Journal of Functional Analysis 282 (2022) 109430

Contents lists available at ScienceDirect

Journal of Functional Analysis

www.elsevier.com/locate/jfa

Reducible KAM tori for two-dimensional nonlinear Schrödinger equations with explicit dependence on the spatial variable ☆

Jiansheng Geng [a], Shuaishuai Xue [b,*]

[a] Department of Mathematics, Nanjing University, Nanjing 210093, PR China
[b] School of Statistics and Mathematics, Nanjing Audit University, Nanjing 211815, PR China

A R T I C L E   I N F O

Article history:
Received 22 July 2020
Accepted 3 January 2022
Available online 19 February 2022
Communicated by F. Béthuel

Keywords:
Schrödinger equation
KAM tori
Quasi–periodic solutions

A B S T R A C T

We study the two-dimensional nonlinear Schrödinger equation

$$iu_t - \triangle u + |u|^2 u + \frac{\partial f(x, u, \bar{u})}{\partial \bar{u}} = 0, \quad t \in \mathbb{R}, x \in \mathbb{T}^2$$

with periodic boundary conditions. The nonlinearity $f(x, u, \bar{u}) = \sum_{j,l,j+l \geq 6} a_{jl}(x) u^j \bar{u}^l$, $a_{jl} = a_{lj}$ is a real analytic function in a neighborhood of the origin. We obtain, through an infinite dimensional KAM theorem, a Whitney smooth family of small–amplitude quasi–periodic solutions.

☆ Jiansheng Geng acknowledges the support from The National Natural Science Foundation of China (Grant No. 11971012); Shuaishuai Xue acknowledges the support from The National Natural Science Foundation of China (Grant No. 12001275, 12001276).

* Corresponding author.
E-mail addresses: jgeng@nju.edu.cn (J. Geng), ssx@nau.edu.cn (S. Xue).

https://doi.org/10.1016/j.jfa.2022.109430
0022-1236/

1. Introduction and main result

There have been extensive study of the dynamics of linear Hamiltonian partial differential equations. One would like to know the persistency of quasi-periodic solutions of linear or integrable equations under Hamiltonian perturbation. KAM theory is not only a collection of specific theorems, but rather a collection of ideas of how to approach certain problems in perturbation theory connected with small divisors. By now, it is a full fledged theory and it provides a systematic tool for the analysis of many dynamical systems.

There have been many remarkable results which reflected some of the main ideas involved in KAM theory. The main difficulties are the fact that one has to deal with the resonances and small divisors. There are two main approaches to deal with difficulties. On the one hand, the approach, based on a combination of a Nash-Moser implicit function iterative scheme and a Lyapunov-Schmidt bifurcation, was greatly developed by Craig, Wayne, Bourgain [11–17,20,39]. The scheme of Craig-Wayne-Bourgain (CWB for brevity) was used originally as a substitute of the usual KAM-scheme in situations involving multiplicities or near-multiplicities of normal frequencies. The key point of these papers is to resounce to use the cumbersome second Melnikov conditions by solving angle dependent homological equations. The advantage is less Hamiltonian and more flexible than the KAM scheme to deal with resonant cases. This approach is particularly inspiring for PDEs in higher space dimension but to a high cost: the approximate linear equations are variable (quasi-periodic in time) coefficients. Moreover, it only establishes persistence of the Invariant tori but no reducibility and no information on linear stability. On the other hand, the approach, based on a combination of KAM algorithm and Birkhoff normal form, has been designed for a variety of models [4,19,21–38,40,42,43,45]. KAM machinery is built up with infinite many KAM iteration steps. Roughly speaking, each KAM step is a change of variables which transforms the Hamiltonian into a nice normal form plus a smaller perturbation. For this purpose one has to solve some homological equations which forces us to assume that the tangential frequencies and normal frequencies (infinitely many) together satisfy some non-resonant relations. In order to satisfy this condition, one must discard some parameters. Thus KAM machinery includes two parts: analytic part which deals with the iteration and proves convergence under some small divisor conditions, and geometric part which proves that the parameter set after infinitely many times iteration has positive Lebesgue measure. 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 of the corresponding equations. For example, one sees the linear stability and zero Lyapunov exponents. Both approaches are quadratic iteration schemes generalizing Newton's steepest descent method. By now, KAM theory for 1-d PDEs has reached a satisfactory level. Limited work has been done in multidimensional PDEs, more precisely, when considering PDEs especially in space dimension larger than one, a significant problem appears due to the presence of clusters of normal frequencies. A satisfactory future is under construction.

There are plenty of works along this line. The first breakthrough result in this direction is made by Bourgain [13,16], extending the Craig-Wayne approach. Bourgain's technique is a multiscale inductive analysis based on the repeated use of the resolvent identity, which proved quasi-periodic solutions in arbitrary dimensions. Another stream of remarkable result for multidimensional PDEs is made by Eliasson–Kuksin [22]. The authors developed a modified KAM method to construct quasi-periodic solutions for a more interesting higher dimensional Schrödinger equation with a convolution potential on $\mathbb{T}^d$, which proved linear stability. Eliasson–Kuksin introduced the concept of Töplitz-Lipschitz matrices in order to extract asymptotic information on the eigenvalues, and so verify the second Melnikov non-resonance conditions. An essential ingredient in [22] is that finitely many Lipschitz domains cover a neighborhood of ∞. Other results have been proved for the higher dimensional nonlinear beam equations with a constant mass potential and nonlocal Schrödinger equations by Geng–You [26,27]. Geng–You [28] proved that the higher dimensional nonlinear Schrödinger equations with the multiplier M_ξ admit small-amplitude linearly stable quasi-periodic solutions. Chen–Geng [18] proved that the higher dimensional nonlocal wave equations with the multiplier M_ξ admit small-amplitude linearly stable quasi-periodic solutions. We also mention that Pöschel [36] described the construction of almost-periodic solutions for a particular Schrödinger equation on a finite x-interval, depending on some potential V.

All the above results are, however, for nonresonant, typically using the convolution operator $V(x)$ or the multiplier M_ξ as external parameters. One usually studies simplified modes, namely parameter PDEs with the parameters chosen in some way as to avoid resonances. Despite under these simplifying conditions the problems are in general complicated. When the equations do not have external parameters, one must deal with the resonances. In order to overcome this problem Bourgain proposed the idea of choosing an appropriate set of tangential sites S wisely, such that the Birkhoff normal form Hamiltonian admits quasi-periodic solutions for the bifurcation equation where only the Fourier indexes of the tangential sites are excited. This strategy was used by Geng–Xu–You in [29] to prove that two dimensional cubic Schrödinger equation

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

with periodic boundary conditions admits a family of small-amplitude quasi-periodic solutions (see also [37,38]). This strategy was generalized by Wang in [39] to study the NLS on $\mathbb{T}^d$ and prove that the energy supercritical nonlinear Schrödinger equations admit small–amplitude quasi–periodic solutions. The proof used a bifurcation analysis to prove the invertibility of appropriate linearized operators. We also mention the existence results of large families of stable and unstable quasi-periodic solutions of Procesi and Procesi in [37,38] to arbitrary dimensions for the translationally invariant cubic NLS (H has no explicit x-dependence) by a systematic study of the Birkhoff normal form.

In [29], the authors carefully chose tangential sites $\{i_1, \cdots, i_b\} \subset \mathbb{Z}^2$ in order to make the normal form as simple as possible. This strategy was generalized in [37,38], the authors give the concept of generic on the tangential sites. In [39], a essential ingredient is that the genericity condition stem from bounding the sizes of these block diagonal matrix with finite types of blocks. The concept of generic in [37,38] seems to bear a certain resemblance to the conditions in [39]. The proof of these above conditions is rather complex and takes finer combinatorial analysis.

We are concerned in this paper with the NLS,

$$iu_t - \triangle u + |u|^2 u + \frac{\partial f(x, u, \bar{u})}{\partial \bar{u}} = 0, \quad t \in \mathbb{R}, x \in \mathbb{T}^2 \tag{1.1}$$

with periodic boundary conditions, where the nonlinearity $f(x, u, \bar{u}) = \sum_{j,l,j+l \geq 6} a_{jl}(x) u^j \bar{u}^l$, $a_{jl} = a_{lj}$ is a real analytic function in a neighborhood of the origin.

Since there are no external parameters and the only freedom is in choice of the initial data, a subtle problem is that we must deal with the resonances. Note that the results by Geng–Xu–You in [29] on the NLS imply existence of quasi-periodic solutions. When the Hamiltonian nonlinearity does not depend on the space variable x, the equation is translation invariant and Geng–Xu–You [29] (see also [37,38]) were able to exploit the corresponding "Special form" conservation, which is preserved along the KAM iteration, to fulfill the nonresonance conditions. A key issue for the [29] study is that such symmetry enables to prove that many monomials are never present along the KAM iteration. Note that the reducibility result for Eq. (1.1) is not a simple transposition of the method developed in [29,37–39]. There are two main motivations for the study of the Eq. (1.1). The first one is that Eq. (1.1) is explicit dependent on the spatial variable and the eigenvalues appear in clusters of unbounded size. In such a case the reducibility result that one could look for is to block-diagonalize the linearized operators (with blocks of increasing dimensions) (see also [22]). We emphasize that in [22] the convolution potential V plays the role of "external parameters". In the case of Eq. (1.1) which is parameter independent, we must use a Birkhoff normal form analysis. The other is that Wang's [39] technique is a bifurcation analysis, by the Nash-Moser method, and [39] does not prove the reducibility.

We encounter two major problems:

- Eq. (1.1) is explicit dependent on the spatial variable and the eigenvalues appear in clusters of unbounded size. Thanks to the results of "Block Decomposition" of [22], the normal frequency clusters are separated into clusters which are sufficiently distant from one another.
- The goal of the "Töplitz-Lipschitz" property is to extract asymptotic information on the eigenvalues, and to verify the second Melnikov non-resonance conditions. In this paper, we use the elementary repeated limit to substitute Lipschitz domain by

Eliasson–Kuksin [22], thus our measure estimates are easier and the whole proof is more KAM–like.

In all steps we need to combine the concept of "Admissible Set" S of [29] with the structure of "Block Decomposition" of [22]. Very recently, M. Berti et al. [7] used "clusterization properties" of the eigenvalues and suitable separation properties of the singular sites to prove long time dynamics of Schrödinger and wave equations on flat tori, which extend [9,10] abstract Nash–Moser theorem to construct quasi-periodic solutions (see also [5,6] for growth of Sobolev norms for quasi-periodic solutions on T^d and [8] on compact Lie groups and homogeneous spaces; see [1–3] for quasi–linear equations).

Before ending this introduction, we define the concept of "Admissible Set" and the structure of "Block Decomposition".

Definition 1.1. A finite set $S = \{i_1, \cdots, i_b\} \subset \mathbb{Z}^2$ is called *admissible* if

1. Any three of them are not vertices of a rectangle.
2. For any $n \in \mathbb{Z}^2 \setminus S$, there exists at most one triplet $\{i, j, m\}$ with $i, j \in S$, $m \in \mathbb{Z}^2 \setminus S$ 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 first type, denoted $n, m \in \mathcal{L}_1$. By definition, n, m are mutually uniquely determined. We say that (n, m) is a resonant pair of first type. Geometrically, (n, m, i, j) forms a rectangle with n, m being two adjacent vertices.
3. For any $n \in \mathbb{Z}^2 \setminus S$, there exists at most one triplet $\{i, j, m\}$ with $i, j \in S$, $m \in \mathbb{Z}^2 \setminus S$ 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, denoted $n, m \in \mathcal{L}_2$. By definition, n, m are mutually uniquely determined. We say that (n, m) is a resonant pair of second type. Geometrically, (n, m, i, j) forms a rectangle with n, m being two diagonal vertices.
4. Any $n \in \mathbb{Z}^2 \setminus S$ is not resonant of both the first type and the second type, i.e., there exist no $i, j, f, g \in S$ and $m, m' \in \mathbb{Z}^2 \setminus S$, 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 $\mathbb{Z}^2 \setminus S$.

In Appendix A of [29], a concrete way of constructing the admissible set is given. It is plausible that any randomly chosen set S is almost surely admissible.

For given b vectors in $\mathbb{Z}^2$, say $\{i_1, \cdots, i_b\}$, we denote $\mathbb{Z}_1^2 = \mathbb{Z}^2 \setminus \{i_1, \cdots, i_b\}$. In order to have a compact formulation when solving homological equations, we need to decompose $\mathbb{Z}_1^2 \setminus \mathcal{L}_2$.

Decomposition of $\mathbb{Z}_1^2\backslash\mathcal{L}_2$: For a nonnegative integer Δ we define an equivalence relation on $\mathbb{Z}_1^2\backslash\mathcal{L}_2$ generated by the pre-equivalence relation

$$a \sim b \Longleftrightarrow \{|a|^2 = |b|^2, |a-b| \leq \Delta\}.$$

Let $[a]_\Delta$ denote the equivalence class (block) and let $(\mathbb{Z}_1^2\backslash\mathcal{L}_2)_\Delta$ be the set of equivalence classes. It is trivial that each block $[a]_\Delta$ is finite (we will write $[\cdot]$ for $[\cdot]_\triangle$).

- Case 1: $|a| \leq \Delta$, we know $\sharp\{b : |a| = |b|, b \in \mathbb{Z}^2\} \leq e^{\frac{\log\Delta}{\log\log\Delta}} \ll \Delta^\varepsilon$;
- Case 2: $|a| > \Delta$, we have $\sharp\{b : |a| = |b|, |a-b| \leq \Delta^{\frac{1}{3}}, b \in \mathbb{Z}^2\} \leq 2$.

Now we state the main theorem as follows.

Theorem 1. *Let $S = \{i_1, \cdots, i_b\} \subset \mathbb{Z}^2$, $(b > 2)$ be an admissible set. For any $0 < \gamma \ll 1$, there exists a Cantor set $\mathcal{O}_\gamma \subset \mathcal{O}$ with $\mathrm{meas}(\mathcal{O} \setminus \mathcal{O}_\gamma) = O(\gamma^{\frac{1}{4}})$ such that for any $\xi = (\xi_1, \cdots, \xi_b) \in \mathcal{O}_\gamma$, the nonlinear Schrödinger equation (1.1) admits a small–amplitude, quasi–periodic solution of the form*

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

This paper is organized as follows: In section 2 we give an infinite dimensional KAM theorem; in section 3, we give its application to two-dimensional Schrödinger equations. The proof of the KAM theorem is given in section 4, 5, 6. Some technical lemmas are given in the Appendix.

2. An infinite dimensional KAM theorem for Hamiltonian partial differential equations

In this section, we will formulate an infinite dimensional KAM theorem that can be applied to two-dimensional Schrödinger equations under periodic boundary conditions.

We start by introducing some notations. Let $w = (\cdots, w_n, \cdots)_{n\in\mathbb{Z}_1^2}$, and its complex conjugate $\bar{w} = (\cdots, \bar{w}_n, \cdots)_{n\in\mathbb{Z}_1^2}$. We introduce the weighted norm

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

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

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

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 $\alpha \equiv (\cdots, \alpha_n, \cdots)_{n\in\mathbb{Z}_1^2}$, $\beta \equiv (\cdots, \beta_n, \cdots)_{n\in\mathbb{Z}_1^2}$, α_n and $\beta_n \in \mathbb{N}$ with finitely many non-zero components of positive integers. The product $w^\alpha \bar{w}^\beta$ denotes $\prod_n w_n^{\alpha_n}\bar{w}_n^{\beta_n}$. For any given function

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

where $F_{\alpha\beta} = \sum\limits_{k\in\mathbb{Z}^b, l\in\mathbb{N}^b} F_{kl\alpha\beta}(\xi) I^l e^{\mathrm{i}\langle k,\theta\rangle}$. We define its finite weighted norm

$$\|F\|_{D_\rho(r,s),\mathcal{O}} = \sup_{D_\rho(r,s)} \sum_{\alpha,\beta,k,l} |F_{kl\alpha\beta}|_{\mathcal{O}}\ |I^l| e^{|k||\mathrm{Im}\theta|}\, |w^\alpha||\bar{w}^\beta|, \tag{2.2}$$

where $|F_{kl\alpha\beta}|_{\mathcal{O}}$ is short for[1]

$$|F_{kl\alpha\beta}|_{\mathcal{O}} \equiv \sup_{\xi\in\mathcal{O}} \sum_{0\le d\le 4} |\partial_\xi^d F_{kl\alpha\beta}|. \tag{2.3}$$

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

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

Its weighted norm is defined by[2]

$$\begin{aligned} \|X_F\|_{D_\rho(r,s),\mathcal{O}} &\equiv \|F_I\|_{D_\rho(r,s),\mathcal{O}} + \frac{1}{s^2}\|F_\theta\|_{D_\rho(r,s),\mathcal{O}} \\ &\quad + \sup_{D_\rho(r,s)} [\frac{1}{s}\sum_{n\in\mathbb{Z}_1^2} \|F_{w_n}\|_{\mathcal{O}} e^{|n|\rho} + \frac{1}{s}\sum_{n\in\mathbb{Z}_1^2} \|F_{\bar{w}_n}\|_{\mathcal{O}} e^{|n|\rho}]. \end{aligned} \tag{2.4}$$

Suppose that S is an admissible set. We now describe a family of Hamiltonians studied in this paper. Let

$$H_0 = N + \mathcal{B} + \bar{\mathcal{B}},$$

$$N = \langle \omega(\xi), I\rangle + \sum_{n\in\mathbb{Z}_1^2\setminus\mathcal{L}_2} \Omega_n(\xi) w_n \bar{w}_n + \sum_{n'\in\mathcal{L}_2} (\Omega_{n'}(\xi) - \omega_{i'}(\xi)) w_{n'}\bar{w}_{n'}.$$

Recall that (i', j') is uniquely determined by the corresponding resonant pair (n', m') in $\mathcal{L}_2$, and

$$\mathcal{B} = \sum_{n'\in\mathcal{L}_2} a_{n'}(\xi) w_{n'} w_{m'},$$

[1] The derivatives with respect to ξ are in the sense of Whitney. In other words, $F_{\alpha\beta}$ is C_W^4 function. (2.3) is $C_W^4(\mathcal{O})$-norm of $F_{kl\alpha\beta}$.

[2] 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 $\|G\|_{D_\rho(r,s),\mathcal{O}} = \sum_{i=1}^m \|G_i\|_{D_\rho(r,s),\mathcal{O}}$.

$$\bar{\mathcal{B}} = \sum_{n' \in \mathcal{L}_2} \bar{a}_{n'}(\xi) \bar{w}_{n'} \bar{w}_{m'},$$

where $\xi \in \mathcal{O}$ is a parameter, the phase space is endowed with the symplectic structure $dI \wedge d\theta + \mathrm{i} \sum_{n \in \mathbb{Z}_1^2} dw_n \wedge d\bar{w}_n$.

For each $\xi \in \mathcal{O}$, the Hamiltonian equation for H_0 admits special solutions $(\theta, 0, 0, 0) \to (\theta + \omega t, 0, 0, 0)$ that corresponds to an invariant torus on the phase space.

Consider now the perturbed Hamiltonian

$$H = H_0 + P = N + \mathcal{B} + \bar{\mathcal{B}} + P(\theta, I, w, \bar{w}, \xi). \tag{2.5}$$

Our goal is to prove that, for most values of parameter $\xi \in \mathcal{O}$ (in Lebesgue measure sense), the Hamiltonians $H = N + \mathcal{B} + \bar{\mathcal{B}} + P$ still admit invariant tori provided that $\|X_P\|_{D_\rho(r,s),\mathcal{O}}$ is sufficiently small.

In order to have a compact formulation when solving homological equations, we rewrite H into matrix form. Let $z_{[n]} = (w_i)_{i \in [n]}$, $\bar{z}_{[n]} = (\bar{w}_i)_{i \in [n]}$; else $z_n = w_n, \bar{z}_n = \bar{w}_n$.

$$\begin{aligned} H &= \langle \omega(\xi), I \rangle + \sum_{n \in \mathbb{Z}_1^2 \setminus \mathcal{L}_2} \Omega_n(\xi) w_n \bar{w}_n + \sum_{n' \in \mathcal{L}_2} (\Omega_{n'}(\xi) - \omega_{i'}(\xi)) w_{n'} \bar{w}_{n'} + \mathcal{B} + \bar{\mathcal{B}} + P \\ &= \langle \omega(\xi), I \rangle + \sum_{[n]} \langle A_{[n]} z_{[n]}, \bar{z}_{[n]} \rangle + \sum_{n' \in \mathcal{L}_2} (\Omega_{n'}(\xi) - \omega_{i'}(\xi)) z_{n'} \bar{z}_{n'} + \mathcal{B} + \bar{\mathcal{B}} + P \end{aligned}$$

where $A_{[n]}$ is $\sharp[n] \times \sharp[n]$ matrix.

We consider Hamiltonian H satisfying the following hypotheses:

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

$(A2)$ *Asymptotics of normal frequencies:*

$$\Omega_n = \varepsilon^{-a} |n|^2 + \tilde{\Omega}_n, a \geq 0, n \in \mathbb{Z}_1^2 \setminus \mathcal{L}_2, \tag{2.6}$$

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

$(A3)$ *Melnikov's non–resonance conditions:* For $n \in \mathbb{Z}_1^2 \setminus \mathcal{L}_2$, let

$$A_{[n]} = \Omega_{[n]} + (P_{ij}^{011})_{i \in [n], j \in [n]} = (\Omega_{ij} + P_{ij}^{011})_{i \in [n], j \in [n]},$$

where if $i \neq j$ then $\Omega_{ij} = 0$; if $i = j$ then $\Omega_{ij} = \Omega_i$. When $|i - j| > K$, $P_{ij}^{011} = 0$.

$A_{[n]}{}'s$ are C^4_W functions of ξ with C^4_W-norm bounded by some positive constant L, that is to say

$$\sup_{\xi\in\mathcal{O}}\max_{0<d\leq 4}\|\partial^d_\xi A_{[n]}\|\leq L.$$

We assume that $\omega(\xi)$, $A_{[n]}(\xi)\in C^4_W(\mathcal{O})$ and there exist $\gamma,\tau>0$ such that, for $|k|\leq K$,

$$\begin{aligned}
|\langle k,\omega\rangle| &\geq \frac{\gamma}{K^\tau}, k\neq 0,\\
|\langle k,\omega\rangle\pm\widetilde{\lambda}_j| &\geq \frac{\gamma}{K^\tau}, j\in[n],\\
|\langle k,\omega\rangle\pm\widetilde{\lambda}_i\pm\widetilde{\lambda}_j| &\geq \frac{\gamma}{K^\tau}, i\in[m], j\in[n],
\end{aligned}$$

where $\widetilde{\lambda}_i,\widetilde{\lambda}_j$ are $A_{[n]}$ and $A_{[m]}$'s eigenvalues respectively.

Let

$$\mathcal{A}_n=\begin{pmatrix}\Omega_n-\omega_i & -\frac{1}{2\pi^2}\sqrt{\xi_i\xi_j}\\ \frac{1}{2\pi^2}\sqrt{\xi_i\xi_j} & -(\Omega_m-\omega_j)\end{pmatrix}, n\in\mathcal{L}_2,$$

where (n,m) are resonant pairs, (i,j) are uniquely determined by (n,m) in $\mathcal{L}_2$.

We assume that $\omega(\xi)$, $\mathcal{A}_n(\xi)\in C^4_W(\mathcal{O})$ and there exist $\gamma,\tau>0$ such that[3] (here I_2 is 2×2 identity matrix)

$$|\langle k,\omega\rangle\pm\widetilde{\lambda}_i\pm\mu_j|\geq\frac{\gamma}{K^\tau}, k\neq 0, i\in[n], j\in\{1,2\},$$

$$|\langle k,\omega\rangle\pm\mu_j|\geq\frac{\gamma}{K^\tau}, k\neq 0, j\in\{1,2\},$$

where $\widetilde{\lambda}_i$ is $A_{[n]}$ eigenvalue and μ_1,μ_2 are $\mathcal{A}_{n'}$ eigenvalues.

$$|det(\langle k,\omega\rangle I\pm\mathcal{A}_n\otimes I_2\pm I_2\otimes\mathcal{A}_{n'})|\geq\frac{\gamma}{K^\tau}, k\neq 0, n,n'\in\mathcal{L}_2.$$

(*A4*) *Regularity of* $\mathcal{B}+\bar{\mathcal{B}}+P$: $\mathcal{B}+\bar{\mathcal{B}}+P$ is real analytic in $I,\theta,w,\bar{w}$ and C^4_W Whitney smooth in ξ; in addition

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

[3] The tensor product (or direct product) of two $m\times n, k\times l$ matrices $A=(a_{ij}), B$ is a $(mk)\times(nl)$ matrix defined by

$$A\otimes B=(a_{ij}B)=\begin{pmatrix}a_{11}B & \cdots & a_{1n}B\\ \cdots & \cdots & \cdots\\ a_{m1}B & \cdots & a_{mn}B\end{pmatrix}\cdots$$

$\|\cdot\|$ for matrix denotes the operator norm, i.e., $\|M\|=\sup_{|y|=1}|My|$. Recall that ω and $\mathcal{A}_n,\mathcal{A}'_n$ depend on ξ.

$(A5)$ *Töplitz-Lipschitz property:* For any fixed $n, m \in \mathbb{Z}^2$, $c \in \mathbb{Z}^2 \setminus \{0\}$, the limits

$$\lim_{p\to\infty} \frac{\partial^2(\mathcal{B}+P)}{\partial w_{n+pc}\partial w_{m-pc}}, \quad \lim_{p\to\infty} \frac{\partial^2(\sum_{n\in\mathbb{Z}_1^2}\tilde{\Omega}_n w_n\bar{w}_n + P)}{\partial w_{n+pc}\partial \bar{w}_{m+pc}}, \quad \lim_{p\to\infty} \frac{\partial^2(\bar{\mathcal{B}}+P)}{\partial \bar{w}_{n+pc}\partial \bar{w}_{m-pc}}$$

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

$$\|\frac{\partial^2(\mathcal{B}+P)}{\partial w_{n+pc}\partial w_{m-pc}} - \lim_{p\to\infty} \frac{\partial^2(\mathcal{B}+P)}{\partial w_{n+pc}\partial w_{m-pc}}\|_{D_\rho(r,s),\mathcal{O}} \leq \frac{\varepsilon}{|p|} e^{-|n+m|\rho},$$

$$\|\frac{\partial^2(\sum_{n\in\mathbb{Z}_1^2}\tilde{\Omega}_n w_n\bar{w}_n + P)}{\partial w_{n+pc}\partial \bar{w}_{m+pc}} - \lim_{p\to\infty} \frac{\partial^2(\sum_{n\in\mathbb{Z}_1^2}\tilde{\Omega}_n w_n\bar{w}_n + P)}{\partial w_{n+pc}\partial \bar{w}_{m+pc}}\|_{D_\rho(r,s),\mathcal{O}} \leq \frac{\varepsilon}{|p|} e^{-|n-m|\rho},$$

$$\|\frac{\partial^2(\bar{\mathcal{B}}+P)}{\partial \bar{w}_{n+pc}\partial \bar{w}_{m-pc}} - \lim_{p\to\infty} \frac{\partial^2(\bar{\mathcal{B}}+P)}{\partial \bar{w}_{n+pc}\partial \bar{w}_{m-pc}}\|_{D_\rho(r,s),\mathcal{O}} \leq \frac{\varepsilon}{|p|} e^{-|n+m|\rho}.$$

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

Theorem 2. *Assume that the Hamiltonian $H_0 + P$ in (2.5) satisfies $(A1)$–$(A5)$. Let $\gamma > 0$ be small enough, there exists a positive constant $\varepsilon = \varepsilon(b, K, \tau, \gamma, r, s, \rho)$. Such that if $\|X_P\|_{D_\rho(r,s),\mathcal{O}} < \varepsilon$, then the following holds true: There exist a Cantor set $\mathcal{O}_\gamma \subset \mathcal{O}$ 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. Then 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 $H = H_0 + P$.

Remark. As far as we know Theorem 2 seems to bear a certain similarity to the abstract KAM theorems existing in the literature. Some of the key ideas follow closely to the ones of [22,29,38]. However, in order to prove the Theorem 2 which we can apply to study small-amplitude quasi-periodic solutions of the (1.1), we need to remove all the cubic terms that do not commute with the linear part. We have underlined that in (1.1) there is no external parameters to modulate in order to fulfill non-degeneracy. We have to modify that strategy in various non-trivial ways, which we shall watch in the section 4, 5, 6. This is the source of the specific problems and complexity for the NLS (1.1).

3. Application to the two-dimensional Schrödinger equations

We now turn to the mathematical formulation of the equation

$$\mathrm{i}u_t - \Delta u + |u|^2 u + \frac{\partial f(x, u, \bar{u})}{\partial \bar{u}} = 0, \qquad x \in \mathbb{T}^2,\ t \in \mathbb{R} \tag{3.1}$$

with periodic boundary conditions

$$u(t, x_1 + 2\pi, x_2) = u(t, x_1, x_2 + 2\pi) = u(t, x_1, x_2),$$

where $f(x, u, \bar{u}) = \sum\limits_{j,l,j+l\geq 6} a_{jl}(x)u^j\bar{u}^l, a_{jl} = a_{lj}$ is a real analytic function in a neighborhood of the origin.

The operator $A = -\triangle$ with periodic boundary conditions has eigenvalues $\{\lambda_n\}$ satisfying

$$\lambda_n = |n|^2 = |n_1|^2 + |n_2|^2, n = (n_1, n_2) \in \mathbb{Z}^2$$

and the corresponding eigenfunctions $\phi_n(x) = \frac{1}{2\pi}e^{\mathrm{i}\langle n,x\rangle}$ form a basis in the domain of the operator.

Equation (3.1) can be rewritten as a Hamiltonian equation

$$u_t = \mathrm{i}\frac{\partial H}{\partial \bar{u}} \tag{3.2}$$

and the corresponding Hamiltonian is

$$H = \langle Au, u\rangle + \frac{1}{2}\int\limits_{\mathbb{T}^2} |u|^4\ dx + \int\limits_{\mathbb{T}^2} f(x, u, \bar{u})\ dx, \tag{3.3}$$

where $\langle\cdot,\cdot\rangle$ denotes the inner product in L^2.

Let

$$u(x) = \sum_{n\in\mathbb{Z}^2} q_n\phi_n(x).$$

System (3.2) is then equivalent to the lattice Hamiltonian equations

$$\dot{q}_n = \mathrm{i}(\lambda_n q_n + \frac{\partial G}{\partial \bar{q}_n}), \quad G \equiv \frac{1}{8\pi^2}\sum_{i-j+n-m=0} q_i\bar{q}_j q_n\bar{q}_m + \int\limits_{\mathbb{T}^2} f(x, u, \bar{u})\ dx, \tag{3.4}$$

with corresponding Hamiltonian function

$$\begin{aligned} H &= \sum_{n\in\mathbb{Z}^2} \lambda_n q_n\bar{q}_n + \frac{1}{8\pi^2}\sum_{i-j+n-m=0} q_i\bar{q}_j q_n\bar{q}_m + \int\limits_{\mathbb{T}^2} f(x, \sum_{n\in\mathbb{Z}^2} q_n\phi_n(x), \sum_{n\in\mathbb{Z}^2}\bar{q}_n\bar{\phi}_n(x))\ dx \\ &= \sum_{n\in\mathbb{Z}^2}\lambda_n|q_n|^2 + G, \end{aligned} \tag{3.5}$$

then $G = \frac{1}{8\pi^2}\sum_{i-j+n-m=0} q_i\bar{q}_j q_n\bar{q}_m + \int_{\mathbb{T}^2} f(x, \sum_{n\in\mathbb{Z}^2} q_n\phi_n(x), \sum_{n\in\mathbb{Z}^2}\bar{q}_n\bar{\phi}_n(x))\ dx$.

As in [25,33,34], the perturbation G in (3.4) has the following regularity property.

Lemma 3.1. *For any fixed $\rho > 0$, the gradient $G_{\bar{q}}$ is real analytic as a map in a neighborhood of the origin with*

$$\|G_{\bar{q}}\|_\rho \leq c\|q\|_\rho^3. \tag{3.6}$$

Proof. By definition (2.4),

$$\begin{aligned}\|G_{\bar{q}}\|_\rho &= \sum_{n\in\mathbb{Z}^2} |G_{\bar{q}_n}|e^{|n|\rho} \\ &\leq c \sum_{n,\alpha,\beta-e_n,|\alpha|+|\beta-e_n|=3} |q^\alpha \bar{q}^{\beta-e_n}|e^{|n|\rho} \\ &\leq c \sum_{\alpha,\beta-e_n,|\alpha|+|\beta-e_n|=3} |q^\alpha \bar{q}^{\beta-e_n}|e^{|\alpha|\rho}e^{|\beta-e_n|\rho} \\ &\leq c\|q\|_\rho^3.\end{aligned}$$

The proof is completed. □

For an admissible set of tangential site $S = \{i_1, \cdots, i_b\} \subset \mathbb{Z}^2$, using Proposition 1, we first show that we have a nice normal form for H. In order to prove the proposition, we need introduce standard action-angle variables in the tangential space

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

and

$$q_n = w_n, \bar{q}_n = \bar{w}_n, n \in \mathbb{Z}_1^2,$$

after a symplectic transformation.

Proposition 1. *Let S be admissible. For Hamiltonian function (3.5), there is a symplectic transformation Ψ, such that*

$$H \circ \Psi = \langle\omega, I\rangle + \langle\Omega w, w\rangle + \mathcal{A} + \mathcal{B} + \bar{\mathcal{B}} + P \tag{3.7}$$

with

$$\begin{cases} \omega_i(\xi) = \varepsilon^{-3}|i|^2 - \dfrac{1}{4\pi^2}\xi_i + \displaystyle\sum_{j\in S} \dfrac{1}{2\pi^2}\xi_j, \\ \Omega_n = \varepsilon^{-3}|n|^2 + \displaystyle\sum_{j\in S} \dfrac{1}{2\pi^2}\xi_j, \end{cases}$$

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

$$\mathcal{B} = \frac{1}{2\pi^2} \sum_{n'\in\mathcal{L}_2} \sqrt{\xi_{i'}\xi_{j'}}w_{n'}w_{m'}e^{-i\theta_{i'} - i\theta_{j'}},$$

$$\bar{\mathcal{B}} = \frac{1}{2\pi^2}\sum_{n'\in\mathcal{L}_2}\sqrt{\xi_{i'}\xi_{j'}}\bar{w}_{n'}\bar{w}_{m'}e^{i\theta_{i'}+i\theta_{j'}},$$

$$\begin{aligned}|P| = \ & O(\varepsilon^2|I|^2+\varepsilon^2|I|\|w\|_\rho^2+\varepsilon\xi^{\frac{1}{2}}\|w\|_\rho^3+\varepsilon^2\|w\|_\rho^4+\varepsilon\xi^3 \\ & +\varepsilon^2\xi^{\frac{5}{2}}\|w\|_\rho+\varepsilon^3\xi^2\|w\|_\rho^2+\varepsilon^4\xi^{\frac{3}{2}}\|w\|_\rho^3).\end{aligned} \tag{3.8}$$

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

$$F = \sum_{\substack{i-j+n-m=0\\ |i|^2-|j|^2+|n|^2-|m|^2\neq 0\\ \sharp S\cap\{i,j,n,m\}\geq 2}} \frac{i}{8\pi^2(\lambda_i-\lambda_j+\lambda_n-\lambda_m)}q_i\bar{q}_j q_n\bar{q}_m \tag{3.9}$$

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

$$\begin{aligned}H\circ X_F^1 &= H+\{H,F\}+\int_0^1(1-t)\{\{H,F\},F\}\circ\phi_F^t dt \\ &= \sum_{i\in S}\lambda_i|q_i|^2+\sum_{i\in\mathbb{Z}_1^2}\lambda_i|w_i|^2+\sum_{i\in S}\frac{1}{8\pi^2}|q_i|^4 && (3.10)\\ &+\sum_{i,j\in S,i\neq j}\frac{1}{2\pi^2}|q_i|^2|q_j|^2+\sum_{i\in S,j\in\mathbb{Z}_1^2}\frac{1}{2\pi^2}|q_i|^2|w_j|^2 && (3.11)\\ &+\sum_{n\in\mathcal{L}_1}\frac{1}{2\pi^2}q_i\bar{q}_j w_n\bar{w}_m+\sum_{n'\in\mathcal{L}_2}\frac{1}{2\pi^2}(q_{i'}q_{j'}\bar{w}_{n'}\bar{w}_{m'}+\bar{q}_{i'}\bar{q}_{j'}w_{n'}w_{m'}) && (3.12)\\ &+O(|q|\|w\|_\rho^3+\|w\|_\rho^4+|q|^6+|q|^5\|w\|_\rho+|q|^4\|w\|_\rho^2+|q|^3\|w\|_\rho^3).\end{aligned}$$

We remind that (n,m) are resonant pairs and (i,j) is uniquely determined by (n,m); (n',m') are resonant pairs and (i',j') is uniquely determined by (n',m') in (3.12).

The next thing to do in the proof is introduce standard action-angle variables in the tangential space

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

and

$$q_n=w_n,\bar{q}_n=\bar{w}_n,n\in\mathbb{Z}_1^2,$$

then

$$\begin{aligned}H\circ X_F^1 = \ & \sum_{i\in S}\lambda_i(I_i+\xi_i)+\sum_{i\in\mathbb{Z}_1^2}\lambda_i|w_i|^2+\sum_{i\in S}\frac{1}{8\pi^2}(I_i+\xi_i)^2 \\ & +\frac{1}{2\pi^2}\sum_{i,j\in S,i\neq j}(I_i+\xi_i)(I_j+\xi_j)+\frac{1}{2\pi^2}\sum_{i\in S,j\in\mathbb{Z}_1^2}(I_i+\xi_i)|w_j|^2\end{aligned}$$

$$
\begin{aligned}
&+\frac{1}{2\pi^2}\sum_{n\in\mathcal{L}_1}\sqrt{(I_i+\xi_i)(I_j+\xi_j)}w_n\bar{w}_m e^{i\theta_i-i\theta_j}\\
&+\frac{1}{2\pi^2}\sum_{n'\in\mathcal{L}_2}\sqrt{(I_{i'}+\xi_{i'})(I_{j'}+\xi_{j'})}w_{n'}w_{m'}e^{-i\theta_{i'}-i\theta_{j'}}\\
&+\frac{1}{2\pi^2}\sum_{n'\in\mathcal{L}_2}\sqrt{(I_{i'}+\xi_{i'})(I_{j'}+\xi_{j'})}\bar{w}_{n'}\bar{w}_{m'}e^{i\theta_{i'}+i\theta_{j'}}\\
&+O(\xi^{\frac{1}{2}}\|w\|_\rho^3+\|w\|_\rho^4+\xi^3+\xi^{\frac{5}{2}}\|w\|_\rho+\xi^2\|w\|_\rho^2+\xi^{\frac{3}{2}}\|w\|_\rho^3)\\
= \quad &\sum_{i\in S}\lambda_i I_i+\sum_{i\in\mathbb{Z}_1^2}\lambda_i|w_i|^2+\sum_{i\in S}\frac{1}{4\pi^2}\xi_i I_i+\sum_{i,j\in S,i\neq j}\frac{1}{2\pi^2}\xi_i I_j\\
&+\sum_{i\in S,j\in\mathbb{Z}_1^2}\frac{1}{2\pi^2}\xi_i|w_j|^2\\
&+\frac{1}{2\pi^2}\sum_{n\in\mathcal{L}_1}\sqrt{\xi_i\xi_j}w_n\bar{w}_m e^{i\theta_i-i\theta_j}\\
&+\frac{1}{2\pi^2}\sum_{n'\in\mathcal{L}_2}\sqrt{\xi_{i'}\xi_{j'}}w_{n'}w_{m'}e^{-i\theta_{i'}-i\theta_{j'}}\\
&+\frac{1}{2\pi^2}\sum_{n'\in\mathcal{L}_2}\sqrt{\xi_{i'}\xi_{j'}}\bar{w}_{n'}\bar{w}_{m'}e^{i\theta_{i'}+i\theta_{j'}}\\
&+O(|I|^2+|I|\|w\|_\rho^2+\xi^{\frac{1}{2}}\|w\|_\rho^3+\|w\|_\rho^4+\xi^3+\xi^{\frac{5}{2}}\|w\|_\rho+\xi^2\|w\|_\rho^2+\xi^{\frac{3}{2}}\|w\|_\rho^3)\\
= \quad & N+\mathcal{A}+\mathcal{B}+\bar{\mathcal{B}}+P,
\end{aligned}
$$

where

$$N=\sum_{i\in S}\lambda_i I_i+\sum_{j\in\mathbb{Z}_1^2}\lambda_j|w_j|^2-\sum_{i\in S}\frac{1}{4\pi^2}\xi_i I_i+\sum_{i,j\in S}\frac{1}{2\pi^2}\xi_i I_j+\sum_{i\in S,j\in\mathbb{Z}_1^2}\frac{1}{2\pi^2}\xi_i|w_j|^2,$$

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

$$\mathcal{B}=\frac{1}{2\pi^2}\sum_{n'\in\mathcal{L}_2}\sqrt{\xi_{i'}\xi_{j'}}w_{n'}w_{m'}e^{-i\theta_{i'}-i\theta_{j'}},$$

$$\bar{\mathcal{B}}=\frac{1}{2\pi^2}\sum_{n'\in\mathcal{L}_2}\sqrt{\xi_{i'}\xi_{j'}}\bar{w}_{n'}\bar{w}_{m'}e^{i\theta_{i'}+i\theta_{j'}}.$$

Finally, by the scaling in time

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

we finally arrive at the rescaled Hamiltonian

$$H = \varepsilon^{-8} H(\varepsilon^3 \xi, \varepsilon^5 I, \theta, \varepsilon^{\frac{5}{2}} w, \varepsilon^{\frac{5}{2}} \bar{w}) = \langle \omega, I \rangle + \langle \Omega w, w \rangle + \mathcal{A} + \mathcal{B} + \bar{\mathcal{B}} + P,$$

where

$$\begin{cases} \omega_i(\xi) = \varepsilon^{-3}|i|^2 - \dfrac{1}{4\pi^2}\xi_i + \displaystyle\sum_{j\in S} \dfrac{1}{2\pi^2}\xi_j, \\ \Omega_n = \varepsilon^{-3}|n|^2 + \displaystyle\sum_{j\in S} \dfrac{1}{2\pi^2}\xi_j, \end{cases}$$

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

$$\mathcal{B} = \frac{1}{2\pi^2} \sum_{n'\in\mathcal{L}_2} \sqrt{\xi_{i'} \xi_{j'}} w_{n'} w_{m'} e^{-i\theta_{i'} - i\theta_{j'}},$$

$$\bar{\mathcal{B}} = \frac{1}{2\pi^2} \sum_{n'\in\mathcal{L}_2} \sqrt{\xi_{i'} \xi_{j'}} \bar{w}_{n'} \bar{w}_{m'} e^{i\theta_{i'} + i\theta_{j'}}.$$

$$\begin{aligned} |P| = \ & O(\varepsilon^2|I|^2 + \varepsilon^2|I|\|w\|_\rho^2 + \varepsilon\xi^{\frac{1}{2}}\|w\|_\rho^3 + \varepsilon^2\|w\|_\rho^4 + \varepsilon\xi^3 \\ & + \varepsilon^2\xi^{\frac{5}{2}}\|w\|_\rho + \varepsilon^3\xi^2\|w\|_\rho^2 + \varepsilon^4\xi^{\frac{3}{2}}\|w\|_\rho^3). \end{aligned} \tag{3.13}$$

This completes the proof of Proposition 1. □

We will show that, by a nonlinear symplectic coordinates transformation, the normal form in Proposition 1 can be transformed into the more elegant form. For this purpose, we need the following lemma from [41].

Lemma 3.2. *For any $k_1, k_2, \cdots, k_m \in \mathbb{Z}^b$, non-singular $m \times m$ matrix S with $S^T \bar{S} = I$, the map $\Phi_0 : (\theta, I, w, w) \to (\theta_+, I_+, z, z)$ defined by*

$$\begin{cases} \theta_+ = \theta \\ I_+ = I - \sum_{j=1}^m w_j \bar{w}_j k_j \\ z = SEw \\ \bar{z} = \bar{S}\bar{E}\bar{w} \end{cases}$$

is symplectic with diagonal matrix

$$E = E(k_1, k_2, \cdots, k_m) = diag(e^{i\langle k_1, \theta\rangle}, e^{i\langle k_2, \theta\rangle}, \cdots, e^{i\langle k_m, \theta\rangle}).$$

The proof of the above lemma refers to [41]. Although the proof is trivial, the consequences of this result are of importance.

By a nonlinear symplectic coordinates transformation Φ:

$$\begin{cases} \theta_+ = \theta \\ I_+ = I - \sum_{n\in\mathcal{L}_1}(w_n\bar{w}_n e_i + w_m\bar{w}_m e_j) + \sum_{n'\in\mathcal{L}_2}(w_{n'}\bar{w}_{n'}e_{i'} + w_{m'}\bar{w}_{m'}e_{j'}) \\ \begin{pmatrix} z_n \\ z_m \end{pmatrix} = S\begin{pmatrix} e^{i\langle k_i,\theta\rangle} & 0 \\ 0 & e^{i\langle k_j,\theta\rangle}\end{pmatrix}\begin{pmatrix} w_n \\ w_m\end{pmatrix}, \\ \begin{pmatrix} \bar{z}_n \\ \bar{z}_m \end{pmatrix} = \bar{S}\begin{pmatrix} e^{-i\langle k_i,\theta\rangle} & 0 \\ 0 & e^{-i\langle k_j,\theta\rangle}\end{pmatrix}\begin{pmatrix} \bar{w}_n \\ \bar{w}_m\end{pmatrix}, n\in\mathcal{L}_1 \\ z_{n'} = w_{n'}e^{-i\theta_{i'}}, \bar{z}_{n'} = \bar{w}_{n'}e^{i\theta_{i'}}; z_{m'} = w_{m'}e^{-i\theta_{j'}}, \bar{z}_{m'} = \bar{w}_{m'}e^{i\theta_{j'}}, n'\in\mathcal{L}_2 \\ z_n = w_n, \bar{z}_n = \bar{w}_n, n\in\mathbb{Z}_1^2\setminus(\mathcal{L}_1\cup\mathcal{L}_2), \end{cases}$$

we get Hamiltonian systems with the Hamiltonian

$$\begin{aligned} H\circ\Psi\circ\Phi = {} & \langle\omega(\xi), I_+\rangle + \sum_{n\in\mathbb{Z}_1^2\setminus(\mathcal{L}_1\cup\mathcal{L}_2)}\Omega_n(\xi)z_n\bar{z}_n \\ & + \sum_{n\in\mathcal{L}_1}[(\varepsilon^{-3}(|n|^2+|i|^2) + \sum_{j\in S}\frac{1}{\pi^2}\xi_j - \frac{1}{8\pi^2}(\xi_i+\xi_j) \\ & + \frac{1}{8\pi^2}\sqrt{\xi_i^2+14\xi_i\xi_j+\xi_j^2})z_n\bar{z}_n \\ & + (\varepsilon^{-3}(|m|^2+|j|^2) + \sum_{j\in S}\frac{1}{\pi^2}\xi_j - \frac{1}{8\pi^2}(\xi_i+\xi_j) - \frac{1}{8\pi^2}\sqrt{\xi_i^2+14\xi_i\xi_j+\xi_j^2})z_m\bar{z}_m] \\ & + \sum_{n'\in\mathcal{L}_2}[(\Omega_{n'}-\omega_{i'})z_{n'}\bar{z}_{n'} + (\Omega_{m'}-\omega_{j'})z_{m'}\bar{z}_{m'}] \\ & + \frac{1}{2\pi^2}\sum_{n'\in\mathcal{L}_2}\sqrt{\xi_{i'}\xi_{j'}}z_{n'}z_{m'} + \frac{1}{2\pi^2}\sum_{n'\in\mathcal{L}_2}\sqrt{\xi_{i'}\xi_{j'}}\bar{z}_{n'}\bar{z}_{m'} \\ & + P(\theta_+, I_+, z, \bar{z}, \xi) \\ = {} & N + \mathcal{B} + \bar{\mathcal{B}} + P, \end{aligned} \tag{3.14}$$

where

$$\begin{aligned} N = {} & \langle\omega(\xi), I_+\rangle + \sum_{n\in\mathbb{Z}_1^2\setminus\mathcal{L}_2}\Omega_n(\xi)z_n\bar{z}_n \\ & + \sum_{n'\in\mathcal{L}_2}[(\Omega_{n'}-\omega_{i'})z_{n'}\bar{z}_{n'} + (\Omega_{m'}-\omega_{j'})z_{m'}\bar{z}_{m'}] \end{aligned}$$

$$\begin{cases} \omega_i(\xi) = \varepsilon^{-3}|i|^2 - \dfrac{1}{4\pi^2}\xi_i + \displaystyle\sum_{j\in S}\dfrac{1}{2\pi^2}\xi_j, \\ \Omega_n = \varepsilon^{-3}|n|^2 + \displaystyle\sum_{j\in S}\dfrac{1}{2\pi^2}\xi_j, n \in \mathbb{Z}_1^2 \setminus \mathcal{L}_1, \\ \Omega_n = \varepsilon^{-3}(|n|^2+|i|^2) + \displaystyle\sum_{j\in S}\dfrac{1}{\pi^2}\xi_j - \dfrac{1}{8\pi^2}(\xi_i+\xi_j) + \dfrac{1}{8\pi^2}\sqrt{\xi_i^2+14\xi_i\xi_j+\xi_j^2}, n \in \mathcal{L}_1, \\ \Omega_m = \varepsilon^{-3}(|m|^2+|j|^2) + \sum_{j\in S}\frac{1}{\pi^2}\xi_j - \frac{1}{8\pi^2}(\xi_i+\xi_j) - \frac{1}{8\pi^2}\sqrt{\xi_i^2+14\xi_i\xi_j+\xi_j^2}, n \in \mathcal{L}_1, \end{cases}$$

$$\mathcal{B} = \frac{1}{2\pi^2}\sum_{n'\in\mathcal{L}_2}\sqrt{\xi_{i'}\xi_{j'}}z_{n'}z_{m'},$$

$$\bar{\mathcal{B}} = \frac{1}{2\pi^2}\sum_{n'\in\mathcal{L}_2}\sqrt{\xi_{i'}\xi_{j'}}\bar{z}_{n'}\bar{z}_{m'}.$$

For the notational simplicity, I, θ, H refer to $I_+, \theta_+, H \circ \Psi \circ \Phi$. Where P is just G with the $(q_{i_1}, \cdots, q_{i_b}, \bar{q}_{i_1}, \cdots, \bar{q}_{i_b}, q_n, \bar{q}_n)$-variables expressed in terms of the $(\theta, I, z_n, \bar{z}_n)$ variables.

$$\begin{aligned} H = \quad & \langle\omega, I\rangle + \sum_{[n]}\langle A_{[n]}z_{[n]}, \bar{z}_{[n]}\rangle \\ & + \sum_{n'\in\mathcal{L}_2}[(\Omega_{n'}-\omega_{i'})z_{n'}\bar{z}_{n'} + (\Omega_{m'}-\omega_{j'})z_{m'}\bar{z}_{m'}] \\ & + \frac{1}{2\pi^2}\sum_{n'\in\mathcal{L}_2}\sqrt{\xi_{i'}\xi_{j'}}z_{n'}z_{m'} + \frac{1}{2\pi^2}\sum_{n'\in\mathcal{L}_2}\sqrt{\xi_{i'}\xi_{j'}}\bar{z}_{n'}\bar{z}_{m'} \\ & + P(\theta, I, z, \bar{z}, \xi) \\ = \quad & N + \mathcal{B} + \bar{\mathcal{B}} + P, \end{aligned} \tag{3.15}$$

where $A_{[n]}$ is $\sharp[n]\times\sharp[n]$ matrix in (3.15)

$$A_{[n]} = \Omega_{[n]} + (P^{011}_{ij})_{i\in[n],j\in[n]} = (\Omega_{ij} + P^{011}_{ij})_{i\in[n],j\in[n]}$$

where if $i \neq j$ then $\Omega_{ij} = 0$; if $i = j$ then $\Omega_{ij} = \Omega_i$. When $|i-j| > K$, $P^{011}_{ij} = 0$.

Next let us verify that $H = N + \mathcal{B} + \bar{\mathcal{B}} + P$ satisfies the assumptions $(A1)$–$(A5)$.
Verification of $(A1)$:

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

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

Verification of $(A2)$: Take $a = 3$, the proof is obvious.
Verification of $(A3)$: The ideas follow the ones of [29]. In the following, we only give the proof for the most complicated case. Let

$$\mathcal{A}_n = \begin{pmatrix} \Omega_n - \omega_i & -\frac{1}{2\pi^2}\sqrt{\xi_i\xi_j} \\ \frac{1}{2\pi^2}\sqrt{\xi_i\xi_j} & -(\Omega_m - \omega_j) \end{pmatrix}, n \in \mathcal{L}_2,$$

where (m, i, j) is uniquely determined by n. We only verify $(A3)$ for $det(\langle k, \omega\rangle I \pm \mathcal{A}_n \otimes I_2 \pm I_2 \otimes \mathcal{A}_{n'})$ 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 3.3.

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

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

The proof of this lemma may be found in standard matrix theory textbooks.

Case 1. $n, n' \in \mathcal{L}_1$**.** For ease of notations, we define $\alpha = \varepsilon^{-3}(|i_1|^2, |i_2|^2, \cdots, |i_b|^2)$, $\xi = (\xi_{i_1}, \xi_{i_2}, \cdots, \xi_{i_b})$, $\beta = \frac{1}{4\pi^2}(2, 2, \cdots, 2)$, and notice that $|n|^2 + |i|^2 = |m|^2 + |j|^2$, $|n'|^2 + |i'|^2 = |m'|^2 + |j'|^2$. The eigenvalues of $\langle k, \omega\rangle \pm \Omega_n \pm \Omega_{n'}$ are

$$\begin{aligned}&\langle k, \alpha\rangle \pm \varepsilon^{-3}(|n|^2 + |i|^2) \pm \varepsilon^{-3}(|n'|^2 + |i'|^2) + \langle Ak \pm 2\beta \pm 2\beta, \xi\rangle \\ \pm &\frac{1}{8\pi^2}[(-\xi_i - \xi_j \pm \sqrt{{\xi_i}^2 + 14\xi_i\xi_j + {\xi_j}^2}) \pm (-\xi_{i'} - \xi_{j'} \pm \sqrt{{\xi_{i'}}^2 + 14\xi_{i'}\xi_{j'} + {\xi_{j'}}^2})].\end{aligned}$$

- 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
  1. if the eigenvalue is

$$\begin{aligned}&\langle k, \alpha\rangle + \varepsilon^{-3}(|n|^2 + |i|^2) - \varepsilon^{-3}(|n'|^2 + |i|^2) + \langle Ak + 2\beta - 2\beta, \xi\rangle \\ &+ \frac{1}{8\pi^2}[(-\xi_i - \xi_j + \sqrt{{\xi_i}^2 + 14\xi_i\xi_j + {\xi_j}^2}) - (-\xi_i - \xi_j + \sqrt{{\xi_i}^2 + 14\xi_i\xi_j + {\xi_j}^2})] \\ &= \langle k, \alpha\rangle + \varepsilon^{-3}(|n|^2 - |n'|^2) + \langle Ak, \xi\rangle,\end{aligned}$$

  then for $k \neq 0$, we have $Ak \neq 0$;
  2. if the eigenvalue is

$$\begin{aligned}&\langle k, \alpha\rangle + \varepsilon^{-3}(|n|^2 + |i|^2) + \varepsilon^{-3}(|n'|^2 + |i|^2) + \langle Ak + 2\beta + 2\beta, \xi\rangle \\ + &\frac{1}{8\pi^2}[(-\xi_i - \xi_j + \sqrt{{\xi_i}^2 + 14\xi_i\xi_j + {\xi_j}^2}) + (-\xi_i - \xi_j - \sqrt{{\xi_i}^2 + 14\xi_i\xi_j + {\xi_j}^2})]\end{aligned}$$

$$= \langle k, \alpha\rangle + \varepsilon^{-3}(|n|^2 + |i|^2) + \varepsilon^{-3}(|n'|^2 + |i|^2) + \langle Ak + 2\beta + 2\beta, \xi\rangle + \frac{1}{4\pi^2}(-\xi_i - \xi_j)$$
$$= \langle k, \alpha\rangle + \varepsilon^{-3}(|n|^2 + |i|^2) + \varepsilon^{-3}(|n'|^2 + |i|^2) + \langle Ak + 2\beta + 2\beta + \frac{1}{4\pi^2}(-e_i - e_j), \xi\rangle,$$

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 $(2b-1)(k+e_i+e_j)_1 + 8 = 0 (b \geq 2)$, 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 are

$$\langle k, \alpha\rangle \pm \varepsilon^{-3}(|n|^2 + |i|^2) \pm \varepsilon^{-3}(|n'|^2 - |i'|^2) + \langle Ak \pm 2\beta, \xi\rangle$$
$$\pm \frac{1}{8\pi^2}[(-\xi_i - \xi_j \pm \sqrt{{\xi_i}^2 + 14\xi_i\xi_j + {\xi_j}^2}) \pm (\xi_{i'} - \xi_{j'} \pm \sqrt{{\xi_{i'}}^2 - 14\xi_{i'}\xi_{j'} + {\xi_{j'}}^2})].$$

- If the presence of the square root terms, all the eigenvalues are not identically zero.
- If $i = i', \xi_j = \xi_{j'} - 14\xi_{i'}$, the square root terms non-existent, hence if the eigenvalue is

$$\langle k, \alpha\rangle + \varepsilon^{-3}(|n|^2 + |i|^2) + \varepsilon^{-3}(|n'|^2 - |i'|^2) + \langle Ak + 2\beta, \xi\rangle$$
$$+ \frac{1}{8\pi^2}[(-\xi_i - \xi_j + \sqrt{{\xi_i}^2 + 14\xi_i\xi_j + {\xi_j}^2}) + (\xi_{i'} - \xi_{j'} - \sqrt{{\xi_{i'}}^2 - 14\xi_{i'}\xi_{j'} + {\xi_{j'}}^2})]$$
$$= \langle k, \alpha\rangle + \varepsilon^{-3}(|n|^2 + |i|^2) + \varepsilon^{-3}(|n'|^2 - |i'|^2) + \langle Ak + 2\beta + \frac{1}{4\pi^2}(-7e_i - e_j), \xi\rangle,$$

then when $Ak + 2\beta + \frac{1}{4\pi^2}(-7e_i - e_j) = 0$, this equation has no integer solutions. Thus all eigenvalues are not identically zero. In other cases, the proof is similar, so we omit it.

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

$$\langle k, \alpha\rangle \pm \varepsilon^{-3}(|n|^2 - |i|^2) \pm \varepsilon^{-3}(|n'|^2 - |i'|^2) + \langle Ak, \xi\rangle$$
$$\pm \frac{1}{8\pi^2}[(\xi_i - \xi_j \pm \sqrt{{\xi_i}^2 - 14\xi_i\xi_j + {\xi_j}^2}) \pm (\xi_{i'} - \xi_{j'} \pm \sqrt{{\xi_{i'}}^2 - 14\xi_{i'}\xi_{j'} + {\xi_{j'}}^2})].$$

- 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
  1. if the eigenvalue is

$$\langle k, \alpha\rangle + \varepsilon^{-3}(|n|^2 - |i|^2) - \varepsilon^{-3}(|n'|^2 - |i|^2) + \langle Ak, \xi\rangle$$
$$+ \frac{1}{8\pi^2}[(\xi_i - \xi_j + \sqrt{{\xi_i}^2 - 14\xi_i\xi_j + {\xi_j}^2}) - (\xi_i - \xi_j + \sqrt{{\xi_i}^2 - 14\xi_i\xi_j + {\xi_j}^2})]$$
$$= \langle k, \alpha\rangle + \varepsilon^{-3}(|n|^2 - |n'|^2) + \langle Ak, \xi\rangle,$$

then for $k \neq 0$, we have $Ak \neq 0$;

2. if the eigenvalue is

$$\begin{aligned}&\langle k,\alpha\rangle+\varepsilon^{-3}(|n|^2-|i|^2)+\varepsilon^{-3}(|n'|^2-|i|^2)+\langle Ak,\xi\rangle\\&+\frac{1}{8\pi^2}[(\xi_i-\xi_j+\sqrt{\xi_i{}^2-14\xi_i\xi_j+\xi_j{}^2})+(\xi_i-\xi_j-\sqrt{\xi_i{}^2-14\xi_i\xi_j+\xi_j{}^2})]\\&=\langle k,\alpha\rangle+\varepsilon^{-3}(|n|^2-|i|^2)+\varepsilon^{-3}(|n'|^2-|i|^2)+\langle Ak,\xi\rangle+\frac{1}{4\pi^2}(\xi_i-\xi_j)\\&=\langle k,\alpha\rangle+\varepsilon^{-3}(|n|^2-|i|^2)+\varepsilon^{-3}(|n'|^2-|i|^2)+\langle Ak+\frac{1}{4\pi^2}(e_i-e_j),\xi\rangle,\end{aligned}$$

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

$$\begin{aligned}&\langle e_i-e_j,\alpha\rangle+\varepsilon^{-3}(|n|^2-|i|^2)+\varepsilon^{-3}(|n'|^2-|i|^2)\\&=\varepsilon^{-3}(|i|^2-|j|^2+|n|^2-|i|^2+(-|m'|^2+|j|^2))\\&=\varepsilon^{-3}(|n|^2-|m'|^2)\neq 0.\end{aligned}$$

To summarize what we have proved, all the eigenvalues are not identically zero. In other cases, the proof is similar, so we omit it. Due to Lemma 3.3, $det(\langle k,\omega\rangle I\pm\mathcal{A}_n\otimes I_2\pm I_2\otimes\mathcal{A}_{n'})$ is polynomial function in ξ of order at most four. Thus

$$|\partial_\xi^4(det(\langle k,\omega\rangle I\pm\mathcal{A}_n\otimes I_2\pm I_2\otimes\mathcal{A}_{n'}))|\geq\frac{1}{2}|k|\neq 0.$$

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

$$|det(\langle k,\omega\rangle I\pm\mathcal{A}_n\otimes I_2\pm I_2\otimes\mathcal{A}_{n'})|\geq\frac{\gamma}{K^\tau},k\neq 0,n,n'\in\mathcal{L}_2.$$

Thus $(A3)$ is verified.

Verification of $(A4)$: For a given $0<r<1$ and $s=\varepsilon^{\frac{1}{2}}$, according to Lemma 3.1, $\|G_{\bar{q}}\|_\rho\leq c\|q\|_\rho^3$, then

$$\sum_{n\in\mathbb{Z}_1^2}\|P_{w_n}\|_{\mathcal{O}}e^{|n|\rho}+\sum_{n\in\mathbb{Z}_1^2}\|P_{\bar{w}_n}\|_{\mathcal{O}}e^{|n|\rho}=\|P_w\|_\rho+\|P_{\bar{w}}\|_\rho\leq c\|q\|_\rho^3\leq c(|I|^{\frac{3}{2}}+\|w\|_\rho^3).$$

In addition,

$$\sup_{\|q\|_\rho<2s}\|G\|_{\mathcal{O}}\leq c\sup_{\|q\|_\rho<2s}\|q\|_\rho^4\leq cs^4,$$

thus

$$\|P\|_{D_\rho(2r,2s),\mathcal{O}}=\sup_{D_\rho(2r,2s)}\|P\|_{\mathcal{O}}\leq cs^4.$$

According to Cauchy estimates,

$$\|P_I\|_{D_\rho(r,s),\mathcal{O}} \leq cs^2, \|P_\theta\|_{D_\rho(r,s),\mathcal{O}} \leq cs^4,$$

then

$$\begin{aligned}\|X_P\|_{D_\rho(r,s),\mathcal{O}} &= \|P_I\|_{D_\rho(r,s),\mathcal{O}} + \frac{1}{s^2}\|P_\theta\|_{D_\rho(r,s),\mathcal{O}} \\ &+ \sup_{D_\rho(r,s)}[\frac{1}{s}\sum_{n\in\mathbb{Z}_1^2}\|P_{w_n}\|_{\mathcal{O}}e^{|n|\rho} + \frac{1}{s}\sum_{n\in\mathbb{Z}_1^2}\|P_{\bar{w}_n}\|_{\mathcal{O}}e^{|n|\rho}] \\ &\leq cs^2 + \frac{cs^4}{s^2} + c\sup_{D_\rho(r,s)}\frac{1}{s}(|I|^{\frac{3}{2}} + \|z\|_\rho^3) \\ &\leq cs^2 \leq c\varepsilon.\end{aligned}$$

Thus $(A4)$ is verified.

Verification of $(A5)$: We only need to check P satisfies $(A5)$. Recall that (3.9), F is given as

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

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

$$\begin{aligned}&\sum_{i,j,n,m,p}\frac{i}{8\pi^2(\lambda_i - \lambda_j + \lambda_{n+pc} - \lambda_{m+pc})}q_i\bar{q}_j w_{n+pc}\bar{w}_{m+pc} \\ =&\sum_{i,j,n,m,p}\frac{i}{8\pi^2(|i|^2 - |j|^2 + |n|^2 - |m|^2 + 2p\langle n-m, c\rangle)}q_i\bar{q}_j w_{n+pc}\bar{w}_{m+pc}.\end{aligned}$$

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

$$\frac{\partial^2 F}{\partial w_{n+pc}\partial \bar{w}_{m+pc}} = \frac{\partial^2 F}{\partial w_n \partial \bar{w}_m};$$

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

$$\|\frac{\partial^2 F}{\partial w_{n+pc}\partial \bar{w}_{m+pc}} - 0\| \leq \frac{\varepsilon}{|p|}e^{-|n-m|\rho}.$$

Similarly,

$$\begin{aligned}&\|\frac{\partial^2 F}{\partial w_{n+pc}\partial w_{m-pc}} - \lim_{p\to\infty}\frac{\partial^2 F}{\partial w_{n+pc}\partial w_{m-pc}}\|, \|\frac{\partial^2 F}{\partial \bar{w}_{n+pc}\partial \bar{w}_{m-pc}} - \lim_{p\to\infty}\frac{\partial^2 F}{\partial \bar{w}_{n+pc}\partial \bar{w}_{m-pc}}\| \\ \leq&\frac{\varepsilon}{|p|}e^{-|n+m|\rho}.\end{aligned}$$

That is to say, F satisfies Töplitz-Lipschitz property. Recalling the construction of Hamiltonian (3.5), we only need to check that $\{G, F\}$ also satisfies the Töplitz-Lipschitz property. Lemma 4.9 in the next section shows that Poisson bracket preserves Töplitz-Lipschitz property. Thus $N + \mathcal{B} + \bar{\mathcal{B}} + P$ satisfies $(A5)$. Thus $(A5)$ is verified.

So we have verified all the assumptions of Theorem 2 for (3.14). By applying Theorem 2, we get Theorem 1.

4. KAM step

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 of 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 ν–step of the KAM iteration, we consider Hamiltonian function

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

where N_ν is an "integrable normal form", $\mathcal{B}_\nu + \bar{\mathcal{B}}_\nu + P_\nu$ defined in $D_{\rho_\nu}(r_\nu, s_\nu) \times \mathcal{O}_\nu$ with satisfying $(A1)$–$(A5)$.

Our goal is to construct a map

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

and

$$H_{\nu+1} = H_\nu \circ \Phi_\nu = N_{\nu+1} + \mathcal{B}_{\nu+1} + \bar{\mathcal{B}}_{\nu+1} + P_{\nu+1} \tag{4.1}$$

satisfies all the above iterative assumptions $(A1) - (A5)$ on $D_{\rho_{\nu+1}}(r_{\nu+1}, s_{\nu+1}) \times \mathcal{O}_\nu$. Moreover,

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

To simplify notations, in what follows, the quantities without subscripts and superscripts refer to quantities at the ν^{th} step, while the quantities with subscript $+$ or superscript $+$ denote the corresponding quantities at the $(\nu+1)^{\text{th}}$ step. Let us then consider the Hamiltonian

$$\begin{aligned} H = \quad & N + \mathcal{B} + \bar{\mathcal{B}} + P \\ = \quad & \langle \omega, I \rangle + \sum_{[n]} \langle A_{[n]} z_{[n]}, \bar{z}_{[n]} \rangle \end{aligned} \tag{4.2}$$

$$+\sum_{n'\in\mathcal{L}_2}[(\Omega_{n'}-\omega_{i'})z_{n'}\bar{z}_{n'}+(\Omega_{m'}-\omega_{j'})z_{m'}\bar{z}_{m'}]$$
$$+\frac{1}{2\pi^2}\sum_{n'\in\mathcal{L}_2}\sqrt{\xi_{i'}\xi_{j'}}z_{n'}z_{m'}+\frac{1}{2\pi^2}\sum_{n'\in\mathcal{L}_2}\sqrt{\xi_{i'}\xi_{j'}}\bar{z}_{n'}\bar{z}_{m'}$$
$$+P(\theta,I,z,\bar{z},\xi)$$

defined in $D_\rho(r,s)\times\mathcal{O}$. We assume that $|k|\le K$,

$$|\langle k,\omega\rangle|\ge\frac{\gamma}{K^\tau},k\ne 0,$$
$$|\langle k,\omega\rangle+\widetilde{\lambda}_j|\ge\frac{\gamma}{K^\tau},j\in[n],$$
$$|\langle k,\omega\rangle\pm\widetilde{\lambda}_i\pm\widetilde{\lambda}_j|\ge\frac{\gamma}{K^\tau},i\in[m],j\in[n],$$

where $\widetilde{\lambda}_i,\widetilde{\lambda}_j$ are eigenvalues.

$$|\langle k,\omega\rangle\pm\widetilde{\lambda}_i\pm\mu_j|\ge\frac{\gamma}{K^\tau},k\ne 0,i\in[n],j\in\{1,2\},$$

$$|\langle k,\omega\rangle\pm\mu_j|\ge\frac{\gamma}{K^\tau},k\ne 0,j\in\{1,2\},$$

$$|det(\langle k,\omega\rangle I\pm\mathcal{A}_n\otimes I_2\pm I_2\otimes\mathcal{A}_{n'})|\ge\frac{\gamma}{K^\tau},k\ne 0,n,n'\in\mathcal{L}_2,$$

where

$$\mathcal{A}_n=\begin{pmatrix}\Omega_n-\omega_i & -\frac{1}{2\pi^2}\sqrt{\xi_i\xi_j}\\ \frac{1}{2\pi^2}\sqrt{\xi_i\xi_j} & -(\Omega_m-\omega_j)\end{pmatrix},n\in\mathcal{L}_2.$$

Recall that (n,m) are resonant pairs, (i,j) are uniquely determined by (n,m) in $\mathcal{L}_2$. Moreover, $N+\mathcal{B}+\bar{\mathcal{B}}+P$ satisfies $(A4),(A5)$.

Remark. The assumption $(A5)$ makes the measure estimate available at each KAM step.

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

$$s_+=\frac{1}{4}s\varepsilon^{\frac{1}{3}},\quad \varepsilon_+=c\gamma^{-5}K^{5(\tau+1)}(r-r_+)^{-c}\varepsilon^{\frac{4}{3}}. \tag{4.3}$$

Here and later, the letter c denotes suitable (possibly different) constants that do not depend on the iteration steps.

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

4.1. Solving the linearized equations

Expand P into the Fourier-Taylor series

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

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 $\alpha\equiv(\cdots,\alpha_n,\cdots)_{n\in\mathbb{Z}_1^2}$, $\beta\equiv(\cdots,\beta_n,\cdots)_{n\in\mathbb{Z}_1^2}$ with finitely many nonzero components of positive integers.

Let R be the truncation of P given by

$$R(\theta,I,z,\bar{z})=R_0+R_1+R_2,$$

where

$$R_0=\sum_{|k|\le K,|l|\le 1}P_{kl00}e^{\mathrm{i}\langle k,\theta\rangle}\ I^l,$$

$$\begin{aligned}R_1=&\sum_{|k|\le K,n'\in\mathcal{L}_2}(P^{k10}_{n'}z_{n'}+P^{k10}_{m'}z_{m'}+P^{k01}_{n'}\bar{z}_{n'}+P^{k01}_{m'}\bar{z}_{m'})e^{\mathrm{i}\langle k,\theta\rangle}\\&+\sum_{|k|\le K,[n]}(\langle R^{k10}_{[n]},z_{[n]}\rangle+\langle R^{k01}_{[n]},\bar{z}_{[n]}\rangle)e^{\mathrm{i}\langle k,\theta\rangle}\ ,\end{aligned}$$

$$\begin{aligned}R_2=&\sum_{|k|\le K,n\in\mathcal{L}_2,n'\in\mathcal{L}_2}(P^{k11}_{nn'}z_n\bar{z}_{n'}+P^{k11}_{mn'}z_m\bar{z}_{n'}+P^{k11}_{nm'}z_n\bar{z}_{m'}+P^{k11}_{mm'}z_m\bar{z}_{m'}\\&+P^{k11}_{n'n}z_{n'}\bar{z}_n+P^{k11}_{m'n}z_{m'}\bar{z}_n+P^{k11}_{n'm}z_{n'}\bar{z}_m+P^{k11}_{m'm}z_{m'}\bar{z}_m)e^{\mathrm{i}\langle k,\theta\rangle}\\&+\sum_{|k|\le K,n\in\mathcal{L}_2,n'\in\mathcal{L}_2}(P^{k20}_{nn'}z_nz_{n'}+P^{k20}_{mn'}z_mz_{n'}+P^{k20}_{nm'}z_nz_{m'}+P^{k20}_{mm'}z_mz_{m'})e^{\mathrm{i}\langle k,\theta\rangle}\\&+\sum_{|k|\le K,n\in\mathcal{L}_2,n'\in\mathcal{L}_2}(P^{k02}_{nn'}\bar{z}_n\bar{z}_{n'}+P^{k02}_{mn'}\bar{z}_m\bar{z}_{n'}+P^{k02}_{nm'}\bar{z}_n\bar{z}_{m'}+P^{k02}_{mm'}\bar{z}_m\bar{z}_{m'})e^{\mathrm{i}\langle k,\theta\rangle}\\&+\sum_{|k|\le K,[n],[m]}(\langle R^{k20}_{[m][n]}z_{[n]},z_{[m]}\rangle+\langle R^{k02}_{[m][n]}\bar{z}_{[n]},\bar{z}_{[m]}\rangle)e^{\mathrm{i}\langle k,\theta\rangle}\\&+\sum_{|k|\le K,[n],[m]}\langle R^{k11}_{[m][n]}z_{[n]},\bar{z}_{[m]}\rangle e^{\mathrm{i}\langle k,\theta\rangle}\\&+\sum_{|k|\le K,n\in\mathbb{Z}_1^2\setminus\mathcal{L}_2,n'\in\mathcal{L}_2}(P^{k20}_{nn'}z_nz_{n'}+P^{k20}_{nm'}z_nz_{m'})e^{\mathrm{i}\langle k,\theta\rangle}\\&+\sum_{|k|\le K,n\in\mathbb{Z}_1^2\setminus\mathcal{L}_2,n'\in\mathcal{L}_2}(P^{k02}_{nn'}\bar{z}_n\bar{z}_{n'}+P^{k02}_{nm'}\bar{z}_n\bar{z}_{m'})e^{\mathrm{i}\langle k,\theta\rangle}\\&+\sum_{|k|\le K,n\in\mathbb{Z}_1^2\setminus\mathcal{L}_2,n'\in\mathcal{L}_2}(P^{k11}_{nn'}z_n\bar{z}_{n'}+P^{k11}_{nm'}z_n\bar{z}_{m'}+P^{k11}_{n'n}z_{n'}\bar{z}_n+P^{k11}_{m'n}z_{m'}\bar{z}_n)e^{\mathrm{i}\langle k,\theta\rangle}\ ,\end{aligned}$$

where $P_n^{k10} = P_{kl\alpha\beta}$ with $\alpha = e_n, \beta = 0$, here e_n denotes the vector with the n^{th} component being 1 and the other components being zero; $P_n^{k01} = P_{kl\alpha\beta}$ with $\alpha = 0, \beta = e_n$; $P_{nm}^{k20} = P_{kl\alpha\beta}$ with $\alpha = e_n + e_m, \beta = 0$; $P_{nm}^{k11} = P_{kl\alpha\beta}$ with $\alpha = e_n, \beta = e_m$; $P_{nm}^{k02} = P_{kl\alpha\beta}$ with $\alpha = 0, \beta = e_n + e_m$.

$R_{[n]}^{k10}$, $R_{[n]}^{k01}$, $R_{[m][n]}^{k20}$, $R_{[m][n]}^{k02}$ and $R_{[m][n]}^{k11}$ are, respectively, $\sharp[n] \times 1, \sharp[n] \times 1, \sharp[m] \times \sharp[n], \sharp[m] \times \sharp[n], \sharp[m] \times \sharp[n]$ matrices

$$R_{[n]}^{k10} = (P_i^{k10})_{i\in[n]}, R_{[n]}^{k01} = (P_i^{k01})_{i\in[n]}, |[n]| \le K,$$

$$R_{[m][n]}^{k20} = (R_{ij}^{k20})_{i\in[m], j\in[n]},$$

where if $|i+j| \le K$, $R_{ij}^{k20} = P_{ij}^{k20}$; if $|i+j| > K$, $R_{ij}^{k20} = 0$;

$$R_{[m][n]}^{k02} = (R_{ij}^{k02})_{i\in[m], j\in[n]},$$

where if $|i+j| \le K$, $R_{ij}^{k02} = P_{ij}^{k02}$; if $|i+j| > K$, $R_{ij}^{k02} = 0$;

$$R_{[m][n]}^{k11} = (R_{ij}^{k11})_{i\in[m], j\in[n]},$$

where if $|i-j| \le K$, $R_{ij}^{k11} = P_{ij}^{k11}$; if $|i-j| > K$, $R_{ij}^{k11} = 0$.

Rewrite H as $H = N + \mathcal{B} + \bar{\mathcal{B}} + R + (P - R)$. By the choice of s_+ in (4.3) and the definition of the norms, it follows immediately that

$$\|X_R\|_{D_\rho(r,s),\mathcal{O}} \le \|X_P\|_{D_\rho(r,s),\mathcal{O}} \le \varepsilon \tag{4.4}$$

for any $\frac{r_0}{2} < \rho \le r$. In the next, we prove that for $\frac{r_0}{2} < \rho \le r_+$

$$\|H_{(P-R)}\|_{D_\rho(r_+,s),\mathcal{O}} < c\varepsilon_+.$$

In fact, $P - R = P^* + h.o.t.$, where

$$\begin{aligned} P^* = &\sum_{|n|>K} [P_n^{k10}(\theta) w_n + P_n^{k01}(\theta)\bar{w}_n] \\ &+ \sum_{|n+m|>K} [P_{nm}^{k20}(\theta) w_n w_m + P_{nm}^{k02}(\theta)\bar{w}_n\bar{w}_m] + \sum_{|n-m|>K} P_{nm}^{k11}(\theta) w_n \bar{w}_m \end{aligned}$$

be the linear and quadratic terms in the perturbation. By virtue of (4.3), the decay property of P, $\|X_P\|_{D_\rho(r,s),\mathcal{O}} \le \varepsilon$, and Cauchy estimates, one has that for $\rho \le r_+$

$$\begin{aligned} &\|X_{P_*}\|_{D_\rho(r_+,s),\mathcal{O}} \\ &\le (r - r_+)^{-1}\Big(\sum_{|n|>K} \varepsilon e^{-|n|r} e^{|n|\rho} + \sum_{|n+m|>K} \varepsilon e^{-|n+m|r} |\bar{w}_m| e^{|n+m|\rho} \\ &\quad + \sum_{|n-m|>K} \varepsilon e^{-|n-m|r} |w_m| e^{|n-m|\rho}\Big) \end{aligned}$$

$$
\begin{aligned}
&\leq (r-r_+)^{-1}\Big(\sum_{|n|>K} \varepsilon e^{-|n|r} e^{|n|\rho} + \sum_{|n|>K,m} \varepsilon e^{-|n|r} |w_m| e^{|n|\rho} e^{m\rho}\Big) \\
&\leq (r-r_+)^{-1} \sum_{|n|>K} \varepsilon e^{-|n|(r-\rho)} \\
&\leq (r-r_+)^{-1} \varepsilon e^{-K(r-\rho)} \\
&\leq \varepsilon_+ .
\end{aligned}
$$

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

$$
\|X_{(P-R)}\|_{D_\rho(r,s_+)} \;<\; \mathrm{c}\, \varepsilon_+ . \tag{4.5}
$$

In the following, we will look for an F, 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

$$
\begin{aligned}
H \circ \phi_F^1 &= (N + \mathcal{B} + \bar{\mathcal{B}} + R) \circ \phi_F^1 + (P - R) \circ \phi_F^1 \\
&= N + \mathcal{B} + \bar{\mathcal{B}} + \{N + \mathcal{B} + \bar{\mathcal{B}}, F\} + R \\
&+ \int_0^1 (1-t)\{\{N + \mathcal{B} + \bar{\mathcal{B}}, F\}, F\} \circ \phi_F^t dt \\
&+ \int_0^1 \{R, F\} \circ \phi_F^t dt + (P - R) \circ \phi_F^1 \\
&= N_+ + \mathcal{B}_+ + \bar{\mathcal{B}}_+ + P_+ + \{N + \mathcal{B} + \bar{\mathcal{B}}, F\} + R \\
&- P_{0000} - \langle \hat{\omega}, I\rangle - \sum_{[n]} \langle P^{011}_{[n][n]} z_{[n]}, \bar{z}_{[n]}\rangle - \sum_{n' \in \mathcal{L}_2} (P^{011}_{n'n'} z_{n'} \bar{z}_{n'} + P^{011}_{m'm'} z_{m'} \bar{z}_{m'}) \\
&- \hat{\mathcal{B}} - \hat{\bar{\mathcal{B}}},
\end{aligned} \tag{4.6}
$$

where

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

$$
\hat{\mathcal{B}} = \sum_{n' \in \mathcal{L}_2} P^{020}_{n'm'} z_{n'} z_{m'},
$$

$$
\hat{\bar{\mathcal{B}}} = \sum_{n' \in \mathcal{L}_2} P^{002}_{n'm'} \bar{z}_{n'} \bar{z}_{m'},
$$

$$
N_+ = N + P_{0000} + \langle \hat{\omega}, I\rangle + \sum_{[n]} \langle P^{011}_{[n][n]} z_{[n]}, \bar{z}_{[n]}\rangle + \sum_{n' \in \mathcal{L}_2} (P^{011}_{n'n'} z_{n'} \bar{z}_{n'} + P^{011}_{m'm'} z_{m'} \bar{z}_{m'}), \tag{4.7}
$$

$$\mathcal{B}_+ = \mathcal{B} + \hat{\mathcal{B}}, \tag{4.8}$$

$$\bar{\mathcal{B}}_+ = \bar{\mathcal{B}} + \hat{\bar{\mathcal{B}}} = \bar{\mathcal{B}} + \bar{\hat{\mathcal{B}}}, \tag{4.9}$$

$$P_+ = \int_0^1 (1-t)\{\{N + \mathcal{B} + \bar{\mathcal{B}}, F\}, F\} \circ \phi_F^t dt + \int_0^1 \{R, F\} \circ \phi_F^t dt + (P - R) \circ \phi_F^1. \tag{4.10}$$

We shall find a function F:

$$F(\theta, I, z, \bar{z}) = F_0 + F_1 + F_2,$$

where

$$F_0 = \sum_{0<|k|\le K, |l|\le 1} F_{kl00} e^{\mathrm{i}\langle k,\theta\rangle}\ I^l,$$

$$\begin{aligned}
F_1 = &\sum_{|k|\le K, n'\in\mathcal{L}_2} (F_{n'}^{k10} z_{n'} + F_{m'}^{k10} z_{m'} + F_{n'}^{k01} \bar{z}_{n'} + F_{m'}^{k01} \bar{z}_{m'}) e^{\mathrm{i}\langle k,\theta\rangle} \\
&+ \sum_{|k|\le K, [n]} (\langle F_{[n]}^{k10}, z_{[n]}\rangle + \langle F_{[n]}^{k01}, \bar{z}_{[n]}\rangle) e^{\mathrm{i}\langle k,\theta\rangle}\ ,
\end{aligned}$$

$$\begin{aligned}
F_2 = &\sum_{|k|\le K, n\in\mathcal{L}_2, n'\in\mathcal{L}_2, |k|+|n-n'|\ne 0} (F_{n'n}^{k11} z_{n'} \bar{z}_n + F_{nn'}^{k11} z_n \bar{z}_{n'}) e^{\mathrm{i}\langle k,\theta\rangle} \\
&+ \sum_{|k|\le K, n\in\mathcal{L}_2, n'\in\mathcal{L}_2, |k|+|n-m'|\ne 0} (F_{m'n}^{k11} z_{m'} \bar{z}_n + F_{nm'}^{k11} z_n \bar{z}_{m'}) e^{\mathrm{i}\langle k,\theta\rangle} \\
&+ \sum_{|k|\le K, n\in\mathcal{L}_2, n'\in\mathcal{L}_2, |k|+|m-n'|\ne 0} (F_{n'm}^{k11} z_{n'} \bar{z}_m + F_{mn'}^{k11} z_m \bar{z}_{n'}) e^{\mathrm{i}\langle k,\theta\rangle} \\
&+ \sum_{|k|\le K, n\in\mathcal{L}_2, n'\in\mathcal{L}_2, |k|+|m-m'|\ne 0} (F_{m'm}^{k11} z_{m'} \bar{z}_m + F_{mm'}^{k11} z_m \bar{z}_{m'}) e^{\mathrm{i}\langle k,\theta\rangle} \\
&+ \sum_{|k|\le K, n\in\mathcal{L}_2, n'\in\mathcal{L}_2, |k|+|n-m'|\ne 0 (or) |k|+|n'-m|\ne 0} (F_{n'n}^{k20} z_{n'} z_n + F_{n'n}^{k02} \bar{z}_{n'} \bar{z}_n) e^{\mathrm{i}\langle k,\theta\rangle} \\
&+ \sum_{|k|\le K, n\in\mathcal{L}_2, n'\in\mathcal{L}_2, |k|+|m-m'|\ne 0 (or) |k|+|n'-n|\ne 0} (F_{m'n}^{k20} z_{m'} z_n + F_{m'n}^{k02} \bar{z}_{m'} \bar{z}_n) e^{\mathrm{i}\langle k,\theta\rangle} \\
&+ \sum_{|k|\le K, n\in\mathcal{L}_2, n'\in\mathcal{L}_2, |k|+|m-m'|\ne 0 (or) |k|+|n'-n|\ne 0} (F_{n'm}^{k20} z_{n'} z_m + F_{n'm}^{k02} \bar{z}_{n'} \bar{z}_m) e^{\mathrm{i}\langle k,\theta\rangle} \\
&+ \sum_{|k|\le K, n\in\mathcal{L}_2, n'\in\mathcal{L}_2, |k|+|n-m'|\ne 0 (or) |k|+|n'-m|\ne 0} (F_{m'm}^{k20} z_{m'} z_m + F_{m'm}^{k02} \bar{z}_{m'} \bar{z}_m) e^{\mathrm{i}\langle k,\theta\rangle} \\
&+ \sum_{|k|\le K, [n], [m]} (\langle F_{[m][n]}^{k20} z_{[n]}, z_{[m]}\rangle + \langle F_{[m][n]}^{k02} \bar{z}_{[n]}, \bar{z}_{[m]}\rangle) e^{\mathrm{i}\langle k,\theta\rangle} \\
&+ \sum_{|k|\le K, [n], [m], |k|+||n|-|m||\ne 0} \langle F_{[m][n]}^{k11} z_{[n]}, \bar{z}_{[m]}\rangle e^{\mathrm{i}\langle k,\theta\rangle}
\end{aligned}$$

$$+\sum_{|k|\le K, n\in\mathbb{Z}_1^2\setminus\mathcal{L}_2, n'\in\mathcal{L}_2}(F^{k20}_{nn'}z_nz_{n'}+F^{k20}_{nm'}z_nz_{m'})e^{\mathrm{i}\langle k,\theta\rangle}$$
$$+\sum_{|k|\le K, n\in\mathbb{Z}_1^2\setminus\mathcal{L}_2, n'\in\mathcal{L}_2}(F^{k02}_{nn'}\bar z_n\bar z_{n'}+F^{k02}_{nm'}\bar z_n\bar z_{m'})e^{\mathrm{i}\langle k,\theta\rangle}$$
$$+\sum_{|k|\le K, n\in\mathbb{Z}_1^2\setminus\mathcal{L}_2, n'\in\mathcal{L}_2}(F^{k11}_{nn'}z_n\bar z_{n'}+F^{k11}_{nm'}z_n\bar z_{m'}+F^{k11}_{n'n}z_{n'}\bar z_n+F^{k11}_{m'n}z_{m'}\bar z_n)e^{\mathrm{i}\langle k,\theta\rangle}\ .$$

$F^{k10}_{[n]}$, $F^{k01}_{[n]}$, $F^{k20}_{[m][n]}$, $F^{k02}_{[m][n]}$ and $F^{k11}_{[m][n]}$ are, respectively, $\sharp[n]\times 1, \sharp[n]\times 1, \sharp[m]\times\sharp[n], \sharp[m]\times\sharp[n], \sharp[m]\times\sharp[n]$ matrices

$$F^{k10}_{[n]}=(F^{k10}_i)_{i\in[n]}, F^{k01}_{[n]}=(F^{k01}_i)_{i\in[n]}, |[n]|\le K,$$

$$F^{k20}_{[m][n]}=(f^{k20}_{ij})_{i\in[m],j\in[n]},$$

where if $|i+j|\le K$, $f^{k20}_{ij}=F^{k20}_{ij}$; if $|i+j|>K$, $f^{k20}_{ij}=0$;

$$F^{k02}_{[m][n]}=(f^{k02}_{ij})_{i\in[m],j\in[n]},$$

where if $|i+j|\le K$, $f^{k02}_{ij}=F^{k02}_{ij}$; if $|i+j|>K$, $f^{k02}_{ij}=0$;

$$F^{k11}_{[m][n]}=(f^{k11}_{ij})_{i\in[m],j\in[n]},$$

where if $|i-j|\le K$, $f^{k11}_{ij}=F^{k11}_{ij}$; if $|i-j|>K$, $f^{k11}_{ij}=0$, satisfying the equation

$$\begin{aligned}&\{N+\mathcal{B}+\bar{\mathcal{B}},F\}+R-P_{0000}-\langle\hat\omega,I\rangle-\sum_{[n]}\langle P^{011}_{[n][n]}z_{[n]},\bar z_{[n]}\rangle-\hat{\mathcal{B}}-\hat{\bar{\mathcal{B}}}\\&-\sum_{n'\in\mathcal{L}_2}(P^{011}_{n'n'}z_{n'}\bar z_{n'}+P^{011}_{m'm'}z_{m'}\bar z_{m'})=0.\end{aligned}\tag{4.11}$$

To find the function F, we need several lemmas.

Lemma 4.1. *F satisfies (4.11) if the Fourier coefficients of F_0, F_1 are defined by the following equations*

$$\begin{aligned}(\langle k,\omega\rangle)F_{kl00}&=\mathrm{i}P_{kl00},\quad |l|\le 1, 0<|k|\le K,\\(\langle k,\omega\rangle I-A_{[n]})F^{k10}_{[n]}&=\mathrm{i}P^{k10}_{[n]},\quad |k|\le K, n\in\mathbb{Z}_1^2\setminus\mathcal{L}_2,\\(\langle k,\omega\rangle I+A_{[n]})F^{k01}_{[n]}&=\mathrm{i}R^{k01}_{[n]},\quad |k|\le K, n\in\mathbb{Z}_1^2\setminus\mathcal{L}_2,\\(\langle k,\omega\rangle I-\mathcal{A}_{n'})(F^{k10}_{n'},F^{k01}_{m'})^T&=i(P^{k10}_{n'},P^{k01}_{m'})^T, |k|\le K, n'\in\mathcal{L}_2,\\(\langle k,\omega\rangle I+\mathcal{A}_{n'})(F^{k01}_{n'},F^{k10}_{m'})^T&=i(P^{k01}_{n'},P^{k10}_{m'})^T, |k|\le K, n'\in\mathcal{L}_2.\end{aligned}\tag{4.12}$$

The Fourier coefficients of F_2 are defined by the following Lemmas:

Case 1: $n,m\in\mathbb{Z}_1^2\setminus\mathcal{L}_2$

Lemma 4.2. *F satisfies (4.11) if the Fourier coefficients of F_2 are defined by the following equations*

$$\begin{aligned}
(\langle k,\omega\rangle I - A_{[m]})F^{k20}_{[m][n]} - F^{k20}_{[m][n]}A_{[n]} &= \mathrm{i}R^{k20}_{[m][n]},\\
(\langle k,\omega\rangle I - A_{[m]})F^{k11}_{[m][n]} + F^{k11}_{[m][n]}A_{[n]} &= \mathrm{i}R^{k11}_{[m][n]}, \quad |k| + ||n|-|m|| \neq 0,\\
(\langle k,\omega\rangle I + A_{[m]})F^{k02}_{[m][n]} + F^{k02}_{[m][n]}A_{[n]} &= \mathrm{i}R^{k02}_{[m][n]}.
\end{aligned} \tag{4.13}$$

Case 2: $n \in \mathbb{Z}^2_1 \setminus \mathcal{L}_2, n' \in \mathcal{L}_2$

Lemma 4.3. *F satisfies (4.11) if the Fourier coefficients of F_2 are defined by the following equations*

$$\begin{aligned}
\{I_2 \otimes [\langle k,\omega\rangle I - A_{[n]}] - \mathcal{A}_{n'} \otimes I\}(F^{k20}_{[n]n'}, F^{k11}_{[n]m'})^T &= i(P^{k20}_{[n]n'}, P^{k11}_{[n]m'})^T,\\
\{I_2 \otimes [\langle k,\omega\rangle I + A_{[n]}] + \mathcal{A}_{n'} \otimes I\}(F^{k02}_{[n]n'}, F^{k11}_{m'[n]})^T &= i(P^{k02}_{[n]n'}, P^{k11}_{m'[n]})^T,\\
\{I_2 \otimes [\langle k,\omega\rangle I - A_{[n]}] + \mathcal{A}_{n'} \otimes I\}(F^{k11}_{[n]n'}, F^{k20}_{[n]m'})^T &= i(P^{k11}_{[n]n'}, P^{k20}_{[n]m'})^T,\\
\{I_2 \otimes [\langle k,\omega\rangle I + A_{[n]}] - \mathcal{A}_{n'} \otimes I\}(F^{k11}_{n'[n]}, F^{k02}_{m'[n]})^T &= i(P^{k11}_{n'[n]}, P^{k02}_{m'[n]})^T,
\end{aligned} \tag{4.14}$$

where I is $\sharp[n] \times \sharp[n]$ identity matrix.

Case 3: $n, n' \in \mathcal{L}_2$

Lemma 4.4. *F satisfies (4.11) if the Fourier coefficients of F_2 are defined by the following equations*

$$\begin{aligned}
(\langle k,\omega\rangle I - \mathcal{A}_n \otimes I + I \otimes \mathcal{A}_{n'})(F^{k11}_{nn'}, F^{k20}_{nm'}, F^{k02}_{mn'}, F^{k11}_{m'm})^T &= i(P^{k11}_{nn'}, P^{k20}_{nm'}, P^{k02}_{mn'}, P^{k11}_{m'm})^T,\\
(\langle k,\omega\rangle I + \mathcal{A}_n \otimes I - I \otimes \mathcal{A}_{n'})(F^{k11}_{n'n}, F^{k02}_{m'n}, F^{k20}_{n'm}, F^{k11}_{mm'})^T &= i(P^{k11}_{n'n}, P^{k02}_{m'n}, P^{k20}_{n'm}, P^{k11}_{mm'})^T,\\
(\langle k,\omega\rangle I - \mathcal{A}_n \otimes I - I \otimes \mathcal{A}_{n'})(F^{k20}_{nn'}, F^{k11}_{nm'}, F^{k11}_{n'm}, F^{k02}_{mm'})^T &= i(P^{k20}_{nn'}, P^{k11}_{nm'}, P^{k11}_{n'm}, P^{k02}_{mm'})^T,\\
(\langle k,\omega\rangle I + \mathcal{A}_n \otimes I + I \otimes \mathcal{A}_{n'})(F^{k02}_{nn'}, F^{k11}_{m'n}, F^{k11}_{mn'}, F^{k20}_{m'm})^T &= i(P^{k02}_{nn'}, P^{k11}_{m'n}, P^{k11}_{mn'}, P^{k20}_{m'm})^T.
\end{aligned}$$

In the following, we only give the proof for the most complicated case.

Proof. Inserting F into (4.11). By comparing the Fourier coefficients, more precisely, if (n', m') is a resonant pair in $\mathcal{L}_2$, it can easily be shown that

$$\begin{aligned}
&\sum_{|k|\le K, n'\in\mathcal{L}_2} [\langle k,\omega\rangle - (\Omega_{n'} - \omega_{i'})]F^{k10}_{n'} z_{n'} e^{\mathrm{i}\langle k,\theta\rangle} + \frac{1}{2\pi^2}\sqrt{\xi_{i'}\xi_{j'}} F^{k01}_{m'} z_{n'} e^{\mathrm{i}\langle k,\theta\rangle}\\
&= i \sum_{|k|\le K, n'\in\mathcal{L}_2} P^{k10}_{n'} z_{n'} e^{\mathrm{i}\langle k,\theta\rangle},\\
&\sum_{|k|\le K, n'\in\mathcal{L}_2} [\langle k,\omega\rangle + (\Omega_{m'} - \omega_{j'})]F^{k01}_{m'} z_{m'} e^{\mathrm{i}\langle k,\theta\rangle} - \frac{1}{2\pi^2}\sqrt{\xi_{i'}\xi_{j'}} F^{k10}_{n'} z_{m'} e^{\mathrm{i}\langle k,\theta\rangle}
\end{aligned}$$

$$= i \sum_{|k|\le K, n'\in\mathcal{L}_2} P^{k01}_{m'} z_{m'} e^{\mathrm{i}\langle k,\theta\rangle} .$$

It is easy to rewrite in matrix form

$$(\langle k,\omega\rangle I - \mathcal{A}_{n'})(F^{k10}_{n'}, F^{k01}_{m'})^T = i(P^{k10}_{n'}, P^{k01}_{m'})^T, |k|\le K, n'\in\mathcal{L}_2,$$

similarly, form

$$(\langle k,\omega\rangle I + \mathcal{A}_{n'})(F^{k01}_{n'}, F^{k10}_{m'})^T = i(P^{k01}_{n'}, P^{k10}_{m'})^T, |k|\le K, n'\in\mathcal{L}_2.$$

The proofs are almost identical, the major change being the substitution of the Fourier coefficients for equations. If (n,m) and (n',m') are resonant pairs in $\mathcal{L}_2$, it can be checked that $(F^{k11}_{nn'}, F^{k20}_{nm'}, F^{k02}_{mn'}, F^{k11}_{m'm})^T$ satisfy

$$\begin{aligned}
&[\langle k,\omega\rangle - (\Omega_n-\omega_i) + (\Omega_{n'}-\omega_{i'})]F^{k11}_{nn'}e^{\mathrm{i}\langle k,\theta\rangle} - \frac{1}{2\pi^2}\sqrt{\xi_{i'}\xi_{j'}}F^{k20}_{nm'}e^{\mathrm{i}\langle k,\theta\rangle}\\
&\quad + \frac{1}{2\pi^2}\sqrt{\xi_i\xi_j}F^{k02}_{mn'}e^{\mathrm{i}\langle k,\theta\rangle}\\
&= iP^{k11}_{nn'}e^{\mathrm{i}\langle k,\theta\rangle} ,
\end{aligned}$$

similarly,

$$\begin{aligned}
&[\langle k,\omega\rangle - (\Omega_n-\omega_i) - (\Omega_{m'}-\omega_{j'})]F^{k20}_{nm'}e^{\mathrm{i}\langle k,\theta\rangle} + \frac{1}{2\pi^2}\sqrt{\xi_{i'}\xi_{j'}}F^{k11}_{nn'}e^{\mathrm{i}\langle k,\theta\rangle}\\
&\quad + \frac{1}{2\pi^2}\sqrt{\xi_i\xi_j}F^{k11}_{m'm}e^{\mathrm{i}\langle k,\theta\rangle}\\
&= iP^{k20}_{nm'}e^{\mathrm{i}\langle k,\theta\rangle} ,\\
&[\langle k,\omega\rangle + (\Omega_m-\omega_j) + (\Omega_{n'}-\omega_{i'})]F^{k02}_{mn'}e^{\mathrm{i}\langle k,\theta\rangle} - \frac{1}{2\pi^2}\sqrt{\xi_{i'}\xi_{j'}}F^{k11}_{m'm}e^{\mathrm{i}\langle k,\theta\rangle}\\
&\quad - \frac{1}{2\pi^2}\sqrt{\xi_i\xi_j}F^{k11}_{nn'}e^{\mathrm{i}\langle k,\theta\rangle}\\
&= iP^{k02}_{mn'}e^{\mathrm{i}\langle k,\theta\rangle} ,\\
&[\langle k,\omega\rangle + (\Omega_m-\omega_j) - (\Omega_{m'}-\omega_{j'})]F^{k11}_{m'm}e^{\mathrm{i}\langle k,\theta\rangle} + \frac{1}{2\pi^2}\sqrt{\xi_{i'}\xi_{j'}}F^{k02}_{mn'}e^{\mathrm{i}\langle k,\theta\rangle}\\
&\quad - \frac{1}{2\pi^2}\sqrt{\xi_i\xi_j}F^{k20}_{nm'}e^{\mathrm{i}\langle k,\theta\rangle}\\
&= iP^{k11}_{m'm}e^{\mathrm{i}\langle k,\theta\rangle} .
\end{aligned}$$

We rewrite them into matrix form

$$\begin{aligned}
&(\langle k,\omega\rangle I - \mathcal{A}_n\otimes I + I\otimes\mathcal{A}_{n'})(F^{k11}_{nn'}, F^{k20}_{nm'}, F^{k02}_{mn'}, F^{k11}_{m'm})^T = i(P^{k11}_{nn'}, P^{k20}_{nm'}, P^{k02}_{mn'}, P^{k11}_{m'm})^T,\\
&\quad |k|\le K, n, n'\in\mathcal{L}_2,
\end{aligned}$$

similarly, form

$$
\begin{aligned}
&(\langle k,\omega\rangle I+\mathcal{A}_n\otimes I-I\otimes\mathcal{A}_{n'})(F^{k11}_{n'n},F^{k02}_{m'n},F^{k20}_{n'm},F^{k11}_{mm'})^T=i(P^{k11}_{n'n},P^{k02}_{m'n},P^{k20}_{n'm},P^{k11}_{mm'})^T,\\
&\quad |k|\le K, n,n'\in\mathcal{L}_2,\\
&(\langle k,\omega\rangle I-\mathcal{A}_n\otimes I-I\otimes\mathcal{A}_{n'})(F^{k20}_{nn'},F^{k11}_{nm'},F^{k11}_{n'm},F^{k02}_{mm'})^T=i(P^{k20}_{nn'},P^{k11}_{nm'},P^{k11}_{n'm},P^{k02}_{mm'})^T,\\
&\quad |k|\le K, n,n'\in\mathcal{L}_2,\\
&(\langle k,\omega\rangle I+\mathcal{A}_n\otimes I+I\otimes\mathcal{A}_{n'})(F^{k02}_{nn'},F^{k11}_{m'n},F^{k11}_{mn'},F^{k20}_{m'm})^T=i(P^{k02}_{nn'},P^{k11}_{m'n},P^{k11}_{mn'},P^{k20}_{m'm})^T,\\
&\quad |k|\le K, n,n'\in\mathcal{L}_2.
\end{aligned}
$$

In other cases, the proof is similar, so we omit it. Thus these Lemmas are obtained. □

We are now in a position to consider the equations

$$Q^T_{[n]}(\langle k,\omega\rangle I-A_{[n]})F^{k10}_{[n]}=iQ^T_{[n]}R^{k10}_{[n]}, |k|\le K.$$

Matrix $Q_{[n]}$ is the $A_{[n]}$'s orthogonal matrix. It is easy to see that

$$(\langle k,\omega\rangle I-Q^T_{[n]}A_{[n]}Q_{[n]})Q^T_{[n]}F^{k10}_{[n]}=iQ^T_{[n]}R^{k10}_{[n]}, |k|\le K,$$

that is

$$(\langle k,\omega\rangle I-\Lambda_{[n]})\hat{F}^{k10}_{[n]}=i\hat{R}^{k10}_{[n]}, |k|\le K.$$

Similarly, form

$$
\begin{aligned}
(\langle k,\omega\rangle I+\Lambda_{[n]})\hat{F}^{k01}_{[n]}&=i\hat{R}^{k01}_{[n]}, |k|\le K,\\
(\langle k,\omega\rangle I-\Lambda_{[m]})\hat{F}^{k20}_{[m][n]}-\hat{F}^{k20}_{[m][n]}\Lambda_{[n]}&=i\hat{R}^{k20}_{[m][n]}, |k|\le K,\\
(\langle k,\omega\rangle I-\Lambda_{[m]})\hat{F}^{k11}_{[m][n]}+\hat{F}^{k11}_{[m][n]}\Lambda_{[n]}&=i\hat{R}^{k11}_{[m][n]}, |k|\le K, |k|+||n|-|m||\ne 0,\\
(\langle k,\omega\rangle I+\Lambda_{[m]})\hat{F}^{k02}_{[m][n]}+\hat{F}^{k02}_{[m][n]}\Lambda_{[n]}&=i\hat{R}^{k02}_{[m][n]}, |k|\le K.
\end{aligned}
$$

$A_{[n]}$ can be diagonalized by orthogonal matrix $Q_{[n]}$, that is $\Lambda_{[n]}=Q^T_{[n]}A_{[n]}Q_{[n]}$.

$$
\begin{aligned}
\hat{R}^{kx}_{[n]}&=Q^T_{[n]}R^{kx}_{[n]}, x=10,01,\\
\hat{R}^{kx}_{[m][n]}&=Q^T_{[m]}R^{kx}_{[m][n]}Q_{[n]}, x=20,11,02.\\
\hat{F}^{kx}_{[n]}&=Q^T_{[n]}F^{kx}_{[n]}, x=10,01,\\
\hat{F}^{kx}_{[m][n]}&=Q^T_{[m]}F^{kx}_{[m][n]}Q_{[n]}, x=20,11,02.
\end{aligned}
$$

Now our problem reduces to focus on the following equations:

$$
\begin{aligned}
(\langle k,\omega\rangle-\widetilde{\lambda}_j)\hat{F}^{k10}_{[n],j}&=i\hat{R}^{k10}_{[n],j},|k|\le K,j\in[n],\\
(\langle k,\omega\rangle+\widetilde{\lambda}_j)\hat{F}^{k01}_{[n],j}&=i\hat{R}^{k01}_{[n],j},|k|\le K,j\in[n],\\
(\langle k,\omega\rangle-\widetilde{\lambda}_i-\widetilde{\lambda}_j)\hat{F}^{k20}_{[m][n],ij}&=i\hat{R}^{k20}_{[m][n],ij},|k|\le K,i\in[m],j\in[n],\\
(\langle k,\omega\rangle-\widetilde{\lambda}_i+\widetilde{\lambda}_j)\hat{F}^{k11}_{[m][n],ij}&=i\hat{R}^{k11}_{[m][n],ij},|k|\le K,|k|+||n|-|m||\neq 0,i\in[m],j\in[n],\\
(\langle k,\omega\rangle+\widetilde{\lambda}_i+\widetilde{\lambda}_j)\hat{F}^{k02}_{[m][n],ij}&=i\hat{R}^{k02}_{[m][n],ij},|k|\le K,i\in[m],j\in[n].
\end{aligned}
$$

The second part of equations follows in a similar manner, then

$$
\begin{aligned}
&(I_2\otimes Q_{[n]})^T\{I_2\otimes[\langle k,\omega\rangle I-A_{[n]}]-\mathcal{A}_{n'}\otimes I\}(I_2\otimes Q_{[n]})\\
&(I_2\otimes Q_{[n]})^T(F^{k20}_{[n]n'},F^{k11}_{[n]m'})^T=i(I_2\otimes Q_{[n]})^T(P^{k20}_{[n]n'},P^{k11}_{[n]m'})^T,|k|\le K.
\end{aligned}
$$

Matrix $Q_{[n]}$ is the $A_{[n]}$'s orthogonal matrix. It is easy to show that

$$
\begin{aligned}
&\{I_2\otimes[\langle k,\omega\rangle I-Q^T_{[n]}A_{[n]}Q_{[n]}]-\mathcal{A}_{n'}\otimes I\}\\
&(I_2\otimes Q_{[n]})^T(F^{k20}_{[n]n'},F^{k11}_{[n]m'})^T=i(I_2\otimes Q_{[n]})^T(P^{k20}_{[n]n'},P^{k11}_{[n]m'})^T,|k|\le K,
\end{aligned}
$$

that is

$$\{I_2\otimes[\langle k,\omega\rangle I-\Lambda_{[n]}]-\mathcal{A}_{n'}\otimes I\}(\widehat{F}^{k20}_{[n]n'},\widehat{F}^{k11}_{[n]m'})^T=i(\widehat{P}^{k20}_{[n]n'},\widehat{P}^{k11}_{[n]m'})^T,|k|\le K.$$

$A_{[n]}$ can be diagonalized by orthogonal matrix $Q_{[n]}$, that is $\Lambda_{[n]}=Q^T_{[n]}A_{[n]}Q_{[n]}$. Thus

$$
\begin{aligned}
(\widehat{F}^{k20}_{[n]n'},\widehat{F}^{k11}_{[n]m'})^T&=(I_2\otimes Q_{[n]})^T(F^{k20}_{[n]n'},F^{k11}_{[n]m'})^T,\\
(\widehat{P}^{k20}_{[n]n'},\widehat{P}^{k11}_{[n]m'})^T&=(I_2\otimes Q_{[n]})^T(P^{k20}_{[n]n'},P^{k11}_{[n]m'})^T.
\end{aligned}
$$

We are considering the equations

$$
\begin{aligned}
&(Q_{n'}\otimes I)^{-1}\{I_2\otimes[\langle k,\omega\rangle I-\Lambda_{[n]}]-\mathcal{A}_{n'}\otimes I\}(Q_{n'}\otimes I)\\
&(Q_{n'}\otimes I)^{-1}(\widehat{F}^{k20}_{[n]n'},\widehat{F}^{k11}_{[n]m'})^T=i(Q_{n'}\otimes I)^{-1}(\widehat{P}^{k20}_{[n]n'},\widehat{P}^{k11}_{[n]m'})^T,|k|\le K,
\end{aligned}
$$

there exists a invertible matrix $Q_{n'}$ such that $Q^{-1}_{n'}\mathcal{A}_{n'}Q_{n'}=J_{n'}$ is a Jordan normal form, that is

$$\{I_2\otimes[\langle k,\omega\rangle I-\Lambda_{[n]}]-J_{n'}\otimes I\}(\widetilde{F}^{k20}_{[n]n'},\widetilde{F}^{k11}_{[n]m'})^T=i(\widetilde{P}^{k20}_{[n]n'},\widetilde{P}^{k11}_{[n]m'})^T,|k|\le K.$$

We define

$$
\begin{aligned}
(\widetilde{F}^{k20}_{[n]n'},\widetilde{F}^{k11}_{[n]m'})^T&=(Q_{n'}\otimes I)^{-1}(\widehat{F}^{k20}_{[n]n'},\widehat{F}^{k11}_{[n]m'})^T,\\
(\widetilde{P}^{k20}_{[n]n'},\widetilde{P}^{k11}_{[n]m'})^T&=(Q_{n'}\otimes I)^{-1}(\widehat{P}^{k20}_{[n]n'},\widehat{P}^{k11}_{[n]m'})^T.
\end{aligned}
$$

Similarly, form

$$\{I_2 \otimes [\langle k,\omega\rangle I + \Lambda_{[n]}] + J_{n'} \otimes I\}(\widetilde{F}^{k02}_{[n]n'}, \widetilde{F}^{k11}_{m'[n]})^T = i(\widetilde{P}^{k02}_{[n]n'}, \widetilde{P}^{k11}_{m'[n]})^T,$$
$$\{I_2 \otimes [\langle k,\omega\rangle I - \Lambda_{[n]}] + J_{n'} \otimes I\}(\widetilde{F}^{k11}_{[n]n'}, \widetilde{F}^{k20}_{[n]m'})^T = i(\widetilde{P}^{k11}_{[n]n'}, \widetilde{P}^{k20}_{[n]m'})^T,$$
$$\{I_2 \otimes [\langle k,\omega\rangle I + \Lambda_{[n]}] - J_{n'} \otimes I\}(\widetilde{F}^{k11}_{n'[n]}, \widetilde{F}^{k02}_{m'[n]})^T = i(\widetilde{P}^{k11}_{n'[n]}, \widetilde{P}^{k02}_{m'[n]})^T,$$

where

$$(\widetilde{F}^{k02}_{[n]n'}, \widetilde{F}^{k11}_{m'[n]})^T = (Q^{-1}_{n'} \otimes Q^T_{[n]})(F^{k02}_{[n]n'}, F^{k11}_{m'[n]})^T,$$
$$(\widetilde{P}^{k02}_{[n]n'}, \widetilde{P}^{k11}_{m'[n]})^T = (Q^{-1}_{n'} \otimes Q^T_{[n]})(P^{k02}_{[n]n'}, P^{k11}_{m'[n]})^T,$$
$$(\widetilde{F}^{k11}_{[n]n'}, \widetilde{F}^{k20}_{[n]m'})^T = (Q^{-1}_{n'} \otimes Q^T_{[n]})(F^{k11}_{[n]n'}, F^{k20}_{[n]m'})^T,$$
$$(\widetilde{P}^{k11}_{[n]n'}, \widetilde{P}^{k20}_{[n]m'})^T = (Q^{-1}_{n'} \otimes Q^T_{[n]})(P^{k11}_{[n]n'}, P^{k20}_{[n]m'})^T,$$
$$(\widetilde{F}^{k11}_{n'[n]}, \widetilde{F}^{k02}_{m'[n]})^T = (Q^{-1}_{n'} \otimes Q^T_{[n]})(F^{k11}_{n'[n]}, F^{k02}_{m'[n]})^T,$$
$$(\widetilde{P}^{k11}_{n'[n]}, \widetilde{P}^{k02}_{m'[n]})^T = (Q^{-1}_{n'} \otimes Q^T_{[n]})(P^{k11}_{n'[n]}, P^{k02}_{m'[n]})^T.$$

- When Jordan normal form is

$$J_{n'} = \begin{pmatrix} \mu_1 & 0 \\ 0 & \mu_2 \end{pmatrix}, n' \in \mathcal{L}_2.$$

Now we focus on the following equations

$$(\langle k,\omega\rangle - \widetilde{\lambda}_j - \mu_1)(\widetilde{F}^{k20}_{[n]n'})_j = i(\widetilde{P}^{k20}_{[n]n'})_j,$$

$$(\langle k,\omega\rangle - \widetilde{\lambda}_j - \mu_2)(\widetilde{F}^{k11}_{[n]m'})_j = i(\widetilde{P}^{k11}_{[n]m'})_j.$$

- When Jordan normal form

$$J_{n'} = \begin{pmatrix} \mu & 1 \\ 0 & \mu \end{pmatrix}, n' \in \mathcal{L}_2.$$

We focus on the following equations

$$\begin{pmatrix} \langle k,\omega\rangle - \widetilde{\lambda}_j - \mu & -1 \\ 0 & \langle k,\omega\rangle - \widetilde{\lambda}_j - \mu \end{pmatrix} \begin{pmatrix} (\widetilde{F}^{k20}_{[n]n'})_j \\ (\widetilde{F}^{k11}_{[n]m'})_j \end{pmatrix} = i \begin{pmatrix} (\widetilde{P}^{k20}_{[n]n'})_j \\ (\widetilde{P}^{k11}_{[n]m'})_j \end{pmatrix}.$$

In the other cases, the proof is similar, so we omit it. In order to solve the last three equations, we need the following elementary algebraic result from matrix theory.

Lemma 4.5. *Let A, B, C be, respectively, $n \times n, m \times m, n \times m$ matrices, and let X be an $n \times m$ unknown matrix. The matrix equation*

$$AX - XB = C,$$

is solvable if and only if $I_m \otimes A - B \otimes I_n$ is nonsingular.

For a detailed proof, we refer the reader to the Appendix in [44].

Remark. Taking the transpose of the fourth equation in Lemma 4.2, one sees that $(F^{k20}_{[m][n]})^T$ satisfies the same equation as $(F^{k20}_{[n][m]})$. Then (by the uniqueness of the solution) it follows that $(F^{k02}_{[n][m]}) = (F^{k02}_{[m][n]})^T$, $(F^{-k11}_{[n][m]}) = \overline{(F^{k11}_{[m][n]})^T}$.

4.2. Estimate for coefficients of F

Let us consider $F^{k20}_{[m][n]}$ and $(F^{k20}_{[n]n'}, F^{k11}_{[n]m'})^T$ for instance, and the other terms can be treated in an analogous way. By the construction above, one sees that

$$F^{k20}_{[m][n],ij} = i \sum_{m_1,n_1} \frac{Q_{[m],im_1} \hat{R}^{k20}_{[m][n],m_1,n_1} Q^T_{[n],n_1 j}}{\langle k, \omega\rangle - \widetilde{\lambda}_i - \widetilde{\lambda}_j},$$

$$\begin{aligned}\begin{pmatrix} F^{k20}_{[n]n'} \\ F^{k11}_{[n]m'} \end{pmatrix} &= i \sum_{0<j\leq 2\sharp[n]} (Q_{n'} \otimes Q_{[n]})_j \begin{pmatrix} \widetilde{P}^{k20}_{[n]n'} \\ \widetilde{P}^{k11}_{[n]m'} \end{pmatrix} \frac{1}{\langle k, \omega\rangle - \widetilde{\lambda}_j - \mu} \\ &+ i \sum_{0<j\leq \sharp[n]} (Q_{n'} \otimes Q_{[n]})_{j+\sharp} \begin{pmatrix} \widetilde{P}^{k20}_{[n]n'} \\ \widetilde{P}^{k11}_{[n]m'} \end{pmatrix} \frac{1}{\langle k, \omega\rangle - \widetilde{\lambda}_j - \mu}.\end{aligned}$$

Then

$$|F^{k20}_{[m][n],ij}| \leq c\varepsilon \frac{K^\tau}{\gamma} e^{K^{1+3\varepsilon}\rho} e^{-\rho|i+j|} e^{-|k|r},$$

$$|\begin{pmatrix} F^{k20}_{[n]n'} \\ F^{k11}_{[n]m'} \end{pmatrix}_j| \leq c\varepsilon \frac{K^\tau}{\gamma} e^{K^{1+3\varepsilon}\rho} e^{-\rho|j-m'|} e^{-|k|r}, \sharp[n] < j \leq 2\sharp[n],$$

$$|\begin{pmatrix} F^{k20}_{[n]n'} \\ F^{k11}_{[n]m'} \end{pmatrix}_j| \leq c\varepsilon \frac{K^{2\tau}}{\gamma^2} e^{K^{1+3\varepsilon}\rho} e^{-\rho|j+n'|} e^{-|k|r}, 0 < j \leq \sharp[n].$$

We use the factor $e^{K^{1+3\varepsilon}\rho}$ to recover the exponential decay under the assumption

$$K^{1+3\varepsilon}\rho = 1.$$

Observe that

$$\|F^{k20}_{[m][n]}\| \leq cK_\nu^{3\varepsilon}\varepsilon_{\nu+1}\frac{K_\nu^{5(\tau+1)}}{\gamma^5}K_\nu^{3\varepsilon} \leq \varepsilon_{\nu+1}^{\frac{1}{3}},$$

$$\|\begin{pmatrix} F^{k20}_{[n]n'} \\ F^{k11}_{[n]m'} \end{pmatrix}\| \leq cK_\nu^{3\varepsilon}\varepsilon_{\nu+1}\frac{K_\nu^{5(\tau+1)}}{\gamma^5} \leq \varepsilon_{\nu+1}^{\frac{1}{3}},$$

under the assumption

$$\varepsilon_{\nu+1} = c\gamma^{-5}(r_\nu - r_{\nu+1})^{-c}K_\nu^{5(\tau+1)}\varepsilon_\nu^{\frac{4}{3}}.$$

4.3. Estimate for the coordinate transformation

We proceed to estimate X_F and ϕ_F^1. We start with the following

Lemma 4.6. *Let $D_i = D(r_+ + \frac{i}{4}(r - r_+), \frac{i}{4}s)$, $0 < i \leq 4$, then*

$$\|X_F\|_{D_3,\mathcal{O}} \leq c\gamma^{-5}K^{5(\tau+1)}(r - r_+)^{-c}\varepsilon. \tag{4.15}$$

The proof of this lemma is not particularly difficult but will not be reproduced here. In the next lemma, we give some estimates for ϕ_F^t. The formula (4.16) will be used to prove our coordinate transformation is well defined. Inequality (4.17) will be used to check the convergence of the iteration.

Lemma 4.7. *Let $\eta = \varepsilon^{\frac{1}{3}}, D_{i\eta} = D(r_+ + \frac{i}{4}(r - r_+), \frac{i}{4}\eta s), 0 < i \leq 4$. If $\varepsilon \ll \frac{1}{2}\gamma^{\frac{15}{2}}K^{-\frac{15}{2}(\tau+1)}(r - r_+)^c$, we then have*

$$\phi_F^t : D_{2\eta} \to D_{3\eta}, \quad -1 \leq t \leq 1. \tag{4.16}$$

Moreover,

$$\|D\phi_F^t - Id\|_{D_{1\eta}} < c\gamma^{-5}K^{5(\tau+1)}(r - r_+)^{-c}\varepsilon. \tag{4.17}$$

Proof. Let

$$\|D^m F\|_{D,\mathcal{O}} = \max\{\|\frac{\partial^{|i|+|l|+|\alpha|+|\beta|}}{\partial\theta^i\partial I^l\partial z^\alpha\partial\bar{z}^\beta}F\|_{D,\mathcal{O}}, |i| + |l| + |\alpha| + |\beta| = m \geq 2\}.$$

Notice that F is a polynomial of degree 1 in I and degree 2 in z, $\bar{z}$. From (2.4), (4.15) and the Cauchy inequality, it follows that

$$\|D^m F\|_{D_2,\mathcal{O}} < c\gamma^{-5}K^{5(\tau+1)}(r - r_+)^{-c}\varepsilon, \tag{4.18}$$

for any $m \geq 2$.

To get the estimates for ϕ_F^t, we start from the integral equation,

$$\phi_F^t = id + \int_0^t X_F \circ \phi_F^s \, ds$$

so that $\phi_F^t : D_{2\eta} \to D_{3\eta}, \quad -1 \le t \le 1$, which follows directly from (4.18). Since

$$D\phi_F^t = Id + \int_0^t (DX_F) D\phi_F^s \, ds = Id + \int_0^t J(D^2 F) D\phi_F^s \, ds,$$

where J denotes the standard symplectic matrix $\begin{pmatrix} 0 & -I \\ I & 0 \end{pmatrix}$, it follows that

$$\| D\phi_F^t - Id \| \le 2 \| D^2 F \| < c\gamma^{-5} K^{5(\tau+1)} (r - r_+)^{-c} \varepsilon. \tag{4.19}$$

Consequently Lemma 4.7 follows. □

4.4. Estimate for the new normal form

The map ϕ_F^1 defined above transforms H into $H_+ = N_+ + \mathcal{B}_+ + \bar{\mathcal{B}}_+ + P_+$ (see (4.6) and (4.11)) with the normal form N_+

$$\begin{aligned} N_+ &= N + P_{0000} + \langle \hat{\omega}, I \rangle + \sum_{[n]} \langle P^{011}_{[n][n]} z_{[n]}, \bar{z}_{[n]} \rangle + \sum_{n' \in \mathcal{L}_2} (P^{011}_{n'n'} z_{n'} \bar{z}_{n'} + P^{011}_{m'm'} z_{m'} \bar{z}_{m'}) \\ &= \langle \omega_+, I \rangle + \sum_{[n]} \langle A^+_{[n]} z_{[n]}, \bar{z}_{[n]} \rangle + \sum_{n' \in \mathcal{L}_2} [(\Omega^+_{n'} - \omega_{i'}) z_{n'} \bar{z}_{n'} + (\Omega^+_{m'} - \omega_{j'}) z_{m'} \bar{z}_{m'})], \end{aligned}$$

where

$$\omega_+ = \omega + P_{0l00}(|l| = 1), \tag{4.20}$$

$$A^+_{[n]} = A_{[n]} + R^{011}_{[n][n]} = A_{[n]} + (R^{011}_{ij})_{i \in [n], j \in [n]}, |i-j| > K, R^{011}_{ij} = 0; |i-j| \le K, R^{011}_{ij} = P^{011}_{ij}$$

$$\Omega^+_{n'} = \Omega_{n'} + P^{011}_{n'n'}, \Omega^+_{m'} = \Omega_{m'} + P^{011}_{m'm'}, n' \in \mathcal{L}_2.$$

Now we prove that N_+ shares the same properties as N. By the regularity of X_P and by Cauchy estimates, then we have

$$|\omega_+ - \omega| < \varepsilon, \quad |P^{011}_{ij+} - P^{011}_{ij}| < \varepsilon e^{-|i-j|\rho}. \tag{4.21}$$

It follows that for $|k| \le K$,

$$|\langle k, \omega + P_{0l00} \rangle| \ge |\langle k, \omega \rangle| - \varepsilon K \ge \frac{\gamma}{K^\tau} - \varepsilon K \ge \frac{\gamma}{K_+^\tau},$$

$$|\langle k, \omega + P_{0l00}\rangle + \widetilde{\lambda}_j^+| \geq |\langle k, \omega\rangle + \widetilde{\lambda}_j| - \varepsilon K \geq \frac{\gamma}{K^\tau} - \varepsilon K \geq \frac{\gamma}{K_+^\tau}.$$

Similarly, we have

$$|\langle k, \omega + P_{0l00}\rangle + \widetilde{\lambda}_i^+ \pm \widetilde{\lambda}_j^+| \geq \frac{\gamma}{K_+^\tau}.$$

In other cases, the proof is similar, so we omit it. This means that in the next KAM step, small denominator conditions are automatically satisfied for $|k| \leq K$. The following bounds will be used for the measure estimates:

$$\sup_{\xi\in\mathcal{O}} \max_{d\leq 4} \|\partial_\xi^d (A_{[n]}^+ - A_{[n]})\| \leq c\varepsilon,$$

$$\sup_{\xi\in\mathcal{O}} \max_{d\leq 4} |\partial_\xi^d (\Omega_{n'}^+ - \Omega_{n'})| \leq \varepsilon,$$

$$\sup_{\xi\in\mathcal{O}} \max_{d\leq 4} |\partial_\xi^d (\omega_+ - \omega)| \leq \varepsilon,$$

and

$$|P_{ij+}^{011} - P_{ij}^{011}|_{\mathcal{O}} \leq \varepsilon e^{-|i-j|\rho}.$$

4.5. Estimate for the new perturbation

Since

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

where $R(t) = (1-t)(N_+ + \mathcal{B}_+ + \bar{\mathcal{B}}_+ - N - \mathcal{B} - \bar{\mathcal{B}}) + tR$, hence

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

According to Lemma 4.7,

$$\|D\phi_F^t - Id\|_{D_{1\eta}} < c\gamma^{-5} K^{5(\tau+1)} (r - r_+)^{-c} \varepsilon, \quad -1 \leq t \leq 1,$$

thus

$$\|D\phi_F^t\|_{D_{1\eta}} \leq 1 + \|D\phi_F^t - Id\|_{D_{1\eta}} \leq 2, \quad -1 \leq t \leq 1.$$

Due to Lemma 7.3,

$$\|X_{\{R(t),F\}}\|_{D_{2\eta}} \leq c\gamma^{-5} K^{5(\tau+1)} (r - r_+)^{-c} \eta^{-2} \varepsilon^2,$$

and

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

we have

$$\|X_{P_+}\|_{D_\rho(r_+,s_+)} \le c\eta\varepsilon + c\gamma^{-5}K^{5(\tau+1)}(r-r_+)^{-c}\eta^{-2}\varepsilon^2 \le c\varepsilon_+.$$

4.6. Verification of (A5) after one step of KAM iteration

Recall that

$$P_+ = P - R + \{P, F\} + \frac{1}{2!}\{\{N+\mathcal{B}+\bar{\mathcal{B}}, F\}, F\} + \frac{1}{2!}\{\{P, F\}, F\}$$
$$+ \cdots + \frac{1}{n!}\{\cdots\{N+\mathcal{B}+\bar{\mathcal{B}}, \underbrace{F\}\cdots, F}_{n}\} + \frac{1}{n!}\{\cdots\{P, \underbrace{F\}\cdots, F}_{n}\} + \cdots.$$

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

$$\|\frac{\partial^2(P-R)}{\partial z_{n+pc}\partial \bar{z}_{m+pc}} - \lim_{p\to\infty}\frac{\partial^2(P-R)}{\partial z_{n+pc}\partial \bar{z}_{m+pc}}\| \le \frac{\varepsilon}{|p|}e^{-|n-m|\rho} \le \frac{\varepsilon_+}{|p|}e^{-|n-m|\rho_+}.$$

That is to say, $P-R$ satisfies $(A5)$ with $K_+, \varepsilon_+, \rho_+$ in place of K, ε, ρ. The proof of the remaining terms satisfying $(A5)$ is composed by the following two lemmas.

Lemma 4.8. *F satisfies $(A5)$ with $\varepsilon^{\frac{2}{3}}$ in place of ε.*

For the proof see [29].

Lemma 4.9. *Assume that P satisfies $(A5)$, F satisfies $(A5)$ with $\varepsilon^{\frac{2}{3}}$ in place of ε and*

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

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

For the proof see [29].

A KAM-step cycle is now completed.

5. Iteration lemma and convergence

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

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

$$\varepsilon_{\nu+1} = c\gamma^{-5}(r_\nu - r_{\nu+1})^{-c} K_\nu^{5(\tau+1)} \varepsilon_\nu^{\frac{4}{3}}, \tag{5.1}$$

$$\eta_{\nu+1} = \varepsilon_{\nu+1}^{\frac{1}{3}}, L_{\nu+1} = L_\nu + \varepsilon_\nu,$$

$$s_{\nu+1} = 2^{-2}\eta_\nu s_\nu = 2^{-2(\nu+1)}(\prod_{i=0}^{\nu} \varepsilon_i)^{\frac{1}{3}} s_0,$$

$$K_{\nu+1}^{1+3\varepsilon}\rho_{\nu+1} = 1,$$

$$K_{\nu+1} = 3K_\nu = 3^{\nu+1}K_0,$$

$$\Delta_{\nu+1} = K_\nu^3,$$

where c is a constant, $\gamma = \varepsilon_0^{\frac{1}{50}} \gg \varepsilon_0$, and the parameters r_0, ε_0, s_0 and K_0 are defined to be r, ε, s and $K_0^2 e^{-K_0(r_0-r_1)} = \varepsilon_0^{\frac{1}{3}}$ respectively.

5.1. Iteration lemma

The preceding analysis can be summarized as follows.

Lemma 5.1. *Let ε be small enough and $\nu \geq 0$. Suppose that*
(1). $N_\nu + \mathcal{B}_\nu + \bar{\mathcal{B}}_\nu$ is a normal form with parameters ξ satisfying

$$|\langle k, \omega_\nu\rangle| \geq \frac{\gamma}{K_\nu^\tau}, k \neq 0,$$

$$|\langle k, \omega_\nu\rangle \pm \widetilde{\lambda}_j^\nu| \geq \frac{\gamma}{K_\nu^\tau}, j \in [n],$$

$$|\langle k, \omega_\nu\rangle \pm \widetilde{\lambda}_i^\nu \pm \widetilde{\lambda}_j^\nu| \geq \frac{\gamma}{K_\nu^\tau}, i \in [m], j \in [n],$$

$$|\langle k, \omega_\nu\rangle \pm \widetilde{\lambda}_i^\nu \pm \mu_j^\nu| \geq \frac{\gamma}{K_\nu^\tau}, k \neq 0, i \in [n], j \in \{1,2\},$$

$$|\langle k, \omega_\nu\rangle \pm \mu_j^\nu| \geq \frac{\gamma}{K_\nu^\tau}, j \in \{1,2\},$$

$$|det(\langle k, \omega_\nu\rangle I \pm \mathcal{A}_n^\nu \otimes I_2 \pm I_2 \otimes \mathcal{A}_{n'}^\nu)| \geq \frac{\gamma}{K_\nu^\tau}, k \neq 0, n, n' \in \mathcal{L}_2, \tag{5.2}$$

on a closed set $\mathcal{O}_\nu$ of $\mathbb{R}^b$ for all $0 < |k| \leq K_\nu$. Moreover, suppose that $\omega_\nu(\xi)$, $P_{ij\nu}^{011}(\xi)$, $A_{[n]}^\nu(\xi)$ are C_W^4 smooth and satisfy

$$\sup_{\xi\in\mathcal{O}_\nu} \max_{d\leq 4} \|\partial_\xi^d (A_{[n]}^\nu - A_{[n]}^{\nu-1})\| \leq c\varepsilon_{\nu-1},$$

$$\sup_{\xi\in\mathcal{O}_\nu}\max_{d\le 4}|\partial_\xi^d(\Omega_{n'}^{\nu}-\Omega_{n'}^{\nu-1})|\le\varepsilon_{\nu-1},$$

$$\sup_{\xi\in\mathcal{O}_\nu}\max_{d\le 4}|\partial_\xi^d(\omega_\nu-\omega_{\nu-1})|\le\varepsilon_{\nu-1},$$

and

$$|P_{ij\nu}^{011}-P_{ij(\nu-1)}^{011}|_{\mathcal{O}_\nu}\le\varepsilon_{\nu-1}e^{-|i-j|\rho},$$

in the sense of Whitney.
(2). $N_\nu+\mathcal{B}_\nu+\bar{\mathcal{B}}_\nu+P_\nu$ satisfies (A5) with $K_\nu,\varepsilon_\nu,\rho_\nu$ and

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

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

$$\mathcal{O}_{\nu+1}=\mathcal{O}_\nu\setminus(\mathcal{R}_k^{\nu+1}),$$

$$\mathcal{R}^{\nu+1}=\bigcup_{K_\nu<|k|\le K_{\nu+1},[n],[m],n,n'}(\mathcal{R}_k^{\nu+1}\bigcup\mathcal{R}_{kn}^{\nu+1}\bigcup\mathcal{R}_{k[n][m]}^{\nu+1}\bigcup\mathcal{R}_{k[n]n'}^{\nu+1}(\gamma)\bigcup\mathcal{C}_{knn'}^{\nu+1}(\gamma)),$$

where

$$\mathcal{R}_k^{\nu+1}=\{\xi\in\mathcal{O}_\nu:|\langle k,\omega_{\nu+1}\rangle|<\frac{\gamma}{K_{\nu+1}^\tau},k\ne 0\},$$

$$\mathcal{R}_{kn}^{\nu+1}=\{\xi\in\mathcal{O}_\nu:\begin{array}{l}|\langle k,\omega_{\nu+1}\rangle\pm\widetilde{\lambda}_j^{\nu+1}|<\frac{\gamma}{K_{\nu+1}^\tau}\\ |\langle k,\omega_{\nu+1}\rangle\pm\mu_j^{\nu+1}|<\frac{\gamma}{K_{\nu+1}^\tau}\end{array}\},$$

$$\mathcal{R}_{k[n][m]}^{\nu+1}=\{\xi\in\mathcal{O}_\nu:\ |\langle k,\omega_{\nu+1}\rangle\pm\widetilde{\lambda}_i^{\nu+1}\pm\widetilde{\lambda}_j^{\nu+1}|<\frac{\gamma}{K_{\nu+1}^\tau},i\in[m],j\in[n]\},$$

$$\mathcal{R}_{k[n]n'}^{\nu+1}=\{\xi\in\mathcal{O}_\nu:|\langle k,\omega_{\nu+1}\rangle\pm\widetilde{\lambda}_i^{\nu+1}\pm\mu_j^{\nu+1}|<\frac{\gamma}{K_{\nu+1}^\tau},k\ne 0,i\in[n],j\in\{1,2\}\},$$

$$\mathcal{C}_{knn'}^{\nu+1}=\{\xi\in\mathcal{O}_\nu:\ |det(\langle k,\omega_{\nu+1}\rangle I\pm\mathcal{A}_n^{\nu+1}\otimes I_2\pm I_2\otimes\mathcal{A}_{n'}^{\nu+1})|<\frac{\gamma}{K_{\nu+1}^\tau},k\ne 0,n,n'\in\mathcal{L}_2\},$$

with $\omega_{\nu+1}=\omega_\nu+P_{0l00}^\nu$, and a symplectic transformation of variables

$$\Phi_\nu:D_{\rho_\nu}(r_{\nu+1},s_{\nu+1})\times\mathcal{O}_\nu\to D_{\rho_\nu}(r_\nu,s_\nu),\tag{5.3}$$

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

$$\begin{aligned}H_{\nu+1}&=e_{\nu+1}+\langle\omega_{\nu+1},I\rangle+\sum_{[n]}\langle A_{[n]}^{\nu+1}(\xi)z_{[n]},\bar{z}_{[n]}\rangle\\&\quad+\sum_{n'\in\mathcal{L}_2}[(\Omega_{n'}^{\nu+1}-\omega_{i'})z_{n'}\bar{z}_{n'}+(\Omega_{m'}^{\nu+1}-\omega_{j'})z_{m'}\bar{z}_{m'}]+\mathcal{B}_{\nu+1}+\bar{\mathcal{B}}_{\nu+1}+P_{\nu+1},\end{aligned}$$

with

$$\sup_{\xi\in\mathcal{O}_\nu}\max_{d\le 4}\|\partial_\xi^d(A_{[n]}^{\nu+1}-A_{[n]}^{\nu})\|\le c\varepsilon_\nu,$$

$$\sup_{\xi\in\mathcal{O}_\nu}\max_{d\le 4}|\partial_\xi^d(\Omega_{n'}^{\nu+1}-\Omega_{n'}^{\nu})|\le \varepsilon_\nu,$$

$$\sup_{\xi\in\mathcal{O}_\nu}\max_{d\le 4}|\partial_\xi^d(\omega_{\nu+1}-\omega_{\nu})|\le \varepsilon_\nu,$$

$$|P_{ij(\nu+1)}^{011}-P_{ij\nu}^{011}|_{\mathcal{O}_\nu}\le \varepsilon_\nu e^{-|i-j|\rho},$$

$$\|X_{P_{\nu+1}}\|_{D(r_{\nu+1},s_{\nu+1}),\mathcal{O}_{\nu+1}}\le \varepsilon_{\nu+1},$$

in the sense of Whitney.

5.2. Convergence

Suppose that the assumptions of Theorem 2 are satisfied to apply the iteration Lemma with $\nu=0$, recall that

$$\varepsilon_0=\varepsilon, r_0=r, s_0=s, L_0=L, N_0=N, \mathcal{B}_0=\mathcal{B}, P_0=P, \gamma=\varepsilon^{\frac{1}{50}}, K_0^2e^{-K_0(r_0-r_1)}=\varepsilon_0^{\frac{1}{3}}$$

$$\mathcal{O}_0=\left\{\xi\in\mathcal{O}:\begin{array}{l}|\langle k,\omega\rangle|\ge\frac{\gamma}{K_0^\tau}, k\ne 0\\ |\langle k,\omega\rangle\pm\widetilde{\lambda}_j|\ge\frac{\gamma}{K_0^\tau}, j\in[n]\\ |\langle k,\omega\rangle\pm\widetilde{\lambda}_i\pm\widetilde{\lambda}_j|\ge\frac{\gamma}{K_0^\tau}, i\in[m], j\in[n]\\ |\langle k,\omega\rangle\pm\widetilde{\lambda}_i\pm\mu_j|\ge\frac{\gamma}{K_0^\tau}, i\in[n], j\in\{1,2\}\\ |\langle k,\omega\rangle\pm\mu_j|\ge\frac{\gamma}{K_0^\tau}, j\in\{1,2\}\\ |det(\langle k,\omega\rangle I\pm\mathcal{A}_n\otimes I_2\pm I_2\otimes\mathcal{A}_{n'})|\ge\frac{\gamma}{K_0^\tau}, n,n'\in\mathcal{L}_2\end{array}\right\},$$

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:

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

Let $\tilde{\mathcal{O}}=\cap_{\nu=0}^{\infty}\mathcal{O}_\nu$. As in [34,35], thanks to Lemma 4.7, it concludes that $N_\nu, \Psi^\nu, D\Psi^\nu$, ω_ν converge uniformly on $D_{\frac{1}{2}r}(\frac{1}{2}r,0)\times\tilde{\mathcal{O}}$ with

$$\begin{aligned}N_\infty+\mathcal{B}_\infty+\bar{\mathcal{B}}_\infty&=e_\infty+\langle\omega_\infty,I\rangle+\sum_{[n]}\langle A_{[n]}^\infty(\xi)z_{[n]},\bar{z}_{[n]}\rangle\\&\quad+\sum_{n'\in\mathcal{L}_2}[(\Omega_{n'}^\infty-\omega_{i'})z_{n'}\bar{z}_{n'}+(\Omega_{m'}^\infty-\omega_{j'})z_{m'}\bar{z}_{m'}]+\mathcal{B}_\infty+\bar{\mathcal{B}}_\infty.\end{aligned}$$

Since

$$\varepsilon_{\nu+1} = c\gamma^{-5} K_\nu^{5(\tau+1)} (r_\nu - r_{\nu-1})^{-c} \varepsilon_\nu^{\frac{4}{3}},$$

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

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

$$\phi_H^t \circ \Psi^\nu = \Psi^\nu \circ \phi_{H_{\nu+1}}^t. \tag{5.4}$$

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 (5.4). Hence, on $D_{\frac{1}{2}r}(\frac{1}{2}r, 0) \times \tilde{\mathcal{O}}$ we get

$$\phi_H^t \circ \Psi^\infty = \Psi^\infty \circ \phi_{H_\infty}^t \tag{5.5}$$

and

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

It follows from (5.5) that

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

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 $A_{[n]}^\infty, \Omega_{n'}^\infty$. $\square$

6. Measure estimates

This section is the essential part for this paper. For notational convenience, let $\mathcal{O}_{-1} = \mathcal{O}$, $K_{-1} = 0$. Then at ν^{th} step of KAM iteration, we have to exclude the following resonant set

$$\mathcal{R}^{\nu+1} = \bigcup_{K_\nu < |k| \leq K_{\nu+1}, [n], [m], n, n'} (\mathcal{R}_k^{\nu+1} \bigcup \mathcal{R}_{kn}^{\nu+1} \bigcup \mathcal{R}_{k[n][m]}^{\nu+1} \bigcup \mathcal{R}_{k[n]n'}^{\nu+1}(\gamma) \bigcup \mathcal{C}_{knn'}^{\nu+1}(\gamma)),$$

where

$$\mathcal{R}_k^{\nu+1} = \{\xi \in \mathcal{O}_\nu : |\langle k, \omega_{\nu+1}\rangle| < \frac{\gamma}{K_{\nu+1}^\tau}, k \neq 0\},$$

$$\mathcal{R}_{kn}^{\nu+1} = \{\xi \in \mathcal{O}_\nu : \begin{array}{l} |\langle k, \omega_{\nu+1}\rangle \pm \widetilde{\lambda}_j^{\nu+1}| < \frac{\gamma}{K_{\nu+1}^\tau} \\ |\langle k, \omega_{\nu+1}\rangle \pm \mu_j^{\nu+1}| < \frac{\gamma}{K_{\nu+1}^\tau} \end{array} \},$$

$$\mathcal{R}_{k[n][m]}^{\nu+1} = \{\xi \in \mathcal{O}_\nu :\ |\langle k, \omega_{\nu+1}\rangle \pm \widetilde{\lambda}_i^{\nu+1} \pm \widetilde{\lambda}_j^{\nu+1}| < \frac{\gamma}{K_{\nu+1}^\tau}, i \in [m], j \in [n]\},$$

$$\mathcal{R}^{\nu+1}_{k[n]n'} = \{\xi \in \mathcal{O}_\nu : |\langle k, \omega_{\nu+1}\rangle \pm \widetilde{\lambda}^{\nu+1}_i \pm \mu^{\nu+1}_j| < \frac{\gamma}{K^\tau_{\nu+1}}, k \neq 0, i \in [n], j \in \{1,2\}\},$$

$$\mathcal{C}^{\nu+1}_{knn'} = \{\xi \in \mathcal{O}_\nu :\ |det(\langle k, \omega_{\nu+1}\rangle I \pm \mathcal{A}^{\nu+1}_n \otimes I_2 \pm I_2 \otimes \mathcal{A}^{\nu+1}_{n'})| < \frac{\gamma}{K^\tau_{\nu+1}}, k \neq 0, n, n' \in \mathcal{L}_2\},$$

recall that $\omega_{\nu+1}(\xi) = \omega(\xi) + \sum_{j=0}^{\nu} P_{0l00}(\xi)$ with $|\sum_{j=0}^{\nu} P^j_{0l00}(\xi)|_{\mathcal{O}_\nu} < \varepsilon$, and

$$\|A^{\nu+1}_{[n]}(\xi) - A_{[n]}(\xi)\|_{\mathcal{O}_\nu} \leq \sum_{j=0}^{\nu} \|R^{011,j}_{[n][n]}\| \leq \varepsilon,$$

$$|\Omega^{\nu+1}_{n'}(\xi) - \Omega_{n'}(\xi)|_{\mathcal{O}_\nu} \leq \sum_{j=0}^{\nu} |R^{011,j}_{n'n'}| \leq \varepsilon.$$

Remark. From the section 4.4, one has that at $(\nu+1)^{\text{th}}$ step, small divisor conditions are automatically satisfied for $|k| \leq K_\nu$. Hence, we only need to excise the above resonant set $\mathcal{R}^{\nu+1}$.

In the following, we only give the proof for the most complicated case $\{\xi \in \mathcal{O}_\nu : |\langle k, \omega_{\nu+1}\rangle + \widetilde{\lambda}^{\nu+1}_n - \widetilde{\lambda}^{\nu+1}_{n'}| < \frac{\gamma}{K^\tau_{\nu+1}}, n, n' \in \mathcal{L}_1\}$ and $\{\xi \in \mathcal{O}_\nu : |det(\langle k, \omega_{\nu+1}\rangle I + \mathcal{A}^{\nu+1}_n \otimes I_2 - I_2 \otimes \mathcal{A}^{\nu+1}_{n'})| < \frac{\gamma}{K^\tau_{\nu+1}}, n, n' \in \mathcal{L}_2\}$. In other cases, the proof is similar, so we omit it. For simplicity, set $M^{\nu+1} = |\langle k, \omega_{\nu+1}\rangle + \widetilde{\lambda}^{\nu+1}_n - \widetilde{\lambda}^{\nu+1}_{n'}|$ and $Y^{\nu+1} = \langle k, \omega_{\nu+1}\rangle I + \mathcal{A}^{\nu+1}_n \otimes I_2 - I_2 \otimes \mathcal{A}^{\nu+1}_{n'}$, $Y^\nu = \langle k, \omega_\nu\rangle I + \mathcal{A}^\nu_n \otimes I_2 - I_2 \otimes \mathcal{A}^\nu_{n'}$, then for $|k| \leq K_\nu$

$$\begin{aligned}\|(Y^{\nu+1})^{-1}\| &= \|(Y^\nu + (Y^{\nu+1} - Y^\nu))^{-1}\| \\ &= \|(I + (Y^\nu)^{-1}(Y^{\nu+1} - Y^\nu))^{-1}(Y^\nu)^{-1}\| \\ &\leq 2\|(Y^\nu)^{-1}\| \leq 2\frac{K^\tau_\nu}{\gamma} \leq \frac{K^\tau_{\nu+1}}{\gamma}.\end{aligned}$$

Lemma 6.1. *For any given $n, n' \in \mathbb{Z}^2_1$ with $|n - n'| \leq K_{\nu+1}$, either $|\langle k, \omega_{\nu+1}\rangle + \widetilde{\lambda}^{\nu+1}_n - \widetilde{\lambda}^{\nu+1}_{n'}| > 1$ or there are $n_0, n'_0, c \in \mathbb{Z}^2$ with $|n_0|, |n'_0|, |c| \leq 3K^2_{\nu+1}$ and $p \in \mathbb{Z}$, such that $n = n_0 + pc$, $n' = n'_0 + pc$.*

Proof. Since $|n - n'| \leq K_{\nu+1}$, with an elementary calculation

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

If $|\langle n - n', n'\rangle| > K^2_{\nu+1}$, we have $|\langle k, \omega_{\nu+1}\rangle + \widetilde{\lambda}^{\nu+1}_n - \widetilde{\lambda}^{\nu+1}_{n'}| > 1$, there will be no small divisor.

In the case that $|\langle n - n', n'\rangle| \leq K^2_{\nu+1}$, clearly $n - n' = 0$ is trivial. Assume $n - n' \neq 0$, without loss of generality, we assume that the first component $(n - n')_1$ of $n - n'$ is not zero. Let

$$c = (-(n - n')_2, (n - n')_1),$$

then

$$c \perp (n-n')$$

and $c \in \mathbb{Z}^2 \setminus \{0\}$ with $|c| \leq |n-n'| \leq K_{\nu+1}$. Clearly, $c, n-n'$ are linearly independent, hence there exist $x_1, x_2 \in \mathbb{R}$ such that

$$n' = x_1 c + x_2(n-n').$$

Set (here $[\cdot]$ denotes the integer part of $\cdot$)

$$p = [x_1],$$

then $p \in \mathbb{Z}$ and $|n'-pc| \leq 2K_{\nu+1}^2$. Take $n_0' = n'-pc \in \mathbb{Z}^2$ and $n_0 = n_0'+n-n' \in \mathbb{Z}^2$. We have $|n_0'| \leq 2K_{\nu+1}^2$ and

$$|n_0| \leq |n_0'| + |n-n'| \leq 3K_{\nu+1}^2. \quad \square$$

Lemma 6.2.

$$\cup_{n,n' \in \mathcal{L}_1} \mathcal{R}_{k[n][n']}^{\nu+1} \subset \cup_{n_0,n_0',c \in \mathbb{Z}^2, p \in \mathbb{Z}} \mathcal{R}_{k,n_0+pc,n_0'+pc}^{\nu+1}$$

where $|n_0|, |n_0'|, |c| \leq 3K_{\nu+1}^2$.

Proof. If $|\langle n-n', n'\rangle| > K_{\nu+1}^2$, $\mathcal{R}_{k[n][n']}^{\nu+1} = \emptyset$. If $|\langle n-n', n'\rangle| \leq K_{\nu+1}^2$, there exist $n_0, n_0', c \in \mathbb{Z}^2, p \in \mathbb{Z}$ with $|n_0|, |n_0'|, |c| \leq 3K_{\nu+1}^2$ such that $n = n_0+pc$, $n' = n_0'+pc$. Hence

$$\cup_{n,n' \in \mathcal{L}_1} \mathcal{R}_{k[n][n']}^{\nu+1} \subset \cup_{n_0,n_0',c \in \mathbb{Z}^2, p \in \mathbb{Z}} \mathcal{R}_{k,n_0+pc,n_0'+pc}^{\nu+1},$$

where $|n_0|, |n_0'|, |c| \leq 3K_{\nu+1}^2$. $\square$

Lemma 6.3. *For fixed* k, n_0, n_0', c, *one has*

$$\text{meas}(\cup_{p \in \mathbb{Z}} \mathcal{R}_{k,n_0+pc,n_0'+pc}^{\nu+1}) < c\frac{\gamma}{K_{\nu+1}^{\frac{\tau}{2}}}.$$

Proof. Due to Töplitz-Lipschitz property of $N_\nu + \mathcal{B}_\nu + \bar{\mathcal{B}}_\nu + P_\nu$, then

$$|M^{\nu+1}(p) - \lim_{p\to\infty} M^{\nu+1}(p)| < \frac{\varepsilon_0}{|p|}.$$

We define resonant set

$$\mathcal{R}_{kn_0n_0'c\infty}^{\nu+1} = \{\xi \in \mathcal{O}_\nu : |\lim_{p\to\infty} M^{\nu+1}(p))| < \frac{\gamma}{K_{\nu+1}^{\frac{\tau}{2}}}\}. \tag{6.1}$$

For fixed k, n_0, n_0', c,

$$\text{meas}(\mathcal{R}_{kn_0n_0'c\infty}^{\nu+1}) < \frac{\gamma}{K_{\nu+1}^{\frac{\tau}{2}}}.$$

Then for $\xi \in \mathcal{O}_\nu \backslash \mathcal{R}^{\nu+1}_{kn_0n_0'c\infty}$, we have

$$|\lim_{p\to\infty} M^{\nu+1}(p))| \geq \frac{\gamma}{K_{\nu+1}^{\frac{\tau}{2}}}.$$

Case 1: When $|p| > K_{\nu+1}^{\frac{\tau}{2}}$, for $\xi \in \mathcal{O}_\nu \backslash \mathcal{R}^{\nu+1}_{kn_0n_0'c\infty}$, we have

$$\begin{aligned}
&|M^{\nu+1}(p)| \\
&\geq |\lim_{p\to\infty} M^{\nu+1}(p)| - \frac{\varepsilon_0}{|p|} \\
&\geq \frac{\gamma}{K_{\nu+1}^{\frac{\tau}{2}}} - \frac{\varepsilon_0}{K_{\nu+1}^{\frac{\tau}{2}}} \\
&\geq \frac{\gamma}{2K_{\nu+1}^{\frac{\tau}{2}}}.
\end{aligned}$$

Case 2: When $|p| \leq K_{\nu+1}^{\frac{\tau}{2}}$, we define resonant set

$$\mathcal{R}^{\nu+1}_{kn_0n_0'cp} = \{\xi \in \mathcal{O}_\nu : |M^{\nu+1}(p)| < \frac{\gamma}{K_{\nu+1}^{\tau}}\}. \tag{6.2}$$

For fixed k, n_0, n_0', c, p,

$$\text{meas}(\mathcal{R}^{\nu+1}_{kn_0n_0'cp}) < \frac{\gamma}{K_{\nu+1}^{\tau}},$$

then

$$\text{meas}\{\cup_{|p|\leq K_{\nu+1}^{\frac{\tau}{2}}} \mathcal{R}^{\nu+1}_{kn_0n_0'cp}\} < K_{\nu+1}^{\frac{\tau}{2}} \frac{\gamma}{K_{\nu+1}^{\tau}} \leq \frac{\gamma}{K_{\nu+1}^{\frac{\tau}{2}}}.$$

As a consequence,

$$\text{meas}(\cup_{p\in\mathbb{Z}} \mathcal{R}^{\nu+1}_{k,n_0+pc,n_0'+pc}) < c\frac{\gamma}{K_{\nu+1}^{\frac{\tau}{2}}}. \quad \square$$

For $K_\nu < |k| \leq K_{\nu+1}$, we consider $n, n' \in \mathcal{L}_2$ as an example, the other cases can be proved analogously. Assume that (n, m) and (n', m') are resonant pairs in $\mathcal{L}_2$, then

Lemma 6.4. *For any given $n, n' \in \mathbb{Z}_1^2$ with $|n - n'| \leq K_{\nu+1}$, either $|det(\langle k, \omega_{\nu+1}\rangle I + \mathcal{A}_n^{\nu+1} \otimes I_2 - I_2 \otimes \mathcal{A}_{n'}^{\nu+1})| > 1$ or there are $n_0, n_0', c \in \mathbb{Z}^2$ with $|n_0|, |n_0'|, |c| \leq 3K_{\nu+1}^2$ and $p \in \mathbb{Z}$, such that $n = n_0 + pc$, $n' = n_0' + pc$.*

Lemma 6.5.

$$\cup_{n,n'\in\mathbb{Z}_1^2} \mathcal{C}^{\nu+1}_{knn'} \subset \cup_{n_0,n_0',c\in\mathbb{Z}^2,p\in\mathbb{Z}} \mathcal{C}^{\nu+1}_{k,n_0+pc,n_0'+pc}$$

where $|n_0|, |n_0'|, |c| \leq 3K_{\nu+1}^2$.

The proof of this result is quite similar to that given earlier and so is omitted.

Lemma 6.6. *For fixed k, n_0, n_0', c, one has*

$$\text{meas}(\cup_{p\in\mathbb{Z}}\mathcal{C}^{\nu+1}_{k,n_0+pc,n_0'+pc}) < c\frac{\gamma^{\frac{1}{4}}}{K_{\nu+1}^{\frac{\tau}{20}}}.$$

Proof. Due to the analysis above and Töplitz-Lipschitz property of $N+\mathcal{B}+\bar{\mathcal{B}}+P$, the coefficient matrix $Y^{\nu+1}(p)$ has a limit as $p\to\infty$,

$$\|Y^{\nu+1}(p)-\lim_{p\to\infty}Y^{\nu+1}(p)\|\leq\frac{\varepsilon_0}{p}.$$

We define resonant set

$$\mathcal{C}^{\nu+1}_{kn_0n_0'c\infty}=\left\{\xi\in\mathcal{O}_\nu : |det\lim_{p\to\infty}Y^{\nu+1}(p)|<\frac{\gamma}{K_{\nu+1}^{\frac{\tau}{5}}}\right\}.$$

Then for $\xi\in\mathcal{O}_\nu\backslash\mathcal{C}^{\nu+1}_{kn_0n_0'c\infty}$, we have

$$\|(\lim_{p\to\infty}Y^{\nu+1}(p))^{-1}\|\leq\frac{K_{\nu+1}^{\frac{\tau}{5}}}{\gamma}.$$

Since

$$\|Y^{\nu+1}(p)-\lim_{p\to\infty}Y^{\nu+1}(p)\|\leq\frac{\varepsilon_0}{p},$$

for $|p|>K_{\nu+1}^{\frac{\tau}{5}}$, we have

$$\|(Y^{\nu+1}(p))^{-1}\|\leq 2\frac{K_{\nu+1}^{\frac{\tau}{5}}}{\gamma}\leq\frac{K_{\nu+1}^{\tau}}{\gamma}.$$

For $|p|\leq K_{\nu+1}^{\frac{\tau}{5}}$, we define resonant set

$$\mathcal{C}^{\nu+1}_{kn_0n_0'cp}=\{\xi\in\mathcal{O}_\nu : |detY^{\nu+1}(p)|<\frac{\gamma}{K_{\nu+1}^{\tau}}\}. \tag{6.3}$$

In addition

$$\inf_{\xi\in\mathcal{O}}\max_{0<d\leq 4}|\partial_\xi^d(detY^{\nu+1}(p))|\geq\frac{1}{2}|k|\geq\frac{1}{2}K.$$

For fixed k, n_0, n_0', c, p,

$$\text{meas}(\mathcal{C}^{\nu+1}_{kn_0n_0'cp})<(\frac{\gamma}{K_{\nu+1}^{\tau}})^{\frac{1}{4}},$$

then

$$\text{meas}\{\cup_{|p|\leq K_{\nu+1}^{\frac{\tau}{5}}}\mathcal{C}^{\nu+1}_{kn_0n_0'cp}\}<K_{\nu+1}^{\frac{\tau}{5}}(\frac{\gamma}{K_{\nu+1}^{\tau}})^{\frac{1}{4}}\leq\frac{\gamma^{\frac{1}{4}}}{K_{\nu+1}^{\frac{\tau}{20}}}.$$

As a consequence,

$$\text{meas}(\cup_{p\in\mathbb{Z}}\mathcal{C}^{\nu+1}_{k,n_0+pc,n_0'+pc}) < c\frac{\gamma^{\frac{1}{4}}}{K_{\nu+1}^{\frac{\tau}{20}}}. \quad \square$$

The following lemma is now a direct consequence of what we have proved.

Lemma 6.7.

$$\text{meas}(\bigcup_{K_\nu<|k|\le K_{\nu+1}}\mathcal{R}_k^{\nu+1}) \le cK^b_{\nu+1}\frac{\gamma}{K^\tau_{\nu+1}} = c\frac{\gamma}{K^{\tau-b}_{\nu+1}}$$

$$\text{meas}(\bigcup_{K_\nu<|k|\le K_{\nu+1},[n],n}\mathcal{R}_{kn}^{\nu}) \le cK^{2+b}_{\nu+1}\frac{\gamma}{K^\tau_{\nu+1}} = c\frac{\gamma}{K^{\tau-2-b}_{\nu+1}}$$

$$\text{meas}(\bigcup_{K_\nu<|k|\le K_{\nu+1},[n],[m]}\mathcal{R}_{k[n][m]}^{\nu+1}) \le c\frac{\gamma}{K^{\frac{\tau}{2}-12-b}_{\nu+1}}$$

$$\text{meas}(\bigcup_{K_\nu<|k|\le K_{\nu+1},[n],n'}\mathcal{R}_{k[n]n'}^{\nu}) \le c\frac{\gamma^{\frac{1}{2}}}{K^{\frac{\tau}{6}-12-b}_{\nu+1}}$$

$$\text{meas}(\bigcup_{K_\nu<|k|\le K_{\nu+1},n,n'}\mathcal{C}_{knn'}^{\nu}) \le c\frac{\gamma^{\frac{1}{4}}}{K^{\frac{\tau}{20}-12-b}_{\nu+1}}$$

Lemma 6.8. *Let $\tau > 20(12+b+1)$, then the total measure need to exclude along the KAM iteration is*

$$\begin{aligned}
&\text{meas}(\bigcup_{\nu\ge 0}\mathcal{R}^{\nu+1})\\
&= \text{meas}[\bigcup_{\nu\ge 0}\bigcup_{K_\nu<|k|\le K_{\nu+1},[n],[m],n,n'}(\mathcal{R}_k^{\nu+1}\bigcup\mathcal{R}_{k[n]}^{\nu+1}\bigcup\mathcal{R}_{k[n][m]}^{\nu+1}\bigcup\mathcal{R}_{k[n]n'}^{\nu+1}(\gamma)\bigcup\mathcal{C}_{knn'}^{\nu+1}(\gamma))]\\
&\le c\sum_{\nu\ge 0}\frac{\gamma^{\frac{1}{4}}}{K_{\nu+1}} \le c\gamma^{\frac{1}{4}}.
\end{aligned}$$

7. Appendix

Lemma 7.1.

$$\|FG\|_{D_\rho(r,s),\mathcal{O}} \le \|F\|_{D_\rho(r,s),\mathcal{O}}\|G\|_{D_\rho(r,s),\mathcal{O}}.$$

Proof. Since $(FG)_{kl\alpha\beta} = \sum_{k',l',\alpha',\beta'} F_{k-k',l-l',\alpha-\alpha',\beta-\beta'}G_{k'l'\alpha'\beta'}$, we have

$$\|FG\|_{D_\rho(r,s),\mathcal{O}} = \sup_{D_\rho(r,s)}\sum_{k,l,\alpha,\beta}|(FG)_{kl\alpha\beta}|_{\mathcal{O}}|I^l||z^\alpha||\bar{z}^\beta|e^{|k||\text{Im}\theta|}$$

$$\leq \sup_{D_\rho(r,s)} \sum_{k,l,\alpha,\beta} \sum_{k',l',\alpha',\beta'} |F_{k-k',l-l',\alpha-\alpha',\beta-\beta'} G_{k'l'\alpha'\beta'}|_{\mathcal{O}} |I^l||z^\alpha||\bar{z}^\beta| e^{|k||\mathrm{Im}\theta|}$$
$$\leq \|F\|_{D_\rho(r,s),\mathcal{O}} \|G\|_{D_\rho(r,s),\mathcal{O}}$$

and the proof is finished. □

Lemma 7.2. *(Generalized Cauchy inequalities)*

$$\|F_\theta\|_{D_\rho(r-\sigma,s),\mathcal{O}} \leq \frac{c}{\sigma} \|F\|_{D_\rho(r,s),\mathcal{O}},$$

$$\|F_I\|_{D_\rho(r,\frac{1}{2}s),\mathcal{O}} \leq \frac{c}{s^2} \|F\|_{D_\rho(r,s),\mathcal{O}},$$

and

$$\|F_z\|_{D_\rho(r,\frac{1}{2}s),\mathcal{O}} \leq \frac{c}{s} \|F\|_{D_\rho(r,s),\mathcal{O}},$$

$$\|F_{\bar{z}}\|_{D_\rho(r,\frac{1}{2}s),\mathcal{O}} \leq \frac{c}{s} \|F\|_{D_\rho(r,s),\mathcal{O}}.$$

Proof. We only prove the third inequality, the others can be proved similarly. Let $w \neq 0$, then $f(t) = F(z + tw)$ is an analytic map from the complex disc $|t| < \frac{s}{\|w\|_\rho}$ in $\mathbb{C}$ into $D_\rho(r,s)$. Hence

$$\|f'(0)\|_{D_\rho(r,\frac{1}{2}s),\mathcal{O}} = \|F_z w\|_{D_\rho(r,\frac{1}{2}s),\mathcal{O}} \leq \frac{c}{s} \|F\|_{D_\rho(r,s),\mathcal{O}} \cdot \|w\|_\rho,$$

by the usual Cauchy inequality. Since $w \neq 0$, so

$$\frac{\|F_z w\|_{D_\rho(r,\frac{1}{2}s),\mathcal{O}}}{\|w\|_\rho} \leq \frac{c}{s} \|F\|_{D_\rho(r,s),\mathcal{O}},$$

thus

$$\|F_z\|_{D_\rho(r,\frac{1}{2}s),\mathcal{O}} = \sup_{w \neq 0} \frac{\|F_z w\|_{D_\rho(r,\frac{1}{2}s),\mathcal{O}}}{\|w\|_\rho} \leq \frac{c}{s} \|F\|_{D_\rho(r,s),\mathcal{O}}. \quad \square$$

Let $\{\cdot,\cdot\}$ denote the Poisson bracket of smooth functions, i.e.,

$$\{F,G\} = \langle \frac{\partial F}{\partial I}, \frac{\partial G}{\partial \theta} \rangle - \langle \frac{\partial F}{\partial \theta}, \frac{\partial G}{\partial I} \rangle + \mathrm{i}(\langle \frac{\partial F}{\partial z}, \frac{\partial G}{\partial \bar{z}} \rangle - \langle \frac{\partial F}{\partial \bar{z}}, \frac{\partial G}{\partial z} \rangle),$$

then we have the following lemma:

Lemma 7.3. *If*

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

then

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

In particular, if $\eta \sim \varepsilon^{\frac{1}{3}}$, $\varepsilon', \varepsilon'' \sim \varepsilon$, *we have* $\|X_{\{F,G\}}\|_{D_\rho(r-\sigma,\eta s),\mathcal{O}} \sim \varepsilon^{\frac{4}{3}}$.

Proof. By Lemma 7.1 and Lemma 7.2,

$$
\begin{aligned}
&\|\frac{\partial^2 F}{\partial I \partial I}\frac{\partial G}{\partial \theta}\|_{D_\rho(r-\sigma,\frac{1}{2}s)} < c\sigma^{-1}s^{-2}\|\frac{\partial F}{\partial I}\|_{D_\rho(r,s)} \cdot \|\frac{\partial G}{\partial \theta}\|_{D_\rho(r,s)},\\
&\|\frac{\partial^2 F}{\partial I \partial \theta}\frac{\partial G}{\partial \theta}\|_{D_\rho(r-\sigma,\frac{1}{2}s)} < c\sigma^{-1}\|\frac{\partial F}{\partial I}\|_{D_\rho(r,s)} \cdot \|\frac{\partial G}{\partial \theta}\|_{D_\rho(r,s)},\\
&\|\frac{\partial^2 F}{\partial I \partial z}\frac{\partial G}{\partial \theta}\|_{D_\rho(r-\sigma,\frac{1}{2}s)} < c\sigma^{-1}s^{-1}\|\frac{\partial F}{\partial I}\|_{D_\rho(r,s)} \cdot \|\frac{\partial G}{\partial \theta}\|_{D_\rho(r,s)},\\
&\|\frac{\partial^2 F}{\partial I \partial \bar{z}}\frac{\partial G}{\partial \theta}\|_{D_\rho(r-\sigma,\frac{1}{2}s)} < c\sigma^{-1}s^{-1}\|\frac{\partial F}{\partial I}\|_{D_\rho(r,s)} \cdot \|\frac{\partial G}{\partial \theta}\|_{D_\rho(r,s)},\\
&\|\frac{\partial^2 F}{\partial z \partial I}\frac{\partial G}{\partial \bar{z}}\|_{D_\rho(r-\sigma,\frac{1}{2}s)} < c\sigma^{-1}s^{-2}\|\frac{\partial F}{\partial z}\|_{D_\rho(r,s)} \cdot \|\frac{\partial G}{\partial \bar{z}}\|_{D_\rho(r,s)},\\
&\|\frac{\partial^2 F}{\partial z \partial \theta}\frac{\partial G}{\partial \bar{z}}\|_{D_\rho(r-\sigma,\frac{1}{2}s)} < c\sigma^{-1}\|\frac{\partial F}{\partial z}\|_{D_\rho(r,s)} \cdot \|\frac{\partial G}{\partial \bar{z}}\|_{D_\rho(r,s)},\\
&\|\frac{\partial^2 F}{\partial z \partial z}\frac{\partial G}{\partial \bar{z}}\|_{D_\rho(r-\sigma,\frac{1}{2}s)} < c\sigma^{-1}s^{-1}\|\frac{\partial F}{\partial z}\|_{D_\rho(r,s)} \cdot \|\frac{\partial G}{\partial \bar{z}}\|_{D_\rho(r,s)},\\
&\|\frac{\partial^2 F}{\partial z \partial \bar{z}}\frac{\partial G}{\partial \bar{z}}\|_{D_\rho(r-\sigma,\frac{1}{2}s)} < c\sigma^{-1}s^{-1}\|\frac{\partial F}{\partial z}\|_{D_\rho(r,s)} \cdot \|\frac{\partial G}{\partial \bar{z}}\|_{D_\rho(r,s)}.
\end{aligned}
$$

The other cases can be obtained analogously, hence

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

Acknowledgments

We are grateful to the anonymous referees and Jiangsu Provincial Double-Innovation Doctor Program for useful comments and suggestions.

References

[1] P. Baldi, M. Berti, R. Montalto, KAM for quasi-linear and fully nonlinear forced perturbations of Airy equation, Math. Ann. 359 (2014) 471–536.

[2] P. Baldi, M. Berti, R. Montalto, KAM for autonomous quasi-linear perturbations of KdV, Ann. Inst. Henri Poincaré, Anal. Non Linéaire 33 (2016) 1589–1638.

[3] P. Baldi, M. Berti, E. Haus, R. Montalto, Time quasi-periodic gravity water waves in finite depth, Invent. Math. 214 (2018) 739–911.

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

[5] M. Berti, P. Bolle, Sobolev quasi-periodic solutions of multidimensional wave equations with a multiplicative potential, Nonlinearity 25 (2012) 2579–2613.

[6] M. Berti, P. Bolle, Quasi-periodic solutions with Sobolev regularity of NLS on T^d with a multiplicative potential, J. Eur. Math. Soc. 15 (2013) 229–286.

[7] M. Berti, A. Maspero, Long time dynamics of Schrödinger and wave equations on flat tori, J. Differ. Equ. 267 (2019) 1167–1200.

[8] M. Berti, M. Procesi, Nonlinear wave and Schrödinger equations on compact Lie groups and homogeneous spaces, Duke Math. J. 159 (2011) 479–538.

[9] M. Berti, P. Bolle, M. Procesi, An abstract Nash–Moser theorem with parameters and applications to PDEs, Ann. Inst. Henri Poincaré, Anal. Non Linéaire 27 (2010) 377–399.

[10] M. Berti, L. Corsi, M. Procesi, An abstract Nash–Moser theorem and quasi-periodic solutions for NLW and NLS on compact Lie groups and homogeneous manifolds, Commun. Math. Phys. 334 (2015) 1413–1454.

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

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

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

[14] J. Bourgain, Nonlinear Schrödinger Equations, Park City Series, vol. 5, American Mathematical Society, Providence, Rhode Island, 1999.

[15] J. Bourgain, On diffusion in high–dimensional Hamiltonian systems and PDE, J. Anal. Math. 80 (2000) 1–35.

[16] J. Bourgain, Green's Function Estimates for Lattice Schrödinger Operators and Applications, Annals of Mathematics Studies, vol. 158, Princeton University Press, Princeton, NJ, 2005.

[17] J. Bourgain, W.-M. Wang, Quasi–periodic solutions of nonlinear random Schrödinger equations, J. Eur. Math. Soc. 10 (2008) 1–45.

[18] Y. Chen, J. Geng, A KAM theorem for higher dimensional wave equations under nonlocal perturbation, J. Dyn. Differ. Equ. 32 (2020) 419–440.

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

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

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

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

[23] L.H. Eliasson, B. Grebert, S.B. Kuksin, KAM for the nonlinear beam equation, Geom. Funct. Anal. 26 (2016) 1588–1715.

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

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

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

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

[28] J. Geng, J. You, A KAM theorem for higher dimensional nonlinear Schrodinger equation, J. Dyn. Differ. Equ. 25 (2013) 451–476.

[29] J. Geng, X. Xu, J. You, An infinite dimensional KAM theorem and its application to the two dimensional cubic Schrödinger equation, Adv. Math. 226 (2011) 5361–5402.

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

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

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

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

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

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

[36] J. Pöschel, On the construction of almost periodic solutions for a nonlinear Schrödinger equations, Ergod. Theory Dyn. Syst. 22 (2002) 1537–1549.

[37] C. Procesi, M. Procesi, A KAM algorithm for the resonant nonlinear Schrödinger equation, Adv. Math. 272 (2015) 399–470.
[38] M. Procesi, C. Procesi, A normal form for the Schrödinger equation with analytic nonlinearities, Commun. Math. Phys. 312 (2012) 501–557.
[39] W.-M. Wang, Energy supercritical nonlinear Schrödinger equations: quasiperiodic solutions, Duke Math. J. 165 (2016) 1129–1192.
[40] C.E. Wayne, Periodic and quasi-periodic solutions for nonlinear wave equations via KAM theory, Commun. Math. Phys. 127 (1990) 479–528.
[41] 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.
[42] J. Xu, Q. Qiu, J. You, A KAM theorem of degenerate infinite dimensional Hamiltonian systems (I), Sci. China Ser. A 39 (1996) 372–383.
[43] J. Xu, Q. Qiu, J. You, A KAM theorem of degenerate infinite dimensional Hamiltonian systems (II), Sci. China Ser. A 39 (1996) 384–394.
[44] J. You, Perturbations of lower dimensional tori for Hamiltonian systems, J. Differ. Equ. 152 (1999) 1–29.
[45] X. Yuan, Quasi-periodic solutions of completely resonant nonlinear wave equations, J. Differ. Equ. 230 (2006) 213–274.