ON POINCARÉ - TRESHCHEV TORI IN HAMILTONIAN SYSTEMS

YONG LI AND YINGFEI YI

Abstract. We study the persistence of Poincaré - Treshchev tori on a resonant surface of a nearly integrable Hamiltonian system in which the unperturbed Hamiltonian needs not satisfy the Kolmogorov non-degenerate condition. The persistence of the majority of invariant tori associated to g-non-degenerate relative equilibria on the resonant surface will be shown under a Rüssmann like condition.

1. Introduction

Consider a nearly integrable Hamiltonian system

$$H(x, y) = N(y) + \varepsilon P(x, y, \varepsilon), \tag{1.1}$$

where $x \in T^d$, $y \in G \subset R^d$, N and P are real analytic functions defined on a complex neighborhood of a bounded, closed, connected region G and $T^d \times G \times [-1, 1]$, respectively, and ε is a small parameter. Corresponding to the standard symplectic structure on $T^d \times G$, the unperturbed motion associated to (1.1) reads

$$\begin{cases} \dot{x} &= \omega(y), \\ \dot{y} &= 0, \end{cases}$$

where $\omega(y) = \frac{\partial N}{\partial y}(y)$. Hence the phase space $T^d \times G$ is foliated into unperturbed, invariant d-tori $\{T_y = T^d \times \{y\}\}$ with the toral frequencies $\{\omega(y)\}$.

On one hand, with the Kolmogorov non-degenerate condition:

K) $\dfrac{\partial^2 N}{\partial y^2}(y) = \dfrac{\partial \omega}{\partial y}(y)$ *is non-singular on* G,

the celebrated KAM theorem [1, 8, 12] asserts the persistence of the majority of unperturbed, non-resonant d-tori in the sense that there is a family of Cantor-like sets G_ε of almost full Lebesgue measure (i.e., $|G \backslash G_\varepsilon| \to 0$) such that for any $y \in G_\varepsilon$, T_y is non-resonant and persistent under small perturbations. The KAM theorem is recently shown to be true even under the Rüssmann non-degenerate condition:

R) $\text{rank}\{\partial^\alpha \omega(y) : \ \forall |\alpha| \le d-1\} \equiv d$, *on* G

(see [16, 17, 20] for details). On the other hand, it is well known that the unperturbed, resonant d-tori tend to be destroyed via arbitrary generic perturbations and give rise to a resonance zone containing both stochastic trajectories and regular

---

1991 *Mathematics Subject Classification.* 58F14, 58F05, 58F30, 37J40.

Key words and phrases. Hamiltonian systems, KAM theory, Resonant tori, Rüssmann condition.

The first author is partially supported by NSFC 19971042, National 973 Key Project of China: Nonlinearity, and National outstanding young's award of China. The second author was partially supported by NSF grant DMS0204119.

orbits. To characterize regular orbits in the resonance zone, two important problems arise: 1) Whether the majority of resonant tori associated to a given resonant type will be 'completely' destroyed via generic perturbations; 2) If not, then under what mechanisms can certain fractions of the resonant tori survive under generic perturbations.

To formulate the problems more precisely, we let g be a rank m $(0 < m < d)$ subgroup of Z^d. Then the set

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

referred to as the *g-resonant surface with multiplicity* m, characterizes a unique class of resonant tori $\{T_y : y \in O(g, G)\}$ associated to the resonance type determined by g. It is clear that if the Kolmogorov non-degenerate condition K) holds, then $O(g, G)$ is a $n = d - m$ dimensional, real analytic sub-manifold of G (with boundary). The group g also determines a splitting of the resonant tori in the class as follows. Let $\{\tau_1, \tau_2, \cdots, \tau_m\}$ and $\{\tau'_1, \tau'_2, \cdots, \tau'_n\}$ be bases of g and the quotient group Z^d/g, respectively, such that

$$K_0 = (K_1, K_2)$$

is unimodular, i.e., $\det K_0 = 1$, where

$$K_1 = (\tau'_1, \tau'_2, \cdots, \tau'_n), \quad K_2 = (\tau_1, \tau_2, \cdots, \tau_m).$$

Then the toral automorphism K_0 defines a new coordinate $\binom{\psi}{\varphi} = K_0^\top x$ on T^d, where $\psi = K_1^\top x \in T^n, \varphi = K_2^\top x \in T^m$, under which the g-resonant surface becomes

$$O(g, G) = \{y \in G : K_2^\top \omega(y) = 0\},$$

and moreover, for each $y \in O(g, G)$, the resonant torus T_y is foliated into invariant n-tori

$$T_y(\varphi) = T^n \times \{\varphi\} \times \{y\}, \quad \varphi \in T^m$$

corresponding to relative equilibria (φ, y) of the reduced system, each carries parallel flow with the toral frequency $K_1^\top \omega(y)$.

With respect to the given resonance type determined by the group g, the above problems become finding mechanisms under which the majority of the resonant tori on $O(g, G)$ will not be completely destroyed via generic perturbations in the sense that certain classes of n dimensional sub-tori $\{T_y(\varphi)\}$ of T_y for the majority of $y \in O(g, G)$ will persist under generic perturbations.

One such mechanism was proposed in a classical work of Poincaré ([13]) as follows. Define

$$h(\varphi, y) = \int_{T^n} \bar{P}(\psi, \varphi, y) \mathrm{d}\psi,$$

where,

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

Then for a fixed $y \in O(g, G)$, $h(\cdot, y)$ is a real analytic function on the m-torus, hence admits at least $m+1$ critical points and generically at least 2^m critical points which are all non-degenerate. An n-torus $\{T_y(\varphi)\}$ is said to be *Poincaré non-degenerate* if φ is a non-degenerate critical point of $h(\cdot, y)$. For a given resonant surface $O(g, G)$ as above, we note that the Poincaré non-degeneracy is independent of coordinate (ψ, φ) on T^n and it is also a generic condition in the sense that for a fixed $r > 0$ there is a residual subset $\mathcal{P}$ of the set of real analytic functions on the complex

r-neighborhood of $T^d \times G$ with the sup-norm such that for each perturbation $p \in \mathcal{P}$ $h(\cdot, y) : T^m \to R^1$ is a Morse function (i.e., all its critical points are non-degenerate).

In [13], Poincaré considered the maximal resonance case, i.e., $m = d - 1$. With respect to the terminology above, Poincaré's theorem simply says that if $\operatorname{rank} g = d - 1$ then all unperturbed Poincaré non-degenerate 1-tori of (1.1) on $O(g, G)$ will persist if ε is sufficiently small.

A breakthrough along the direction of Poincaré's theorem was made by Treshchev in [19] who considered the general multiplicity $m = d - n$ of a resonant surface $O(g, G)$ and showed the persistence of all Diophantine, Poincaré non-degenerate, hyperbolic n-tori on $O(g, G)$. More precisely, Treshchev's theorem states the following: Assume that the unperturbed Hamiltonian of (1.1) satisfies the Kolmogorov non-degenerate condition K). If a Poincaré non-degenerate n-torus $\{T_{y_0}(\varphi_0)\}$ on $O(g, G)$ is Diophantine (i.e., its toral frequency $K_1^\top \omega(y_0)$ is Diophantine), and hyperbolic in the sense that

T) *no eigenvalue of* $\dfrac{\partial^2 h}{\partial \varphi^2}(\varphi_0, y_0) K_2^\top \dfrac{\partial^2 N}{\partial y^2}(y_0) K_2$ *is positive or zero,*

then it persists under sufficiently small ε with unchanged toral frequency. We note that the condition T) implies the so-called *g-non-resonant* condition that

G) $K_2^\top \dfrac{\partial^2 N}{\partial y^2}(y_0) K_2$ *is non-singular.*

Similar persistence results for the multiplicity one resonant case (i.e., $m = 1$) were later obtained in the works of Eliasson ([7]), Cheng ([3]), Chierchia and Gallavotti ([4]), Rudnev and Wiggins ([15]). As $h(\cdot, y)$ is a function on the 1-tours in these cases, the persisted $d - 1$-tori or their associated critical points of $h(\cdot, y)$ are either elliptic or hyperbolic.

The case of general multiplicity and general toral types were recently studied by Cong *et al* ([6]) and the authors ([10]). The result in [6] implies that if both the Kolmogorov non-degenerate condition K) and the g-non-resonant condition G) hold on $O(g, G)$, then there exists a family of Cantor-like sets $O_\varepsilon(g, G) \subset O(g, G)$ with $|O(g, G) \setminus O_\varepsilon(g, G)| \to 0$ such that for any $y \in O_\varepsilon(g, G)$ all Poincaré non-degenerate n-tori $T_y(\varphi)$ will persist as ε sufficiently small. The same result was shown in [10] by the authors under the Kolmogorov non-degenerate condition K) only.

In this paper, we shall still consider the case of general multiplicity of the resonant surface with respect to general toral types, however, allowing more degenerate unperturbed Hamiltonian. More precisely, we first assume the g-non-resonant condition on $O(g, G)$, i.e.,

G1) $K_2^\top \dfrac{\partial^2 N}{\partial y^2}(y) K_2$ *is non-singular on* $O(g, G)$.

Under this condition, the map $K_2^\top \omega : G \to R^n$ is of maximal rank, hence $O(g, G)$ is an n dimensional, real analytic submanifold of G. Instead of the Kolmogorov non-degenerate condition K), we then assume the following Rüssmann condition on $O(g, G)$:

R1) $\operatorname{rank}\{\partial_\lambda^\alpha K_1^\top \omega(\lambda) : \ \forall |\alpha| \le n - 1\} \equiv n$, *where* λ *is a local coordinate on* $O(g, G)$.

It is clear that the Rüssmann condition R1) on $O(g, G)$ holds if N itself satisfies the Rüssmann condition R) on G.

Our main result states as follows.

Theorem. *Consider* (1.1) *and let* $O(g,G)$ *be a fixed* n $(<d)$ *dimensional resonant surface. If both the* g*-non-degenerate condition* G1) *and the Rüssmann condition* R1) *hold on* $O(g,G)$*, then there is a family of Cantor-like sets* $O_\varepsilon(g,G)\subset O(g,G)$ *with* $|O(g,G)\setminus O_\varepsilon(g,G)|\to 0$ *as* $\varepsilon\to 0$ *such that for any* ε *sufficiently small and any* $y\in O_\varepsilon(g,G)$*, all Poincaré non-degenerate* n*-tori* $T_y(\varphi)$ *will persist. Moreover, for each fixed* ε *small, the perturbed* n*-tori form a finite number of Whitney smooth families on* $O(g,G)$*.*

Combining the above result with that in [10], one concludes that the majority of Poincaré non-degenerate tori on a resonant surface will persist either under the Kolmogorov non-degeneracy K) on G or under the g-non-degeneracy G1) together with the Rüssmann non-degeneracy R1) on $O(g,G)$. After perturbations, the frequencies of Poincaré non-degenerate tori in the former are kept unchanged due to the Kolmogorov non-degeneracy, which is not necessarily true in the later due to the Rüssmann non-degeneracy.

The above theorem will be proved based on the normal form reduction procedure introduced in [19] and a linear KAM iterative scheme contained in [11].

Through the paper, we shall use the same symbol $|\cdot|$ to denote the sup-norm of vectors and its induced matrix norm, the standard l_1 norm of a lattice Z^p, absolute value of functions, and measure of sets etc.. Also, $[\cdot]$ will denote both the average of a function on T^n and the integral part of a real number. For any (vector, matrix valued) function f defined on a domain D, $|f|_D$ stands for $\sup_D|f|$, and, for any two complex column vectors ξ,ζ of the same dimension, $\langle\xi,\zeta\rangle$ means the transpose of ξ times ζ.

The remaining sections are devoted to the proof of the Theorem. In Section 2, we reduce (1.1) to a normal form on a resonant surface by using the Treshchev reduction technique ([19]). In Section 3, we describe the linear iterative scheme for one KAM step with respect to (1.1) and give necessary estimates for the symplectic transformation and the new Hamiltonian. In Section 4, we state an iteration lemma which checks the validity of all KAM steps and complete the proof.

2. Normal Form

In the sequel, we let $g,G,h,m,n,O(g,G),K_0,K_1,K_2$ be as in the introduction. If there is no Poincaré non-degenerate n-torus for any $y\in O(g,G)$, then there is nothing to be proved in the Theorem. Otherwise, it follows from the compactness, connectness of $O(g,G)$, and the implicit function theorem that the Poincaré non-degenerate n-tori form a finite number of real analytic families $\{T_y(\varphi_j(y)):y\in O(g,G)\}$, where for each j, $\varphi_j:O(g,G)\to T^m$ is a real analytic function and $\varphi_j(y)$ is a non-degenerate critical point of $h(\cdot,y)$ for any $y\in O(g,G)$.

We first consider the persistence problem for a fixed family of Poincaré non-degenerate n-tori $\{T_y(\varphi(y)):y\in O(g,G)\}$, where $\varphi=\varphi_j$ for some j.

Let $\Gamma_{11}(y),\Gamma_{12}(y),\Gamma_{21}(y),\Gamma_{22}(y)$ be the $n\times n, n\times m, m\times n, m\times m$ blacks of $\Gamma(y)=K_0^\top\frac{\partial^2 N}{\partial y^2}(y)K_0$, respectively. In particular, $\Gamma_{22}(y)={K_2}^\top\frac{\partial^2 N}{\partial y^2}(y)K_2$. For any $y_0\in O(g,G)$, the Hamiltonian (1.1), up to a constant, admits the following Taylor expansion

$$H(x,y,\varepsilon)=\langle\omega(y_0),y-y_0\rangle+\frac{1}{2}\langle y-y_0,\frac{\partial^2 N}{\partial y^2}(y_0)(y-y_0)\rangle+\varepsilon P(x,y,\varepsilon)+O(|y-y_0|^3).$$

Consider the symplectic transformation $y - y_0 = K_0 p, q = K_0^\top (x - x_0)$, where x_0 is such that $K_2^\top x_0 = \varphi(y_0)$. Then the above Hamiltonian becomes

$$(2.1)\quad H(q,p,\varepsilon) = \langle \omega^*(y_0), p'\rangle + \frac{1}{2}\langle p, \Gamma(y_0)p\rangle + \varepsilon \bar{P}(q,p) + O(\varepsilon^2) + O(|p|^3),$$

where $\omega^*(y_0) = K_1^\top \omega(y_0)$, p' and p'' are the fist n and the last m components of p respectively, and $\bar{P}(q,p) = P(x_0 + (K_0^\top)^{-1} q, y_0 + K_0 p, 0)$. We also let q' and q'' be the first n and the last m components of q respectively.

We fix $l_0 > n, \tau > \max\{n(n-1)-1, 0\}$ and let $\gamma = \varepsilon^{\frac{1}{48m^2(l_0+1)}}$. Consider the set $\hat{O}$ of $y_0 \in O(g,G)$ such that $\omega^*(y_0)$ is Diophantine of the Diophantine type (γ, τ). Since the frequency map $\omega^* : O(g,G) \to R^1$ satisfies the Rüssmann non-degenerate condition R1), it follows from [20] that $|O(g,G) \setminus \hat{O}| = O(\gamma^{\frac{1}{n-1}}) \to 0$, as $\varepsilon \to 0$.

For $y_0 \in \hat{O}$, we separate the first order resonant terms from the perturbation of (2.1) by considering the symplectic transformation

$$(p, q \,\mathrm{mod} 2\pi) \longrightarrow (Y, X \,\mathrm{mod} 2\pi): \quad p = \frac{\partial S(q,Y)}{\partial q}, X = \frac{\partial S(q,Y)}{\partial Y},$$

where S is a generating function defined by

$$S = \langle Y, q\rangle + \varepsilon \sum_{k \in Z^n \setminus \{0\}} \frac{\sqrt{-1}\hat{h}_k(q'', y_0)}{\langle \omega^*(y_0), k\rangle} e^{\sqrt{-1}\langle k, q'\rangle},$$

with

$$\hat{h}_k(q'', y_0) = \int_{T^m} \bar{P}(q, y_0) e^{-\sqrt{-1}\langle k, q'\rangle} \mathrm{d}q'.$$

Since

$$p' = Y' + \sqrt{-1}\varepsilon \sum_{k \in Z^n} k S_k e^{\sqrt{-1}\langle k, q'\rangle}, \;\; p'' = Y'' + O(\frac{\varepsilon}{\gamma}), \;\; X = q,$$

the transformed Hamiltonian, up to a constant, becomes

$$\begin{aligned} H(X,Y,\varepsilon) &= \langle \omega^*(y_0), Y'\rangle + \frac{1}{2}\langle Y, \Gamma(y_0)Y\rangle + \frac{\varepsilon}{2}\langle X'', \partial_\varphi^2 h(\varphi(y_0), y_0)X''\rangle \\ &\quad + \varepsilon(O(|X''|^3) + O(\frac{\varepsilon}{\gamma}|Y|) + O(\frac{\varepsilon^2}{\gamma}) + O(|Y|^3)), \end{aligned}$$

where $X = (X', X'') \in R^n \times R^m$ (mod 2π), $Y = (Y', Y'') \in R^n \times R^m$. Using the Whitney extension theorem, the above transformation can be extended to depend C^{l_0} smoothly on $y_0 \in O(g,G)$.

Consider the rescaling $Y = \sqrt{\varepsilon}\bar{Y}$. Then the rescaled Hamiltonian reads

$$\begin{aligned} H(X,\bar{Y}) &= \frac{H(X, \sqrt{\varepsilon}\bar{Y})}{\sqrt{\varepsilon}} = \langle \omega^*(y_0), \bar{Y}'\rangle + \frac{\sqrt{\varepsilon}}{2}(\langle \bar{Y}, \Gamma(y_0)\bar{Y}\rangle \\ &+ \langle X'', \partial_\varphi^2 h(\varphi(y_0), y_0)X''\rangle) + \sqrt{\varepsilon}O(|X''|^3) + \sqrt{\varepsilon}(O(\frac{\sqrt{\varepsilon}}{\gamma}|\bar{Y}|) \\ &+ O(\frac{\varepsilon}{\gamma^2}) + \sqrt{\varepsilon}O(|\bar{Y}|^3)). \end{aligned}$$

By passing to a new symplectic transformation

$$(\bar{Y}, X \,\mathrm{mod} 2\pi) \to (\hat{Y}, \hat{X} \,\mathrm{mod} 2\pi)$$

using the generating function $W(\hat{Y},X) = \langle \hat{Y}, X\rangle - \langle \hat{Y}, \Gamma_{12}\Gamma_{22}^{-1}X''\rangle$, we obtain the Hamiltonian

$$\begin{aligned} H(\hat{X},\hat{Y}) &= \langle \omega^*(y_0), \hat{Y}'\rangle + \frac{\sqrt{\varepsilon}}{2}(\langle \hat{Y}', A(y_0)\hat{Y}'\rangle + \langle \hat{Y}'', \Gamma_{22}(y_0)\hat{Y}''\rangle \\ &+ \langle X'', \partial_\varphi^2 h(\varphi(y_0), y_0)X''\rangle) + \sqrt{\varepsilon}O(|\hat{X}''|^3) \\ &+ \sqrt{\varepsilon}(\frac{\sqrt{\varepsilon}}{\gamma}O(|\hat{Y}|) + O(\frac{\varepsilon}{\gamma^2}) + \sqrt{\varepsilon}O(|\hat{Y}|^3)), \end{aligned}$$

where $\hat{Y}', \hat{Y}'', \hat{X}', \hat{X}''$ are defined similarly as Y', Y'', X', X'', respectively, and,

$$\begin{aligned} A(y_0) &= K_1^\top \frac{\partial^2 N}{\partial y^2}(y_0)K_1 \\ &-(K_1^\top \frac{\partial^2 N}{\partial y^2}(y_0)K_2)(K_2^\top \frac{\partial^2 N}{\partial y^2}(y_0)K_2)^{-1}(K_1^\top \frac{\partial^2 N}{\partial y^2}(y_0)K_2)^\top. \end{aligned} \tag{2.2}$$

Now, treating y_0 as the parameter $\lambda \in \hat{O}$ and replacing $\omega^*(y_0)$, y_0, $\hat{X}'$, $\hat{Y}'$, $(\hat{Y}'', \hat{X}'')$ by $\Omega(\lambda)$, λ, x, y, z, respectively, we arrive at the Hamiltonian normal form

$$H = e(\lambda) + \langle \omega(\lambda), y\rangle + \frac{\sqrt{\varepsilon}}{2}\langle \begin{pmatrix} y \\ z \end{pmatrix}, M(\lambda)\begin{pmatrix} y \\ z \end{pmatrix}\rangle + \sqrt{\varepsilon}Q(z,\lambda) + P(x,y,z,\varepsilon,\lambda), \tag{2.3}$$

where $(x,y,z) \in T^n \times R^n \times R^{2m}$, $\lambda \in \hat{O}$, $Q(z,\lambda) = O(|z|^3)$,

$$\begin{aligned} M(\lambda) &= \operatorname{diag}\{A(\lambda), M_{22}(\lambda)\}, \\ M_{22}(\lambda) &= \operatorname{diag}\{\Gamma_{22}(\lambda), \partial_\varphi^2 h(\varphi(\lambda),\lambda)\}, \\ P &= \sqrt{\varepsilon}(\frac{\sqrt{\varepsilon}}{\gamma}O(|(y,z)|) + O(\frac{\varepsilon}{\gamma^2}) + \sqrt{\varepsilon}O(|y|^3)), \end{aligned}$$

Q, P are real analytic in $(y,z) \in D(s) = \{(y,z):\ |y| < s, |z| < s\}$, $(x,y,z) \in D(r,s) = \{(x,y,z):\ |\mathrm{Im}x| < r, |y| < s, |z| < s\}$, respectively, for some $r,s > 0$, Whitney smooth in $\lambda \in O(g,G)$, and e,ω,M are Whitney smooth in $\lambda \in O(g,G)$. By the g-non-degenerate condition, M_{22} is clearly non-singular on $O(g,G)$. Due to the Whitney smoothness, we assume without loss of generality that e, M, Q, P are C^{l_0} in $\lambda \in O(g,G)$.

3. KAM Step

We consider KAM iteration to the normal form (2.3) in which all terms are assumed to be C^{l_0} in $\lambda \in O(g,G)$. To begin with the induction, we initially set $N_0 = N$, $e_0 = e$, $\Omega_0 = \Omega$, $M^0 = M$, $M_{22}^0 = M_{22}$, $Q_0 = Q$, $P_0 = P$, $\mathcal{O}_0 = O(g,G)$, $r_0 = r$, $\beta_0 = s$, $\gamma_0 = \gamma = \varepsilon^{\overline{\frac{1}{48m^2(l_0+1)}}}$, $a = 4m^2(l_0+\sigma_0)$, $\mu_0 = \gamma_0^a\varepsilon^{\frac{1}{36}}$, and $s_0 = {\gamma_0}^a\varepsilon^{\frac{1}{38}}$ for a fixed $\sigma_0 \in (0, \frac{1}{3})$.

Then it is easy to see that

$$|\partial_\lambda^l P_0|_{D(r_0,s_0)\times\mathcal{O}_0} \le \sqrt{\varepsilon}\gamma_0^a s_0^2\mu_0,\ |l| \le l_0.$$

Suppose that after a νth KAM step, we arrive at a real analytic, parameter-dependent Hamiltonian

$$\begin{aligned} H &= H_\nu = N + P, \\ N &= N_\nu = e + \langle \Omega(\lambda), y\rangle + \frac{\sqrt{\varepsilon}}{2}\langle \begin{pmatrix} y \\ z \end{pmatrix}, M(\lambda)\begin{pmatrix} y \\ z \end{pmatrix}\rangle + \sqrt{\varepsilon}Q(y,z,\lambda), \end{aligned} \tag{3.1}$$

where $(x,y,z)\in D=D_\nu=D(r,s)$, $r=r_\nu\le r_0, s=s_\nu\le s_0$, $\gamma=\gamma_\nu\le\gamma_0$, $\lambda\in\mathcal{O}=\mathcal{O}_\nu\subset\mathcal{O}_0$, $e(\lambda)=e_\nu(\lambda),\Omega(\lambda)=\Omega_\nu(\lambda)$ are C^{l_0} smooth on $\mathcal{O}$, $M(\lambda)=M^\nu(\lambda)$ is real symmetric and C^{l_0} smooth on $\mathcal{O}$ whose right lower $2m\times 2m$ block $M=M^\nu$ is non-singular on $\mathcal{O}$, $Q=Q_\nu(y,z,\lambda)=O(|(y,z)|^3)$, Q and $P=P_\nu(x,y,z,\lambda)$ are real analytic in $(y,z)\in\mathcal{D}=\mathcal{D}(s)=\{(y,z):|y|<s,|z|<s\}$ and in $(x,y,z)\in D$ respectively and C^{l_0} smooth in $\lambda\in\mathcal{O}$, and moreover,

$$|\partial_\lambda^l P|_{D\times\mathcal{O}}\le\sqrt{\varepsilon}\gamma^a s^2\mu,\ \ |l|\le l_0, \tag{3.2}$$

for some $\mu=\mu_\nu>0$.

We now construct a symplectic transformation $\Phi_+=\Phi_{\nu+1}$, which, in smaller frequency and phase domains, transforms the Hamiltonian (3.1) into a similar form but 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 + (=$\nu+1$). All constants c_i, $i=1,2,\cdots,5$, 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,σ,δ be positive constants such that

$$\delta<1,\ \ \sigma-(b+\sigma)(2b+3\sigma)>0,\ \ \delta(1+b+\sigma)>1$$

and define

$$\begin{aligned}
r_+ &= \delta r+(1-\delta)(1-\frac{\delta^2}{2})r_0,\\
s_+ &= s^{1+b+\sigma},\\
\beta_+ &= \frac{\beta}{2}+\frac{\beta_0}{4},\\
\gamma_+ &= \frac{\gamma}{2}+\frac{\gamma_0}{4},\\
K_+ &= ([\log\frac{1}{s}]+1)^3,\\
\mathcal{D}_+ &= \mathcal{D}(\beta_+),\\
D_+ &= D(r_+,s_+),\\
\tilde{D}_+ &= D(r_++\frac{5}{8}(r-r_+),\beta_+),\\
D_i &= D(r_++\frac{i-1}{8}(r-r_+),is_+),\ \ i=1,2,3,4,\\
D(\eta) &= D(r_++\frac{7}{8}(r-r_+),\eta),\\
\Gamma(\eta) &= e^{\frac{r_0(1-\delta)\delta^2}{16}}\sum_{0<|k|\le K_+}|k|^\chi e^{-\frac{\eta}{8}},
\end{aligned}$$

where $\eta>0$ and $\chi=2(l_0+1)(4m^2+1)\tau$.

3.1. **Truncation and linear equations.** Let R be the truncation of the Taylor-Fourier series of P up to quadratic order, i.e.,

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

If

$$s_+ \le \frac{s}{16}, \quad \int_{K_+}^{\infty} \lambda^n e^{-\lambda \frac{r-r_+}{16}} \mathrm{d}\lambda \le s, \tag{A}$$

then it follows from a standard argument (e.g., [10]) that there is a constant c_1 such that

$$|\partial_\lambda^l (P-R)|_{D_8} \le c_1 \sqrt{\varepsilon}\gamma^a (s^3 + \frac{s_+^3}{s})\mu, \ |l| \le l_0. \tag{3.3}$$

We first seek for an averaging transformation as the time-1 map ϕ_F^1 of the Hamiltonian flow ϕ_F^t associated to a Hamiltonian of the form

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

in which F_{kij}, $0 \le i+j \le 2$, $0 < |k| \le K_+$, will be solved from the linear homological equations

$$L_{0k} F_{k00} = P_{k00}, \tag{3.5}$$

$$L_{0k} F_{k10} = P_{k10} - \sqrt{\varepsilon} M_{12} J F_{k01}, \tag{3.6}$$

$$L_{1k} F_{k01} = P_{k01}, \tag{3.7}$$

$$L_{0k} F_{k20} = P_{k20} + \frac{\sqrt{\varepsilon}}{2} (F_{k11} J M_{21} - M_{12} J F_{k11}^\top), \tag{3.8}$$

$$L_{1k} F_{k11} = P_{k11} - \sqrt{\varepsilon} (F_{k02}^\top + F_{k02}) J M_{21}, \tag{3.9}$$

$$L_{2k} F_{k02} = P_{k02}, \tag{3.10}$$

where

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

$M_{11}, M_{12}, M_{21}, M_{22}$ are $n \times n, n \times 2m, 2m \times n, 2m \times 2m$ blocks of M respectively, and, $\otimes$ denotes the standard tensor product of matrices.

Consider the set

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

Then, on $\mathcal{O}_+$, equations (3.5)-(3.10) can be solved uniquely to obtain solutions F_{kij}, depending C^{l_0} smoothly on $\lambda \in \mathcal{O}_+$ and satisfying $\bar{F}_{kij} = F_{-kij}$, for all

$0<|k|\le K_+$, $0\le i+j\le 2$, which uniquely determine the Hamiltonian F in (3.4). Moreover, it is easy to see that there is a constant $c>0$ such that

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

for all $q=0,1,2$, from which it follows that there is a constant $c_2>0$ such that

$$|\partial_\lambda^l F|,\ |\partial_\lambda^l F_x|,\ s|\partial_\lambda^l F_y|,\ s|\partial_\lambda^l F_z|\le c_2 s^2\mu\Gamma(r-r_+),\ |l|\le l_0,\ \text{on}\, D(s)\times\mathcal{O}_+, \tag{3.11}$$

$$|\partial_\lambda^l\partial_{(x,y,z)}^i F|\le c_2\mu\Gamma(r-r_+),\ |i|\le 4,\ |l|\le l_0,\ \text{on}\, D(\beta)\times\mathcal{O}_+. \tag{3.12}$$

Let

$$\begin{aligned}\hat{Q} &= -\sqrt{\varepsilon}\sqrt{-1}\sum_{0<|k|\le K_+}\langle k, M_{11}y+M_{12}z+\frac{\partial Q}{\partial y}\rangle(F_{k00}+\langle F_{k10},y\rangle+\langle F_{k01},z\rangle\\ &\quad+\langle y,F_{k20}y\rangle+\langle z,F_{k11}y\rangle+\langle z,F_{k02}z\rangle)e^{\sqrt{-1}\langle k,x\rangle}\\ &\quad-\sqrt{\varepsilon}\sum_{0<|k|\le K_+}\langle\frac{\partial Q}{\partial z},J(F_{k01}+F_{k11}y+F_{k02}z+F_{k02}^\top z)\rangle e^{\sqrt{-1}\langle k,x\rangle}.\end{aligned} \tag{3.13}$$

Then

$$\{N,F\}+R-[R]-\hat{Q}=0. \tag{3.14}$$

We note that $\hat{Q}$ consists of all terms in $\{N,F\}$ of size $O(s^3\mu)$ and of order $O(y^\imath z^\jmath)$ for $|\imath|+|\jmath|\ge 3$.

It follows that

$$\begin{aligned}H\circ\varphi_F^1 &= (N+R)\circ\varphi_F^1+(P-R)\circ\varphi_F^1\\ &= \bar{N}_++\bar{P}_+,\end{aligned}$$

where

$$\begin{aligned}\bar{N}_+ &= N+[R],\\ \bar{P}_+ &= \int_0^1\{R_t,F\}\circ\varphi_F^t\mathrm{d}t+(P-R)\circ\varphi_F^1+\hat{Q},\end{aligned}$$

with

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

If

$$|\partial_\lambda^l(M-M^0)|_{\mathcal{O}}\le\sqrt{\varepsilon}\gamma_0^a\mu_0^{\frac14},\ |l|\le l_0 \tag{B}$$

holds, then by the implicit function theorem, there is a $z_0\in C^{l_0}(\mathcal{O},R^{2m})$ such that

$$(\mathrm{diag}(O,M_{22}(\lambda))+\partial_{(y,z)}^2Q(0,z_0,\lambda))\begin{pmatrix}0\\ z_0\end{pmatrix}+\partial_{(y,z)}Q(0,z_0,\lambda)=-\frac{1}{\sqrt{\varepsilon}}\begin{pmatrix}0\\ P_{001}\end{pmatrix},$$

and moreover, there is a positive constant, again denoted by c_2, such that

$$|\partial_\lambda^l z_0|_{\mathcal{O}_+}\le c_2\gamma^a s\mu,\ |l|\le l_0. \tag{3.15}$$

This defines a translation

$$\varphi: x\to x,\ y\to y,\ z\to z+z_0$$

so that $\Phi_+=\varphi_F^1\circ\varphi$ will transform H into the Hamiltonian H_+ in the next KAM cycle, i.e.,

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

where

$$\begin{aligned} N_+ &= e_+ + \langle \Omega_+(\lambda), y\rangle + \frac{\sqrt{\varepsilon}}{2}\langle \begin{pmatrix} y \\ z \end{pmatrix}, M^+ \begin{pmatrix} y \\ z \end{pmatrix}\rangle + \sqrt{\varepsilon} Q_+(y,z,\lambda), \\ P_+ &= \bar{P}_+ \circ \varphi + 2\langle \begin{pmatrix} y \\ z \end{pmatrix}, \begin{pmatrix} P_{020} & P_{011} \\ P_{011}^\top & P_{002} \end{pmatrix} \begin{pmatrix} 0 \\ z_0 \end{pmatrix}\rangle \end{aligned}$$

with

$$\begin{aligned} e_+ &= e + P_{000} + \sqrt{\varepsilon} Q(0, z_0, \lambda) + \langle P_{001}, z_0\rangle \\ &\quad + \frac{1}{2}\langle \begin{pmatrix} 0 \\ z_0 \end{pmatrix}, \begin{pmatrix} P_{020} & P_{011} \\ P_{011}^\top & P_{002} \end{pmatrix} \begin{pmatrix} 0 \\ z_0 \end{pmatrix}\rangle, \\ \Omega_+ &= \Omega + P_{010}, \\ M^+ &= M + \frac{2}{\sqrt{\varepsilon}} \begin{pmatrix} P_{020} & P_{011} \\ P_{011}^\top & P_{002} \end{pmatrix}, \\ Q_+ &= Q(y, z+z_0, \lambda) - Q(0, z_0, \lambda) - \langle \partial_{(y,z)} Q(0, z_0, \lambda), \begin{pmatrix} y \\ z \end{pmatrix}\rangle \\ &\quad - \frac{1}{2}\langle \begin{pmatrix} y \\ z \end{pmatrix}, \partial^2_{(y,z)} Q(0, z_0, \lambda) \begin{pmatrix} y \\ z \end{pmatrix}\rangle. \end{aligned}$$

It is clear that there is a constant $c_3 > 0$ such that

$$\begin{aligned} |\partial_\omega^l e_+ - \partial_\omega^l e|_{\mathcal{O}_+} &\leq \sqrt{\varepsilon} c_3 \gamma^a s^2 \mu, \ |l| \leq l_0, \\ |\partial_\omega^l M^+ - \partial_\omega^l M|_{\mathcal{O}_+} &\leq \sqrt{\varepsilon} c_3 \gamma^a \mu, \ |l| \leq l_0, \\ |\partial_\omega^l \partial_{(y,z)}^j Q_+ - \partial_\omega^l \partial_{(y,z)}^j Q|_{\mathcal{D}_+ \times \mathcal{O}_+} &\leq c_3 \gamma^a \mu, \ |j| \leq l_0, \ |l| \leq l_0. \end{aligned}$$

By [11], Lemma 3.6, we have that if

$$3c_3 \mu K_+^{8m^2\tau + 8m^2} < \min\{\frac{\gamma - \gamma_+}{\gamma_0}, \frac{\gamma^{2m} - \gamma_+^{2m}}{\gamma_0^{2m}}, \frac{\gamma^{4m^2} - \gamma_+^{4m^2}}{\gamma_0^{4m^2}}\}, \tag{C}$$

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

$$|L_{0k}^+| > \frac{\gamma_+}{|k|^\tau}, \quad |\det L_{1k}^+| > \frac{\gamma_+^{2m}}{|k|^{2m\tau}}, \quad |\det L_{2k}^+| > \frac{\gamma_+^{4m^2}}{|k|^{4m^2\tau}}. \tag{3.16}$$

3.2. **Estimates on the transformation and new perturbation.**

Lemma 3.1. *If*

$$\begin{aligned} c_2 \mu \Gamma(r - r_+) &< \frac{r - r_+}{8}, \\ c_2 s \mu \Gamma(r - r_+) &< s_+, \\ c_2 \mu \Gamma(r - r_+) + c_2 \mu &< \beta - \beta_+, \end{aligned} \tag{D}$$

then the following holds.

1) *For all* $0 \leq t \leq 1$

$$\begin{aligned} \varphi_F^t : D_3 &\to D_4, \\ \varphi : D_1 &\to D_3 \end{aligned}$$

are well defined, real analytic and depend C^{l_0} *smoothly on* $\lambda \in \mathcal{O}_+$.

2) $\Phi_+ : D_+ \times \mathcal{O}_+ \to D \times \mathcal{O}$ *can be extended to a* C^{l_0+1} *function* $\Phi_+ : \tilde{D}_+ \times \mathcal{O}_+ \to D(r, \beta) \times \mathcal{O}$.

3) *There is a constant c_4 such that for all $0 \le t \le 1$, $|l| \le l_0$,*

$$|\partial_x^i \partial_{(y,z)}^j \partial_\lambda^l \varphi_F^t|_{D_3\times\mathcal{O}_+} \le \begin{cases} c_4 s\mu\Gamma(r-r_+), & |i|+|j|=0, |l|\ge 1, \\ c_4\mu\Gamma(r-r_+), & 2\le |l|+|i|+|j|\le l_0, \\ c_4, & \text{otherwise}, \end{cases}$$

$$|\partial_{(x,y,z)}^p \partial_\lambda^l(\Phi_+ - id)|_{\tilde{D}_+\times\mathcal{O}_+} \le c_3\mu\Gamma(r-r_+), \ |p| \le l_0 - |l| + 1,$$

$$|\partial_\lambda^l \varphi|_{D_+\times\mathcal{O}_+} \le c_3\gamma^a s\mu.$$

Proof. The proof easily follows from (3.11)-(3.12) and the Whitney extension theorem. □

Lemma 3.2. *Let* $\Delta = \dfrac{\gamma^a\Gamma(r-r_+)^3}{r-r_+}(s^2\mu^2 + s_+s^2\mu)$. *Then there is a constant $c_5 > 0$ such that $|\partial_\lambda^l P_+|_{D_+\times\mathcal{O}_+} \le c_5\sqrt{\varepsilon}\Delta$, $|l| \le l_0$. Thus, if*

$$c_5\Delta \le \gamma_+^a s_+^2 \mu_+, \tag{E}$$

then $|\partial_\lambda^l P_+|_{D_+\times\mathcal{O}_+} \le \sqrt{\varepsilon}\gamma_+^a s_+^2\mu_+, \ |l| \le l_0$.

Proof. The proof follows from (3.2), (3.3), (3.11)-(3.15), and Lemma 3.1. See [10, 11] for details. □

4. Iteration Lemma

Let $r_0, s_0, \mu_0, \gamma_0, \mathcal{O}_0, H_0, N_0, e_0, \Omega_0, M^0, M_{22}^0, Q_0, P_0$ be 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 $\nu = 0, 1\cdots$, we index all index free quantities in Section 3 by ν and index all " + "-indexed quantities in Section 3 by $\nu + 1$. This yields the following sequences:

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

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

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

Lemma 4.1. (**Iteration Lemma**) *For sufficiently small $\mu_0 = \mu_0(r_0, \beta_0, l_0)$, the KAM steps described in* Section 3 *are valid for all $\nu = 0, 1, \cdots$, and the following holds for all $\nu = 1, 2, \cdots$.*

1) *$e_\nu = e_\nu(\lambda)$, $\Omega_\nu = \Omega_\nu(\lambda)$ are C^{l_0} smooth on $\mathcal{O}_\nu$, $M^\nu = M^\nu(\lambda)$ is real symmetric, C^{l_0} smooth on $\mathcal{O}_\nu$ with the $2m \times 2m$ lower right block M_{22}^ν being non-singular on $\mathcal{O}_\nu$, $Q_\nu = Q_\nu(y, z, \lambda) = O(|(y, z)|^3)$ and $P_\nu = P_\nu(x, y, z, \lambda)$ are real analytic in $(y, z) \in \mathcal{D}_\nu$ and in $(x, y, z) \in \tilde{D}_\nu$ respectively and C^{l_0}*

smooth in $\lambda \in \mathcal{O}_\nu$*. Moreover, for all* $|l| \le l_0$,

$$
\begin{aligned}
&|\partial_\lambda^l e_\nu - \partial_\lambda^l e_{\nu-1}|_{\mathcal{O}_\nu} \le \frac{\gamma_0^{2(1+\sigma_0)a+a}\mu_0^{\frac14}}{2^\nu},\\
&|\partial_\lambda^l e_\nu - \partial_\lambda^l e_0|_{\mathcal{O}_\nu} \le \gamma_0^{2(1+\sigma_0)a+a}\mu_0^{\frac14},\\
&|\partial_\lambda^l \Omega_\nu - \partial_\lambda^l \Omega_{\nu-1}|_{\mathcal{O}_\nu} \le \frac{\gamma_0^{(1+\sigma_0)a+a}\mu_0^{\frac14}}{2^\nu},\\
&|\partial_\lambda^l (\Omega_\nu - id)|_{\mathcal{O}_\nu} \le \gamma_0^{(1+\sigma_0)a+a}\mu_0^{\frac14},\\
&|\partial_\lambda^l M^\nu - \partial_\lambda^l M^{\nu-1}|_{\mathcal{O}_\nu} \le \frac{\gamma_0^a\mu_0^{\frac14}}{2^\nu},\\
&|\partial_\lambda^l M^\nu - \partial_\lambda^l M^0|_{\mathcal{O}_\nu} \le \gamma_0^a\mu_0^{\frac14},\\
&|\partial_\lambda^l \partial_{(y,z)}^j (Q_\nu - Q_{\nu-1})|_{\mathcal{D}_\nu\times\mathcal{O}_\nu} \le \frac{\gamma_0^a\mu_0^{\frac14}}{2^\nu},\ |j| \le l_0,\\
&|\partial_\lambda^l (Q_\nu - Q_0)|_{\mathcal{D}_\nu\times\mathcal{O}_\nu} \le \gamma_0^a\mu_0^{\frac14},\ |j| \le l_0,\\
&|\partial_\lambda^l P_\nu|_{D_\nu\times\mathcal{O}_\nu} \le \sqrt{\varepsilon}\gamma_\nu^a s_\nu^2\mu_\nu.
\end{aligned}
$$

2)

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

3) $\Phi_\nu : D_\nu \times \mathcal{O}_\nu \to D_{\nu-1}, \tilde{D}_\nu \times \mathcal{O}_\nu \to \tilde{D}_{\nu-1}$ *is symplectic for each* $\lambda \in \mathcal{O}_\nu$*, real analytic in* $(x, y, z) \in \tilde{D}_\nu$ *and* C^{l_0+1} *smooth in* $\lambda \in \mathcal{O}_\nu$*, and,*

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

Moreover,

$$|\partial_\lambda^l \partial_{(x,y,z)}^p (\Phi_\nu - id)|_{\tilde{D}_\nu\times\mathcal{O}_\nu} \le \frac{\mu^{\frac14}}{2^\nu},\ |p| + |l| \le l_0 + 1. \tag{4.1}$$

Proof. Similar to [10, 11], one can verify the conditions (A)-(E) inductively. Hence the KAM step described in Section 3 is valid for all ν.

The estimates in 1) follow from the definition of $e_\nu, \Omega_\nu, M^\nu, Q_\nu, P_\nu$ and Lemma 3.2, 2) follows from the induction and (3.16), and 3) follows from Lemma 3.1 (see [10, 11] for details). □

5. Proof of Theorem

We first consider (2.3). Let

$$
\begin{aligned}
&\Psi_\nu = \Phi_0 \circ \Phi_1 \circ \cdots \circ \Phi_\nu : \tilde{D}_\nu \times \mathcal{O}_\nu \to \tilde{D}_0,\\
&H_0 \circ \Psi_\nu = H_\nu = N_\nu + P_\nu,\\
&N_\nu = e_\nu + \langle \Omega_\nu, y\rangle + \frac{\sqrt{\varepsilon}}{2}\langle \begin{pmatrix} y\\ z\end{pmatrix}, M^\nu \begin{pmatrix} y\\ z\end{pmatrix}\rangle + \sqrt{\varepsilon} Q_\nu,
\end{aligned}
$$

$\nu = 0, 1, \cdots$, and,

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

Using the Whitney extension theorem, the frequencies Ω_ν can be extended C^n smoothly on $\mathcal{O}_0 = O(g, G)$ and the extended frequencies also satisfy the corresponding estimates in Lemma 4.1 1). It follows that Ω_ν satisfies the Rüssmann condition R) on $\mathcal{O}_0$ as ε sufficiently small, and therefore from the measure estimate in [20] (see also [5, 9]) that $|O(g, G) \setminus O_\infty| \to 0$ as $\varepsilon \to 0$. Hence if $O_* = O_\infty \cap \hat{O}$, then $|O(g, G) \setminus O_*| \to 0$ as $\varepsilon \to 0$.

By Lemma 4.1, we obtain the uniform convergence of Ψ_ν on $D(\frac{r_0}{2}, \frac{\beta_0}{2}) \times O_*$, say, to Ψ_∞, which is uniformly close to the identity, real analytic in (x, y, z) and C^n Whitney smooth in λ. Also, on $\mathcal{D}(\frac{\beta_0}{2}) \times O_*$, Q_ν is uniformly convergent to a function $Q_\infty = O((|y| + |z|)^3)$ which is C^{l_0} in (y, z) and C^{l_0} Whitney smooth in λ. It follows that, on $D(\frac{r_0}{2}, \frac{\beta_0}{2}) \times O_*$, N_ν converge uniformly to a function

$$N_\infty = e_\infty(\lambda) + \langle \Omega_\infty, y \rangle + \frac{\sqrt{\varepsilon}}{2} \langle \begin{pmatrix} y \\ z \end{pmatrix}, M^\infty(\lambda) \begin{pmatrix} y \\ z \end{pmatrix} \rangle + \sqrt{\varepsilon} Q_\infty$$

which is real analytic in (y, z) and C^{l_0} Whitney smooth in λ. Hence

$$P_\nu = H_0 \circ \Psi_\nu - N_\nu$$

can be extended to $D(\frac{r_0}{2}, \frac{\beta_0}{2}) \times O_*$ and converge uniformly to the function

$$P_\infty = H_0 \circ \Psi_\infty - N_\infty$$

which is real analytic in (x, y, z) and C^n Whitney smooth in λ.

For any $\nu \in Z_+$, $\lambda \in \Lambda$, $j \in Z_+^n$, $k \in Z_+^{2m}$ with $|j| + |k| \le 2$, by applying the last inequality in Lemma 4.1 1) and Cauchy's estimate on $D(r_\nu, \frac{1}{2} s_\nu)$, we have that

$$|\partial_y^j \partial_z^k P_\nu| \le 2^{j+2} \gamma_\nu^a \mu_\nu .$$

Hence

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

for all $x \in T^n$, $\lambda \in \Lambda$, $j \in Z_+^n$, $k \in Z_+^{2m}$ with $|j| + |k| \le 2$. Thus, for each $\lambda \in O_*$, the system (2.3) admits an analytic, quasi-periodic, invariant torus with the Diophantine toral frequency $\Omega_\infty(\lambda)$, which is slightly deformed from the unperturbed torus corresponding to $\Omega_0(\lambda)$. Moreover, these perturbed tori form a C^{l_0} Whitney smooth family. Tracing back to the normal form, we then obtain a Whitney smooth family of invariant, analytic, quasi-periodic n-tori $\{T_{y,\varepsilon}(\varphi^j(y)) : y \in O_*\}$ corresponding to the family of Poincaré non-degenerate, unperturbed n-tori $\{T_y(\varphi^j(y)) : y \in O_*\}$ on $O(g, G)$.

Denote $O_*^j = O_*$. Since by the compactness and connectness of $O(g, G)$ there is only a finite number of such real analytic families of Poincaré non-degenerate, unperturbed n-tori $\{T_y(\varphi^j(y)) : y \in O(g, G)\}$ over $O(g, G)$, the proof of the theorem is now complete by taking $O_\varepsilon(g, G)$ as the intersection of all the O_*^j's.

References

[1] V. I. Arnold, Proof of A. N. Kolmogorov's theorem on the preservation of quasi periodic motions under small perturbations of the Hamiltonian, *Usp. Math. USSR* **18**(1963), 13-40.

[2] H. Broer, G. Huitema and M. Sevryuk, Quasi-periodic motions in families of dynamical systems, *Lect. Notes Math.* **1645**, Springer-Verlag, 1996

[3] C.-Q. Cheng, Birkhoff-Kolmogorov-Arnold-Moser tori in convex Hamiltonian systems, *Commun. Math. Phys.* **177** (1996), 529-559

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

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

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

[7] L. H. Eliasson, Biasymptotic solutions of perturbed integrable Hamiltonian systems, *Bol. Soc. Mat.* **25**(1994), 57-76

[8] A. N. Kolmogorov, On quasi-periodic motions under small perturbations of the Hamiltonian, *Doklady Akad. Nauk. USSR* **98** (1954), 527-530

[9] Y. Li and Y. Yi, Persistence of invariant tori in generalized Hamiltonian systems, *Erg. Th. Dyn. Sys.* **22** (2002), 1233-1261

[10] Y. Li and Y. Yi, A quasi-periodic Poincaré's theorem, *Math. Ann.,* to appear

[11] Y. Li and Y. Yi, Persistence of lower dimensional tori in Hamiltonian systems with degeneracy, submitted

[12] J. Moser, On invariant curves of area preserving mappings of an annulus, *Nachr. Akad. Wiss. Gott. Math. Phys K1,* **2** (1962), 1-20

[13] H. Poincaré, Les Méthodes Nouvelles de la Mécaniques Céleste, I-III, Gauthier-Villars, 1892, 1893, 1899 (The English translation: New Methods of Celestial Mechanics, AIP Press, Williston, 1992)

[14] J. Pöschel, Integrability of Hamiltonian systems on Cantor-like sets, *Comm. Pure Appl. Math.* **35** (1982), 653-696

[15] M. Rudnev and S. Wiggins, KAM theory near multiplicity one resonant surfaces in perturbations of A-priori stable Hamiltonian systems, *J. Nonl. Sci.* **7**(1997), 177-209.

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

[17] M. B. Sevryuk, KAM-stable Hamiltonians, *J. Dyn. Control Syst.* **1** (1995), 351-366

[18] C. L. Siegel and J. K. Moser, Lectures on Celestial Mechanics, *Grundlehren* **187***,* Springer, Berlin, 1971

[19] D. V. Treshchev, Mechanism for destroying resonance tori of Hamiltonian systems, *Mat. USSR. Sb.* **180**(1989), 1325-1346

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

Y. Li: Department of Mathematics, Jilin University, Changchun 130012, P. R. China
E-mail address: `ylimd@email.jlu.edu.cn`

Y. Yi: School of Mathematics, Georgia Institute of Technology, Atlanta GA 30332, USA
E-mail address: `yi@math.gatech.edu`