Available online at www.sciencedirect.com
ScienceDirect
Advances in Mathematics 226 (2011) 5361–5402
ADVANCES IN Mathematics
www.elsevier.com/locate/aim

An infinite dimensional KAM theorem and its application to the two dimensional cubic Schrödinger equation ☆

Jiansheng Geng *, Xindong Xu, Jiangong You

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

Received 1 December 2009; accepted 13 January 2011

Available online 26 January 2011

Communicated by Gang Tian

Abstract

We prove an infinite dimensional KAM theorem. As an application, we use the theorem to study the two dimensional nonlinear Schrödinger equation

$$\mathrm{i}u_t - \Delta u + |u|^2 u = 0, \quad t \in \mathbb{R},\ x \in \mathbb{T}^2$$

with periodic boundary conditions. We obtain for the equation a Whitney smooth family of small-amplitude quasi-periodic solutions corresponding to finite dimensional invariant tori of an associated infinite dimensional dynamical system.

MSC: 37K55; 35B10

Keywords: Schrödinger equation; KAM tori; Birkhoff normal form

☆ This work is partially supported by NSFC grants 10531050, 10771098 and 973 projects 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), xuxdnju@gmail.com (X. Xu), jyou@nju.edu.cn (J. You).

0001-8708/$ – see front matter
doi:10.1016/j.aim.2011.01.013

1. Introduction and main result

There have been many remarkable results in KAM (Kolmogorov–Arnold–Moser) theory of Hamiltonian PDEs achieved either by methods from the finite dimensional KAM theory [1,8, 10,11,13–15,12,16–26,28], or by a Newtonian scheme developed by Craig, Wayne, Bourgain [4, 3,5,7,6,2,9]. The advantage of the method from the finite dimensional KAM theory is the construction of a local normal form in a neighborhood of the obtained solutions in addition to the existence of quasi-periodic solutions. The normal form is helpful to understand the dynamics. For example, one sees the linear stability and zero Lyapunov exponents. The scheme of CWB avoids the cumbersome second Melnikov conditions by solving angle dependent homological equations. All those methods are well developed for one dimensional Hamiltonian PDEs. However, they meet difficulties in higher dimensional Hamiltonian PDEs. Bourgain [5] made the first breakthrough by proving that the two dimensional nonlinear Schrödinger equations admit small-amplitude quasi-periodic solutions. Later he improved in [7] his method and proved that the higher dimensional nonlinear Schrödinger and wave equations admit small-amplitude quasi-periodic solutions.

Constructing quasi-periodic solutions of higher dimensional Hamiltonian PDEs by method from the finite dimensional KAM theory appeared later. Geng and You [14,15] proved that the higher dimensional nonlinear beam equations and nonlocal Schrödinger equations admit small-amplitude linearly-stable quasi-periodic solutions. The breakthrough of constructing quasi-periodic solutions for more interesting higher dimensional Schrödinger equation by modified KAM method was made recently by Eliasson–Kuksin. They proved in [11] that the higher dimensional nonlinear Schrödinger equations admit small-amplitude linearly-stable quasi-periodic solutions.

However, all the above results on higher dimensional Schrödinger equation need artificial parameters, and hence do not apply to classical equations such as the higher dimensional cubic Schrödinger equation. To obtain quasi-periodic solutions of Hamiltonian PDEs with physical background such as the cubic Schrödinger equation, it is necessary to use the Birkhoff normal form techniques to get amplitude-frequency modulation. When the space dimension is greater than one, due to complicated resonances between the corresponding eigenvalues, it is difficult to get a nice integrable Birkhoff normal form. So far there are only two results available for the physical backgrounded higher dimensional Hamiltonian PDEs. In [5,6], Bourgain proved the existence of *two-frequency* quasi-periodic solutions for the two dimensional Schrödinger equation with constant potential

$$\mathrm{i}u_t - \Delta u + mu + u|u|^2 = 0. \tag{1.1}$$

More concretely, for two fixed distinguished lattice points $i_1, i_2 \in \mathbb{Z}^2$ on a circle

$$|i_1| = |i_2| = R, \quad i_1 \neq -i_2,$$

where $|\cdot|$ denotes Euclid-norm, Bourgain proved that (1.1) possesses quasi-periodic solutions of the form

$$u(t,x) = \sum_{j=1}^{2} \xi_j e^{\mathrm{i}(\omega_j t + \langle i_j, x\rangle)} + O\big(|\xi|^3\big)$$

with frequencies $\omega = (\omega_1, \omega_2)$ satisfying

$$\omega_j = |i_j|^2 + m + O\big(|\xi|^2\big), \quad j = 1, 2.$$

Here $\xi = (\xi_1, \xi_2)$ are in a Cantor set $\mathcal{O}$ of positive measure. As observed by Bourgain [5], the normal form analysis in the case of b sites with $b > 2$ involves additional difficulties, which leads to the generalization of his result widely open. In [15], Geng–You considered d dimensional nonlinear beam equations

$$u_{tt} + \Delta^2 u + \sigma u + f(u) = 0, \quad x \in \mathbb{T}^d,\ t \in \mathbb{R},$$
$$u(t, x_1 + 2\pi, \dots, x_d) = \cdots = u(t, x_1, \dots, x_d + 2\pi) = u(t, x_1, \dots, x_d),$$

where $\sigma \in \mathcal{I} \equiv [\sigma_1, \sigma_2]$ are parameters, and $f(u)$ is a real-analytic function near $u = 0$ with $f(0) = f'(0) = 0$. Then for carefully-chosen tangential sites $\{i_1, \dots, i_b\} \in \mathbb{Z}^d$, the above nonlinear beam equation admits a family of small-amplitude, linearly-stable quasi-periodic solution. Unfortunately the KAM theorem in [15] cannot be applied to the cubic Schrödinger equations.

In this paper, we will consider the two dimensional nonlinear Schrödinger equation

$$\mathrm{i}u_t - \Delta u + |u|^2 u = 0, \quad x \in \mathbb{T}^2,\ t \in \mathbb{R}, \tag{1.2}$$

with the periodic boundary conditions

$$u(t, x_1, x_2) = u(t, x_1 + 2\pi, x_2) = u(t, x_1, x_2 + 2\pi). \tag{1.3}$$

Eq. (1.2) is equivalent to (1.1) by a simple change of variables. We shall prove that the above equation admits a family of small-amplitude quasi-periodic solutions. Our results extend the Bourgain's existence result [5,6] to arbitrary finite dimensional invariant tori. We emphasize that, besides the existence of quasi-periodic solutions, we also get a nice linear normal form, which can be used to study the linear stability of the obtained solutions.

Now we state the main results of this paper. Let $\phi_n = \sqrt{\frac{1}{4\pi^2}} e^{\mathrm{i}\langle n, x\rangle}$ be the orthonormal eigenfunctions of operator $-\Delta$ with periodic boundary conditions (1.3), and $\lambda_n = |n|^2 = n_1^2 + n_2^2$, $n = (n_1, n_2) \in \mathbb{Z}^2$ the corresponding eigenvalues.

A finite set $S = \{i_1 = (x_1, y_1), \dots, i_b = (x_b, y_b)\} \subset \mathbb{Z}^2$ is called *admissible* if

1. *Any three of them are not vertices of a rectangle.*
2. *For any $n \in \mathbb{Z}^2 \backslash S$, there exists at most one triplet $\{i, j, m\}$ with $i, j \in S$, $m \in \mathbb{Z}^2 \backslash S$ such that $n - m + i - j = 0$ and $|i|^2 - |j|^2 + |n|^2 - |m|^2 = 0$. If such triplet exists, we say that n, m are resonant of the first type. By definition, n, m are mutually uniquely determined. We say that (n, m) is a resonant pair of the first type. Geometrically, (m, n, i, j) forms a rectangle with n, m being two adjacent vertices.*
3. *For any $n \in \mathbb{Z}^2 \backslash S$, there exists at most one triplet $\{i, j, m\}$ with $i, j \in S$, $m \in \mathbb{Z}^2 \backslash S$ such that $n + m - i - j = 0$ and $|n|^2 + |m|^2 - |i|^2 - |j|^2 = 0$. If such triplet exists, we say that n, m are resonant of the second type. By the definition n, m are mutually uniquely determined. We say that (n, m) is a resonant pair of the second type. Geometrically, (m, n, i, j) forms a rectangle with n, m being two diagonal vertices.*

4. *Any $n \in \mathbb{Z}^2 \backslash S$ is not resonant of both the first type and the second type, i.e., there exist no $i, j, f, g \in S$ and $m, m' \in \mathbb{Z}^2 \backslash S$, such that*

$$\begin{cases} n + m - i - j = 0, \\ |n|^2 + |m|^2 - |i|^2 - |j|^2 = 0, \\ n - m' + f - g = 0, \\ |n|^2 - |m'|^2 + |f|^2 - |g|^2 = 0. \end{cases}$$

Geometrically, any two of the above defined rectangles cannot share vertex in $\mathbb{Z}^2 \backslash S$.

In Appendix A, a concrete way of constructing the admissible set will be given. It is plausible that any randomly chosen set S is almost surely admissible.

Theorem 1. *Let $S = \{i_1, i_2, \dots, i_b\} \in \mathbb{Z}^2$ be an admissible set. There exists a Cantor set C of positive-measure such that for any $\xi = (\xi_1, \dots, \xi_b) \in C$, the nonlinear Schrödinger equation* (1.2) *with* (1.3) *admits a small-amplitude analytic quasi-periodic solution of the form*

$$u(t,x) = \sum_{j=1}^{b} \sqrt{\xi_j} e^{\mathrm{i}\omega_j t} \phi_{i_j} + O\big(|\xi|^{\frac{3}{2}}\big), \quad \omega_j = |i_j|^2 + O\big(|\xi|\big).$$

We shall prove the theorem by a KAM theorem given in Section 2. One knows that the KAM theory applies to perturbations of a nice normal form. The nice normal form is not only an important outcome of the KAM theory, but also a very important ingredient in the proof. For Hamiltonian systems without external parameters, one has to use normal form theory to put the Hamiltonian system into a small perturbation of a nice normal form (usually twisted and integrable). This would be difficult for the Hamiltonian system coming from the nonlinear Schrödinger equation (1.2) since the linear part is completely resonant. This difficulty is avoided in [5,6,11] by introducing external parameters.

Since the linear part of the cubic Schrödinger equation is completely resonant, an integrable normal form is not available to (1.2). Some θ-dependent quadratic terms $\sum_{|n| \neq |m|} P_{nm}(\theta) z_n \bar{z}_m$ will be kept in the normal form part, thus the KAM theorem in [11] cannot be applied directly to our case. Our strategy is to choose the tangential sites, to make the non-integrable terms in the normal form as sparse as possible so that the homological equations in KAM iteration is easy to be solved. Similar idea has been used in [15]. In the next section, we shall prove an infinite dimensional KAM theorem which allows sparsing θ-dependent terms in the normal form.

To prove the KAM theorem, we will incorporate with methods in [11] (Töplitz–Lipschitz property) and [27] (solving angle dependent homological equations). A major innovation in [11] is the introduction of the concept of Töplitz–Lipschitz property which allows them to deal with the measure estimate caused by $\langle k, \omega \rangle + \Omega_n - \Omega_m$. In this paper we shall use Eliasson–Kuksin's Töplitz–Lipschitz property at the conceptual level. Our proof is close to the standard KAM. Since the normal form part is much more simpler than Bourgain's although it depends on the angle variables θ, as in [27] the homological equations can be decomposed into a set of linear equations of dimension at most four. As a result, the homological equations are easier to solve in each KAM iteration steps. Finally, we give a few more remarks on Theorem 1.

Remark 1.1. The quasi-periodic solutions we obtained above are probably partially-hyperbolic. For example, if we choose the amplitudes ξ_1, ξ_2 such that $\xi_1^2+\xi_2^2 < 14\xi_1\xi_2$ for the two-frequency case, the corresponding quasi-periodic solutions are partially-hyperbolic.

Remark 1.2. Theorem 1 holds for more general two dimensional nonlinear Schrödinger equation

$$\mathrm{i}u_t - \Delta u + f\big(|u|^2\big)u = 0, \quad x \in \mathbb{T}^2,\ t \in \mathbb{R},$$

with periodic boundary conditions (1.3), f is a real analytic function in some neighborhood of the origin satisfying $f(0)=0,\ f'(0) \neq 0$. However three or higher dimensional nonlinear Schrödinger equation is significantly different from two dimensional case because it is difficult to construct the admissible tangential sites S.

The rest of the paper is organized as follows: We state an abstract infinite dimensional KAM theorem (Theorem 2) suitable for the application to two dimensional Schrödinger equation in Section 2; in Section 3, we prove Theorem 1 by using Theorem 2. In Section 4, Theorem 2 is proved. A concrete way of constructing the tangential sites is given in Appendix A.

2. An infinite dimensional KAM theorem

In this section, we give an infinite dimensional KAM theorem which allows a few θ dependent terms in the normal form part. The KAM can be applied to two dimensional Schrödinger equation with periodic boundary conditions.

We start by introducing some notations. For b vectors in $\mathbb{Z}^2$, say $\{i_1,\dots,i_b\}$, we denote $\mathbb{Z}_1^2 = \mathbb{Z}^2 \setminus \{i_1,\dots,i_b\}$. Let $z=(\dots,z_n,\dots)_{n\in\mathbb{Z}_1^2}$, and its complex conjugate $\bar z=(\dots,\bar z_n,\dots)_{n\in\mathbb{Z}_1^2}$. We introduce the weighted norm

$$\|z\|_\rho = \sum_{n\in\mathbb{Z}_1^2} |z_n| e^{|n|\rho},$$

where $|n| = \sqrt{n_1^2+n_2^2}$, $n=(n_1,n_2)$ and $\rho>0$. Denote a neighborhood of $\mathbb{T}^b \times \{I=0\} \times \{z=0\} \times \{\bar z = 0\}$ by

$$D_\rho(r,s) = \big\{(\theta,I,z,\bar z)\colon |\mathrm{Im}\,\theta| < r,\ |I| < s^2,\ \|z\|_\rho < s,\ \|\bar z\|_\rho < s\big\},$$

where $|\cdot|$ denotes the sup-norm of complex vectors. Moreover, we denote by $\mathcal{O}$ a positive-measure parameter set in $\mathbb{R}^b$.

Let $\alpha \equiv (\dots,\alpha_n,\dots)_{n\in\mathbb{Z}_1^2}$, $\beta \equiv (\dots,\beta_n,\dots)_{n\in\mathbb{Z}_1^2}$, α_n and $\beta_n \in \mathbb{N}$ with finitely many nonzero components of positive integers. The product $z^\alpha \bar z^\beta$ denotes $\prod_n z_n^{\alpha_n} \bar z_n^{\beta_n}$. Let

$$F(\theta,I,z,\bar z) = \sum_{\alpha,\beta} F_{\alpha\beta}(\theta,I) z^\alpha \bar z^\beta, \tag{2.1}$$

where $F_{\alpha\beta} = \sum_{k,l} F_{kl\alpha\beta} I^l e^{\mathrm{i}\langle k,\theta\rangle}$ are C_W^4 functions in parameter ξ in the sense of Whitney. Define the weighted norm of F by

$$\|F\|_{D_\rho(r,s),\mathcal{O}} \equiv \sup_{\substack{\|z\|_\rho<s \\ \|\bar z\|_\rho<s}} \sum_{\alpha,\beta} \|F_{\alpha\beta}\| \left|z^\alpha\right| \left|\bar z^\beta\right|, \tag{2.2}$$

where, if $F_{\alpha\beta} = \sum_{k\in\mathbb{Z}^b, l\in\mathbb{N}^b} F_{kl\alpha\beta}(\xi) I^l e^{\mathrm{i}\langle k,\theta\rangle}$ ($\langle\cdot,\cdot\rangle$ being the standard inner product), $\|F_{\alpha\beta}\|$ is short for

$$\|F_{\alpha\beta}\| \equiv \sum_{k,l} |F_{kl\alpha\beta}|_{\mathcal{O}} s^{2|l|} e^{|k|r}, \qquad |F_{kl\alpha\beta}|_{\mathcal{O}} \equiv \sup_{\xi\in\mathcal{O}} \sum_{0\leqslant d\leqslant 4} \left|\partial_\xi^d F_{kl\alpha\beta}\right| \tag{2.3}$$

(the derivatives with respect to ξ are in the sense of Whitney).

To a function F, we associate a Hamiltonian vector field defined by

$$X_F = \left(F_I, -F_\theta, \{\mathrm{i}F_{z_n}\}_{n\in\mathbb{Z}_1^2}, \{-\mathrm{i}F_{\bar z_n}\}_{n\in\mathbb{Z}_1^2}\right).$$

Its weighted norm is defined by[1]

$$\|X_F\|_{D_\rho(r,s),\mathcal{O}} \equiv \|F_I\|_{D_\rho(r,s),\mathcal{O}} + \frac{1}{s^2}\|F_\theta\|_{D_\rho(r,s),\mathcal{O}} + \frac{1}{s}\left(\sum_{n\in\mathbb{Z}_1^2} \|F_{z_n}\|_{D_\rho(r,s),\mathcal{O}} e^{|n|\rho} + \sum_{n\in\mathbb{Z}_1^2} \|F_{\bar z_n}\|_{D_\rho(r,s),\mathcal{O}} e^{|n|\rho}\right). \tag{2.4}$$

Suppose that S is an admissible set. Let $\mathcal{L}_1$ be the subset of $\mathbb{Z}_1^2$ with the following property: for each $n\in\mathcal{L}_1$, there exists a unique triplet (i,j,m) with $m\in\mathbb{Z}_1^2$, $i,j\in S$ such that

$$i - j + n - m = 0, \qquad |i|^2 - |j|^2 + |n|^2 - |m|^2 = 0.$$

In this case, we say that (n,m) is a resonant pair of the first type. $\mathcal{L}_1$ is composed of resonant pairs of the first type.

Let $\mathcal{L}_2$ be the subset of $\mathbb{Z}_1^2$ with the similar property: for each $n\in\mathcal{L}_2$, a unique triplet (i,j,m) with $m\in\mathbb{Z}_1^2$, $i,j\in S$ such that

$$-i - j + n + m = 0, \qquad -|i|^2 - |j|^2 + |n|^2 + |m|^2 = 0.$$

In this case, we say that (n,m) is a resonant pair of the second type. $\mathcal{L}_2$ is composed of finitely many resonant pairs of the second type. We assume that $\mathcal{L}_1 \cap \mathcal{L}_2 = \emptyset$.

We now describe the family of Hamiltonians studied in this paper. Let

[1] The norm $\|\cdot\|_{D_\rho(r,s),\mathcal{O}}$ for scalar functions is defined in (2.2). The vector function $G: D_\rho(r,s)\times\mathcal{O}\to\mathbb{C}^m$, ($m<\infty$) is similarly defined as $\|G\|_{D_\rho(r,s),\mathcal{O}} = \sum_{i=1}^m \|G_i\|_{D_\rho(r,s),\mathcal{O}}$.

$$
\begin{aligned}
H_0 &= N + \mathcal{A} + \mathcal{B} + \bar{\mathcal{B}}, \\
N &= \langle \omega(\xi), I \rangle + \sum_{n \in \mathbb{Z}_1^2} \Omega_n(\xi) z_n \bar{z}_n, \\
\mathcal{A} &= \sum_{n \in \mathcal{L}_1} a_n(\xi) z_n \bar{z}_m e^{\mathrm{i}(\theta_i - \theta_j)}, \\
\mathcal{B} &= \sum_{n \in \mathcal{L}_2} a_n(\xi) z_n z_m e^{\mathrm{i}(-\theta_i - \theta_j)}, \\
\bar{\mathcal{B}} &= \sum_{n \in \mathcal{L}_2} \bar{a}_n(\xi) \bar{z}_n \bar{z}_m e^{\mathrm{i}(\theta_i + \theta_j)},
\end{aligned} \tag{2.5}
$$

where $\xi \in \mathcal{O}$ is a parameter. Recall that (i, j) is uniquely determined by the corresponding resonant pair (n, m). Let $\mathcal{L}_1'$, $\mathcal{L}_2'$ be subsets of $\mathcal{L}_1$, $\mathcal{L}_2$ which contains one element in each resonant pair. We re-write the $\mathcal{A}$, $\mathcal{B}$ in resonant pairs by

$$
\begin{aligned}
\mathcal{A} &= \sum_{n \in \mathcal{L}_1'} \left\langle \begin{pmatrix} 0 & a_m \\ a_n & 0 \end{pmatrix} \begin{pmatrix} z_n e^{\mathrm{i}\theta_i} \\ z_m e^{\mathrm{i}\theta_j} \end{pmatrix}, \begin{pmatrix} \bar{z}_n e^{-\mathrm{i}\theta_i} \\ \bar{z}_m e^{-\mathrm{i}\theta_j} \end{pmatrix} \right\rangle, \\
\mathcal{B} &= \sum_{n \in \mathcal{L}_2'} \left\langle \begin{pmatrix} 0 & a_m \\ a_n & 0 \end{pmatrix} \begin{pmatrix} z_n e^{-\mathrm{i}\theta_i} \\ z_m e^{-\mathrm{i}\theta_j} \end{pmatrix}, \begin{pmatrix} z_n e^{-\mathrm{i}\theta_i} \\ z_m e^{-\mathrm{i}\theta_j} \end{pmatrix} \right\rangle.
\end{aligned} \tag{2.6}
$$

For each $\xi \in \mathcal{O}$, the Hamiltonian equations for H_0 are

$$
\begin{aligned}
\frac{d\theta}{dt} &= \omega, \\
\frac{dI}{dt} &= -\frac{\partial(\mathcal{A} + \mathcal{B} + \bar{\mathcal{B}})}{\partial \theta}, \\
\frac{d\binom{z_n}{z_m}}{dt} &= \mathrm{i} \begin{pmatrix} \Omega_n & a_m e^{\mathrm{i}(-\theta_i + \theta_j)} \\ a_n e^{\mathrm{i}(\theta_i - \theta_j)} & \Omega_m \end{pmatrix} \begin{pmatrix} z_n \\ z_m \end{pmatrix}, \quad n \in \mathcal{L}_1, \\
\frac{d\binom{z_n}{\bar{z}_m}}{dt} &= \mathrm{i} \begin{pmatrix} \Omega_n & \bar{a}_m e^{\mathrm{i}(\theta_i + \theta_j)} \\ -a_n e^{\mathrm{i}(-\theta_i - \theta_j)} & -\Omega_m \end{pmatrix} \begin{pmatrix} z_n \\ \bar{z}_m \end{pmatrix}, \quad n \in \mathcal{L}_2, \\
\frac{dz_n}{dt} &= \mathrm{i}\Omega_n z_n, \qquad \frac{d\bar{z}_n}{dt} = -\mathrm{i}\Omega_n \bar{z}_n, \quad n \in \mathbb{Z}_1^2 \setminus (\mathcal{L}_1 \cup \mathcal{L}_2).
\end{aligned} \tag{2.7}
$$

The system admits special solutions $(\theta, 0, 0, 0) \to (\theta + \omega t, 0, 0, 0)$ that corresponds to an invariant torus in the phase space. Consider now the perturbed Hamiltonian

$$
H = H_0 + P = N + \mathcal{A} + \mathcal{B} + \bar{\mathcal{B}} + P(\theta, I, z, \bar{z}, \xi). \tag{2.8}
$$

Our goal is to prove that, for most values of parameter $\xi \in \mathcal{O}$ (in Lebesgue measure sense), the Hamiltonians $H = N + \mathcal{A} + \mathcal{B} + \bar{\mathcal{B}} + P$ still admit invariant tori provided that $\|X_P\|_{D_\rho(r,s),\mathcal{O}}$ is sufficiently small. One should not expect a KAM theorem for general infinite dimensional Hamiltonian systems. So we consider Hamiltonian H satisfying the following hypotheses:

(A1) *Nondegeneracy:* The map $\xi \to \omega(\xi)$ is a C_W^4 diffeomorphism between $\mathcal{O}$ and its image.

(A2) *Asymptotics of normal frequencies:*

$$\Omega_n = \varepsilon^{-a}|n|^2 + \tilde{\Omega}_n, \quad a \geqslant 0 \tag{2.9}$$

where $\tilde{\Omega}_n$'s are C_W^4 functions of ξ with C_W^4-norm bounded by some small positive constant L.

(A3) *Melnikov's nondegeneracy:* Let $A_n = \Omega_n$ for $n \in \mathbb{Z}_1^2 \setminus (\mathcal{L}_1 \cup \mathcal{L}_2)$, and let

$$A_n = \begin{pmatrix} \Omega_n + \omega_i & a_n \\ a_m & \Omega_m + \omega_j \end{pmatrix}, \quad n \in \mathcal{L}_1',$$

$$A_n = \begin{pmatrix} \Omega_n - \omega_i & -a_n \\ \bar{a}_m & -\Omega_m + \omega_j \end{pmatrix}, \quad n \in \mathcal{L}_2',$$

where (n, m) are resonant pairs, (i, j) are uniquely determined by (n, m). We assume that $\omega(\xi)$, $A_n(\xi) \in C_W^4(\mathcal{O})$ and there exist $\gamma, \tau > 0$ (here I_2 is 2×2 identity matrix)

$$\bigl|\langle k, \omega\rangle\bigr| \geqslant \frac{\gamma}{|k|^\tau}, \quad k \neq 0,$$

$$\bigl|\det\bigl(\langle k, \omega\rangle I + A_n\bigr)\bigr| \geqslant \frac{\gamma}{|k|^\tau},$$

$$\bigl|\det\bigl(\langle k, \omega\rangle I \pm A_n \otimes I_2 \pm I_2 \otimes A_{n'}\bigr)\bigr| \geqslant \frac{\gamma}{|k|^\tau}, \quad k \neq 0.$$

(A4) *Regularity of $\mathcal{A} + \mathcal{B} + \bar{\mathcal{B}} + P$:* $\mathcal{A} + \mathcal{B} + \bar{\mathcal{B}} + P$ is real analytic in $I, \theta, q, \bar{q}$ and Whitney smooth in ξ; in addition

$$\|X_{\mathcal{A}}\|_{D_\rho(r,s),\mathcal{O}} + \|X_{\mathcal{B}}\|_{D_\rho(r,s),\mathcal{O}} < 1, \qquad \|X_P\|_{D_\rho(r,s),\mathcal{O}} < \varepsilon.$$

(A5) *Special form:* $\mathcal{A} + \mathcal{B} + \bar{\mathcal{B}} + P$ admits a special form of the following

$$\mathcal{D} = \biggl\{ \mathcal{A} + \mathcal{B} + \bar{\mathcal{B}} + P\colon \mathcal{A} + \mathcal{B} + \bar{\mathcal{B}} + P = \sum_{k \in \mathbb{Z}^b, l \in \mathbb{N}^b, \alpha, \beta} (\mathcal{A} + \mathcal{B} + \bar{\mathcal{B}} + P)_{kl\alpha\beta}(\xi) I^l e^{\mathrm{i}\langle k, \theta\rangle} z^\alpha \bar{z}^\beta \biggr\}$$

where k, α, β have the following relation

$$\sum_{j=1}^b k_j i_j + \sum_{n \in \mathbb{Z}_1^2} (\alpha_n - \beta_n) n = 0. \tag{2.10}$$

(A6) *Töplitz–Lipschitz property:* For any fixed $n, m \in \mathbb{Z}^2$, $c \in \mathbb{Z}^2 \setminus \{0\}$, the limits

$$\lim_{t\to\infty} \frac{\partial^2(\mathcal{B}+P)}{\partial z_{n+tc}\partial z_{m-tc}}, \qquad \lim_{t\to\infty} \frac{\partial^2(\sum_{n\in\mathbb{Z}_1^2} \tilde{\Omega}_n z_n \bar{z}_n + \mathcal{A} + P)}{\partial z_{n+tc}\partial \bar{z}_{m+tc}}, \qquad \lim_{t\to\infty} \frac{\partial^2(\bar{\mathcal{B}}+P)}{\partial \bar{z}_{n+tc}\partial \bar{z}_{m-tc}}$$

exist. Moreover, there exists $K > 0$, such that when $t > K$, $N + \mathcal{A} + \mathcal{B} + \bar{\mathcal{B}} + P$ satisfies

$$\left\| \frac{\partial^2(\mathcal{B}+P)}{\partial z_{n+tc}\partial z_{m-tc}} - \lim_{t\to\infty} \frac{\partial^2(\mathcal{B}+P)}{\partial z_{n+tc}\partial z_{m-tc}} \right\|_{D_\rho(r,s),\mathcal{O}} \leqslant \frac{\varepsilon}{t} e^{-|n+m|\rho},$$

$$\left\| \frac{\partial^2(\sum_{n\in\mathbb{Z}_1^2} \tilde{\Omega}_n z_n \bar{z}_n + \mathcal{A} + P)}{\partial z_{n+tc}\partial \bar{z}_{m+tc}} - \lim_{t\to\infty} \frac{\partial^2(\sum_{n\in\mathbb{Z}_1^2} \tilde{\Omega}_n z_n \bar{z}_n + \mathcal{A} + P)}{\partial z_{n+tc}\partial \bar{z}_{m+tc}} \right\|_{D_\rho(r,s),\mathcal{O}}$$
$$\leqslant \frac{\varepsilon}{t} e^{-|n-m|\rho},$$

$$\left\| \frac{\partial^2(\bar{\mathcal{B}}+P)}{\partial \bar{z}_{n+tc}\partial \bar{z}_{m-tc}} - \lim_{t\to\infty} \frac{\partial^2(\bar{\mathcal{B}}+P)}{\partial \bar{z}_{n+tc}\partial \bar{z}_{m-tc}} \right\|_{D_\rho(r,s),\mathcal{O}} \leqslant \frac{\varepsilon}{t} e^{-|n+m|\rho}.$$

Now we are ready to state an infinite dimensional KAM theorem.

Theorem 2. *Assume that the Hamiltonian* $N + \mathcal{A} + \mathcal{B} + \bar{\mathcal{B}} + P$ *in* (2.8) *satisfies* (A1)–(A6). *Let* $\gamma > 0$ *be small enough. Then there is a positive constant* ε*, depending on* $b, L, K, \tau, \gamma, r, s$ *and* ρ *such that if* $\|X_P\|_{D_\rho(r,s),\mathcal{O}} < \varepsilon$*, the following holds: There exist a Cantor subset* $\mathcal{O}_\gamma \subset \mathcal{O}$ *with* $\mathrm{meas}(\mathcal{O} \setminus \mathcal{O}_\gamma) = O(\gamma^{\frac{1}{4}})$ *and two maps (analytic in* θ *and* C_W^4 *in* ξ*)*

$$\Psi : \mathbb{T}^b \times \mathcal{O}_\gamma \to D_\rho(r,s), \qquad \tilde{\omega} : \mathcal{O}_\gamma \to \mathbb{R}^b,$$

where Ψ *is* $\frac{\varepsilon}{\gamma^4}$*-close to the trivial embedding* $\Psi_0 : \mathbb{T}^b \times \mathcal{O} \to \mathbb{T}^b \times \{0,0,0\}$ *and* $\tilde{\omega}$ *is* ε*-close to the unperturbed frequency* ω*, such that for any* $\xi \in \mathcal{O}_\gamma$ *and* $\theta \in \mathbb{T}^b$*, the curve* $t \to \Psi(\theta + \tilde{\omega}(\xi)t, \xi)$ *is a quasi-periodic solution of the Hamiltonian equations governed by* $H = N + \mathcal{A} + \mathcal{B} + \bar{\mathcal{B}} + P$.

3. Proof of Theorem 1

3.1. Hamiltonian and Birkhoff normal form

With scaling $u \to \varepsilon^{\frac{1}{2}} u$, we consider equation $\mathrm{i}u_t - \Delta u + \varepsilon |u|^2 u = 0$. The associated Hamiltonian is

$$H = \langle -\Delta u, u \rangle + \frac{\varepsilon}{2} \int_{\mathbb{T}^2} |u|^4 \, dx,$$

where $\langle \cdot,\cdot \rangle$ is the inner product in L^2. The operator $-\Delta$ under periodic boundary conditions (1.3) has a family of orthonormal eigenfunctions $\phi_n(x) = \sqrt{\frac{1}{4\pi^2}} e^{\mathrm{i}\langle n,x\rangle}$, $n \in \mathbb{Z}^2$ and the corresponding eigenvalues are $\lambda_n = |n|^2$. Let $u = \sum_{n\in\mathbb{Z}^2} q_n \phi_n(x)$, we have

$$H=\sum_{n\in\mathbb{Z}^2}\lambda_n|q_n|^2+\frac{\varepsilon}{8\pi^2}\sum_{i-j+n-m=0}q_i\bar{q}_jq_n\bar{q}_m. \tag{3.1}$$

For an admissible set of tangential sites S, we have a nice normal form for H.

Proposition 1. *Let S be admissible. For Hamiltonian function* (3.1)*, there is a symplectic transformation Ψ, such that*

$$H\circ\Psi=\langle\omega,I\rangle+\langle\Omega z,z\rangle+\mathcal{A}+\mathcal{B}+\bar{\mathcal{B}}+P \tag{3.2}$$

with

$$\begin{cases}\omega_i(\xi)=\varepsilon^{-4}|i|^2-\frac{1}{4\pi^2}\xi_i+\sum_{j\in S}\frac{1}{2\pi^2}\xi_j,\\ \Omega_n=\varepsilon^{-4}|n|^2+\sum_{j\in S}\frac{1}{2\pi^2}\xi_j,\end{cases}$$

$$\mathcal{A}=\frac{1}{2\pi^2}\sum_{n\in\mathcal{L}_1}\sqrt{\xi_i\xi_j}z_n\bar{z}_me^{\mathrm{i}\theta_i-\mathrm{i}\theta_j},$$

$$\mathcal{B}=\frac{1}{2\pi^2}\sum_{n\in\mathcal{L}_2}\sqrt{\xi_i\xi_j}z_nz_me^{-\mathrm{i}\theta_i-\mathrm{i}\theta_j},$$

$$\bar{\mathcal{B}}=\frac{1}{2\pi^2}\sum_{n\in\mathcal{L}_2}\sqrt{\xi_i\xi_j}\bar{z}_n\bar{z}_me^{\mathrm{i}\theta_i+\mathrm{i}\theta_j},$$

$$\begin{aligned}|P|=O\big(&\varepsilon^2|I|^2+\varepsilon^2|I|\|z\|_\rho^2+\varepsilon\xi^{\frac12}\|z\|_\rho^3+\varepsilon^2\|z\|_\rho^4+\varepsilon^2\xi^3\\&+\varepsilon^3\xi^{\frac52}\|z\|_\rho+\varepsilon^4\xi^2\|z\|_\rho^2+\varepsilon^5\xi^{\frac32}\|z\|_\rho^3\big).\end{aligned} \tag{3.3}$$

Proof. The proof consists of several symplectic change of variables. Firstly, let

$$F=\sum_{\substack{i-j+n-m=0\\|i|^2-|j|^2+|n|^2-|m|^2\neq0\\ \sharp S\cap\{i,j,n,m\}\geqslant2}}\frac{\mathrm{i}\varepsilon}{8\pi^2(\lambda_i-\lambda_j+\lambda_n-\lambda_m)}q_i\bar{q}_jq_n\bar{q}_m, \tag{3.4}$$

and X_F^1 be the time one map of the flow of the associated Hamiltonian systems. The change of variables X_F^1 sends H to

$$H\circ X_F^1=\sum_{i\in S}\lambda_i|q_i|^2+\sum_{i\in\mathbb{Z}_1^2}\lambda_i|z_i|^2+\sum_{i\in S}\frac{\varepsilon}{8\pi^2}|q_i|^4 \tag{3.5}$$

$$+\sum_{i,j\in S,i\neq j}\frac{\varepsilon}{2\pi^2}|q_i|^2|q_j|^2+\sum_{i\in S,j\in\mathbb{Z}_1^2}\frac{\varepsilon}{2\pi^2}|q_i|^2|z_j|^2 \tag{3.6}$$

$$+\sum_{n\in\mathcal{L}_1}\frac{\varepsilon}{2\pi^2}q_i\bar{q}_jz_n\bar{z}_m+\sum_{n\in\mathcal{L}_2}\frac{\varepsilon}{2\pi^2}(q_iq_j\bar{z}_n\bar{z}_m+\bar{q}_i\bar{q}_jz_nz_m) \tag{3.7}$$

$$+O\big(\varepsilon|q|\|z\|_\rho^3+\varepsilon\|z\|_\rho^4+\varepsilon^2|q|^6+\varepsilon^2|q|^5\|z\|_\rho+\varepsilon^2|q|^4\|z\|_\rho^2+\varepsilon^2|q|^3\|z\|_\rho^3\big).$$

We remind that (n,m) are resonant pairs and (i,j) is uniquely determined by (n,m) in (3.7).

Introducing the action-angle variable in the tangential space

$$q_j = \sqrt{I_j + \xi_j}\, e^{\mathrm{i}\theta_j}, \qquad \bar{q}_j = \sqrt{I_j + \xi_j}\, e^{-\mathrm{i}\theta_j}, \quad j \in S, \tag{3.8}$$

we have

$$\begin{aligned}
H \circ X_F^1 &= \sum_{i\in S} \lambda_i (I_i + \xi_i) + \sum_{i\in\mathbb{Z}_1^2} \lambda_i |z_i|^2 + \sum_{i\in S} \frac{\varepsilon}{8\pi^2} (I_i + \xi_i)^2 \\
&\quad + \frac{\varepsilon}{2\pi^2} \sum_{i,j\in S, i\neq j} (I_i + \xi_i)(I_j + \xi_j) + \frac{\varepsilon}{2\pi^2} \sum_{i\in S, j\in\mathbb{Z}_1^2} (I_i + \xi_i)|z_j|^2 \\
&\quad + \frac{\varepsilon}{2\pi^2} \sum_{n\in\mathcal{L}_1} \sqrt{(I_i + \xi_i)(I_j + \xi_j)}\, z_n \bar{z}_m e^{\mathrm{i}\theta_i - \mathrm{i}\theta_j} \\
&\quad + \frac{\varepsilon}{2\pi^2} \sum_{n\in\mathcal{L}_2} \sqrt{(I_i + \xi_i)(I_j + \xi_j)}\, z_n z_m e^{-\mathrm{i}\theta_i - \mathrm{i}\theta_j} \\
&\quad + \frac{\varepsilon}{2\pi^2} \sum_{n\in\mathcal{L}_2} \sqrt{(I_i + \xi_i)(I_j + \xi_j)}\, \bar{z}_n \bar{z}_m e^{\mathrm{i}\theta_i + \mathrm{i}\theta_j} \\
&\quad + O\big(\varepsilon \xi^{\frac12} \|z\|_\rho^3 + \varepsilon \|z\|_\rho^4 + \varepsilon^2 \xi^3 + \varepsilon^2 \xi^{\frac52} \|z\|_\rho + \varepsilon^2 \xi^2 \|z\|_\rho^2 + \varepsilon^2 \xi^{\frac32} \|z\|_\rho^3\big) \\
&= \sum_{i\in S} \lambda_i I_i + \sum_{i\in\mathbb{Z}_1^2} \lambda_i |z_i|^2 + \sum_{i\in S} \frac{\varepsilon}{4\pi^2} \xi_i I_i \\
&\quad + \frac{\varepsilon}{2\pi^2} \sum_{i,j\in S, i\neq j} \xi_i I_j + \frac{\varepsilon}{2\pi^2} \sum_{i\in S, j\in\mathbb{Z}_1^2} \xi_i |z_j|^2 \\
&\quad + \frac{\varepsilon}{2\pi^2} \sum_{n\in\mathcal{L}_1} \sqrt{\xi_i \xi_j}\, z_n \bar{z}_m e^{\mathrm{i}\theta_i - \mathrm{i}\theta_j} \\
&\quad + \frac{\varepsilon}{2\pi^2} \sum_{n\in\mathcal{L}_2} \sqrt{\xi_i \xi_j}\, z_n z_m e^{-\mathrm{i}\theta_i - \mathrm{i}\theta_j} \\
&\quad + \frac{\varepsilon}{2\pi^2} \sum_{n\in\mathcal{L}_2} \sqrt{\xi_i \xi_j}\, \bar{z}_n \bar{z}_m e^{\mathrm{i}\theta_i + \mathrm{i}\theta_j} \\
&\quad + O\big(\varepsilon |I|^2 + \varepsilon |I| \|z\|_\rho^2 + \varepsilon \xi^{\frac12} \|z\|_\rho^3 + \varepsilon \|z\|_\rho^4 + \varepsilon^2 \xi^3 \\
&\quad + \varepsilon^2 \xi^{\frac52} \|z\|_\rho + \varepsilon^2 \xi^2 \|z\|_\rho^2 + \varepsilon^2 \xi^{\frac32} \|z\|_\rho^3\big) \\
&= N + \mathcal{A} + \mathcal{B} + \bar{\mathcal{B}} + P,
\end{aligned}$$

where

$$N = \sum_{i\in S} \lambda_i I_i + \sum_{j\in\mathbb{Z}_1^2} \lambda_j |z_j|^2 - \sum_{i\in S} \frac{\varepsilon}{4\pi^2} \xi_i I_i + \sum_{i,j\in S} \frac{\varepsilon}{2\pi^2} \xi_i I_j + \sum_{i\in S, j\in\mathbb{Z}_1^2} \frac{\varepsilon}{2\pi^2} \xi_i |z_j|^2,$$

$$\mathcal{A} = \frac{\varepsilon}{2\pi^2} \sum_{n \in \mathcal{L}_1} \sqrt{\xi_i \xi_j} z_n \bar{z}_m e^{\mathrm{i}\theta_i - \mathrm{i}\theta_j},$$

$$\mathcal{B} = \frac{\varepsilon}{2\pi^2} \sum_{n \in \mathcal{L}_2} \sqrt{\xi_i \xi_j} z_n z_m e^{-\mathrm{i}\theta_i - \mathrm{i}\theta_j},$$

$$\bar{\mathcal{B}} = \frac{\varepsilon}{2\pi^2} \sum_{n \in \mathcal{L}_2} \sqrt{\xi_i \xi_j} \bar{z}_n \bar{z}_m e^{\mathrm{i}\theta_i + \mathrm{i}\theta_j},$$

$$|P| = O\big(\varepsilon |I|^2 + \varepsilon |I| \|z\|_\rho^2 + \varepsilon \xi^{\frac{1}{2}} \|z\|_\rho^3 + \varepsilon \|z\|_\rho^4 + \varepsilon^2 \xi^3 + \varepsilon^2 \xi^{\frac{5}{2}} \|z\|_\rho + \varepsilon^2 \xi^2 \|z\|_\rho^2 + \varepsilon^2 \xi^{\frac{3}{2}} \|z\|_\rho^3\big).$$

By the scaling in time

$$\xi \to \varepsilon^3 \xi, \quad I \to \varepsilon^5 I, \quad \theta \to \theta, \quad z \to \varepsilon^{\frac{5}{2}} z, \quad \bar{z} \to \varepsilon^{\frac{5}{2}} \bar{z}$$

we finally arrive at the rescaled Hamiltonian

$$H = \varepsilon^{-9} H\big(\varepsilon^3 \xi, \varepsilon^5 I, \theta, \varepsilon^{\frac{5}{2}} z, \varepsilon^{\frac{5}{2}} \bar{z}\big) = \langle \omega, I \rangle + \langle \Omega z, z \rangle + \mathcal{A} + \mathcal{B} + \bar{\mathcal{B}} + P, \tag{3.9}$$

where

$$\begin{cases} \omega_i(\xi) = \varepsilon^{-4} |i|^2 - \frac{1}{4\pi^2} \xi_i + \sum_{j \in S} \frac{1}{2\pi^2} \xi_j, \\ \Omega_n = \varepsilon^{-4} |n|^2 + \sum_{j \in S} \frac{1}{2\pi^2} \xi_j, \end{cases}$$

$$\mathcal{A} = \frac{1}{2\pi^2} \sum_{n \in \mathcal{L}_1} \sqrt{\xi_i \xi_j} z_n \bar{z}_m e^{\mathrm{i}\theta_i - \mathrm{i}\theta_j},$$

$$\mathcal{B} = \frac{1}{2\pi^2} \sum_{n \in \mathcal{L}_2} \sqrt{\xi_i \xi_j} z_n z_m e^{-\mathrm{i}\theta_i - \mathrm{i}\theta_j},$$

$$\bar{\mathcal{B}} = \frac{1}{2\pi^2} \sum_{n \in \mathcal{L}_2} \sqrt{\xi_i \xi_j} \bar{z}_n \bar{z}_m e^{\mathrm{i}\theta_i + \mathrm{i}\theta_j},$$

$$|P| = O\big(\varepsilon^2 |I|^2 + \varepsilon^2 |I| \|z\|_\rho^2 + \varepsilon \xi^{\frac{1}{2}} \|z\|_\rho^3 + \varepsilon^2 \|z\|_\rho^4 + \varepsilon^2 \xi^3 + \varepsilon^3 \xi^{\frac{5}{2}} \|z\|_\rho$$
$$+ \varepsilon^4 \xi^2 \|z\|_\rho^2 + \varepsilon^5 \xi^{\frac{3}{2}} \|z\|_\rho^3\big). \qquad \square$$

3.2. Verifying (A1)–(A6) for (3.2)

Verifying (A1): In view of (3.3), we have

$$\frac{\partial \omega}{\partial \xi} \triangleq A = \frac{1}{4\pi^2} \begin{pmatrix} 1 & 2 & \cdots & 2 \\ 2 & 1 & \cdots & 2 \\ \cdots & \cdots & \cdots & \cdots \\ 2 & 2 & \cdots & 1 \end{pmatrix}_{b \times b}. \tag{3.10}$$

It is easy to check that $\det A \neq 0$, thus (A1) is verified.

Verifying (A2): Take $a = 4$, the proof is obvious.

Verifying (A3): For (3.2), A_n read as follows

$$A_n = \Omega_n, \quad n \in \mathbb{Z}_1^2 \setminus (\mathcal{L}_1 \cup \mathcal{L}_2),$$
$$A_n = \begin{pmatrix} \Omega_n + \omega_i & \frac{1}{2\pi^2}\sqrt{\xi_i\xi_j} \\ \frac{1}{2\pi^2}\sqrt{\xi_i\xi_j} & \Omega_m + \omega_j \end{pmatrix}, \quad n \in \mathcal{L}_1',$$
$$A_n = \begin{pmatrix} \Omega_n - \omega_i & -\frac{1}{2\pi^2}\sqrt{\xi_i\xi_j} \\ \frac{1}{2\pi^2}\sqrt{\xi_i\xi_j} & -\Omega_m + \omega_j \end{pmatrix}, \quad n \in \mathcal{L}_2',$$

where (m, i, j) is uniquely determined by n. We only verify (A3) for $\det[\langle k, \omega\rangle I \pm A_n \otimes I_2 \pm I_2 \otimes A_{n'}]$ which is the most complicated. Let A, B be 2×2 matrices. We know that $\lambda I + A \otimes I - I \otimes B = (\lambda I + A) \otimes I - I \otimes B$. Moreover, we have

Lemma 3.1.

$$|A \otimes I \pm I \otimes B| = \big(|A| - |B|\big)^2 + |A|\big(\operatorname{tr}(B)\big)^2 + |B|\big(\operatorname{tr}(A)\big)^2 \pm \big(|A| + |B|\big)\operatorname{tr}(A)\operatorname{tr}(B)$$

where $|\cdot|$ denotes the determinant of the corresponding matrices.

Case 1. $n, n' \in \mathcal{L}_1$.

$$\begin{aligned}
&\langle k, \omega\rangle I \pm A_n \otimes I_2 \pm I_2 \otimes A_{n'} \\
&\quad = \operatorname{Diag}\begin{pmatrix} \langle k, \omega\rangle \pm (\Omega_n + \omega_i) \pm (\Omega_{n'} + \omega_{i'}) \\ \langle k, \omega\rangle \pm (\Omega_n + \omega_i) \pm (\Omega_{m'} + \omega_{j'}) \\ \langle k, \omega\rangle \pm (\Omega_m + \omega_j) \pm (\Omega_{n'} + \omega_{i'}) \\ \langle k, \omega\rangle \pm (\Omega_m + \omega_j) \pm (\Omega_{m'} + \omega_{j'}) \end{pmatrix} \\
&\qquad \pm \begin{pmatrix} 0 & \frac{1}{2\pi^2}\sqrt{\xi_i\xi_j} \\ \frac{1}{2\pi^2}\sqrt{\xi_i\xi_j} & 0 \end{pmatrix} \otimes I_2 \pm I_2 \otimes \begin{pmatrix} 0 & \frac{1}{2\pi^2}\sqrt{\xi_{i'}\xi_{j'}} \\ \frac{1}{2\pi^2}\sqrt{\xi_{i'}\xi_{j'}} & 0 \end{pmatrix}.
\end{aligned}$$

Set $\alpha = \varepsilon^{-4}(|i_1|^2, |i_2|^2, \ldots, |i_b|^2)$, $\xi = (\xi_{i_1}, \xi_{i_2}, \ldots, \xi_{i_b})$, $\beta = \frac{1}{4\pi^2}(2, 2, \ldots, 2)$, and notice that $|n|^2 + |i|^2 = |m|^2 + |j|^2$, $|n'|^2 + |i'|^2 = |m'|^2 + |j'|^2$. We have

$$\begin{aligned}
&\langle k, \omega\rangle I \pm A_n \otimes I_2 \pm I_2 \otimes A_{n'} \\
&\quad = \big(\langle k, \alpha\rangle \pm \varepsilon^{-4}\big(|n|^2 + |i|^2\big) \pm \varepsilon^{-4}\big(\big|n'\big|^2 + \big|i'\big|^2\big) + \langle Ak \pm 2\beta \pm 2\beta, \xi\rangle\big) I \\
&\qquad \pm \begin{pmatrix} -\frac{1}{4\pi^2}\xi_i & \frac{1}{2\pi^2}\sqrt{\xi_i\xi_j} \\ \frac{1}{2\pi^2}\sqrt{\xi_i\xi_j} & -\frac{1}{4\pi^2}\xi_j \end{pmatrix} \otimes I_2 \pm I_2 \otimes \begin{pmatrix} -\frac{1}{4\pi^2}\xi_{i'} & \frac{1}{2\pi^2}\sqrt{\xi_{i'}\xi_{j'}} \\ \frac{1}{2\pi^2}\sqrt{\xi_{i'}\xi_{j'}} & -\frac{1}{4\pi^2}\xi_{j'} \end{pmatrix}.
\end{aligned} \tag{3.11}$$

Its eigenvalues are

$$\begin{aligned}
&\langle k, \alpha\rangle \pm \varepsilon^{-4}\big(|n|^2 + |i|^2\big) \pm \varepsilon^{-4}\big(\big|n'\big|^2 + \big|i'\big|^2\big) + \langle Ak \pm 2\beta \pm 2\beta, \xi\rangle \\
&\quad \pm \frac{1}{8\pi^2}\Big[\Big(-\xi_i - \xi_j \pm \sqrt{\xi_i^2 + 14\xi_i\xi_j + \xi_j^2}\Big) \pm \Big(-\xi_{i'} - \xi_{j'} \pm \sqrt{\xi_{i'}^2 + 14\xi_{i'}\xi_{j'} + \xi_{j'}^2}\Big)\Big].
\end{aligned}$$

If $i \neq i'$, all the eigenvalues are not identically zero due to the presence of the square root terms. If $i = i'$, consequently $j = j'$, hence if the eigenvalue is

$$\begin{aligned}
&\langle k,\alpha\rangle + \varepsilon^{-4}\big(|n|^2+|i|^2\big) - \varepsilon^{-4}\big(\big|n'\big|^2+|i|^2\big) + \langle Ak+2\beta-2\beta,\xi\rangle \\
&\quad + \frac{1}{8\pi^2}\Big[\Big(-\xi_i-\xi_j+\sqrt{\xi_i^2+14\xi_i\xi_j+\xi_j^2}\Big) - \Big(-\xi_i-\xi_j+\sqrt{\xi_i^2+14\xi_i\xi_j+\xi_j^2}\Big)\Big] \\
&= \langle k,\alpha\rangle + \varepsilon^{-4}\big(|n|^2-\big|n'\big|^2\big) + \langle Ak,\xi\rangle
\end{aligned}$$

then $Ak \neq 0$ for $k \neq 0$; if the eigenvalue is

$$\begin{aligned}
&\langle k,\alpha\rangle + \varepsilon^{-4}\big(|n|^2+|i|^2\big) + \varepsilon^{-4}\big(\big|n'\big|^2+|i|^2\big) + \langle Ak+2\beta+2\beta,\xi\rangle \\
&\quad + \frac{1}{8\pi^2}\Big[\Big(-\xi_i-\xi_j+\sqrt{\xi_i^2+14\xi_i\xi_j+\xi_j^2}\Big) + \Big(-\xi_i-\xi_j-\sqrt{\xi_i^2+14\xi_i\xi_j+\xi_j^2}\Big)\Big] \\
&= \langle k,\alpha\rangle + \varepsilon^{-4}\big(|n|^2+|i|^2\big) + \varepsilon^{-4}\big(\big|n'\big|^2+|i|^2\big) + \langle Ak+2\beta+2\beta,\xi\rangle + \frac{1}{4\pi^2}(-\xi_i-\xi_j) \\
&= \langle k,\alpha\rangle + \varepsilon^{-4}\big(|n|^2+|i|^2\big) + \varepsilon^{-4}\big(\big|n'\big|^2+|i|^2\big) + \left\langle Ak+2\beta+2\beta+\frac{1}{4\pi^2}(-e_i-e_j),\xi\right\rangle
\end{aligned}$$

then when $Ak+2\beta+2\beta+\frac{1}{4\pi^2}(-e_i-e_j)=0$, all components of $k-e_i-e_j$ are equal and $(2b-1)(k-e_i-e_j)_1+8=0$ ($b \geqslant 2$), this equation has no integer solutions. Thus all eigenvalues are not identically zero.

Case 2. $n \in \mathcal{L}_1$, $n' \in \mathcal{L}_2$. In this case, the eigenvalues of $\langle k,\omega\rangle I \pm A_n \otimes I_2 \pm I_2 \otimes A_{n'}$ are

$$\begin{aligned}
&\langle k,\alpha\rangle \pm \varepsilon^{-4}\big(|n|^2+|i|^2\big) \pm \varepsilon^{-4}\big(\big|n'\big|^2-\big|i'\big|^2\big) + \langle Ak\pm 2\beta,\xi\rangle \\
&\quad \pm \frac{1}{8\pi^2}\Big[\Big(-\xi_i-\xi_j\pm\sqrt{\xi_i^2+14\xi_i\xi_j+\xi_j^2}\Big) \pm \Big(\xi_{i'}-\xi_{j'}\pm\sqrt{\xi_{i'}^2-14\xi_{i'}\xi_{j'}+\xi_{j'}^2}\Big)\Big].
\end{aligned}$$

Hence all the eigenvalues are not identically zero due to the presence of the square root terms.

Case 3. $n, n' \in \mathcal{L}_2$. In this case, the eigenvalues of $\langle k,\omega\rangle I \pm A_n \otimes I_2 \pm I_2 \otimes A_{n'}$ are

$$\begin{aligned}
&\langle k,\alpha\rangle \pm \varepsilon^{-4}\big(|n|^2-|i|^2\big) \pm \varepsilon^{-4}\big(\big|n'\big|^2-\big|i'\big|^2\big) + \langle Ak,\xi\rangle \\
&\quad \pm \frac{1}{8\pi^2}\Big[\Big(\xi_i-\xi_j\pm\sqrt{\xi_i^2-14\xi_i\xi_j+\xi_j^2}\Big) \pm \Big(\xi_{i'}-\xi_{j'}\pm\sqrt{\xi_{i'}^2-14\xi_{i'}\xi_{j'}+\xi_{j'}^2}\Big)\Big].
\end{aligned}$$

If $i \neq i'$, all the eigenvalues are not identically zero due to the presence of the square root terms. If $i = i'$, consequently $j = j'$, hence if the eigenvalue is

$$\begin{aligned}
&\langle k,\alpha\rangle + \varepsilon^{-4}\big(|n|^2-|i|^2\big) - \varepsilon^{-4}\big(\big|n'\big|^2-|i|^2\big) + \langle Ak,\xi\rangle \\
&\quad + \frac{1}{8\pi^2}\Big[\Big(\xi_i-\xi_j+\sqrt{\xi_i^2-14\xi_i\xi_j+\xi_j^2}\Big) - \Big(\xi_i-\xi_j+\sqrt{\xi_i^2-14\xi_i\xi_j+\xi_j^2}\Big)\Big] \\
&= \langle k,\alpha\rangle + \varepsilon^{-4}\big(|n|^2-\big|n'\big|^2\big) + \langle Ak,\xi\rangle
\end{aligned}$$

then $Ak \neq 0$ for $k \neq 0$; if the eigenvalue is

$$\begin{aligned}
&\langle k,\alpha\rangle + \varepsilon^{-4}\big(|n|^2-|i|^2\big) + \varepsilon^{-4}\big(\big|n'\big|^2-|i|^2\big) + \langle Ak,\xi\rangle \\
&\quad + \frac{1}{8\pi^2}\Big[\Big(\xi_i-\xi_j+\sqrt{\xi_i^2-14\xi_i\xi_j+\xi_j^2}\Big)+\Big(\xi_i-\xi_j-\sqrt{\xi_i^2-14\xi_i\xi_j+\xi_j^2}\Big)\Big] \\
&= \langle k,\alpha\rangle + \varepsilon^{-4}\big(|n|^2-|i|^2\big) + \varepsilon^{-4}\big(\big|n'\big|^2-|i|^2\big) + \langle Ak,\xi\rangle + \frac{1}{4\pi^2}(\xi_i-\xi_j) \\
&= \langle k,\alpha\rangle + \varepsilon^{-4}\big(|n|^2-|i|^2\big) + \varepsilon^{-4}\big(\big|n'\big|^2-|i|^2\big) + \Big\langle Ak+\frac{1}{4\pi^2}(e_i-e_j),\xi\Big\rangle
\end{aligned}$$

then when $Ak + \frac{1}{4\pi^2}(e_i-e_j)=0$, all components of $k+e_i-e_j$ are equal and $(2b-1)(k-e_i+e_j)_1=0$ $(b \geqslant 2)$, the integer solutions to this equation are $k=e_i-e_j$. While at this time, when $|n| \neq |m'|$,

$$\begin{aligned}
&\langle e_i-e_j,\alpha\rangle + \varepsilon^{-4}\big(|n|^2-|i|^2\big) + \varepsilon^{-4}\big(\big|n'\big|^2-|i|^2\big) \\
&\quad = \varepsilon^{-4}\big(|i|^2-|j|^2+|n|^2-|i|^2+\big(-\big|m'\big|^2+|j|^2\big)\big) \\
&\quad = \varepsilon^{-4}\big(|n|^2-\big|m'\big|^2\big) \neq 0.
\end{aligned}$$

Thus all eigenvalues are not identically zero. Due to Lemma 3.1, $\det(\langle k,\omega\rangle I \pm A_n \otimes I_2 \pm I_2 \otimes A_{n'})$ is polynomial function in ξ of order at most four. Thus

$$\big|\partial_\xi^4\big(\det\big(\langle k,\omega\rangle I \pm A_n \otimes I_2 \pm I_2 \otimes A_{n'}\big)\big)\big| \geqslant \frac{1}{2}|k| \neq 0.$$

By excluding some parameter set with measure $O(\gamma^{\frac{1}{4}})$, we have

$$\big|\det\big(\langle k,\omega\rangle I \pm A_n \otimes I_2 \pm I_2 \otimes A_{n'}\big)\big| \geqslant \frac{\gamma}{|k|^\tau}, \quad k \neq 0.$$

(A3) is verified.

Verifying (A4): See [14].

Verifying (A5): See [14].

Verifying (A6): We only need to check P satisfies (A6). Recall (3.4). F is given as

$$F = \sum_{\substack{i-j+n-m=0\\ |i|^2-|j|^2+|n|^2-|m|^2\neq 0\\ \sharp S\cap\{i,j,n,m\}\geqslant 2}} \frac{\mathrm{i}\varepsilon}{8\pi^2(\lambda_i-\lambda_j+\lambda_n-\lambda_m)} q_i\bar{q}_j z_n \bar{z}_m.$$

Then for t sufficient large and $\forall c \in \mathbb{Z}^2 \setminus \{0\}$, we have

$$\begin{aligned}
&\sum_{i,j,n,m,t} \frac{\mathrm{i}\varepsilon}{8\pi^2(\lambda_i-\lambda_j+\lambda_{n+tc}-\lambda_{m+tc})} q_i\bar{q}_j z_{n+tc}\bar{z}_{m+tc} \\
&\quad = \sum_{i,j,n,m,t} \frac{\mathrm{i}\varepsilon}{8\pi^2(|i|^2-|j|^2+|n|^2-|m|^2+2t\langle n-m,c\rangle)} q_i\bar{q}_j z_{n+tc}\bar{z}_{m+tc}.
\end{aligned}$$

Hence, when $\langle n-m, c\rangle = 0$,

$$\frac{\partial^2 F}{\partial z_{n+tc}\partial \bar{z}_{m+tc}} = \frac{\partial^2 F}{\partial z_n \partial \bar{z}_m};$$

when $\langle n-m, c\rangle \neq 0$,

$$\left\| \frac{\partial^2 F}{\partial z_{n+tc}\partial \bar{z}_{m+tc}} - 0 \right\| \leqslant \frac{\varepsilon}{t} e^{-|n-m|\rho}.$$

Similarly,

$$\left\| \frac{\partial^2 F}{\partial z_{n+tc}\partial z_{m-tc}} - \lim_{t\to\infty} \frac{\partial^2 F}{\partial z_{n+tc}\partial z_{m-tc}} \right\|, \left\| \frac{\partial^2 F}{\partial \bar{z}_{n+tc}\partial \bar{z}_{m-tc}} - \lim_{t\to\infty} \frac{\partial^2 F}{\partial \bar{z}_{n+tc}\partial \bar{z}_{m-tc}} \right\| \leqslant \frac{\varepsilon}{t} e^{-|n+m|\rho}.$$

That is to say, F satisfies Töplitz–Lipschitz property. Recalling the construction of Hamiltonian (3.1), we only need to check that $\{\frac{\varepsilon}{8\pi^2}\sum_{i-j+n-m=0} q_i \bar{q}_j z_n \bar{z}_m, F\}$ also satisfies the Töplitz–Lipschitz property. Lemma 4.4 in the next section shows that Poisson bracket preserves Töplitz–Lipschitz property. Thus $N + \mathcal{A} + \mathcal{B} + \bar{\mathcal{B}} + P$ satisfies (A6).

By applying Theorem 2, we get Theorem 1.

4. Proof of Theorem 2

Theorem 2 will be proved by a KAM iteration which involves an infinite sequence of change of variables. Each step of KAM iteration makes the perturbation smaller than that in the previous step at the cost of excluding a small set of parameters and contraction of weight. We have to prove the convergence of the iteration and estimate the measure of the excluded set after infinite KAM steps.

At the νth step of the KAM iteration, we consider a Hamiltonian vector field with

$$H_\nu = N_\nu + \mathcal{A}_\nu + \mathcal{B}_\nu + \bar{\mathcal{B}}_\nu + P_\nu = \langle \omega_\nu, I\rangle + \sum_{n\in\mathbb{Z}_1^2} \Omega_n^\nu z_n \bar{z}_n + \mathcal{A}_\nu + \mathcal{B}_\nu + \bar{\mathcal{B}}_\nu + P_\nu,$$

where $\mathcal{A}_\nu + \mathcal{B}_\nu + \bar{\mathcal{B}}_\nu + P_\nu \in \mathcal{A}$ is defined in $D_{\rho_\nu}(r_\nu, s_\nu) \times \mathcal{O}_{\nu-1}$ and satisfies (A1)–(A6).

We will construct a symplectic change of variables

$$\Phi_\nu : D_{\rho_\nu}(r_{\nu+1}, s_{\nu+1}) \times \mathcal{O}_\nu \to D_{\rho_\nu}(r_\nu, s_\nu) \times \mathcal{O}_{\nu-1}$$

such that the vector field $X_{H_\nu\circ\Phi_\nu}$ defined on $D_{\rho_{\nu+1}}(r_{\nu+1}, s_{\nu+1})$ satisfies

$$\|X_{P_{\nu+1}}\|_{D_{\rho_{\nu+1}}(r_{\nu+1},s_{\nu+1}),\mathcal{O}_\nu} = \|X_{H_\nu\circ\Phi_\nu} - X_{N_{\nu+1}+\mathcal{A}_{\nu+1}+\mathcal{B}_{\nu+1}+\bar{\mathcal{B}}_{\nu+1}}\|_{D_{\rho_{\nu+1}}(r_{\nu+1},s_{\nu+1}),\mathcal{O}_\nu} \leqslant \varepsilon_\nu^\kappa$$

with some $\kappa > 1$ and some new normal form $N_{\nu+1}$, $\mathcal{A}_{\nu+1}$, $\mathcal{B}_{\nu+1}$, $\bar{\mathcal{B}}_{\nu+1}$. Moreover, the new Hamiltonian still satisfies (A1)–(A6).

For simplicity, in what follows the quantities without subscripts and superscripts refer to quantities at the νth step, while the quantities with subscript $+$ or superscript $+$ denote the corresponding quantities at the $(\nu+1)$th step. Thus we consider now Hamiltonian

$$\begin{aligned} H &= N + \mathcal{A} + \mathcal{B} + \bar{\mathcal{B}} + P \\ &\equiv e + \langle \omega(\xi), I \rangle + \sum_{n \in \mathbb{Z}_1^2} \Omega_n(\xi) z_n \bar{z}_n + \mathcal{A} + \mathcal{B} + \bar{\mathcal{B}} + P(\theta, I, z, \bar{z}, \xi, \varepsilon) \end{aligned} \tag{4.1}$$

defined in $D_\rho(r, s) \times \mathcal{O}$.

We assume that for $\xi \in \mathcal{O}$, $|k| \leqslant K$,

$$\begin{gathered} \big|\langle k, \omega(\xi)\rangle\big| \geqslant \frac{\gamma}{K^\tau}, \quad k \neq 0, \\ \big|\det\big(\langle k, \omega\rangle I + A_n\big)\big| \geqslant \frac{\gamma}{K^\tau}, \\ \big|\det\big(\langle k, \omega\rangle I \pm A_n \otimes I_2 \pm I_2 \otimes A_{n'}\big)\big| \geqslant \frac{\gamma}{K^\tau}, \quad k \neq 0, \end{gathered} \tag{4.2}$$

where $A_n = \Omega_n$ for $n \in \mathbb{Z}_1^2 \setminus (\mathcal{L}_1 \cup \mathcal{L}_2)$,

$$\begin{aligned} A_n &= \begin{pmatrix} \Omega_n + \omega_i & a_n \\ a_m & \Omega_m + \omega_j \end{pmatrix}, \quad n \in \mathcal{L}_1', \\ A_n &= \begin{pmatrix} \Omega_n - \omega_i & -a_n \\ \bar{a}_m & -\Omega_m + \omega_j \end{pmatrix}, \quad n \in \mathcal{L}_2', \end{aligned}$$

where (n, m) are resonant pairs, and (i, j) is uniquely determined by (n, m). Moreover, $N + \mathcal{A} + \mathcal{B} + \bar{\mathcal{B}} + P$ satisfies (A4), (A5), (A6).

Expand P into the Fourier–Taylor series $P = \sum_{k,l,\alpha,\beta} P_{kl\alpha\beta} I^l e^{\mathrm{i}\langle k,\theta\rangle} z^\alpha \bar{z}^\beta$. (A5) implies that

$$P_{kl\alpha\beta} = 0 \quad \text{if} \sum_{j=1}^{b} k_j i_j + \sum_{n \in \mathbb{Z}_1^2} (\alpha_n - \beta_n) n \neq 0. \tag{4.3}$$

We now let $0 < r_+ < r$ and define

$$s_+ = \frac{1}{4} s \varepsilon^{\frac{1}{3}}, \qquad \varepsilon_+ = c\big(\gamma^{-1} K^\tau\big)^4 \varepsilon^{\frac{4}{3}}. \tag{4.4}$$

Here and later, the letter c denotes suitable (possibly different) constant not depending on the iteration steps.

We will construct a set $\mathcal{O}_+ \subset \mathcal{O}$ and a change of variables $\Phi : D_+ \times \mathcal{O}_+ = D_\rho(r_+, s_+) \times \mathcal{O}_+ \to D_\rho(r, s) \times \mathcal{O}$ such that the transformed Hamiltonian $H_+ = N_+ + \mathcal{A}_+ + \mathcal{B}_+ + \bar{\mathcal{B}}_+ + P_+ \equiv H \circ \Phi$ satisfies all the above iterative assumptions with new parameters $s_+, \varepsilon_+, r_+, \rho_+$ and with $\xi \in \mathcal{O}_+$.

4.1. Solving the linearized equations

Expand P into the Fourier–Taylor series

$$P = \sum_{k,l,\alpha,\beta} P_{kl\alpha\beta} e^{\mathrm{i}\langle k,\theta\rangle} I^l z^\alpha \bar{z}^\beta$$

where $k \in \mathbb{Z}^b, l \in \mathbb{N}^b$ and the multi-indices α and β run over the set of all infinite dimensional vectors $\alpha \equiv (\dots, \alpha_n, \dots)_{n \in \mathbb{Z}_1^2}$ with finitely many nonzero components of positive integers.

Let R be the truncation of P given by

$$\begin{aligned} R(\theta, I, z, \bar{z}) &= R^0 + R^1 + R^{10} + R^{01} + R^{20} + R^{11} + R^{02} \\ &= R^0(\theta) + \langle R^1(\theta), I\rangle + \langle R^{10}(\theta), z\rangle + \langle R^{01}(\theta), \bar{z}\rangle \\ &\quad + \langle R^{20}(\theta) z, z\rangle + \langle R^{11}(\theta) z, \bar{z}\rangle + \langle R^{02}(\theta)\bar{z}, \bar{z}\rangle \\ &= \sum_{|k| \leqslant K} P_k^0 e^{\mathrm{i}\langle k,\theta\rangle} + \sum_{|k| \leqslant K} \langle P_k^1, I\rangle e^{\mathrm{i}\langle k,\theta\rangle} + \sum_{|k| \leqslant K, n} P_{k,n}^{10} z_n e^{\mathrm{i}\langle k,\theta\rangle} \\ &\quad + \sum_{|k| \leqslant K, n} P_{k,n}^{01} \bar{z}_n e^{\mathrm{i}\langle k,\theta\rangle} + \sum_{|k| \leqslant K, n, m} P_{k,nm}^{20} z_n z_m e^{\mathrm{i}\langle k,\theta\rangle} \\ &\quad + \sum_{|k| \leqslant K, n, m} P_{k,nm}^{11} z_n \bar{z}_m e^{\mathrm{i}\langle k,\theta\rangle} + \sum_{|k| \leqslant K, n, m} P_{k,nm}^{02} \bar{z}_n \bar{z}_m e^{\mathrm{i}\langle k,\theta\rangle} \end{aligned} \tag{4.5}$$

where $P_{k,n}^{10} = P_{k0\alpha\beta}$ with $\alpha = e_n, \beta = 0$, here e_n denotes the vector with the nth component being 1 and the other components being zero; $P_{k,n}^{01} = P_{k0\alpha\beta}$ with $\alpha = 0$, $\beta = e_n$; $P_{k,nm}^{20} = P_{k0\alpha\beta}$ with $\alpha = e_n + e_m$, $\beta = 0$; $P_{k,nm}^{11} = P_{k0\alpha\beta}$ with $\alpha = e_n$, $\beta = e_m$; $P_{k,nm}^{02} = P_{k0\alpha\beta}$ with $\alpha = 0$, $\beta = e_n + e_m$.

Rewrite H as $H = N + \mathcal{A} + \mathcal{B} + \bar{\mathcal{B}} + R + (P - R)$. By the choice of s_+ in (4.4) and the definition of the norms, it follows immediately that

$$\|X_R\|_{D_\rho(r,s),\mathcal{O}} \leqslant \|X_P\|_{D_\rho(r,s),\mathcal{O}} \leqslant \varepsilon. \tag{4.6}$$

Moreover, we take $s_+ \ll s$ such that in a domain $D_\rho(r, s_+)$,

$$\|X_{(P-R)}\|_{D_\rho(r,s_+)} \leqslant \varepsilon_+. \tag{4.7}$$

In the following, we will construct an F satisfying (A5), defined in a domain $D_+ = D_\rho(r_+, s_+)$, such that the time one map ϕ_F^1 of the Hamiltonian vector field X_F defines a map from $D_+ \to D$ and transforms H into H_+. More precisely, by second order Taylor formula, we have

$$\begin{aligned} H \circ \phi_F^1 &= (N + \mathcal{A} + \mathcal{B} + \bar{\mathcal{B}} + R) \circ \phi_F^1 + (P - R) \circ \phi_F^1 \\ &= N + \mathcal{A} + \mathcal{B} + \bar{\mathcal{B}} + \{N + \mathcal{A} + \mathcal{B} + \bar{\mathcal{B}}, F\} + R \\ &\quad + \int_0^1 (1-t)\big\{\{N + \mathcal{A} + \mathcal{B} + \bar{\mathcal{B}}, F\}, F\big\} \circ \phi_F^t \, dt \\ &\quad + \int_0^1 \{R, F\} \circ \phi_F^t \, dt + (P - R) \circ \phi_F^1 \\ &= N_+ + \mathcal{A}_+ + \mathcal{B}_+ + \bar{\mathcal{B}}_+ + P_+ + \{N + \mathcal{A} + \mathcal{B} + \bar{\mathcal{B}}, F\} + R \end{aligned}$$

$$- P_{0000} - \langle \hat{\omega}, I \rangle - \sum_n P_{nn}^{011} z_n \bar{z}_n - \hat{\mathcal{A}} - \hat{\mathcal{B}} - \bar{\hat{\mathcal{B}}}, \tag{4.8}$$

where

$$\hat{\omega} = \int \frac{\partial P}{\partial I} d\theta |_{z=\bar{z}=0, I=0},$$

$$\hat{\mathcal{A}} = \sum_{n \in \mathcal{L}_1} P_{e_i - e_j, nm}^{11} z_n \bar{z}_m e^{\mathrm{i}(\theta_i - \theta_j)},$$

$$\hat{\mathcal{B}} = \sum_{n \in \mathcal{L}_2} P_{-e_i - e_j, nm}^{20} z_n z_m e^{\mathrm{i}(-\theta_i - \theta_j)},$$

$$\bar{\hat{\mathcal{B}}} = \sum_{n \in \mathcal{L}_2} P_{e_i + e_j, nm}^{02} \bar{z}_n \bar{z}_m e^{\mathrm{i}(\theta_i + \theta_j)},$$

$$N_+ = N + P_{0000} + \langle \hat{\omega}, I \rangle + \sum_n P_{nn}^{011} z_n \bar{z}_n, \tag{4.9}$$

$$\mathcal{A}_+ = \mathcal{A} + \hat{\mathcal{A}}, \tag{4.10}$$

$$\mathcal{B}_+ = \mathcal{B} + \hat{\mathcal{B}}, \tag{4.11}$$

$$\bar{\mathcal{B}}_+ = \bar{\mathcal{B}} + \bar{\hat{\mathcal{B}}} = \bar{\mathcal{B}} + \hat{\bar{\mathcal{B}}}, \tag{4.12}$$

$$P_+ = \int_0^1 (1-t) \big\{ \{ N + \mathcal{A} + \mathcal{B} + \bar{\mathcal{B}}, F \}, F \big\} \circ \phi_F^t \, dt$$

$$+ \int_0^1 \{R, F\} \circ \phi_F^t \, dt + (P - R) \circ \phi_F^1. \tag{4.13}$$

We shall construct a function F of the form

$$\begin{aligned} F(\theta, I, z, \bar{z}) &= F^0 + F^1 + F^{10} + F^{01} + F^{20} + F^{11} + F^{02} \\ &= F^0(\theta) + \big\langle F^1(\theta), I \big\rangle + \big\langle F^{10}(\theta), z \big\rangle + \big\langle F^{01}(\theta), \bar{z} \big\rangle \\ &\quad + \big\langle F^{20}(\theta) z, z \big\rangle + \big\langle F^{11}(\theta) z, \bar{z} \big\rangle + \big\langle F^{02}(\theta) \bar{z}, \bar{z} \big\rangle \end{aligned} \tag{4.14}$$

which satisfies the equation

$$\{ N + \mathcal{A} + \mathcal{B} + \bar{\mathcal{B}}, F \} + R - P_{0000} - \langle \hat{\omega}, I \rangle - \sum_n P_{nn}^{011} z_n \bar{z}_n - \hat{\mathcal{A}} - \hat{\mathcal{B}} - \hat{\bar{\mathcal{B}}} = 0. \tag{4.15}$$

(4.15) is equivalent to

$$\big\{ N, F^0 + F^1 \big\} + R^0 + R^1 - P_{0000} - \langle \hat{\omega}, I \rangle = 0, \tag{4.16}$$

$$\big\{ N + \mathcal{A} + \mathcal{B} + \bar{\mathcal{B}}, F^{10} + F^{01} \big\} + R^{10} + R^{01} = 0, \tag{4.17}$$

$$\begin{aligned}&\big\{N+\mathcal{A}+\mathcal{B}+\bar{\mathcal{B}},F^{20}+F^{11}+F^{02}\big\}+R^{20}+R^{11}+R^{02}\\&\quad-\sum_n P^{011}_{nn}z_n\bar{z}_n-\hat{\mathcal{A}}-\hat{\mathcal{B}}-\hat{\bar{\mathcal{B}}}=0.\end{aligned}\tag{4.18}$$

Solving (4.16). $F^0(\theta)=\sum_{0<|k|\leqslant K}F^0_k e^{\mathrm{i}\langle k,\theta\rangle}$, $F^1(\theta)=\sum_{0<|k|\leqslant K}F^1_k e^{\mathrm{i}\langle k,\theta\rangle}$ is constructed by setting

$$F^j_k=\frac{1}{i\langle k,\omega\rangle}P^j_k,\quad j=0,1,\ 0<|k|\leqslant K.$$

From the assumption

$$\big|\langle k,\omega(\xi)\rangle\big|\geqslant\frac{\gamma}{K^\tau},\quad \xi\in\mathcal{O},$$

we have

$$\big|F^j_k\big|_{\mathcal{O}}\leqslant\gamma^{-2}K^{2\tau}\big|P^j_k\big|_{\mathcal{O}},\quad 0<|k|\leqslant K.$$

Solving (4.17). Comparing the Fourier coefficients, (4.17) is decomposed into a set of linear systems of order 1 or 2. More precisely, we have

(1) If $n\in\mathbb{Z}^2_1\backslash\{\mathcal{L}_1\cup\mathcal{L}_2\}$, we have

$$\begin{aligned}&\big(\langle k,\omega\rangle+\Omega_n\big)F^{10}_{k,n}=-\mathrm{i}R^{10}_{k,n},\\&\big(\langle k,\omega\rangle-\Omega_n\big)F^{01}_{k,n}=-\mathrm{i}R^{01}_{k,n}.\end{aligned}\tag{4.19}$$

(2) If (n,m) is a resonant pair in $\mathcal{L}_1$, we have

$$\begin{aligned}&\big(\langle k+e_i,\omega\rangle+\Omega_n\big)F^{10}_{k+e_i,n}+a_nF^{10}_{k+e_j,m}=-\mathrm{i}R^{10}_{k+e_i,n},\\&\big(\langle k+e_j,\omega\rangle+\Omega_m\big)F^{10}_{k+e_j,m}+a_mF^{10}_{k+e_i,n}=-\mathrm{i}R^{10}_{k+e_j,m}.\end{aligned}\tag{4.20}$$

(3) If (n,m) is a resonant pair in $\mathcal{L}_2$, we have

$$\begin{aligned}&\big(\langle k-e_i,\omega\rangle+\Omega_n\big)F^{10}_{k-e_i,n}-a_nF^{01}_{k+e_j,m}=-\mathrm{i}R^{10}_{k-e_i,n},\\&\big(\langle k+e_j,\omega\rangle-\Omega_m\big)F^{01}_{k+e_j,m}+\bar{a}_mF^{10}_{k-e_i,n}=-\mathrm{i}R^{01}_{k+e_j,m}.\end{aligned}\tag{4.21}$$

(4.19), (4.20) and (4.21) are linear systems with coefficient matrix

$$\langle k,\omega\rangle I+A_n.\tag{4.22}$$

By the small divisor assumption

$$\big|\det\big(\langle k,\omega\rangle I+A_n\big)\big|\geqslant\frac{\gamma}{K^\tau},\quad |k|\leqslant K$$

we have

$$\left|F_{k,n}^{10}\right|_{\mathcal{O}} \leqslant \left(\frac{K^{\tau}}{\gamma}\right)^4 \varepsilon e^{-|k|r} e^{-|n|\rho}, \qquad \left|F_{k,m}^{01}\right|_{\mathcal{O}} \leqslant \left(\frac{K^{\tau}}{\gamma}\right)^4 \varepsilon e^{-|k|r} e^{-|m|\rho}.$$

Solving (4.18). Similarly, by comparing the Fourier coefficient, (4.18) is decomposed into a set of linear systems of order 1, 2 and 4 with coefficient matrix

$$\langle k,\omega\rangle I \pm A_n \otimes I \pm I \otimes A_{n'}, \quad n, n' \in \mathbb{Z}_1^2.$$

For example, in case that $n, n' \in \mathbb{Z}_1^2 \backslash (\mathcal{L}_1 \cup \mathcal{L}_2)$, we have

$$\begin{aligned}
&\big(\langle k,\omega\rangle + \Omega_n - \Omega_{n'}\big) F_{k,nn'}^{11} = -\mathrm{i} R_{k,nn'}^{11},\\
&\big(\langle k,\omega\rangle + \Omega_n + \Omega_{n'}\big) F_{k,nn'}^{20} = -\mathrm{i} R_{k,nn'}^{20},\\
&\big(\langle k,\omega\rangle - \Omega_n - \Omega_{n'}\big) F_{k,nn'}^{02} = -\mathrm{i} R_{k,nn'}^{02}.
\end{aligned} \tag{4.23}$$

In case that $n \in \mathbb{Z}_1^2 \backslash (\mathcal{L}_1 \cup \mathcal{L}_2)$ and (n', m') is a resonant pair in $\mathcal{L}_1$, we have

$$\begin{aligned}
&\big(\langle k - e_{i'},\omega\rangle + \Omega_n - \Omega_{n'}\big) F_{k-e_{i'},nn'}^{11} - a_{n'} F_{k-e_{j'},nm'}^{11} = -\mathrm{i} R_{k-e_{i'},nn'}^{11},\\
&\big(\langle k - e_{j'},\omega\rangle + \Omega_n - \Omega_{m'}\big) F_{k-e_{j'},nm'}^{11} - a_{m'} F_{k-e_{i'},nn'}^{11} = -\mathrm{i} R_{k-e_{j'},nm'}^{11}.
\end{aligned}$$

In case that (n, m) is a resonant pair in $\mathcal{L}_1$ and (n', m') is a resonant pair in $\mathcal{L}_2$. Comparing the Fourier coefficients, we have that $F_{k+e_i+e_{i'},nn'}^{11}$, $F_{k+e_i-e_{j'},nm'}^{20}$, $F_{k+e_j+e_{i'},mn'}^{11}$, $F_{k+e_j-e_{j'},mm'}^{20}$ satisfy

$$\begin{aligned}
&\big(\langle k + e_i + e_{i'},\omega\rangle + \Omega_n - \Omega_{n'}\big) F_{k+e_i+e_{i'},nn'}^{11} + a_{n'} F_{k+e_i-e_{j'},nm'}^{20} + a_n F_{k+e_j+e_{i'},mn'}^{11}\\
&\quad = -\mathrm{i} R_{k+e_i+e_{i'},nn'}^{11},\\
&\big(\langle k + e_i - e_{j'},\omega\rangle + \Omega_n + \Omega_{m'}\big) F_{k+e_i-e_{j'},nm'}^{20} - \bar{a}_{m'} F_{k+e_i+e_{i'},nn'}^{11} + a_n F_{k+e_j-e_{j'},mm'}^{20}\\
&\quad = -\mathrm{i} R_{k+e_i-e_{j'},nm'}^{20},\\
&\big(\langle k + e_j + e_{i'},\omega\rangle + \Omega_m - \Omega_{n'}\big) F_{k+e_j+e_{i'},mn'}^{11} + a_m F_{k+e_i+e_{i'},nn'}^{11} + a_n F_{k+e_i-e_{j'},mm'}^{20}\\
&\quad = -\mathrm{i} R_{k+e_j+e_{i'},mn'}^{11},\\
&\big(\langle k + e_j - e_{j'},\omega\rangle + \Omega_m + \Omega_{m'}\big) F_{k+e_j-e_{j'},mm'}^{20} + a_m F_{k+e_i-e_{j'},nm'}^{20} - \bar{a}_{m'} F_{k+e_j+e_{i'},mn'}^{11}\\
&\quad = -\mathrm{i} R_{k+e_j-e_{j'},mm'}^{20}.
\end{aligned}$$

By small divisor assumption

$$\big|\det\big(\langle k,\omega\rangle I \pm A_n \otimes I \pm I \otimes A_{n'}\big)\big| \geqslant \frac{\gamma}{K^{\tau}},$$

we have estimates

$$\left|F^{11}_{k,nn'}\right|_{\mathcal{O}} \leqslant \left(\frac{K^\tau}{\gamma}\right)^4 \varepsilon e^{-|k|r} e^{-|n-n'|\rho}, \qquad \left|F^{20}_{k,nn'}\right|_{\mathcal{O}} + \left|F^{02}_{k,nn'}\right|_{\mathcal{O}} \leqslant \left(\frac{K^\tau}{\gamma}\right)^4 \varepsilon e^{-|k|r} e^{-|n+n'|\rho}.$$

4.2. Estimation on the coordinate transformation

With the previous section, we give the estimate to X_F and ϕ^1_F.

Lemma 4.1. *Let $D_i = D(r_+ + \frac{i}{4}(r-r_+), \frac{i}{4}s)$, $0 < i \leqslant 4$. Then*

$$\|X_F\|_{D_3,\mathcal{O}} \leqslant c\big(\gamma^{-1}K^\tau\big)^4\varepsilon. \tag{4.24}$$

Lemma 4.2. *Let $\eta = \varepsilon^{\frac{1}{3}}$, $D_{i\eta} = D(r_+ + \frac{i}{4}(r-r_+), \frac{i}{4}\eta s)$, $0 < i \leqslant 4$. If $\varepsilon \ll (\frac{1}{2}\gamma K^{-\tau})^6$, we then have*

$$\phi^t_F : D_{2\eta} \to D_{3\eta}, \quad -1 \leqslant t \leqslant 1. \tag{4.25}$$

Moreover,

$$\left\|D\phi^t_F - Id\right\|_{D_{1\eta}} \leqslant c\big(\gamma^{-1}K^\tau\big)^4\varepsilon. \tag{4.26}$$

4.3. Estimation for the new perturbation

The map ϕ^1_F defined above transforms H into

$$H_+ = N_+ + \mathcal{A}_+ + \mathcal{B}_+ + \bar{\mathcal{B}}_+ + P_+$$

(see (4.8) and (4.15)).

Since

$$\begin{aligned} P_+ &= \int_0^1 (1-t)\big\{\{N + \mathcal{A} + \mathcal{B} + \bar{\mathcal{B}}, F\}, F\big\} \circ \phi^t_F\, dt \\ &\quad + \int_0^1 \{R, F\} \circ \phi^t_F\, dt + (P-R)\circ\phi^1_F \\ &= \int_0^1 \big\{R(t), F\big\} \circ \phi^t_F\, dt + (P-R)\circ\phi^1_F, \end{aligned}$$

where $R(t) = (1-t)(N_+ - N) + tR$. Hence

$$X_{P_+} = \int_0^1 \left(\phi_F^t\right)^* X_{\{R(t),F\}}\, dt + \left(\phi_F^1\right)^* X_{(P-R)}.$$

According to Lemma 4.2,

$$\left\| D\phi_F^t - Id \right\|_{D_{1\eta}} \leqslant c\left(\gamma^{-1} K^\tau\right)^4 \varepsilon, \quad -1 \leqslant t \leqslant 1,$$

thus

$$\left\| D\phi_F^t \right\|_{D_{1\eta}} \leqslant 1 + \left\| D\phi_F^t - Id \right\|_{D_{1\eta}} \leqslant 2, \quad -1 \leqslant t \leqslant 1,$$
$$\| X_{\{R(t),F\}} \|_{D_{2\eta}} \leqslant c\left(\gamma^{-1} K^\tau\right)^4 \eta^{-2} \varepsilon^2,$$

and

$$\| X_{(P-R)} \|_{D_{2\eta}} \leqslant c\eta\varepsilon,$$

we have

$$\| X_{P_+} \|_{D(r_+,s_+)} \leqslant c\eta\varepsilon + c\left(\gamma^{-1} K^\tau\right)^4 \eta^{-2} \varepsilon^2 \leqslant \varepsilon_+.$$

4.4. Verification of (A5) and (A6) after one step of KAM iteration

(A5) after one step of KAM iteration is proved by Geng–You in Lemma 4.4 [14]. In the following, we have to check that the new error term P_+ satisfies (A6) with $K_+, \varepsilon_+, \rho_+$ in place of K, ε, ρ. Since

$$\begin{aligned} P_+ = P - R + \{P, F\} + \frac{1}{2!}\big\{\{N, F\}, F\big\} + \frac{1}{2!}\big\{\{P, F\}, F\big\} \\ + \cdots + \frac{1}{n!}\big\{\cdots\{N, \underbrace{F\}\cdots, F}_{n}\big\} + \frac{1}{n!}\big\{\cdots\{P, \underbrace{F\}\cdots, F}_{n}\big\} + \cdots \end{aligned}$$

then for a fixed $c \in \mathbb{Z}^2 \setminus \{0\}$, and $|n-m| > K$ with $K \geqslant \frac{1}{\rho-\rho_+} \ln(\frac{\varepsilon}{\varepsilon_+})$,

$$\left\| \frac{\partial^2 (P-R)}{\partial z_{n+tc} \partial \bar{z}_{m+tc}} - \lim_{t\to\infty} \frac{\partial^2 (P-R)}{\partial z_{n+tc} \partial \bar{z}_{m+tc}} \right\| \leqslant \frac{\varepsilon}{t} e^{-|n-m|\rho} \leqslant \frac{\varepsilon_+}{t} e^{-|n-m|\rho_+}.$$

That is to say, $P - R$ satisfies (A6) with $K_+, \varepsilon_+, \rho_+$ in place of K, ε, ρ. The proof of the remaining terms satisfying (A6) is composed by the following two lemmas.

Lemma 4.3. *F satisfies* (A6) *with $\varepsilon^{\frac{2}{3}}$ in place of ε.*

Proof. In the case that $n, n' \in \mathbb{Z}_1^2 \setminus (\mathcal{L}_1 \cup \mathcal{L}_2)$, we have

$$F_{knn'}^{11} = \frac{-\mathrm{i}}{\langle k,\omega\rangle + \Omega_n - \Omega_{n'}} R_{knn'}^{11}$$

where

$$\Omega_n = |n|^2 + \tilde{\Omega}_n,$$

and $\lim_{t\to\infty} \tilde{\Omega}_{n+tc}$ exists for all $n, c \in \mathbb{Z}^2$ with

$$\left|\Omega_{n+tc} - \lim_{t\to\infty} \tilde{\Omega}_{n+tc}\right| < \frac{\varepsilon_0}{t}. \tag{4.27}$$

Notice that

$$\langle k,\omega\rangle + \varepsilon_0^{-a}\big(|n+tc|^2 - \left|n'+tc\right|^2\big) = \langle k,\omega\rangle + \varepsilon_0^{-a}\big(|n|^2 - \left|n'\right|^2 + 2t\langle n-n',c\rangle\big).$$

In the case that $\langle n-n', c\rangle = 0$,

$$F_{k,n+tc,n'+tc}^{11} = \frac{-\mathrm{i}}{\langle k,\omega\rangle + \varepsilon_0^{-a}(|n|^2 - |n'|^2) + \tilde{\Omega}_{n+tc} - \tilde{\Omega}_{n'+tc}} R_{k,n+tc,n'+tc}^{11}.$$

Thus

$$\lim_{t\to\infty} F_{k,n+tc,n'+tc}^{11} = \frac{-\mathrm{i}}{\langle k,\omega\rangle + \varepsilon_0^{-a}(|n|^2 - |n'|^2) + \lim_{t\to\infty}(\tilde{\Omega}_{n+tc} - \tilde{\Omega}_{n'+tc})} \lim_{t\to\infty} R_{k,n+tc,n'+tc}^{11}.$$

By (4.27),

$$\left|F_{k,n+tc,n'+tc}^{11} - \lim_{t\to\infty} F_{k,n+tc,n'+tc}^{11}\right| \leqslant \gamma^{-8} K^{8\tau} \frac{\varepsilon}{t} e^{-|n-n'|\rho} e^{-|k|r} \leqslant \frac{\varepsilon^{\frac{2}{3}}}{t} e^{-|n-n'|\rho} e^{-|k|r}.$$

Thus

$$\left\| \frac{\partial^2 F}{\partial z_{n+tc}\partial \bar{z}_{n'+tc}} - \lim_{t\to\infty} \frac{\partial^2 F}{\partial z_{n+tc}\partial \bar{z}_{n'+tc}} \right\| \leqslant \frac{\varepsilon^{\frac{2}{3}}}{t} e^{-|n-n'|\rho}.$$

If $\langle n-n', c\rangle \neq 0$ and $t > K$, it is easy to see that

$$\left\| \frac{\partial^2 F}{\partial z_{n+tc}\partial \bar{z}_{n'+tc}} - 0 \right\| \leqslant \frac{\varepsilon}{t} e^{-|n-n'|\rho}.$$

Similarly, we have

$$g\left\|\frac{\partial^2 F}{\partial z_{n+tc}\partial z_{n'-tc}}-\lim_{t\to\infty}\frac{\partial^2 F}{\partial z_{n+tc}\partial z_{n'-tc}}\right\|,\left\|\frac{\partial^2 F}{\partial \bar{z}_{n+tc}\partial \bar{z}_{n'-tc}}-\lim_{t\to\infty}\frac{\partial^2 F}{\partial \bar{z}_{n+tc}\partial \bar{z}_{n'-tc}}\right\|$$
$$\leqslant \frac{\varepsilon^{\frac{2}{3}}}{t}e^{-|n+n'|\rho}.$$

In the case that $n\in\mathcal{L}_1$, $n'\in\mathcal{L}_2$, we let

$$Q_{k,nn'}=\left(F^{11}_{k+e_i+e_{i'},nn'},F^{20}_{k+e_i-e_{j'},nm'},F^{11}_{k+e_j+e_{i'},mn'},F^{20}_{k+e_j-e_{j'},mm'}\right),$$
$$T_{k,nn'}=\left(R^{11}_{k+e_i+e_{i'},nn'},R^{20}_{k+e_i-e_{j'},nm'},R^{11}_{k+e_j+e_{i'},mn'},R^{20}_{k+e_j-e_{j'},mm'}\right).$$

Then,

$$\left(\langle k,\omega\rangle I+A_{n+tc}\otimes I_2-I_2\otimes A_{n'+tc}\right)Q_{k,n+tc,n'+tc}=-\mathrm{i}T_{k,n+tc,n'+tc}.$$

Recall (4.2)

$$\left|\det\left(\langle k,\omega\rangle I\pm A_n\otimes I_2\pm I_2\otimes A_{n'}\right)\right|\geqslant\frac{\gamma}{K^\tau},\quad k\neq 0,$$

and notice that

$$\langle k,\omega\rangle+\varepsilon_0^{-a}\left(|n+i+tc|^2-\left|n'-i'+tc\right|^2\right)$$
$$=\langle k,\omega\rangle+\varepsilon_0^{-a}\left(|n+i|^2-\left|n'-i'\right|^2+2t\left\langle n+i-n'+i',c\right\rangle\right).$$

Hence if $\langle n+i-n'+i',c\rangle=0$,

$$\lim_{t\to\infty}Q_{k,n+tc,n'+tc}=-\mathrm{i}\Bigl(\langle k,\omega\rangle I+\lim_{t\to\infty}(A_{n+tc}\otimes I_2-I_2\otimes A_{n'+tc})\Bigr)^{-1}\lim_{t\to\infty}T_{k,(n+tc)(n'+tc)},$$

exists.

Notice that

$$\left\|\lim_{t\to\infty}Q_{k,n+tc,n'+tc}\right\|\leqslant\gamma^{-4}K^{4\tau}\varepsilon e^{-|k|r}e^{-|n-n'|\rho},$$

$$\left(\langle k,\omega\rangle I+A_{n+tc}\otimes I_2-I_2\otimes A_{n'+tc}\right)\Bigl(Q_{k,n+tc,n'+tc}-\lim_{t\to\infty}Q_{k,n+tc,n'+tc}\Bigr)$$
$$=-\mathrm{i}\Bigl(T_{k,n+tc,n'+tc}-\lim_{t\to\infty}T_{k,n+tc,n'+tc}\Bigr)$$
$$-\Bigl(A_{n+tc}\otimes I_2-I_2\otimes A_{n'+tc}-\lim_{t\to\infty}(A_{n+tc}\otimes I_2-I_2\otimes A_{n'+tc})\Bigr)\lim_{t\to\infty}Q_{k,n+tc,n'+tc},$$

and

$$\begin{aligned}&\left\| \Big(A_{n+tc} \otimes I_2 - I_2 \otimes A_{n'+tc} - \lim_{t\to\infty} (A_{n+tc} \otimes I_2 - I_2 \otimes A_{n'+tc}) \Big) \lim_{t\to\infty} Q_{k,n+tc,n'+tc} \right\| \\ &\quad \leqslant \frac{\varepsilon_0}{t} \gamma^{-4} K^{4\tau} \varepsilon e^{-|k|r} e^{-|n-n'|\rho},\end{aligned}$$

we have

$$\left\| Q_{k,(n+tc)(n'+tc)} - \lim_{t\to\infty} Q_{k,n+tc,n'+tc} \right\| \leqslant \gamma^{-8} K^{8\tau} \frac{\varepsilon}{t} e^{-|k|r} e^{-|n-n'|\rho} \leqslant \frac{\varepsilon^{\frac{2}{3}}}{t} e^{-|k|r} e^{-|n-n'|\rho}.$$

As a consequence,

$$\left\| \frac{\partial^2 F}{\partial z_{n+tc} \partial \bar{z}_{n'+tc}} - \lim_{t\to\infty} \frac{\partial^2 F}{\partial z_{n+tc} \partial \bar{z}_{n'+tc}} \right\| \leqslant \frac{\varepsilon^{\frac{2}{3}}}{t} e^{-|n-n'|\rho}.$$

If $\langle n + i - n' + i', c\rangle \neq 0$ and $t > K$, it is easy to see that

$$\left\| \frac{\partial^2 F}{\partial z_{n+tc} \partial \bar{z}_{n'+tc}} - 0 \right\| \leqslant \frac{\varepsilon}{t} e^{-|n-n'|\rho}.$$

Similarly, we have

$$\begin{aligned}&\left\| \frac{\partial^2 F}{\partial z_{n+tc} \partial z_{n'-tc}} - \lim_{t\to\infty} \frac{\partial^2 F}{\partial z_{n+tc} \partial z_{n'-tc}} \right\|, \left\| \frac{\partial^2 F}{\partial \bar{z}_{n+tc} \partial \bar{z}_{n'-tc}} - \lim_{t\to\infty} \frac{\partial^2 F}{\partial \bar{z}_{n+tc} \partial \bar{z}_{n'-tc}} \right\| \\ &\quad \leqslant \frac{\varepsilon^{\frac{2}{3}}}{t} e^{-|n+n'|\rho}.\end{aligned}$$

As a result, F satisfies Töplitz–Lipschitz property (A6) with $\varepsilon^{\frac{2}{3}}$ in place of ε. $\square$

Lemma 4.4. *Assume that P satisfies* (A6)*, F satisfies* (A6) *with $\varepsilon^{\frac{2}{3}}$ in place of ε and*

$$\frac{\partial^2 F}{\partial z_n \partial z_m} = 0 \quad \big(|n+m| > K\big), \qquad \frac{\partial^2 F}{\partial z_n \partial \bar{z}_m} = 0 \quad \big(|n-m| > K\big),$$
$$\frac{\partial^2 F}{\partial \bar{z}_n \partial \bar{z}_m} = 0 \quad \big(|n+m| > K\big),$$

then $\{P, F\}$ satisfies (A6) *with ε_+ in place of ε.*

Proof. Set $f^{11}_{nm} = \lim_{t\to\infty} \frac{\partial^2 F}{\partial z_{n+tc} \partial \bar{z}_{m+tc}}$, $p^{11}_{nm} = \lim_{t\to\infty} \frac{\partial^2 P}{\partial z_{n+tc} \partial \bar{z}_{m+tc}}$, we have

$$\left\| \frac{\partial^2 F}{\partial z_{n+tc} \partial \bar{z}_{m+tc}} - f^{11}_{nm} \right\| < \frac{\varepsilon^{\frac{2}{3}} e^{-|n-m|\rho}}{t},$$
$$\left\| \frac{\partial^2 P}{\partial z_{n+tc} \partial \bar{z}_{m+tc}} - p^{11}_{nm} \right\| < \frac{\varepsilon e^{-|n-m|\rho}}{t}.$$

The notations f_{nm}^{20}, f_{nm}^{02}, p_{nm}^{20}, p_{nm}^{02} are defined similarly. This leads to

$$\begin{aligned}
&\left\| \frac{\partial^2\{F, P\}}{\partial z_{n+tc}\partial \bar{z}_{m+tc}} - \sum_j \left(f_{nj}^{11} p_{jm}^{11} - f_{mj}^{11} p_{jn}^{11}\right)\right. \\
&\qquad \left. - \sum_j \left(f_{nj}^{20} p_{jm}^{02} - f_{nj}^{02} p_{jm}^{20}\right) - \sum_j \left(f_{mj}^{20} p_{jn}^{02} - f_{mj}^{02} p_{jn}^{20}\right)\right\| \\
&\quad \leqslant \sum_j \left\| \frac{\partial^2 F}{\partial z_{n+tc}\partial \bar{z}_{j+tc}} - f_{nj}^{11}\right\| \left\| p_{jm}^{11}\right\| + \left\| \frac{\partial^2 F}{\partial z_{m+tc}\partial \bar{z}_{j+tc}} - f_{mj}^{11}\right\| \left\| p_{jn}^{11}\right\| \\
&\qquad + \sum_j \left\| \frac{\partial^2 P}{\partial z_{n+tc}\partial \bar{z}_{j+tc}} - p_{nj}^{11}\right\| \left\| f_{jm}^{11}\right\| + \left\| \frac{\partial^2 P}{\partial z_{m+tc}\partial \bar{z}_{j+tc}} - p_{mj}^{11}\right\| \left\| f_{jn}^{11}\right\| \\
&\qquad + \sum_j \left\| \frac{\partial^2 P}{\partial z_{n+tc}\partial z_{j+tc}} - p_{nj}^{20}\right\| \left\| f_{jm}^{02}\right\| + \left\| \frac{\partial^2 P}{\partial z_{m+tc}\partial z_{j+tc}} - p_{mj}^{20}\right\| \left\| f_{jn}^{02}\right\| \\
&\qquad + \sum_j \left\| \frac{\partial^2 F}{\partial z_{n+tc}\partial z_{j+tc}} - f_{nj}^{20}\right\| \left\| p_{jm}^{02}\right\| + \left\| \frac{\partial^2 F}{\partial z_{m+tc}\partial z_{j+tc}} - f_{mj}^{20}\right\| \left\| p_{jn}^{02}\right\| \\
&\qquad + \sum_j \left\| \frac{\partial^2 P}{\partial \bar{z}_{n+tc}\partial \bar{z}_{j+tc}} - p_{nj}^{02}\right\| \left\| f_{jm}^{20}\right\| + \left\| \frac{\partial^2 P}{\partial \bar{z}_{m+tc}\partial \bar{z}_{j+tc}} - p_{mj}^{02}\right\| \left\| f_{jn}^{20}\right\| \\
&\qquad + \sum_j \left\| \frac{\partial^2 F}{\partial \bar{z}_{n+tc}\partial \bar{z}_{j+tc}} - f_{nj}^{02}\right\| \left\| p_{jm}^{20}\right\| + \left\| \frac{\partial^2 F}{\partial \bar{z}_{m+tc}\partial \bar{z}_{j+tc}} - f_{mj}^{02}\right\| \left\| p_{jn}^{20}\right\| \\
&\qquad + \sum_j \left\| \frac{\partial^2 F}{\partial z_{n+tc}\partial \bar{z}_{j+tc}} - f_{nj}^{11}\right\| \left\| \frac{\partial^2 P}{\partial z_{m+tc}\partial \bar{z}_{j+tc}} - p_{mj}^{11}\right\| \\
&\qquad + \sum_j \left\| \frac{\partial^2 F}{\partial z_{m+tc}\partial \bar{z}_{j+tc}} - f_{mj}^{11}\right\| \left\| \frac{\partial^2 P}{\partial z_{n+tc}\partial \bar{z}_{j+tc}} - p_{nj}^{11}\right\| \\
&\qquad + \sum_j \left\| \frac{\partial^2 P}{\partial z_{n+tc}\partial z_{j+tc}} - p_{nj}^{20}\right\| \left\| \frac{\partial^2 F}{\partial \bar{z}_{m+tc}\partial \bar{z}_{j+tc}} - f_{mj}^{02}\right\| \\
&\qquad + \sum_j \left\| \frac{\partial^2 P}{\partial z_{m+tc}\partial z_{j+tc}} - p_{mj}^{20}\right\| \left\| \frac{\partial^2 F}{\partial \bar{z}_{n+tc}\partial \bar{z}_{j+tc}} - f_{nj}^{02}\right\| \\
&\qquad + \sum_j \left\| \frac{\partial^2 P}{\partial \bar{z}_{n+tc}\partial \bar{z}_{j+tc}} - p_{nj}^{02}\right\| \left\| \frac{\partial^2 F}{\partial z_{m+tc}\partial z_{j+tc}} - f_{mj}^{20}\right\| \\
&\qquad + \sum_j \left\| \frac{\partial^2 P}{\partial \bar{z}_{m+tc}\partial \bar{z}_{j+tc}} - p_{mj}^{20}\right\| \left\| \frac{\partial^2 F}{\partial z_{n+tc}\partial z_{j+tc}} - f_{nj}^{20}\right\| \\
&\quad \leqslant \frac{K^2 \varepsilon^{\frac{5}{3}}}{t} e^{-|n-m|\rho} + \frac{K^2 \varepsilon^{\frac{5}{3}}}{t^2} e^{-|n-m|\rho} \leqslant \frac{\varepsilon_+}{t} e^{-|n-m|\rho}.
\end{aligned}$$

In the above inequalities, j is bounded by K^2 due to

$$|j+m| \leqslant K \quad \text{and} \quad |j+n| \leqslant K, \quad \text{or} \quad |j-n| \leqslant K \quad \text{and} \quad |j-m| \leqslant K.$$

The other cases are proved similarly. □

4.5. Iteration lemma and convergence

For any given $s, \varepsilon, r, \gamma$ and for all $\nu \geqslant 1$, we define the following sequences

$$\begin{aligned}
r_\nu &= r\left(1 - \sum_{i=2}^{\nu+1} 2^{-i}\right), \\
\varepsilon_\nu &= c\big(\gamma^{-1} K_{\nu-1}^{\tau}\big)^4 \varepsilon_{\nu-1}^{\frac{4}{3}}, \\
\eta_\nu &= \varepsilon_\nu^{\frac{1}{3}}, \qquad L_\nu = L_{\nu-1} + \varepsilon_{\nu-1}, \\
s_\nu &= \frac{1}{4}\eta_{\nu-1} s_{\nu-1} = 2^{-2\nu}\left(\prod_{i=0}^{\nu-1} \varepsilon_i\right)^{\frac{1}{3}} s_0, \\
\rho_\nu &= \rho\left(1 - \sum_{i=2}^{\nu+1} 2^{-i}\right), \\
K_\nu &= c\big((\rho_{\nu-1} - \rho_\nu)^{-1} \ln \varepsilon_\nu^{-1}\big),
\end{aligned} \tag{4.28}$$

where c is a constant, and the parameters r_0, ε_0, L_0, s_0 and K_0 are defined to be r, ε, L, s and $\ln \frac{1}{\varepsilon}$ respectively.

4.5.1. Iteration lemma

The preceding analysis is summarized as follows.

Lemma 4.5. *Let ε be small enough and $\nu \geqslant 0$. Suppose that*

(1) $N_\nu + \mathcal{A}_\nu + \mathcal{B}_\nu + \bar{\mathcal{B}}_\nu = e_\nu + \langle \omega^\nu(\xi), I\rangle + \sum_n \Omega_n^\nu(\xi) z_n \bar{z}_n + \mathcal{A}_\nu + \mathcal{B}_\nu + \bar{\mathcal{B}}_\nu$ *is a normal form with parameters ξ on a closed set $\mathcal{O}_\nu$ of $\mathbb{R}^b$ satisfying*

$$\begin{gathered}
\big|\langle k, \omega^\nu\rangle\big| \geqslant \frac{\gamma}{K_\nu^\tau}, \quad 0 < |k| \leqslant K_\nu, \\
\big|\det\big(\langle k, \omega^\nu\rangle I + A_n^\nu\big)\big| \geqslant \frac{\gamma}{K_\nu^\tau}, \quad |k| \leqslant K_\nu, \\
\big|\det\big(\langle k, \omega^\nu\rangle I \pm A_n^\nu \otimes I \pm I \otimes A_{n'}^\nu\big)\big| \geqslant \frac{\gamma}{K_\nu^\tau}, \quad 0 < |k| \leqslant K_\nu,
\end{gathered}$$

where $A_n^\nu = \Omega_n^\nu$ for $n \in \mathbb{Z}_1^2 \setminus (\mathcal{L}_1 \cup \mathcal{L}_2)$,

$$A_n^\nu = \begin{pmatrix} \Omega_n^\nu + \omega_i^\nu & a_n^\nu \\ a_m^\nu & \Omega_m^\nu + \omega_j^\nu \end{pmatrix}, \quad n \in \mathcal{L}_1',$$
$$A_n^\nu = \begin{pmatrix} \Omega_n^\nu - \omega_i^\nu & -a_n^\nu \\ \bar{a}_m^\nu & -\Omega_m^\nu + \omega_j^\nu \end{pmatrix}, \quad n \in \mathcal{L}_2'.$$

Here (n, m) *are resonant pairs, and* (i, j) *is uniquely determined by* (n, m).

(2) $\omega^\nu(\xi)$, $\Omega_n^\nu(\xi)$ *are* C_W^4 *smooth in* ξ *satisfying*

$$\left|\omega^\nu - \omega^{\nu-1}\right|_{\mathcal{O}_\nu} \leqslant \varepsilon_{\nu-1}, \qquad \left|\Omega_n^\nu - \Omega_n^{\nu-1}\right|_{\mathcal{O}_\nu} \leqslant \varepsilon_{\nu-1}.$$

(3) $N_\nu + \mathcal{A}_\nu + \mathcal{B}_\nu + \bar{\mathcal{B}}_\nu + P_\nu$ *satisfies* (A5), (A6) *with* $K_\nu, \varepsilon_\nu, \rho_\nu$ *and*

$$\|X_{P_\nu}\|_{D(r_\nu, s_\nu), \mathcal{O}_\nu} \leqslant \varepsilon_\nu.$$

Then there is a subset $\mathcal{O}_{\nu+1} \subset \mathcal{O}_\nu$,

$$\mathcal{O}_{\nu+1} = \mathcal{O}_\nu \Big\backslash \Big(\bigcup_{K_\nu < |k| \leqslant K_{\nu+1}} \mathcal{R}_k^{\nu+1}(\gamma) \Big),$$

where $\mathcal{R}_k^{\nu+1}(\gamma)$ *is given in* (4.35) *with* $\omega^{\nu+1} = \omega^\nu + P_{0l00}^\nu$, *and a symplectic transformation of variables*

$$\Phi_\nu : D_{\rho_\nu}(r_{\nu+1}, s_{\nu+1}) \times \mathcal{O}_\nu \to D_{\rho_\nu}(r_\nu, s_\nu), \tag{4.29}$$

such that on $D_{\rho_{\nu+1}}(r_{\nu+1}, s_{\nu+1}) \times \mathcal{O}_{\nu+1}$, $H_{\nu+1} = H_\nu \circ \Phi_\nu$ *has the form*

$$H_{\nu+1} = e_{\nu+1} + \langle \omega^{\nu+1}, I \rangle + \sum_n \Omega_n^{\nu+1} z_n \bar{z}_n + \mathcal{A}_{\nu+1} + \mathcal{B}_{\nu+1} + \bar{\mathcal{B}}_{\nu+1} + P_{\nu+1}, \tag{4.30}$$

with

$$\left|\omega^{\nu+1} - \omega^\nu\right|_{\mathcal{O}_{\nu+1}} \leqslant \varepsilon_\nu, \qquad \left|\Omega_n^{\nu+1} - \Omega_n^\nu\right|_{\mathcal{O}_{\nu+1}} \leqslant \varepsilon_\nu. \tag{4.31}$$

And for ξ *in a closed subset* $\mathcal{O}_{\nu+1}$ *of* $\mathbb{R}^b$, *satisfies Diophantine condition*

$$\left|\langle k, \omega^{\nu+1} \rangle\right| \geqslant \frac{\gamma}{K_{\nu+1}^\tau}, \quad 0 < |k| \leqslant K_{\nu+1},$$
$$\left|\det\left(\langle k, \omega^{\nu+1} \rangle I + A_n^{\nu+1}\right)\right| \geqslant \frac{\gamma}{K_{\nu+1}^\tau}, \quad |k| \leqslant K_{\nu+1},$$
$$\left|\det\left(\langle k, \omega^{\nu+1} \rangle I \pm A_n^{\nu+1} \otimes I \pm I \otimes A_{n'}^{\nu+1}\right)\right| \geqslant \frac{\gamma}{K_{\nu+1}^\tau}, \quad 0 < |k| \leqslant K_{\nu+1},$$

where $A_n^{\nu+1} = \Omega_n^{\nu+1}$ *for* $n \in \mathbb{Z}_1^2 \setminus (\mathcal{L}_1 \cup \mathcal{L}_2)$,

$$A_n^{\nu+1} = \begin{pmatrix} \Omega_n^{\nu+1} + \omega_i^{\nu+1} & a_n^{\nu+1} \\ a_m^{\nu+1} & \Omega_m^{\nu+1} + \omega_j^{\nu+1} \end{pmatrix}, \quad n \in \mathcal{L}_1',$$

$$A_n^{\nu+1} = \begin{pmatrix} \Omega_n^{\nu+1} - \omega_i^{\nu+1} & -a_n^{\nu+1} \\ \bar{a}_m^{\nu+1} & -\Omega_m^{\nu+1} + \omega_j^{\nu+1} \end{pmatrix}, \quad n \in \mathcal{L}_2'.$$

Here (n,m) are resonant pairs, (i,j) is uniquely determined by (n,m).

And also $N_{\nu+1} + \mathcal{A}_{\nu+1} + \mathcal{B}_{\nu+1} + \bar{\mathcal{B}}_{\nu+1} + P_{\nu+1}$ has the special form defined in (A5), (A6) *with $K_{\nu+1}, \varepsilon_{\nu+1}, \rho_{\nu+1}$ in place of $K_\nu, \varepsilon_\nu, \rho_\nu$ and*

$$\|X_{P_{\nu+1}}\|_{D_{\rho_{\nu+1}}(r_{\nu+1}, s_{\nu+1}), \mathcal{O}_{\nu+1}} \leqslant \varepsilon_{\nu+1}. \tag{4.32}$$

4.5.2. Convergence

Suppose that the assumptions of Theorem 2 are satisfied. Recall that

$$\varepsilon_0 = \varepsilon, \qquad r_0 = r, \qquad s_0 = s, \qquad \rho_0 = \rho, \qquad L_0 = L, \qquad N_0 = N,$$
$$\mathcal{A}_0 = \mathcal{A}, \qquad \mathcal{B}_0 = \mathcal{B}, \qquad P_0 = P,$$

$\mathcal{O}$ is a bounded positive-measure set. The assumptions of the iteration lemma are satisfied when $\nu = 0$ if ε_0 and γ are sufficiently small. Inductively, we obtain the following sequences:

$$\mathcal{O}_{\nu+1} \subset \mathcal{O}_\nu,$$
$$\Psi^\nu = \Phi_0 \circ \Phi_1 \circ \cdots \circ \Phi_\nu : D_{\rho_\nu}(r_{\nu+1}, s_{\nu+1}) \times \mathcal{O}_\nu \to D_{\rho_0}(r_0, s_0), \quad \nu \geqslant 0,$$
$$H \circ \Psi^\nu = H_{\nu+1} = N_{\nu+1} + \mathcal{A}_{\nu+1} + \mathcal{B}_{\nu+1} + \bar{\mathcal{B}}_{\nu+1} + P_{\nu+1}.$$

Let $\tilde{\mathcal{O}} = \bigcap_{\nu=0}^{\infty} \mathcal{O}_\nu$. As in [21,22], thanks to Lemma 4.2, it concludes that $N_\nu, \Psi^\nu, D\Psi^\nu, \omega_\nu$ converge uniformly on $D_{\frac{1}{2}\rho}(\frac{1}{2}r, 0) \times \tilde{\mathcal{O}}$ with

$$N_\infty + \mathcal{A}_\infty + \mathcal{B}_\infty + \bar{\mathcal{B}}_\infty = e_\infty + \langle \omega^\infty, I \rangle + \sum_n \Omega_n^\infty z_n \bar{z}_n + \mathcal{A}_\infty + \mathcal{B}_\infty + \bar{\mathcal{B}}_\infty.$$

Since

$$\varepsilon_{\nu+1} = c\bigl(\gamma^{-1} K_\nu^\tau\bigr)^4 \varepsilon_\nu^{\frac{4}{3}}$$

it follows that $\varepsilon_{\nu+1} \to 0$ provided that ε is sufficiently small. And we also have $\sum_{\nu=0}^{\infty} \varepsilon_\nu \leqslant 2\varepsilon$.

Let ϕ_H^t be the flow of X_H. Since $H \circ \Psi^\nu = H_{\nu+1}$, we have

$$\phi_H^t \circ \Psi^\nu = \Psi^\nu \circ \phi_{H_{\nu+1}}^t. \tag{4.33}$$

The uniform convergence of $\Psi^\nu, D\Psi^\nu, \omega_\nu$ and X_{H_ν} implies that the limits can be taken on both sides of (4.33). Hence, on $D_{\frac{1}{2}\rho}(\frac{1}{2}r, 0) \times \tilde{\mathcal{O}}$ we get

$$\phi_H^t \circ \Psi^\infty = \Psi^\infty \circ \phi_{H_\infty}^t \tag{4.34}$$

and

$$\Psi^\infty : D_{\frac{1}{2}\rho}\left(\frac{1}{2}r, 0\right) \times \tilde{\mathcal{O}} \to D_\rho(r,s) \times \mathcal{O}.$$

It follows from (4.34) that

$$\phi_H^t\left(\Psi^\infty\left(\mathbb{T}^b \times \{\xi\}\right)\right) = \Psi^\infty\left(\mathbb{T}^b \times \{\xi\}\right)$$

for $\xi \in \tilde{\mathcal{O}}$. This means that $\Psi^\infty(\mathbb{T}^b \times \{\xi\})$ is an embedded torus which is invariant for the original perturbed Hamiltonian system at $\xi \in \tilde{\mathcal{O}}$. We remark here that the frequencies $\omega^\infty(\xi)$ associated to $\Psi^\infty(\mathbb{T}^b \times \{\xi\})$ are slightly different from $\omega(\xi)$. The normal behavior of the invariant torus is governed by normal frequencies Ω_n^∞.

4.6. Measure estimates

For notational convenience, let $\mathcal{O}_{-1} = \mathcal{O}$, $K_{-1} = 0$. Then at νth step of KAM iteration, we have to exclude the following resonant set

$$\mathcal{R}^\nu = \bigcup_{K_{\nu-1} < |k| \leqslant K_\nu, n, n'} \left(\mathcal{R}_k^\nu \cup \mathcal{R}_{kn}^\nu \cup \mathcal{R}_{knn'}^\nu\right) \tag{4.35}$$

where

$$\mathcal{R}_k^\nu = \left\{\xi \in \mathcal{O}_{\nu-1}\colon \left|\left\langle k, \omega^\nu(\xi)\right\rangle\right| < \frac{\gamma}{K_\nu^\tau}\right\}, \tag{4.36}$$

$$\mathcal{R}_{kn}^\nu = \left\{\xi \in \mathcal{O}_{\nu-1}\colon \left|\det\left(\left\langle k, \omega^\nu\right\rangle I + A_n^\nu\right)\right| < \frac{\gamma}{K_\nu^\tau}\right\}, \tag{4.37}$$

$$\mathcal{R}_{knn'}^\nu = \left\{\xi \in \mathcal{O}_{\nu-1}\colon \left|\det\left(\left\langle k, \omega^\nu\right\rangle I \pm A_n^\nu \otimes I \pm I \otimes A_{n'}^\nu\right)\right| < \frac{\gamma}{K_\nu^\tau}\right\}. \tag{4.38}$$

In the following, we only give the proof for the most complicated case: $\{\xi \in \mathcal{O}_{\nu-1}\colon |\det(\langle k, \omega^\nu\rangle I + A_n^\nu \otimes I - I \otimes A_{n'}^\nu)| < \frac{\gamma}{K_\nu^\tau}\}$.

Set $M^\nu = \langle k, \omega^\nu\rangle I + A_n^\nu \otimes I - I \otimes A_{n'}^\nu$, $M^{\nu-1} = \langle k, \omega^{\nu-1}\rangle I + A_n^{\nu-1} \otimes I - I \otimes A_{n'}^{\nu-1}$, then for $|k| \leqslant K_{\nu-1}$

$$\begin{aligned}\left\|\left(M^\nu\right)^{-1}\right\| &= \left\|\left(M^{\nu-1} + \left(M^\nu - M^{\nu-1}\right)\right)^{-1}\right\| \\ &= \left\|\left(I + \left(M^{\nu-1}\right)^{-1}\left(M^\nu - M^{\nu-1}\right)\right)^{-1}\left(M^{\nu-1}\right)^{-1}\right\| \\ &\leqslant 2\left\|\left(M^{\nu-1}\right)^{-1}\right\| \leqslant 2\frac{K_{\nu-1}^\tau}{\gamma} \leqslant \frac{K_\nu^\tau}{\gamma}.\end{aligned}$$

For $K_{\nu-1} < |k| \leqslant K_\nu$, we consider $n, n' \in \mathcal{L}_1$ as an example, the other cases can be proved analogously. Assume that (n, m) and (n', m') are resonant pairs in $\mathcal{L}_1$, then

$$\begin{aligned}
&\langle k,\omega^\nu\rangle I + A_n^\nu\otimes I_2 - I_2\otimes A_{n'}^\nu\\
&\quad = \mathrm{Diag}\begin{pmatrix}
\langle k,\omega^\nu\rangle + (\Omega_n^\nu+\omega_i^\nu) - (\Omega_{n'}^\nu+\omega_{i'}^\nu)\\
\langle k,\omega^\nu\rangle + (\Omega_n^\nu+\omega_i^\nu) - (\Omega_{m'}^\nu+\omega_{j'}^\nu)\\
\langle k,\omega^\nu\rangle + (\Omega_m^\nu+\omega_j^\nu) - (\Omega_{n'}^\nu+\omega_{i'}^\nu)\\
\langle k,\omega^\nu\rangle + (\Omega_m^\nu+\omega_j^\nu) - (\Omega_{m'}^\nu+\omega_{j'}^\nu)
\end{pmatrix}\\
&\qquad + \begin{pmatrix} 0 & a_n^\nu\\ a_m^\nu & 0\end{pmatrix}\otimes I_2 - I_2\otimes\begin{pmatrix} 0 & a_{n'}^\nu\\ a_{m'}^\nu & 0\end{pmatrix}.
\end{aligned}$$

Lemma 4.6. *For any given $n,n'\in\mathbb{Z}_1^2$ with $|n-n'|\leqslant K_\nu$, either $|\det(\langle k,\omega^\nu\rangle I + A_n^\nu\otimes I_2 - I_2\otimes A_{n'}^\nu)|>1$ or there are n_0, n_0', c with $|n_0|,|n_0'|,|c|\leqslant 3K_\nu$ and $t_0\in\mathbb{Z}$, such that $n=n_0+tc$, $n'=n_0'+tc$.*

Proof. Since $|n-n'|\leqslant K_\nu$, with an elementary calculation

$$|n|^2-\left|n'\right|^2=\left|n-n'\right|^2+2\langle n-n',n'\rangle.$$

If $|\langle n-n',n'\rangle|>K_\nu^2$, we have $|\det(\langle k,\omega^\nu\rangle I + A_n^\nu\otimes I_2 - I_2\otimes A_{n'}^\nu)|>1$, there will be no small divisor problem.

In the case that $|\langle n-n',n'\rangle|\leqslant K_\nu^2$, we choose $c\in\mathbb{Z}^2$ such that $c\perp n-n'$ and $|c|=|n-n'|\leqslant K_\nu$. It is easy to see that there is a $t_0\in\mathbb{Z}$ such that $|n'-ct_0|\leqslant 2K_\nu$. Take $n_0'=n'-ct_0$ and $n_0=n_0'+n-n'$. We have $|n_0'|\leqslant 2K_\nu$ and

$$|n_0|\leqslant\left|n-n'\right|+\left|n_0'\right|\leqslant 3K_\nu. \qquad\square$$

Lemma 4.7.

$$\bigcup_{n,n'\in\mathbb{Z}_1^2}\mathcal{R}_{knn'}^\nu\subset\bigcup_{n_0,n_0',c\in\mathbb{Z}^2,t\in\mathbb{Z}}\mathcal{R}_{k,n_0+tc,n_0'+tc}^\nu \tag{4.39}$$

where $|n_0|,|n_0'|,|c|\leqslant 3K_\nu$.

Proof. If $|\langle n-n',n'\rangle|>K_\nu^2$, $\mathcal{R}_{knn'}^\nu=\emptyset$. If $|\langle n-n',n'\rangle|\leqslant K_\nu^2$, there exist n_0,n_0',c with $|n_0|,|n_0'|,|c|\leqslant 3K_\nu$ such that $n=n_0+tc$, $n'=n_0'+tc$. Hence

$$\bigcup_{n,n'\in\mathbb{Z}_1^2}\mathcal{R}_{knn'}^\nu\subset\bigcup_{n_0,n_0',c\in\mathbb{Z}^2,t\in\mathbb{Z}}\mathcal{R}_{k,n_0+tc,n_0'+tc}^\nu \tag{4.40}$$

where $|n_0|,|n_0'|,|c|\leqslant 3K_\nu$. $\square$

Lemma 4.8. *(See Lemma 8.4 of [1].) Let $g:\mathcal{I}\to\mathbb{R}$ be $b+3$ times differentiable, and assume that*

(1) $\forall\sigma\in\mathcal{I}$ *there exists $s\leqslant b+2$ such that $g^{(s)}(\sigma)>B$.*

(2) *There exists A such that $|g^{(s)}(\sigma)| \leqslant A$ for $\forall \sigma \in \mathcal{I}$ and $\forall s$ with $1 \leqslant s \leqslant b+3$.*
Define

$$\mathcal{I}_h \equiv \{\sigma \in \mathcal{I}\colon\ |g(\sigma)| \leqslant h\},$$

then

$$\frac{\mathrm{meas}(\mathcal{I}_h)}{\mathrm{meas}(\mathcal{I})} \leqslant \frac{A}{B} 2\bigl(2+3+\cdots+(b+3)+2B^{-1}\bigr) h^{\frac{1}{b+3}}.$$

For a proof see [1].

Lemma 4.9. *For fixed k, n_0, n_0', c,*

$$\mathrm{meas}\biggl(\bigcup_{t\in\mathbb{Z}} \mathcal{R}^{\nu}_{k,n_0+tc,n_0'+tc}\biggr) < \frac{\gamma^{\frac{1}{4}}}{K_\nu^{\frac{\tau}{20}}}.$$

Proof. Due to the analysis above and Töplitz–Lipschitz property of $N+\mathcal{A}+\mathcal{B}+\bar{\mathcal{B}}+P$, the coefficient matrix $M^\nu(t)$ has a limit as $t \to \infty$,

$$\Bigl\| M^\nu(t) - \lim_{t\to\infty} M^\nu(t)\Bigr\| \leqslant \frac{\varepsilon_0}{t}.$$

We define resonant set

$$\mathcal{R}^{\nu}_{kn_0n_0'c\infty} = \biggl\{\xi \in \mathcal{O}_{\nu-1}\colon\ \Bigl|\det \lim_{t\to\infty} M^\nu(t)\Bigr| < \frac{\gamma}{K_\nu^{\frac{\tau}{5}}}\biggr\}. \tag{4.41}$$

Then for $\xi \in \mathcal{O}_{\nu-1} \backslash \mathcal{R}^{\nu}_{kn_0n_0'c\infty}$, we have $\|(\lim_{t\to\infty} M^\nu(t))^{-1}\| \leqslant \frac{K_\nu^{\frac{\tau}{5}}}{\gamma}$.
Since

$$\Bigl\| M^\nu(t) - \lim_{t\to\infty} M^\nu(t)\Bigr\| \leqslant \frac{\varepsilon_0}{t},$$

for $|t| > K_\nu^{\frac{\tau}{5}}$, we have

$$\bigl\|\bigl(M^\nu\bigr)^{-1}(t)\bigr\| \leqslant 2\frac{K_\nu^{\frac{\tau}{5}}}{\gamma} \leqslant \frac{K_\nu^{\tau}}{\gamma}.$$

For $|t| \leqslant K_\nu^{\frac{\tau}{5}}$, we define the resonant set

$$\mathcal{R}^{\nu}_{kn_0n_0'ct} = \biggl\{\xi \in \mathcal{O}_{\nu-1}\colon\ \bigl|\det M^\nu(t)\bigr| < \frac{\gamma}{K_\nu^{\tau}}\biggr\}. \tag{4.42}$$

In addition

$$\inf_{\xi\in\mathcal{O}} \max_{0<d\leqslant 4} \bigl|\partial_\xi^d\bigl(\det M^\nu(t)\bigr)\bigr| \geqslant \frac{1}{2}|k| \geqslant \frac{1}{2}K.$$

In view of Lemma 4.8, we have

$$\mathrm{meas}\{\mathcal{R}^{\nu}_{kn_0n'_0c\infty}\} < \left(\frac{\gamma}{K_\nu^{\frac{\tau}{5}}}\right)^{\frac{1}{4}} = \frac{\gamma^{\frac{1}{4}}}{K_\nu^{\frac{\tau}{20}}},$$

$$\mathrm{meas}\left\{\bigcup_{|t|\leqslant K_\nu^{\frac{\tau}{5}}}\mathcal{R}_{kn_0n'_0ct}\right\} < K_\nu^{\frac{\tau}{5}}\left(\frac{\gamma}{K_\nu^{\tau}}\right)^{\frac{1}{4}} \leqslant \frac{\gamma^{\frac{1}{4}}}{K_\nu^{\frac{\tau}{20}}}.$$

Hence

$$\mathrm{meas}\left(\bigcup_t \mathcal{R}^{\nu}_{k,n_0+tc,n'_0+tc}\right) < \frac{\gamma^{\frac{1}{4}}}{K_\nu^{\frac{\tau}{20}}}. \qquad \square$$

Lemma 4.10.

$$\mathrm{meas}\left(\bigcup_{K_{\nu-1}<|k|\leqslant K_\nu} R^{\nu}_{k}\right) \leqslant K_\nu^{b}\left(\frac{\gamma}{K_\nu^{\frac{\tau}{5}}}\right)^{\frac{1}{4}} = \frac{\gamma^{\frac{1}{4}}}{K_\nu^{\frac{\tau}{20}-b}},$$

$$\mathrm{meas}\left(\bigcup_{K_{\nu-1}<|k|\leqslant K_\nu,n} R^{\nu}_{kn}\right) \leqslant K_\nu^{3+b}\left(\frac{\gamma}{K_\nu^{\frac{\tau}{5}}}\right)^{\frac{1}{4}} = \frac{\gamma^{\frac{1}{4}}}{K_\nu^{\frac{\tau}{20}-3-b}},$$

$$\mathrm{meas}\left(\bigcup_{K_{\nu-1}<|k|\leqslant K_\nu,n,n'} R^{\nu}_{knn'}\right) \leqslant \frac{\gamma^{\frac{1}{4}}}{K_\nu^{\frac{\tau}{20}-3-b}}.$$

Lemma 4.11. *Let $\tau > 20(b+4)$, then the total measure need to exclude along the KAM iteration is*

$$\mathrm{meas}\left(\bigcup_{\nu\geqslant 0}\mathcal{R}^{\nu}\right) = \mathrm{meas}\left[\bigcup_{\nu\geqslant 0}\left(\bigcup_{K_{\nu-1}<|k|\leqslant K_\nu,n,n'}\mathcal{R}^{\nu}_{k}\cup\mathcal{R}^{\nu}_{kn}\cup\mathcal{R}^{\nu}_{knn'}\right)\right]$$
$$\leqslant \sum_{\nu\geqslant 0}\frac{\gamma^{\frac{1}{4}}}{K_\nu} \leqslant \gamma^{\frac{1}{4}}.$$

Appendix A. A precise way to construct the admissible tangential sites

For any given positive integer b, we give a concrete way to construct the admissible tangential sites $S = \{i_1 = (x_1, y_1), i_2 = (x_2, y_2), \dots, i_b = (x_b, y_b)\}$. Firstly we choose x_1, y_1, x_2, y_2 such that $x_1 > b^2$, $y_1 = x_1^5$, $x_2 = y_1^5$, $y_2 = x_2^5$, the others are defined inductively by

$$x_{j+1} = y_j^5 \prod_{2\leqslant m\leqslant j,1\leqslant l<m}\left((x_m - x_l)^2 + (y_m - y_l)^2\right), \quad 2\leqslant j\leqslant b-1,$$

$$y_{j+1} = x_{j+1}^5, \quad 2\leqslant j\leqslant b-1.$$

Lemma A.1. *The tangential sites given above is admissible.*

Proof. The proof is elementary but cumbersome. Firstly, because for any three points $c, d, f \in S$, we have

$$\frac{c_1 - d_1}{c_2 - d_2} > 0, \qquad \frac{d_2 - f_2}{d_1 - f_1} > 0,$$

hence

$$\frac{c_1 - d_1}{c_2 - d_2} + \frac{d_2 - f_2}{d_1 - f_1} > 0,$$

thus

$$\langle c - d, d - f\rangle = (c_1 - d_1)(d_1 - f_1) + (c_2 - d_2)(d_2 - f_2) \neq 0.$$

As a result, any three points in S cannot be three vertices of a rectangle.

Note that $n - m + i - j = 0$, $|n|^2 - |m|^2 + |i|^2 - |j|^2 = 0$ implies $\langle n - j, j - i\rangle = 0$ and $n + m - c - d = 0$, $|n|^2 + |m|^2 - |c^2| - |d|^2 = 0$ implies $\langle n - c, n - d\rangle = 0$. To prove that S is admissible, it suffices to prove that

$$\begin{cases} \langle n - g, g - f\rangle = 0, \\ \langle n - c, c - d\rangle = 0, \end{cases} \tag{A.1}$$

$$\begin{cases} \langle n - g, n - f\rangle = 0, \\ \langle n - c, n - d\rangle = 0, \end{cases} \tag{A.2}$$

$$\begin{cases} \langle n - g, g - f\rangle = 0, \\ \langle n - c, n - d\rangle = 0, \end{cases} \tag{A.3}$$

have no solution in $\mathbb{Z}_1^2$ for $c, d, f, g \in S$ and $\{c, d\} \neq \{f, g\}$.

Write Eqs. (A.1) in detail

$$\begin{cases} (n_1 - g_1)(g_1 - f_1) + (n_2 - g_2)(g_2 - f_2) = 0, \\ (n_1 - c_1)(c_1 - d_1) + (n_2 - c_2)(c_2 - d_2) = 0. \end{cases} \tag{A.4}$$

We prove that (A.4) has no solution in $\mathbb{Z}_1^2$ by contradiction.

(I) We consider the case that only one of $\{|c|, |d|, |f|, |g|\}$ reaches the maximum value of them.

(1) $|d| = \max\{|c|, |d|, |f|, |g|\}$.

By an elementary calculation, we have

$$\begin{aligned} n_2 &= \frac{(g_1 - c_1)(g_1 - f_1)(c_1 - d_1) + g_2(g_2 - f_2)(c_1 - d_1) - c_2(c_2 - d_2)(g_1 - f_1)}{(c_1 - d_1)(g_2 - f_2) - (c_2 - d_2)(g_1 - f_1)} \\ &= c_2 + \frac{(g_1 - c_1)(g_1 - f_1)(c_1 - d_1) + (g_2 - c_2)(c_1 - d_1)(g_2 - f_2)}{(c_1 - d_1)(g_2 - f_2) - (c_2 - d_2)(g_1 - f_1)}. \end{aligned}$$

According to the choice of the tangential sites S, if $g = c$, then $n = c \notin \mathbb{Z}_1^2$; if $g \neq c$, the numerator is of order d_1 and the divisor is of order d_2, which concludes that $n_2 \notin \mathbb{Z}$.

(2) $|f| = \max\{|c|, |d|, |f|, |g|\}$.

We have

$$n_2 = \frac{(g_1 - c_1)(g_1 - f_1)(c_1 - d_1) + g_2(g_2 - f_2)(c_1 - d_1) - c_2(c_2 - d_2)(g_1 - f_1)}{(c_1 - d_1)(g_2 - f_2) - (c_2 - d_2)(g_1 - f_1)}$$
$$= g_2 + \frac{(g_1 - c_1)(g_1 - f_1)(c_1 - d_1) + (g_2 - c_2)(c_2 - d_2)(g_1 - f_1)}{(c_1 - d_1)(g_2 - f_2) - (c_2 - d_2)(g_1 - f_1)}.$$

By same analysis as in the above case, we have $n \notin \mathbb{Z}_1^2$.

(3) $|g| = \max\{|c|, |d|, |f|, |g|\}$.

We have

$$\begin{aligned}
n_2 &= \frac{(g_1 - c_1)(g_1 - f_1)(c_1 - d_1) + g_2(g_2 - f_2)(c_1 - d_1) - c_2(c_2 - d_2)(g_1 - f_1)}{(c_1 - d_1)(g_2 - f_2) - (c_2 - d_2)(g_1 - f_1)} \\
&= g_2 + \frac{(g_1 - c_1)(g_1 - f_1)(c_1 - d_1) + (g_2 - c_2)(c_2 - d_2)(g_1 - f_1)}{(c_1 - d_1)(g_2 - f_2) - (c_2 - d_2)(g_1 - f_1)} \\
&= g_2 + \frac{(g_1 - c_1)(g_1 - f_1)(c_1 - d_1) + (g_1^5 - c_1^5)(c_1^5 - d_1^5)(g_1 - f_1)}{(c_1 - d_1)(g_1^5 - f_1^5) - (c_1^5 - d_1^5)(g_1 - f_1)} \\
&= g_2 + \frac{(g_1 - c_1) + (g_1^5 - c_1^5)(c_1^4 + c_1^3 d_1 + c_1^2 d_1^2 + c_1 d_1^3 + d_1^4)}{(g_1^4 + g_1^3 f_1 + g_1^2 f_1^2 + g_1 f_1^3 + f_1^4) - (c_1^4 + c_1^3 d_1 + c_1^2 d_1^2 + c_1 d_1^3 + d_1^4)} \\
&= g_2 + \left(c_1^4 + c_1^3 d_1 + c_1^2 d_1^2 + c_1 d_1^3 + d_1^4\right) g_1 - \left(c_1^4 + c_1^3 d_1 + c_1^2 d_1^2 + c_1 d_1^3 + d_1^4\right) f_1 \\
&\quad + \frac{g_1 - c_1 + (g_1 - f_1)(c_1^4 + c_1^3 d_1 + c_1^2 d_1^2 + c_1 d_1^3 + d_1^4)^2 + (f_1^5 - c_1^5)(c_1^4 + c_1^3 d_1 + c_1^2 d_1^2 + c_1 d_1^3 + d_1^4)}{(g_1^4 + g_1^3 f_1 + g_1^2 f_1^2 + g_1 f_1^3 + f_1^4) - (c_1^4 + c_1^3 d_1 + c_1^2 d_1^2 + c_1 d_1^3 + d_1^4)}.
\end{aligned}$$

Note that

$$\frac{g_1 - c_1 + (g_1 - f_1)(c_1^4 + c_1^3 d_1 + c_1^2 d_1^2 + c_1 d_1^3 + d_1^4)^2 + (f_1^5 - c_1^5)(c_1^4 + c_1^3 d_1 + c_1^2 d_1^2 + c_1 d_1^3 + d_1^4)}{(g_1^4 + g_1^3 f_1 + g_1^2 f_1^2 + g_1 f_1^3 + f_1^4) - (c_1^4 + c_1^3 d_1 + c_1^2 d_1^2 + c_1 d_1^3 + d_1^4)} \in (0, 1).$$

Hence $n_2 \notin \mathbb{Z}$.

(4) $|c| = \max\{|c|, |d|, |f|, |g|\}$.

This proof is the same as **(3)**.

(II) Secondly, we consider the case that two of $\{|c|, |d|, |f|, |g|\}$ reach the maximum of them.

(1) $|d| = |g| = \max\{|c|, |d|, |f|, |g|\}$.

We have $d = g$. So we only need to prove that

$$\langle n-g, g-f\rangle = 0, \qquad \langle n-c, c-g\rangle = 0$$

have no solution in $\mathbb{Z}^2$.

From above equation, we have

$$(n_1-g_1)(g_1-f_1)+(n_2-g_2)(g_2-f_2)=0,$$
$$(n_1-c_1)(c_1-g_1)+(n_2-c_2)(c_2-g_2)=0.$$

By elementary calculation, we have

$$\begin{aligned}
n_2 &= \frac{g_2(g_2-f_2)(c_1-g_1)-c_2(c_2-g_2)(g_1-f_1)-(g_1-c_1)^2(g_1-f_1)}{(g_2-f_2)(c_1-g_1)-(c_2-g_2)(g_1-f_1)}\\
&= \frac{g_1^5(g_1^5-f_1^5)(c_1-g_1)-c_1^5(c_1^5-g_1^5)(g_1-f_1)-(g_1-c_1)^2(g_1-f_1)}{(g_1^5-f_1^5)(c_1-g_1)-(c_1^5-g_1^5)(g_1-f_1)}\\
&= \frac{g_1^5(g_1^4+g_1^3f_1+g_1^2f_1^2+g_1f_1^3+f_1^4)-c_1^5(g_1^4+g_1^3c_1+g_1^2c_1^2+g_1c_1^3+c_1^4)+(g_1-c_1)}{(g_1^4+g_1^3f_1+g_1^2f_1^2+g_1f_1^3+f_1^4)-(g_1^4+g_1^3c_1+g_1^2c_1^2+g_1c_1^3+c_1^4)}.
\end{aligned}$$

Without loss of generality, we assume that $|c| < |f|$. According to the choice of tangential sites S,

$$c_1^4 \mid \big[\big(g_1^4+g_1^3f_1+g_1^2f_1^2+g_1f_1^3+f_1^4\big)-\big(g_1^4+g_1^3c_1+g_1^2c_1^2+g_1c_1^3+c_1^4\big)\big],$$
$$c_1^4 \mid \big[g_1^5\big(g_1^4+g_1^3f_1+g_1^2f_1^2+g_1f_1^3+f_1^4\big)-c_1^5\big(g_1^4+g_1^3c_1+g_1^2c_1^2+g_1c_1^3+c_1^4\big)+g_1\big], \quad c_1^4 \nmid c_1,$$

hence

$$\frac{g_1^5(g_1^4+g_1^3f_1+g_1^2f_1^2+g_1f_1^3+f_1^4)-c_1^5(g_1^4+g_1^3c_1+g_1^2c_1^2+g_1c_1^3+c_1^4)+(g_1-c_1)}{(g_1^4+g_1^3f_1+g_1^2f_1^2+g_1f_1^3+f_1^4)-(g_1^4+g_1^3c_1+g_1^2c_1^2+g_1c_1^3+c_1^4)} \notin \mathbb{Z}.$$

(2) $|d| = |f| = \max\{|c|, |d|, |f|, |g|\}$.

We have $d = f$. So

$$\langle n-g, g-f\rangle = \langle n-g, g-d\rangle = 0, \qquad \langle n-c, c-d\rangle = 0.$$

Hence $c, d, g \in S$ are three vertices of a rectangle, which is impossible.

(3) $|c| = |g| = \max\{|c|, |d|, |f|, |g|\}$.

We have $c = g$. From

$$\langle n-g, g-f\rangle = 0, \qquad \langle n-c, c-d\rangle = \langle n-g, g-d\rangle = 0,$$

thus the vectors $g - f$ and $g - d$ is parallel each other. Therefore $d, f, g \in S$ lie on the same line, which contradicts with the choice of tangential sites S.

Now we prove that (A.3) has no solution in $\mathbb{Z}^2$.

(III) Firstly, we consider the case that only one of $\{|c|,|d|,|f|,|g|\}$ reaches the maximum value of them and c,d,f,g are different from each other.

(1) $|d|=\max\{|c|,|d|,|f|,|g|\}$.

From the above two equations, we have

$$\langle n-c,n-d\rangle=0,\qquad \langle n-g,g-f\rangle=0.$$

We take g to be the origin. Then n,c,d,f will change to be $n-g,c-g,d-g,f-g$, however $|d-g|\gg|f-g|+|c-g|$. This condition is enough in this part. For simplicity we still use n,c,d,f to substitute $n-g,c-g,d-g,f-g$.

From $\langle n-c,n-d\rangle=0$, $\langle n,f\rangle=0$, we have

$$\bigl(f_1^2+f_2^2\bigr)n_1^2+\bigl((c_2+d_2)f_2f_1-f_2^2(c_1+d_1)\bigr)n_1+f_2^2c_2d_2+f_2^2c_1d_1=0.$$

Let

$$\begin{aligned}\Delta&=\bigl((c_2+d_2)f_2f_1-(c_1+d_1)f_2^2\bigr)^2-4\bigl(f_1^2+f_2^2\bigr)\bigl(f_2^2c_2d_2+f_2^2c_1d_1\bigr)\\&=\bigl(f_1f_2d_2+\bigl(c_2f_2f_1-f_2^2c_1-f_2^2d_1\bigr)\bigr)^2-4\bigl(f_1^2+f_2^2\bigr)\bigl(f_2^2c_2d_2+f_2^2c_1d_1\bigr)\\&=\left(f_1f_2d_2+\left(c_2f_2f_1-f_2^2c_1-f_2^2d_1-\frac{2c_2(f_1^2+f_2^2)f_2}{f_1}\right)\right)^2-\left(\frac{2c_2(f_1^2+f_2^2)f_2}{f_1}\right)^2\\&\quad+\frac{4c_2(f_1^2+f_2^2)f_2}{f_1}\bigl(c_2f_2f_1-f_2^2c_1-f_2^2d_1\bigr)-4\bigl(f_1^2+f_2^2\bigr)f_2^2c_1d_1.\end{aligned}$$

Since $\frac{4c_2(f_1^2+f_2^2)f_2}{f_1}(c_2f_2f_1-f_2^2c_1-f_2^2d_1)-4(f_1^2+f_2^2)f_2^2c_1d_1-(\frac{2c_2(f_1^2+f_2^2)f_2}{f_1})^2\in\mathbb{Z}$ is of order $\frac{4c_2f_2^5d_1}{f_1}$ which is far less than d_2, we may assume

$$\Delta=\left(f_1f_2d_2+\left(c_2f_2f_1-f_2^2c_1-f_2^2d_1-\frac{2c_2(f_1^2+f_2^2)f_2}{f_1}\right)-\alpha\right)^2$$

where $\alpha\sim\frac{d_1}{d_2}\ll\frac{1}{f_1}$. Thus

$$\begin{aligned}n_1&=\frac{-((c_2+d_2)f_2f_1-f_2^2(c_1+d_1))\pm\sqrt{\Delta}}{2(f_1^2+f_2^2)}\\&=\frac{-((c_2+d_2)f_2f_1-f_2^2(c_1+d_1))\pm(f_1f_2d_2+(c_2f_2f_1-f_2^2c_1-f_2^2d_1-\frac{2c_2(f_1^2+f_2^2)f_2}{f_1})-\alpha)}{2(f_1^2+f_2^2)}.\end{aligned}$$

Since $0<\frac{\alpha}{2(f_1^2+f_2^2)}\ll\frac{1}{2f_1(f_1^2+f_2^2)}$, we have $n_1\notin\mathbb{Z}$.

(2) $|f|=\max\{|c|,|d|,|f|,|g|\}$.

As before, we take g to be the origin. Solving (A.3), we get

$$\bigl(f_1^2+f_2^2\bigr)n_1^2+\bigl((c_2+d_2)f_2f_1-f_2^2(c_1+d_1)\bigr)n_1+f_2^2c_2d_2+f_2^2c_1d_1=0.\tag{A.5}$$

Eq. (A.5) has no solution since

$$\begin{aligned}\Delta &= \big((c_2+d_2)f_2f_1-(c_1+d_1)f_2^2\big)^2-4\big(f_1^2+f_2^2\big)\big(f_2^2c_2d_2+f_2^2c_1d_1\big)\\ &=(c_1-d_1)^2f_2^4-4c_2d_2f_2^4-2(c_1+d_1)(c_2+d_2)f_2^3f_1\\ &\quad+(c_2-d_2)^2f_1^2f_2^2-4f_1^2f_2^2c_1d_1<0.\end{aligned}$$

(3) $|g|=\max\{|c|,|d|,|f|,|g|\}$.

From $|n|^2-|m'|^2+|f|^2-|g|^2=0$, we have $|n|^2+|f|^2=|g|^2+|m'|^2$, which lead to $|n|^2>|g|^2-|f|^2$. Finally, we get

$$|n|^2+|m|^2-|c|^2-|d|^2>0.$$

(IV) Secondly, we consider the case that only one of $\{|c|,|d|,|f|,|g|\}$ reach the maximum of them and two of the remaining are same.

(1) $|d|=\max\{|c|,|d|,|f|,|g|\}$.

If $c=g$, we should solve

$$\langle n-c,n-d\rangle=0,\qquad \langle n-c,f-c\rangle=0.$$

With an elementary calculation, we have

$$\begin{aligned}n_2&=\frac{(f_1-c_1)^2d_2+c_2(f_2-c_2)^2-(f_2-c_2)(f_1-c_1)(d_1-c_1)}{(f_1-c_1)^2+(f_2-c_2)^2}\\ &=c_2+\frac{(f_1-c_1)^2d_2-(f_2-c_2)(f_1-c_1)d_1}{(f_1-c_1)^2+(f_2-c_2)^2}+\frac{-c_2(f_1-c_1)^2+(f_2-c_2)(f_1-c_1)c_1}{(f_1-c_1)^2+(f_2-c_2)^2}\\ &=c_2+\frac{(f_1-c_1)^2d_2-(f_2-c_2)(f_1-c_1)d_1}{(f_1-c_1)^2+(f_2-c_2)^2}+\frac{(f_1-c_1)(c_1f_2-c_2f_1)}{(f_1-c_1)^2+(f_2-c_2)^2}\\ &=c_2+\frac{(f_1-c_1)^2d_2-(f_2-c_2)(f_1-c_1)d_1}{(f_1-c_1)^2+(f_2-c_2)^2}+\frac{(f_1-c_1)(c_1f_1^5-c_1^5f_1)}{(f_1-c_1)^2+(f_1^5-c_1^5)^2}\\ &=c_2+\frac{(f_1-c_1)^2d_2-(f_2-c_2)(f_1-c_1)d_1}{(f_1-c_1)^2+(f_2-c_2)^2}+\frac{c_1f_1(f_1^3+f_1^2c_1+f_1c_1^2+c_1^3)}{1+(f_1^4+f_1^3c_1+f_1^2c_1^2+f_1c_1^3+c_1^4)^2}.\end{aligned}$$

According to the choice of tangential sites S,

$$\frac{(f_1-c_1)^2d_2-(f_2-c_2)(f_1-c_1)d_1}{(f_1-c_1)^2+(f_2-c_2)^2}\in\mathbb{Z},\qquad \frac{c_1f_1(f_1^3+f_1^2c_1+f_1c_1^2+c_1^3)}{1+(f_1^4+f_1^3c_1+f_1^2c_1^2+f_1c_1^3+c_1^4)^2}\in(0,1),$$

hence $n_2\notin\mathbb{Z}$.

If $c=f$, the proof is similar to the case **III(1)**.

(2) $|f|=\max\{|c|,|d|,|f|,|g|\}$.

If $g=c$, we have

$$\langle n-c,n-d\rangle=0,\qquad \langle n-c,c-f\rangle=0.$$

If we take c to be the origin as before, we have

$$n_2 = \frac{f_1^2 d_2 - f_2 f_1 d_1}{f_1^2 + f_2^2} \notin \mathbb{Z}.$$

The proof is the same as the case $g = d$.

(3) $|g| = \max\{|c|, |d|, |f|, |g|\}$.

The proof is the same as the case **III(3)**.

(V) Thirdly, we consider the case when two of $\{|c|, |d|, |f|, |g|\}$ reach the maximum.

(1) $|g| = |d| = \max\{|c|, |d|, |f|, |g|\}$.

Obviously we have $g = d$. In this case, we have the following relations

$$\langle n - c, n - d\rangle = (n_1 - c_1)(n_1 - d_1) + (n_2 - c_2)(n_2 - d_2) = 0,$$
$$\langle d - f, n - d\rangle = (d_1 - f_1)(n_1 - d_1) + (d_2 - f_2)(n_2 - d_2) = 0.$$

Hence

$$\frac{n_1 - d_1}{n_2 - d_2} = -\frac{n_2 - c_2}{n_1 - c_1} = -\frac{d_2 - f_2}{d_1 - f_1},$$

i.e.,

$$(d_1 - f_1)(n_2 - c_2) - (d_2 - f_2)(n_1 - c_1) = 0.$$

Thus we have

$$(d_1 - f_1)(n_1 - d_1) + (d_2 - f_2)(n_2 - d_2) = 0,$$
$$(d_1 - f_1)(n_2 - c_2) - (d_2 - f_2)(n_1 - c_1) = 0.$$

An elementary calculation shows that

$$\begin{aligned} n_1 &= \frac{d_1(d_1 - f_1)^2 + c_1(d_2 - f_2)^2 + (d_2 - c_2)(d_1 - f_1)(d_2 - f_2)}{(d_1 - f_1)^2 + (d_2 - f_2)^2} \\ &= c_1 + \frac{(d_1 - f_1)^2(d_1 - c_1) + (d_2 - c_2)(d_1 - f_1)(d_2 - f_2)}{(d_1 - f_1)^2 + (d_2 - f_2)^2} \\ &= c_1 + d_1 - f_1 + \frac{(d_1 - f_1)^2(d_1 - c_1) + (f_2 - c_2)(d_1 - f_1)(d_2 - f_2) - (d_1 - f_1)^3}{(d_1 - f_1)^2 + (d_2 - f_2)^2}. \end{aligned}$$

So from the choice of tangential sites S, we have $n_1 \notin \mathbb{Z}$.

(2) $|f| = |d| = \max\{|c|, |d|, |f|, |g|\}$.

Obviously we have $f = d$. We encounter the equation as before

$$\langle n - c, n - d\rangle = 0,$$
$$\langle n - g, g - d\rangle = 0.$$

Without loss of generality, we take g to be the origin. As above we use n, c, d to denote $n-g, c-g, d-g$.

Then we have

$$(n_1-c_1)(n_1-d_1)+(n_2-c_2)(n_2-d_2)=0,$$
$$n_1d_1+n_2d_2=0.$$

So

$$\left(d_1^2+d_2^2\right)n_1^2+\left(d_1d_2c_2-d_2^2c_1\right)n_1+c_1d_1d_2^2+c_2d_2^3=0$$

has no solution since

$$\Delta=\left(d_1d_2c_2-d_2^2c_1\right)^2-4\left(d_1^2+d_2^2\right)\left(c_1d_1d_2^2+c_2d_2^3\right)<0.$$

Finally, we prove that (A.2) has no solution in $\mathbb{Z}^2$. Note that $\{c,d\}\neq\{f,g\}$. If $\sharp\{c,d\}\cap\{f,g\}=1$, we may assume $c=f$. It suffices to prove that

$$\begin{cases}\langle n-g, g-d\rangle=0,\\ \langle n-c, n-d\rangle=0,\end{cases} \tag{A.6}$$

has no solution in $\mathbb{Z}^2$, while we have proved in **V(2)** that this case has no solution in $\mathbb{Z}^2$.

So we only need to consider the case that c, d, f, g are different from each other. For simplicity, we assume that $|d|>\max\{|c|,|f|,|g|\}$. From the first equation in (A.2), we have $|n|^2\ll d_1$. Moreover, from (A.2), we have

$$\langle n, c+d-f-g\rangle=\langle c,d\rangle-\langle f,g\rangle.$$

By elementary computation,

$$\begin{aligned}n_2&=\frac{c_2d_2-\langle f,g\rangle+c_1d_1-n_1(c_1+d_1-f_1-g_1)}{c_2+d_2-f_2-g_2}\\&=c_2+\frac{-c_2(c_2-f_2-g_2)-\langle f,g\rangle+c_1d_1-n_1(c_1+d_1-f_1-g_1)}{c_2+d_2-f_2-g_2}\\&=c_2+\frac{(c_1-n_1)d_1-c_2(c_2-f_2-g_2)-\langle f,g\rangle-n_1(c_1-f_1-g_1)}{c_2+d_2-f_2-g_2}.\end{aligned}$$

From the choice of the tangential sites S: if $c_1-n_1=0$, we have

$$0<\left|-c_2(c_2-f_2-g_2)-\langle f,g\rangle-n_1(c_1-f_1-g_1)\right|<d_1;$$

if $c_1-n_1\neq0$, we have

$$0<\left|(c_1-n_1)d_1-c_2(c_2-f_2-g_2)-\langle f,g\rangle-n_1(c_1-f_1-g_1)\right|<\frac{1}{2}d_2.$$

So we have proved that $n_2\notin\mathbb{Z}$. $\square$

References

[1] D. Bambusi, On long time stability in Hamiltonian perturbations of non-resonant linear PDEs, Nonlinearity 12 (1999) 823–850.

[2] J. Bourgain, Construction of quasi-periodic solutions for Hamiltonian perturbations of linear equations and applications to nonlinear PDE, Int. Math. Res. Not. IMRN (1994) 475–497.

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

[4] J. Bourgain, Construction of approximative and almost-periodic solutions of perturbed linear Schrödinger and wave equations, Geom. Funct. Anal. 6 (1996) 201–230.

[5] J. Bourgain, Quasiperiodic solutions of Hamiltonian perturbations of 2D linear Schrödinger equations, Ann. of Math. 148 (1998) 363–439.

[6] J. Bourgain, Nonlinear Schrödinger Equations, Park City Ser., vol. 5, American Mathematical Society, Providence, RI, 1999.

[7] J. Bourgain, Green's Function Estimates for Lattice Schrödinger Operators and Applications, Ann. of Math. Stud., vol. 158, Princeton University Press, Princeton, NJ, 2005.

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

[9] W. Craig, C.E. Wayne, Newton's method and periodic solutions of nonlinear wave equations, Comm. Pure Appl. Math. 46 (1993) 1409–1498.

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

[11] L.H. Eliasson, S.B. Kuksin, KAM for the nonlinear Schrödinger equation, Ann. of Math. 172 (2010) 371–435.

[12] J. Geng, Y. Yi, Quasi-periodic solutions in a nonlinear Schrödinger equation, J. Differential Equations 233 (2007) 512–542.

[13] J. Geng, J. You, A KAM theorem for one dimensional Schrödinger equation with periodic boundary conditions, J. Differential Equations 209 (2005) 1–56.

[14] J. Geng, J. You, A KAM theorem for Hamiltonian partial differential equations in higher dimensional spaces, Comm. Math. Phys. 262 (2006) 343–372.

[15] J. Geng, J. You, KAM tori for higher dimensional beam equations with constant potentials, Nonlinearity 19 (2006) 2405–2423.

[16] T. Kappeler, J. Pöschel, KdV & KAM, Springer, 2003.

[17] S.B. Kuksin, Hamiltonian perturbations of infinite-dimensional linear systems with an imaginary spectrum, Funct. Anal. Appl. 21 (1987) 192–205.

[18] S.B. Kuksin, Nearly Integrable Infinite Dimensional Hamiltonian Systems, Lecture Notes in Math., vol. 1556, Springer, Berlin, 1993.

[19] S.B. Kuksin, A KAM-theorem for equations of the Korteweg–de Vries type, Rev. Math. Phys. 10 (1998) 1–64.

[20] S.B. Kuksin, J. Pöschel, Invariant Cantor manifolds of quasiperiodic oscillations for a nonlinear Schrödinger equation, Ann. of Math. 143 (1996) 149–179.

[21] J. Pöschel, On elliptic lower dimensional tori in Hamiltonian systems, Math. Z. 202 (1989) 559–608.

[22] J. Pöschel, Quasi-periodic solutions for a nonlinear wave equation, Comment. Math. Helv. 71 (1996) 269–296.

[23] J. Pöschel, A KAM theorem for some nonlinear partial differential equations, Ann. Sc. Norm. Sup. Pisa Cl. Sci. 23 (1996) 119–148.

[24] C.E. Wayne, Periodic and quasi-periodic solutions for nonlinear wave equations via KAM theory, Comm. Math. Phys. 127 (1990) 479–528.

[25] J. Xu, Q. Qiu, J. You, A KAM theorem of degenerate infinite dimensional Hamiltonian systems (I), Sci. China Ser. A 39 (1996) 372–383.

[26] J. Xu, Q. Qiu, J. You, A KAM theorem of degenerate infinite dimensional Hamiltonian systems (II), Sci. China Ser. A 39 (1996) 384–394.

[27] 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.

[28] X. Yuan, Quasi-periodic solutions of completely resonant nonlinear wave equations, J. Differential Equations 230 (2006) 213–274.