leq \sum_{|k| \geq K_+} \gamma^{a_j} |k|^{|i|} s^{\kappa_j} \mu^{d_j} e^{-|k|r} e^{|k|(r_+ + \frac{7}{8}(r-r_+))} \\ &\leq \gamma^{a_j} s^{\kappa_j} \mu^{d_j} \sum_{\eta=K_+}^{\infty} \eta^{n+l_0} e^{-\eta \frac{r-r_+}{8}} \\ &\leq \gamma^{a_j} s^{\kappa_j} \mu^{d_j} \int_{K_+}^{\infty} \lambda^{n+l_0} e^{-\lambda \frac{r-r_+}{16}} d\lambda \\ &\leq \gamma^{a_j} s^{\kappa_j+1} \mu^{d_j}, \quad |l| + |i| + |j| \leq l_0. \end{aligned}$$

It follows that

$$\begin{aligned} |\partial_\omega^l \partial_x^i \partial_{(y,z)}^j (P - I)|_{D(s)} &\leq |\partial_\omega^l \partial_x^i \partial_{(y,z)}^j P|_{D(r,s)} + |\partial_\omega^l \partial_x^i \partial_{(y,z)}^j I|_{D(s)} \\ &\leq 2\gamma^{a_j} s^{\kappa_j} \mu^{d_j}, \quad |l| + |i| + |j| \leq l_0. \end{aligned}$$

By performing Cauchy's estimate of  $\partial_\omega^l \partial_x^i \partial_{(y,z)}^j (P - I)$  on  $D(s)$ , we have

$$\begin{aligned}
& |\partial_\omega^l \partial_x^i \partial_{(y,z)}^j II|_{D_8} \\
& \leq |\partial_\omega^l \partial_x^i \partial_{(y,z)}^j \int \frac{\partial^{(p,q)}}{\partial y^p \partial z^q} \sum_{|k| \leq K_+, |i|+|j| \geq 3} p_{kij} y^i z^j e^{\sqrt{-1}\langle k, x \rangle} dy dz|_{D_8} \\
& \leq \left| \int \left| \frac{\partial^{(p,q)}}{\partial y^p \partial z^q} \partial_\omega^l \partial_x^i \partial_{(y,z)}^j (P - I) \right| dy dz \right|_{D_8} \\
& \leq 2 \left( \frac{1}{s - 8s_+} \right)^3 \gamma^{a_j} s^{\kappa_j} \mu^{d_j} \left| \int dy dz \right|_{D_8} \\
& \leq c \gamma^{a_j} \mu^{d_j} s^{\kappa_j - 3} s_+^3, \quad |l| + |i| + |j| \leq l_0.
\end{aligned}$$

Thus,

$$|\partial_\omega^l \partial_x^i \partial_{(y,z)}^j (P - R)|_{D_8} \leq c \gamma^{a_j} (s^{\kappa_j + 1} + s^{\kappa_j - 3} s_+^3) \mu^{d_j}, \quad |l| + |i| + |j| \leq l_0.$$

□

#### 3.4. Estimate on the transformation $\Phi_+$ .

**Lemma 3.2.** *Assume H3) and also that*

$$\mathbf{H4)} \quad |\partial_\omega^l M - \partial_\omega^l M^0|_{\mathcal{O}}, \quad |\partial_\omega^l (\Omega - id)|_{\mathcal{O}} \leq \mu_*^{\frac{1}{4}}, \quad |l| \leq l_0.$$

*If  $F$  is the Hamiltonian defined in Sections 3.1, 3.2, then there is a constant  $c_2$  such that*

$$(3.32) \quad |\partial_\omega^l \partial_x^i \partial_{(y,z)}^j F|_{D(s) \times \mathcal{O}_+} \leq c_2 s^{\kappa_j} \mu \Gamma(r - r_+),$$

$$(3.33) \quad |\partial_\omega^l \partial_x^i \partial_{(y,z)}^j F|_{D(\beta) \times \mathcal{O}_+} \leq c_2 \mu \Gamma(r - r_+)$$

for all  $0 < |k| \leq K_+$ ,  $|l| \leq l_0$ ,  $|i| + |j| \leq l_0 + 1$ .

*Proof.* Let  $q = 0, 1, 2$ ,  $0 < |k| \leq K_+$ . By H4),

$$(3.34) \quad |\partial_\omega^{l'} L_{qk}|_{\mathcal{O}_+} \leq c |k|, \quad |l'| \leq l_0,$$

and by (3.31),

$$\begin{aligned}
|L_{0k}^{-1}|_{\mathcal{O}_+} &= \left| \frac{1}{\langle k, \Omega(\omega) \rangle} \right|_{\mathcal{O}_+} \leq c \frac{|k|^\tau}{\gamma}, \\
|L_{1k}^{-1}|_{\mathcal{O}_+} &= \left| \frac{\text{adj} L_{1k}}{\det L_{1k}} \right|_{\mathcal{O}_+} \leq c \frac{|k|^{2m\tau + 2m - 1}}{\gamma^{2m}}, \\
|L_{2k}^{-1}|_{\mathcal{O}_+} &= \left| \frac{\text{adj} L_{2k}}{\det L_{2k}} \right|_{\mathcal{O}_+} \leq c \frac{|k|^{4m^2\tau + 4m^2 - 1}}{\gamma^{4m^2}},
\end{aligned}$$

i.e.,

$$(3.35) \quad |L_{qk}^{-1}|_{\mathcal{O}_+} \leq c \frac{|k|^{(2m)^q \tau + (2m)^q - 1}}{\gamma^{(2m)^q}}.$$

Using (3.34), (3.35) and applying the identity

$$\partial_\omega^{l'} L_{qk}^{-1} = - \sum_{|l''|=1}^{|l'|} \binom{l'}{l''} (\partial_\omega^{l' - l''} L_{qk}^{-1} \partial_\omega^{l''} L_{qk}) L_{qk}^{-1}$$

inductively, it is easy to see that

$$(3.36) \quad |\partial_\omega^{l'} L_{qk}^{-1}| \leq c \frac{|k|^{|l'|+(|l'+1)(2m)^q\tau}}{\gamma^{(|l'+1)(2m)^q}}, \quad |l'| \leq l_0.$$

By Cauchy's estimate, we also have

$$(3.37) \quad |\partial_\omega^l P_{kij}|_{\mathcal{O}} \leq |\partial_\omega^l P|_{D(r,s) \times \mathcal{O}} s^{-(i+j)} e^{-|k|r} \leq \gamma^a s^{2-i-j} \mu e^{-|k|r}, \quad 0 \leq i+j \leq 2.$$

It now follows from (3.26)-(3.30), (3.36), (3.37) that

$$\begin{aligned} |\partial_\omega^l F_{kij}|_{\mathcal{O}_+} &\leq c \frac{|k|^{|l|+(|l+1)4m^2\tau}}{\gamma^{(|l+1)4m^2}} \gamma^a s^{2-i-j} \mu e^{-|k|r} \\ &\leq c |k|^{|l|+(|l+1)4m^2\tau} s^{2-i-j} \mu e^{-|k|r}, \quad |l| \leq l_0, \quad 0 \leq i+j \leq 2. \end{aligned}$$

Since  $F$  is of quadratic orders in  $y, z$ , we have, on  $D(s) \times \mathcal{O}_+$ , that

$$\begin{aligned} |\partial_\omega^l \partial_x^i \partial_{(y,z)}^j F| &\leq c \sum_{0 < |k| \leq K_+} |k|^{|i|} (|\partial_\omega^l F_{k00}| + |\partial_\omega^l F_{k10}| s^{1-\text{sgn}(|j|)} \\ &\quad + |\partial_\omega^l F_{k01}| s^{1-\text{sgn}(|j|)} + |\partial_\omega^l F_{k20}| s^{1-\text{sgn}(|j|-1)} + |\partial_\omega^l F_{k11}| s^{1-\text{sgn}(|j|-1)} \\ &\quad + |\partial_\omega^l F_{k02}| s^{1-\text{sgn}(|j|-1)}) e^{|k|(r_+ + \frac{7}{8}(r-r_+))} \\ &\leq c s^{\kappa_j} \mu \sum_{0 < |k| \leq K_+} |k|^{\chi} e^{-|k|\frac{r-r_+}{8}} \leq c s^{\kappa_j} \mu \Gamma(r-r_+), \quad |l| \leq l_0, \quad |i| + |j| \leq l_0 + 1. \end{aligned}$$

This proves (3.32). The proof of (3.33) is similar.  $\square$

**Lemma 3.3.** *Assume H1)-H4) and also that*

$$\mathbf{H5)} \quad c_2 \mu \Gamma(r-r_+) < \frac{1}{8}(r-r_+),$$

$$\mathbf{H6)} \quad c_2 s \mu \Gamma(r-r_+) < s_+,$$

$$\mathbf{H7)} \quad c_2 \mu \Gamma(r-r_+) + c_2 \mu < \beta - \beta_+.$$

1) Let  $\phi_F^t$  be the flow generated by  $F$ . Then for all  $0 \leq t \leq 1$

$$\begin{aligned} \phi_F^t : D_3 &\rightarrow D_4, \\ \phi : D_1 &\rightarrow D_3 \end{aligned}$$

are well defined, real analytic and depend smoothly on  $\omega \in \mathcal{O}_+$ .

2) Let  $\Phi_+ = \phi_F^1 \circ \phi$ . Then for all  $\omega \in \mathcal{O}_+$ ,

$$(3.38) \quad \Phi_+ : \begin{cases} D_+ \rightarrow D, \\ \tilde{D}_+ \rightarrow D(r, \beta). \end{cases}$$

3) There is a constant  $c_3$  such that for all  $0 \leq t \leq 1$ ,  $|l| \leq l_0$ ,

$$(3.39) \quad |\partial_\omega^l \partial_x^i \partial_{(y,z)}^j \phi_F^t|_{D_3 \times \mathcal{O}_+} \leq \begin{cases} c_3 s \mu \Gamma(r-r_+), & |i| + |j| = 0, \quad |l| \geq 1, \\ c_3 \mu \Gamma(r-r_+), & 2 \leq |l| + |i| + |j| \leq l_0 + 1, \\ c_3, & \text{otherwise,} \end{cases}$$

$$(3.40) \quad |\partial_\omega^l \partial_\xi^p (\Phi_+ - id)|_{\tilde{D}_+ \times \mathcal{O}_+} \leq c_3 \mu \Gamma(r-r_+), \quad |p| \leq l_0 + 1,$$

$$(3.41) \quad |\partial_\omega^l \phi|_{D_+ \times \mathcal{O}_+} \leq c_3 \gamma^a s \mu,$$

where  $\xi = (x, y, z)$ .

*Proof.* Let  $\phi_{F1}^t, \phi_{F2}^t, \phi_{F3}^t$  be components of  $\phi_F^t$  in  $x, y, z$  planes respectively. We note that

$$(3.42) \quad \phi_F^t = \text{id} + \int_0^t X_F \circ \phi_F^\lambda d\lambda,$$

where  $X_F = (F_y, -F_x, JF_z)^\top$  denotes the vector field generated by  $F$ .

For any  $(x, y, z) \in D_3$ , let  $t_* = \sup\{t \in [0, 1] : \phi_F^t(x, y, z) \in D_4\}$ . Since, by H2),  $D_4 \subset D(s)$ , it follows from H5)-H7) and (3.32) that

$$\begin{aligned} |\phi_{F1}^t(x, y, z)| &= |x| + \left| \int_0^t F_y \circ \phi_F^\lambda d\lambda \right| \leq |x| + |F_y|_{D(s)} \\ &\leq r_+ + \frac{2}{8}(r - r_+) + c_2 s \mu \Gamma(r - r_+) < r_+ + \frac{3}{8}(r - r_+), \\ |\phi_{F2}^t(x, y, z)| &= |y| + \left| - \int_0^t F_x \circ \phi_F^\lambda d\lambda \right| \leq |y| + |F_x|_{D(s)} \\ &\leq 3s_+ + c_2 s^2 \mu \Gamma(r - r_+) < 4s_+, \\ |\phi_{F3}^t(x, y, z)| &= |z| + \left| \int_0^t JF_z \circ \phi_F^\lambda d\lambda \right| \leq |z| + |F_z|_{D(s)} \\ &\leq 3s_+ + c_2 s^2 \mu \Gamma(r - r_+) < 4s_+, \end{aligned}$$

i.e.,  $\phi_F^t(x, y, z) \in D_4$  for all  $0 \leq t \leq t_*$ . Thus,  $t_* = 1$  and  $\phi_F^t(x, y, z) \in D_4$ ,  $0 \leq t \leq 1$ .

Next, we note by H3), H4) and (3.37) that

$$|\partial_\omega^l z_0|_{\mathcal{O}_+}, |\partial_\omega^l y_0|_{\mathcal{O}_+} \leq c_2 \gamma^a s \mu, \quad |l| \leq l_0,$$

where  $y_0, z_0$  are as in (3.8). It follows from H2) that  $\phi : D_1 \rightarrow D_3$  is well defined and satisfies (3.41). Thus,  $\Phi_+ : D_+ \rightarrow D$  is well defined. Using H5), H7), (3.33) and a similar argument as above, one sees that  $\phi_F^t : \hat{D}_+ \rightarrow D(r, \beta)$  is well defined for all  $0 \leq t \leq 1$ , where  $\hat{D}_+ = D(r_+ + \frac{5}{8}(r - r_+), \beta_+ + c_2 \mu)$ . Thus,  $\Phi_+ : \tilde{D}_+ \rightarrow D(r, \beta)$  is also well defined.

As the proofs for (3.39) and (3.40) are similar, we only consider (3.40).

Observe that

$$\Phi_+ - \text{id} = (\phi_F^1 - \text{id}) \circ \phi + \begin{pmatrix} 0 \\ y_0 \\ z_0 \end{pmatrix}.$$

Let  $\omega \in \mathcal{O}_+$ ,  $0 \leq t \leq 1$ . By (3.41), to prove (3.40), it suffices to show that

$$|\partial_\omega^l \partial_\xi^p (\phi_F^t - \text{id})|_{\hat{D}_+} \leq c \mu \Gamma(r - r_+), \quad |p| \leq l_0 + 1, \quad |l| \leq l_0.$$

Note that

$$(3.43) \quad |D^{l'} X_F|_{\tilde{D}_0 \times \mathcal{O}_+} \leq c |D_\xi^{l'+1} F|_{\tilde{D}_0 \times \mathcal{O}_+}, \quad |l'| \leq l_0.$$

Using (3.42) and (3.33), we immediately have

$$|\phi_F^t - \text{id}|_{\hat{D}_+} \leq c \mu \Gamma(r - r_+).$$

Differentiating (3.42) yields

$$D_\xi \phi_F^t = I_{2n+2m} + \int_0^t (DX_F) D_\xi \phi_F^\lambda d\lambda = I_{2n+2m} + \int_0^t J(D_\xi^2 F) D_\xi \phi_F^\lambda d\lambda.$$



It follows from (3.33) and the Gronwall's inequality that

$$\begin{aligned} |D_\xi \phi_F^t - I_{2n+2m}|_{\hat{D}_+} &\leq \int_0^t |D_\xi^2 F|_{D(\beta)} |D_\xi \phi_F^\lambda - I_{2n+2m}|_{\hat{D}_+} d\lambda + |D_\xi^2 F|_{D(\beta)} \\ &\leq \int_0^t |D_\xi^2 F|_{D(\beta)} |D_\xi \phi_F^\lambda - I_{2n+2m}|_{\hat{D}_+} d\lambda + c_2 \mu \Gamma(r - r_+) \leq c \mu \Gamma(r - r_+). \end{aligned}$$

Similarly, by using (3.33), the Gronwall's inequality and (3.43) inductively, we have

$$|\partial_\omega^l \partial_\xi^p \phi_F^t|_{\hat{D}_+ \times \mathcal{O}_+} \leq c \mu \Gamma(r - r_+), \quad 2 \leq |p| + |l| \leq l_0 + 1, \quad |l| \leq l_0, \quad \text{or,} \quad |p| = 0, \quad |l| \geq 1.$$

□

**3.5. Estimate on the new Hamiltonian.** We now estimate the new Hamiltonian

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

where  $N_+$ ,  $P_+$  are as in (3.13), (3.14).

**Lemma 3.4.** *There is a constant  $c_4$  such that for all  $|l| \leq l_0$ ,*

$$\begin{aligned} |\partial_\omega^l \Omega_+ - \partial_\omega^l \Omega|_{\mathcal{O}_+} &\leq c_4 \gamma^a s \mu, \\ |\partial_\omega^l e_+ - \partial_\omega^l e|_{\mathcal{O}_+} &\leq c_4 \gamma^a s^2 \mu, \\ |\partial_\omega^l M^+ - \partial_\omega^l M|_{\mathcal{O}_+} &\leq c_4 \gamma^a \mu, \\ |\partial_\omega^l \partial_{(y,z)}^j (h_+ - h)|_{\mathcal{D}_+ \times \mathcal{O}_+} &\leq c_4 \gamma^a \mu, \quad |j| \leq l_0. \end{aligned}$$

*Proof.* The lemmas follows easily from (3.10)-(3.12) and (3.37). □

Let

$$\begin{aligned} (3.44) \quad \Delta_{|j|} &= (\gamma^a s^{\kappa_j} + s^{2-\text{sgn}(|j|)-\text{sgn}(|j|-2)}) \mu^2 \Gamma^3(r - r_+) \\ &\quad + s_+^{1-\text{sgn}(|j|)} s^{2-\text{sgn}(|j|)\text{sgn}(|j|-1)-\text{sgn}(|j|)\text{sgn}(|j|-1)\text{sgn}(|j|-2)} \mu \Gamma(r - r_+) \\ &\quad + \gamma^{a_j} (s^{\kappa_j+1} + s^{\kappa_j-3} s_+^3) \mu^{d_j}. \end{aligned}$$

**Lemma 3.5.** *Assume H1)-H6). Then there is a constant  $c_5$  such that*

$$|\partial_\omega^l \partial_x^i \partial_{(y,z)}^j P_+|_{\mathcal{D}_+ \times \mathcal{O}_+} \leq c_5 \Delta_{|j|}, \quad |l| + |i| + |j| \leq l_0.$$

*Thus, if*

$$\text{H8)} \quad c_5 \Delta_{|j|} \leq \gamma_+^{a_j} s_+^{\kappa_j} \mu_+^{d_j},$$

*then*

$$|\partial_\omega^l \partial_x^i \partial_{(y,z)}^j P_+|_{\mathcal{D}_+ \times \mathcal{O}_+} \leq \gamma_+^{a_j} s_+^{\kappa_j} \mu_+^{d_j}, \quad |l| + |i| + |j| \leq l_0.$$

*Proof.* Let  $|l| + |i| + |j| \leq l_0$ . Similar to the proof of (3.32), we have, by H4), (3.15), (3.38) and Lemma 3.4 that

$$\begin{aligned} (3.45) \quad |\partial_\omega^l \partial_x^i \partial_{(y,z)}^j Q|_{\mathcal{D}_3 \times \mathcal{O}_+} &\leq c s_+^{1-\text{sgn}(|j|)} s^{2-\text{sgn}(|j|)\text{sgn}(|j|-1)-\text{sgn}(|j|)\text{sgn}(|j|-1)\text{sgn}(|j|-2)} \cdot \mu \Gamma(r - r_+), \end{aligned}$$

$$(3.46) \quad |\partial_\omega^l \partial_x^i \partial_{(y,z)}^j Q|_{\mathcal{D}(s) \times \mathcal{O}_+} \leq c s^{2-\text{sgn}(|j|)-\text{sgn}(|j|-2)} \mu \Gamma(r - r_+).$$

Using the expression of  $R$  in (3.3), we also have

$$(3.47) \quad |\partial_\omega^l \partial_x^i \partial_{(y,z)}^j [R]|_{\mathcal{D}(s) \times \mathcal{O}} + |\partial_\omega^l \partial_x^i \partial_{(y,z)}^j R|_{\mathcal{D}(s) \times \mathcal{O}} \leq c \gamma^a s^{\kappa_j} \mu \Gamma(r - r_+).$$

It follows from (3.32) and (3.47) that

$$(|\partial_\omega^l \partial_x^i \partial_{(y,z)}^j \{[R], F\}| + |\partial_\omega^l \partial_x^i \partial_{(y,z)}^j \{R, F\}|)_{D(s) \times \mathcal{O}_+} \leq c\gamma^a s^{\kappa_j} \mu^2 \Gamma^2(r - r_+),$$

and from (3.32) and (3.46) that

$$|\partial_\omega^l \partial_x^i \partial_{(y,z)}^j \{Q, F\}|_{D(s) \times \mathcal{O}_+} \leq cs^{2-\text{sgn}(|j|)-\text{sgn}(|j|-2)} \mu^2 \Gamma^2(r - r_+).$$

Hence, by (3.39), we have on  $D_2 \times \mathcal{O}_+$  that

$$(3.48) \quad \begin{aligned} |\partial_\omega^l \partial_x^i \partial_{(y,z)}^j \int_0^1 \{R_t, F\} \circ \phi_F^t dt| &= \left| \int_0^1 \{(1-t)(Q + [R] - R) + R, F\} \circ \phi_F^t dt \right| \\ &\leq c(\gamma^a s^{\kappa_j} + s^{2-\text{sgn}(|j|)-\text{sgn}(|j|-2)}) \mu^2 \Gamma^3(r - r_+). \end{aligned}$$

Combining (3.45), (3.48) and Lemma 3.1 with (3.7), we have

$$\begin{aligned} |\partial_\omega^l \partial_x^i \partial_{(y,z)}^j \bar{P}_+|_{D_2 \times \mathcal{O}_+} &\leq c((\gamma^a s^{\kappa_j} + s^{2-\text{sgn}(|j|)-\text{sgn}(|j|-2)}) \mu^2 \Gamma^3(r - r_+) \\ &\quad + s_+^{1-\text{sgn}(|j|)} s^{2-\text{sgn}(|j|)\text{sgn}(|j|-1)-\text{sgn}(|j|)\text{sgn}(|j|-1)\text{sgn}(|j|-2)} \mu \Gamma(r - r_+) \\ &\quad + \gamma^{a_j} (s^{\kappa_j+1} + s^{\kappa_j-3} s_+^3) \mu^{d_j}), \end{aligned}$$

which, together with (3.45) and (3.41), implies that

$$|\partial_\omega^l \partial_x^i \partial_{(y,z)}^j P_+|_{D_+ \times \mathcal{O}_+} \leq c\Delta_{|j|}.$$

□

**3.6. Estimate on the new frequency domain.** Define

$$\begin{aligned} L_{0k}^+ &= \sqrt{-1} \langle k, \Omega_+ \rangle, \\ L_{1k}^+ &= \sqrt{-1} \langle k, \Omega_+ \rangle I_{2m} - M_{22}^+ J, \\ L_{2k}^+ &= \sqrt{-1} \langle k, \Omega_+ \rangle I_{4m^2} - (M_{22}^+ J) \otimes I_{2m} - I_{2m} \otimes (M_{22}^+ J), \end{aligned}$$

$\omega \in \mathcal{O}$ ,  $k \in \mathbb{Z}^n \setminus \{0\}$ .

**Lemma 3.6.** *If*

$$\text{H9)} \quad 3c_4 \mu K_+^{8m^2\tau+8m^2} < \min\left\{ \frac{\gamma_+}{\gamma_0}, \frac{\gamma_+^{2m} - \gamma_+^{2m}}{\gamma_0^{2m}}, \frac{\gamma_+^{4m^2} - \gamma_+^{4m^2}}{\gamma_0^{4m^2}} \right\},$$

then for all  $0 < |k| \leq K_+$ ,  $\omega \in \mathcal{O}_+$ ,

$$|L_{0k}^+| > \frac{\gamma_+}{|k|^\tau}, \quad |\det L_{1k}^+| > \frac{\gamma_+^{2m}}{|k|^{2m\tau}}, \quad |\det L_{2k}^+| > \frac{\gamma_+^{4m^2}}{|k|^{4m^2\tau}}.$$

*Proof.* Let  $0 < |k| \leq K_+$ ,  $\omega \in \mathcal{O}_+$ . We note by (3.31) that

$$|L_{0k}| > \frac{\gamma}{|k|^\tau}, \quad |\det L_{1k}| > \frac{\gamma^{2m}}{|k|^{2m\tau}}, \quad |\det L_{2k}| > \frac{\gamma^{4m^2}}{|k|^{4m^2\tau}}.$$

It follows from Lemma 3.4 and H9) that

$$\begin{aligned} |L_{0k}^+| &> |L_{0k}| - c_4 \mu K_+ > \frac{\gamma_+}{|k|^\tau}, \\ |\det L_{1k}^+| &> |\det L_{1k}| - 2c_4 \nu K_+^{2m} > \frac{\gamma_+^{2m}}{|k|^{2m\tau}}, \\ |\det L_{2k}^+| &> |\det L_{2k}| - 3c_4 \nu K_+^{4m^2} > \frac{\gamma_+^{4m^2}}{|k|^{4m^2\tau}}. \end{aligned}$$

□

## 4. ITERATION LEMMA

In this section, we shall prove an iteration lemma which guarantees the inductive construction of the transformations in all KAM steps.

Let  $r_0, s_0, \mu_0, \gamma_0, \mathcal{O}_0, H_0, N_0, e_0, \Omega_0, M^0, M_{22}^0, h_0, P_0$  be as defined at the beginning of Section 3 and define  $\mathcal{D}_0 = \mathcal{D}(\beta_0)$ ,  $\tilde{D}_0 = D(r_0, \beta_0)$ ,  $D_0 = D(r_0, s_0)$ ,  $K_0 = 0$ ,  $\Phi_0 = id$ . For any  $\nu = 0, 1, \dots$ , we index all index-free quantities in Section 3 by  $\nu$  and index all “+”-indexed quantities in Section 3 by  $\nu+1$ . This yields the following sequences:

$$\begin{aligned} & r_\nu, s_\nu, \mu_\nu, K_\nu, \mathcal{O}_\nu, \mathcal{D}_\nu, D_\nu, \tilde{D}_\nu, H_\nu, N_\nu, \\ & e_\nu, \Omega_\nu, M^\nu, M_{22}^\nu, L_{0k}^\nu, L_{1k}^\nu, L_{2k}^\nu, h_\nu, P_\nu, \Phi_\nu \end{aligned}$$

for  $\nu = 1, 2, \dots$ . In particular,

$$\begin{aligned} H_\nu &= N_\nu + P_\nu, \\ N_\nu &= e_\nu + \langle \Omega_\nu, y \rangle + \frac{1}{2} \left\langle \begin{pmatrix} y \\ z \end{pmatrix}, M^\nu \begin{pmatrix} y \\ z \end{pmatrix} \right\rangle + h_\nu(y, z, \omega), \\ M^\nu &= \begin{pmatrix} M_{11}^\nu & M_{12}^\nu \\ (M_{12}^\nu)^\top & M_{22}^\nu \end{pmatrix}, \\ L_{0k}^\nu &= \sqrt{-1} \langle k, \Omega_\nu \rangle, \\ L_{1k}^\nu &= \sqrt{-1} \langle k, \Omega_\nu \rangle I_{2m} - M_{22}^\nu J, \\ L_{2k}^\nu &= \sqrt{-1} \langle k, \Omega_\nu \rangle I_{4m^2} - (M_{22}^\nu J) \otimes I_{2m} - I_{2m} \otimes (M_{22}^\nu J), \\ r_\nu &= r_0 \left( 1 - \frac{1}{2} (1 - \delta) \sum_{i=1}^{\nu} \delta^{i+1} \right), \\ s_\nu &= s_{\nu-1}^{1+b+\sigma}, \\ \beta_\nu &= \beta_0 \left( 1 - \sum_{i=1}^{\nu} \frac{1}{2^{i+1}} \right), \\ \mu_\nu &= c_0 s_{\nu-1}^\sigma \mu_{\nu-1}, \quad c_0 = \max\{1, c_1, \dots, c_5\}, \\ \gamma_\nu &= \gamma_0 \left( 1 - \sum_{i=1}^{\nu} \frac{1}{2^{i+1}} \right), \\ K_\nu &= (\lceil \log \frac{1}{s_{\nu-1}} \rceil + 1)^3, \\ \mathcal{O}_\nu &= \{ \omega \in \mathcal{O}_{\nu-1} : |L_{0k}^{\nu-1}| > \frac{\gamma_{\nu-1}}{|k|^\tau}, |\det L_{1k}^{\nu-1}| > \frac{\gamma_{\nu-1}^{2m}}{|k|^{2m\tau}}, \\ & \quad |\det L_{2k}^{\nu-1}| > \frac{\gamma_{\nu-1}^{4m^2}}{|k|^{4m^2\tau}}, 0 < |k| \leq K_\nu \}, \\ \mathcal{D}_\nu &= \mathcal{D}(\beta_\nu), \\ D_\nu &= D(r_\nu, s_\nu), \\ \tilde{D}_\nu &= D\left(r_\nu + \frac{5}{8}(r_{\nu-1} - r_\nu), \beta_\nu\right). \end{aligned}$$

We note that  $c_0$  only depends on  $r_0, \beta_0, l_0$ .

**Lemma 4.1** (Iteration Lemma). *If (2.1) holds for sufficiently small  $\mu = \mu(r_0, \beta_0, l_0)$ , then the KAM steps described in Section 3 are valid for all  $\nu = 0, 1, \dots$ , and the following holds for all  $\nu = 1, 2, \dots$ .*

- 1)  $e_\nu = e_\nu(\omega)$ ,  $\Omega_\nu = \Omega_\nu(\omega)$  are smooth on  $\mathcal{O}_\nu$ ,  $M^\nu = M^\nu(\omega)$  is symmetric and smooth on  $\mathcal{O}_\nu$ ,  $h_\nu = h_\nu(y, z, \omega) = O(|(y, z)|^3)$  and  $P_\nu = P_\nu(x, y, z, \omega)$  are real analytic in  $(y, z) \in \mathcal{D}_\nu$  and in  $(x, y, z) \in \tilde{D}_\nu$  respectively and smooth in  $\omega \in \mathcal{O}_\nu$ . Moreover, for all  $|l| \leq l_0$ ,

$$(4.1) \quad |\partial_\omega^l e_\nu - \partial_\omega^l e_{\nu-1}|_{\mathcal{O}_\nu} \leq \frac{\gamma^{2(1+\sigma_0)a+a} \mu^{\frac{1}{4}}}{2^\nu},$$

$$(4.2) \quad |\partial_\omega^l e_\nu - \partial_\omega^l e_0|_{\mathcal{O}_\nu} \leq \gamma^{2(1+\sigma_0)a+a} \mu^{\frac{1}{4}},$$

$$(4.3) \quad |\partial_\omega^l \Omega_\nu - \partial_\omega^l \Omega_{\nu-1}|_{\mathcal{O}_\nu} \leq \frac{\gamma^{(1+\sigma_0)a+a} \mu^{\frac{1}{4}}}{2^\nu},$$

$$(4.4) \quad |\partial_\omega^l (\Omega_\nu - id)|_{\mathcal{O}_\nu} \leq \gamma^{(1+\sigma_0)a+a} \mu^{\frac{1}{4}},$$

$$(4.5) \quad |\partial_\omega^l M^\nu - \partial_\omega^l M^{\nu-1}|_{\mathcal{O}_\nu} \leq \frac{\gamma^a \mu^{\frac{1}{4}}}{2^\nu},$$

$$(4.6) \quad |\partial_\omega^l M^\nu - \partial_\omega^l M^0|_{\mathcal{O}_\nu} \leq \gamma^a \mu^{\frac{1}{4}},$$

$$(4.7) \quad |\partial_\omega^l \partial_{(y,z)}^j (h_\nu - h_{\nu-1})|_{\mathcal{D}_\nu \times \mathcal{O}_\nu} \leq \frac{\gamma^a \mu^{\frac{1}{4}}}{2^\nu}, \quad |j| \leq l_0,$$

$$(4.8) \quad |\partial_\omega^l \partial_{(y,z)}^j (h_\nu - h_0)|_{\mathcal{D}_\nu \times \mathcal{O}_\nu} \leq \gamma^a \mu^{\frac{1}{4}}, \quad |j| \leq l_0,$$

$$(4.9) \quad |\partial_\omega^l \partial_x^i \partial_{(y,z)}^j P_\nu|_{D_\nu \times \mathcal{O}_\nu} \leq \gamma_\nu^{\alpha_j} s_\nu^{\kappa_j} \mu_\nu^{d_j}, \quad |l| + |i| \leq l_0, |j| \leq l_0.$$

- 2) If  $M^0$  is non-singular on  $\mathcal{O}_0$ , then  $\Omega_\nu(\omega) \equiv \omega$ .  
 3)  $\Phi_\nu : D_\nu \times \mathcal{O}_\nu \rightarrow D_{\nu-1}, \tilde{D}_\nu \times \mathcal{O}_\nu \rightarrow \tilde{D}_{\nu-1}$  is symplectic for each  $\omega \in \mathcal{O}_\nu$ , real analytic in  $(x, y, z) \in \tilde{D}_\nu$  and smooth in  $\omega \in \mathcal{O}_\nu$ , and

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

Moreover,

$$(4.10) \quad |\partial_\omega^l \partial_\xi^p (\Phi_\nu - id)|_{\tilde{D}_\nu \times \mathcal{O}_\nu} \leq \frac{\mu^{\frac{1}{4}}}{2^\nu}, \quad |p| \leq l_0 + 1, \quad |l| \leq l_0,$$

where  $\xi = (x, y, z)$ .

4)

$$\begin{aligned} \mathcal{O}_\nu &= \left\{ \omega \in \mathcal{O}_{\nu-1} : |L_{0k}^{\nu-1}| > \frac{\gamma_{\nu-1}}{|k|^\tau}, \quad |\det L_{1k}^{\nu-1}| > \frac{\gamma_{\nu-1}^{2m}}{|k|^{2m\tau}}, \right. \\ &\quad \left. |\det L_{2k}^{\nu-1}| > \frac{\gamma_{\nu-1}^{4m^2}}{|k|^{4m^2\tau}}, \quad \text{for all } K_{\nu-1} < |k| \leq K_\nu \right\}. \end{aligned}$$

*Proof.* We need to verify the conditions H1)-H9) in Section 3 for all  $\nu = 0, 1, \dots$ .

Note that

$$(4.11) \quad \mu_\nu = (c_0)^\nu \mu_0 s_0^{\frac{\sigma}{b+\sigma}((1+b+\sigma)^\nu - 1)},$$

$$(4.12) \quad s_\nu = s_0^{(1+b+\sigma)^\nu},$$

$$(4.13) \quad s_0 = \gamma^{(1+\sigma_0)a} \mu^{\frac{1}{4}}.$$

By (4.12), we see that if  $\mu < (\frac{1}{16})^{\frac{4}{b+\sigma}}$ , then

$$s_{\nu+1} \leq s_0^{b+\sigma} s_\nu \leq \frac{s_\nu}{16},$$

i.e., H2) holds.

To verify H3), we denote

$$E_\nu = \frac{r_\nu - r_{\nu+1}}{16} = \frac{1}{32} r_0 (1 - \delta) \delta^{\nu+2}.$$

Since  $\delta(1 + b + \sigma) > 1$ , we see that if  $\mu$  is small, then

$$\begin{aligned} \frac{E_\nu}{2} \log \frac{1}{s_\nu} &= -\frac{1}{32} r_0 (1 - \delta) \delta^2 (\delta(1 + b + \sigma))^\nu \log s_0 \\ (4.14) \quad &> -\frac{1}{32} r_0 (1 - \delta) \delta^2 \log s_0 > 1. \end{aligned}$$

It follows from (4.12), (4.13), (4.14) that

$$\begin{aligned} &\log(n + l_0 + 1)! + 3(n + l_0) \log([\log \frac{1}{s_\nu}] + 1) - E_\nu([\log \frac{1}{s_\nu}] + 1)^3 - (n + l_0) \log E_\nu \\ &\leq \log(n + l_0 + 1)! + 3(n + l_0) \log(\log \frac{1}{s_\nu} + 2) - (\log \frac{1}{s_\nu})^2 + (n + l_0) \log \frac{1}{s_\nu} \\ &\leq -\log \frac{1}{s_\nu}, \end{aligned}$$

provided that  $\mu$  (hence  $s_0$ ) is sufficiently small. Thus,

$$\int_{K_{\nu+1}}^\infty \lambda^{n+l_0} e^{-\lambda E_\nu} d\lambda \leq (n + l_0 + 1)! \frac{K_{\nu+1}^{n+l_0}}{E_\nu^{n+l_0}} e^{-K_{\nu+1} E_\nu} \leq s_\nu,$$

i.e., H3) holds.

Denote

$$\eta = n + 2(l_0 + 1)(2m^2[\tau] + 2m^2 + 1)$$

and fix an  $0 < \varepsilon \ll 1$  such that

$$(4.15) \quad \begin{aligned} (1 + \sigma_0)(1 + b + \sigma)(1 - b - 2\sigma - 2\varepsilon) &> 1, \\ \frac{\sigma}{b + \sigma} &> (4b + 5\sigma + 3\varepsilon) + 3\varepsilon, \quad \lambda_0 - (1 - \lambda_0)\sigma - 2\varepsilon > 0. \end{aligned}$$

We let  $\mu$  (hence  $s_0$ ) be sufficiently small so that

$$\begin{aligned} \frac{s_\nu^\varepsilon}{E_\nu^{4\eta}} &= \left( \frac{32}{r_0(1 - \delta)\delta^2} \right)^{4\eta} \frac{s_0^{\varepsilon(1+b+\sigma)^\nu}}{\delta^{4\eta\nu}} \leq \left( \frac{32}{r_0(1 - \delta)\delta^2} \right)^{4\eta} s_0^\varepsilon \left( \frac{s_0^{b\varepsilon}}{\delta^{4\eta}} \right)^\nu \\ (4.16) \quad &\leq \left( \frac{32}{r_0(1 - \delta)\delta^2} \right)^{4\eta} s_0^\varepsilon \leq \frac{1}{(\eta!)^4 e^{8E_0}}. \end{aligned}$$

Using (4.16) and the fact that

$$1 \leq \Gamma_\nu = \Gamma(r_\nu - r_{\nu+1}) = \Gamma\left(\frac{1}{2} r_0 \delta^{\nu+2} (1 - \delta)\right) \leq e^{2E_0} \int_1^\infty \lambda^\eta e^{-\lambda E_\nu} d\lambda \leq \frac{\eta! e^{2E_0}}{E_\nu^\eta},$$

we have

$$(4.17) \quad s_\nu^\varepsilon \Gamma_\nu^4 \leq \frac{(\eta!)^4 e^{8E_0} s_\nu^\varepsilon}{E_\nu^{4\eta}} \leq 1.$$

Since

$$\sigma_0 - \frac{\sigma}{b + \sigma} (1 + \sigma_0) a > 0,$$

it follows that if  $\mu$  is small, then

$$\begin{aligned}
(4.18) \quad \frac{c_0 \mu_\nu \Gamma_\nu}{E_\nu} &\leq c_0 \mu_\nu s_\nu^{-2\varepsilon} \left( \frac{s_\nu^\varepsilon}{E_\nu^{3\eta}} \right) (s_\nu^\varepsilon \Gamma_\nu^3) \\
&\leq c_0 \mu_\nu s_\nu^{-2\varepsilon} = c_0 \mu_0 s_0^{-\frac{\sigma}{b+\sigma}} c_0^\nu s_0^{(1+b+\sigma)^\nu (\frac{\sigma}{b+\sigma} - 2\varepsilon)} \\
&\leq c_0 \mu_0 s_0^{-\frac{\sigma}{b+\sigma}} (c_0 s_0^{\sigma - 2\varepsilon(b+\sigma)})^\nu \leq c_0 \mu_0 s_0^{-\frac{\sigma}{b+\sigma}} \leq c_0 \mu^{\frac{1}{4}} < 1,
\end{aligned}$$

i.e., H5) holds.

(4.11)-(4.13), (4.15) and  $\nu \geq 1$  yield that

$$\begin{aligned}
(4.19) \quad \frac{\mu_\nu}{s_\nu^{4b+5\sigma+2\varepsilon}} &= \frac{(c_0)^\nu \mu_0 s_0^{\frac{\sigma}{b+\sigma}((1+b+\sigma)^\nu - 1)}}{s_0^{(1+b+\sigma)^\nu (4b+5\sigma+2\varepsilon)}} \\
&= s_0^{-\frac{\sigma}{b+\sigma}} \mu_0 (c_0)^\nu s_0^{(\frac{\sigma}{b+\sigma} - (4b+5\sigma+2\varepsilon))(1+b+\sigma)^\nu} \\
&\leq s_0^{-\frac{\sigma}{b+\sigma}} \mu_0 (c_0 s_0^\varepsilon)^\nu \leq \mu^{\frac{1}{4}},
\end{aligned}$$

provided that  $\mu$  is small so that  $c_0 s_0^\varepsilon \leq 1$ . This together with (4.17) implies that

$$\frac{c_0 s_\nu \mu_\nu \Gamma_\nu}{s_{\nu+1}} \leq \frac{c_0 \mu_\nu}{s_\nu^{b+\sigma+\varepsilon}} \leq c_0 \mu^{\frac{1}{4}} < 1,$$

i.e., H6) holds. H7) is obvious as  $\mu$  small.

To verify H8), we let  $\Delta_{|j|} = \Delta_{|j|}^\nu$  be the sequence defined in (3.44) for the  $\nu$ th KAM step. Then,

$$\begin{aligned}
c_0 \Delta_1 &= c_0 ((\gamma_\nu^a s_\nu^2 + s_\nu^3) \mu_\nu^2 \Gamma_\nu^3 + s_{\nu+1} s_\nu^2 \mu_\nu \Gamma_\nu + \gamma_\nu^a (s_\nu^3 + s_\nu^{-1} s_{\nu+1}^3) \mu_\nu), \\
c_0 \Delta_2 &= c_0 ((\gamma_\nu^a s_\nu + s_\nu^2) \mu_\nu^2 \Gamma_\nu^3 + s_\nu^2 \mu_\nu \Gamma_\nu + \gamma_\nu^a (s_\nu^2 + s_\nu^{-2} s_{\nu+1}^3) \mu_\nu), \\
c_0 \Delta_3 &= c_0 ((\gamma_\nu^a + s_\nu) \mu_\nu^2 \Gamma_\nu^3 + s_\nu \mu_\nu \Gamma_\nu + \gamma_\nu^a (s_\nu + s_\nu^{-3} s_{\nu+1}^3) \mu_\nu), \\
c_0 \Delta_{|j|} &= c_0 ((\gamma_\nu^a + 1) \mu_\nu^2 \Gamma_\nu^3 + \mu_\nu \Gamma_\nu + \gamma_\nu^a (s_\nu + s_\nu^{-3} s_{\nu+1}^3) \mu_\nu^{1-\lambda_0}), \quad |j| \geq 3.
\end{aligned}$$

Using (4.17), (4.19), we have

$$\begin{aligned}
\frac{c_0 \Delta_1 \Gamma_\nu}{\gamma_{\nu+1}^a s_{\nu+1}^2 \mu_{\nu+1}} &\leq (2s_\nu^{2(b+\sigma)} + \gamma_0^{-a} s_\nu^{1-b-2\sigma-2\varepsilon} + s_\nu^{1-2b-3\sigma} + s_\nu^b) \Gamma_\nu, \\
\frac{c_0 \Delta_2 \Gamma_\nu}{\gamma_{\nu+1}^a s_{\nu+1} \mu_{\nu+1}} &\leq (2s_\nu^{3(b+\sigma)} + \gamma_0^{-a} s_\nu^{1-b-2\sigma-2\varepsilon} + s_\nu^{1-b-2\sigma} + s_\nu^{2b+\sigma}) \Gamma_\nu, \\
\frac{c_0 \Delta_3 \Gamma_\nu}{\gamma_{\nu+1}^a \mu_{\nu+1}} &\leq (2s_\nu^{4(b+\sigma)} + \gamma_0^{-a} s_\nu^{1-b-\sigma-2\varepsilon} + s_\nu^{1-\sigma} + s_\nu^{3b+2\sigma}) \Gamma_\nu, \\
\frac{c_0 \Delta_{|j|}}{\mu_{\nu+1}} &\leq 2s_\nu^{4(b+\sigma)} + s_\nu^{\lambda_0 - (1-\lambda_0)\sigma - 2\varepsilon} + s_\nu^{1-\sigma} + s_\nu^{3b+2\sigma}, \quad |j| \geq 3.
\end{aligned}$$

Since, by (4.15) and  $\nu \geq 1$ ,

$$\begin{aligned}
s_\nu^{1-b-2\sigma-2\varepsilon} &\leq s_0^{(1+b+\sigma)(1-b-2\sigma-2\varepsilon)} \\
Y &= \gamma_0^{(1+\sigma_0)a(1+b+\sigma)(1-b-2\sigma-2\varepsilon)} \mu^{\frac{1-b-2\sigma-2\varepsilon}{4}} \leq \gamma_0^a \mu^{\frac{1-b-2\sigma-2\varepsilon}{4}},
\end{aligned}$$

it is clear that one can make  $\mu$  small such that

$$\gamma_0^{-a} s_\nu^{1-b-\sigma-2\varepsilon} \leq \gamma_0^{-a} s_\nu^{1-b-2\sigma-2\varepsilon} < \frac{1}{4}.$$

By (4.12), (4.13), all other terms in the above can also be made smaller than  $\frac{1}{4}$  by assuming  $\mu$  small. This verifies H8). As  $\mu$  sufficiently small, we have by (4.19) that

$$\begin{aligned} 3c_0\mu_\nu K_{\nu+1}^{8m^2\tau+8m^2} &< \frac{3c_0}{2^{(\nu+2)(4m^2)}} s_\nu^{4b+5\sigma+2\varepsilon} \left(\left[\log \frac{1}{\mu_\nu}\right] + 1\right)^{3(4m^2\tau+4m^2)} \\ &< \frac{1}{2^{(\nu+2)(4m^2)}} < \frac{1}{2^{(\nu+2)(4m)}} < \frac{1}{2^{\nu+2}}. \end{aligned}$$

This verifies H9).

Since by (4.17), (4.19),

$$\mu_\nu \Gamma_\nu \leq \mu^{\frac{1}{4}} (c_0 s_0^\varepsilon)^\nu (s_\nu^\varepsilon \Gamma_\nu) s_\nu^{4b+5\sigma} \leq \mu^{\frac{1}{4}} (c_0 s_0^\varepsilon)^\nu,$$

we can make  $\mu$  small so that

$$(4.20) \quad c_0 \mu_\nu \Gamma_\nu \leq \frac{\mu^{\frac{1}{4}}}{2^{\nu+1}},$$

$$(4.21) \quad c_0 \gamma^a \mu_\nu \Gamma_\nu \leq \frac{\gamma^a \mu^{\frac{1}{4}}}{2^{\nu+1}}$$

for all  $\nu$ .

We verify H1), H4) by induction. As H1), H4) trivially hold for  $\nu = 0$ , the KAM step described in Section 3 is valid for  $\nu = 0$ . We now assume that for some positive integer  $\nu_*$  H1), H4) hold for all  $\nu = 0, 1, \dots, \nu_*$ , i.e., the KAM steps are valid for all  $\nu = 1, 2, \dots, \nu_*$ . Applying Lemma 3.4 and (4.21) for all  $\nu = 0, 1, \dots, \nu_*$ , we have

$$\begin{aligned} (4.22) \quad |\partial_\omega^l M_{22}^{\nu_*+1} - \partial_\omega^l M_{22}^0|_{\mathcal{O}_{\nu_*+1}} &\leq |\partial_\omega^l M^{\nu_*+1} - \partial_\omega^l M^0|_{\mathcal{O}_{\nu_*+1}} \\ &\leq \sum_{i=0}^{\nu_*} |\partial_\omega^l (M^{i+1} - M^i)|_{\mathcal{O}_{i+1}} \leq \sum_{i=0}^{\nu_*} \frac{\gamma^a \mu^{\frac{1}{4}}}{2^{i+1}} \leq \gamma^a \mu^{\frac{1}{4}}, \\ |\partial_\omega^l (\Omega_{\nu_*+1} - id)|_{\mathcal{O}_{\nu_*+1}} &\leq \sum_{i=0}^{\nu_*} |\partial_\omega^l \Omega_{i+1} - \partial_\omega^l \Omega_i|_{\mathcal{O}_{i+1}} \leq \sum_{i=0}^{\nu_*} \frac{\gamma^{(1+\sigma_0)a+a} \mu^{\frac{1}{4}}}{2^{i+1}} \\ &\leq \gamma^{(1+\sigma_0)a+a} \mu^{\frac{1}{4}}, \end{aligned}$$

i.e., H4) holds for  $\nu = \nu_* + 1$ . Denote  $\hat{M}_{\nu_*+1} = M^{\nu_*+1}, M_{22}^{\nu_*+1}, \hat{M}_0 = M^0, M_{22}^0$ , respectively. If  $\mu$  is small so that  $s_0 < \frac{1}{2^{|\hat{M}_0|^{-1}|_{\mathcal{O}_0}}}$ , then by (4.22) and the invertibility of  $\hat{M}_0$  on  $\mathcal{O}_0$ , we see that  $\hat{M}_{\nu_*+1}$  is non-singular on  $\mathcal{O}_{\nu_*+1}$ , and

$$|(\hat{M}_{\nu_*+1})^{-1}|_{\mathcal{O}_{\nu_*+1}} \leq \frac{|(\hat{M}_0)^{-1}|_{\mathcal{O}_0}}{1 - |\hat{M}_{\nu_*+1} - \hat{M}_0|_{\mathcal{O}_{\nu_*+1}} |(\hat{M}_0)^{-1}|_{\mathcal{O}_0}} \leq 2|(\hat{M}_0)^{-1}|_{\mathcal{O}_0}.$$

Hence H1) holds for  $\nu = \nu_* + 1$ . Therefore, all H1)-H9) hold and the KAM steps described in Section 3 are valid for all  $\nu$ .

By performing the KAM steps described in Section 3 inductively, we then obtain the desired sequences stated in the lemma. Now, 2) clearly follows from (3.11), and (4.1), (4.3), (4.5), (4.9), (4.10) follow from Lemmas 3.3-3.5 and (4.20), (4.21) accordingly. By the same argument as in (4.22), we also obtain (4.2), (4.4), (4.6) and (4.8) from (4.1), (4.3), (4.5) and (4.7) respectively.

Note that 4) automatically holds for  $\nu = 1$ . We now let  $\nu > 1$ . By Lemma 3.6, it is clear that

$$\mathcal{O}_{\nu-1} = \left\{ \omega \in \mathcal{O}_{\nu-1} : |L_{0k}^{\nu-1}| > \frac{\gamma_{\nu-1}}{|k|^\tau}, |\det L_{1k}^{\nu-1}| > \frac{\gamma_{\nu-1}^{2m}}{|k|^{2m\tau}} \right\},$$

$$|\det L_{2k}^{\nu-1}| > \frac{\gamma_{\nu-1}^{4m^2}}{|k|^{4m^2\tau}}, \quad \text{for all } 0 < |k| \leq K_{\nu-1}.$$

Denote

$$\begin{aligned} \hat{\mathcal{O}}_\nu &= \{\omega \in \mathcal{O}_{\nu-1} : |\langle k, \omega \rangle| > \frac{\gamma_{\nu-1}}{|k|^\tau}, |\det L_{1k}^{\nu-1}| > \frac{\gamma_{\nu-1}^{2m}}{|k|^{2m\tau}}, \\ &\quad |\det L_{2k}^{\nu-1}| > \frac{\gamma_{\nu-1}^{4m^2}}{|k|^{4m^2\tau}}, \quad \text{for all } K_{\nu-1} < |k| \leq K_\nu\}. \end{aligned}$$

Then

$$\begin{aligned} \mathcal{O}_\nu &= \{\omega \in \mathcal{O}_{\nu-1} : |\langle k, \omega \rangle| > \frac{\gamma_{\nu-1}}{|k|^\tau}, |\det L_{1k}^{\nu-1}| > \frac{\gamma_{\nu-1}^{2m}}{|k|^{2m\tau}}, \\ &\quad |\det L_{2k}^{\nu-1}| > \frac{\gamma_{\nu-1}^{4m^2}}{|k|^{4m^2\tau}}, \quad \text{for all } 0 < |k| \leq K_\nu\} \\ (4.23) \quad &= \mathcal{O}_{\nu-1} \cap \hat{\mathcal{O}}_\nu = \hat{\mathcal{O}}_\nu. \end{aligned}$$

The lemma is now complete.  $\square$

## 5. PROOF OF THE THEOREM

**5.1. Convergence.** By assuming  $\mu = \mu(r, s, l_0)$  small, one can apply Lemma 4.1 inductively to obtain the following sequences:

$$\begin{aligned} \Psi_\nu &= \Phi_0 \circ \Phi_1 \circ \cdots \circ \Phi_\nu : \tilde{D}_\nu \times \mathcal{O}_\nu \rightarrow \tilde{D}_0, \\ H \circ \Psi_\nu &= H_\nu = N_\nu + P_\nu, \\ N_\nu &= e_\nu + \langle \Omega_\nu, y \rangle + \frac{1}{2} \left\langle \begin{pmatrix} y \\ z \end{pmatrix}, M^\nu(\omega) \begin{pmatrix} y \\ z \end{pmatrix} \right\rangle + h_\nu(y, z, \omega), \end{aligned}$$

for  $\nu = 0, 1, \dots$ .

Let

$$\mathcal{O}_* = \bigcap_{\nu=0}^{\infty} \mathcal{O}_\nu.$$

First, we show the uniform convergence of  $\Psi_\nu$  on  $D(\frac{r_0}{2}, \frac{\beta_0}{2}) \times \mathcal{O}_*$ . Note that

$$(5.1) \quad \Psi_\nu = \Psi_0 + \sum_{i=1}^{\nu} (\Psi_i - \Psi_{i-1}),$$

where, for each  $i = 1, 2, \dots$ ,

$$(5.2) \quad \begin{aligned} \Psi_i - \Psi_{i-1} &= \Phi_0 \circ \cdots \circ \Phi_i - \Phi_0 \circ \cdots \circ \Phi_{i-1} \\ &= \int_0^1 D(\Phi_0 \circ \cdots \circ \Phi_{i-1})(id + \theta(\Phi_i - id)) d\theta(\Phi_i - id). \end{aligned}$$

Since, by (4.10), on  $D(\frac{r_0}{2}, \frac{\beta_0}{2}) \times \mathcal{O}_*$ ,

$$(5.3) \quad \begin{aligned} &|D(\Phi_0 \circ \cdots \circ \Phi_{i-1})(id + \theta(\Phi_i - id))| \\ &\leq |D\Phi_0(\Phi_1 \circ \cdots \circ \Phi_{i-1})(id + \theta(\Phi_i - id))| \cdots |D\Phi_{i-1}(id + \theta(\Phi_i - id))| \\ &\leq (1 + \mu^{\frac{1}{4}})(1 + \frac{\mu^{\frac{1}{4}}}{2}) \cdots (1 + \frac{\mu^{\frac{1}{4}}}{2^{i-1}}) \leq e^{1 + \frac{1}{2} + \cdots + \frac{1}{2^{i-1}}} \leq e^2, \end{aligned}$$



we have

$$|\Psi_i - \Psi_{i-1}|_{D(\frac{r_0}{2}, \frac{\beta_0}{2}) \times \mathcal{O}_*} \leq e^2 |\Phi_i - id|_{D(\frac{r_0}{2}, \frac{\beta_0}{2}) \times \mathcal{O}_*} \leq e^2 \frac{\mu^{\frac{1}{4}}}{2^i},$$

for all  $i = 1, 2, \dots$ . Thus,  $\Psi_\nu$  converges uniformly on  $D(\frac{r_0}{2}, \frac{\beta_0}{2}) \times \mathcal{O}_*$ . We denote its limit by  $\Psi_\infty$ . Then

$$(5.4) \quad \Psi_\infty = \Psi_0 + \sum_{i=0}^{\infty} (\Psi_i - \Psi_{i-1}) = id + \sum_{i=0}^{\infty} (\Psi_i - \Psi_{i-1}).$$

It follows that  $\Psi_\infty$  is real analytic in  $\xi = (x, y, z)$  and uniformly close to the identity. In fact, using a similar argument,  $\Psi_\infty$  can be shown to be  $C^{l_0}$  uniformly close to the identity. To show the Whitney smoothness of  $\Psi_\infty$  in  $\omega$ , one can apply the standard Whitney extension theorem (see [31, 36]) to uniformly extend all  $\Phi_\nu$ , with respect to  $\omega$ , to functions on  $\tilde{D}_\nu \times \mathcal{O}_0$  of class  $C^{l_0 - \sigma_0}$  in  $\omega \in \mathcal{O}_0$  for some fixed  $0 < \sigma_0 < 1$ , whose Hölder norms satisfy the same estimate (4.10), up to multiplication of a constant. This results in a sequence of extended transformations  $\Psi_\nu$  defined on  $\tilde{D}_\nu \times \mathcal{O}_0$ . Similar to (5.2), (5.3), one can use (4.10) to obtain the estimates

$$|\partial_\omega^l \Psi_\nu - \partial_\omega^l \Psi_{\nu-1}|_{D(\frac{r_0}{2}, \frac{\beta_0}{2}) \times \mathcal{O}_0} \leq c \frac{\mu^{\frac{1}{4}}}{2^\nu}$$

for all  $\nu = 1, 2, \dots$ , which shows the uniform convergence of  $\partial_\omega^l \Psi_\nu$ , on  $D(\frac{r_0}{2}, \frac{\beta_0}{2}) \times \mathcal{O}_0$ . Let  $\Psi_\infty$  be defined as in (5.4) in terms of the extended transformations  $\Psi_\nu$ . Then a similar argument shows that  $\partial_\omega^l \Psi_\nu$  are equally (Hölder) continuous, and  $\partial_\omega^l \Psi_\nu \rightarrow \partial_\omega^l \Psi_\infty$ ,  $1 \leq |l| \leq l_0 - 1$ , uniformly on  $D(\frac{r_0}{2}, \frac{\beta_0}{2}) \times \mathcal{O}_0$ . Thus, the original limit  $\Psi_\infty$  is  $C^{l_0-1}$  Whitney smooth in  $\omega \in \mathcal{O}_*$ .

Next, we show the convergence of the Hamiltonians  $H_\nu$ . By Lemma 4.1 1),  $e_\nu, \Omega_\nu, M^\nu$  converge uniformly on  $\mathcal{O}_*$  and  $h_\nu$  converges uniformly on  $\mathcal{D}(\frac{\beta_0}{2}) \times \mathcal{O}_*$ , as  $\nu \rightarrow \infty$ . We denote their limits by  $e_\infty, \Omega_\infty, M^\infty, h_\infty$  respectively. Clearly,  $h_\infty = O(|(y, z)|^3)$ . By a similar application of the Whitney extension theorem, one can uniformly extend all  $P_\nu$ , with respect to  $\omega$ , to functions on  $D_\nu \times \mathcal{O}_0$  of class  $C^{l_0 - \sigma_0}$  in  $\omega \in \mathcal{O}_0$ , whose Hölder norms satisfy the same estimate (4.9), up to multiplication of a constant. For such extended functions  $P_\nu$ , we then use the same formula (3.12) to define smooth extensions of  $M^\nu$  on  $\mathcal{O}_0$ . Then the estimates in (4.5), (4.6), up to multiplication of a constant, are still valid for all extended matrices  $M^\nu$  on  $\mathcal{O}_0$ , which implies the uniform convergence of  $\partial_\omega^l M^\nu$  on  $\mathcal{O}_0$  for all  $1 \leq |l| \leq l_0 - 1$ . Using the Hölder norms of the extended functions  $P_\nu$  one further shows the Hölder continuities of  $\partial_\omega^l M^\nu$  which implies that the limit of  $M^\nu$  on  $\mathcal{O}_0$  is of class  $C^{l_0-1}$ . Thus,  $M^\infty$  is  $C^{l_0-1}$  Whitney smooth on  $\mathcal{O}_*$ , and

$$|\partial_\omega^l M^\infty - \partial_\omega^l M^0|_{\mathcal{O}_*} = O(\gamma^a \mu^{\frac{1}{4}}), \quad |l| \leq l_0 - 1,$$

in the sense of Whitney (see [31]). Similarly,  $e_\infty(\omega), \Omega_\infty(\omega)$  are  $C^{l_0-1}$  Whitney smooth on  $\mathcal{O}_*$ ,  $h_\infty(y, z, \omega)$  is real analytic in  $(y, z) \in \mathcal{D}(\frac{\beta_0}{2})$  and  $C^{l_0-1}$  Whitney smooth in  $\omega \in \mathcal{O}_*$ , and

$$\begin{aligned} |\partial_\omega^l e_\infty - \partial_\omega^l e_0|_{\mathcal{O}_*} &= O(\gamma^{2(1+\sigma_0)a+a} \mu^{\frac{1}{4}}), \\ |\partial_\omega^l (\Omega_\infty - id)|_{\mathcal{O}_*} &= O(\gamma^{(1+\sigma_0)a+a} \mu^{\frac{1}{4}}), \\ |\partial_\omega^l \partial_{(y,z)}^j h_\infty - \partial_\omega^l \partial_{(y,z)}^j h_0|_{\mathcal{D}(\frac{\beta_0}{2}) \times \mathcal{O}_*} &= O(\gamma^a \mu^{\frac{1}{4}}), \end{aligned}$$

for all  $|j| \leq l_0$ ,  $|l| \leq l_0 - 1$ . It follows that, on  $D(\frac{r_0}{2}, \frac{\beta_0}{2}) \times \mathcal{O}_*$ ,  $N_\nu$  and all their possible derivatives converge uniformly to

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

and its corresponding derivatives, in the sense of Whitney. By the definition of  $\mathcal{O}_*$ , it is easy to see that  $\Omega_\infty(\mathcal{O}_*) \subset \hat{\mathcal{O}}_\gamma$ .

Let

$$P_\infty = H \circ \Psi_\infty - N_\infty.$$

Using the uniform convergence of  $N_\nu$ ,  $\Psi_\nu$  to  $N_\infty$ ,  $\Psi_\infty$ , respectively, we see that, for all  $|l| \leq l_0 - 1$ ,  $\partial_\omega^l P_\nu \rightarrow \partial_\omega^l P_\infty$  uniformly on  $D(\frac{r_0}{2}, \frac{\beta_0}{2}) \times \mathcal{O}_*$  in the sense of Whitney, and  $P_\infty$  is real analytic in  $(x, y, z) \in D(\frac{r_0}{2}, \frac{s_0}{2})$  and  $C^{l_0-1}$  Whitney smooth in  $\omega \in \mathcal{O}_*$ .

For any  $\nu \in Z_+$ ,  $\omega \in \mathcal{O}_*$ ,  $j \in Z_+^n$ ,  $k \in Z_+^{2m}$  with  $|j| + |k| \leq 2$ , by applying the inequality

$$|P_\nu|_{D_\nu} \leq \gamma^a s_\nu^2 \mu_\nu$$

and performing Cauchy's estimate of  $P_\nu$  on  $D(r_\nu, \frac{1}{2}s_\nu)$ , we see that

$$|\partial_y^j \partial_z^k P_\nu| \leq 2^{j+2} \gamma^a \mu_\nu,$$

for all  $|j| + |k| \leq 2$  and  $\nu = 1, 2, \dots$ . By (4.11), it is easy to see that the right hand side of the above converges to 0 as  $\nu \rightarrow \infty$ . Hence, on  $D(\frac{r_0}{2}, 0) \times \mathcal{O}_*$ ,

$$\partial_y^j \partial_z^k P_\infty|_{(y,z)=0} = 0$$

for all  $x \in T^n$ ,  $\omega \in \mathcal{O}_*$ ,  $j \in Z_+^n$ ,  $k \in Z_+^{2m}$  with  $|j| + |k| \leq 2$ . It follows that for each  $\omega \in \mathcal{O}_*$ ,  $T_\omega = T^n \times \{0\} \times \{0\}$  is an analytic, quasi-periodic, invariant  $n$ -torus associated to the Hamiltonian  $H_\infty = H \circ \Psi_\infty$  with the Diophantine toral frequency  $\Omega_\infty(\omega)$ , which corresponds to a perturbed invariant  $n$ -torus of (1.1) with the same properties. Moreover, these perturbed tori form a  $C^{l_0-1}$  Whitney smooth family. In the case that  $M^0$  is non-singular on  $\mathcal{O}_0$ , it follows from Lemma 4.1 2) that  $\Omega_\infty(\omega) \equiv \omega$ , i.e., toral frequencies are also preserved in this case.

**5.2. Measure estimates.** We wish to show that

$$(5.5) \quad |\mathcal{O}_0 \setminus \mathcal{O}_*| \rightarrow 0, \text{ as } \gamma \rightarrow 0.$$

The proof of (5.5) will be based on the following two lemmas. The first lemma shows the connection between the non-resonance condition NR) with the small divisor condition involved in (3.31) for  $\nu = 0$ . The second lemma deals with measure estimates.

**Lemma 5.1.** *Let  $\lambda_j(\omega)$ ,  $j = 1, 2, \dots, 2m$ , be eigenvalues of  $JM_{22}^0(\omega)$ . Then the following holds.*

1) For all  $k \in Z^n$ ,

$$\begin{aligned} \det L_{1k}^0 &= \prod_{i=1}^{2m} (\sqrt{-1} \langle k, \omega \rangle - \lambda_i), \\ \det L_{2k}^0 &= \prod_{i,j=1}^{2m} (\sqrt{-1} \langle k, \Omega \rangle - \lambda_i - \lambda_j). \end{aligned}$$

2) *The set*

$$\{\omega \in \mathcal{O}_0 : \langle k, \omega \rangle \neq 0, \det L_{1k}^0 \neq 0, \det L_{2k}^0 \neq 0, \text{ for all } k \in \mathbb{Z}^n \setminus \{0\}\}$$

*admits full Lebesgue measure relative to  $\mathcal{O}_0$ .*

*Proof.* 1) Let  $E$  be the Jordan canonical form of  $M_{22}^0 J$  and let  $T$  be the non-singular matrix such that  $T^{-1}(M_{22}^0 J)T = E$ . Then

$$\begin{aligned} \det L_{1k}^0 &= \det(\sqrt{-1}\langle k, \omega \rangle I_{2m} - M_{22}^0 J) = \det(\sqrt{-1}\langle k, \omega \rangle I_{2m} - E) \\ &= \prod_{i=1}^{2m} (\sqrt{-1}\langle k, \omega \rangle - \lambda_i). \end{aligned}$$

Since, for any square matrices  $A, B, C, D$  of the same dimension,

$$(A \otimes B)(C \otimes D) = (AC \otimes BD),$$

we have

$$(T^{-1} \otimes T^{-1})((M_{22}^0 J) \otimes I_{2m} + I_{2m} \otimes (M_{22}^0 J))(T \otimes T) = E \otimes I_{2m} + I_{2m} \otimes E.$$

It follows that

$$\begin{aligned} \det L_{2k}^0 &= \det(\sqrt{-1}\langle k, \Omega \rangle I_{4m^2} - M_{22}^0 J \otimes I_{2m} - I_{2m} \otimes (M_{22}^0 J)) \\ &= \det(\sqrt{-1}\langle k, \Omega \rangle I_{4m^2} - E \otimes I_{2m} - I_{2m} \otimes E) \\ &= \prod_{i,j=1}^{2m} (\sqrt{-1}\langle k, \Omega \rangle - \lambda_i - \lambda_j). \end{aligned}$$

The proof of 1) is now complete since  $JM_{22}^0, M_{22}^0 J$  have the same eigenvalues.

2) Denote

$$O_1 = \{\omega \in \mathcal{O}_0 : \sqrt{-1}\langle k, \omega \rangle - \lambda_i \neq 0 \text{ for all } k \in \mathbb{Z}^n \setminus \{0\}, 1 \leq i \leq 2m\},$$

$$O_2 = \{\omega \in \mathcal{O}_0 : \sqrt{-1}\langle k, \omega \rangle - \lambda_i - \lambda_j \neq 0 \text{ for all } k \in \mathbb{Z}^n \setminus \{0\}, 1 \leq i, j \leq 2m\}.$$

For any  $\omega \in O_2, k \in \mathbb{Z}^n \setminus \{0\}, 1 \leq i \leq 2m$ , we have that  $\sqrt{-1}\langle 2k, \omega \rangle - \lambda_i - \lambda_i \neq 0$ , i.e.,  $\omega \in O_1$ . Thus,  $O_2 \subset O_1$ , and 2) easily follows from NR) and 1).  $\square$

**Lemma 5.2.** *Suppose that  $g \in C^p(\bar{I})$ ,  $p \geq 2$ , where  $I \subset \mathbb{R}^1$  is a finite interval. Let  $I_h = \{x \in I : |g(x)| \leq h\}$ ,  $h > 0$ . If, on  $I$ ,  $|g^{(p)}(x)| \geq c > 0$  for some constant  $c$ , then  $|I_h| \leq c'h^{\frac{1}{p}}$ , where  $c' = p + 2 + \frac{2}{c}$ .*

*Proof.* See [39], Lemma 2.1.  $\square$

Let

$$R_k^{\nu+1}(\gamma) = \{\omega \in \mathcal{O}_\nu : |L_{0k}^\nu| \leq \frac{\gamma}{|k|^\tau}, \text{ or } |\det L_{1k}^\nu| \leq \frac{\gamma^{2m}}{|k|^{2m\tau}}, \text{ or } |\det L_{2k}^\nu| \leq \frac{\gamma^{4m^2}}{|k|^{4m^2\tau}}\}.$$

We note by Lemma 4.1 4) that

$$\mathcal{O}_{\nu+1} = \mathcal{O}_\nu \setminus \bigcup_{K_\nu < |k| \leq K_{\nu+1}} R_k^{\nu+1}(\gamma_\nu),$$

for all  $\nu = 0, 1, \dots$ , which implies that

$$\mathcal{O}_0 \setminus \mathcal{O}_* = \bigcup_{\nu=0}^{\infty} \bigcup_{K_\nu < |k| \leq K_{\nu+1}} R_k^{\nu+1}(\gamma_\nu).$$

Thus, the measure estimate (5.5) amounts to the estimate of  $R_k^{\nu+1}(\gamma)$  for all  $\nu$  and  $k$ .

Given  $k = (k_1, k_2, \dots, k_n)^\top \in Z^n \setminus \{0\}$ ,  $\nu = 0, 1, \dots$ . Without loss of generality, we assume that  $k_1 = \max\{|k_i|\}$ . For any  $(\omega_2, \dots, \omega_n)$ , we consider the sets  $I = \{\omega_1 : \omega = (\omega_1, \omega_2, \dots, \omega_n)^\top \in \mathcal{O}_\nu\}$ , and

$$S_1 = \{\omega_1 \in I : \omega = (\omega_1, \omega_2, \dots, \omega_n)^\top, |g(\omega)| \leq \frac{\gamma_\nu^{4m^2}}{|k|^{4m^2\tau}}\},$$

where  $g(\omega) = \det(L_{2k}^\nu(\omega))$ . By (4.6), we have that, on  $I$ ,

$$A_k = \left| \frac{\partial^{4m^2}}{\partial \omega_1^{4m^2}} g(\omega) \right| = |k_1|^{4m^2} ((4m^2)! + O(\frac{1}{|k|+1}) + O(\mu^{\frac{1}{4}})),$$

where  $O(\frac{1}{|k|+1})$ ,  $O(\mu^{\frac{1}{4}})$  are independent of  $\nu, \omega, l_0$ . Thus, there is a positive integer  $n_0$  such that

$$A_k \geq |k_1| \geq 1,$$

provided that  $|k| \geq n_0$  and  $\mu$  is small. Hence by Lemma 5.2,

$$(5.6) \quad |S_1| \leq (4m^2 + 3) \frac{\gamma_\nu}{|k|^\tau},$$

provided that  $|k| \geq n_0$ . It follows from (5.6) and Fubini's theorem that if  $|k| \geq n_0$ , then

$$|\{\omega \in \mathcal{O}_\nu : |g(\omega)| \leq \frac{\gamma_\nu^{4m^2}}{|k|^{4m^2\tau}}\}| \leq c|S_1| \leq c \frac{\gamma}{|k|^\tau}.$$

Similarly, by making  $n_0$  larger if necessary, we have that

$$\begin{aligned} |\{\omega \in \mathcal{O}_\nu : |L_{0k}^\nu| \leq \frac{\gamma_\nu}{|k|^\tau}\}| &\leq c \frac{\gamma}{|k|^\tau}, \\ |\{\omega \in \mathcal{O}_\nu : |\det L_{1k}^\nu| \leq \frac{\gamma_\nu^{2m}}{|k|^{2m\tau}}\}| &\leq c \frac{\gamma}{|k|^\tau} \end{aligned}$$

for all  $|k| \geq n_0$ . As all constants above are independent of  $k, \nu$ , we conclude that

$$(5.7) \quad |R_k^{\nu+1}(\gamma_\nu)| \leq c \frac{\gamma}{|k|^\tau}$$

for all  $\nu$  and all  $|k| \geq n_0$ . Let  $\nu_0$  be such that  $K_\nu \geq n_0$  as  $\nu \geq \nu_0$ . Then

$$(5.8) \quad \left| \bigcup_{\nu=\nu_0}^{\infty} \bigcup_{K_\nu < |k| \leq K_{\nu+1}} R_k^{\nu+1}(\gamma_\nu) \right| \leq c\gamma \sum_{\nu=\nu_0}^{\infty} \sum_{K_\nu < |k| \leq K_{\nu+1}} \frac{1}{|k|^\tau} = O(\gamma).$$

We now estimate  $R_k^{\nu+1}(\gamma_\nu)$  for  $0 < |k| \leq K_\nu$ ,  $\nu \leq \nu_0$ . Since, by Lemma 4.1 1),  $|\Omega_\nu - id|_{\mathcal{O}_\nu}$ ,  $|M^\nu - M^0|_{\mathcal{O}_\nu} = O(\gamma^{2a})$ , one can make  $\gamma$  small such that  $R_k^{\nu+1}(\gamma_\nu)$  is contained in the set

$$\{\omega \in \mathcal{O}_0 : |\langle k, \omega \rangle| \leq \frac{2\gamma}{|k|^\tau}, \text{ or } |\det L_{1k}^0| \leq \frac{2\gamma^{2m}}{|k|^{2m\tau}}, \text{ or } |\det L_{2k}^0| \leq \frac{2\gamma^{4m^2}}{|k|^{4m^2\tau}}\}$$

for all  $0 < |k| \leq K_\nu$ ,  $\nu \leq \nu_0$ . It follows from NR) and Lemma 5.1 that

$$|R_k^{\nu+1}(\gamma)| \rightarrow 0, \text{ as } \gamma \rightarrow 0,$$

uniformly for all  $0 < |k| \leq K_\nu$ ,  $\nu \leq \nu_0$ .

Consequently,

$$(5.9) \quad \left| \bigcup_{\nu=0}^{\nu_0} \bigcup_{0 < |k| \leq K_\nu} R_k^{\nu+1}(\gamma_\nu) \right| \rightarrow 0,$$

as  $\gamma \rightarrow 0$ .

Thus, by (5.8) and (5.9),

$$|\mathcal{O} \setminus \mathcal{O}_*| \leq \left| \bigcup_{\nu=0}^{\nu_0} \bigcup_{0 < |k| \leq K_\nu} R_k^{\nu+1}(\gamma_\nu) \right| + \left| \bigcup_{\nu=\nu_0}^{\infty} \bigcup_{K_\nu < |k| \leq K_{\nu+1}} R_k^{\nu+1}(\gamma) \right| \rightarrow 0,$$

as  $\gamma \rightarrow 0$ .

This completes the theorem.

## 6. APPLICATIONS

**6.1. Quasi-periodic solutions near an equilibrium.** Consider a real analytic Hamiltonian  $H(p, q)$ ,  $p \in R^d$ ,  $q \in R^d$ , with the standard symplectic structure. We assume that the origin is an equilibrium point and  $J\partial^2 H(0)$  admits  $n$  ( $1 \leq n < d$ ) distinct pairs of purely imaginary eigenvalues  $\pm\sqrt{-1}\sigma_1, \dots, \pm\sqrt{-1}\sigma_n$ , where  $\sigma_k > 0$ ,  $k = 1, 2, \dots, n$ , and  $J$  denotes the standard  $d \times d$  symplectic matrix. Let

$$\Sigma_0 = (\sqrt{-1}\sigma_1, \dots, \sqrt{-1}\sigma_n)^\top, \quad \Lambda_0 = (\lambda_1^0, \dots, \lambda_m^0)^\top,$$

where  $m = d - n$  and  $\pm\lambda_1^0, \dots, \pm\lambda_m^0$  denote the rest of the eigenvalues of  $J\partial^2 H(0)$ .

We assume the following joint non-resonance conditions:

**A1)**  $\langle k, \Sigma_0 \rangle + \langle l, \Lambda_0 \rangle \neq 0$  for all  $k \in Z^n$ ,  $l \in Z^m$  with  $1 \leq |k| \leq K$ ,  $1 \leq |k| + |l| \leq K$ , where  $K = 24m^2$ .

**A2)** There are  $\gamma_0 > 0$ ,  $\tau > n - 1$  such that

$$|\langle k, \Sigma_0 \rangle + \langle l, \Lambda_0 \rangle| > \frac{\gamma_0}{|k|^\tau}$$

for all  $k \in Z^n \setminus \{0\}$ ,  $l \in Z^m$  with  $|l| = 2$ .

The condition A1) particularly implies that  $\{\sigma_1, \sigma_2, \dots, \sigma_n\}$  are non-resonant up to order  $K$ .

We now derive a normal form of  $H$  near the origin. Clearly, by using a linear symplectic transformation,  $H$  can be reduced to a Hamiltonian  $H(p, q)$  of the form

$$(6.1) \quad H(p, q) = \sum_{k=1}^n \lambda_k p_k q_k + H_2(Z) + H_3(p, q, Z) + \dots,$$

where  $p, q \in R^n$ ,  $Z = (p_{n+1}, \dots, p_d, q_{n+1}, \dots, q_d)^\top \in R^{2m}$ , and for each  $i \geq 2$ ,  $H_i$  is a homogeneous polynomial of degree  $i$ . Let  $\xi_k = p_k q_k$ ,  $k = 1, 2, \dots, n$ ,  $\xi = (\xi_1, \xi_2, \dots, \xi_n)^\top$ . It follows from A1) and [8] (pp 132) that the Hamiltonian (6.1) in a neighborhood of 0 can be normalized to

$$(6.2) \quad H = \langle \Lambda_0, \xi \rangle + H_2(Z) + \sum_{k=3}^K H_k(\xi, Z) + P(p, q),$$

where  $P$  consists of monomials of  $p, q$  of degrees  $\geq K + 1$ , and for each  $k = 1, 2, \dots, K$ ,  $H_k(\xi, Z)$  is a homogeneous polynomial which is uniquely determined

and linearly combines monomials  $\xi^\alpha Z^\beta$ ,  $0 \leq |\alpha| \leq [\frac{k}{2}]$ ,  $2|\alpha| + |\beta| = k$ . Consider the following symplectic change of coordinate

$$p = \frac{1}{\sqrt{2}}(p' + \sqrt{-1}q'), \quad q = \frac{1}{\sqrt{2}}(p' - \sqrt{-1}q')$$

and let  $(Y, x) \in R^n \times T^n$  be the standard action-angle variables associated to  $p', q'$ , i.e.,  $Y = (Y_1, Y_2, \dots, Y_n)^\top$ ,  $x = (x_1, x_2, \dots, x_n)^\top$  with  $Y_k = \frac{1}{2}(p'_k{}^2 + q'_k{}^2)$ ,  $p'_k = \sqrt{Y_k} \cos x_k$  and  $q'_k = \sqrt{Y_k} \sin x_k$ ,  $k = 1, 2, \dots, n$ . Then the action-angle form of the Hamiltonian (6.2) reads

$$(6.3) \quad \begin{aligned} H(x, Y, Z) &= N(Y, Z) + P(x, Y, Z), \\ N(Y, Z) &= \langle \omega_0, Y \rangle + \frac{1}{2} \langle Y, AY \rangle + \langle Y, BZ \rangle + \frac{1}{2} \langle Z, CZ \rangle + h(Y, Z), \end{aligned}$$

where  $\omega_0 = (\sigma_1, \sigma_2, \dots, \sigma_n)^\top$ ,  $C = D^2 H_2(0)$ ,  $A$  is contributed by the monomials  $\{Y^\alpha : |\alpha| = 2\}$  in  $H_4$ ,  $B$  is contributed by the monomials  $\{Y^\alpha Z^\beta : |\alpha| = 1, |\beta| = 1\}$  in  $H_3$ ,  $P(x, Y, Z)$  is defined by  $P(p, q)$  through the above changes of variables, and  $h(Y, Z)$  is a polynomial which is associated to the rest of terms in (6.2) of degree 3 or higher.

Viewing  $P$  in the above as a perturbation, we note that the equation of motion associated to the unperturbed part of (6.3) reads

$$(6.4) \quad \begin{cases} \dot{x} &= \omega_0 + AY + BZ + \frac{\partial h(Y, Z)}{\partial Y}, \\ \dot{Y} &= 0, \\ \dot{Z} &= JCZ + JB^\top Y + J \frac{\partial h(Y, Z)}{\partial Z}, \end{cases}$$

where  $J$  denotes the standard  $2m \times 2m$  symplectic matrix. Let

$$\mathcal{M} = \begin{pmatrix} A & B \\ B^\top & C \end{pmatrix}$$

and assume that

**A3)**  $\mathcal{M}$  is non-singular.

Then by the implicit function theorem, the parameterized equation

$$(6.5) \quad \mathcal{M} \begin{pmatrix} Y \\ Z \end{pmatrix} + \nabla h(Y, Z) = \begin{pmatrix} \omega - \omega_0 \\ 0 \end{pmatrix}$$

has an analytic family of solutions  $(Y(\omega), Z(\omega))^\top$  with  $|(Y(\omega), Z(\omega))| = O(|\omega - \omega_0|)$  as  $|\omega - \omega_0| \rightarrow 0$ , which clearly corresponds to an analytic family of invariant  $n$ -tori  $T_\omega$  of (6.4) with toral frequencies parameterized by  $\omega$ . Thus, with the above setting, the persistence problem of lower dimensional tori of  $H(p, q)$  in the vicinity of the origin becomes a perturbation problem for such a family of invariant  $n$ -tori. We note that the hypothesis A3) above can be weakened by assuming the existence of solutions of (6.5) instead. In this case, some branches of solutions  $(Y(\omega), Z(\omega))^\top$  may only depend on  $\omega$  smoothly.

Now, by introducing the translation of coordinates

$$y = Y - Y(\omega), \quad z = Z - Z(\omega),$$

the Hamiltonian (6.3) takes the form of (1.1), i.e.,

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

where

$$\begin{aligned} e(\omega) &= \frac{1}{2} \left\langle \begin{pmatrix} Y(\omega) \\ Z(\omega) \end{pmatrix}, \mathcal{M} \begin{pmatrix} Y(\omega) \\ Z(\omega) \end{pmatrix} \right\rangle + h(Y(\omega), Z(\omega)), \\ M(\omega) &= \mathcal{M} + D^2 h(Y(\omega), Z(\omega)), \\ h(x, y, z, \omega) &= h(y + Y(\omega), z + Z(\omega)) - h(Y(\omega), Z(\omega)) - \nabla h(Y(\omega), Z(\omega)) \begin{pmatrix} y \\ z \end{pmatrix} \\ &\quad - \frac{1}{2} \left\langle \begin{pmatrix} y \\ z \end{pmatrix}, D^2 h(Y(\omega), Z(\omega)) \begin{pmatrix} y \\ z \end{pmatrix} \right\rangle, \\ P(x, y, z, \omega) &= P(x, y + Y(\omega), z + Z(\omega)). \end{aligned}$$

To apply our main results, we let  $s > 0$  be sufficiently small and choose

$$\gamma = s^{1+a_0}, \quad \mu_* = s^{a_0}, \quad \mathcal{O} = \{\omega \in \mathbb{R}^n : 2s \leq |\omega - \omega_0| \leq 4s\},$$

where  $0 < a_0 \ll 1$ . Such choice of  $\gamma$  ensures that the set

$$\hat{\mathcal{O}}_\gamma = \{\omega \in \mathcal{O} : |\langle k, \omega \rangle| > \frac{\gamma}{|k|^\tau}, \quad k \neq 0\}$$

is of positive Lebesgue measure. In fact,

$$|\mathcal{O} \setminus \hat{\mathcal{O}}_\gamma| = O(\gamma) = O(s^{1+a_0}).$$

Since

$$P(x, y, z, \omega) = O((|y| + |z| + |\omega - \omega_0|)^{\frac{\kappa+1}{2}}),$$

there is a constant  $c > 0$  such that

$$|P|_{D(r,s) \times \mathcal{O}} \leq c\gamma^{3(1+\sigma_0)a} \mu_*,$$

where

$$a = 4m^2, \quad 0 < \sigma_0 < \frac{\frac{1}{2} - a_0 - 3aa_0}{3a(1+a_0)}.$$

Therefore, the smallness condition (2.3) of the perturbation is satisfied with  $\mu = c\mu_*$ . By the definition of  $M(\omega)$ , it is clear that  $M(\omega)$  is non-singular on  $\mathcal{O}$  if  $s$  is sufficiently small.

We now verify the non-resonance condition NR). Let  $M_{22} = M_{22}(\omega)$  be the  $2m \times 2m$  lower right block of  $M(\omega)$  and denote the eigenvalues of  $JM_{22}(\omega)$  by  $\pm\lambda_i(\omega)$ ,  $i = 1, 2, \dots, m$ . Consider

$$\begin{aligned} R_k(\gamma) &= \{\omega \in \mathcal{O} : |\det(\sqrt{-1}\langle k, \omega \rangle I_{4m^2} - (M_{22}J) \otimes I_{2m} \\ &\quad - I_{2m} \otimes (M_{22}J))| \leq \frac{\gamma^{4m^2}}{|k|^{4m^2\tau}}\}, \end{aligned}$$

$\gamma > 0$ ,  $k \in \mathbb{Z}^n \setminus \{0\}$ . Then Lemma 5.1 implies that

$$(6.6) \quad R_k(\gamma) = \{\omega \in \mathcal{O} : \left| \prod_{l \in \mathbb{Z}^m, |l|=2} (\sqrt{-1}\langle k, \omega \rangle + \langle l, \Lambda(\omega) \rangle) \right| \leq \frac{\gamma^{4m^2}}{|k|^{4m^2\tau}}\},$$

$\gamma > 0$ ,  $k \in \mathbb{Z}^n \setminus \{0\}$ , where  $\Lambda(\omega) = (\lambda_1(\omega), \dots, \lambda_m(\omega))^\top$ .

By a similar measure estimate as in Section 5.2, we first observe that there is an integer  $k_0 > 0$  and a constant  $c > 0$  such that

$$|R_k(\gamma)| \leq c \frac{\gamma}{|k|^\tau},$$

for all  $|k| \geq k_0$ ,  $\gamma > 0$ . It follows that

$$(6.7) \quad \left| \bigcup_{|k| \geq k_0} R_k(\gamma) \right| = O(\gamma).$$

Since  $\lambda_i(\omega) \rightarrow \lambda_i^0$ ,  $i = 1, 2, \dots, m$ ,  $\omega \in \mathcal{O}$ , uniformly as  $s \rightarrow 0$ , A2) and (6.6) imply that there is a  $\gamma_0 > 0$  such that

$$(6.8) \quad R_k(\gamma) = \emptyset,$$

for all  $|k| < k_0$  and  $\gamma \leq \gamma_0$ , provided that  $s$  is sufficiently small. Let

$$\tilde{\mathcal{O}}_\gamma = \{\omega \in \mathcal{O} : |\sqrt{-1}\langle k, \omega \rangle + \langle l, \Lambda(\omega) \rangle| > \frac{\gamma^{4m^2}}{|k|^{4m^2\tau}}, k \in Z^n \setminus \{0\}, |l| = 2\}.$$

It follows from (6.6)-(6.8) that

$$|\mathcal{O} \setminus \tilde{\mathcal{O}}_\gamma| = O(\gamma),$$

as  $\gamma \rightarrow 0$ , which particularly implies that the set

$$\{\omega \in \mathcal{O} : \sqrt{-1}\langle k, \omega \rangle + \langle l, \Lambda(\omega) \rangle \neq 0, k \in Z^n \setminus \{0\}, |l| = 2\}$$

admits full Lebesgue measure relative to  $\mathcal{O}$ , i.e., NR) holds as long as  $s$  is sufficiently small. Thus, by Theorem 2, we have the following.

**Proposition 6.1.** *If conditions A1)-A3) hold, then there is a  $s > 0$  sufficiently small and Cantor-like sets  $\mathcal{O}_\gamma \subset \tilde{\mathcal{O}}_\gamma$  with*

$$(6.9) \quad |\mathcal{O} \setminus \mathcal{O}_\gamma| = O(\gamma) = O(s^{1+a_0})$$

*such that the Hamiltonian  $H(p, q)$  admits a Whitney smooth family of invariant, quasi-periodic  $n$ -tori in an  $O(s)$ -neighborhood of the origin with the toral frequency  $\omega \in \mathcal{O}_\gamma$ .*

We note that in this particular case the measure estimate (6.9) is more concrete than the one stated in the theorem.

**Remark 6.1** In the above, there are no restrictions on the type, multiplicity or singularity of the normal eigenvalues  $\{\pm\lambda_1^0, \dots, \pm\lambda_m^0\}$ . Also, the matrix  $M(\omega)$  (or  $\mathcal{M}(\omega)$ ) is in general a non-diagonal matrix due to the existence of the term  $H_3$  in (6.3). In the case of multiple or singular normal eigenvalues, neither can such a term be eliminated via canonical transformations nor can it be treated as a perturbative term via re-scalings.

**Example 6.1** (*Normal Degeneracy*). As a special case, we consider the case that  $\Lambda_0 = 0$ , i.e., the normal eigenvalues of  $JD^2H(0)$  are completely degenerate. Let  $\Sigma_0$  be Diophantine. Then both A1) and A2) hold automatically. Hence the conclusions of Proposition 6.1 hold if A3) is satisfied.



**6.2. Coupled oscillators.** Consider a system of near-neighboring coupled oscillators with the following real analytic Hamiltonian

$$(6.10) \quad H(p, q) = \sum_{j=1}^d \left( \frac{1}{2} p_j^2 + U_j(q_j) \right) + \sum_{j=1}^{d-1} \eta_j P_j(q_{j+1} - q_j),$$

where  $p = (p_1, p_2, \dots, p_d)^\top$ ,  $q = (q_1, q_2, \dots, q_d)^\top$ , and  $\eta_j$ ,  $j = 1, 2, \dots, d$ , are coupling coefficients. We first assume that

- A4)** the potential energies  $U_j$  admit critical points  $q_j^0$  for  $j = 1, 2, \dots, d$  respectively such that the first  $n$  ( $0 < n < d$ ) of them are elliptic, i.e.,  $U_j'(q_j^0) = 0$ ,  $j = 1, 2, \dots, d$ , and  $U_j''(q_j^0) > 0$ ,  $j = 1, 2, \dots, n$ .

Then by the Liouville-Arnold integrability theorem, in a small annulus  $G$  around  $\{0\} \times \{(q_1^0, q_2^0, \dots, q_n^0)^\top\} \in R^{2n}$ , one can express the first  $n$  oscillators into action-angle variables  $(I_j, \phi_j)$ ,  $j = 1, 2, \dots, n$ , such that

$$\frac{1}{2} p_j^2 + U_j(q_j) \equiv H_j(I_j),$$

for  $I = (I_1, I_2, \dots, I_n)^\top$  lying in some small annulus  $O$  around the origin of  $R^n$ , where  $H_j$ 's are real analytic functions defined near the origin with  $H_j''(I_j) \neq 0$ ,  $j = 1, 2, \dots, n$ . As  $I^0 = (I_1^0, I_2^0, \dots, I_n^0)^\top$  varies in  $O$ , we treat  $\omega = (H_1'(I_1^0), H_2'(I_2^0), \dots, H_n'(I_n^0))^\top$  as a parameter in a bounded closed region  $\mathcal{O}$  and define  $A(\omega) = \text{diag}\{H_1''(I_1^0), H_2''(I_2^0), \dots, H_n''(I_n^0)\}$ .

Let  $m = d - n$ ,  $y = I - I^0$ ,  $x = (\phi_1, \phi_2, \dots, \phi_n)^\top$ ,  $z = (p_{n+1}, \dots, p_d, q_{n+1}, \dots, q_d)^\top$ . Then the Hamiltonian (6.10) reduces to the form

$$(6.11) \quad \begin{aligned} H_1(x, y, z, \omega) &= e_1(\omega) + \langle \omega, y \rangle + \frac{1}{2} \left\langle \begin{pmatrix} y \\ z \end{pmatrix}, M_1(\omega) \begin{pmatrix} y \\ z \end{pmatrix} \right\rangle \\ &+ O((|y| + |z|)^3) + P_1(x, y, z, \omega), \end{aligned}$$

where  $\omega \in \mathcal{O}$ ,  $x \in T^n$ ,  $(y, z)$  lies in a small neighborhood of the origin in  $R^n \times R^{2m}$ ,  $M_1(\omega) = \text{diag}\{A(\omega), C_1\}$ ,  $C_1 = \text{diag}\{I_m, U_{n+1}''(q_{n+1}^0), \dots, U_d''(q_d^0)\}$ ,

$$P_1(x, y, z, \omega) = \sum_{j=1}^{d-1} \eta_j P_j(q_{j+1} - q_j).$$

Thus,  $M_1(\omega)$  is non-singular on  $\mathcal{O}$  as long as  $U_j''(q_j^0) \neq 0$  for all  $j = n+1, n+2, \dots, d$ , and  $P_1$  is small as long as the coupling coefficients are small. Since  $C_1$  is a constant matrix, a straightforward measure estimate shows that NR) holds on  $\mathcal{O}$  automatically. Applying Theorem 2 to (6.11), we then have the following.

**Proposition 6.2.** *Assume A4) and that  $U_j''(q_j^0) \neq 0$ ,  $j = n+1, n+2, \dots, d$ . Then as  $\eta = (\eta_1, \eta_2, \dots, \eta_{d-1})$  sufficiently small, there are Cantor-like sets  $G_\eta \subset G \subset R^{2n}$  with  $|G \setminus G_\eta| \rightarrow 0$  as  $|\eta| \rightarrow 0$  such that for each  $(p^0, q^0) \in G_\eta$  the unperturbed torus associated to  $(p^0, q^0)$  persists and gives rise to a slightly deformed, invariant, quasi-periodic  $n$ -torus of (6.10) with the same toral frequency.*

The application of our results also allows some strong couplings among the remaining  $m$  oscillators. Instead of assuming the weak couplings, we assume the following:

- A5)**  $q_{n+1}^0 = \dots = q_d^0 = q^0$ ,  $F'_{n+1}(0) = \dots = F'_{d-1}(0) = 0$ .

Denote  $\alpha_j = U_j''(q^0)$ ,  $j = n+1, n+2, \dots, d$ ,

$$\beta_j = \begin{cases} 0, & j = n, \\ \eta_j F_j''(q^0), & n+1 \leq j \leq d-1, \\ 0, & j = d, \end{cases}$$

$j = n, n+1, \dots, d$ , and let  $C_2 = \text{diag}\{I_m, C\}$ , where  $C = (c_{ij})$  is the  $m \times m$  symmetric matrix defined by

$$c_{ij} = \begin{cases} \alpha_{n+i} + \beta_{n+i-1} + \beta_{n+i}, & j = i, \\ -\beta_{n+i}, & j = i+1, \\ 0, & j > i+1. \end{cases}$$

Then with respect to the new coordinate  $(x, y, z)$  above, the Hamiltonian (6.10) becomes

$$(6.12) \quad H_2(x, y, z, \omega) = e_2(\omega) + \langle \omega, y \rangle + \frac{1}{2} \left\langle \begin{pmatrix} y \\ z \end{pmatrix}, M_2(\omega) \begin{pmatrix} y \\ z \end{pmatrix} \right\rangle + O((|y| + |z|)^3) + P_2(x, y, z, \omega),$$

where  $\omega \in \mathcal{O}$ ,  $M_2(\omega) = \text{diag}\{A(\omega), C_2\}$ ,

$$P_2(x, y, z, \omega) = \sum_{j=1}^n \eta_j P_j(q_{j+1} - q_j).$$

Thus, Theorem 2 can be also applied to (6.12) to yield the following result:

**Proposition 6.3.** *Assume A5) and that  $C$  is non-singular. Then as  $\eta = (\eta_1, \eta_2, \dots, \eta_n)$  sufficiently small, the conclusion of Proposition 6.2 holds on  $G$ .*

**Remark 6.2** It is easy to see that the invertibility of  $C$  holds in the following two particular situations:

- i)  $\alpha_j = 0$ ,  $\beta_j \neq 0$ ,  $j = n+1, \dots, d-1$ . In this case, the remaining  $m$  oscillators are completely degenerate at  $q^0$ .
- ii)  $\alpha_j > 0$ ,  $\beta_j \geq 0$ ,  $j = n+1, \dots, d$ . In this case,  $q^0$  is an elliptic critical point for all remaining  $m$  oscillators, and  $C$  is diagonally dominant (hence non-singular).

**Example 6.2** (*Coupled Pendulums*). As a special case, we consider  $d$  coupled mathematical pendulums of the following Hamiltonian:

$$(6.13) \quad H(p, q) = \sum_{j=1}^d \left( \frac{1}{2} p_j^2 + 1 - \cos q_j \right) + \sum_{j=1}^{d-1} \frac{1}{2} \eta_j (q_{j+1} - q_j)^2,$$

where for each  $j = 1, 2, \dots, d$ ,  $q_j$  and  $p_j = \dot{q}_j$  denote the oscillating angle and the angular velocity of the  $j$ th pendulum respectively, and  $\eta_j \geq 0$  is the spring constant of the spring which links the  $j$ th pendulum with the  $(j+1)$ th one, for  $j = 1, 2, \dots, d-1$  respectively.

Such a system has been used to model a row of idealized pendulums coupled with horizontal springs (see [14] for extensive studies on higher order resonance cases) and a finite lattice of nearest neighboring coupled particles (the Frenkel-Kontorova model, see [4]). It can also be thought of as a spatial discretization of the 1D sine Gordon equation.

Let  $U_j(q) = 1 - \cos q$ ,  $j = 1, 2, \dots, d$ , and  $F_j(q) = q^2$ ,  $j = 1, 2, \dots, d - 1$ . Then  $U'_j(q_j^0) = 0$ ,  $q_j^0 = 0$  or  $\pi$ , and

$$U''_j(q_j^0) = \alpha_j = \begin{cases} 1, & \text{if } q_j^0 = 0; \\ -1, & \text{if } q_j^0 = \pi, \end{cases}$$

for  $j = 1, 2, \dots, d$ .

**Case 1** (*soft springs*). Assume that the pendulums are all coupled with soft springs, i.e.,  $\eta_j$ ,  $j = 1, 2, \dots, d$ , are sufficiently small. Let  $n$  be such that  $q_j^0 = 0$ ,  $j = 1, 2, \dots, n$ , and  $q_j^0 = 0$ , or  $\pi$ ,  $j = n + 1, \dots, d$ . Then Proposition 6.2 is immediately applicable to yield a positive measure set of quasi-periodic, invariant  $n$ -tori near  $\{0\} \times (0, \dots, 0, q_{n+1}^0, \dots, q_d^0) \in R^{2d}$ .

**Case 2** (*hard springs*). Assume that for some  $0 < n < d$ , the first  $n$  pendulums are coupled with soft springs, i.e.,  $\eta_j$ ,  $j = 1, 2, \dots, n$ , are sufficiently small, but the remaining  $d - n$  pendulums are coupled with arbitrary springs. Consider the elliptic case that  $q_j^0 = 0$  for all  $j = 1, 2, \dots, d$ . Then conditions in Remark 6.2 ii) hold and thus Proposition 6.3 is immediately applicable to yield a positive measure set of quasi-periodic, invariant  $n$ -tori near the origin of  $R^{2m}$ .

#### REFERENCES

- [1] V. I. Arnold, Proof of a theorem by A.N. Kolmogorov on the preservation of quasi-periodic motions under small perturbations of the Hamiltonian, *Usp. Mat. Nauk.***18** (1963), 13-40
- [2] J. Bourgain, Construction of quasi-periodic solution for Hamiltonian perturbations of linear equations and applications to nonlinear PDE, *Int. Math. Res. Notices* **11** (1994), 475-497
- [3] J. Bourgain, On Melnikov's persistency problem, *Math. Res. Lett.* **4** (1997), 445-458
- [4] O. M. Braun and Y. S. Kivshar, Nonlinear dynamics of the Frenkel-Kontorova model, *Physics Reports* **306** (1998), 1-108
- [5] H. Broer, G. Huitema and F. Takens, Unfoldings of quasi-periodic tori, *Mem. Amer. Math. Soc.* **83** (1990), 13-42
- [6] H. Broer, G. Huitema and M. B. Sevryuk, Families of quasi-periodic motions in dynamical systems depending on parameters, *Nonlinear Dynamical Systems and Chaos (Proc. dyn. syst. conf., H. W. Broer et. al., eds.)*, Birkhäuser, Basel, 1996, 171-211
- [7] H. Broer, G. Huitema and M. Sevryuk, Quasi-periodic motions in families of dynamical systems, *Lect. Notes Math.* **1645**, Springer-Verlag, 1996
- [8] S.-N. Chow, C. Li and D. Wang, Normal forms and bifurcation of planar vector fields, Cambridge University Press, 1994
- [9] 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
- [10] L. H. Eliasson, Floquet solutions for the 1-dimensional quasi-periodic Schrödinger equation, *Comm. Math. Phys.* **146** (1992), 447-482
- [11] L. H. Eliasson, Perturbations of stable invariant tori for Hamiltonian systems, *Ann. Scuola Norm. Sup. Pisa Cl. Sci. Ser.IV* **15** (1988), 115-147
- [12] S. M. Graff, On the continuation of hyperbolic invariant tori for Hamiltonian systems, *J. Differential Equations* **15** (1974), 1-69
- [13] H. He and J. You, On Eliasson's full measure reducibility problem, preprint 2002
- [14] H. H. Jong, Quasiperiodic breathers in systems of weakly coupled pendulums, Ph. D thesis, Groningen University, 1999
- [15] A. Jorba and C. Simo, On the reducibility of linear differential equations with quasiperiodic coefficients, *J. Differential Equations* **98** (1992), 111-124
- [16] A. Jorba and J. Villanueva, On the persistence lower dimensional invariant tori under quasiperiodic perturbations, *J. Nonl. Sci.* **7** (1997), 427-473
- [17] A. N. Kolmogorov, On the conservation of conditionally periodic motions for a small change in Hamilton's function, *Dolk. Akad. Nauk. SSSR* **98** (1954), 525-530

- [18] R. Krikorian, Global density of reducible quasi-periodic cocycles on  $T^1 \times SU(2)$ , *Annals of Math.* **154** (2001), 269-326
- [19] S. B. Kuksin, Hamiltonian perturbations of infinite dimensional linear systems with an imaginary spectrum, *Funct. Anal. Appl.* **21** (1987), 192-205
- [20] S. B. Kuksin, Nearly integrable infinite-dimensional Hamiltonian systems, *Lect. Notes Math.*, **1556**, Springer-Verlag, 1993
- [21] Y. Li and Y. Yi, On Poincaré - Treshchev tori in Hamiltonian systems, submitted
- [22] Y. Li and Y. Yi, Persistence of hyperbolic tori in Hamiltonian systems, to appear in *J. Differential Equations*
- [23] Y. Li and Y. Yi, A quasi-periodic Poincaré's theorem, *Math. Ann.* **326** (2003), 649-690
- [24] Y. Li and Y. Yi, Persistence of invariant tori in generalized Hamiltonian systems, *Erg. Th. Dyn. Sys.* **22** (2002), 1233-1261
- [25] V. K. Melnikov, On some cases of conservation of conditionally periodic motions under a small change of the Hamiltonian function, *Sov. Math. Dokl.* **6** (1965), 1592-1596
- [26] V. K. Melnikov, A family of conditionally periodic solutions of a Hamiltonian system, *Sov. Math. Dokl.* **9** (1968), 882-886
- [27] J. Moser, Convergent series expansions for quasi-periodic motions, *Math. Ann.* **169** (1967), 136-176
- [28] J. Moser, On invariant curves of area preserving mappings of an annulus, *Nachr. Akad. Wiss. Gött. Math. Phys.* **K1** (1962), 1-20
- [29] J. Moser and J. Pöschel, An extension of a result by Dinaburg and Sinai on quasi-periodic potentials, *Comment. Math. Helv.* **59** (1984), 39-85
- [30] I. O. Parasyuk, On preservation of multidimensional invariant tori of Hamiltonian systems, *Ukrain Mat. Zh.* **36** (1984), 467-473
- [31] J. Pöschel, Integrability of Hamiltonian systems on Cantor-like sets, *Comm. Pure Appl. Math.* **35** (1982), 653-696
- [32] J. Pöschel, On the elliptic lower dimensional tori in Hamiltonian systems, *Math. Z.* **202** (1989), 559-608
- [33] J. Pöschel, A KAM-theorem for some nonlinear partial differential equations, *Ann. Scuola Norm. Sup. Pisa Cl. Sci. Ser. IV* **23** (1996), 119-148
- [34] M. B. Sevryuk, Excitation of elliptic normal models of invariant tori in Hamiltonian systems, *Amer. Math. Soc. Transl.* **180** (1997), 209-218
- [35] C. L. Siegel and J. Moser, Lectures on celestial mechanics, Berlin-Heidelberg-New York, Springer, 1971
- [36] E. M. Stein, Singular integrals and differentiability properties of functions, Princeton University Press, Princeton, New Jersey, 1970
- [37] D. V. Treshchev, The mechanism of destruction of resonance tori of Hamiltonian systems, *Math. USSR Sb.* **68** (1991), 181-203
- [38] J. Xu and J. You, On reducibility of linear differential equations with almost periodic coefficients, *Chinese J. Contemp. Math.* **17** (1996), 375-386
- [39] J. X. Xu, J. G. You and Q. J. Qiu, Invariant tori for nearly integrable Hamiltonian systems with degeneracy, *Math. Z.* **226** (1997), 375-387
- [40] J. G. You, Perturbations of lower dimensional tori for Hamiltonian systems, *J. Differential Equations* **152** (1999), 1-29
- [41] E. Zehnder, Generalized implicit function theorem with applications to some small divisors I and II, *Comm. Pure Appl. Math.* **28** (1975), 91-140; **29** (1976), 49-111

Y. LI: DEPARTMENT OF MATHEMATICS, JILIN UNIVERSITY, CHANGCHUN 130012, P. R. CHINA,  
*E-mail address:* YLIMD@EMAIL.JLU.EDU.CN

SCHOOL OF MATHEMATICS, GEORGIA INSTITUTE OF TECHNOLOGY, ATLANTA, GA 30332, USA,  
*E-mail address:* YI@MATH.GATECH.EDU