Qualitative Theory of Dynamical Systems (2024) 23:92
https://doi.org/10.1007/s12346-023-00938-1

A Formal KAM Theorem for Hamiltonian Systems and Its Application to Hyperbolic Lower Dimensional Invariant Tori

Qi Li[1] · Junxiang Xu[1]

Received: 11 July 2023 / Accepted: 8 December 2023 / Published online: 30 January 2024

Abstract
In this paper we reformulate a formal KAM theorem for Hamiltonian systems with parameters under Bruno-Rüssmann condition. The proof is based on KAM iteration and the key is to adjust the parameters for small divisors after KAM iteration instead of in each KAM step. By this formal KAM theorem we can follow some well known KAM-type results for hyperbolic tori. Moreover, it can also be applied to the persistence of invariant tori with prescribed frequencies.

Keywords Formal KAM theorem · Bruno-Rüssmann condition · KAM iteration

1 Introduction

With the development of KAM theory, there are many well known KAM theorems [1, 4, 10, 12–14, 16, 21, 22]. The classical KAM theorem [1, 10, 16] asserts that if the frequency mapping satisfies Kolmogorov non-degeneracy condition, then the Lagrangian invariant tori with Diophantine frequencies can persist under small perturbations. Kolmogorov non-degeneracy condition can be weakened to Bruno non-degeneracy condition and Rüssmann non-degeneracy condition [6, 19, 23, 26], in particular, Rüssmann non-degeneracy condition is sharpest one for KAM theorems. Moreover, the Diophantine condition can be weakened to the Bruno-Rüssmann condition [2, 8, 17–20]. In addition, a similar problem for non-Hamiltonian vector fields with Bruno frequency vectors is studied in [9]. In particular, as an alternative to the KAM method, the renormalization method is used in [8, 9].

In this paper we are concerned about lower dimensional invariant tori with Bruno frequency vectors in Hamiltonian systems. Consider the following real analytic nearly

✉ Junxiang Xu
xujun@seu.edu.cn

Qi Li
qili0529@126.com

[1] Department of Mathematics, Southeast University, Nanjing 210096, China

integrable Hamiltonians

$$H_{\pm}(x, y, u, v) = \langle \omega, y \rangle + \frac{1}{2}\sum_{j=1}^{m} \Omega_j \left(u_j^2 \pm v_j^2\right) + P(x, y, u, v). \tag{1.1}$$

The phase space is $T^n \times \mathbb{R}^n \times \mathbb{R}^m \times \mathbb{R}^m$ associated with the symplectic structure

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

where $T^n = \mathbb{R}^n/2\pi\mathbb{Z}^n$ is the n-torus. The tangential frequency ω is regarded as a parameter and is usually implied for simplicity of notations. Assume $\Omega_j \neq 0,\ \forall j = 1, 2, \ldots m$, which usually depend on ω. P is a small perturbation. If $P = 0$, then Hamiltonian H_+ (H_-) becomes a normal form and has a parameterized family of elliptic (hyperbolic) lower dimensional invariant tori $\mathcal{T}_\omega = T^n \times \{0\} \times \{0\} \times \{0\}$ with frequencies ω.

Melnikov [12, 13] concluded that if P is sufficiently small, for most of the frequency parameters ω, the invariant tori $\mathcal{T}_\omega$ for Hamiltonian H_+ can persist under the following non-resonance conditions:

$$\langle \omega, k \rangle + \Omega_j(\omega) \neq 0, \quad \forall k \in \mathbb{Z}^n,\ j = 1, 2, \cdots m, \tag{1.2}$$

$$\langle \omega, k \rangle + \Omega_i(\omega) + \Omega_j(\omega) \neq 0, \quad \forall k \in \mathbb{Z}^n,\ i, j = 1, 2, \cdots m, \tag{1.3}$$

$$\langle \omega, k \rangle + \Omega_i(\omega) - \Omega_j(\omega) \neq 0, \quad \forall k \in \mathbb{Z}^n,\ |k| + |i - j| \neq 0, \tag{1.4}$$

where (1.2) is called the first Melnikov condition, while (1.3) and (1.4) are called the second Melnikov condition. Later the result is improved by Pöschel and Bourgain [3, 17].

As to hyperbolic invariant tori for Hamiltonian H_-, there are many well known KAM theorems [5, 7, 11, 15], which are essentially some extension of Lagrangian invariant tori. Actually, hyperbolic case is much simpler than elliptic case since there is no problem of Melnikov conditions.

Recently, Xu and Lu [24] developed some new KAM techniques to prove two formal KAM theorems, which can be used to prove various kinds of KAM theorems for Lagrangian tori and elliptic lower dimensional tori. Note that the frequency considered in [24] is Diophantine. By motivation of [24], in this paper we want to give a formal KAM theorem for hyperbolic invariant tori under Bruno-Rüssmann non-resonance. By this formal KAM theorem, many previous results can be direct corollaries.

2 Main Result

For $s, r > 0$, let $T_s = \left\{x \in \mathbb{C}^n/2\pi\mathbb{Z}^n \mid |\mathrm{Im} x| \le s\right\}$ and

$$D_{s,r} = \left\{w \in \mathbb{C}^n/2\pi\mathbb{Z}^n \times \mathbb{C}^n \times \mathbb{C}^m \times \mathbb{C}^m : |\mathrm{Im} x| \le s, |y|_1 \le r^2, |u|_2 \le r, |v|_2 \le r\right\},$$

where $|\cdot|$ is the sup-norm, $|\cdot|_1$ is the l^1-norm, and $|\cdot|_2$ indicates the Euclidean norm. Let $U \subset \mathbb{R}^n$ be a domain and $\ell \geq 0$ be an integer.

Consider a parameterized Hamiltonian

$$H(\xi; w) = \langle \omega(\xi), y \rangle + \langle \Omega u, v \rangle + P(\xi; w), \tag{2.1}$$

where $w = (x, y, u, v)$ is the phase variable and ξ is a parameter. It is easy to see that $(u, v) = (0, 0)$ is a hyperbolic equilibrium for Hamiltonian H if $P = 0$. Here we should note that under the symplectic mapping, $\frac{u-v}{\sqrt{2}} = \tilde{u}$, $\frac{u+v}{\sqrt{2}} = \tilde{v}$, $\langle \Omega \tilde{u}, \tilde{v} \rangle = \frac{1}{2} \sum_{j=1}^{m} \Omega_j (u_j^2 - v_j^2)$. So we use the normal form in (2.1) for convenience.

Assume that $H(\xi; w)$ is analytic in w on $D_{s,r}$ and C^ℓ-smooth in ξ on U. Then $P(\xi; w)$ can be expanded as Fourier series with respect to x with

$$P(\xi; w) = \sum_{k \in \mathbb{Z}^n} P_k(\xi; \bar{w}) e^{\sqrt{-1}\langle k, x \rangle},$$

where $P_k(\xi; \bar{w}) = \sum_{i \in \mathbb{Z}_+^n, j, l \in \mathbb{Z}_+^m} P_{ijlk}(\xi) y^i u^j v^l$, where $\mathbb{Z}_+^n$ is composed of all the integer vectors with nonnegative components, and $\mathbb{Z}_+^m$ has the same meaning.

Denote by $C^{\ell;a}(U \times D_{s,r})$ the set which consists of functions that are analytic in w on $D_{s,r}$ and C^ℓ-smooth in ξ on U. For $P \in C^{\ell;a}(U \times D_{s,r})$, we define

$$\|P\|_{U \times D_{s,r}} = \sum_k \|P_k\|_{U;r} e^{|k|s},$$

where

$$\|P_k\|_{U;r} = \sup_{|y|_1 \leq r^2, |z|_2 \leq r, |\bar{z}|_2 \leq r} \left| \sum_{i \in \mathbb{Z}_+^n, j, l \in \mathbb{Z}_+^m} \|P_{ijlk}\|_{\alpha, C^\ell(U)} y^i u^j v^l \right|,$$

with the weighted norm

$$\|P_{ijlk}\|_{\alpha, C^\ell(U)} = \max_{|\beta| \leq \ell} \alpha^{|\beta|} \max_{\xi \in U} \left| \frac{\partial^\beta P_{ijlk}(\xi)}{\partial^\beta \xi} \right|,$$

where $\beta \in \mathbb{Z}_+^n$ and α is a constant in (2.4).

2.1 Bruno-Rüssmann Condition

Let $\Xi : [0, +\infty) \to [1, +\infty)$ be a nondecreasing unbounded function. Ξ is called an approximating function if

$$\Xi(0) = 1, \quad \frac{\log(\Xi(t))}{t} \to 0, \quad 0 \leq t \to \infty, \tag{2.2}$$

and

$$\int^{+\infty} t^{-2}\log(\Xi(t))\,dt < \infty. \tag{2.3}$$

Moreover, assume that the approximation function $\Xi(t)$ is sufficiently increasing, which is absolutely continuous and satisfies the condition (5.3) in the Appendix.

If

$$|\langle k, \omega\rangle| \geq \frac{\alpha}{\Xi(|k|)}, \quad 0 \neq k \in \mathbb{Z}^n, \tag{2.4}$$

where $0 < \alpha \leq 1$, we call ω satisfies Bruno-Rüssmann condition.

Theorem 2.1 *(The formal KAM theorem) Let $H \in C^{\ell;a}(U \times D_{s,r})$ be given in (2.1). Then for $0 < \sigma \leq s/2$, there exists a sufficiently small $\gamma > 0$, such that if*

$$\|P\|_{U\times D_{s,r}} \leq \epsilon = \alpha\gamma r^2, \tag{2.5}$$

there exist a $C^\ell(U)$-smooth family of parameterized symplectic mappings $\{\Psi(\xi;\cdot)\}_{\xi\in U}$ and a family of Hamiltonians $\{H_*(\xi;\cdot)\}_{\xi\in U}$ with the following conclusions holding true:

(1) $\Psi_* \in C^{\ell;a}(U \times D_{s/2,r/2})$ *with*

$$\|W(\Psi_* - id)\|_{U\times D_{s/2,r/2}} \leq c\Delta(\sigma)\gamma,$$

where $W = diag(\sigma^{-1}Id, r^{-2}Id, r^{-1}Id, r^{-1}Id)$, and $\Delta(\sigma)$ is as shown in (5.2).

(2)

$$H_*(\xi; w) = N_*(\xi; w) + P_*(\xi; w), \tag{2.6}$$

where $N_*(\xi; w) = \langle\omega_*(\xi), y\rangle + \langle\Omega u, v\rangle + \langle Q_*(\xi; x)z, z\rangle$ with $z = (u, v)^T$, and

$$P_*(w) = \sum_{2|i|+|j|+|l|>2} P_{*\beta}(x)\bar{w}^\beta, \quad \bar{w}^\beta = y^i u^j v^l.$$

Furthermore,

$$\|\omega_* - \omega\|_{\mathcal{C}^\ell(U)} \leq 2\alpha\gamma, \quad \|Q_*\|_{\mathcal{C}^\ell(U)\times T_{s/2}} \leq c\Delta(\sigma)\gamma. \tag{2.7}$$

(3) *If for some $\xi \in U$, $\omega_*(\xi)$ satisfies (2.4), then*

$$H \circ \Psi_*(\xi; w) = H_*(\xi; w),$$

therefore, $H(\xi;\cdot)$ has an invariant torus $\Psi_*(\xi; T^n \times \{0\} \times \{0\} \times \{0\})$ with frequencies $\omega_*(\xi)$.

Remark 2.1 Note that in Theorem 2.1 we use the Bruno-Rüssmann condition, which is a little weaker than the Diophantine condition in [24]. Moreover, we can have a similar result for elliptic lower dimensional tori. For simplicity we do not mention elliptic case in this paper.

3 Applications of Theorem 2.1

In this section we give some applications of Theorem 2.1 in two non-degenerate cases and delay the proof to the next section.

(1) *Bruno non-degenerate case* Consider a real analytic Hamiltonian

$$H(q,p,u,v) = h(p) + \langle \Omega u, v\rangle + f(q,p,u,v), \tag{3.1}$$

where $\Omega = \mathrm{diag}(\Omega_1, \cdots \Omega_m)$ with $\Omega_j \neq 0$, for $\forall j = 1, 2, \cdots m$ and f is a sufficiently small perturbation. The phase space is $T^n \times D \times \mathbb{R}^m \times \mathbb{R}^m$, where $D \subset \mathbb{R}^n$ is an open domain.

By introducing parameters, we consider an equivalent system. Let $q = x$, $p = y + \xi$, $w = (x, y, u, v)$, then

$$\begin{aligned} H(q,p,u,v) &= h(y+\xi) + \big\langle \Omega u, v\big\rangle + f(x, \xi+y, u, v) \\ &= e + \big\langle \omega(\xi), y\big\rangle + \big\langle \Omega u, v\big\rangle + P(\xi; w), \end{aligned} \tag{3.2}$$

where $e = h(\xi)$ is an energy constant, which is usually ignored, $\omega(\xi) = h_p(\xi)$, and $P(\xi; w) = O(y^2) + f(\xi + y; x, y, u, v)$, where $O(y^2) = h(\xi + y) - h(\xi) - \big\langle \omega(\xi), y\big\rangle$.

Consider the parameterized Hamiltonian (3.2), which is real analytic in w on $D_{s,r}$ and C^ℓ-smooth in ξ on U, where $U = \{\xi \in D \mid \mathrm{dist}(x, \partial D) \geq \delta_0 > 0\}$. Suppose the Bruno non-degeneracy condition holds:

$$\mathrm{rank}(\partial_\xi \omega) = n - 1, \quad \mathrm{rank}(\partial_\xi \omega^T, \omega^T) = n, \quad \forall \xi \in U. \tag{3.3}$$

Let

$$|f(q,p,u,v)| \leq \varepsilon, \ \forall q \in T_s, \ \ p \in D, \ |u| \leq \delta, \ |v| \leq \delta.$$

Let $r = \varepsilon^{\frac{1}{4}} \leq \min\{\delta_0, \delta\}$. Then

$$\|P\|_{U \times D_{s,r}} \leq \varepsilon + cr^4 \leq c\varepsilon = \epsilon = \alpha\gamma r^2,$$

where $\gamma = \frac{c\varepsilon^{\frac{1}{2}}}{\alpha}$. If ε is sufficiently small, Theorem 2.1 holds for Hamiltonian (3.2).

Obviously, γ is sufficiently small if ε is sufficiently small. By measure estimate it follows that for most of $\xi \in U$, $\omega(\xi)$ satisfies (2.4).

Moreover, $\omega_*(\xi)$ is a small perturbation of ω. Since $\omega(\xi)$ is Bruno non-degenerate and ω_* is a small perturbation of ω, by measure estimate as in [17, 24], we can prove that for most of ξ in the sense of Lebesgue measure, $\omega_*(\xi)$ satisfies (2.4). By Theorem 2.1, for $\xi \in U$ such that $\omega_*(\xi)$ satisfies (2.4), then the original Hamiltonian

$$H(\xi; w) = \langle \omega(\xi), y \rangle + \langle \Omega u, v \rangle + P(\xi; w), \quad \xi \in U$$

can be normalized to

$$H_*(\xi; w) = \langle \omega_*(\xi), y \rangle + \langle \Omega_* u, v \rangle + P_*(\xi; w), \quad \xi \in U,$$

and then it admits a lower dimensional invariant torus with frequencies $\omega_*(\xi)$. However, in this paper we are interested in the persistence of an invariant torus of the unperturbed system with frequency $\omega_0 = \omega(\xi_0)$ ($\xi_0 \in U$). If ω_0 satisfies (2.4), since $\omega(\xi)$ is Bruno non-degenerate in the sense of (3.3), in the same way as in [24] (Here we refer to Proposition 1 in [24] for details), there exist a $\xi \in U$ and a small constant $\lambda = O(\varepsilon)$ such that $\omega_*(\xi) = (1+\lambda)\omega_0$. By Theorem 2.1, Hamiltonian H has an invariant torus with the frequency $\omega_*(\xi)$, which is a small dilation of ω_0.

(2) *Rüssmann non-degenerate case* In this case we can also obtain many invariant tori by standard KAM method if f is sufficiently small, but we cannot get more information about their frequencies. Here we are concerned about the persistence of KAM tori with prescribed frequencies. Consider the Hamiltonian H in (3.1) with

$$h(y) = \langle \omega_0, y \rangle + y_1^{2l_1} + \cdots + y_n^{2l_n}, \quad |y| \le 2\delta_0, \quad l_1, l_2, \cdots l_n \ge 2.$$

Then $\omega(\xi) = \omega_0 + (2l_1\xi_1^{2l_1-1}, \cdots 2l_n\xi_n^{2l_n-1})$. Assume that ω_0 satisfies (2.4). Obviously, $\deg(\omega, U, \omega_0) \ne 0$, where $U = \{\xi \in \mathbb{R}^n \mid |\xi| \le \delta_0\}$. By Theorem 2.1, if f is sufficiently small, $\deg(\omega_*, U, \omega_0) \ne 0$. Then there exists $\xi_* \in U$, such that $\omega_*(\xi_*) = \omega_0$ and so $H(\xi_*; \cdot)$ has a hyperbolic lower dimensional invariant torus with frequencies ω_0.

4 Proof of Theorem 2.1

In this section we are going to prove Theorem 2.1. Our KAM iteration is divided into several parts. Let

$$\Gamma(\sigma) = \sup_{t \ge 0} (1+t)^{\ell+2} \Xi^{\ell+2}(t) e^{-\sigma t}.$$

By the property of approximation functions, $\Gamma(\sigma)$ is well defined.

4.1 KAM Step and Iteration Lemma

Our KAM step is summarized in the following iteration lemma.

Lemma 4.1 (Iteration Lemma) *Consider* $H(\xi; w) = N(\xi; w) + P(\xi; w)$, *where*

$$N(\xi; w) = \big\langle \omega(\xi), y \big\rangle + \big\langle \Omega u, v \big\rangle + \big\langle Q(\xi; x) z, z \big\rangle$$

is a normal form, with $z = (u, v)^T$, $Q(\xi; x)$ *is a small* $2m$*-order symmetric matrix, and* P *is a perturbation.*

Let $H \in C^{\ell;a}(U \times D_{s,r})$, *and* $\|Q\|_{U \times T_s} \ll 1$. *Suppose*

$$\|P\|_{U \times D_{s,r}} \leq \epsilon = \alpha r^2 E.$$

Let $r_+ = \eta r$, $s_+ = s - 4\sigma$. *If* $\epsilon > 0$ *is sufficiently small, then the following results hold true:*

(1) *There exists a parameterized family of symplectic mappings* $\{\Phi(\xi; \cdot),\ \xi \in U\}$, *such that* $\Phi \in C^{\ell;a}(U \times D_{s_+,r_+})$ *with*

$$\Phi(\xi; \cdot) : D_{s_+,r_+} \to D_{s,r}.$$

Moreover,

$$||W(\Phi - id)||_{U \times D_{s_+,r_+}} \leq c\Gamma E$$

and

$$||W(\mathcal{D}\Phi - Id)W^{-1}||_{U \times D_{s_+,r_+}} \leq c\Gamma E,$$

where $W = diag(\sigma^{-1} Id, r^{-2} Id, r^{-1} Id, r^{-1} Id)$ *and* $\mathcal{D}$ *denotes the differential operator with respect to* w.

(2) *There exists a Hamiltonian* $H_+ \in C^{\ell;a}(U \times D_{s_+,r_+})$ *with*

$$H_+(\xi; w) = N_+(\xi; w) + P_+(\xi; w),$$

where $N_+(\xi; w) = \langle \omega_+(\xi), y \rangle + \langle \Omega u, v \rangle + \big\langle Q_+(\xi; x) z, z \big\rangle$, *and* $\omega_+ = \omega + \hat{\omega}$. *Moreover,*

$$\|\hat{\omega}\| \leq \frac{\epsilon}{r^2}, \quad \|Q_+ - Q\|_{U \times T_s} \leq c\Gamma E. \tag{4.1}$$

Furthermore, P_+ *satisfies*

$$\|P_+\|_{U \times D_{s-4\sigma,\eta r}} \leq c\Gamma E \epsilon + c e^{-K\sigma} \epsilon + c\eta^3 \epsilon.$$

(3) *Set*

$$R_\alpha^K = \left\{\omega \in \mathbb{R}^n \mid |\langle k, \omega\rangle| \geq \frac{\alpha}{\Xi(|k|)},\ 0 < |k| \leq K\right\}$$

and

$$\tilde{U} = \{\xi \in U \mid \omega(\xi) \in R_\alpha^K\}. \tag{4.2}$$

Then,

$$H \circ \Phi(\xi; w) = H_+(\xi; w) = N_+(\xi; w) + P_+(\xi; w),\ \forall\, \xi \in \tilde{U}.$$

Moreover, define

$$\tilde{U}_+ = \left\{\xi \in U \mid \omega_+(\xi) \in R_{\alpha_+}^{K_+}\right\}, \tag{4.3}$$

where $K_+ > K$. *If* $2K\,\Xi(K)\epsilon \leq (\alpha_+ - \alpha)r^2$, *then* $\tilde{U}_+ \subset \tilde{U}$.

4.1.1 Proof of Iteration Lemma

1. *Truncation* Let

$$P = \sum_{i,j,l} P_{ijl}(\xi; x) y^i u^j v^l.$$

Make a truncation for the perturbation P and let

$$R = P_{000}(\xi; x) + \left\langle P_{100}(\xi; x), y\right\rangle + \left\langle P_{010}(\xi; x), u\right\rangle + \left\langle P_{001}(\xi; x), v\right\rangle$$

and

$$\left\langle \hat{Q}_1(\xi; x)z, z\right\rangle = \left\langle P_{020}(x)u, u\right\rangle + \left\langle P_{011}(x)u, v\right\rangle + \left\langle P_{002}(x)v, v\right\rangle,$$

here and below ξ is implied without confusion.

Let

$$R^K = P_{000}^K(x) + \left\langle P_{100}^K(x), y\right\rangle + \left\langle P_{010}^K(x), u\right\rangle + \left\langle P_{001}^K(x), v\right\rangle,$$

where

$$P_{ijl}^K(x) = \sum_{k \in \mathbb{Z}^n, |k| \leq K} P_{ijlk} e^{\mathrm{i}\langle k, x\rangle},\ \ \mathrm{i} = \sqrt{-1}.$$

Since R is composed of the zero-order terms and the one-order terms of P, by Cauchy's estimate we have $\|R\|_{U\times D_{s,r}} \leq 4\epsilon$. Then we truncate the Fourier series of R at order K to obtain R^K. By the definition of the norm, we have

$$\|R^K\|_{U\times D_{s,r}} \leq 4\epsilon,$$

and

$$\begin{aligned}\|R - R^K\|_{U\times D_{s-\sigma,r}} &\leq \sum_{|k|>K} \|R_k\| e^{|k|(s-\sigma)} \\ &\leq e^{-K\sigma} \sum_{|k|>K} \|R_k\| e^{|k|s} \leq 4e^{-K\sigma}\epsilon. \end{aligned} \tag{4.4}$$

2. *Construction of symplectic transformations* The symplectic mapping Φ is the flow X_F^t at 1-time, where F will be decided later. Let

$$F = F_{000}(x) + \langle F_{100}(x), y\rangle + \langle F_{010}(x), u\rangle + \langle F_{001}(x), v\rangle.$$

Let $G = (F_{010}, F_{001})^T$ and J be the standard $2m$-th symplectic matrix. Let $H = N + R + (P - R)$, it follows that

$$\begin{aligned} N\circ\Phi &= N + \{N, F\} + \int_0^1 \{(1-t)\{N, F\}, F\}\circ X_F^t\, dt, \\ R\circ\Phi &= R + \int_0^1 \{R, F\}\circ X_F^t\, dt, \\ \langle \hat{Q}_1(\xi; x)z, z\rangle \circ\Phi &= \langle \hat{Q}_1(\xi; x)z, z\rangle + \int_0^1 \{\langle \hat{Q}_1(\xi; x)z, z\rangle, F\}\circ X_F^t\, dt, \end{aligned}$$

where $\{\cdot, \cdot\}$ denotes the Poisson bracket.

Then

$$\{N, F\} = \langle\langle Q_x, F_{100}\rangle z, z\rangle - \langle \omega, F_x\rangle + \langle \Omega v, F_{001}\rangle - \langle \Omega u, F_{010}\rangle + \langle JG, 2Qz\rangle.$$

It follows that

$$\begin{aligned} H\circ X_F^1 = N &- \partial_\omega F - \langle \Omega u, F_{010}\rangle + \langle \Omega v, F_{001}\rangle \\ &+ R^K + \langle \hat{Q}_1 z, z\rangle + \langle\langle Q_x, F_{100}\rangle z, z\rangle + \langle JG, 2Qz\rangle + P_+, \end{aligned}$$

where $\partial_\omega F \stackrel{\text{def}}{=} \langle \omega, F_x\rangle$ and

$$P_+ = \left(R - R^K\right) + \left(P - R - \langle \hat{Q}_1(\xi; x)z, z\rangle\right)\circ X_F^1 + \tilde{P}, \tag{4.5}$$

where

$$\tilde{P} = \int_0^1 \{(1-t)\{N, F\} + R + \langle \hat{Q}_1(\xi; x)z, z\rangle, F\} \circ X_F^t \, dt.$$

Then, we need to solve the equations:

$$\begin{cases} \partial_\omega F_{000} = P_{000}^K - [P_{000}] \\ \partial_\omega F_{100} = P_{100}^K - [P_{100}] \\ \partial_\omega G - MG - 2Q(\xi; x)JG = g \end{cases}, \tag{4.6}$$

where $M = \mathrm{diag}(-\Omega, \Omega)$, $g = (P_{010}, P_{001})^T$, $[\,\cdot\,]$ denotes the mean value over T^n.

3. *Extension of small divisors* Take a $C^\infty(\mathbb{R})$-smooth function $\psi(t)$ such that

$$\psi(t) = \begin{cases} 0, & |t| \le \frac{1}{2}, \\ 1, & |t| \ge 1. \end{cases}$$

For $h > 0$, set $\psi_h(t) = \psi(\frac{t}{h})$. Then $\psi_h(t) \in C^\infty(\mathbb{R})$ with the estimate:

$$|\frac{d^l}{dt^l}\psi_h(t)| \le \frac{c_l}{h^l}, \quad \forall t \in \mathbb{R}, \ \ \forall l \ge 1,$$

where c_l is a constant depending on l.

Set

$$h = \frac{\alpha}{\Xi(|k|)}, \quad t_k(\xi) = \langle k, \omega(\xi)\rangle, \quad f_k(\xi) = \frac{\psi_h(t_k(\xi))}{\mathrm{i}\langle k, \omega(\xi)\rangle}.$$

Recall the definition of $\tilde{U}$, it follows easily that for $\xi \in \tilde{U}$, $f_k(\xi) = \frac{1}{\mathrm{i}\langle k,\omega(\xi)\rangle}$. Here, we observe that even though $\tilde{U} = \emptyset$, the extension $f_k(\xi)$ is still well defined on U. Then $f_k(\xi) \in C^\ell(U)$, which satisfies

$$\Big|\frac{\partial^\beta f_k}{\partial \xi^\beta}(\xi)\Big| \le ch^{-|\beta|-1}|k|^{|\beta|}, \quad \xi \in U, \ \ \forall |\beta| \le \ell.$$

Set

$$F_{\flat k}(\xi; \bar{w}) = f_k(\xi)(P_{\flat k} - [P_{\flat k}]) = \frac{\psi_h(t_k(\xi))}{\mathrm{i}\langle k, \omega(\xi)\rangle}(P_{\flat k} - [P_{\flat k}]),$$

where the subscript $\flat = 000, 100, \ 0 < |k| \le K$. Then we extend $F_{\flat k}(\xi; \bar{w})$ for ξ from $\tilde{U}$ to the whole set U.

4. *Solving the homological equations* The first two equations for (4.6) are standard. By the extension of small divisors, in the same way as in [24, 25], we have F_{000} and F_{100} such that

$$\begin{aligned} \|F_{000}\|_{U\times D_{s-2\sigma,r}} &\leq c\alpha^{-1}\sum_k \Xi^{\ell+1}(|k|)|k|^\ell e^{-|k|\sigma}\|P_{000k}\|e^{|k|s} \\ &\leq c\alpha^{-1}\sup_{t\geq 0} t^\ell\,\Xi^{\ell+1}(t)e^{-\sigma t}\|P_{000}\|_{U\times D_{s-\sigma,r}} \\ &\leq c\alpha^{-1}\epsilon\sup_{t\geq 0}(1+t)^\ell\,\Xi^{\ell+1}(t)e^{-\sigma t} \\ &\leq c\alpha^{-1}\Gamma_{\ell+1}\epsilon \end{aligned}$$

and

$$\|F_{100}\|_{U\times D_{s-2\sigma,r/2}} \leq c\alpha^{-1}r^{-2}\Gamma_{\ell+1}\epsilon,$$

where $\Gamma_{\ell+1}(\sigma)=\sup_{t\geq 0}(1+t)^{\ell+1}\,\Xi^{\ell+1}(t)e^{-\sigma t}$. Moreover, for $\xi\in\tilde U$, where $\tilde U$ is given in (4.2), F_{000} and F_{100} are solutions of the equations.

For $G=(F_{010},F_{001})^T$, we apply Lemma 5.1 with $Q_0=M,\ \hat Q=2QJ$ to have G satisfying

$$\|F_{010}\|_{U\times D_{s-2\sigma,r/2}} \leq cr^{-1}\epsilon,\quad \|F_{001}\|_{U\times D_{s-2\sigma,r/2}} \leq cr^{-1}\epsilon. \tag{4.7}$$

Therefore,

$$\|F\|_{U\times D_{s-2\sigma,r/2}} \leq c\alpha^{-1}\Gamma_{\ell+1}\epsilon, \tag{4.8}$$

and

$$\begin{aligned} \|\partial_x F\|_{U\times D_{s-2\sigma,r/2}} &\leq \sum_k |k|\cdot\|F_k\|e^{|k|(s-\sigma)} \\ &\leq c\alpha^{-1}\sup_{t\geq 0} t\cdot(1+t)^{\ell+1}\,\Xi^{\ell+1}(t)e^{-\sigma t}\epsilon \leq c\alpha^{-1}\Gamma\epsilon, \end{aligned}$$

where $\Gamma(\sigma)=\sup_{t\geq 0}(1+t)^{\ell+2}\,\Xi^{\ell+2}(t)e^{-\sigma t}$.

5. *Estimates of the symplectic mapping* Write the symplectic mapping as

$$\Phi(\xi;w)=X_F^1=(\tilde a(\xi;x),\tilde b(\xi;w),\tilde d(\xi;x,u),\tilde e(\xi;x,v)).$$

By the construction of F and Φ, it follows that $\tilde b$ is affine in y,u,v, $\tilde d$ and $\tilde e$ are the translations of u,v, respectively. Moreover, by the estimates of F, we have

$$\begin{aligned} &\|\tilde a-id\|_{U\times D_{s-2\sigma,r/2}} \leq c\Gamma_{\ell+1}E,\quad \|\tilde b-id\|_{U\times D_{s-2\sigma,r/4}} \leq c\alpha^{-1}\Gamma\epsilon, \\ &\|\tilde d-id\|_{U\times D_{s-2\sigma,r/4}} \leq cr^{-1}\epsilon,\quad \|\tilde e-id\|_{U\times D_{s-2\sigma,r/4}} \leq cr^{-1}\epsilon. \end{aligned}$$

And

$$\mathcal{D}\Phi = \begin{pmatrix} \tilde{a}_x & 0 & 0 & 0 \\ \tilde{b}_x & \tilde{b}_y & \tilde{b}_u & \tilde{b}_v \\ \tilde{d}_x & 0 & Id & 0 \\ \tilde{e}_x & 0 & 0 & Id \end{pmatrix}.$$

By Lemma 5.4, we have

$$\frac{\Gamma_{\ell+1}(\sigma)}{\sigma} \le \frac{\Gamma_{\ell+2}(\sigma)}{2(\ell+1)} = \frac{\Gamma(\sigma)}{2(\ell+1)}.$$

Assume that

$$c\Gamma E \le \sigma < \eta^2 \le \frac{1}{4}. \tag{4.9}$$

Then the symplectic mapping $\Phi : D_{s-4\sigma,\eta r} \to D_{s-3\sigma,2\eta r}$, with estimates

$$\|W(\Phi - id)\|_{U\times D_{s-4\sigma,\eta r}} \le c\Gamma E \tag{4.10}$$

and

$$\|W(\mathcal{D}\Phi - Id)W^{-1}\|_{U\times D_{s-4\sigma,\eta r}} \le c\Gamma E, \tag{4.11}$$

where the weight matrix $W = \mathrm{diag}(\sigma^{-1}Id, r^{-2}Id, r^{-1}Id, r^{-1}Id)$.

6. *Estimates of the new error terms* Recall that $N = \langle\omega(\xi), y\rangle + \langle\Omega u, v\rangle + \langle Qz, z\rangle$. Let $\hat{Q}_2 = Q_x \cdot F_{100}$. Then it follows that $H \circ \Phi = N_+ + P_+$, where $N_+ = N + \hat{N}$, with

$$\hat{\omega}(\xi) = [P_{100}], \quad \hat{Q}(\xi; x) = \hat{Q}_1 + \hat{Q}_2.$$

By standard estimate, we get

$$\|\hat{\omega}\|_U \le \alpha E, \quad \|\hat{Q}\|_{U\times T_{s-4\sigma}} \le c\Gamma E.$$

Also note that P_+ is given in (4.5). By (4.4), we have

$$\left\| R - R^K \right\|_{U\times D_{s-\sigma,r}} \le 4e^{-K\sigma}\epsilon.$$

By Taylor's formula with remainder and Cauchy's estimate, we have

$$\left\| P - R - \left\langle \hat{Q}_1(\xi; x)z, z \right\rangle \right\|_{U\times D_{s,2\eta r}} \le c\eta^3\epsilon,$$

Combining with the estimates of F, R and $\hat{Q}_1$, we get

$$\|P_+\|_{U\times D_{s-4\sigma,\eta r}} \le c\Gamma E\epsilon + ce^{-K\sigma}\epsilon + c\eta^3\epsilon, \tag{4.12}$$

where c is a constant independent of KAM steps.

In the same way as [17], we will choose iteration parameters such that the KAM step can iterate. The idea is as follows. By some suitable choices of $K,\ \eta,\ \epsilon_+,\ r_+$ as

$$e^{-K\sigma} \sim \Gamma E,\quad \eta^3 \sim \Gamma E,\quad \epsilon_+ \sim \Gamma E\epsilon,\quad r_+ \sim \eta r,$$

we can have

$$\|P_+\|_{U\times D_{s-4\sigma,\eta r}} \le c\Gamma E\epsilon + ce^{-K\sigma}\epsilon + c\eta^3\epsilon \le \epsilon_+ = \alpha_+ r_+^2 E_+.$$

Moreover, it follows that

$$\frac{\epsilon_+}{r_+^2} \sim \frac{\Gamma E\epsilon}{\eta^2 r^2} \sim \frac{\Gamma E^2}{\eta^2} \sim \Gamma^{\frac{1}{3}} E^{\frac{4}{3}} \sim E_+.$$

In KAM step, E will decrease rapidly and it will be so small that $\Gamma^{\frac{1}{3}} E^{\frac{4}{3}}$ becomes much smaller.

4.1.2 KAM Iteration

Recall that

$$\Gamma(\sigma) = \sup_{t\ge 0}(1+t)^{\ell+2}\Xi^{\ell+2}(t)e^{-\sigma t}.$$

By Lemma 5.3, for $\sigma = s/2$, there exists a sequence $\sigma_0 \ge \sigma_1 \ge \sigma_2 \ge \cdots > 0$, such that $\sigma_0 + \sigma_1 + \sigma_2 + \cdots = \sigma$ and

$$\Delta(\sigma) = \prod_{j=0}^{\infty} \Gamma(\sigma_j)^{\kappa_j},\quad \kappa_j = \frac{\kappa - 1}{\kappa^{j+1}} \ \text{ with } \ \kappa = \frac{4}{3}.$$

At the initial step, let $H_0 = H$ and set $s_0 = s$, $r_0 = r$, $E_0 = \gamma$. For $i \ge 0$, define

$$\alpha_{i+1} = (1 - \frac{1}{2^{i+3}})\alpha,\quad \Theta_i = \prod_{j=0}^{i-1}\big(a2^j\Gamma(\sigma_j)\big)^{\kappa_j},\quad E_i = (\Theta_i E_0)^{\kappa^i},$$

where $\Theta_0 = 1$, $a = (2c)^3$, and c is the constant in the estimate of P_+. Let $\epsilon_i = \alpha_i r_i^2 E_i$. Moreover, define K_i and η_i by

$$e^{-K_i\sigma_i} = 2^{i+4}\Gamma(\sigma_i)E_i,\quad \eta_i^3 = 2^i\Gamma(\sigma_i)E_i.$$

Here the multipliers 2^{i+4} and 2^i are required for small divisor conditions. Define $r_{i+1} = \eta_i r_i,\ s_{i+1} = s_i - 4\sigma_i$.

Note that

$$\Delta(\sigma) = \prod_{j=0}^{\infty} \Gamma(\sigma_j)^{\kappa_j}, \ \sum_{j=0}^{\infty} \kappa_j = 1, \ \sum_{j=0}^{\infty} j\kappa_j = \frac{1}{\kappa - 1}.$$

Then we get

$$\Theta_i \to 8a\Delta(\sigma), \quad i \to \infty.$$

Note that $\Gamma(\sigma_i) \leq \Gamma(\sigma_j)$ for all $j \geq i$. It follows that

$$a2^i\Gamma(\sigma_i) = \prod_{j=i}^{\infty} \left(a2^i\Gamma(\sigma_i)\right)^{\kappa_j\kappa^i} \leq \left(\prod_{j=i}^{\infty} \left(a2^j\Gamma(\sigma_j)\right)^{\kappa_j}\right)^{\kappa^i}$$

and then

$$a2^i\Gamma(\sigma_i) \cdot E_i \leq \left(\left(\prod_{j=0}^{\infty} a2^j\Gamma(\sigma_j)\right)^{\kappa_j} E_0\right)^{\kappa^i} = (a\Delta(\sigma)E_0)^{\kappa^i}. \tag{4.13}$$

Denote by $D_i = D_{s_i,r_i}$. By Lemma 4.1, there exists a sequence of Hamiltonians $\{H_i(\xi; w),\ \xi \in U,\ w \in D_i\}$ such that $H_i \in C^{\ell;a}(U \times D_i)$ and $H_i = N_i + P_i$, where

$$N_i = \langle \omega_i, y\rangle + \langle \Omega u, v\rangle + \langle Q_i(x)z, z\rangle,$$

and P_i satisfies that

$$\|P_i\|_{U\times D_i} \leq \epsilon_i = \alpha_i r_i^2 E_i. \tag{4.14}$$

Moreover, there exists a sequence of parameterized symplectic transformations $\{\Phi_i(\xi; w),\ \xi \in U,\ w \in D_{i+1}\}$, such that for each $\xi \in U$, $\Phi_i(\xi; w) : D_{i+1} \to D_i$. Moreover, and $\Phi_i \in C^{\ell;a}(U \times D_{i+1})$ with estimates:

$$\|W_i(\Phi_i - id)\|_{U\times D_{i+1}} \leq c\Gamma(\sigma_i)E_i \tag{4.15}$$

and

$$\|W_i(\mathcal{D}\Phi_i - Id)W_i^{-1}\|_{U\times D_{i+1}} \leq c\Gamma(\sigma_i)E_i. \tag{4.16}$$

Let

$$\tilde{U}_i = \left\{\xi \in U \mid |\langle k, \omega_i(\xi)\rangle| \geq \frac{\alpha_i}{\Xi(|k|)},\ 0 < |k| \leq K_i\right\}.$$

Then for $\xi \in \tilde{U}_i$,

$$H_{i+1} = H_i \circ \Phi_i = N_{i+1} + P_{i+1}, \tag{4.17}$$

where

$$N_{i+1} = \langle \omega_{i+1}, y \rangle + \langle \Omega u, v \rangle + \langle Q_{i+1}(x)z, z \rangle,$$

and by (4.9) and (4.12),

$$\begin{aligned} \|P_{i+1}\|_{U \times D_{i+1}} &\leq c\Gamma(\sigma_i)E_i\epsilon_i + ce^{-K_i\sigma_i}\epsilon_i + c\eta_i^3\epsilon_i \\ &\leq c \cdot 2^i \Gamma(\sigma_i)E_i\epsilon_i. \end{aligned}$$

By the definitions of η_i, ϵ_i, E_i and $r_{i+1} = \eta_i r_i$, it follows that

$$\frac{c2^i\Gamma(\sigma_i)E_i\epsilon_i}{r_{i+1}^2\alpha_{i+1}} \leq \frac{2c2^i\Gamma(\sigma_i)E_i^2}{\eta_i^2} \leq \left((2c)^3 2^i\Gamma(\sigma_i)\right)^{\frac{1}{3}} E_i^{\frac{4}{3}} \leq \left(a2^i\Gamma(\sigma_i)\right)^{\frac{1}{3}} E_i^{\frac{4}{3}} \leq E_{i+1},$$

where $a = (2c)^3$. Then

$$\|P_{i+1}\|_{U \times D_{i+1}} \leq \epsilon_{i+1} = \alpha_{i+1} r_{i+1}^2 E_{i+1}. \tag{4.18}$$

In addition,

$$|\hat{\omega}_i| \leq \alpha_i E_i, \quad \|\hat{Q}_i\| \leq c\Gamma(\sigma_i)E_i, \tag{4.19}$$

where $\hat{\omega}_i = \omega_{i+1} - \omega_i$.

Let $\Psi_0 = id$, $\Psi_i = \Phi_0 \circ \Phi_1 \circ \cdots \circ \Phi_{i-1}$, $i \geq 1$. By Lemma 4.1 again, if $2K_i \Xi(K_i)\epsilon_i \leq (\alpha_{i+1} - \alpha_i)r_i^2$, $\forall i \geq 0$, we have $\tilde{U}_i \supset \tilde{U}_{i+1}$, $\forall i \geq 0$. The monotonousness of $\{\tilde{U}_i\}$ implies that for $\xi \in \tilde{U}_i$, $H_i = H \circ \Psi_i$.

Now we verify the assumption $2K_i \Xi(K_i)\epsilon_i \leq (\alpha_{i+1} - \alpha_i)r_i^2$, which is equivalent to $2^{i+4}K_i \Xi(K_i)\epsilon_i / r_i^2 \leq \alpha$. By the definition of ϵ_i, we need to prove

$$E_i \leq \frac{1}{(2^{i+4} - 4)K_i \Xi(K_i)}. \tag{4.20}$$

Recall $e^{-K_i\sigma_i} = 2^{i+4}\Gamma(\sigma_i)E_i$. By the definition of $\Gamma(\sigma_i)$, it follows that

$$\begin{aligned} \frac{1}{K_i \Xi(K_i)} &= \frac{e^{-K_i\sigma_i}}{K_i \Xi(K_i)e^{-K_i\sigma_i}} = \frac{2^{i+4}\Gamma(\sigma_i)E_i}{K_i \Xi(K_i)e^{-K_i\sigma_i}} \\ &\geq \frac{2^{i+4}\Gamma(\sigma_i)E_i}{\Gamma(\sigma_i)} = 2^{i+4}E_i, \end{aligned}$$

then

$$E_i \leq \frac{1}{2^{i+4} K_i \Xi(K_i)} \leq \frac{1}{(2^{i+4}-4) K_i \Xi(K_i)},$$

which shows (4.20).

4.1.3 Convergence

Now we consider the convergence of the KAM iteration. Note that $r_i \to 0$ as $i \to \infty$. Let $D_i \to D_* = D_{s/2,0}$ as $i \to \infty$. We first consider the convergence of $\{\Psi_i\}$. The proof is the same way as in [17]. First we have

$$\begin{aligned} \|W_0 \mathcal{D}\Psi_{i-1} W_{i-1}^{-1}\| \leq & \|W_0 \mathcal{D}\Phi_0 W_0^{-1}\| \|W_0 W_1^{-1}\| \|W_1 \mathcal{D}\Phi_1 W_1^{-1}\| \cdots \|W_{i-3} W_{i-2}^{-1}\| \\ & \|W_{i-2} \mathcal{D}\Phi_{i-2} W_{i-2}^{-1}\| \|W_{i-2} W_{i-1}^{-1}\| \\ \leq & \prod_{j=0}^{i-2} \left(1 + c\Gamma\left(\sigma_j\right) E_j\right). \end{aligned}$$

By (4.13), if $a\Delta(\sigma)E_0 < 1$, then $\prod_{j=0}^{\infty}(1+c\Gamma(\sigma_j)E_j) < \infty$. (4.15) and (4.16) imply that

$$\begin{aligned} \|W_0(\Psi_i - \Psi_{i-1})\|_{U \times D_i} &= \|W_0(\Psi_{i-1} \circ \Phi_{i-1} - \Psi_{i-1})\|_{U \times D_i} \\ &\leq \|W_0 \mathcal{D}\Psi_{i-1} W_{i-1}^{-1}\|_{U \times D_i} \cdot \|W_{i-1}(\Phi_{i-1} - id)\|_{U \times D_i} \\ &\leq c\Gamma(\sigma_{i-1}) E_{i-1}, \end{aligned} \tag{4.21}$$

and so $\{\Psi_i\}$ is convergent on D_*.

Note that Ψ_i has the same structure as Φ_i, and recall $z = (u, v)^T$. Let

$$\Psi_i(\xi; w) = (A_i(x), y + B_i(x) + C_i(x)y + D_i(x)z, z + E_i(x)).$$

Since $\{\Psi_i(\xi; w)\}$ is convergent for $x \in T_{s/2}, y = 0, z = 0$, then $\{A_i(x)\}$ and $\{B_i(x)\}$, $\{E_i(x)\}$ are convergent as $i \to \infty$ for $x \in T_{s/2}$. Below we prove that $\{C_i(x)\}$, $\{D_i(x)\}$ are also convergent on $T_{s/2}$.

Let

$$\Phi_i(\xi; w) = (a_i(x), y + b_i(x) + c_i(x)y + d_i(x)z, z + e_i(x)).$$

Then $a_i : x \in T_{s_{i+1}} \to a_i(x) \in T_{s_i}$. By the estimate for $\mathcal{D}\Phi$ in Lemma 4.1, it follows that

$$\|c_i(x)\| \leq c\Gamma(\sigma_i)E_i, \ \|d_i(x)\| \leq c\Gamma(\sigma_i)E_i r_i, \ x \in T_{s_{i+1}}.$$

Moreover,

$$\begin{pmatrix} Id_n + C_i(x) & D_i(x) \\ 0 & Id_{2m} \end{pmatrix} = \prod_{j=0}^{i-1} \begin{pmatrix} Id_n + c_j(\tilde{x}_j(x)) & d_j(\tilde{x}_j(x)) \\ 0 & Id_{2m} \end{pmatrix},$$

where Id_k indicates the k-th unit matrix and

$$\tilde{x}_j(x) = a_j \circ a_{j-1} \circ \cdots \circ a_{i-1}(x), \quad \tilde{x}_j : x \in T_{s_i} \to \tilde{x}_j(x) \in T_{s_{j+1}}.$$

Then we have

$$\|c_j(\tilde{x}_j(x))\| \le c\Gamma(\sigma_i)E_i, \; \|d_j(\tilde{x}_j(x))\| \le c\Gamma(\sigma_i)E_i r_i, \; x \in T_{s_i}.$$

Thus, as $i \to \infty$, $\{C_i(x)\}$ and $\{D_i(x)\}$ are convergent on $T_{s/2}$. So Ψ_i is actually convergent on $D_{s/2,r/2}$. Let $\Psi_* = \lim_{i\to\infty} \Psi_i$.

Note that $\omega_i = \omega_0 + \sum_{j=0}^{i-1} \hat{\omega}_j$. By (4.19), it follows $\omega_i \to \omega_*$ as $i \to \infty$, moreover,

$$|\omega_* - \omega_i| \le \sum_{j=i}^{\infty} \alpha_j E_j \le 2\alpha_i E_i.$$

In particular, noting $\omega_0 = \omega$, we have

$$|\omega_* - \omega| \le 2\alpha E_0.$$

Also note $Q_i = \sum_{j=0}^{i-1} \hat{Q}_j$. (4.19) implies $Q_i \to Q_*$ as $i \to \infty$. Recall that $E_0 = \gamma$. If γ is sufficiently small,

$$\|Q_i\|_{U\times T_{s_i}} \le \sum_{j=0}^{i-1} c\Gamma(\sigma_j)E_j \le c\Delta(\sigma)E_0 = c\Delta(\sigma)\gamma \ll 1.$$

Thus $\lim_{i\to\infty} N_i = N_*$, where

$$N_* = \langle \omega_*, y \rangle + \langle \Omega u, v \rangle + \langle Q_*(\xi; x)z, z \rangle. \tag{4.22}$$

Let $P_i \to P_*$, then $P_* \in C^{\ell;a}(U \times D_{s/2,r/2})$. By (4.14) and Cauchy's estimate we have $\partial_y P_* = 0$, $\partial_z P_* = 0$, $\partial^2_{zz} P_* = 0$ for $(y, z) = (0, 0)$. Thus,

$$P_*(\xi; w) = \sum_{2|i|+|j|+|l|>2} P_{*\beta}(\xi; x)\bar{w}^\beta, \quad \bar{w}^\beta = y^i u^j v^l. \tag{4.23}$$

Let $\tilde{U}_* = \{\xi \in \tilde{U} \mid \omega_*(\xi) \in R_\alpha^K\}$. We are going to prove that $\tilde{U}_* \subset \tilde{U}_i$ for $\forall i \geq 0$. Recall that $2^{i+4} K_i \Xi(K_i)\epsilon_i / r_i^2 \leq \alpha$. For $\xi \in \tilde{U}_*$ and $0 < |k| \leq K_i$,

$$|\langle k, \omega_i(\xi)\rangle| \geq |\langle k, \omega_*(\xi)\rangle| - |\langle k, \omega_*(\xi) - \omega_i(\xi)\rangle| \geq \frac{\alpha}{\Xi(|k|)} - \frac{2\epsilon_i}{r_i^2} K_i \geq \frac{\alpha_i}{\Xi(|k|)},$$

thus $\omega_i(\xi) \in R_{\alpha_i}^{K_i}$, $\forall i \geq 0$ and so $\tilde{U}_* \subset \tilde{U}_i$, $\forall i \geq 0$.

By (4.17), it follows that

$$H \circ \Psi_i = N_i + P_i, \ (\xi; w) \in \tilde{U}_* \times D_i.$$

Taking the limit (as $i \to \infty$) in the above equation, we get

$$H \circ \Psi_* = N_* + P_*, \ (\xi; w) \in \tilde{U}_* \times D_{s/2, r/2},$$

where N_* and P_* are given in (4.22) and (4.23). Thus, we finish the proof.

Author Contributions J.X. suggested the study of this problem and Qi Li gave the details of the proof. All authors read and approved the final manuscript.

Declarations

Conflict of interest The authors declare no competing interests.

Appendix

Lemma 5.1 *Let $\lambda_1, \lambda_2, \cdots \lambda_{2m}$ be the eigenvalues of matrix Q_0 with $|\mathrm{Re}\lambda_i| \geq \delta_0 > 0$, for any $i = 1, 2, \cdots 2m$. Set $g(x) \in \mathcal{A}$, where $\mathcal{A}$ denotes the analytic function space defined on the strip T_s, among which $T_s = \{x \in \mathbb{C}^n/2\pi\mathbb{Z}^n : |\mathrm{Im}x|_\infty \leq s\}$. There exists a sufficiently small $\epsilon_0 > 0$, such that for $\hat{Q}(x) \in \mathcal{A}$, if $||\hat{Q}||_{T_s} \leq \epsilon_0$, then the equation*

$$\langle \omega, \partial_x f(x)\rangle - (Q_0 + \hat{Q}(x)) f(x) = g(x)$$

has a unique solution $f(x) \in \mathcal{A}$, with

$$||f||_{T_s} \leq c||g||_{T_s}.$$

Proof Let $\mathcal{L}f = \langle \omega, \partial_x f(x)\rangle - Q_0 f(x), \ f \in \mathcal{A}$. Then $\mathcal{L} : \mathcal{A} \to \mathcal{A}$ is a linear operator. By assumption, the matrix Q_0 is hyperbolic, the operator $\mathcal{L}$ has a bounded inverse with $\|\mathcal{L}^{-1}\| \leq \frac{1}{\delta_0}$. By Banach fixed point theorem, it is easy to follow this lemma and we omit the details. □

Let $\bar{\Xi}$ be an approximation function satisfying (2.2) and (2.3). Let

$$\bar{\Gamma}(\sigma) = \sup_{t \geq 0} \bar{\Xi}(t) e^{-\sigma t}.$$

Define

$$\bar{\Delta}(\sigma) = \inf \prod_{j=0}^{\infty} \bar{\Gamma}(\sigma_j)^{\kappa_j}, \quad \kappa_j = \frac{\kappa - 1}{\kappa^{j+1}} \text{ with } \kappa = \frac{4}{3}, \tag{5.1}$$

where the infimum is taken for sequences $\{\sigma_j\}$ satisfying

$$\sigma_0 \geq \sigma_1 \geq \sigma_2 \geq \cdots > 0 \text{ and } \sigma_0 + \sigma_1 + \sigma_2 + \cdots \leq \sigma.$$

We first state a lemma which is proved in [17]. We also refer it to [18].

Lemma 5.2 (Lemma A.1 [17]) *For all $\sigma > 0$, the function $\bar{\Delta}(\sigma)$ is finite. More precisely, if*

$$\frac{1}{\log \kappa} \int_T^{\infty} \frac{\log\big(\bar{\Xi}(t)\big)}{t^2}\, dt \leq \sigma,$$

then

$$\bar{\Delta}(\sigma) \leq e^{(\kappa-1)\sigma T}.$$

Let Ξ be an approximation function satisfying (2.2) and (2.3). Let $\ell \geq 0$ and

$$\Gamma(\sigma) = \sup_{t \geq 0} (1+t)^{\ell+2} \Xi^{\ell+2}(t) e^{-\sigma t}.$$

Define

$$\Delta(\sigma) = \inf \prod_{j=0}^{\infty} \Gamma(\sigma_j)^{\kappa_j}, \quad \kappa_j = \frac{\kappa - 1}{\kappa^{j+1}} \text{ with } \kappa = \frac{4}{3}, \tag{5.2}$$

where the infimum is taken for sequences $\{\sigma_j\}$ satisfying

$$\sigma_0 \geq \sigma_1 \geq \sigma_2 \geq \cdots > 0 \text{ and } \sigma_0 + \sigma_1 + \sigma_2 + \cdots \leq \sigma.$$

Since Ξ is an approximation function, it is easy to check that $(1+t)^{\ell+2}\Xi^{\ell+2}(t)$ is also an approximation function. By Lemma 5.2 with $\bar{\Xi}(t) = (1+t)^{\ell+2}\Xi^{\ell+2}(t)$, it follows that for all $\sigma > 0$, $\Delta(\sigma)$ is finite.

Lemma 5.3 *The supremum in the definition of $\Gamma(\sigma)$ can be attained. Moreover, the infimum in the definition of $\Delta(\sigma)$ can also be attained. More precisely, for any $\sigma > 0$, there exists a sequence $\sigma_0^* \geq \sigma_1^* \geq \sigma_2^* \geq \cdots > 0$ such that $\sum_{i=0}^{\infty} \sigma_i^* = \sigma$ and*

$$\Delta(\sigma) = \prod_{j=0}^{\infty} \Gamma(\sigma_j^*)^{\kappa_j}.$$

Proof This lemma is actually proved in [17, 18]. However, because of small divisor conditions, our definitions of Γ and Δ are different. For the convenience of readers, we give the proof, but the idea is the same as in [17, 18].

At first, by assumption (2.2) we have $\frac{\log(\Xi(t))}{t} \to 0,\ 0 \le t \to \infty$. Then it is easy to see that the supremum $\Gamma(\sigma) = \sup_{t\ge 0}(1+t)^{\ell+2}\Xi^{\ell+2}(t)e^{-\sigma t}$ is attained and finite.

Note that

$$\Delta(\sigma) = \inf \prod_{j=0}^{\infty} \Gamma(\sigma_j)^{\kappa_j},\ \ \kappa_j = \frac{\kappa - 1}{\kappa^{j+1}} \text{ with } \kappa = \frac{4}{3}.$$

Let

$$f(\tilde{\sigma}) = \prod_{j=0}^{\infty} \Gamma(\sigma_j)^{\kappa_j},$$

where

$$\tilde{\sigma} = (\sigma_0, \sigma_1, \sigma_2, \cdots \sigma_n, \cdots) \in l^1.$$

We consider $f(\tilde{\sigma})$ as a functional on l^1.

Note that the weakly convergence in l^1 implies the pointwise convergence. Then $f(\tilde{\sigma})$ is weakly lower semi-continuous on the set:

$$A = \left\{ (\sigma_0, \sigma_1, \sigma_2, \cdots \sigma_n, \cdots) \mid \sum_{j\ge 0} \sigma_j \le \sigma, \sigma_j > 0, \forall j \ge 0 \right\}.$$

In fact, let $\tilde{\sigma}_k \rightharpoonup \tilde{\sigma}$, that is,

$$\sigma_{kj} \to \sigma_j,\ k \to \infty,\ \forall j = 0, 1, 2 \cdots.$$

Moreover,

$$\varliminf_{k\to\infty} \Gamma(\sigma_{kj}) \ge \varliminf_{k\to\infty} (1+t)^{\ell+2}\Xi^{\ell+2}(t)e^{-\sigma_{kj}t} = (1+t)^{\ell+2}\Xi^{\ell+2}(t)e^{-\sigma_j t},\ \forall t \ge 0,$$

thus,

$$\varliminf_{k\to\infty} \Gamma(\sigma_{kj}) \ge \sup_{t\ge 0}(1+t)^{\ell+2}\Xi^{\ell+2}(t)e^{-\sigma_j t} = \Gamma(\sigma_j).$$

Then

$$\varliminf_{k\to\infty} f(\tilde{\sigma}_k) \ge \prod_{j=0}^{\infty} \varliminf_{k\to\infty} \Gamma(\sigma_{kj}) \ge \prod_{j=0}^{\infty} \Gamma(\sigma_j) = f(\tilde{\sigma}).$$

Also note that if $\sigma_j \to 0$, we have $f(\tilde{\sigma}) \to +\infty$. And A is a bounded set of l^1. Then the infimum $\inf_{\tilde{\sigma}\in A} f(\tilde{\sigma})$ can be attained. Thus, there exists a sequence $\sigma_0^* \geq \sigma_1^* \geq \sigma_2^* \geq \cdots > 0$ such that $\sum_{i=0}^{\infty} \sigma_i^* \leq \sigma$ and

$$\Delta(\sigma) = f(\tilde{\sigma}_*) = \prod_{j=0}^{\infty} \Gamma\left(\sigma_j^*\right)^{\kappa_j},$$

where $\tilde{\sigma}_* = (\sigma_0^*, \sigma_1^*, \sigma_2^*, \cdots)$. Now we can prove $\sum_{i=0}^{\infty} \sigma_i^* = \sigma$ by contradiction. If not so, that is, $\sum_{i=0}^{\infty} \sigma_i^* < \sigma$, then there exists a sequence $\{\sigma_i'\}$ such that $\sigma_i^* < \sigma_i'$, $\forall i \geq 0$, and $\sum_{i=0}^{\infty} \sigma_i' = \sigma$. Obviously, $f(\tilde{\sigma}') < f(\tilde{\sigma}_*)$, which is a contradiction. Thus, $\sum_{i=0}^{\infty} \sigma_i^* = \sigma$.

In addition, the infimum is not only attainable, but also finite. This conclusion can follow from Lemma 5.2. □

If the approximation function $\Xi(t)$ is absolutely continuous and for almost every $t \geq 0$, assume that

$$\frac{d}{dt}\log\big(\Xi(t)\big) \geq \frac{1}{1+t}, \tag{5.3}$$

then $\Xi(t)$ is called a sufficiently increasing function. Without saying it directly, we assume that all approximation functions in this paper are sufficiently increasing.

Lemma 5.4 *If the approximation function $\Xi(t)$ is sufficiently increasing, then it follows*

$$\Gamma_{\ell+1}(\sigma) \leq \frac{\sigma}{2(\ell+1)}\Gamma_{\ell+2}(\sigma), \quad \ell \geq 0,$$

where $\Gamma_{\ell+1}(\sigma) = \sup_{t\geq 0}(1+t)^{\ell+1}\Xi^{\ell+1}(t)e^{-\sigma t}$.

Proof Assume that the approximation function $\Xi(t)$ is sufficiently increasing, if $\sigma(1+t) \leq 2(\ell+1)$, we can get

$$\frac{d}{dt}\log\big((1+t)^{\ell+1}\Xi^{\ell+1}(t) - \sigma t\big) \geq \frac{d}{dt}\log\big((1+t)^{\ell+1}\Xi^{\ell+1}(t)\big) - \frac{2(\ell+1)}{1+t} \geq 0,$$

then $(1+t)^{\ell+1}\Xi^{\ell+1}(t)e^{-\sigma t}$ can get the supremum at some point t_* and the inequality $\sigma(1+t_*) \geq 2(\ell+1)$ holds true. Thus, it follows

$$\Gamma_{\ell+1}(\sigma) = (1+t_*)^{\ell+1}\Xi^{\ell+1}(t_*)e^{-\sigma t_*} \le \frac{\sigma}{2(\ell+1)}(1+t_*)^{\ell+2}\Xi^{\ell+1}(t_*)e^{-\sigma t_*} \le \frac{\sigma}{2(\ell+1)}\Gamma_{\ell+2}(\sigma).$$

The conclusion is proven. □

References

1. Arnold, V.I.: Proof of a theorem of A. N. Kolmogorov on the persistence of quasi-perodic motions under small perturbations of the Hamiltonian. Russ. Math. Surv. **18**(5), 9–36 (1963)
2. Bounemoura, A., Fischler, S.: The classical KAM theorem for Hamiltonian systems via rational approximations. Regul. Chaotic Dyn. **19**(2), 251–265 (2014)
3. Bourgain, J.: On Melnikov's persistency problem. Math. Res. Lett. **4**(4), 445–458 (1997)
4. Eliasson, L.H., Fayad, B., Krikorian, R.: KAM tori near an analytic elliptic fixed point. Regul. Chaotic Dyn. **18**(6), 801–831 (2013)
5. Gallavotti, G., Gentile, G.: Hyperbolic low-dimensional invariant tori and summations of divergent series. Commun. Math. Phys. **227**, 421–460 (2002)
6. Gentile, G.: Degenerate lower-dimensional tori under the Bryuno condition. Ergodic Theory Dyn. Syst. **27**(2), 427–457 (2007)
7. Graff, S.M.: On the conservation of hyperbolic invariant tori for Hamiltonian systems. J. Differ. Equ. **15**(1), 1–69 (1974)
8. Koch, H., Kocić, S.: A renormalization approach to lower-dimensional tori with Brjuno frequency vectors. J. Differ. Equ. **249**(8), 1986–2004 (2010)
9. Koch, H., Kocić, S.: A renormalization group approach to quasiperiodic motion with Brjuno frequencies. Ergodic Theory Dyn. Syst. **30**(4), 1131–1146 (2010)
10. Kolmogorov, A.N.: On conservation of conditionally perodic motions for a small change in Hamilton's function. Dokl. Akad. Nauk SSSR **98**(4), 527–530 (1954)
11. Li, Y., Yi, Y.: Persistence of hyperbolic tori in Hamiltonian systems. J. Differ. Equ. **208**(2), 344–387 (2005)
12. Melnikov, V.K.: On certain cases of conservation of almost periodic motions with a small change of the Hamiltonian function. Dokl. Akad. Nauk SSSR **165**, 1245–1248 (1965)
13. Melnikov, V.K.: A certain family of conditionally periodic solutions of a Hamiltonian system. Dokl. Akad. Nauk SSSR **181**, 546–549 (1968)
14. Moser, J.: On invariant curves of area-preserving mappings of an annulus. Nachr. Akad. Wiss. Gött. Math. Phys. Kl. **II**(1962), 1–20 (1962)
15. Moser, J.: Convergent series expansions for quasi-periodic motions. Math. Ann. **169**, 136–176 (1967)
16. Pöschel, J.: A lecture on the classical KAM theorem. Proc. Sympos. Pure Math. **69**, 707–732 (2001)
17. Pöschel, J.: On elliptic lower-dimensional tori in Hamiltonian systems. Math. Z. **202**(4), 559–608 (1989)
18. Rüssmann, H.: On the one-dimensional Schrödinger equation with a quasi-periodic potential. Ann. N. Y. Acad. Sci. **357**, 90–107 (1980)
19. Rüssmann, H.: Invariant tori in non-degenerate nearly integrable Hamiltonian systems. Regul. Chaotic Dyn. **6**(2), 119–204 (2001)
20. Rüssmann, H.: Stability of elliptic fixed points of analytic area-preserving mappings under the Bruno condition. Ergodic Theory Dyn. Syst. **22**(5), 1551–1573 (2002)
21. Servyuk, M.B.: KAM-stable Hamiltonians. J. Dyn. Control Syst. **1**(3), 351–366 (1995)
22. Sevryuk, M.B.: Partial preservation of frequencies in KAM theory. Nonlinearity **19**(5), 1099–1140 (2006)
23. Xu, J., You, J., Qiu, Q.: Invariant tori of nearly integrable Hamiltonian systems with degeneracy. Math. Z. **226**, 375–386 (1997)
24. Xu, J., Lu, X.: General KAM theorems and their applications to invariant tori with prescribed frequencies. Regul. Chaotic Dyn. **21**(1), 107–125 (2016)
25. Xu, J.: Persistence of lower dimensional degenerate invariant tori with prescribed frequencies in Hamiltonian systems with small parameter. Nonlinearity **34**, 8192–8247 (2021)

26. Yoccoz, J.C.: Analytic Linearization of Circle Diffeomorphisms, Dynamical Systems and Small Divisors (Cetraro 1998). Lecture Notes in Math. 1784, pp. 125–173. Springer, Berlin (2002)