J. Differential Equations 249 (2010) 2796–2821

Contents lists available at ScienceDirect

Journal of Differential Equations

www.elsevier.com/locate/jde

Lower dimensional invariant tori with prescribed frequency for nonlinear wave equation ☆

Jiansheng Geng *, Xiufang Ren

Department of Mathematics and Institute of Mathematical Science, Nanjing University, Nanjing 210093, PR China

ARTICLE INFO

Article history:
Received 27 January 2010
Revised 1 April 2010
Available online 18 April 2010

Keywords:
Wave equation
Hamiltonian system
Birkhoff normal form
KAM theory
Invariant tori

ABSTRACT

In this paper, one-dimensional (1D) nonlinear wave equation $u_{tt} - u_{xx} + mu + u^3 = 0$, subject to Dirichlet boundary conditions is considered. We show that for each given $m > 0$, and each prescribed integer $b > 1$, the above equation admits a Whitney smooth family of small-amplitude quasi-periodic solutions with b-dimensional Diophantine frequencies, which correspond to b-dimensional invariant tori of an associated infinite-dimensional dynamical system. In particular, these Diophantine frequencies are the small dilation of a prescribed Diophantine vector. The proof is based on a partial Birkhoff normal form reduction and an improved KAM method.

1. Introduction and main result

The main conclusion we obtain in this paper is that there exist some quasi-periodic solutions, whose frequencies are the small dilation of a fixed Diophantine frequency ω^*, with the dilation factor λ, i.e.,

$$\omega = \lambda\omega^*, \quad \lambda \in \mathbb{R},\ \lambda \approx 1, \tag{1.1}$$

of the one-dimensional (1D) nonlinear wave equation

$$u_{tt} - u_{xx} + mu + u^3 = 0, \quad x \in [0, \pi],\ t \in \mathbb{R},\ m \in \mathbb{R}_+, \tag{1.2}$$

☆ This work is partially supported by NSFC grants 10531050, 10771098, 973 project of China 2007CB814800. This work is also partially supported by Program for New Century Excellent Talents in University.
* Corresponding author.
E-mail addresses: jgeng@nju.edu.cn (J. Geng), xiufangren@gmail.com (X. Ren).

0022-0396/$ – see front matter
doi:10.1016/j.jde.2010.04.003

subject to Dirichlet boundary conditions

$$u(t,0)=0=u(t,\pi). \tag{1.3}$$

Based on Eliasson [4], Melnikov [13] and Pöschel [14], the KAM method has been extensively developed in finite dimensions concerning the persistence of lower dimensional invariant tori in Hamiltonian systems (see also Bourgain [1], Li and Yi [11], Xu and You [17], You [18]). In recent years, the KAM method has been extended to infinite dimensions in works of [9,15], in studying quasi-periodic solutions for 1D nonlinear beam, wave and Schrödinger equations with constant potentials or parameterized potentials under Dirichlet boundary conditions or periodic boundary conditions (see also [3,5,10,12,16]), as to higher dimensional case, see Bourgain [2], Geng and You [7,8].

As we know, in [1], Bourgain combined the KAM method with the Nash–Moser type methods to obtain the persistence of the invariant torus $\mathbb{T}^b\times\{0\}\times\{0\}$ in $\mathbb{R}^{2b}\times\mathbb{R}^{2r}$-phase space, with perturbed frequency vector ω of the form (1.1) under the first Melnikov's non-resonance condition. In [4], Eliasson proved this result under the first, the second and the third Melnikov's non-resonance conditions and the non-degenerate conditions

$$\det\bigl(D\omega(y)\bigr)\neq 0,\qquad \bigl\langle l,\Omega(y)-\omega(y)\bigl(D\omega(y)\bigr)^{-1}D\Omega(y)\bigr\rangle\neq 0, \tag{1.4}$$

for all $y\in\mathbb{R}^n$, $l\in\mathbb{Z}^m\setminus 0$, $|l|\leqslant 3$.

Nonetheless, so far, such results have not yet been extended to infinite dimensions, i.e., the persistence of lower invariant tori, whose perturbed frequency vectors are of the form (1.1) in some infinite-dimensional phase space. The aim of this paper is to show that there exist many quasi-periodic solutions with the frequencies having the form (1.1) for Eq. (1.2), under conditions (1.3) and certain non-degenerate conditions similar to (1.4), we thus will give a positive answer to a question posed by J. Bourgain in [1] that the form (1.1) for finite-dimensional case can really be generalized to an infinite-dimensional phase space setting.

In [4], Eliasson imposed the third Melnikov's non-resonance condition in order that the frequencies keep the form (1.1). However in our case, because the normal frequencies have the form of $\mu_n=n+O(\frac{1}{n}), n\in\mathbb{Z}_+$, we can only assume the first and the second Melnikov's non-resonance conditions, since

$$\mu_i\pm\mu_j\pm\mu_k\to 0,\quad \text{as } i\pm j\pm k=0\quad\text{and}\quad \min\{i,j,k\}\to\infty.$$

As a result, we cannot eliminate all terms involving three normal variables in the perturbation. To overcome this difficulty, we make use of the idea in [11], to treat the tangential variable y and the normal variable w in the same scale rather than the traditional way of treating y as much smaller variable than w (see details in (2.4)). Therefore, the normal form of the Hamiltonian becomes more complicated, since there is a twist term $\langle yA,y\rangle$ in it, however, this twist term plays an essential role in ensuring the form (1.1) of the tangential frequencies. In fact, at each KAM step, we make a translation to extract a frequencies' rectification term from $\langle yA,y\rangle$ to eliminate the frequencies' drift. Consequently, after infinitely many KAM steps, we will have infinitely many parameters, however, they can be transformed into the same one-dimensional parameter, i.e., dilation factor λ. Although it is just one-dimensional, it will add to the hardship of the measure estimates (see details in Remark 3.3).

For any prescribed integer $b>1$, and any ordered b-index integer set $J_b=\{\{i_1,\dots,i_b\}\in\mathbb{Z}_+\colon\ 0<i_1<\cdots<i_b,\ \min_{1\leqslant j<b} i_{j+1}-i_j\leqslant b-1\}$, it is clear that the linearized equation associated with (1.2) with the same boundary conditions (1.3) has some small-amplitude quasi-periodic solutions of the form

$$u(t,x)=\sum_{j=1}^{b}\sqrt{\xi_j}\cos(\mu_{i_j}t)\sin i_jx,\quad \mu_{i_j}=\sqrt{i_j^2+m},\ 0<\xi_j\ll 1,$$

taking $\xi = (\xi_1, \ldots, \xi_b) \in \mathcal{O} \subset \mathbb{R}^b_+$ as parameters, in addition, we call J_b as an admissible tangential set with respect to b, and denote by $\mathbb{N}_1 = \mathbb{Z}_+ \setminus \{i_1, \ldots, i_b\}$.

Our main result states as follows:

Main Theorem. *Consider one-dimensional nonlinear wave equation*

$$u_{tt} - u_{xx} + mu + u^3 = 0, \quad x \in [0, \pi],\ t \in \mathbb{R},\ m \in \mathbb{R}_+,$$

subject to Dirichlet boundary conditions

$$u(t, 0) = 0 = u(t, \pi).$$

For any prescribed integer $b > 1$, choose $\{i_1, \ldots, i_b\} \in J_b$, then linearized equation has solutions

$$u(t, x) = \sum_{j=1}^{b} \sqrt{\xi_j} \cos(\mu_{i_j} t) \sin i_j x, \quad \mu_{i_j} = \sqrt{i_j^2 + m},\ 0 < \xi_j \ll 1,$$

taking $\xi = (\xi_1, \ldots, \xi_b) \in \mathcal{O} \subset \mathbb{R}^b_+$ as parameters, there exists a positive-measure Cantor subset $\widetilde{\mathcal{O}} \subset \mathcal{O}$, such that for any $\xi \in \widetilde{\mathcal{O}}$, the above nonlinear wave equation has a real analytic quasi-periodic solution

$$u(t, x) = \sum_{j=1}^{b} \sqrt{\xi_j} \cos(\omega_j t) \sin i_j x + O\big(|\xi|^{\frac{3}{2}}\big),$$

with

$$\omega_j = \lambda \omega_j^*, \quad \lambda \in \mathbb{R},\ \lambda \approx 1, \qquad \omega_j^* = \mu_{i_j} + \frac{6}{\pi \mu_{i_j}} \left(-\frac{\xi_j}{\mu_{i_j}} + \sum_{l=1}^{b} \frac{4\xi_l}{\mu_{i_l}} \right), \quad 1 \leqslant j \leqslant b.$$

Remark 1.1. The assumption of the set J_b is consistent with that of [16], which is made to ensure the existence of the small-amplitude quasi-periodic solutions for all positive m. Otherwise, one might have to exclude some set of m-values, which is discrete in every compact interval in $(0, \infty)$.

Remark 1.2. The result remains true if the nonlinearities u^3 is replaced by an odd function of the form $f(x, u) = au^3 + \sum_{k \geqslant 5} f_k(x) u^k$, $a \neq 0$, where the coefficients f_k are real analytic in x, or in some Sobolev space $H^s([0, \pi])$, $s > \frac{1}{2}$, with norms growing at most exponentially to ensure analyticity in u. We may also add a general odd perturbation term $\varepsilon g(x, u) = \varepsilon \sum_{k \geqslant 0} g_k(x) u^k$ to the above nonlinearity $f(x, u)$, with coefficients g_k of the same type as the f_k.

Remark 1.3. The frequency vector $\omega^* = (\omega_1^*, \ldots, \omega_b^*)$ is a Diophantine vector, i.e., there exist $\tau > 0$ and $\gamma > 0$ such that the frequency vector ω^* satisfies the following Diophantine conditions,

$$\big|\langle k, \omega^* \rangle\big| \geqslant \frac{\gamma}{|k|^\tau}, \quad \text{for all } k \in \mathbb{Z}^b \setminus \{0\}.$$

The rest of the paper is devoted to the proof of the Main Theorem. In Section 2, we define the weighted norms and study the basic properties, then we derive a partial Birkhoff normal form of order four for the lattice Hamiltonian (2.2), and then we extract the parameters from amplitude-frequency modulation. In Section 3, we give details for one step of KAM iteration. In Section 4, we show an iteration lemma and convergence. Proof of the theorem is completed in Section 5 by conducting measure estimates.

2. Normal form

First, we introduce some notations. Let $\ell^{a,\rho}$ be the Hilbert space of all real-valued sequences $q = (q_1, q_2, \ldots)$, endowed with the finite weighted norm

$$\|q\|_{a,\rho} = \sum_{n \geqslant 1} |q_n| n^a e^{n\rho} < \infty.$$

Introduce $v = u_t$ and $B = -\partial_{xx} + m$, then (1.2) reads

$$\begin{aligned} u_t &= \frac{\partial H}{\partial v} = v, \\ v_t &= -\frac{\partial H}{\partial u} = -Bu - u^3, \end{aligned} \tag{2.1}$$

where

$$H = \frac{1}{2}\langle v, v\rangle + \frac{1}{2}\langle Bu, u\rangle + \frac{1}{4}\int_0^\pi |u|^4 \, dx.$$

Let

$$u(t,x) = \sum_{n \geqslant 1} \frac{1}{\sqrt{\mu_n}} q_n(t)\phi_n(x), \qquad v(t,x) = \sum_{n \geqslant 1} \sqrt{\mu_n} p_n(t)\phi_n(x),$$

where $\phi_n(x) = \sqrt{\frac{2}{\pi}} \sin nx$, for $n = 1, 2, \ldots$ are the Dirichlet eigenfunctions of the operator B with eigenvalues $\lambda_n = n^2 + m$, setting $\mu_n = \sqrt{\lambda_n}$. Then, associated with the symplectic structure $\sum_{n \geqslant 1} dq_n \wedge dp_n$ on $\ell^{a,\rho} \times \ell^{a,\rho}$, we get the following Hamiltonian equations

$$\begin{aligned} \dot{q}_n &= \frac{\partial H}{\partial p_n}, \qquad \dot{p}_n = -\frac{\partial H}{\partial q_n}, \quad n \geqslant 1, \\ H &= \Lambda + G, \\ \Lambda &= \frac{1}{2}\sum_{n \geqslant 1} \mu_n \left(p_n^2 + q_n^2\right), \\ G &= \frac{1}{4}\int_0^\pi \left|u(x)\right|^4 dx = \frac{1}{4}\sum_{i,j,k,l} G_{ijkl} q_i q_j q_k q_l, \end{aligned} \tag{2.2}$$

where

$$G_{ijkl} = \frac{1}{\sqrt{\mu_i \mu_j \mu_k \mu_l}} \int_0^\pi \phi_i \phi_j \phi_k \phi_l \, dx, \qquad G_{ijkl} = 0 \quad \text{whenever } i \pm j \pm k \pm l \neq 0.$$

Lemma 2.1. *Let $a \geqslant 0$ and $\rho > 0$. If a curve $I \to \ell^{a,\rho} \times \ell^{a,\rho}$, $t \mapsto (q(t), p(t))$ is a real analytic solution of* (2.2)*, then*

$$u(t,x) = \sum_{n \geqslant 1} \frac{1}{\sqrt{\mu_n}} q_n(t)\phi_n(x)$$

is a real analytic solution of (1.2) *on $I \times [0,\pi]$.*

Let ℓ_b^1 and L^1, respectively, be the Hilbert spaces of all bi-infinite, absolute summable sequences with complex coefficients and all absolute-integrable complex-valued functions on $[-\pi,\pi]$. Let

$$\mathcal{G} : \ell_b^1 \to L^1, \quad q \mapsto \mathcal{G}q = \frac{1}{\sqrt{2\pi}} \sum_{n \in \mathbb{Z}} q_n e^{inx}$$

be the inverse discrete Fourier transform, which defines an isometry between the two spaces. For $a \geqslant 0$ and $\rho > 0$, define

$$\ell_b^{a,\rho} = \left\{ q \in \ell_b^1 \colon \|q\|_{a,\rho} = |q_0| + \sum_{n \neq 0} |q_n| n^a e^{n\rho} < \infty \right\},$$

through $\mathcal{G}$ they define subspaces $W^{a,\rho}$ of L^1 that are normed by setting $\|\mathcal{G}q\|_{a,\rho} = \|q\|_{a,\rho}$.

Lemma 2.2. *For $a \geqslant 0$ and $\rho > 0$, the space $\ell_b^{a,\rho}$ is a Banach algebra with respect to convolution of sequences, and*

$$\|q * p\|_{a,\rho} \leqslant c\|q\|_{a,\rho}\|p\|_{a,\rho},$$

with a constant c depending only on a. Consequently, $W^{a,\rho}$ is a Banach algebra with respect to multiplication of functions.

Lemma 2.3. *For $a \geqslant 0$ and $\rho > 0$, the gradient G_q is a real analytic map from a neighborhood of the origin of $\ell^{a,\rho}$ into $\ell^{a+1,\rho}$, with*

$$\|G_q\|_{a+1,\rho} = O\left(\|q\|_{a,\rho}^3\right).$$

By introducing the complex coordinates

$$z_n = \frac{1}{\sqrt{2}}(q_n + ip_n), \qquad \bar{z}_n = \frac{1}{\sqrt{2}}(q_n - ip_n),$$

we obtain a real analytic Hamiltonian $H = \sum_{n \geqslant 1} \mu_n |z_n|^2 + \cdots$ on the now complex Hilbert space $\ell^{a,\rho}$ with the symplectic structure $i\sum_{n \geqslant 1} \mathrm{d}z_n \wedge \mathrm{d}\bar{z}_n$.

Lemma 2.4. *For each finite $b > 1$ and each $m > 0$, there exists a real analytic symplectic change of coordinates Γ in some neighborhood of the origin in $\ell^{a,\rho}$ that takes the Hamiltonian $H = \Lambda + G$ into its partial Birkhoff normal form up to order four, that is*

$$H \circ \Gamma = \Lambda + \overline{G} + \widehat{G} + K,$$

such that $X_{\overline{G}}$, $X_{\widehat{G}}$, X_K are real analytic maps from some neighborhood of the origin in $\ell^{a,\rho}$ to $\ell^{a+1,\rho}$, where

$$\overline{G} = \frac{1}{2} \sum_{J_b \cap \{i,j\} \neq \emptyset} \overline{G}_{ij} |z_i|^2 |z_j|^2, \quad \overline{G}_{ij} = \frac{6}{\pi} \cdot \frac{4 - \delta_{ij}}{\lambda_i \lambda_j},$$

$$|\widehat{G}| = O\big(\|\hat{z}\|^4_{a,\rho}\big), \qquad |K| = O\big(\|z\|^6_{a,\rho}\big), \quad \hat{z} = (z_n)_{n \in \mathbb{N}_1}.$$

Moreover, the neighborhood can be chosen uniformly for every compact m-interval in $(0, \infty)$, and the dependence of Γ on m is real analytic.

For the proof of the above four lemmata, see [16].

Letting $I = (|z_{i_1}|^2, \ldots, |z_{i_b}|^2)$, $Z = (|z_n|^2, \ldots)$, $n \in \mathbb{N}_1 = \mathbb{Z}_+ \setminus \{i_1, \ldots, i_b\}$, by Lemma 2.4, we have

$$\Lambda = \langle \alpha, I \rangle + \langle \beta, Z \rangle, \qquad \overline{G} = \frac{1}{2} \langle IA, I \rangle + \langle IB, Z \rangle,$$

with vectors $\alpha = (\mu_{i_1}, \ldots, \mu_{i_b})$, $\beta = (\mu_n, \ldots)$, $n \in \mathbb{N}_1$, and matrices

$$A = \left(\frac{6}{\pi} \cdot \frac{4 - \delta_{kl}}{\mu_{i_k} \mu_{i_l}} \right)_{1 \leqslant k,l \leqslant b}, \qquad B = \left(\frac{24}{\pi \mu_{i_k} \mu_n} \right)_{1 \leqslant k \leqslant b,\, n \in \mathbb{N}_1}.$$

Next, we introduce the symplectic polar and complex coordinates by setting

$$z_n = \begin{cases} \sqrt{\xi_n + y_n} e^{-i x_n}, & n \in \{i_1, \ldots, i_b\}, \\ w_n, & n \in \mathbb{N}_1, \end{cases}$$

depending on the parameter ξ. We then get

$$i \sum_{n \geqslant 1} \mathrm{d}z_n \wedge \mathrm{d}\bar{z}_n = \sum_{n \in \{i_1, \ldots, i_b\}} \mathrm{d}x_n \wedge \mathrm{d}y_n + i \sum_{n \in \mathbb{N}_1} \mathrm{d}w_n \wedge \mathrm{d}\overline{w}_n,$$

and the new Hamiltonian

$$H = \big\langle \omega(\xi), y \big\rangle + \sum_{n \in \mathbb{N}_1} \Omega_n(\xi) w_n \overline{w}_n + \frac{1}{2} \langle yA, y \rangle + \langle yB, Z \rangle + \widehat{G} + K,$$

with frequencies $\omega(\xi) = \alpha + \xi A$, $\Omega(\xi) = \beta + \xi B$, where

$$Z = \big(|w_n|^2, \ldots\big), \quad n \in \mathbb{N}_1,$$

$$|\widehat{G}| = O\big(\|w\|^4_{a,\rho}\big), \quad w = (w_n, \ldots),\ n \in \mathbb{N}_1,$$

$$\begin{aligned} |K| = {} & O\big(|\xi|^3\big) + O\big(|y|^3\big) + O\big(|\xi|^2|y|\big) + O\big(|\xi||y|^2\big) + O\big(|\xi|^{\frac{5}{2}} \|w\|_{a,\rho}\big) \\ & + O\big(|\xi|^2 \|w\|^2_{a,\rho}\big) + O\big(|\xi||y|\|w\|^2_{a,\rho}\big) + O\big(|y|^2 \|w\|^2_{a,\rho}\big) + O\big(|\xi|^{\frac{3}{2}} \|w\|^3_{a,\rho}\big) \\ & + O\big(|\xi|\|w\|^4_{a,\rho}\big) + O\big(|y|\|w\|^4_{a,\rho}\big) + O\big(|\xi|^{\frac{1}{2}} \|w\|^5_{a,\rho}\big) + O\big(\|w\|^6_{a,\rho}\big). \end{aligned}$$

Rescaling y, w, $\overline{w}$, ξ by $\varepsilon^4 y$, $\varepsilon^2 w$, $\varepsilon^2 \overline{w}$, $\varepsilon^3 \xi$, we obtain

$$\begin{aligned}
\widetilde{H}(x, y, w, \overline{w}, \xi) &= \varepsilon^{-7} H\big(x, \varepsilon^4 y, \varepsilon^2 w, \varepsilon^2 \overline{w}, \varepsilon^3 \xi\big) \\
&= \big\langle \varepsilon^{-3}\alpha + \xi A, y\big\rangle + \big\langle \varepsilon^{-3}\beta + \xi B, w\overline{w}\big\rangle + \frac{1}{2}\langle y\varepsilon A, y\rangle \\
&\quad + \langle y\varepsilon B, Z\rangle + O\big(\varepsilon\|w\|_{a,\rho}^4\big) + O\big(\varepsilon^2|\xi|^3\big) + O\big(\varepsilon^5|y|^3\big) \\
&\quad + O\big(\varepsilon^3|\xi|^2|y|\big) + O\big(\varepsilon^4|\xi||y|^2\big) + O\big(\varepsilon^{\frac{5}{2}}|\xi|^{\frac{5}{2}}\|w\|_{a,\rho}\big) \\
&\quad + O\big(\varepsilon^3|\xi|^2\|w\|_{a,\rho}^2\big) + O\big(\varepsilon^4|\xi||y|\|w\|_{a,\rho}^2\big) + O\big(\varepsilon^5|y|^2\|w\|_{a,\rho}^2\big) \\
&\quad + O\big(\varepsilon^{\frac{7}{2}}|\xi|^{\frac{3}{2}}\|w\|_{a,\rho}^3\big) + O\big(\varepsilon^4|\xi|\|w\|_{a,\rho}^4\big) + O\big(\varepsilon^5|y|\|w\|_{a,\rho}^4\big) \\
&\quad + O\big(\varepsilon^{\frac{9}{2}}|\xi|^{\frac{1}{2}}\|w\|_{a,\rho}^5\big) + O\big(\varepsilon^5\|w\|_{a,\rho}^6\big).
\end{aligned}$$

From now on, we consider a Hamiltonian

$$\begin{gathered}
H = N + P, \\
N = \big\langle \omega(\xi), y\big\rangle + \big\langle \Omega(\xi), w\overline{w}\big\rangle + \frac{1}{2}\big\langle yA', y\big\rangle, \\
\omega(\xi) = \varepsilon^{-3}\alpha + \xi A, \qquad \Omega(\xi) = \varepsilon^{-3}\beta + \xi B, \qquad A' = \varepsilon A, \\
P = \widetilde{H} - N := \varepsilon\widetilde{P}(x, y, w, \overline{w}, \xi, \varepsilon). \qquad (2.3)
\end{gathered}$$

For simplicity, we substitute ξ_{i_j}, x_{i_j}, y_{i_j} by ξ_j, x_j, y_j, $j = 1, \dots, b$, respectively. To avoid confusion, we rewrite the above ε as ε_* in the following context. For given $r, s > 0$, let

$$D(r, s) = \big\{(x, y, w, \overline{w})\colon\ |\operatorname{Im} x| < r,\ |y| < s,\ \|w\|_{a,\rho} < s,\ \|\overline{w}\|_{a,\rho} < s\big\} \qquad (2.4)$$

be the complex neighborhood of $\mathbb{T}^b \times \{y = 0\} \times \{w = 0\} \times \{\overline{w} = 0\}$ in $\mathbb{T}^b \times \mathbb{R}^b \times \ell^{a,\rho} \times \ell^{a,\rho}$, where $|\cdot|$ denotes the sup-norm of complex vectors. Let $\alpha \equiv (\alpha_1, \dots, \alpha_n, \dots)_{n\in\mathbb{N}_1}$, $\beta \equiv (\beta_1, \dots, \beta_n, \dots)_{n\in\mathbb{N}_1}$, α_n and $\beta_n \in \mathbb{N}$ with finitely many nonzero components of positive integers. The product $w^\alpha \overline{w}^\beta$ denotes $\prod_n w_n^{\alpha_n} \overline{w}_n^{\beta_n}$. For any given real analytic function

$$F(x, y, w, \overline{w}) = \sum_{\alpha,\beta} F_{\alpha\beta}(x, y) w^\alpha \overline{w}^\beta,$$

where $F_{\alpha\beta}$ is a C_W^1 function depending on a parameter $\xi \in \mathcal{O}$ in the sense of Whitney (the precise form of the parameter space $\mathcal{O}$ will be specified at the end of this section), we define the weighted norm of F by

$$\begin{gathered}
\|F\|_{D(r,s),\mathcal{O}} \equiv \sup_{\substack{\|w\|_{a,\rho} < s \\ \|\overline{w}\|_{a,\rho} < s}} \sum_{\alpha,\beta} \|F_{\alpha\beta}\| \big|w^\alpha\big| \big|\overline{w}^\beta\big|, \\
F_{\alpha\beta} = \sum_{k\in\mathbb{Z}^b,\, l\in\mathbb{N}^b} F_{kl\alpha\beta}(\xi) y^l e^{i\langle k,x\rangle}, \\
\|F_{\alpha\beta}\| \equiv \sum_{k,l} |F_{kl\alpha\beta}|_{\mathcal{O}} s^{|l|} e^{|k|r}, \qquad |F_{kl\alpha\beta}|_{\mathcal{O}} = \sup_{\xi\in\mathcal{O}} \big|F_{kl\alpha\beta}(\xi)\big| \qquad (2.5)
\end{gathered}$$

($\langle\cdot,\cdot\rangle$ being the standard inner product in $\mathbb{C}^b$). The weighted norm of the Hamiltonian vector field

$$X_F = \big(F_y, -F_x, \{iF_{w_n}\}_{n\in\mathbb{N}_1}, \{-iF_{\overline{w}_n}\}_{n\in\mathbb{N}_1}\big)$$

associated with F on $D(r,s)\times\mathcal{O}$ is defined by[1]

$$\|X_F\|_{D(r,s),\mathcal{O}} \equiv \frac{1}{s}\|F_y\|_{D(r,s),\mathcal{O}} + \frac{1}{s^2}\|F_x\|_{D(r,s),\mathcal{O}} + \frac{1}{s}\bigg(\sum_{n\in\mathbb{N}_1}\|F_{w_n}\|_{D(r,s),\mathcal{O}}n^{\bar a}e^{n\rho} + \sum_{n\in\mathbb{N}_1}\|F_{\overline{w}_n}\|_{D(r,s),\mathcal{O}}n^{\bar a}e^{n\rho}\bigg), \tag{2.6}$$

where $\bar a > a$.

Remark 2.1. From Lemma 2.3, we know that $\bar a = a+1$, i.e., the weight of vector fields is a little heavier than that of w, $\overline{w}$. The boundedness of $\|X_F\|_{D(r,s),\mathcal{O}}$ means that X_F sends a decaying w-sequence to a faster decaying sequence.

For any real analytic functions F and G, define the Poisson bracket by

$$\{F,G\} = \left\langle\frac{\partial F}{\partial x},\frac{\partial G}{\partial y}\right\rangle - \left\langle\frac{\partial F}{\partial y},\frac{\partial G}{\partial x}\right\rangle + i\sum_n\left(\frac{\partial F}{\partial w_n}\frac{\partial G}{\partial \overline{w}_n} - \frac{\partial F}{\partial \overline{w}_n}\frac{\partial G}{\partial w_n}\right).$$

Lemma 2.5. *There exists a constant $c>0$, such that if*

$$\|X_F\|_{D(r,s),\mathcal{O}} < \varepsilon', \qquad \|X_G\|_{D(r,s),\mathcal{O}} < \varepsilon'',$$

for some $\varepsilon', \varepsilon'' > 0$, then for any $0<\sigma<r$ and $0<\eta\ll 1$, we have

$$\|X_{\{F,G\}}\|_{D(r-\sigma,\eta s),\mathcal{O}} < c\sigma^{-1}\eta^{-2}\varepsilon'\varepsilon''.$$

The proof is omitted, since it is just a copy of that in [6].

It is clear that $\mathbb{T}^b\times\{y=0\}\times\{w=0\}\times\{\overline{w}=0\}$ is an invariant torus of the integrable Hamiltonian N in the phase space $\mathbb{T}^b\times\mathbb{R}^b\times\ell^{a,\rho}\times\ell^{a,\rho}$. Our purpose is to prove that the Hamiltonian system determined by Hamiltonian $H=N+P$ still admits invariant tori provided that $\|X_P\|_{D(r,s),\mathcal{O}}$ is sufficiently small. Moreover, we point out that the tangential frequencies of these invariant tori lie in a fixed direction, they are just a multiple of a given Diophantine vector, and the multiple is around 1. However, this calls for imposing some conditions on the frequencies mapping $\xi\mapsto(\omega(\xi),\Omega(\xi))$ and the perturbation P in $\mathcal{O}$. We state them as follows.

(A1) *Regularity of the perturbation*: The perturbation P is *regular* in the sense that $\|X_P\|_{D(r,s),\mathcal{O}}<\infty$, with $\bar a = a+1$.

(A2) *Non-degeneracy*: The tangential frequencies mapping $\xi\mapsto\omega(\xi)$ is a C^1_W diffeomorphism between $\mathcal{O}$ and its image.

(A3) *Asymptotics of normal frequencies*:

$$\Omega_n\neq 0, \qquad \Omega_n(\xi)=\overline{\Omega}_n+\widehat{\Omega}_n, \qquad \overline{\Omega}_n = n+O\big(n^{-1}\big), \qquad \big|\widehat{\Omega}_n(\xi)\big|_{\mathcal{O}} = o\big(n^{-1}\big),$$
$$\overline{\Omega}_n-\overline{\Omega}_m = n-m+O\big(m^{-1}\big), \quad m\leqslant n,$$

for all $n,m\in\mathbb{N}_1$, where $\overline{\Omega}_n$ are real and independent of ξ.

[1] The norm $\|\cdot\|_{D(r,s),\mathcal{O}}$ for scalar functions is defined in (2.5). The vector function $G: D(r,s)\times\mathcal{O}\to\mathbb{C}^m$ ($m<\infty$) is similarly defined as $\|G\|_{D(r,s),\mathcal{O}} = \sum_{i=1}^m\|G_i\|_{D(r,s),\mathcal{O}}$.

(A4) *Non-resonance conditions*: For a given $\gamma > 0$ small enough, choose $2 < a_0 < 48$, and a function $l(a_0)$, satisfying $\frac{48a_0}{48-a_0} < l(a_0) < \infty$, let $\tau > 2b + (3 + \frac{2}{a_0})l(a_0)$ be fixed, we assume that for all $\xi \in \widetilde{\mathcal{O}}$, the frequencies of the obtained invariant tori satisfy the Diophantine conditions:

$$\left|\langle k, \omega(\xi)\rangle + \langle l, \Omega(\xi)\rangle\right| \geqslant \frac{\gamma}{|k|^\tau}, \quad \text{for all } (k,l) \in \mathbb{Z}^b \times \mathbb{Z}^\infty \setminus 0,\ |l| \leqslant 2, \tag{2.7}$$

i.e., $(\omega, \Omega) \in DC(\gamma, \tau)$, where $\widetilde{\mathcal{O}}$ is some subset of $\mathcal{O}$, and $\gamma = \gamma_* = \varepsilon_*^{\frac{1}{48}}$ before the first KAM step.

Remark 2.2. We say that $\omega \in DC_0(\gamma, \tau)$, if (2.7) holds for $k \in \mathbb{Z}^b \setminus 0$, $l \equiv 0$. Note that ω^* in Remark 1.3 belongs to $DC_0(\gamma, \tau)$.

Definition. We say that $H \in NF(r, s, d_1, d_2, d_3)$, if H is defined on $D(r,s) \times \mathcal{O}$, and the lower order terms are of the form

$$H_0 = [H_0], \qquad H_1 = \langle \omega(\xi), y\rangle, \qquad H_2 = [H_2] = \langle \Omega(\xi), w\overline{w}\rangle + \frac{1}{2}\langle yM(\xi), y\rangle,$$

in addition, the tangential frequencies mapping $\xi \mapsto \omega(\xi)$ and the matrix-valued mapping $\xi \mapsto M(\xi)$, defined on $\mathcal{O}$ satisfy

$$\begin{aligned} &C_1. \quad |\omega|_{\mathcal{O}} = \sup\left|\omega(\xi)\right|_{\mathcal{O}} \leqslant d_1 < \infty, \\ &C_2. \quad \left|(D\omega)^{-1}\right|_{\mathcal{O}} = \sup\left|\left(D\omega(\xi)\right)^{-1}\right|_{\mathcal{O}} \leqslant d_2 < \infty, \\ &C_3. \quad \left|M^{-1}\right|_{\mathcal{O}} = \sup\left|\left(M(\xi)\right)^{-1}\right|_{\mathcal{O}} \leqslant d_3 < \infty, \end{aligned}$$

for some $d_1, d_2, d_3 \in \mathbb{R}_+$, where

$$\begin{aligned} H_j(x, y, w, \overline{w}) &= \sum_{|l|+|\alpha|+|\beta|=j} H^j_{kl\alpha\beta}\, y^l e^{i\langle k,x\rangle} w^\alpha \overline{w}^\beta, \\ \left[H_j(x, y, w, \overline{w})\right] &= \sum_{|l|+2|\alpha|=j} H^j_{0l\alpha\alpha}\, y^l (w\overline{w})^\alpha, \end{aligned}$$

$[H_j]$ represents the mean value of H_j, $j = 0, 1, 2, \ldots,$ and M is a $b \times b$ matrix. We say that $H \in DC(\gamma, \tau)$, if the above $(\omega(\xi), \Omega(\xi)) \in DC(\gamma, \tau)$, for some $\xi \in \mathcal{O}$. We define the norm of the matrix $M = (m_{ij})_{a\times b}$ by $\max_{1\leqslant i\leqslant a} \sum_{j=1}^b |m_{ij}|$, and denote by $\widetilde{H}$ the function $H - H_0 - H_1 - H_2$. For convenience, we write $H_{0l00,|l|=2}, H_{0l00,|l|=1}, H_{00ll,|l|=1}$ as $H_{0200}, H_{0100}, H_{0011}$, respectively.

Remark 2.3. In view of the Hamiltonian (2.3) and the non-degeneracy of $A = (\frac{6}{\pi} \cdot \frac{4-\delta_{kl}}{\mu_{i_k}\mu_{i_l}})_{1\leqslant k,l\leqslant b}$, it is reasonable to assume that $\xi \mapsto \omega(\xi)$ and $\xi \mapsto A'(\xi) = \varepsilon_* A$, defined on $[1,2]^b \subset \mathbb{R}^b$, satisfy conditions C_1, C_2, C_3 for some $d_1^*, d_2^*, d_3^* \in \mathbb{R}_+$. We will see later that $\mathcal{O} = [1,2]^b$ before the first KAM step, however, throughout the KAM steps,

$$\mathcal{O} = \left[-\left(1 + \varepsilon_*^2\right) d_1^* d_2^* \varepsilon_*^{3+\frac{1}{a_0}}, \left(1 + \varepsilon_*^2\right) d_1^* d_2^* \varepsilon_*^{3+\frac{1}{a_0}}\right]^b, \tag{2.8}$$

equivalently, $\mathcal{O} = [-\frac{1}{2}, \frac{1}{2}]$.

3. KAM step

To begin with the KAM iteration, we first fix $r, s, \varepsilon_* > 0$ and restrict the Hamiltonian (2.3) to the domain $D(r,s)$, restrict the parameter ξ to the set $[1,2]^b$. We set the initial values $\omega_0 = \omega_*$, $\Omega^0 = \Omega^*$, $A'_0 = A'_* = \varepsilon_* A$, $P^0 = P^*$, $r_0 = r$, $s_0 = s$, $\gamma_0 = \gamma_* = \varepsilon_*^{\frac{1}{48}}$, $d_1^0 = d_1^*$, $d_2^0 = d_2^*$, $d_3^0 = d_3^*$, and

$$H^0 = N_0 + P^0\big(x, y, w, \overline{w}, \xi^*, \varepsilon_*\big),$$
$$N_0 = \langle \omega_0, y\rangle + \big\langle \Omega^0, w\overline{w}\big\rangle + \frac{1}{2}\big\langle yA'_0, y\big\rangle \in NF\big(r_0, s_0, d_1^0, d_2^0, d_3^0\big) \cap DC(\gamma_0, \tau),$$

where $\omega_* = \omega(\xi^*) = \varepsilon_*^{-3}\alpha + \xi^* A$, $\Omega^* = \Omega(\xi^*) = \varepsilon_*^{-3}\beta + \xi^* B$, $P^* = \varepsilon_* \widetilde{P}(x, y, w, \overline{w}, \xi^*, \varepsilon_*)$, $\xi^* \in \mathcal{O}_0$,

$$\mathcal{O}_0 = \left\{\xi \in [1,2]^b\colon \left|\big\langle k, \omega_0(\xi)\big\rangle + \big\langle l, \Omega^0(\xi)\big\rangle\right| \geqslant \frac{\gamma_0}{|k|^\tau},\ |k| + |l| \neq 0\right\}.$$

For convenience, we will fix ξ by ξ^* throughout this paper. It is obvious that there exists a positive constant c_*, such that

$$\|X_{P^0}\|_{D(r_0,s_0),\mathcal{O}_0} \leqslant c_*\varepsilon_* := \varepsilon_0. \tag{3.1}$$

Remark 3.1. For fixed $\varepsilon_* > 0$, and for prescribed integer $b > 1$, the existence of $\mathcal{O}_0$ can be guaranteed by Lemma 6 in [16], since we have chosen the admissible tangential set J_b.

Suppose that after νth KAM step, we arrive at a Hamiltonian

$$H = H_\nu = N + P\big(x, y, w, \overline{w}, \xi^*, \lambda, \varepsilon\big),$$
$$N = N_\nu = \langle \omega, y\rangle + \langle \Omega, w\overline{w}\rangle + \frac{1}{2}\big\langle yA', y\big\rangle \in NF(r, s, d_1, d_2, d_3) \cap DC(\gamma, \tau), \tag{3.2}$$

which is real analytic on $D = D_\nu = D(r_\nu, s_\nu)$, for some $r = r_\nu \leqslant r_0$, $s = s_\nu \leqslant s_0$, and depends on $\lambda = \lambda_\nu \in \Lambda = \Lambda_\nu \subset \mathcal{E} = [-\frac{1}{2}, \frac{1}{2}]$ Whitney smoothly, with $\omega = \omega_\nu = \omega_\nu(\lambda) = (1 + \varepsilon_*^{3+\frac{1}{a_0}}\lambda)\omega_* := t(\lambda)\omega_* = \tilde{\lambda}\omega_*$, $\Omega = \Omega^\nu = \Omega^\nu(\lambda)$, $A' = A'_\nu = A'_\nu(\lambda)$, $d_1 = d_1^\nu$, $d_2 = d_2^\nu$, $d_3 = d_3^\nu$, $\frac{1}{4}\gamma_0 \leqslant \gamma = \gamma_\nu \leqslant \frac{1}{2}\gamma_0$, $P = P^\nu(\lambda)$,

$$\Lambda_\nu = \left\{\lambda\colon \left|\langle k, \omega_\nu\rangle + \big\langle l, \Omega^\nu\big\rangle\right| \geqslant \frac{\gamma_\nu}{|k|^\tau},\ |k| + |l| \neq 0,\ \omega_\nu = t(\lambda)\omega_*\right\}.$$

We also assume that

$$\|X_P\|_{D,\Lambda} \leqslant \varepsilon, \tag{3.3}$$

for some $0 < \varepsilon = \varepsilon_\nu \leqslant \varepsilon_0$.

Remark 3.2. In the last part of this paper, we will find that $\omega_* = \varepsilon_*^{-3}\hat{\omega}_*$, $\hat{\omega}_* = \hat{\omega}(\hat{\xi}^*) = \alpha + \hat{\xi}^* A$, where $\hat{\xi}^* = \varepsilon_*^3 \xi^*$. For each $\lambda^\infty \in \Lambda^\star = \bigcap_{\nu \geqslant 1} \Lambda_\nu$, and each $\hat{\xi}^* \in \widetilde{\mathcal{O}}_0 \subset [\varepsilon_*^3, 2\varepsilon_*^3]^b$, we finally get an invariant torus $\Psi_\infty(\mathbb{T}^b \times \{\lambda^\infty\})$ of the original Hamiltonian H_0 with the tangential frequency of the form $\omega^\star = (1 + \varepsilon_*^{3+\frac{1}{a_0}}\lambda^\infty)\hat{\omega}_*$. Since $\omega_\nu(0) = \omega_*$, we can view Λ_0 as $\{\mathcal{O}_0\colon \lambda \equiv 0\}$. Moreover, in this paper, we index quantities at $(\nu+1)$th step by $+$, and write $< \cdot$ in the estimates to suppress various constants, which do not depend on the iteration steps. In addition, all the constants c_1, c_2, c_3, c_4 below are positive and independent of the iteration steps.

In the following, we look for a special $F = F^\nu$ defined on a smaller domain $D_+ = D(r_+, s_+)$, such that the translation $\phi_+ : (x, y, w, \overline{w}) \to (x, y + y^+, w, \overline{w})$, where y^+ is $(\nu+1)$th new introduced parameter, and the time one map $\Phi_F^1 := \Phi_+$ of the Hamiltonian flow Φ_F^t associated with F carry the above Hamiltonian (3.2) into the next KAM cycle, which means that the new Hamiltonian $H^+ = H \circ \Phi^+ = H \circ \phi_+ \circ \Phi_+ = N_+ + P^+$ satisfies all the above assumptions $(A_1), \ldots, (A_4)$ and has the same estimates as (3.3) with respect to the new parameters r_+, s_+, ε_+, d_1^+, d_2^+, d_3^+, γ_+ and new domains D_+, Λ_+. Moreover, $N_+ \in NF(r_+, s_+, d_1^+, d_2^+, d_3^+) \cap DC(\gamma_+, \tau)$, $\omega_+ = (1 + \varepsilon_*^{3+\frac{1}{a_0}} \lambda)\omega_*$, $\lambda \in \Lambda_+$.

3.1. Truncating the perturbation and solving the homological equation

Let $R = P_0 + P_1 + P_2$ be the truncation of the Taylor–Fourier series of P up to order 2, i.e., $R = \sum_{|l|+|\alpha|+|\beta|\leqslant 2} P_{kl\alpha\beta} y^l e^{i\langle k,x\rangle} w^\alpha \overline{w}^\beta$, we wish to construct a function $F = F_0 + F_1 + F_2$, with $[F] = 0$, such that

$$\{F, N\} = R - [R] - Q, \tag{3.4}$$

where $Q = \{N, F\}_3 = \{\frac{1}{2}\langle yA', y\rangle, F_2\}$. Let Φ_F^t be the Hamiltonian flow of F, then

$$\begin{aligned}
(N + P) \circ \phi_+ \circ \Phi_F^1 &= N_+ + P^+, \\
N_+ &= N + \frac{1}{2}\langle y^+ A, y^+\rangle + \langle y^+ A, y\rangle + [R_0] + [R_1] + [R_2], \\
P^+ &= \int_0^1 \{tR + (1-t)[R], F\} \circ \phi_+ \circ \Phi_F^t \, dt + \widetilde{P} \circ \phi_+ \circ \Phi_F^1 + Q \circ \phi_+.
\end{aligned} \tag{3.5}$$

Therefore, after $(\nu+1)$th KAM step, we arrive at the Hamiltonian

$$\begin{aligned}
H \circ \phi_1 \circ \Phi_1 \circ \cdots \circ \phi_{\nu+1} \circ \Phi_{\nu+1} &= H^{\nu+1} = N_{\nu+1} + P^{\nu+1}, \\
N_{\nu+1} &= \langle \omega_{\nu+1}, y\rangle + \langle \Omega^{\nu+1}, w\overline{w}\rangle + \frac{1}{2}\langle yA'_{\nu+1}, y\rangle, \\
\omega_{\nu+1} &= \omega_* + (y^1 + \cdots + y^{\nu+1})A + P^0_{0100} + \cdots + P^\nu_{0100}, \\
\Omega^{\nu+1} &= \Omega^* + (y^1 + \cdots + y^{\nu+1})B + P^0_{0011} + \cdots + P^\nu_{0011}, \\
A'_{\nu+1} &= A'_* + P^0_{0200} + \cdots + P^\nu_{0200}.
\end{aligned}$$

Define $\phi_j : (x, y, w, \overline{w}) \to (x, y + y^j, w, \overline{w})$, $j \geqslant 1$, with $y^j = -P^{j-1}_{0100} A^{-1}$, $j \geqslant 2$, $y^1 = ((\tilde{\lambda} - 1)\omega_* - P^0_{0100})A^{-1}$, we then have

$$\begin{aligned}
\omega_{\nu+1} &= \omega_* + y^1 A + P^0_{0100} = \tilde{\lambda}\omega_* = \big(1 + \varepsilon_*^{3+\frac{1}{a_0}} \lambda\big)\omega_*, \\
\Omega^{\nu+1} &= \Omega^* + y^1 B - \big(P^1_{0100} + \cdots + P^\nu_{0100}\big)A^{-1}B + P^0_{0011} + \cdots + P^\nu_{0011}, \\
A'_{\nu+1} &= A'_* + P^0_{0200} + \cdots + P^\nu_{0200}.
\end{aligned} \tag{3.6}$$

Remark 3.3. Observe that $\tilde{\lambda} \in [1 - \frac{1}{2}\varepsilon_*^{3+\frac{1}{a_0}}, 1 + \frac{1}{2}\varepsilon_*^{3+\frac{1}{a_0}}]$, $|\omega_*| \leqslant d_1^0 = O(\varepsilon_*^{-3})$, $P^j_{0100} = P^j_{0100}(\xi^*, y^1, \ldots, y^j)$, $1 \leqslant j \leqslant \nu$, $P^0_{0100} = O(\varepsilon_*^3 |\xi^*|^2)$, $\xi^* \in [1, 2]^b$ is fixed, and $|A^{-1}| \leqslant c_{\pi,m,b}$, for some constant $c_{\pi,m,b} > 0$. Therefore, we have $y^1 = y^1(\lambda)$, $y^j = y^j(\xi^*, y^1, \ldots, y^{j-1}) = y^j(y^1) = y^j(\lambda)$, $j \geqslant 2$, with the estimates $|y^1| \leqslant (1+O(\varepsilon_*^{3-\frac{1}{a_0}}))c_{\pi,m,b}d_1^0\varepsilon_*^{3+\frac{1}{a_0}} \leqslant (1+\varepsilon_*^2)c_{\pi,m,b}d_1^0\varepsilon_*^{3+\frac{1}{a_0}} := \chi = O(\varepsilon_*^{\frac{1}{a_0}})$, and $|y^j| \leqslant c_{\pi,m,b}|P_{0100}^{j-1}| \leqslant c_{\pi,m,b}\varepsilon_{j-1}$. Thus, $P_{0100}^j = P_{0100}^j(y^1)$, $P_{0011}^j = P_{0011}^j(y^1)$, $P_{0200}^j = P_{0200}^j(y^1)$, $1 \leqslant j \leqslant \nu$. Consequently, the frequencies and the matrix-valued mapping in (3.6) can be regarded as the mapping from $[-\chi,\chi]^b$ to their images, i.e., $\omega_{\nu+1} = \omega_{\nu+1}(y^1)$, $\Omega^{\nu+1} = \Omega^{\nu+1}(y^1)$, $A'_{\nu+1} = A'_{\nu+1}(y^1)$, $\nu \geqslant 0$, so after we choose ξ in the initial parameter space, the parameter space in the KAM step is $[-\chi,\chi]^b$ in fact, however, since $y^1(\lambda) : \lambda \to (\varepsilon_*^{3+\frac{1}{a_0}}\lambda\omega_* - P_{0100}^0)A^{-1}$ is one–one, it is reasonable to consider the one-dimensional parameter space $\mathcal{E} = [-\frac{1}{2}, \frac{1}{2}]$, rather than the whole one $\Xi := \{(\xi, y^1, \ldots, y^\infty)\colon\ \xi \in [1,2]^b,\ y^1 \in [-\chi,\chi]^b,\ y^j \in [-c_{\pi,m,b}\varepsilon_{j-1}, c_{\pi,m,b}\varepsilon_{j-1}]^b,\ 2 \leqslant j \leqslant \infty\}$, and to assume that H depends on λ Whitney smoothly.

Lemma 3.1. *Define $D_j = D(r_j, s_j) = D(r_+ + j\sigma, \frac{j}{4}s)$, $0 < j \leqslant 4$, then the solution of* (3.4) *satisfies*

$$\|X_F\|_{D_3,\Lambda} < \cdot\gamma^{-4}(r-r_+)^{-(4\tau+5)}\varepsilon,$$

with

$$|F_{kl\alpha\beta}|_\Lambda = \sup_{\lambda\in\Lambda}\max_{|\iota|\leqslant 1}\left(\left|\frac{\partial^\iota F_{kl\alpha\beta}}{\partial\lambda^\iota}\right|\right). \tag{3.7}$$

(*The derivatives with respect to λ are in the sense of Whitney.*)

Proof. We will construct a function $F = F_0 + F_1 + F_2$, where

$$F_0 = \sum_{k\neq 0} F_k^0 e^{i\langle k,x\rangle}, \qquad F_1 = \sum_{k,n}\big(F_{kn}^{10}w_n + F_{kn}^{01}\overline{w}_n\big)e^{i\langle k,x\rangle} + \sum_{k\neq 0,\,|l|=1} F_k^1 y^l e^{i\langle k,x\rangle},$$

$$F_2 = \sum_{|l|=1,\,n}\big(F_{kn}^{110}y^l w_n + F_{kn}^{101}y^l\overline{w}_n\big)e^{i\langle k,x\rangle} + \sum_{|k|+|n-m|\neq 0} F_{knm}^{11}w_n\overline{w}_m e^{i\langle k,x\rangle}$$
$$+ \sum_{k,n,m}\big(F_{knm}^{20}w_nw_m + F_{knm}^{02}\overline{w}_n\overline{w}_m\big)e^{i\langle k,x\rangle} + \sum_{k\neq 0,\,|l|=2} F_k^2 y^l e^{i\langle k,x\rangle},$$

having the same form as P_0, P_1, P_2, such that

$$\{F_j, N\} = P_j - [P_j] - Q_j, \qquad [F_j] = 0, \quad j = 0,1,2. \tag{3.8}$$

This is equivalent to

$$i\langle k,\omega\rangle F_k^0 = P_k^0, \quad k\neq 0,$$
$$i\big(\langle k,\omega\rangle + \Omega_n\big)F_{kn}^{10} = P_{kn}^{10}, \quad |l| = 1,$$
$$i\big(\langle k,\omega\rangle - \Omega_n\big)F_{kn}^{01} = P_{kn}^{01}, \quad |l| = 1,$$
$$i\langle k,\omega\rangle F_k^1 + ikF_k^0 A = P_k^1,$$
$$i\big(\langle k,\omega\rangle + \Omega_n + \Omega_m\big)F_{knm}^{20} + ikF_k^1 A = P_{knm}^{20},$$
$$i\big(\langle k,\omega\rangle + \Omega_n - \Omega_m\big)F_{knm}^{11} = P_{knm}^{11}, \quad |k| + |n-m| \neq 0,$$
$$i\big(\langle k,\omega\rangle - \Omega_n - \Omega_m\big)F_{knm}^{02} = P_{knm}^{02},$$

$$i\big(\langle k,\omega\rangle+\Omega_n\big)F_{kn}^{110}+ikF_{kn}^{10}A=P_{kn}^{110},$$
$$i\big(\langle k,\omega\rangle-\Omega_n\big)F_{kn}^{101}+ikF_{kn}^{01}A=P_{kn}^{101},$$
$$i\langle k,\omega\rangle F_k^2+ikF_k^1A=P_k^2.$$

Since $N\in DC(\gamma,\tau)$, we have

$$\begin{aligned}
\big|F_k^0\big|_\Lambda &\leqslant \gamma^{-1}|k|^\tau\big|P_k^0\big|_\Lambda, \quad k\neq 0,\\
\big|F_{kn}^{10}\big|_\Lambda &\leqslant \gamma^{-1}|k|^\tau\big|P_{kn}^{10}\big|_\Lambda, \quad |l|=1,\\
\big|F_{kn}^{01}\big|_\Lambda &\leqslant \gamma^{-1}|k|^\tau\big|P_{kn}^{01}\big|_\Lambda, \quad |l|=1,\\
\big|F_{knm}^{11}\big|_\Lambda &\leqslant \gamma^{-1}|k|^\tau\big|P_{knm}^{11}\big|_\Lambda, \quad |k|+|n-m|\neq 0,\\
\big|F_{knm}^{02}\big|_\Lambda &\leqslant \gamma^{-1}|k|^\tau\big|P_{knm}^{02}\big|_\Lambda.
\end{aligned}\tag{3.9}$$

It follows that

$$\begin{aligned}
\big|F_k^1\big|_\Lambda &<\cdot \frac{|k||F_k^0|_\Lambda}{|\langle k,\omega\rangle|}<\cdot\gamma^{-2}|k|^{2\tau+1}\big|P_k^0\big|_\Lambda,\\
\big|F_{knm}^{20}\big|_\Lambda &<\cdot \frac{|k||F_k^1|_\Lambda}{|\langle k,\omega\rangle+\Omega_n+\Omega_m|}<\cdot\gamma^{-3}|k|^{3\tau+2}\big|P_k^0\big|_\Lambda,\\
\big|F_{kn}^{110}\big|_\Lambda &<\cdot \frac{|k||F_{kn}^{10}|_\Lambda}{|\langle k,\omega\rangle+\Omega_n|}<\cdot\gamma^{-2}|k|^{2\tau+1}\big|P_{kn}^{10}\big|_\Lambda,\\
\big|F_{kn}^{101}\big|_\Lambda &<\cdot \frac{|k||F_{kn}^{01}|_\Lambda}{|\langle k,\omega\rangle-\Omega_n|}<\cdot\gamma^{-2}|k|^{2\tau+1}\big|P_{kn}^{01}\big|_\Lambda,\\
\big|F_k^2\big|_\Lambda &<\cdot \frac{|k||F_k^1|_\Lambda}{|\langle k,\omega\rangle|}<\cdot\gamma^{-3}|k|^{3\tau+2}\big|P_k^0\big|_\Lambda.
\end{aligned}\tag{3.10}$$

Therefore,

$$\begin{aligned}
\|X_{F_0}\|_{D_3,\Lambda} &\leqslant \frac{1}{s_3^2}\sum_{k\neq 0}\big|F_k^0\big||k|e^{|k|r_3}\leqslant \frac{1}{s_3^2}\sum_{k\neq 0}\gamma^{-1}|k|^\tau|k|\big|P_k^0\big|e^{|k|r_3}\\
&\leqslant \frac{s_4^2}{s_3^2}\gamma^{-1}\sum_{k\neq 0}|k|^\tau e^{|k|r_3}\|X_{P_0}\|_{D_4,\Lambda}e^{-|k|r_4}\\
&<\cdot\gamma^{-1}\sum_{k\neq 0}|k|^\tau e^{-|k|(r_4-r_3)}\varepsilon,
\end{aligned}\tag{3.11}$$

$$\begin{aligned}
\|X_{F_1}\|_{D_3,\Lambda} =& \frac{1}{s_3^2}\bigg\|\sum_{k,n}ik\big(F_{kn}^{10}w_n+F_{kn}^{01}\overline{w}_n\big)e^{i\langle k,x\rangle}+\sum_{k\neq 0,\,|l|=1}ikF_k^1y^le^{i\langle k,x\rangle}\bigg\|\\
&+\frac{1}{s_3}\bigg\|\sum_{k\neq 0}F_k^1e^{i\langle k,x\rangle}\bigg\|+\frac{1}{s_3}\sum_{k,n}\big\|\big(F_{kn}^{10}+F_{kn}^{01}\big)e^{i\langle k,x\rangle}\big\|n^{\bar a}e^{n\rho}
\end{aligned}$$

$$< \frac{s_4}{s_3}\Big(\gamma^{-1}\sum_{k,n}|k|^\tau|k|\big(|P^{10}_{kn}|+|P^{01}_{kn}|\big)e^{|k|r_3}+\gamma^{-2}\sum_{k\neq 0}|k||k|^{2\tau+1}|P^0_k|e^{|k|r_3}$$

$$+\gamma^{-2}\sum_{k\neq 0}|k|^{2\tau+1}|P^0_k|e^{|k|r_3}+\gamma^{-1}\sum_{k,n}|k|^\tau\big(|P^{10}_{kn}|+|P^{01}_{kn}|\big)e^{|k|r_3}n^{\bar a}e^{n\rho}\Big)$$

$$< \cdot\gamma^{-2}\sum_{k\neq 0}|k|^{2\tau+1}e^{-|k|(r_4-r_3)}\varepsilon, \tag{3.12}$$

$$\|X_{F_2}\|_{D_3,\Lambda}=\frac{1}{s_3}\Big\|\sum_{k\neq 0,\,|l'|=1}2F^2_k y^{l'}e^{i\langle k,x\rangle}+\sum_{|l|=1,n}F^{110}_{kn}w_n e^{i\langle k,x\rangle}+\sum_{|l|=1,n}F^{101}_{kn}\overline{w}_n e^{i\langle k,x\rangle}\Big\|$$

$$+\frac{1}{s_3^2}\Big\|\sum_{|l|=2}ikF^2_k y^l e^{i\langle k,x\rangle}+\sum_{|l|=1,n}ikF^{110}_{kn}y^l w_n e^{i\langle k,x\rangle}+\sum_{|l|=1,n}ikF^{101}_{kn}y^l\overline{w}_n e^{i\langle k.x\rangle}$$

$$+\sum_{k,n,m}ik\big(F^{20}_{knm}w_n w_m+F^{02}_{knm}\overline{w}_n\overline{w}_m\big)e^{i\langle k.x\rangle}+\sum_{|k|+|n-m|\neq 0}ikF^{11}_{knm}w_n\overline{w}_m e^{i\langle k,x\rangle}\Big\|$$

$$+\frac{1}{s_3}\sum_n\Big[\Big\|\sum_{|l|=1}F^{110}_{kn}y^l e^{i\langle k,x\rangle}+\sum_{k,m}F^{11}_{knm}\overline{w}_m e^{i\langle k,x\rangle}+\sum_{k,m}F^{20}_{knm}w_m e^{i\langle k,x\rangle}\Big\|$$

$$+\Big\|\sum_{|l|=1}F^{101}_{kn}y^l e^{i\langle k,x\rangle}+\sum_{k,m}F^{02}_{knm}\overline{w}_m e^{i\langle k,x\rangle}\Big\|\Big]n^{\bar a}e^{n\rho}$$

$$\leqslant\Big(2\sum_{k,n}\gamma^{-3}|k|^{3\tau+2}|P^0_k|e^{|k|r_3}+\gamma^{-2}|k|^{2\tau+1}\big(|P^{10}_{kn}|+|P^{01}_{kn}|\big)e^{|k|r_3}$$

$$+\sum_k|k|\gamma^{-3}|k|^{3\tau+2}|P^0_k|e^{|k|r_3}+\sum_{k,n}|k|\gamma^{-2}|k|^{2\tau+1}\big(|P^{10}_{kn}|+|P^{01}_{kn}|\big)e^{|k|r_3}$$

$$+\sum_{k,n,m}|k|\gamma^{-3}|k|^{3\tau+2}|P^{20}_{knm}|e^{|k|r_3}+\sum_{k,n,m}|k|\gamma^{-1}|k|^\tau|P^{02}_{knm}|e^{|k|r_3}$$

$$+\sum_{|k|+|n-m|\neq 0}|k|\gamma^{-1}|k|^\tau|P^{11}_{knm}|e^{|k|r_3}+\sum_n\Big(\sum_{k,m}\gamma^{-1}|k|^\tau|P^{11}_{knm}|e^{|k|r_3}$$

$$+\sum_k\gamma^{-2}|k|^{2\tau+1}\big(|P^{10}_{kn}|+|P^{01}_{kn}|\big)e^{|k|r_3}+\sum_k\gamma^{-3}|k|^{3\tau+2}2|P^0_k|e^{|k|r_3}$$

$$+\sum_{k,m}\gamma^{-1}|k|^\tau|P^{02}_{knm}|e^{|k|r_3}\Big)n^{\bar a}e^{n\rho}\Big)<\cdot\gamma^{-3}\sum_{k\neq 0}|k|^{3\tau+2}e^{-|k|(r_4-r_3)}\varepsilon. \tag{3.13}$$

In conclusion, we get

$$\|X_F\|_{D_3,\Lambda}<\cdot\gamma^{-3}\sum_{l\geqslant 1}l^{3\tau+b+1}e^{-l(r_4-r_3)}\varepsilon<\cdot\gamma^{-3}(r-r_+)^{-(3\tau+4)}\varepsilon.$$

Through elementary calculation, we have

$$|F_{kl\alpha\beta}|_\Lambda+\left|\frac{\partial F_{kl\alpha\beta}}{\partial\lambda}\right|_\Lambda<\cdot\gamma^{-2}|k|^{2\tau+1}\left(|P_{kl\alpha\beta}|_\Lambda+\left|\frac{\partial P_{kl\alpha\beta}}{\partial\lambda}\right|_\Lambda\right),$$

thus the conclusion can be reached, if we multiply the upper bounds of (3.9), (3.10), thus of (3.11), (3.12), (3.13) by $\gamma^{-1}|k|^{\tau+1}$, and redefine the norm of $F_{kl\alpha\beta}$ by the C^1 norm (3.7), instead of the C^0 norm (2.5). □

3.2. Estimating the new lower order terms

By virtue of $\omega_+ = \omega_+(y^1) = \omega_* + y^1 A + P^0_{0100} = \tilde{\lambda}\omega_*$, $\Omega^+ = \Omega + y^+ B + P^\nu_{0011}$, $A'_+ = A'_* + P^0_{0200} + \cdots + P^\nu_{0200}$, for all $\nu \geqslant 0$, and $|y^1| = O(\varepsilon_*^{\frac{1}{a_0}})$, $|y^+| = O(\varepsilon)$, for all $\nu \geqslant 1$, $\tilde{\lambda} \in [1 - \frac{1}{2}\varepsilon_*^{3+\frac{1}{a_0}}, 1 + \frac{1}{2}\varepsilon_*^{3+\frac{1}{a_0}}]$, $|A'^{-1}_*| = O(\varepsilon_*^{-1})$, $P^0_{0200} = O(\varepsilon_*^4|\xi^*|) = O(\varepsilon_*^4)$, we have

$$\begin{aligned} |\omega_+|_{[-\chi,\chi]^b} &\leqslant \left(1 + \frac{1}{2}\varepsilon_*^{3+\frac{1}{a_0}}\right)|\omega_*|_{[1,2]^b} \leqslant \left(1 + \frac{1}{2}\varepsilon_*^{3+\frac{1}{a_0}}\right)d_1^0, \\ \left|(D\omega_+)^{-1}\right|_{[-\chi,\chi]^b} &= \left|A^{-1}\right| = \left|(D\omega_*)^{-1}\right|_{[1,2]^b} \leqslant d_2^0, \\ \left|A'^{-1}_+\right|_{[-\chi,\chi]^b} &= \left|A'^{-1}_*\left[I + \left(P^0_{0200} + \cdots + P^\nu_{0200}\right)A'^{-1}_*\right]^{-1}\right| \\ &\leqslant \left(1 + O\left(\varepsilon_*^3\right)\right)\left|A'^{-1}_*\right|_{[1,2]^b} \leqslant \left(1 + \cdot\varepsilon_*^3\right)d_3^0, \end{aligned} \tag{3.14}$$

where $D\omega_+ = \frac{d\omega_+}{dy^1}$, $D\omega_* = \frac{d\omega_*}{d\xi^*}$. Thus $N_+ \in NF(r_+, s_+, d_1^+, d_2^+, d_3^+)$, with $d_1^+ \equiv (1 + \frac{1}{2}\varepsilon_*^{3+\frac{1}{a_0}})d_1^0$, $d_2^+ \equiv d_2^0$, $d_3^+ \equiv (1 + \cdot\varepsilon_*^3)d_3^0$. Since $|(y^+B)_n|_{\Lambda_+} < \cdot n^{-1}|y^+|_{\Lambda_+} < \cdot n^{-1}\varepsilon$, $|(y^1 B)_n|_{\Lambda_1} < \cdot n^{-1}|y^1|_{\Lambda_1} < \cdot n^{-1}\varepsilon_*^{\frac{1}{a_0}}$, $|P^{n,0}_{0011}|_{\mathcal{O}_0} \leqslant n^{-1}\varepsilon_0$, $|P^n_{0011}|_\Lambda \leqslant n^{-1}\varepsilon$, we have $|\Omega^1_n - \Omega^*_n|_{\Lambda_1} < \cdot n^{-1}\varepsilon_*^{\frac{1}{a_0}}$, $|\Omega^+_n - \Omega_n|_{\Lambda_+} < \cdot n^{-1}\varepsilon$. Moreover, the upper bound $c_{\pi,m,b}$ of $|A^{-1}|$ in Remark 3.3 is d_2^0 in fact. As a result, $|A'_*| = |(\varepsilon_* A)^{-1}| \leqslant \varepsilon_*^{-1}d_2^0$, $y^1 \in [-(1+\varepsilon_*^2)d_1^0 d_2^0 \varepsilon_*^{3+\frac{1}{a_0}}, (1+\varepsilon_*^2)d_1^0 d_2^0 \varepsilon_*^{3+\frac{1}{a_0}}]^b$, $y^j \in [-d_2^0\varepsilon_{j-1}, d_2^0\varepsilon_{j-1}]^b$, $j = 2, \ldots, \infty$. In addition, d_3^0 can be chosen as $\varepsilon_*^{-1}d_2^0$.

3.3. Estimating the new frequency domain

Observe that

$$\begin{aligned} \left|\langle k, \tilde{\lambda}\omega_*\rangle\right| &\geqslant \frac{1}{2}\left|\langle k, \omega_*\rangle\right| \geqslant \frac{1}{2}\frac{\gamma_*}{|k|^\tau} \geqslant \frac{\gamma_+}{|k|^\tau}, \\ \left|\langle k, \tilde{\lambda}\omega_*\rangle + \Omega^+_n\right| &\geqslant \left|\langle k, \tilde{\lambda}\omega_*\rangle + \Omega_n\right| - \left|\left(y^+B\right)_n\right| - \left|P^n_{0011}\right| \\ &\geqslant \frac{\gamma}{|k|^\tau} - \cdot n^{-1}\left|y^+\right| - n^{-1}\varepsilon \geqslant \frac{\gamma_+}{|k|^\tau}, \\ \left|\langle k, \tilde{\lambda}\omega_*\rangle + \Omega^+_n \pm \Omega^+_m\right| &\geqslant \left|\langle k, \tilde{\lambda}\omega_*\rangle + \Omega_n \pm \Omega_m\right| - \left|\left(y^+B\right)_n\right| \\ &\quad - \left|\left(y^+B\right)_m\right| - \left|P^n_{0011}\right| - \left|P^m_{0011}\right| \\ &\geqslant \frac{\gamma}{|k|^\tau} - 2\frac{\cdot|y^+| + \varepsilon}{\min\{|n|, |m|\}} \geqslant \frac{\gamma_+}{|k|^\tau}, \end{aligned}$$

if $\gamma_\nu \leqslant \frac{1}{2}\gamma_*$, $c_1\varepsilon_*^{\frac{1}{a_0}}|k|^\tau \leqslant \gamma_* - \gamma_1$, $c_1\varepsilon_\nu|k|^\tau \leqslant \gamma_\nu - \gamma_{\nu+1}$, for some $c_1 > 0$ and for all $\nu \geqslant 1$. Thus, in the succeeding KAM step, small divisor conditions are automatically satisfied for $|k| \leqslant K$, if for all $\nu \geqslant 1$,

$$\gamma_\nu \leqslant \frac{1}{2}\gamma_*, \qquad c_1\varepsilon_*^{\frac{1}{a_0}}K_0^\tau \leqslant \gamma_* - \gamma_1, \qquad c_1\varepsilon_\nu K_\nu^\tau \leqslant \gamma_\nu - \gamma_{\nu+1}. \tag{3.15}$$

In what follows, we consider some new domains. Let $\eta=\varepsilon^{\frac{1}{3}}$, $D^{\nu}_{j\eta}=D_{j\eta}=D(r_+ + j\sigma, \frac{j}{4}\eta s)$, where $\sigma=\frac{1}{4}(r-r_+)$, $0<j\leqslant 4$, $D_+=D(r_+,s_+)$, $s_+=\frac{1}{8}\eta s$, $D=D(r,s)$. It is clear that $D_+\subset D_{j\eta}\subset D_j\subset D$.

3.4. Estimating the coordinate transformation

Lemma 3.2. *If $\varepsilon\ll(\frac{1}{2}\gamma^2(r-r_+)^{2\tau+3})^3$, we then have*

$$\phi_+\circ\Phi^t_F: D_{2\eta}\times\Lambda_+\to D,\quad -1\leqslant t\leqslant 1.$$

Moreover, denote by $\Phi^+=\phi_+\circ\Phi^1_F$, then for all $\nu\geqslant 1$, we have

$$\left\|\Phi^1-id\right\|_{D^0_{1\eta},\Lambda_1}<\cdot\varepsilon_*^{\frac{1}{a_0}},\qquad \left\|\Phi^+-id\right\|_{D_{1\eta},\Lambda_+}<\cdot\varepsilon_*^{-\frac{1}{12}}\varepsilon,$$

$$\left\|D\Phi^1-Id\right\|_{D^0_{1\eta},\Lambda_1}<\cdot\varepsilon_*^{\frac{1}{a_0}},\qquad \left\|D\Phi^+-Id\right\|_{D_{1\eta},\Lambda_+}<\cdot\varepsilon.$$

Proof. Denote by Φ^t_{F1}, Φ^t_{F2}, Φ^t_{F3}, Φ^t_{F4} the components of Φ^t_F in x, y, w, $\overline{w}$ planes respectively, setting $\beta=\gamma^{-4}(r-r_+)^{-(4\tau+5)}\varepsilon$, by virtue of

$$\Phi^t_F=id+\int_0^t X_F\circ\Phi^s_F\,\mathrm{d}s,\quad X_F=\big(F_y,-F_x,\{iF_{w_n}\}_{n\in\mathbb{N}_1},\{-iF_{\overline{w}_n}\}_{n\in\mathbb{N}_1}\big),$$

we have the estimates

$$\left\|\Phi^t_F-id\right\|_{D_{2\eta},\Lambda_+}\leqslant\|X_F\|_{D_3,\Lambda}<\cdot\beta<1,$$

$$\left|\Phi^t_{F1}\right|_{D_{2\eta},\Lambda_+}\leqslant|x|+\left|\int_0^t F_y\circ\Phi^s_F\,\mathrm{d}s\right|<r_++2\sigma+\cdot s\beta\leqslant r_++3\sigma,$$

$$\left|\Phi^t_{F2}\right|_{D_{2\eta},\Lambda_+}\leqslant|y|+\left|-\int_0^t F_x\circ\Phi^s_F\,\mathrm{d}s\right|<\frac{1}{2}\eta s+\cdot s^2\beta\leqslant\frac{3}{4}\eta s,$$

$$\left|\Phi^t_{F3}\right|_{D_{2\eta},\Lambda_+}\leqslant|w|+\left|\int_0^t F_w\circ\Phi^s_F\,\mathrm{d}s\right|<\frac{1}{2}\eta s+\cdot s\beta\leqslant\frac{3}{4}\eta s,$$

$$\left|\Phi^t_{F4}\right|_{D_{2\eta},\Lambda_+}\leqslant|\overline{w}|+\left|\int_0^t F_{\overline{w}}\circ\Phi^s_F\,\mathrm{d}s\right|<\frac{1}{2}\eta s+\cdot s\beta\leqslant\frac{3}{4}\eta s,$$

provided that

$$\varepsilon\ll\left(\frac{1}{2}\gamma^2(r-r_+)^{2\tau+3}\right)^3,\tag{3.16}$$

this means that $\Phi_F^t : D_{2\eta} \to D_{3\eta}, -1 \leqslant t \leqslant 1$. Due to $|y^1| = O(\varepsilon_*^{\frac{1}{a_0}})$, $|y^+| = O(\varepsilon)$, we know that $\phi_+ : D_{3\eta} \to D$ is well defined, thus, $\Phi^+ = \phi_+ \circ \Phi_F^1 : D_{2\eta} \times \Lambda_+ \to D$ is also well defined. Choosing $\gamma_* = \varepsilon_*^{\frac{1}{48}}$, $\gamma_+ = \frac{1}{2}\gamma + \frac{1}{8}\gamma_0$, since

$$\left|\frac{dy^1}{d\lambda}\right|_{\Lambda_1} = \varepsilon_*^{3+\frac{1}{a_0}} \left|\omega_* A^{-1}\right|_{\mathcal{O}_0} < \cdot \varepsilon_*^{\frac{1}{a_0}}, \qquad \left|\frac{dy^+}{d\lambda}\right|_{\Lambda_+} = \left|\frac{dP_{0100}A^{-1}}{d\lambda}\right|_{\Lambda_+} < \cdot \varepsilon,$$

we have

$$\left|\frac{d\phi_1}{d\lambda}\right|_{D^0_{1\eta},\Lambda_1} < \cdot \varepsilon_*^{\frac{1}{a_0}}, \qquad \left|\frac{d\phi_+}{d\lambda}\right|_{D_{1\eta},\Lambda_+} < \cdot \varepsilon,$$

thus

$$D\Phi_F^t - Id = \int_0^t DX_F D\Phi_F^s \, ds = \int_0^t J\big(D^2 F\big) D\Phi_F^s \, ds$$

together with

$$\Phi^+ - id = \phi_+ \circ \big(\Phi_F^1 - id\big) + (\phi_+ - id)$$

imply that

$$\left\|\Phi^1 - id\right\|_{D^0_{1\eta},\Lambda_1} \leqslant |\phi_1|_{D^0_{1\eta},\Lambda_1} \|X_{F^0}\|_{D_3,\Lambda_1} + \left|y^1\right|_{\Lambda_1} < \cdot \varepsilon_*^{\frac{11}{12}} + \cdot \varepsilon_*^{\frac{1}{a_0}} < \cdot \varepsilon_*^{\frac{1}{a_0}},$$

$$\left\|\Phi^+ - id\right\|_{D_{1\eta},\Lambda_+} \leqslant |\phi_+|_{D_{1\eta},\Lambda_+} \|X_F\|_{D_3,\Lambda_+} + \left|y^+\right|_{\Lambda_+} < \cdot \varepsilon_*^{-\frac{1}{12}} \varepsilon + \cdot \varepsilon < \cdot \varepsilon_*^{-\frac{1}{12}} \varepsilon,$$

$$\left\|d\Phi^1 - Id\right\|_{D^0_{1\eta},\Lambda_1} \leqslant \left|\frac{d\phi_1}{d\lambda}\right|_{D^0_{1\eta},\Lambda_1} \left\|D^2 F^0\right\|_{D_2,\Lambda_1} + \left|\frac{dy^1}{d\lambda}\right|_{\Lambda_1} < \cdot \varepsilon_*^{\frac{1}{a_0}},$$

$$\left\|D\Phi^+ - Id\right\|_{D_{1\eta},\Lambda_+} \leqslant \left|\frac{d\phi_+}{d\lambda}\right|_{D_{1\eta},\Lambda_+} \left\|D^2 F\right\|_{D_2,\Lambda_+} + \left|\frac{dy^+}{d\lambda}\right|_{\Lambda_+} < \cdot \varepsilon,$$

for all $\nu \geqslant 1$. Therefore, Lemma 3.2 follows. □

3.5. Estimating the new perturbation

Since

$$P^+ = \int_0^1 \{G_t, F\} \circ \Phi_t^+ \, dt + \widetilde{P} \circ \Phi^+ + Q \circ \phi_+,$$

where $G_t = tR + (1-t)[R]$, $\Phi_t^+ = \phi_+ \circ \Phi_F^t$, $\Phi^+ = \phi_+ \circ \Phi_F^1$, $\widetilde{P} = P - P_0 - P_1 - P_2$, $Q = \{\frac{1}{2}\langle yA', y\rangle, F_2\}$, we have

$$X_{P^+} = \int_0^1 \left(\Phi_t^+\right)^* X_{\{G_t,F\}}\,\mathrm{d}t + \left(\Phi^+\right)^* X_{\widetilde{P}} + (\phi_+)^* X_Q\,.$$

It follows from Lemma 2.5 that

$$\begin{aligned}
\|X_{\{G_t,F\}}\|_{D_{2\eta},\Lambda_+} &\leqslant \cdot\,\sigma^{-1}\eta^{-2}\|X_R\|_{D_3,\Lambda_+}\|X_F\|_{D_3,\Lambda_+} \\
&\leqslant c_2\gamma^{-4}(r-r_+)^{-(4\tau+6)}\eta^{-2}\varepsilon^2, \\
\|X_{\widetilde{P}}\|_{D_{2\eta},\Lambda_+} &\leqslant c_3\eta\|X_P\|_{D,\Lambda_+} \leqslant c_3\eta\varepsilon, \\
\|X_Q\|_{D_{2\eta},\Lambda_+} &< c_4\gamma^{-4}(r-r_+)^{-(4\tau+5)}\eta^3\varepsilon.
\end{aligned}$$

Let $c = 3\max\{c_1, c_2, c_3, c_4\} > 0$ and $\varepsilon_+ = c\gamma^{-4}(r-r_+)^{-(4\tau+6)}\varepsilon^{\frac{4}{3}}$, we then have $\|X_{P^+}\|_{D_+,\Lambda_+} \leqslant \varepsilon_+$. Moreover, if $\varepsilon_0 = c_*\varepsilon_*$ is sufficiently small, then there exists a constant κ, with $1 < \kappa < \frac{4}{3}$, such that $\varepsilon_\nu = \varepsilon_*^{\kappa^\nu}$, for all $\nu \geqslant 1$. As a result, $y^j \in [-d_2^0\varepsilon_*^{\kappa^{j-1}}, d_2^0\varepsilon_*^{\kappa^{j-1}}]^b$, $|y^j| = o(\varepsilon_*)$, $j = 2, \dots, \infty$. This completes one step of KAM iterations.

4. Iterative lemma and convergence

4.1. Iterative lemma

For any given r, s, c_*, ε_*, $d_1^* = O(\varepsilon_*^{-3})$, $d_2^* = O(1)$, $d_3^* = \varepsilon_*^{-1}d_2^*$, $2 < a_0 < 48$, and for all $\nu \geqslant 1$, we define the following sequences:

$$r_\nu = r_0\left(1 - \sum_{i=2}^{\nu+1} 2^{-i}\right), \quad r_0 = r,$$

$$s_\nu = \frac{1}{8}\varepsilon_{\nu-1}^{\frac{1}{3}}s_{\nu-1} = 2^{-3\nu}\left(\prod_{j=0}^{\nu-1}\varepsilon_j\right)^{\frac{1}{3}}s_0, \quad s_0 = s,$$

$$\gamma_\nu = \frac{1}{2}\gamma_0\left(1 - \sum_{i=2}^{\nu+1} 2^{-i}\right), \quad \gamma_0 = \gamma_* = \varepsilon_*^{\frac{1}{48}},$$

$$\varepsilon_\nu = c\gamma_{\nu-1}^{-4}(r_{\nu-1} - r_\nu)^{-(4\tau+6)}\varepsilon_{\nu-1}^{\frac{4}{3}}, \quad \varepsilon_0 = c_*\varepsilon_*,$$

$$d_1^\nu = \left(1 + \frac{1}{2}\varepsilon_*^{3+\frac{1}{a_0}}\right)d_1^0, \qquad d_2^\nu = d_2^0, \qquad d_3^\nu = \left(\varepsilon_*^{-1} + \cdot\varepsilon_*^2\right)d_2^0,$$

$$d_j^0 = d_j^*, \quad j = 1, 2, 3,$$

$$K_\nu = 2^\nu K_0^\tau, \quad K_0 = \varepsilon_*^{-\frac{1}{l(a_0)\tau}}, \quad \frac{48a_0}{48 - a_0} < l(a_0) < \infty,$$

$$D_{\nu-1} = D(r_{\nu-1}, s_{\nu-1}), \qquad D_{\nu-1}^j = D\left(r_\nu + \frac{j}{4}(r_{\nu-1} - r_\nu), 2js_\nu\right), \quad j = 2, 3,$$

$$\Lambda_\nu = \left\{\lambda \in \Lambda_{\nu-1} \subset \left[-\frac{1}{2}, \frac{1}{2}\right]\colon \left|\langle k, \tilde{\lambda}\omega_*\rangle + \left\langle l, \Omega^\nu\right\rangle\right| \geqslant \frac{\gamma_\nu}{|k|^\tau},\ |k| + |l| \neq 0, \right.$$
$$\left. \tilde{\lambda} = 1 + \varepsilon_*^{3+\frac{1}{a_0}}\lambda\right\}, \quad \Lambda_0 = \{0\}.$$

Note that

$$\Theta(r_0)=\prod_{j=1}^{\infty}\big[(r_{j-1}-r_j)^{-(4\tau+6)}\big]^{(\frac{3}{4})^j}$$

is a well-defined finite function of r_0. Since $\varepsilon_+ \ll \frac{c}{4}\varepsilon^{\frac{2}{3}}$, by the choice of K_ν and γ_ν, we can easily verify that the conditions (3.15) and (3.16) hold automatically for $\nu \geqslant 1$ and $\nu \geqslant 0$, respectively.

We summary the preceding analysis as follows.

Lemma 4.1. *Suppose that for $\nu \geqslant 0$, $H_\nu = N_\nu + P^\nu$ is given on $D_\nu \times \mathcal{O}$, which is real analytic in $(x, y, w, \overline{w}) \in D_\nu$, and Whitney smooth in $\xi^* \in \mathcal{O}_0 \subset \mathcal{O} = [1, 2]^b$ for $\nu = 0$, Whitney smooth in $\lambda \in \Lambda_\nu \subset \mathcal{O} = [-\frac{1}{2}, \frac{1}{2}]$ for $\nu \geqslant 1$, where*

$$H_\nu = N_\nu + P^\nu = \langle\omega_\nu, y\rangle + \sum_{n\in\mathbb{N}_1}\Omega_n^\nu w_n\overline{w}_n + \frac{1}{2}\langle yA'_\nu, y\rangle + P^\nu\big(x, y, w, \overline{w}, \xi^*, \lambda, \varepsilon\big),$$

satisfying

$$\begin{aligned}
&N_\nu \in NF\big(r_\nu, s_\nu, d_1^\nu, d_2^\nu, d_3^\nu\big) \cap DC(\gamma_\nu, \tau), \quad \nu \geqslant 0,\\
&\omega_\nu = \omega_\nu(\lambda) = \big(1 + \varepsilon_*^{3+\frac{1}{a_0}}\lambda\big)\omega_*, \qquad A'_\nu = A'_\nu(\lambda), \quad \nu \geqslant 0,\\
&\big|\Omega_n^1(\lambda) - \Omega_n^*\big|_{\Lambda_1} < \cdot n^{-1}\varepsilon_*^{\frac{1}{a_0}}, \qquad \big|\Omega_n^\nu(\lambda) - \Omega_n^{\nu-1}(\lambda)\big|_{\Lambda_\nu} < \cdot n^{-1}\varepsilon_{\nu-1}, \quad \nu \geqslant 2,\\
&\|X_{P^0}\|_{D_0,\mathcal{O}_0} \leqslant \varepsilon_0, \qquad \|X_{P^\nu}\|_{D_\nu,\Lambda_\nu} \leqslant \varepsilon_\nu, \quad \nu \geqslant 1,
\end{aligned} \tag{4.1}$$

with $\lambda \in \Lambda_\nu$ for $\nu \geqslant 1$, and $\lambda \equiv 0$ for $\nu = 0$, where $\omega_* = \omega(\xi^*) = \varepsilon_*^{-3}\alpha + \xi^* A$, $\alpha = (\mu_{i_1}, \ldots, \mu_{i_b})$, $A = (\frac{6}{\pi}\cdot\frac{4-\delta_{kl}}{\mu_{i_k}\mu_{i_l}})_{1\leqslant k,l\leqslant b}$, $|\omega_*|_{[1,2]^b} \leqslant d_1^0$, $|A^{-1}|_{[1,2]^b} \leqslant d_2^0$, $\Omega^* = \Omega(\xi^*) = \varepsilon_*^{-3}\beta + \xi^* B$, $\beta = (\mu_n, \ldots)_{n\in\mathbb{N}_1}$, $B = (\frac{24}{\pi\mu_{i_k}\mu_n})_{1\leqslant k\leqslant b,\, n\in\mathbb{N}_1}$, $A'_0 = \varepsilon_* A$, $d_3^0 = \varepsilon_*^{-1}d_2^0$, then there exist a symplectic diffeomorphism $\Phi_{\nu+1} : D_\nu^2 \to D_\nu^3$, and a translation

$$\phi_{\nu+1} : D_\nu^3 \to D(r_\nu, s_\nu), \quad (x, y, w, \overline{w}) \mapsto \big(x, y + y^{\nu+1}, w, \overline{w}\big),$$

such that for $H_{\nu+1} = H_\nu \circ \phi_{\nu+1} \circ \Phi_{\nu+1} := H_\nu \circ \Phi^{\nu+1} = N_{\nu+1} + P^{\nu+1}$, the same assumptions (4.1) are satisfied with $\nu + 1$ in place of ν, for some $\frac{1}{4}\gamma_0 \leqslant \gamma_{\nu+1} \leqslant \frac{1}{2}\gamma_0$, where

$$\Lambda_{\nu+1} = \Lambda_\nu \setminus \Bigg(\bigcup_{|k|>K_\nu, l}\mathcal{R}_{k,l}^{\nu+1}(\gamma_{\nu+1})\Bigg),$$

$$\mathcal{R}_{k,l}^{\nu+1}(\gamma_{\nu+1}) = \bigg\{\lambda \in \Lambda_\nu\colon \big|\langle k, \tilde{\lambda}\omega_*\rangle + \big\langle l, \Omega^{\nu+1}\big\rangle\big| < \frac{\gamma_{\nu+1}}{|k|^\tau},\ \tilde{\lambda} = t(\lambda)\bigg\}.$$

Moreover, for all $\nu \geqslant 1$, and for some $1 < \kappa < \frac{4}{3}$, we have

$$\begin{aligned}
&\big\|\Phi^1 - id\big\|_{D_1,\Lambda_1} < \cdot\varepsilon_*^{\frac{1}{a_0}}, \qquad \big\|\Phi^{\nu+1} - id\big\|_{D_{\nu+1},\Lambda_{\nu+1}} \leqslant \cdot\varepsilon_*^{\kappa^\nu - \frac{1}{12}},\\
&\big\|D\Phi^1 - Id\big\|_{D_1,\Lambda_1} < \cdot\varepsilon_*^{\frac{1}{a_0}}, \qquad \big\|D\Phi^{\nu+1} - Id\big\|_{D_{\nu+1},\Lambda_{\nu+1}} \leqslant \cdot\varepsilon_*^{\kappa^\nu}.
\end{aligned}$$

4.2. Convergence

Let $\Psi_\nu(\lambda) = \Phi^1 \circ \cdots \circ \Phi^\nu$, where $\Phi^j = \phi_j \circ \Phi_j$, $1 \leqslant j \leqslant \nu$, and denote by $\Psi_0 = id$. Inductively, we have $\Psi_\nu : D_\nu \times \Lambda_\nu \to D_0$, such that for all $\nu \geqslant 1$,

$$H_\nu = H_0 \circ \Psi_\nu = N_\nu + P^\nu\big(x, y, w, \overline{w}, \xi^*, \lambda, \varepsilon\big).$$

Let $\Lambda^\star = \bigcap_{\nu \geqslant 1} \Lambda_\nu$, we can apply Lemma 4.1 to conclude that H_ν, N_ν, P^ν, ω_ν, Ω_n^ν, A'_ν converge uniformly on $D(\frac{r_0}{2}, 0) \times \Lambda^\star$ to H_∞, N_∞, P^∞, ω_∞, Ω_n^∞, A'_∞, respectively. Clearly, $\omega_\infty = (1 + \varepsilon_*^{3+\frac{1}{a_0}} \lambda^\infty)\omega_*$, $\lambda^\infty \in \Lambda^\star$, and

$$\begin{aligned} N_\infty &= \langle \omega_\infty, y\rangle + \sum_n \Omega_n^\infty w_n \overline{w}_n + \frac{1}{2}\big\langle y A'_\infty, y\big\rangle \\ &\in NF\left(\frac{r_0}{2}, 0, \left(1 + \frac{1}{2}\varepsilon_*^{3+\frac{1}{a_0}}\right) d_1^*, d_2^*, \big(\varepsilon_*^{-1} + \cdot\varepsilon_*^2\big) d_2^*\right) \cap DC\left(\frac{1}{4}\varepsilon_*^{\frac{1}{48}}, \tau\right). \end{aligned}$$

Denote by ϕ_H^t the flow of X_H, since $H_0 \circ \Psi_\nu = H_\nu$, we have

$$\phi_{H_0}^t \circ \Psi_\nu = \Psi_\nu \circ \phi_{H_\nu}^t. \tag{4.2}$$

Note that

$$\Psi_{j+1} - \Psi_j = \int_0^1 D\big(\Phi^1 \circ \cdots \circ \Phi^j\big)\big(id + \theta\big(\Phi^{j+1} - id\big)\big)\, \mathrm{d}\theta \big(\Phi^{j+1} - id\big).$$

Since, by Lemma 4.1, on $D(\frac{r_0}{2}, 0) \times \Lambda^\star$,

$$\begin{aligned} &\big|D\big(\Phi^1 \circ \cdots \circ \Phi^j\big)\big(id + \theta\big(\Phi^{j+1} - id\big)\big)\big| \\ &\quad \leqslant \prod_{i=1}^j \big|D\Phi^i\big(\Phi^{i+1} \circ \cdots \circ \Phi^j\big)\big(id + \theta\big(\Phi^{j+1} - id\big)\big)\big| \\ &\quad \leqslant \big(1 + \cdot\varepsilon_*^{\frac{1}{a_0}}\big)\big(1 + \cdot\varepsilon_*^{\kappa}\big) \cdots \big(1 + \cdot\varepsilon_*^{\kappa^{j-1}}\big) \leqslant e^{\cdot(\varepsilon_*^{\frac{1}{a_0}} + \sum_{1 \leqslant \iota \leqslant j-1} \varepsilon_*^{\kappa^\iota})} < e, \end{aligned}$$

we have

$$\begin{aligned} |\Psi_1 - \Psi_0|_{D(\frac{r_0}{2}, 0) \times \Lambda^\star} &= \big|\Phi^1 - id\big|_{D(\frac{r_0}{2}, 0) \times \Lambda^\star} < \cdot\varepsilon_*^{\frac{1}{a_0}}, \\ |\Psi_{j+1} - \Psi_j|_{D(\frac{r_0}{2}, 0) \times \Lambda^\star} &\leqslant \cdot e\big|\Phi^{j+1} - id\big|_{D(\frac{r_0}{2}, 0) \times \Lambda^\star} < \cdot\varepsilon_*^{\kappa^j - \frac{1}{12}}, \end{aligned}$$

for all $j = 1, 2, \ldots$. This shows that Ψ_ν converges uniformly on $D(\frac{r_0}{2}, 0) \times \Lambda^\star$, we denote by Ψ_∞ its limit. Then

$$\Psi_\infty = \Psi_0 + \sum_{j=0}^\infty (\Psi_{j+1} - \Psi_j) = id + \sum_{j=0}^\infty (\Psi_{j+1} - \Psi_j).$$

It follows that Ψ_∞ is real analytic in $(x, y, w, \overline{w})$ and uniformly close to the identity. In the same way, $D\Psi_\nu$ converges uniformly to $D\Psi_\infty$ on $D(\frac{r_0}{2}, 0) \times \Lambda^\star$. Therefore, we can pass the limit on both sides of (4.2) to conclude that

$$\phi^t_{H_0} \circ \Psi_\infty = \Psi_\infty \circ \phi^t_{H_\infty}, \tag{4.3}$$

and

$$\Psi_\infty : D\left(\frac{r_0}{2}, 0\right) \times \Lambda^\star \to D(r, s).$$

Since

$$\varepsilon_\nu = c\gamma_{\nu-1}^{-4}(r_{\nu-1} - r_\nu)^{-(4\tau+6)}\varepsilon_{\nu-1}^{\frac{4}{3}} \leqslant c^{-3}\big(2^{24}c^3c_*\Theta(r_0)\varepsilon_*^{\frac{3}{4}}\big)^{(\frac{4}{3})^\nu},$$

we conclude that

$$\|X_{P^\infty}\|_{D(\frac{r_0}{2},0)\times\Lambda^\star} \equiv 0.$$

It follows that

$$\phi^t_{H_0} \circ \Psi_\infty\big(\mathbb{T}^b \times \{\lambda^\infty\}\big) = \Psi_\infty \circ \phi^t_{N_\infty}\big(\mathbb{T}^b \times \{\lambda^\infty\}\big) = \Psi_\infty\big(\mathbb{T}^b \times \{\lambda^\infty\}\big)$$

on $D(\frac{r_0}{2}, 0)$, for all $\lambda^\infty \in \Lambda^\star$. Hence, $\Psi_\infty(\mathbb{T}^b \times \{\lambda^\infty\})$ is a b-dimensional embedded invariant torus of the original perturbed Hamiltonian system at $\lambda^\infty \in \Lambda^\star$.

5. Measure estimates

After $(\nu+1)$th KAM step, we get the frequencies mapping $\xi^* \mapsto (\omega(\xi^*), \Omega(\xi^*))$, $y^1 \mapsto (\omega_{\nu+1}(y^1), \Omega^{\nu+1}(y^1))$ and the matrix-valued mapping $\xi^* \mapsto A'(\xi^*)$, $y^1 \mapsto A'_{\nu+1}(y^1)$, $y^1 = y^1(\lambda)$, which satisfy conditions C_1, C_2, C_3 with respect to d^*_j and $d^{\nu+1}_j$, $j = 1, 2, 3$, respectively, where

$$\begin{aligned}
\omega\big(\xi^*\big) &= \varepsilon_*^{-3}\alpha + \xi^* A = \omega_*,\\
\Omega\big(\xi^*\big) &= \varepsilon_*^{-3}\beta + \xi^* B = \Omega^*,\\
A'\big(\xi^*\big) &= \varepsilon_* A = A'_*,\\
\omega_{\nu+1}\big(y^1\big) &= \omega_* + y^1 A + P^0_{0100} = \tilde{\lambda}\omega_*,\\
\Omega^{\nu+1}\big(y^1\big) &= \Omega^* + y^1 B - \big(P^1_{0100} + \cdots + P^\nu_{0100}\big)A^{-1}B + P^0_{0011} + \cdots + P^\nu_{0011},\\
A'_{\nu+1}\big(y^1\big) &= A'_* + P^0_{0200} + \cdots + P^\nu_{0200},
\end{aligned}$$

with $\xi^* \in [1, 2]^b$, $y^1 \in [-(1+\varepsilon_*^2)d_1^* d_2^* \varepsilon_*^{3+\frac{1}{a_0}}, (1+\varepsilon_*^2)d_1^* d_2^* \varepsilon_*^{3+\frac{1}{a_0}}]^b$, $\varepsilon_* = c_*^{-1}\varepsilon_0$, $\tilde{\lambda} = 1 + \varepsilon_*^{3+\frac{1}{a_0}}\lambda = t(\lambda)$, $\lambda \in \Lambda_\nu \subset \mathcal{E} = [-\frac{1}{2}, \frac{1}{2}]$. However, we need to exclude the resonant set

$$\mathcal{R}^{\nu+1} = \bigcup_{|k|\geqslant K_\nu, l} \mathcal{R}^{\nu+1}_{k,l} = \bigcup_{|k|\geqslant K_\nu} \left(\mathcal{R}^{\nu+1}_k \cup \mathcal{R}^{\nu+1}_{kn} \cup \mathcal{R}^{\nu+1}_{kpq}\right)$$

$$= \bigcup_{|k|\geqslant K_\nu, l} \left\{\lambda \in \Lambda_\nu\colon \left|\langle k, \tilde{\lambda}\omega_*\rangle + \langle l, \Omega^{\nu+1}\rangle\right| < \frac{\gamma_{\nu+1}}{|k|^\tau},\ \tilde{\lambda} = t(\lambda)\right\},$$

where

$$\Lambda_\nu = \left\{\lambda\colon \left|\langle k, \tilde{\lambda}\omega_*\rangle + \langle l, \Omega^\nu\rangle\right| \geqslant \frac{\gamma_\nu}{|k|^\tau},\ |k|+|l| \neq 0,\ \tilde{\lambda} = t(\lambda)\right\},$$

$$\mathcal{R}^{\nu+1}_k = \left\{\lambda \in \Lambda_\nu\colon \left|\langle k, \tilde{\lambda}\omega_*\rangle\right| < \frac{\gamma_{\nu+1}}{|k|^\tau}\right\},$$

$$\mathcal{R}^{\nu+1}_{kn} = \left\{\lambda \in \Lambda_\nu\colon \left|\langle k, \tilde{\lambda}\omega_*\rangle \pm \Omega^{\nu+1}_n\right| < \frac{\gamma_{\nu+1}}{|k|^\tau}\right\},$$

$$\mathcal{R}^{\nu+1}_{kpq} = \left\{\lambda \in \Lambda_\nu\colon \left|\langle k, \tilde{\lambda}\omega_*\rangle \pm \left(\Omega^{\nu+1}_p \pm \Omega^{\nu+1}_q\right)\right| < \frac{\gamma_{\nu+1}}{|k|^\tau}\right\}.$$

Lemma 5.1. *When $k \neq 0$, $\langle l, \Omega^{\nu+1}\rangle = 0$, we have $\mathcal{R}^{\nu+1}_{k,l} = \emptyset$.*

Proof. Obviously, $|\langle k, \tilde{\lambda}\omega_*\rangle + \langle l, \Omega^{\nu+1}\rangle| = |\tilde{\lambda}||\langle k, \omega_*\rangle| \geqslant \frac{1}{2}\frac{\gamma_*}{|k|^\tau} \geqslant \frac{\gamma_{\nu+1}}{|k|^\tau}$. Hence, we complete the proof. □

Remark 5.1. Lemma 5.1 implies that for $k \neq 0$, $\mathcal{R}^{\nu+1}_k = \emptyset$ and

$$\mathcal{R}^{\nu+1}_{kpq} = \left\{\lambda \in \Lambda_\nu\colon \left|\langle k, \tilde{\lambda}\omega_*\rangle \pm \left(\Omega^{\nu+1}_p + \Omega^{\nu+1}_q\right)\right| < \frac{\gamma_{\nu+1}}{|k|^\tau}\right\}$$

$$\cup \left\{\lambda \in \Lambda_\nu\colon \left|\langle k, \tilde{\lambda}\omega_*\rangle \pm \left(\Omega^{\nu+1}_p - \Omega^{\nu+1}_q\right)\right| < \frac{\gamma_{\nu+1}}{|k|^\tau},\ p \neq q\right\}.$$

In the following, we consider the set

$$\widehat{\mathcal{R}}^{\nu+1} = \bigcup_{|k|\geqslant K_\nu, l} \left\{\lambda \in \Lambda_\nu\colon \left|\langle k, \omega_*\rangle + \sigma\langle l, \Omega^{\nu+1}\rangle\right| < \frac{\gamma_{\nu+1}}{(1 - \frac{1}{2}\varepsilon_*^{3+\frac{1}{a_0}})|k|^\tau},\ \sigma = \frac{1}{t(\lambda)}\right\}.$$

Obviously, $\mathcal{R}^{\nu+1} \subset \widehat{\mathcal{R}}^{\nu+1}$. Since $\lambda \to \frac{1}{t(\lambda)}$ is a diffeomorphism between $\mathcal{E} = [-\frac{1}{2}, \frac{1}{2}]$ and $1 + \mathcal{F} := [1 - O(\varepsilon_*^{3+\frac{1}{a_0}}), 1 + O(\varepsilon_*^{3+\frac{1}{a_0}})]$, we just need to consider an auxiliary resonant set

$$\widetilde{\mathcal{R}}^{\nu+1} = \bigcup_{|k|\geqslant K_\nu, l} \left\{\sigma = \frac{1}{t(\lambda)} \in \Sigma_\nu\colon \left|\langle k, \omega_*\rangle + \sigma\langle l, \Omega^{\nu+1}\rangle\right| < \frac{\gamma_{\nu+1}}{(1 - \frac{1}{2}\varepsilon_*^{3+\frac{1}{a_0}})|k|^\tau}\right\}$$

$$= \bigcup_{|k|\geqslant K_\nu} \left(\widetilde{\mathcal{R}}^{\nu+1}_{kn} \cup \widetilde{\mathcal{R}}^{\nu+1}_{kpq}\right) \subset 1 + \mathcal{F},$$

where

$$\Sigma_\nu = \left\{\sigma\colon \left|\langle k, \omega_*\rangle + \sigma\left\langle l, \Omega^\nu\right\rangle\right| \geqslant \frac{\gamma_\nu}{(1-\frac{1}{2}\varepsilon_*^{3+\frac{1}{a_0}})|k|^\tau},\ |k|+|l| \neq 0\right\},$$

$$\widetilde{\mathcal{R}}_{kn}^{\nu+1} = \left\{\sigma \in \Sigma_\nu\colon \left|\langle k, \omega_*\rangle \pm \sigma\Omega_n^{\nu+1}\right| < \frac{\gamma_{\nu+1}}{(1-\frac{1}{2}\varepsilon_*^{3+\frac{1}{a_0}})|k|^\tau}\right\},$$

$$\widetilde{\mathcal{R}}_{kpq}^{\nu+1} = \left\{\sigma \in \Sigma_\nu\colon \left|\langle k, \omega_*\rangle \pm \sigma\left(\Omega_p^{\nu+1} + \Omega_q^{\nu+1}\right)\right| < \frac{\gamma_{\nu+1}}{(1-\frac{1}{2}\varepsilon_*^{3+\frac{1}{a_0}})|k|^\tau}\right\}$$

$$\cup \left\{\sigma \in \Sigma_\nu\colon \left|\langle k, \omega_*\rangle \pm \sigma\left(\Omega_p^{\nu+1} - \Omega_q^{\nu+1}\right)\right| < \frac{\gamma_{\nu+1}}{(1-\frac{1}{2}\varepsilon_*^{3+\frac{1}{a_0}})|k|^\tau},\ p \neq q\right\}.$$

Setting $V^{\nu+1} = \langle k, \omega_*\rangle + \sigma\langle l, \Omega^{\nu+1}\rangle$, we then have

$$\frac{dV^{\nu+1}}{d\sigma} = \left\langle l, \Omega^{\nu+1} + \sigma\frac{d[\sum_{j=1}^{\nu}(P_{0011}^j - P_{0100}^j A^{-1}B) + y^1 B]}{d\sigma}\right\rangle := \left\langle l, \Delta^{\nu+1}\right\rangle.$$

Lemma 5.2. *For fixed $\nu \geqslant 0$, $m > 0$, and for all $l \in \mathbb{Z}^\infty$, $1 \leqslant |l| \leqslant 2$, we have*

$$\left|\frac{dV^{\nu+1}}{d\sigma}\right| > \varepsilon_*^{-3}\frac{1}{4\sqrt{2}}\min\left\{1, \frac{3}{\sqrt{m}}\right\}.$$

Proof. From the Hamiltonian (2.3) and Section 3, we notice that $P_{0100}^0 = O(\varepsilon_*^3|\xi^*|^2) = O(\varepsilon_*^3)$, $P_{0011}^0 = O(\varepsilon_*^3|\xi^*|^2) = O(\varepsilon_*^3), \ldots, P_{0100}^\nu = O(\varepsilon_*^{\kappa^\nu})$, $P_{0011}^{\nu,n} = O(n^{-1}\varepsilon_\nu) = O(n^{-1}\varepsilon_*^{\kappa^\nu})$, $1 < \kappa < \frac{4}{3}, \ldots.$ We can draw the conclusion from the following two cases.

Case 1. $|l| = 1$, we have

$$\left|\Delta_n^{\nu+1}\right| = \left|\varepsilon_*^{-3}\sqrt{n^2+m} + \left(\xi^* B + y^1 B\right)_n + \sum_{j=1}^{\nu}\left(P_{0011}^j - P_{0100}^j A^{-1}B\right)_n + P_{0011}^{0,n} + \sigma\frac{d[\sum_{j=1}^{\nu}(P_{0011}^j - P_{0100}^j A^{-1}B) + y^1 B]_n}{d\sigma}\right| \geqslant \frac{1}{2}\varepsilon_*^{-3}.$$

Case 2. $|l| = 2$, if the two nonzero components of l with the same sign, we have

$$\left|\Delta_p^{\nu+1} + \Delta_q^{\nu+1}\right| \geqslant \frac{1}{2}\varepsilon_*^{-3}\left|\sqrt{p^2+m} + \sqrt{q^2+m}\right| > \frac{1}{2}\varepsilon_*^{-3},$$

otherwise, we have $p \neq q$, such that

$$\left|\Delta_p^{\nu+1} - \Delta_q^{\nu+1}\right| \geqslant \frac{1}{2}\varepsilon_*^{-3}\left|\sqrt{p^2+m} - \sqrt{q^2+m}\right| = \frac{1}{2}\varepsilon_*^{-3}\frac{|p+q||p-q|}{\sqrt{p^2+m}+\sqrt{q^2+m}}.$$

Subcase a. $0 < m < \max\{p^2, q^2\}$, we have

$$\left|\Delta_p^{\nu+1} - \Delta_q^{\nu+1}\right| > \frac{1}{2}\varepsilon_*^{-3}\frac{p+q}{2\sqrt{2}\max\{p,q\}} > \frac{1}{4\sqrt{2}}\varepsilon_*^{-3}.$$

Subcase b. $m \geqslant \max\{p^2, q^2\}$, we have

$$\left|\Delta_p^{\nu+1} - \Delta_q^{\nu+1}\right| \geqslant \frac{1}{2}\varepsilon_*^{-3}\frac{p+q}{2\sqrt{2m}} \geqslant \frac{3}{4\sqrt{2m}}\varepsilon_*^{-3}.$$

As a consequence, we arrive at the conclusion. □

Lemma 5.3. *For fixed $\nu \geqslant 0$ and fixed $|k| \geqslant K_\nu$,*

$$\operatorname{meas}\left(\bigcup_l \widetilde{\mathcal{R}}_{k,l}^{\nu+1}\right) < \cdot\, \varepsilon_*^{\frac{3}{2}} \frac{\gamma_{\nu+1}^{\frac{1}{2}}}{|k|^{\frac{\tau}{2}-1}}.$$

Proof. We only consider the case of l with two nonzero components of opposite sign, which is the most complicated. Assume that $p - q = \varsigma \geqslant 1$. If $\varsigma > c'|k|$, then $\widetilde{\mathcal{R}}_{kpq}^{\nu+1} = \emptyset$, where c' is some constant large enough, independent of the iteration steps, if $1 \leqslant \varsigma \leqslant c'|k|$, since $|\Omega_n^{\nu+1} - \Omega_n^*|_{\Lambda_+} = |\sum_{1\leqslant j\leqslant \nu+1}(y^j B)_n + \sum_{0\leqslant j\leqslant \nu} P_{0011}^{j,n}|_{\Lambda_+} < \cdot\, \frac{1}{n}\varepsilon_*^{\frac{1}{q_0}}$, by the assumption (A3), we have

$$\widetilde{\mathcal{R}}_{kpq}^{\nu+1} \subseteq \widetilde{\mathcal{Q}}_{k\varsigma q}^{\nu+1} := \left\{\sigma\colon \left|\langle k, \omega_*\rangle \pm \varepsilon_*^{-3}\sigma\varsigma\right| < \frac{\gamma_{\nu+1}}{(1 - \frac{1}{2}\varepsilon_*^{3+\frac{1}{q_0}})|k|^\tau} + O\left(\varepsilon_*^{-3}q^{-1}\right)\right\}.$$

Due to Lemma 5.2, we get

$$\begin{aligned}\operatorname{meas}\left(\bigcup_{1\leqslant\varsigma\leqslant c'|k|}\ \bigcup_{p-q=\varsigma} \widetilde{\mathcal{R}}_{kpq}^{\nu+1}\right) &\leqslant \sum_{1\leqslant\varsigma\leqslant c'|k|}\left(\sum_{q<q_0}\operatorname{meas}\left(\widetilde{\mathcal{R}}_{kpq}^{\nu+1}\right) + \operatorname{meas}\left(\widetilde{\mathcal{Q}}_{k\varsigma q_0}^{\nu+1}\right)\right) \\ &< \cdot\, \varepsilon_*^3\left(\frac{\gamma_{\nu+1}q_0}{|k|^{\tau-1}} + O\left(\varepsilon_*^{-3}q_0^{-1}|k|\right)\right),\end{aligned}$$

by choosing $\frac{\gamma_{\nu+1}q_0}{|k|^{\tau-1}} = \varepsilon_*^{-3}q_0^{-1}|k|$, i.e., $q_0 = (\frac{|k|^\tau}{\varepsilon_*^3\gamma_{\nu+1}})^{\frac{1}{2}}$, we then arrive at

$$\operatorname{meas}\left(\bigcup_{1\leqslant\varsigma\leqslant c'|k|}\ \bigcup_{p-q=\varsigma} \widetilde{\mathcal{R}}_{kpq}^{\nu+1}\right) < \cdot\, \varepsilon_*^{\frac{3}{2}} \frac{\gamma_{\nu+1}^{\frac{1}{2}}}{|k|^{\frac{\tau}{2}-1}},$$

and the proof is finished. □

Lemma 5.4.

$$\operatorname{meas}\left(\bigcup_{\nu\geqslant 0}\widetilde{\mathcal{R}}^{\nu+1}\right) = \operatorname{meas}\left(\bigcup_{\nu\geqslant 0}\ \bigcup_{|k|\geqslant K_\nu, l}\widetilde{\mathcal{R}}_{k,l}^{\nu+1}\right) < \varepsilon_*^{\frac{1}{100}}\operatorname{meas}(1+\mathcal{F}).$$

Proof.

$$\operatorname{meas}\Bigl(\bigcup_{\nu\geqslant 0}\bigcup_{|k|>K_\nu,l}\widetilde{\mathcal{R}}^{\nu+1}_{k,l}\Bigr)<\cdot\varepsilon_*^{\frac{3}{2}}\sum_{\nu\geqslant 0}\sum_{|k|>K_\nu}\frac{\gamma_{\nu+1}^{\frac{1}{2}}}{|k|^{\frac{\tau}{2}-1}}=\cdot\varepsilon_*^{\frac{3}{2}}\sum_{\nu\geqslant 0}\sum_{l>K_\nu}\frac{\gamma_{\nu+1}^{\frac{1}{2}}}{l^{\frac{\tau}{2}-1}}l^{b-1}$$

$$<\cdot\varepsilon_*^{\frac{3}{2}}\sum_{\nu\geqslant 0}\frac{\gamma_{\nu+1}^{\frac{1}{2}}}{(2^\nu\varepsilon_*^{-\frac{1}{l(a_0)}})^{\frac{\tau}{2}-b}}<\cdot\varepsilon_*^{\frac{3}{2}+\frac{1}{96}+(\frac{\tau}{2}-b)\frac{1}{l(a_0)}},$$

where we have used that $\gamma_\nu\leqslant\frac{1}{2}\gamma_*$, $\gamma_*=\varepsilon_*^{\frac{1}{48}}$, for all $\nu\geqslant 1$. Owing to $2<a_0<48$, and $\tau>2b+(3+\frac{2}{a_0})l(a_0)$, we have

$$\operatorname{meas}\Bigl(\bigcup_{\nu\geqslant 0}\widetilde{\mathcal{R}}^{\nu+1}\Bigr)<\varepsilon_*^{\frac{\frac{\tau}{2}-b}{l(a_0)}-\frac{3}{2}-\frac{1}{a_0}+\frac{1}{100}}\operatorname{meas}(1+\mathcal{F})<\varepsilon_*^{\frac{1}{100}}\operatorname{meas}(1+\mathcal{F}).$$

This completes the proof. □

This means that the total measure of all excluded parameters in $1+\mathcal{F}$ can be as small as we wish. Since $\frac{1}{\sigma}=\tilde{\lambda}=1+\varepsilon_*^{3+\frac{1}{a_0}}\lambda$, we know that $\operatorname{meas}(\bigcup_{\nu\geqslant 0}\mathcal{R}^{\nu+1})/\operatorname{meas}(\mathcal{E})$ can be as small as we wish, thus, we obtain a positive-measure Cantor subset $\Lambda^\star$ of $\mathcal{E}$, such that $\Psi_\infty(\mathbb{T}^b\times\{(\xi^*,\lambda^\infty)\})$ is an embedded invariant torus of the original perturbed Hamiltonian system at $(\xi^*,\lambda^\infty)\in\mathcal{O}_0\times\Lambda^\star$, where $\mathcal{O}_0$ is a positive-measure Cantor subset of $[1,2]^b$. Let $\hat{\xi}^*=\varepsilon_*^3\xi^*$, then $\hat{\xi}^*\in\widetilde{\mathcal{O}}_0$ (a positive-measure Cantor subset of $[\varepsilon_*^3,2\varepsilon_*^3]^b$). Define $\hat{\omega}(\hat{\xi}^*)=\hat{\omega}_*=\alpha+\hat{\xi}^*A$, since $\omega(\xi^*)=\omega_*=\varepsilon_*^{-3}\alpha+\xi^*A$, we have $|\hat{\omega}_*|_{[\varepsilon_*^3,2\varepsilon_*^3]^b}=\varepsilon_*^3|\omega_*|_{[1,2]^b}\leqslant\varepsilon_*^3d_1^*$, $|(D\hat{\omega}(\hat{\xi}^*))^{-1}|_{[\varepsilon_*^3,2\varepsilon_*^3]^b}=|(D\omega(\xi^*))^{-1}|_{[1,2]^b}=|A^{-1}|\leqslant d_2^*$, thus, the tangential frequencies mapping $\hat{\xi}^*\to\hat{\omega}(\hat{\xi}^*)$ satisfies C_1,C_2 for $\hat{d}_1^*=\varepsilon_*^3d_1^*$, $\hat{d}_2^*=d_2^*$. At this time, $y^1\in[-(1+\varepsilon_*^2)\hat{d}_1^*d_2^*\varepsilon_*^{\frac{1}{a_0}},(1+\varepsilon_*^2)\hat{d}_1^*d_2^*\varepsilon_*^{\frac{1}{a_0}}]^b$. Since $|y^1|=O(\varepsilon_*^{\frac{1}{a_0}})$, $|y^j|=o(\varepsilon_*)$, $j=2,\dots,\infty$, at each $(\hat{\xi}^*,\lambda^\infty)\in\widetilde{\mathcal{O}}_0\times\Lambda^\star$, if we let $\omega^\star=\tilde{\lambda}^\infty\hat{\omega}_*=(1+\varepsilon_*^{3+\frac{1}{a_0}}\lambda^\infty)\hat{\omega}_*$, then Eq. (1.2) admits a small-amplitude quasi-periodic solution of the form

$$u(t,x)=\sum_{j=1}^{b}\sqrt{\hat{\xi}_j^*+\varepsilon_*^3y_j^1+\varepsilon_*^3y_j^2+\cdots+\varepsilon_*^3y_j^\infty}\cos(\omega_j^\star t)\sin i_jx+O\bigl(|\hat{\xi}^*|^{\frac{3}{2}}\bigr)$$

$$=\sum_{j=1}^{b}\sqrt{\hat{\xi}_j^*}\cos(\omega_j^\star t)\sin i_jx+O\bigl(|\hat{\xi}^*|^{\frac{3}{2}}\bigr).$$

From the above analysis, we complete the proof of the Main Theorem.

References

[1] J. Bourgain, On Melnikov's persistency problem, Math. Res. Lett. 4 (1997) 445–458.

[2] J. Bourgain, Construction of periodic solutions of nonlinear wave equations in higher dimension, Geom. Funct. Anal. 5 (1995) 629–639.

[3] L. Chierchia, J. You, KAM tori for 1D nonlinear wave equations with periodic boundary conditions, Comm. Math. Phys. 211 (2000) 498–525.

[4] L.H. Eliasson, Perturbations of stable invariant tori for Hamiltonian systems, Ann. Sc. Norm. Super. Pisa Cl. Sci. 15 (1988) 115–147.

[5] J. Geng, J. You, KAM tori of Hamiltonian perturbations of 1D linear beam equations, J. Math. Anal. Appl. 277 (2003) 104–121.
[6] J. Geng, J. You, A KAM theorem for one-dimensional Schrödinger equation with periodic boundary conditions, J. Differential Equations 209 (2005) 1–56.
[7] J. Geng, J. You, A KAM theorem for Hamiltonian partial differential equations in higher dimensional spaces, Comm. Math. Phys. 262 (2006) 343–372.
[8] J. Geng, J. You, KAM tori for higher dimensional beam equations with constant potentials, Nonlinearity 19 (2006) 2405–2423.
[9] S.B. Kuksin, Nearly Integrable Infinite-Dimensional Hamiltonian Systems, Lecture Notes in Math., vol. 1556, Springer-Verlag, Berlin, 1993.
[10] S.B. Kuksin, J. Pöschel, Invariant Cantor manifolds of quasi-periodic oscillations for a nonlinear Schrödinger equation, Ann. of Math. 143 (1996) 149–179.
[11] Y. Li, Y. Yi, Persistence of lower dimensional tori of general types in Hamiltonian systems, Trans. Amer. Math. Soc. 357 (2005) 1565–1600.
[12] Z. Liang, J. You, Quasi-periodic solutions for 1D Schrödinger equations with higher order nonlinearity, SIAM J. Math. Anal. 36 (2005) 1965–1990.
[13] V.K. Melnikov, On some cases of conservation of conditionally periodic motions under a small change of the Hamiltonian function, Soviet Math. Dokl. 6 (1965) 1592–1596.
[14] J. Pöschel, On elliptic lower dimensional tori in Hamiltonian systems, Math. Z. 202 (1989) 559–608.
[15] J. Pöschel, A KAM-theorem for some nonlinear partial differential equations, Ann. Sc. Norm. Super. Pisa Cl. Sci. 23 (1996) 119–148.
[16] J. Pöschel, Quasi-periodic solutions for nonlinear wave equations, Comment. Math. Helv. 71 (1996) 269–296.
[17] J. Xu, J. You, Persistence of lower-dimensional tori under the first Melnikov's non-resonance condition, J. Math. Pures Appl. 80 (2001) 1045–1067.
[18] J. You, Perturbations of lower-dimensional tori for Hamiltonian systems, J. Differential Equations 152 (1999) 1–29.