Available online at www.sciencedirect.com

ScienceDirect

Journal of Differential Equations 364 (2023) 1–52

Journal of Differential Equations

www.elsevier.com/locate/jde

A KAM algorithm for two-dimensional nonlinear Schrödinger equations with spatial variable ☆

Shuaishuai Xue

School of Mathematics, Nanjing Audit University, Nanjing 211815, PR China

Received 9 July 2020; revised 19 February 2023; accepted 21 March 2023

Abstract

We consider the two-dimensional nonlinear Schrödinger equation

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

with periodic boundary conditions, where the nonlinearity $H(x, u, \bar{u}) = \sum_{m=1}^{\infty} \alpha_m(x)|u|^{2p+2m}u$ is a real analytic function in a neighborhood of the origin. We obtain, through a KAM algorithm, a Whitney smooth family of small–amplitude quasi–periodic solutions.

Keywords: Schrödinger equation; KAM tori; Quasi–periodic solutions

1. Introduction and main result

Many PDEs arising in physics can be seen as infinite dimensional Hamiltonian systems. The dynamics of linear Hamiltonian PDEs is quite clear. A natural question is to know what happens of these solutions of linear Hamiltonian PDEs under the effect of the nonlinearity. There have been many remarkable results which reflected some of the main ideas involved in KAM (Kolmogorov–Arnold–Moser) theory on the persistence of quasi-periodic motions under pertur-

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

E-mail address: ssx@nau.edu.cn.

https://doi.org/10.1016/j.jde.2023.03.036

bations. It is a full fledged theory and it provides a systematic tool for the analysis of many dynamical systems and it also has relations with other areas of analysis.

The central question of KAM theory is: do "most" of the solutions of an integrable PDE persist, just slightly deformed, under the effect of a perturbation? The main source of difficulties is the fact that one has to deal with the resonances and small divisors.

There are two main approaches to 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 [11–17,20,40]. 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. These papers renounce 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 use of a combination of KAM algorithm and Birkhoff normal form [4,19,21–23,25–39,41–43,46,47]. 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 we have 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. Therefore, KAM theory is a collection of ideas of how to approach certain problems in perturbation theory connected with small divisors. In order to satisfy this condition, one must discard some parameters. In general, 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 left 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 Birkhoff normal form analysis implies that the frequencies of the expected quasi-periodic solutions vary with their "amplitudes", allowing to prove that the Melnikov non-resonance conditions are satisfied for most amplitudes. The nice normal form is not only an important outcome of the KAM theory, but also a very important ingredient in the proof. 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 Newton's quadratic iteration schemes. By now, KAM theory for 1-dimensional PDEs has reached a satisfactory level while KAM theory for multidimensional PDEs still contains few results, more precisely, when considering PDEs especially in space dimension larger than one, normal frequencies appear in huge clusters of increasing size. A satisfactory future is under construction.

There have been many remarkable works which reflected some of the main ideas involved in KAM theory and usually studied parameter families of PDEs. The first breakthrough result in this direction is made by Bourgain [11,14]. Bourgain's technique is a multiscale inductive analysis based on the repeated use of the resolvent identity. Another stream of results for multidimensional PDEs is made by Eliasson–Kuksin [22]. The authors have developed a tactic to construct quasi-periodic solutions for a higher dimensional Schrödinger equation with a convolution potential on $\mathbb{T}^d$, which proved linear stability. Eliasson–Kuksin require a subtle analysis and the introduction of the concept of "Töplitz-Lipschitz matrices" in order to extract asymptotic information on the eigenvalues, which is important to verify the second Melnikov non-resonance conditions. An essential ingredient in [22] is that finitely many Lipschitz domains cover a neighborhood of ∞. Geng–You [27,28,30] proved that the higher dimensional nonlinear beam equations with a constant mass potential, nonlocal Schrödinger equations and 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. Pöschel [37] 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, for nonresonant, typically using the convolution operator $V(x)$ or the multiplier M_ξ as external parameters in some way as to avoid resonances. Despite under these simplifying conditions the problems are in general complicated. When the equations have no external parameters, one must deal with the resonances. 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 which excite only the Fourier indexes of the tangential sites. A similar strategy was used by Geng–Xu–You [25] to prove that two dimensional cubic Schrödinger equation with periodic boundary conditions admits a family of small-amplitude quasi-periodic solutions. The authors carefully choose the tangential sites $\{i_1, \cdots, i_b\} \subset \mathbb{Z}^2$ in order to make the normal form as sparse as possible so that the homological equations in KAM iteration are easy to be solved. Wang [40], extending the strategy, proves that the energy supercritical nonlinear Schrödinger equations on $\mathbb{T}^d$ admit small–amplitude quasi–periodic solutions. The proof used a bifurcation analysis to prove the invertibility of appropriate linearized operators. The author requires a subtle analysis and the introduction of the concept of "genericity condition" in order to restrict the sizes of these block diagonal matrix with finite types of blocks. We also mention the existence results of large families of stable and unstable quasi-periodic solutions of Procesi and Procesi in [38,39] to arbitrary dimensions for the translational invariant cubic NLS (H has no explicit x-dependence). The authors give the concept of generic on the tangential sites. The concept of generic in [38,39] seems to bear a certain resemblance to the conditions in [40]. The proof of these above conditions is rather complex and takes finer combinatorial analysis. In [25], the existence of admissible set S is proved for $p = 1$ by a very direct and lengthy computation. The present paper (where we consider a larger class of NLS with nonlinearity $|u|^{2p}u$) need a more conceptual proof based on estimates of integral points on algebraic curves, valid for all p. We essentially use the concept of generic in [38,39]. Thus the existence proof of admissible set S in our case refers to [38].

The present paper, we consider NLS,

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

with periodic boundary conditions, where the nonlinearity $H(x, u, \bar{u}) = \sum_{m=1}^{\infty} \alpha_m(x)|u|^{2p+2m}u$ is a real analytic function in a neighborhood of the origin.

When the Hamiltonian nonlinearity does not depend on the space variable x, the equation is translation invariant. Note that the results by Geng–Xu–You [25] on the translational invariant cubic NLS imply existence of quasi-periodic solutions. Geng–Xu–You [25] were able to exploit the corresponding "Special form" conservation, which is preserved along the KAM iteration, to fulfill the nonresonance conditions. A essential ingredient is that such symmetry enables to prove that many monomials are never present along the KAM iteration. In the case of Eq. (1.1), this is a much more difficult situation than the Eq. in [25] because Eq. (1.1) is explicit dependent on the spatial variable and the eigenvalues appear in clusters of unbounded size. In [22] the convolution potential plays the role of "external parameters", while in the case of parameter independent Eq. (1.1) we must use a systematic study of the Birkhoff normal form. We also mention the existence results of large families of stable and unstable quasi-periodic solutions of Procesi and Procesi in [38,39] to arbitrary dimensions for the translationally invariant cubic NLS with no explicit x-dependence. Note that the results by Wang in [40] construct time quasi-periodic solutions to the energy supercritical NLS with explicit dependence on the spatial variable on the torus in arbitrary dimensions. Wang's technique is a bifurcation analysis, by the Nash-Moser method, and does not prove the reducibility.

In this paper we need to address and analyze three issues:

- Choosing an appropriate set of tangential sites $S = \{i_1, \cdots, i_b\} \subset \mathbb{Z}^2$ in order to make the normal form as simple as possible, we need to combine the concept of "Admissible Set" of [25] with the concept of generic of [38,39]. As in [25] the homological equations can be decomposed into a set of linear equations of dimension at most four. As a result, the homological equations are easier to solve in each KAM iteration steps.
- Eq. (1.1) is explicit dependent on the spatial variable and the eigenvalues appear in clusters of unbounded size. We deduce from the results of "Block Decomposition" of [22]. Essential ingredients for Block Decomposition are that the normal frequency clusters are separated into small clusters (of cardinality ≤ 2) which are sufficiently distant from one another and the total number of small clusters is at most $e^{\frac{\log \Delta}{\log\log \Delta}} \ll \Delta^{\varepsilon}$. It should be pointed out that the second property is of course special for the case $x \in \mathbb{T}^2$. We have to modify that strategy in various non-trivial ways, which will become apparent later on in the section 4, 5, 6.
- The "Töplitz-Lipschitz" property of the perturbation. An key step in [22] is that finitely many Lipschitz domains cover a neighborhood of ∞. The goal is to extract asymptotic information on the eigenvalues, so verify the assumption of the second Melnikov non-resonance conditions are objective. 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.

The core of the problem consists in our proof is that the equations have no external parameters and the nonlinearities explicit depend on the spatial variable. In order to avoid the difficulty posed by the violating of the external parameters, we shall use Birkhoff normal form techniques to verify nonresonance conditions. For multidimensional PDEs, it is a challenge due to complicated resonance phenomena. When the nonlinearities depend on the spatial variables, a subtle problem is that there exist normal frequency clusters of unbounded size. A essential ingredient is that the normal frequency clusters are separated into clusters which are sufficiently distant from one another. In all steps we need to combine the concept of "Admissible Set" S of [25] with the structure of "Block Decomposition" of [22]. This is the source of most of the specific problems and complexity for the NLS. Very recently, M. Berti et al. [10] 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 [5,9] abstract Nash–Moser theorem to construct quasi-periodic solutions (see also [7,8] for growth of Sobolev norms for quasi-periodic solutions on T^d and [6] on compact Lie groups and homogeneous spaces; see [1–3] for quasi–linear equations).

More concretely, 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.

It is convenient to define two linear maps:

$$\pi : \mathbb{Z}^b \to Span(S), \pi(e_j) = i_j; \eta : \mathbb{Z}^b \to \mathbb{Z}, \eta(e_j) = 1,$$

where the space $Span(S)$ has a basis over $\mathbb{Z}$ given by the elements $\{i_1, \cdots, i_b\} \subset \mathbb{Z}^2$ and $e_1, \cdots, e_b$ are the elements of the standard basis of the lattice $\mathbb{Z}^b$.

For the definition below, it is again more convenient to define S:

Definition 1.1. Consider the elements

$$X_p := \{l := \sum_{k=1}^{2p} \pm e_{j_k} = \sum_{j=1}^{b} l_j e_j, l \neq 0, -2e_j \forall j, \eta(l) \in \{0, -2\}\}.$$

Notice the constraint $\eta(l) = \sum_{j=1}^{b} l_j \in \{0, -2\}$, we denote all these elements by X_p^0, X_p^{-2} respectively.

The purpose of symbols, such as X_p, X_p^0 and X_p^{-2}, is to impose a list of linear and quadratic constraints on the choice of *Admissible Set*. These constraints make the normal form as simple as possible.

Definition 1.2. A finite set $S = \{i_1, \cdots, i_b\} \subset \mathbb{Z}^2$, $b > 2$ is called *admissible* if
1. For any choice of $2q + 1$ vectors $i_j \in S$ the following holds: If there exists a further vector $w \in \mathbb{Z}^2$ such that

$$\begin{cases} i_1 - i_2 + i_3 - i_4 + \cdots + i_{2q+1} - w = 0, \\ |i_1|^2 - |i_2|^2 + |i_3|^2 - |i_4|^2 + \cdots + |i_{2q+1}|^2 - |w|^2 = 0, \end{cases}$$

then $w \in S$.
2. For all $n_j \in Z, \sum_{j=1}^{b} n_j = 0, 2 \leq \sum_{j=1}^{b} |n_j| \leq 2q + 2$,

$$\sum_{j=1}^{b} n_j i_j \neq 0.$$

3. For any $n \in \mathbb{Z}^2 \setminus S$, there exists at most a point $\{l, m\} \in \mathbb{Z}^{b+2}$ with $l = \sum_{j=1}^{b} l_j e_j \in X_p$, $m \in \mathbb{Z}^2 \setminus S$ such that

$$\begin{cases} n - m + \sum_{j=1}^{b} l_j i_j = 0 \\ |n|^2 - |m|^2 + \sum_{j=1}^{b} l_j |i_j|^2 = 0 \end{cases}, \; l = \sum_{j=1}^{b} l_j e_j \in X_p^0;$$

or

$$\begin{cases} n+m+\sum_{j=1}^{b} l_j i_j = 0 \\ |n|^2+|m|^2+\sum_{j=1}^{b} l_j |i_j|^2 = 0 \end{cases}, \; l=\sum_{j=1}^{b} l_j e_j \in X_p^{-2}.$$

We say respectively that n, m are resonant of the first type (In symbols, $n, m \in \mathcal{L}_1$), $\eta(l)=0$ and the second type (In symbols, $n, m \in \mathcal{L}_2$) $\eta(l)=-2$. By definition, n, m are mutually uniquely determined. Moreover, for all $l=\sum_{j=1}^{b} l_j e_j \in X_p^{-2}$,

$$2\sum_{j=1}^{b} l_j |i_j|^2 + |\sum_{j=1}^{b} l_j i_j|^2 \neq 0.$$

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 $l=\sum_{j=1}^{b} l_j e_j \in X_p^0, l'=\sum_{j=1}^{b} l'_j e_j \in X_p^{-2}$ and $m, m' \in \mathbb{Z}^2 \setminus S$, such that

$$\begin{cases} n-m+\sum_{j=1}^{b} l_j i_j = 0 \\ |n|^2-|m|^2+\sum_{j=1}^{b} l_j |i_j|^2 = 0 \\ n+m'+\sum_{j=1}^{b} l'_j i_j = 0 \\ |n|^2+|m'|^2+\sum_{j=1}^{b} l'_j |i_j|^2 = 0. \end{cases}$$

Remark. In Appendix A of [25], a concrete way of constructing the admissible set is given. It is plausible that any randomly chosen set S is almost surely admissible. The proof of the existence of S of the above Definition which is parallel to special case in [38]. In Proposition 14 of [38], Procesi-Procesi prove that for $d=2$ and every b there exists infinitely many choices of generic tangential sites $S=\{i_1, \cdots, i_b\} \subset \mathbb{Z}^2$.

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 $\mathbb{Z}_1^2 \backslash \mathcal{L}_2$.

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 $[\cdot]_\Delta$).

- 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. There exists a Cantor set $\mathcal{C}$ of positive–measure such that for any $\xi=(\xi_1, \cdots, \xi_b) \in \mathcal{C}$, the nonlinear Schrödinger equation (1.1) admits a small–amplitude, quasi–periodic solution of the form*

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

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.

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. For given b vectors in $\mathbb{Z}^2$, say $\{i_1,\cdots,i_b\}$, we denote $\mathbb{Z}_1^2=\mathbb{Z}^2\setminus\{i_1,\cdots,i_b\}$. 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 $F_{\alpha\beta}=\sum\limits_{k\in\mathbb{Z}^b,l\in\mathbb{N}^b}F_{kl\alpha\beta}(\xi)I^le^{\mathrm{i}\langle k,\theta\rangle}$ is C_W^{4p} function in parameter ξ in the sense of Whitney, we denote

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

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

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

(the derivatives with respect to ξ are in the sense of Whitney). We define the weighted norm of F by

$$\|F\|_{D_\rho(r,s),\mathcal{O}} \equiv \sup_{D_\rho(r,s)} \|F\|_{\mathcal{O}}. \tag{2.3}$$

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[1]

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

Suppose that S is an admissible set.

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

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

$$N = \langle \omega(\xi), I\rangle + \sum_{n\in\mathbb{Z}_1^2\setminus\mathcal{L}_2} \Omega_n(\xi) z_n\bar{z}_n + \sum_{n\in\mathcal{L}_2} [(\Omega_n + \langle l^+, \omega\rangle) z_n\bar{z}_n + (\Omega_m - \langle l^-, \omega\rangle) z_m\bar{z}_m],$$

$$\mathcal{B} = \sum_{\substack{l=l^+-l^-\in X_p^{-2} \\ \alpha\in\mathbb{N}^b, |l^++\alpha|_1=p-1}} \frac{p}{(2\pi)^{2p}} \binom{p-1}{l^++\alpha}\binom{p+1}{l^-+\alpha} \sum_{n\in\mathcal{L}_2} \xi^{\frac{l}{2}+\alpha} z_n z_m,$$

$$\bar{\mathcal{B}} = \sum_{\substack{l=l^+-l^-\in X_p^{-2} \\ \alpha\in\mathbb{N}^b, |l^++\alpha|_1=p-1}} \frac{p}{(2\pi)^{2p}} \binom{p-1}{l^++\alpha}\binom{p+1}{l^-+\alpha} \sum_{n\in\mathcal{L}_2} \xi^{\frac{l}{2}+\alpha} \bar{z}_n \bar{z}_m,$$

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

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

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

[1] The norm $\|\cdot\|_{D_\rho(r,s),\mathcal{O}}$ for scalar functions is defined in (2.3). The vector function $G : D_\rho(r,s) \times \mathcal{O} \to \mathbb{C}^m$, ($m < \infty$) is similarly defined as $\|G\|_{D_\rho(r,s),\mathcal{O}} = \sum_{i=1}^m \|G_i\|_{D_\rho(r,s),\mathcal{O}}$.

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 $z_{[n]} = (z_i)_{i\in[n]}$, $\bar{z}_{[n]} = (\bar{z}_i)_{i\in[n]}$.

$$H = \langle \omega, I\rangle + \sum_{[n]} \langle A_{[n]} z_{[n]}, \bar{z}_{[n]}\rangle + \sum_{n\in\mathcal{L}_2} [(\Omega_n + \langle l^+, \omega\rangle) z_n \bar{z}_n + (\Omega_m - \langle l^-, \omega\rangle) z_m \bar{z}_m] + \mathcal{B} + \bar{\mathcal{B}} + P$$

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

We consider Hamiltonian H satisfying the following hypotheses:

$(A1)$ *Nondegeneracy:* Suppose for $\forall \xi \in \mathcal{O}$,

$$\begin{cases} \operatorname{rank}(\frac{\partial \omega_1}{\partial \xi}, \cdots, \frac{\partial \omega_b}{\partial \xi}) = \kappa, \\ \operatorname{rank}(\frac{\partial^{|\beta|}\omega}{\partial \xi^\beta} | \forall \beta, 1 \le |\beta| \le b - \kappa + 1) = b, \end{cases} \tag{2.6}$$

where κ is a given integer with $1 \le \kappa \le b$, $\frac{\partial \omega_1}{\partial \xi}, \cdots, \frac{\partial \omega_b}{\partial \xi}$ are vectors of all 1–order partial derivatives in ξ, and for a fixed β, $\frac{\partial^{|\beta|}\omega}{\partial \xi^\beta} = (\frac{\partial^{|\beta|}\omega_1}{\partial \xi^\beta}, \cdots, \frac{\partial^{|\beta|}\omega_b}{\partial \xi^\beta})$.

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

$$\Omega_n = \varepsilon^{-a}|n|^2 + \tilde{\Omega}_n, a \ge 0, n \in \mathbb{Z}_1^2 \backslash \mathcal{L}_2, \tag{2.7}$$

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

$(A3)$ *Melnikov's non–resonance conditions:* For $n \in \mathbb{Z}_1^2 \backslash \mathcal{L}_2$, 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 $i \neq j$, $\Omega_{ij} = 0$; if $i = j$, $\Omega_{ij} = \Omega_i$. When $|i - j| > K$, $P_{ij}^{011} = 0$, where $A_{[n]}'s$ are C_W^{4p} functions of ξ with C_W^{4p}-norm bounded by some positive constant L, that is to say

$$\sup_{\xi\in\mathcal{O}} \max_{0<d\le 4p} \|\partial_\xi^d A_{[n]}\| \le L.$$

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

$$|\langle k, \omega\rangle| \ge \frac{\gamma}{K^\tau}, k \neq 0,$$

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

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

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

Let

$$\mathcal{A}_n = \begin{pmatrix} \Omega_n + \langle l^+, \omega\rangle & -C(l,\alpha)\xi^{\frac{l}{2}+\alpha} \\ C(l,\alpha)\xi^{\frac{l}{2}+\alpha} & -(\Omega_m - \langle l^-, \omega\rangle) \end{pmatrix}, n \in \mathcal{L}_2,$$

where $C(l,\alpha) = \sum_{\substack{l=l^+-l^-\in X_p^{-2} \\ \alpha\in\mathbb{N}^b, |l^++\alpha|_1=p-1}} \frac{p}{(2\pi)^{2p}} \begin{pmatrix} p-1 \\ l^++\alpha \end{pmatrix} \begin{pmatrix} p+1 \\ l^-+\alpha \end{pmatrix}$.

We assume that $\omega(\xi)$, $\mathcal{A}_n(\xi) \in C_W^{4p}(\mathcal{O})$ and there exist $\gamma, \tau > 0$ such that[2] (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_W^{4p} Whitney smooth in ξ; in addition

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

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

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

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

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

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

[2] 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.

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

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 $\operatorname{meas}(\mathcal{O}\setminus\mathcal{O}_\gamma)=O(\gamma^{\frac{1}{4p}})$ and two maps (analytic in θ and C_W^{4p} in ξ)*

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

where Ψ is $\frac{\varepsilon}{\gamma^{4p}}$-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 [22,24,25,38]. 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 terms of degree $2p+1$ 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 and nonlinear Schrödinger equations explicit depend on the spatial variable. This requires a subtle analysis and the introduction of the results of "Block Decomposition". Mention that the dimension of tori is two is essential to our analysis for nonlinear Schrödinger equations. We have to modify that strategy in various non-trivial ways, which will become apparent later on 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 consider the two–dimensional nonlinear Schrödinger equations

$$iu_t-\triangle u+|u|^{2p}u+H(x,u,\bar{u})=0,\quad t\in\mathbb{R},x\in\mathbb{T}^2 \tag{3.1}$$

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 $H(x,u,\bar{u})=\sum_{m=1}^{\infty}\alpha_m(x)|u|^{2p+2m}u$ 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

$$H = \langle Au, u\rangle + \frac{1}{p+1}\int_{\mathbb{T}^2} |u|^{2p+2}\,dx + \sum_{m=1}^{\infty}\frac{1}{p+1+m}\int_{\mathbb{T}^2}\alpha_m(x)|u|^{2p+2+2m}\,dx, \tag{3.3}$$

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

$$\begin{aligned}\dot{q}_n &= \mathrm{i}(\lambda_n q_n + \frac{\partial G}{\partial \bar{q}_n}),\\ G &\equiv \frac{1}{(p+1)(2\pi)^{2p}}\sum_{k_1-k_2+k_3-k_4+\cdots+k_{2p+1}-k_{2p+2}=0} q_{k_1}\bar{q}_{k_2}q_{k_3}\bar{q}_{k_4}\cdots q_{k_{2p+1}}\bar{q}_{k_{2p+2}}\\ &\quad+\sum_{m=1}^{\infty}\frac{1}{p+1+m}\int_{\mathbb{T}^2}\alpha_m(x)|\sum_{n\in\mathbb{Z}^2}q_n\phi_n(x)|^{2p+2+2m}\,dx,\end{aligned} \tag{3.4}$$

with corresponding Hamiltonian function

$$\begin{aligned}H &= \sum_{n\in\mathbb{Z}^2}\lambda_n q_n\bar{q}_n + \frac{1}{(p+1)(2\pi)^{2p}}\sum_{k_1-k_2+k_3-k_4+\cdots+k_{2p+1}-k_{2p+2}=0} q_{k_1}\bar{q}_{k_2}q_{k_3}\bar{q}_{k_4}\cdots q_{k_{2p+1}}\bar{q}_{k_{2p+2}}\\ &\quad+\sum_{m=1}^{\infty}\frac{1}{p+1+m}\int_{\mathbb{T}^2}\alpha_m(x)|\sum_{n\in\mathbb{Z}^2}q_n\phi_n(x)|^{2p+2+2m}\,dx\\ &=\sum_{n\in\mathbb{Z}^2}\lambda_n|q_n|^2+G\end{aligned} \tag{3.5}$$

As in [34,35,26], 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*

$$\|G_{\bar{q}}\|_\rho \le c\|q\|_\rho^{2p+1}. \tag{3.6}$$

Proof.

$$
\begin{aligned}
\|G_{\bar q}\|_\rho &= \sum_{n\in\mathbb{Z}^2} |G_{\bar q_n}| e^{|n|\rho} \\
&\le c \sum_{n,\alpha,\beta-e_n,|\alpha|+|\beta-e_n|=2p+1} |q^\alpha \bar q^{\beta-e_n}| e^{|n|\rho} \\
&\le c \sum_{\alpha,\beta-e_n,|\alpha|+|\beta-e_n|=2p+1} |q^\alpha \bar q^{\beta-e_n}| e^{|\alpha|\rho} e^{|\beta-e_n|\rho} \\
&\le c\|q\|_\rho^{2p+1}.
\end{aligned}
$$

The proof is completed. □

For an admissible set of tangential site $S=\{i_1,\cdots,i_b\}\subset\mathbb{Z}^2$, we have a nice normal form for H.

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

$$
\begin{cases}
\omega_i(\xi) = \varepsilon^{-3p}|i|^2 + \dfrac{1}{(p+1)(2\pi)^{2p}}\partial_{\xi_i} A_{p+1}(\xi) \\
\Omega_n = \varepsilon^{-3p}|n|^2 + \dfrac{p+1}{(2\pi)^{2p}} A_p(\xi),
\end{cases}
$$

$$
\mathcal{A} = \sum_{\substack{l=l^+-l^-\in X_p^0\\ \alpha\in\mathbb{N}^b,|l^++\alpha|_1=p}} \frac{p+1}{(2\pi)^{2p}} \binom{p}{l^++\alpha}\binom{p}{l^-+\alpha} \sum_{n\in\mathcal{L}_1} \xi^{\frac{l}{2}+\alpha} e^{i\langle l,\theta\rangle} w_n \bar w_m,
$$

$$
\mathcal{B} = \sum_{\substack{l=l^+-l^-\in X_p^{-2}\\ \alpha\in\mathbb{N}^b,|l^++\alpha|_1=p-1}} \frac{p}{(2\pi)^{2p}} \binom{p-1}{l^++\alpha}\binom{p+1}{l^-+\alpha} \sum_{n\in\mathcal{L}_2} \xi^{\frac{l}{2}+\alpha} e^{i\langle l,\theta\rangle} w_n w_m,
$$

$$
\bar{\mathcal{B}} = \sum_{\substack{l=l^+-l^-\in X_p^{-2}\\ \alpha\in\mathbb{N}^b,|l^++\alpha|_1=p-1}} \frac{p}{(2\pi)^{2p}} \binom{p-1}{l^++\alpha}\binom{p+1}{l^-+\alpha} \sum_{n\in\mathcal{L}_2} \xi^{\frac{l}{2}+\alpha} e^{-i\langle l,\theta\rangle} \bar w_n \bar w_m,
$$

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

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

$$F = \sum_{\substack{\alpha,\beta\in(\mathbb{Z}^2)^N:|\alpha|=|\beta|=p+1 \\ \sum_{k\in\mathbb{Z}^2}(\alpha_k-\beta_k)k=0,\sum_{k\in\mathbb{Z}^2}(\alpha_k-\beta_k)|k|^2\neq 0 \\ \sharp S\cap\{\frac{\pi(\alpha_k)}{\eta(\alpha_k)},\frac{\pi(\beta_k)}{\eta(\beta_k)}\}_{k\in\mathbb{Z}^2}\geq 2p}} \frac{i}{(p+1)(2\pi)^{2p}(\sum_{k\in\mathbb{Z}^2}(\alpha_k-\beta_k)|k|^2)} q^\alpha \bar{q}^\beta, \tag{3.9}$$

and X_F^1 be the time one map of the flow of the associated Hamiltonian systems. For the definition below, it is again more convenient to define $H\circ X_F^1$: we define

$$A_r(\xi_1,\cdots,\xi_b) = \sum_{\sum_j k_j = r} \begin{pmatrix} & r & \\ k_1, & \cdots & ,k_b \end{pmatrix}^2 \prod_j \xi_j^{k_j}.$$

The change of variables X_F^1 sends H to

$$\begin{aligned}
H\circ X_F^1 &= H + \{H,F\} + \int_0^1 (1-t)\{\{H,F\},F\}\circ\phi_F^t dt \\
&= \sum_{i\in S}\lambda_i|q_i|^2 + \sum_{n\in\mathbb{Z}_1^2}\lambda_i|w_n|^2 + \frac{1}{(p+1)(2\pi)^{2p}}A_{p+1}(|q_{i_1}|^2,\cdots,|q_{i_b}|^2) \\
&+ \sum_{n\in\mathbb{Z}_1^2}\frac{p+1}{(2\pi)^{2p}}A_p(|q_{i_1}|^2,\cdots,|q_{i_b}|^2)|w_n|^2 \\
&+ \sum_{\substack{l=l^+-l^-\in X_p^0 \\ \alpha\in\mathbb{N}^b,|l^++\alpha|_1=p}} \frac{p+1}{(2\pi)^{2p}}\begin{pmatrix} p \\ l^++\alpha\end{pmatrix}\begin{pmatrix} p \\ l^-+\alpha\end{pmatrix}\sum_{n\in\mathcal{L}_1} q^{l^+}\bar{q}^{l^-}q^\alpha\bar{q}^\alpha w_n\bar{w}_m + \\
&\sum_{\substack{l=l^+-l^-\in X_p^{-2} \\ \alpha\in\mathbb{N}^b,|l^++\alpha|_1=p-1}} \frac{p}{(2\pi)^{2p}}\begin{pmatrix} p-1 \\ l^++\alpha\end{pmatrix}\begin{pmatrix} p+1 \\ l^-+\alpha\end{pmatrix}\sum_{n\in\mathcal{L}_2}(q^{l^+}\bar{q}^{l^-}q^\alpha\bar{q}^\alpha\bar{w}_n\bar{w}_m + \bar{q}^{l^+}q^{l^-}q^\alpha\bar{q}^\alpha w_n w_m) \\
&+ O(|q|^{2p-1}\|w\|_\rho^3 + |q|^{2p-2}\|w\|_\rho^4 + \cdots + \|w\|_\rho^{2p+2} + |q|^{4p+2} \\
&+ |q|^{4p+1}\|w\|_\rho + |q|^{4p}\|w\|_\rho^2 + \cdots + |q|^{2p+1}\|w\|_\rho^{2p+1}).
\end{aligned} \tag{3.10}$$

We remind that (n,m) are resonant pairs and $l=l^+-l^-\in X_p, \alpha\in\mathbb{N}^b$ are determined by (n,m).

Next we 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

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

we have

$$
\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(I_i+\xi_i)+\sum_{n\in\mathbb{Z}_1^2}\lambda_i|w_n|^2+\frac{1}{(p+1)(2\pi)^{2p}}A_{p+1}(I_1+\xi_1,\cdots,I_b+\xi_b)\\
&+\sum_{n\in\mathbb{Z}_1^2}\frac{p+1}{(2\pi)^{2p}}A_p(I_1+\xi_1,\cdots,I_b+\xi_b)|w_n|^2\\
&+\sum_{\substack{l=l^+-l^-\in X_p^0\\ \alpha\in\mathbb{N}^b,|l^++\alpha|_1=p}}\frac{p+1}{(2\pi)^{2p}}\binom{p}{l^++\alpha}\binom{p}{l^-+\alpha}\sum_{n\in\mathcal{L}_1}(I+\xi)^{\frac{l}{2}+\alpha}e^{i\langle l,\theta\rangle}w_n\bar{w}_m+\\
\sum_{\substack{l=l^+-l^-\in X_p^{-2}\\ \alpha\in\mathbb{N}^b,|l^++\alpha|_1=p-1}}&\frac{p}{(2\pi)^{2p}}\binom{p-1}{l^++\alpha}\binom{p+1}{l^-+\alpha}\sum_{n\in\mathcal{L}_2}(I+\xi)^{\frac{l}{2}+\alpha}(e^{-i\langle l,\theta\rangle}\bar{w}_n\bar{w}_m+e^{i\langle l,\theta\rangle}w_nw_m)\\
&+O(|\xi|^{p-\frac12}\|w\|_\rho^3+|\xi|^{p-1}\|w\|_\rho^4+\cdots+\|w\|_\rho^{2p+2}+|\xi|^{2p+1}+|\xi|^{2p+\frac12}\|w\|_\rho\\
&+|\xi|^{2p}\|w\|_\rho^2+\cdots+|\xi|^{p+\frac12}\|w\|_\rho^{2p+1})\\
&=\sum_{i\in S}\lambda_iI_i+\sum_{n\in\mathbb{Z}_1^2}\lambda_i|w_n|^2+\frac{1}{(p+1)(2\pi)^{2p}}\langle\nabla_\xi A_{p+1}(\xi),I\rangle\\
&+\sum_{n\in\mathbb{Z}_1^2}\frac{p+1}{(2\pi)^{2p}}A_p(\xi)|w_n|^2\\
&+\sum_{\substack{l=l^+-l^-\in X_p^0\\ \alpha\in\mathbb{N}^b,|l^++\alpha|_1=p}}\frac{p+1}{(2\pi)^{2p}}\binom{p}{l^++\alpha}\binom{p}{l^-+\alpha}\sum_{n\in\mathcal{L}_1}\xi^{\frac{l}{2}+\alpha}e^{i\langle l,\theta\rangle}w_n\bar{w}_m+\\
\sum_{\substack{l=l^+-l^-\in X_p^{-2}\\ \alpha\in\mathbb{N}^b,|l^++\alpha|_1=p-1}}&\frac{p}{(2\pi)^{2p}}\binom{p-1}{l^++\alpha}\binom{p+1}{l^-+\alpha}\sum_{n\in\mathcal{L}_2}\xi^{\frac{l}{2}+\alpha}(e^{-i\langle l,\theta\rangle}\bar{w}_n\bar{w}_m+e^{i\langle l,\theta\rangle}w_nw_m)\\
&+O(|\xi|^{p-\frac12}\|w\|_\rho^3+|\xi|^{p-1}\|w\|_\rho^4+\cdots+\|w\|_\rho^{2p+2}+|\xi|^{2p+1}+|\xi|^{2p+\frac12}\|w\|_\rho\\
&+|\xi|^{2p}\|w\|_\rho^2+\cdots+|\xi|^{p+\frac12}\|w\|_\rho^{2p+1}+|\xi|^{p-1}|I|^2+\cdots+|I|^{p+1}\\
&+|\xi|^{p-1}|I|\|w\|_\rho^2+\cdots+|I|^p\|w\|_\rho^2).
\end{aligned}
$$

By the scaling in time

$$\xi\to\varepsilon^3\xi,\ I\to\varepsilon^5I,\ \theta\to\theta,\ w\to\varepsilon^{\frac52}w,\ \bar{w}\to\varepsilon^{\frac52}\bar{w},$$

we finally arrive at the rescaled Hamiltonian

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

where

$$
N = \sum_{i\in S} \varepsilon^{-3p}\lambda_i I_i + \sum_{n\in\mathbb{Z}_1^2} \varepsilon^{-3p}\lambda_i |w_n|^2 + \frac{1}{(p+1)(2\pi)^{2p}} \langle \nabla_\xi A_{p+1}(\xi), I\rangle
$$

$$
+ \sum_{n\in\mathbb{Z}_1^2} \frac{p+1}{(2\pi)^{2p}} A_p(\xi) |w_n|^2,
$$

$$
\begin{cases}
\omega_i(\xi) = \varepsilon^{-3p}|i|^2 + \dfrac{1}{(p+1)(2\pi)^{2p}} \partial_{\xi_i} A_{p+1}(\xi) \\
\Omega_n = \varepsilon^{-3p}|n|^2 + \dfrac{p+1}{(2\pi)^{2p}} A_p(\xi),
\end{cases}
$$

$$
\mathcal{A} = \sum_{\substack{l=l^+-l^-\in X_p^0 \\ \alpha\in\mathbb{N}^b, |l^++\alpha|_1=p}} \frac{p+1}{(2\pi)^{2p}} \binom{p}{l^++\alpha} \binom{p}{l^-+\alpha} \sum_{n\in\mathcal{L}_1} \xi^{\frac{l}{2}+\alpha} e^{i\langle l,\theta\rangle} w_n \bar{w}_m,
$$

$$
\mathcal{B} = \sum_{\substack{l=l^+-l^-\in X_p^{-2} \\ \alpha\in\mathbb{N}^b, |l^++\alpha|_1=p-1}} \frac{p}{(2\pi)^{2p}} \binom{p-1}{l^++\alpha} \binom{p+1}{l^-+\alpha} \sum_{n\in\mathcal{L}_2} \xi^{\frac{l}{2}+\alpha} e^{i\langle l,\theta\rangle} w_n w_m,
$$

$$
\bar{\mathcal{B}} = \sum_{\substack{l=l^+-l^-\in X_p^{-2} \\ \alpha\in\mathbb{N}^b, |l^++\alpha|_1=p-1}} \frac{p}{(2\pi)^{2p}} \binom{p-1}{l^++\alpha} \binom{p+1}{l^-+\alpha} \sum_{n\in\mathcal{L}_2} \xi^{\frac{l}{2}+\alpha} e^{-i\langle l,\theta\rangle} \bar{w}_n \bar{w}_m,
$$

$$
\begin{aligned}
|P| = \quad & O(\varepsilon|\xi|^{p-\frac{1}{2}}\|w\|_\rho^3 + \varepsilon^2|\xi|^{p-1}\|w\|_\rho^4 + \cdots + \varepsilon^{2p}\|w\|_\rho^{2p+2} + \varepsilon^{3p-2}|\xi|^{2p+1} \\
& + \varepsilon^{3p-1}|\xi|^{2p+\frac{1}{2}}\|w\|_\rho + \varepsilon^{3p}|\xi|^{2p}\|w\|_\rho^2 + \cdots + \varepsilon^{5p-1}|\xi|^{p+\frac{1}{2}}\|w\|_\rho^{2p+1} \\
& + \varepsilon^2|\xi|^{p-1}|I|^2 + \cdots + \varepsilon^{2p}|I|^{p+1} + \varepsilon^2|\xi|^{p-1}|I|\|w\|_\rho^2 + \cdots + \varepsilon^{2p}|I|^p\|w\|_\rho^2). \quad \square
\end{aligned}
\tag{3.11}
$$

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 [44].

Lemma 3.2. *For any $k_1, k_2, \cdots, k_m \in \mathbb{Z}^b$, non-singular $m\times m$ matrix S_1 with $S_1^T\bar{S}_1 = I$, the map $\Phi_0 : (\theta, I, w, w) \to (\theta_+, I_+, z, z)$ defined by*

$$
\begin{cases}
\theta_+ = \theta \\
I_+ = I - \sum_{j=1}^m w_j\bar{w}_j k_j \\
z = S_1 E w \\
\bar{z} = \bar{S}_1 \bar{E} \bar{w}
\end{cases}
$$

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 [44]. Although the proof is trivial, the consequences of this result are of importance.

A nonlinear symplectic coordinates transformation Φ:

$$\begin{cases} \theta_+ = \theta \\ I_+ = I - \sum\limits_{n\in\mathcal{L}_1} (w_n\bar{w}_n l^+ + w_m\bar{w}_m l^-) - \sum\limits_{n'\in\mathcal{L}_2} (w_{n'}\bar{w}_{n'} l^+ - w_{m'}\bar{w}_{m'} l^-) \\ \begin{pmatrix} z_n \\ z_m \end{pmatrix} = S_1 \begin{pmatrix} e^{i\langle l^+,\theta\rangle} & 0 \\ 0 & e^{i\langle l^-,\theta\rangle} \end{pmatrix} \begin{pmatrix} w_n \\ w_m \end{pmatrix}, \begin{pmatrix} \bar{z}_n \\ \bar{z}_m \end{pmatrix} = \bar{S}_1 \begin{pmatrix} e^{-i\langle l^+,\theta\rangle} & 0 \\ 0 & e^{-i\langle l^-,\theta\rangle} \end{pmatrix} \begin{pmatrix} \bar{w}_n \\ \bar{w}_m \end{pmatrix}, n \in \mathcal{L}_1 \\ \begin{pmatrix} z_n \\ z_m \end{pmatrix} = \begin{pmatrix} e^{i\langle l^+,\theta\rangle} & 0 \\ 0 & e^{-i\langle l^-,\theta\rangle} \end{pmatrix} \begin{pmatrix} w_n \\ w_m \end{pmatrix}, \begin{pmatrix} \bar{z}_n \\ \bar{z}_m \end{pmatrix} = \begin{pmatrix} e^{-i\langle l^+,\theta\rangle} & 0 \\ 0 & e^{i\langle l^-,\theta\rangle} \end{pmatrix} \begin{pmatrix} \bar{w}_n \\ \bar{w}_m \end{pmatrix}, n \in \mathcal{L}_2, \\ z_n = w_n, \bar{z}_n = \bar{w}_n, n \in \mathbb{Z}_1^2 \setminus (\mathcal{L}_1 \cup \mathcal{L}_2). \end{cases}$$

It is again more convenient to define $H \circ \Psi \circ \Phi$: we define $\iota = \varepsilon^{-3p}(|i_1|^2, |i_2|^2, \cdots, |i_b|^2)$, $\xi = (\xi_{i_1}, \xi_{i_2}, \cdots, \xi_{i_b})$, $|l| = l^+ + l^-$,

$$G = \langle l, \nabla_\xi A_{p+1}(\xi)\rangle^2 + 4 \sum_{\substack{l=l^+-l^-\in X_p^0 \\ \alpha\in\mathbb{N}^b, |l^++\alpha|_1 = p}} (p+1)^4 \left[\begin{pmatrix} p \\ l^+ + \alpha \end{pmatrix} \begin{pmatrix} p \\ l^- + \alpha \end{pmatrix} \right]^2 \xi^{l+2\alpha}.$$

We get Hamiltonian systems with the Hamiltonian

$$\begin{aligned} H \circ \Psi \circ \Phi = {} & \langle \omega(\xi), I_+\rangle + \sum_{n\in\mathbb{Z}_1^2\setminus(\mathcal{L}_1\cup\mathcal{L}_2)} \Omega_n(\xi) z_n\bar{z}_n \\ & + \sum_{n\in\mathcal{L}_1} \Big[\Big(\varepsilon^{-3p}|n|^2 + \langle l^+, \iota\rangle + \frac{p+1}{(2\pi)^{2p}} A_p(\xi) \\ & + \frac{1}{2(p+1)(2\pi)^{2p}} \langle |l|, \nabla_\xi A_{p+1}(\xi)\rangle + \frac{1}{2(p+1)(2\pi)^{2p}} \sqrt{G} \Big) z_n\bar{z}_n \\ & + \Big(\varepsilon^{-3p}|m|^2 + \langle l^-, \iota\rangle + \frac{p+1}{(2\pi)^{2p}} A_p(\xi) + \frac{1}{2(p+1)(2\pi)^{2p}} \langle |l|, \nabla_\xi A_{p+1}(\xi)\rangle \\ & - \frac{1}{2(p+1)(2\pi)^{2p}} \sqrt{G} \Big) z_m\bar{z}_m \Big] \\ & + \sum_{n\in\mathcal{L}_2} [(\Omega_n + \langle l^+, \omega\rangle) z_n\bar{z}_n + (\Omega_m - \langle l^-, \omega\rangle) z_m\bar{z}_m] + \\ \sum_{\substack{l=l^+-l^-\in X_p^{-2} \\ \alpha\in\mathbb{N}^b, |l^++\alpha|_1 = p-1}} & \frac{p}{(2\pi)^{2p}} \begin{pmatrix} p-1 \\ l^+ + \alpha \end{pmatrix} \begin{pmatrix} p+1 \\ l^- + \alpha \end{pmatrix} \sum_{n\in\mathcal{L}_2} \xi^{\frac{l}{2}+\alpha} (\bar{z}_n\bar{z}_m + z_n z_m) \\ & + P(\theta_+, I_+, z, \bar{z}, \xi) \end{aligned}$$

$$= \quad N + \mathcal{B} + \bar{\mathcal{B}} + P \tag{3.12}$$

where

$$N = \langle \omega(\xi), I_+ \rangle + \sum_{n \in \mathbb{Z}_1^2 \setminus \mathcal{L}_2} \Omega_n(\xi) z_n \bar{z}_n + \sum_{n \in \mathcal{L}_2} [(\Omega_n + \langle l^+, \omega \rangle) z_n \bar{z}_n + (\Omega_m - \langle l^-, \omega \rangle) z_m \bar{z}_m]$$

$$\begin{cases} \omega_i(\xi) = \varepsilon^{-3p}|i|^2 + \dfrac{1}{(p+1)(2\pi)^{2p}} \partial_{\xi_i} A_{p+1}(\xi) \\ \Omega_n = \varepsilon^{-3p}|n|^2 + \dfrac{p+1}{(2\pi)^{2p}} A_p(\xi), n \in \mathbb{Z}_1^2 \setminus \mathcal{L}_1 \\ \Omega_n = \varepsilon^{-3p}|n|^2 + \langle l^+, \iota \rangle + \frac{p+1}{(2\pi)^{2p}} A_p(\xi) + \frac{1}{2(p+1)(2\pi)^{2p}} \langle |l|, \nabla_\xi A_{p+1}(\xi) \rangle \\ \qquad + \frac{1}{2(p+1)(2\pi)^{2p}} \sqrt{G}, n \in \mathcal{L}_1 \\ \Omega_m = \varepsilon^{-3p}|m|^2 + \langle l^-, \iota \rangle + \frac{p+1}{(2\pi)^{2p}} A_p(\xi) + \frac{1}{2(p+1)(2\pi)^{2p}} \langle |l|, \nabla_\xi A_{p+1}(\xi) \rangle \\ \qquad - \frac{1}{2(p+1)(2\pi)^{2p}} \sqrt{G}, n \in \mathcal{L}_1 \end{cases}$$

$$\mathcal{B} = \sum_{\substack{l = l^+ - l^- \in X_p^{-2} \\ \alpha \in \mathbb{N}^b, |l^+ + \alpha|_1 = p-1}} \frac{p}{(2\pi)^{2p}} \binom{p-1}{l^+ + \alpha} \binom{p+1}{l^- + \alpha} \sum_{n \in \mathcal{L}_2} \xi^{\frac{l}{2} + \alpha} z_n z_m,$$

$$\bar{\mathcal{B}} = \sum_{\substack{l = l^+ - l^- \in X_p^{-2} \\ \alpha \in \mathbb{N}^b, |l^+ + \alpha|_1 = p-1}} \frac{p}{(2\pi)^{2p}} \binom{p-1}{l^+ + \alpha} \binom{p+1}{l^- + \alpha} \sum_{n \in \mathcal{L}_2} \xi^{\frac{l}{2} + \alpha} \bar{z}_n \bar{z}_m.$$

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.

$$\begin{aligned} H &= \langle \omega, I \rangle + \sum_{[n]} \langle A_{[n]} z_{[n]}, \bar{z}_{[n]} \rangle + \sum_{n \in \mathcal{L}_2} [(\Omega_n + \langle l^+, \omega \rangle) z_n \bar{z}_n + (\Omega_m - \langle l^-, \omega \rangle) z_m \bar{z}_m] \\ &\quad + \mathcal{B} + \bar{\mathcal{B}} + P(\theta, I, z, \bar{z}, \xi) \\ &= N + \mathcal{B} + \bar{\mathcal{B}} + P \end{aligned} \tag{3.13}$$

Note that $A_{[n]}$ is $\sharp[n] \times \sharp[n]$ matrix in (3.13) and

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

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

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

Verification of $(A1)$: The Jacobian of ω is a matrix. The entries in the matrix are polynomials in ξ of degree $p-1$ with integer coefficients. The coefficients of $\frac{\partial \omega_i}{\partial \xi_j}$ are

$$\frac{p}{(2\pi)^{2p}} \binom{p-1}{k_1 \quad , \cdots , \quad k_i - 1 \quad , \cdots , \quad k_j - 1, \quad \cdots , \quad k_b} \binom{p+1}{k_1 \quad , \cdots , \quad k_b}$$

while the coefficients of $\frac{\partial \omega_i}{\partial \xi_i}$ are

$$\frac{p}{(2\pi)^{2p}}\binom{p-1}{k_1 \quad ,\cdots, \quad k_i-2 \quad ,\cdots, \quad k_j, \quad \cdots, \quad k_b}\binom{p+1}{k_1 \quad ,\cdots, \quad k_b}.$$

It is convenient to define a prime r dividing $p+1$, that is $p+1=r^t s$, where r does not divide s and $\binom{p+1}{k_1 \quad ,\cdots, \quad k_b}$. At first we divide the Jacobian of ω by p, then modulo by r. The matrix is diagonal matrix with non-zero terms by a direct computation. Therefore, it is easy to check that $\det \frac{\partial \omega}{\partial \xi} \neq 0$. Thus $(A1)$ is verified.
Verification of $(A2)$: Take $a = 3p$, the proof is obvious.
Verification of $(A3)$: This part is the same as [25], for the sake of completeness, we rewrite it as follows: In the following, we only give the proof for the most complicated case. Let

$$\mathcal{A}_n = \begin{pmatrix} \Omega_n + \langle l^+, \omega\rangle & -C(l,\alpha)\xi^{\frac{l}{2}+\alpha} \\ C(l,\alpha)\xi^{\frac{l}{2}+\alpha} & -(\Omega_m - \langle l^-, \omega\rangle) \end{pmatrix}, n \in \mathcal{L}_2.$$

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

$$\langle k,\omega\rangle \pm \Omega_n \pm \Omega_{n'}$$

Set $\iota = \varepsilon^{-3p}(|i_1|^2, |i_2|^2, \cdots, |i_b|^2)$, $\xi = (\xi_{i_1}, \xi_{i_2}, \cdots, \xi_{i_b})$, $|l| = l^+ + l^-$, and notice that

$$\begin{cases} n - m + \sum_{j=1}^b l_j i_j = 0 \\ |n|^2 - |m|^2 + \sum_{j=1}^b l_j |i_j|^2 = 0 \end{cases}, \ l = l^+ - l^- = \sum_{j=1}^b l_j e_j \in X_p^0.$$

Its eigenvalues are

$$\langle k,\iota\rangle \pm \left(\varepsilon^{-3p}|n|^2 + \langle l^+,\iota\rangle + \frac{p+1}{(2\pi)^{2p}}A_p(\xi)\right) + \frac{1}{2(p+1)(2\pi)^{2p}}\langle 2k \pm |l|, \nabla_\xi A_{p+1}(\xi)\rangle$$
$$\pm \left(\varepsilon^{-3p}|n'|^2 + \langle l'^+,\iota\rangle + \frac{p+1}{(2\pi)^{2p}}A_p(\xi) + \frac{1}{2(p+1)(2\pi)^{2p}}\langle |l'|, \nabla_\xi A_{p+1}(\xi)\rangle\right)$$
$$\pm \frac{1}{2(p+1)(2\pi)^{2p}}(\sqrt{G} \pm \sqrt{G'}).$$

- If $l^+ \neq l'^+$, all the eigenvalues are not identically zero due to the presence of the square root terms.
- $l^+ = l'^+$, consequently $l^- = l'^-$, hence
  1. if the eigenvalue is

$$
\begin{aligned}
&\langle k, \iota\rangle + \left(\varepsilon^{-3p}|n|^2 + \langle l^+, \iota\rangle + \frac{p+1}{(2\pi)^{2p}} A_p(\xi)\right) + \frac{1}{2(p+1)(2\pi)^{2p}}\langle 2k + |l|, \nabla_\xi A_{p+1}(\xi)\rangle \\
&-\left(\varepsilon^{-3p}|n'|^2 + \langle l^+, \iota\rangle + \frac{p+1}{(2\pi)^{2p}} A_p(\xi) + \frac{1}{2(p+1)(2\pi)^{2p}}\langle |l|, \nabla_\xi A_{p+1}(\xi)\rangle\right) \\
&+\frac{1}{2(p+1)(2\pi)^{2p}}(\sqrt{G} - \sqrt{G}) \\
&= \langle k, \iota\rangle + \varepsilon^{-3p}(|n|^2 - |n'|^2) + \frac{1}{p(p+1)(2\pi)^{2p}}\langle H(A_{p+1}(\xi))k, \xi\rangle,
\end{aligned}
$$

     where $H(A_{p+1}(\xi))$ is Hessian matrix of $A_{p+1}(\xi)$. We modulo Hessian matrix by the prime r. The matrix is diagonal matrix with non-zero terms by a direct computation. Therefore, the Hessian matrix is a non-degenerate matrix and it is easy to check that $H(A_{p+1}(\xi))k \neq 0$ for $k \neq 0$;
  2. if the eigenvalue is

$$
\begin{aligned}
&\langle k, \iota\rangle + \left(\varepsilon^{-3p}|n|^2 + \langle l^+, \iota\rangle + \frac{p+1}{(2\pi)^{2p}} A_p(\xi)\right) + \frac{1}{2(p+1)(2\pi)^{2p}}\langle 2k + |l|, \nabla_\xi A_{p+1}(\xi)\rangle \\
&+\left(\varepsilon^{-3p}|n'|^2 + \langle l^+, \iota\rangle + \frac{p+1}{(2\pi)^{2p}} A_p(\xi) + \frac{1}{2(p+1)(2\pi)^{2p}}\langle |l|, \nabla_\xi A_{p+1}(\xi)\rangle\right) \\
&+\frac{1}{2(p+1)(2\pi)^{2p}}(\sqrt{G} - \sqrt{G}) \\
&= \langle k, \iota\rangle + \varepsilon^{-3p}(|n|^2 + |n'|^2) + 2\langle l^+, \iota\rangle + \frac{1}{p(p+1)(2\pi)^{2p}}\langle H(A_{p+1}(\xi))(k + |l|) \\
&+2(p+1)^2\nabla_\xi A_p(\xi), \xi\rangle,
\end{aligned}
$$

     when $H(A_{p+1}(\xi))(k + |l|) + 2(p+1)^2\nabla_\xi A_p(\xi) = 0$, it is easy to check that

$$
\begin{aligned}
&\binom{p+1}{p+1 \quad , \cdots, \quad 0}(k+|l|)_1 + \binom{p+1}{p \quad , 1, \quad \cdots}(k+|l|)_2 + \cdots + \\
&\binom{p+1}{p \quad , \cdots, \quad 1}(k+|l|)_b = -2(p+1)\binom{p}{p \quad , \cdots, \quad 0}, \\
&\binom{p+1}{1 \quad , p, \quad \cdots}(k+|l|)_1 + \binom{p+1}{0 \quad , p+1, \quad \cdots}(k+|l|)_2 + \cdots + \\
&\binom{p+1}{0, p \quad , \cdots, \quad 1}(k+|l|)_b = -2(p+1)\binom{p}{0, p \quad , \cdots, \quad 0}, \\
&\qquad\qquad \vdots \quad \vdots \quad \vdots
\end{aligned}
$$

$$\binom{p+1}{1\ ,\cdots,\ p}(k+|l|)_1+\binom{p+1}{0,1\ ,\cdots,\ p}(k+|l|)_2+\cdots+$$
$$\binom{p+1}{0\ ,\cdots,\ p+1}(k+|l|)_b=-2(p+1)\binom{p}{0\ ,\cdots,\ p}.$$

Thus, all components of $k+|l|$ are equal and $[1+(p+1)(b-1)](k+|l|)_1=-2(p+1)$. This equation has no integer solutions.

Thus all eigenvalues are not identically zero.

Case 2. $n\in\mathcal{L}_1, n'\in\mathcal{L}_2$. Set

$$M=\Big(\langle l,\nabla_\xi A_{p+1}(\xi)\rangle+2(p+1)A_p(\xi)\Big)^2$$
$$+4\sum_{\substack{l=l^+-l^-\in X_p^{-2}\\ \alpha\in\mathbb{N}^b,|l^++\alpha|_1=p-1}} p^2(p+1)^2\left[\binom{p-1}{l^++\alpha}\binom{p+1}{l^-+\alpha}\right]^2\xi^{l+2\alpha}$$

In this case, the eigenvalues are

$$\langle k,\iota\rangle\pm\left(\varepsilon^{-3p}|n|^2+\langle l^+,\iota\rangle+\frac{p+1}{(2\pi)^{2p}}A_p(\xi)\right)+\frac{1}{2(p+1)(2\pi)^{2p}}\langle 2k\pm|l|,\nabla_\xi A_{p+1}(\xi)\rangle$$
$$\pm\left(\varepsilon^{-3p}|n'|^2+\langle l'^+,\iota\rangle+\frac{1}{2(p+1)(2\pi)^{2p}}\langle|l'|,\nabla_\xi A_{p+1}(\xi)\rangle\right)$$
$$\pm\frac{1}{2(p+1)(2\pi)^{2p}}(\sqrt{G}\pm\sqrt{M'}).$$

- If the presence of the square root terms, all the eigenvalues are not identically zero.
- If the square root terms non-existent, hence if the eigenvalue is

$$\langle k,\iota\rangle+\left(\varepsilon^{-3p}(|n|^2\pm|n'|^2)+\langle l^+\pm l'^+,\iota\rangle\right)$$
$$+\frac{1}{2p(p+1)(2\pi)^{2p}}\langle H(A_{p+1}(\xi))(2k+|l|\pm|l'|)+2(p+1)^2\nabla_\xi A_p(\xi),\xi\rangle$$

  when $H(A_{p+1}(\xi))(2k+|l|\pm|l'|)+2(p+1)^2\nabla_\xi A_p(\xi)=0$, we also have a fact that all components of $2k+|l|\pm|l'|$ are equal and $[1+(p+1)(b-1)](2k+|l|\pm|l'|)_1=-2(p+1)$. This equation has no integer solutions.

Thus all eigenvalues are not identically zero.

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

$$\langle k,\iota\rangle\pm\left(\varepsilon^{-3p}|n|^2+\langle l^+,\iota\rangle+\frac{1}{2(p+1)(2\pi)^{2p}}\langle 2k\pm|l|,\nabla_\xi A_{p+1}(\xi)\rangle\right)$$

$$\pm\left(\varepsilon^{-3p}|n'|^2+\langle l'^+,\iota\rangle+\frac{1}{2(p+1)(2\pi)^{2p}}\langle|l'|,\nabla_\xi A_{p+1}(\xi)\rangle\right)$$
$$\pm\frac{1}{2(p+1)(2\pi)^{2p}}(\sqrt{M}\pm\sqrt{M'}).$$

- If $l^+\neq l'^+$, all the eigenvalues are not identically zero due to the presence of the square root terms.
- If $l^+=l'^+$, consequently $l^-=l'^-$, hence
  1. if the eigenvalue is

$$\begin{aligned}&\langle k,\iota\rangle+\left(\varepsilon^{-3p}|n|^2+\langle l^+,\iota\rangle+\frac{1}{2(p+1)(2\pi)^{2p}}\langle 2k+|l|,\nabla_\xi A_{p+1}(\xi)\rangle\right)\\&-\left(\varepsilon^{-3p}|n'|^2+\langle l^+,\iota\rangle+\frac{1}{2(p+1)(2\pi)^{2p}}\langle|l|,\nabla_\xi A_{p+1}(\xi)\rangle\right)\\&+\frac{1}{2(p+1)(2\pi)^{2p}}(\sqrt{M}-\sqrt{M})\\&=\langle k,\iota\rangle+\varepsilon^{-3p}(|n|^2-|n'|^2)+\frac{1}{p(p+1)(2\pi)^{2p}}\langle H(A_{p+1}(\xi))k,\xi\rangle,\end{aligned}$$

  where $H(A_{p+1}(\xi))$ is Hessian matrix of $A_{p+1}(\xi)$. At first we modulo Hessian matrix by the prime r. Then the matrix is diagonal matrix with non-zero terms by a direct computation. Therefore, the Hessian matrix is a non-degenerate matrix and it is easy to check that $H(A_{p+1}(\xi))k\neq 0$ for $k\neq 0$;
  2. if the eigenvalue is

$$\begin{aligned}&\langle k,\iota\rangle+\left(\varepsilon^{-3p}|n|^2+\langle l^+,\iota\rangle+\frac{1}{2(p+1)(2\pi)^{2p}}\langle 2k+|l|,\nabla_\xi A_{p+1}(\xi)\rangle\right)\\&+\left(\varepsilon^{-3p}|n'|^2+\langle l^+,\iota\rangle+\frac{1}{2(p+1)(2\pi)^{2p}}\langle|l|,\nabla_\xi A_{p+1}(\xi)\rangle\right)\\&+\frac{1}{2(p+1)(2\pi)^{2p}}(\sqrt{M}-\sqrt{M})\\&=\langle k,\iota\rangle+\varepsilon^{-3p}(|n|^2+|n'|^2)+2\langle l^+,\iota\rangle+\frac{1}{p(p+1)(2\pi)^{2p}}\langle H(A_{p+1}(\xi))(k+|l|),\xi\rangle,\end{aligned}$$

  when $H(A_{p+1}(\xi))(k+|l|)=0$, then $k=-|l|$ are total possible solutions for that sort of equation. While at this time, when $|n|\neq|m'|$,

$$\langle k,\iota\rangle+\varepsilon^{-3p}(|n|^2+|n'|^2)+2\langle l^+,\iota\rangle=\varepsilon^{-3p}\left(|n|^2-|m'|^2\right)\neq 0.\tag{3.14}$$

Thus all 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 $4p$. Thus

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

By excluding some parameter set with measure $\mathcal{O}(\gamma^{\frac{1}{4p}})$, 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 $s = \varepsilon^{\frac{1}{2p}}$, according to Lemma 3.1, $\|G_{\bar{q}}\|_{\rho} \leq c\|q\|_{\rho}^{2p+1}$, then

$$\sum_{n\in\mathbb{Z}_1^2}\|P_{w_n}\|_{\mathcal{O}}e^{|n|\rho} + \sum_{n\in\mathbb{Z}_1^2}\|P_{\bar{w}_n}\|_{\mathcal{O}}e^{|n|\rho} = \|P_w\|_{\rho} + \|P_{\bar{w}}\|_{\rho} \leq c\|q\|_{\rho}^{2p+1} \leq c(|I|^{p+\frac{1}{2}} + \|w\|_{\rho}^{2p+1}).$$

In addition,

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

thus

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

According to Cauchy estimates,

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

Hence

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

Thus $(A4)$ is verified.

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

$$F = \sum_{\substack{\alpha,\beta\in(\mathbb{Z}^2)^N:|\alpha|=|\beta|=p+1\\ \sum_{k\in\mathbb{Z}^2}(\alpha_k-\beta_k)k=0,\sum_{k\in\mathbb{Z}^2}(\alpha_k-\beta_k)|k|^2\neq 0\\ \sharp S\cap\{\frac{\pi(\alpha_k)}{\eta(\alpha_k)},\frac{\pi(\beta_k)}{\eta(\beta_k)}\}_{k\in\mathbb{Z}^2}\geq 2p}} \frac{i}{(p+1)(2\pi)^{2p}(\sum_{k\in\mathbb{Z}^2}(\alpha_k-\beta_k)|k|^2)} q^\alpha \bar{q}^\beta.$$

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

$$\begin{aligned}
&\sum_{\substack{l=l^+-l^-\in X_p^0\\ \alpha\in\mathbb{N}^b,|l^++\alpha|_1=p}} \frac{i}{(p+1)(2\pi)^{2p}(\sum_{j=1}^b l_j|i_j|^2+\lambda_{n+tc}-\lambda_{m+tc})} q^{l^+}\bar{q}^{l^-}q^\alpha\bar{q}^\alpha w_{n+tc}\bar{w}_{m+tc}\\
=&\sum_{\substack{l=l^+-l^-\in X_p^0\\ \alpha\in\mathbb{N}^b,|l^++\alpha|_1=p}} \frac{i}{(p+1)(2\pi)^{2p}(\sum_{j=1}^b l_j|i_j|^2+|n|^2-|m|^2+2t\langle n-m,c\rangle)}\\
&\times q^{l^+}\bar{q}^{l^-}q^\alpha\bar{q}^\alpha w_{n+tc}\bar{w}_{m+tc}.
\end{aligned}$$

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

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

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

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

Similarly,

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

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 (3.12). 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

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

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,

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

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

$$\begin{aligned}
H = \quad & N + \mathcal{B} + \bar{\mathcal{B}} + P \\
= \quad & \langle \omega, I\rangle + \sum_{[n]} \langle A_{[n]} z_{[n]}, \bar{z}_{[n]}\rangle \qquad (4.2)\\
& + \sum_{n\in\mathcal{L}_2} [(\Omega_n + \langle l^+, \omega\rangle) z_n \bar{z}_n + (\Omega_m - \langle l^-, \omega\rangle) z_m \bar{z}_m] + \\
\sum_{\substack{l=l^+-l^-\in X_p^{-2}\\ \alpha\in\mathbb{N}^b, |l^++\alpha|_1=p-1}} & \frac{p}{(2\pi)^{2p}} \binom{p-1}{l^++\alpha} \binom{p+1}{l^-+\alpha} \sum_{n\in\mathcal{L}_2} \xi^{\frac{l}{2}+\alpha} (\bar{z}_n \bar{z}_m + z_n z_m) \\
& + P(\theta, I, z, \bar{z}, \xi)
\end{aligned}$$

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.

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

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

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

where

$$\mathcal{A}_n = \begin{pmatrix} \Omega_n + \langle l^+, \omega\rangle & -C(l,\alpha)\xi^{\frac{l}{2}+\alpha} \\ C(l,\alpha)\xi^{\frac{l}{2}+\alpha} & -(\Omega_m - \langle l^-, \omega\rangle) \end{pmatrix}, n \in \mathcal{L}_2.$$

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

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

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 $\Phi : D_+ \times \mathcal{O}_+ = D_\rho(r_+, s_+) \times \mathcal{O}_+ \to D_\rho(r,s) \times \mathcal{O}$ such that the transformed Hamiltonian $H_+ = N_+ + \mathcal{B}_+ + \bar{\mathcal{B}}_+ + P_+ \equiv H \circ \Phi$ 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

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

$$R_1 = \sum_{|k|\le K, n\in\mathcal{L}_2} (P_n^{k10} z_n + P_m^{k10} z_m + P_n^{k01}\bar{z}_n + P_m^{k01}\bar{z}_m) e^{\mathrm{i}\langle k,\theta\rangle} + \sum_{|k|\le K,[n]} (\langle R_{[n]}^{k10}, z_{[n]}\rangle + \langle R_{[n]}^{k01}, \bar{z}_{[n]}\rangle) e^{\mathrm{i}\langle k,\theta\rangle}$$

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

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 $\alpha = 0, \beta = e_n + e_m$.

Note that, $R_{[n]}^{k10}$, $R_{[n]}^{k01}$, $R_{[m][n]}^{k20}$, $R_{[m][n]}^{k02}$ 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

$$R_{[n]}^{k10} = (P_i^{k10})_{i\in[n]}, R_{[n]}^{k01} = (P_i^{k01})_{i\in[n]}, |[n]| \le K,$$
$$R_{[m][n]}^{k20} = (R_{ij}^{k20})_{i\in[m], j\in[n]},$$

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

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

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

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

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 $\rho \le r_+$

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

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

$$\begin{aligned}
H\circ\phi_F^1 &= (N+\mathcal{B}+\bar{\mathcal{B}}+R)\circ\phi_F^1+(P-R)\circ\phi_F^1\\
&= N+\mathcal{B}+\bar{\mathcal{B}}+\{N+\mathcal{B}+\bar{\mathcal{B}},F\}+R\\
&+\int_0^1(1-t)\{\{N+\mathcal{B}+\bar{\mathcal{B}},F\},F\}\circ\phi_F^t dt\\
&+\int_0^1\{R,F\}\circ\phi_F^t dt+(P-R)\circ\phi_F^1 \qquad (4.6)\\
&= N_++\mathcal{B}_++\bar{\mathcal{B}}_++P_++\{N+\mathcal{B}+\bar{\mathcal{B}},F\}+R\\
&- P_{0000}-\langle\hat{\omega},I\rangle-\sum_{[n]}\langle P^{011}_{[n][n]}z_{[n]},\bar{z}_{[n]}\rangle-\sum_{n\in\mathcal{L}_2}(P^{011}_{nn}z_n\bar{z}_n+P^{011}_{mm}z_m\bar{z}_m)-\hat{\mathcal{B}}-\hat{\bar{\mathcal{B}}},
\end{aligned}$$

where

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

$$\hat{\mathcal{B}}=\sum_{n\in\mathcal{L}_2}P^{020}_{nm}z_nz_m$$

$$\hat{\bar{\mathcal{B}}}=\sum_{n\in\mathcal{L}_2}P^{002}_{nm}\bar{z}_n\bar{z}_m$$

$$N_+=N+P_{0000}+\langle\hat{\omega},I\rangle+\sum_{[n]}\langle P^{011}_{[n][n]}z_{[n]},\bar{z}_{[n]}\rangle+\sum_{n\in\mathcal{L}_2}(P^{011}_{nn}z_n\bar{z}_n+P^{011}_{mm}z_m\bar{z}_m), \qquad (4.7)$$

$$\mathcal{B}_+=\mathcal{B}+\hat{\mathcal{B}}, \qquad (4.8)$$

$$\bar{\mathcal{B}}_+=\bar{\mathcal{B}}+\hat{\bar{\mathcal{B}}}=\bar{\mathcal{B}}+\bar{\hat{\mathcal{B}}}, \qquad (4.9)$$

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

We shall find a function F:

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

where

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

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

$$
\begin{aligned}
F_2 = & \sum_{|k|\le K, n\in\mathcal{L}_2, n'\in\mathcal{L}_2, |k|+|n-n'|\neq 0} (F^{k11}_{n'n} z_{n'}\bar{z}_n + F^{k11}_{nn'} 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'|\neq 0} (F^{k11}_{m'n} z_{m'}\bar{z}_n + F^{k11}_{nm'} 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'|\neq 0} (F^{k11}_{n'm} z_{n'}\bar{z}_m + F^{k11}_{mn'} 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'|\neq 0} (F^{k11}_{m'm} z_{m'}\bar{z}_m + F^{k11}_{mm'} 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'|\neq 0 (or) |k|+|n'-m|\neq 0} (F^{k20}_{n'n} z_{n'} z_n + F^{k02}_{n'n} \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'|\neq 0 (or) |k|+|n'-n|\neq 0} (F^{k20}_{m'n} z_{m'} z_n + F^{k02}_{m'n} \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'|\neq 0 (or) |k|+|n'-n|\neq 0} (F^{k20}_{n'm} z_{n'} z_m + F^{k02}_{n'm} \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'|\neq 0 (or) |k|+|n'-m|\neq 0} (F^{k20}_{m'm} z_{m'} z_m + F^{k02}_{m'm} \bar{z}_{m'}\bar{z}_m) e^{\mathrm{i}\langle k,\theta\rangle} \\
& + \sum_{|k|\le K, [n], [m]} (\langle F^{k20}_{[m][n]} z_{[n]}, z_{[m]}\rangle + \langle F^{k02}_{[m][n]} \bar{z}_{[n]}, \bar{z}_{[m]}\rangle) e^{\mathrm{i}\langle k,\theta\rangle} \\
& + \sum_{|k|\le K, [n], [m], |k|+||n|-|m||\neq 0} \langle F^{k11}_{[m][n]} z_{[n]}, \bar{z}_{[m]}\rangle e^{\mathrm{i}\langle k,\theta\rangle} \\
& + \sum_{|k|\le K, n\in\mathbb{Z}^2_1\setminus\mathcal{L}_2, n'\in\mathcal{L}_2} (F^{k20}_{nn'} z_n z_{n'} + F^{k20}_{nm'} z_n z_{m'}) e^{\mathrm{i}\langle k,\theta\rangle} \\
& + \sum_{|k|\le K, n\in\mathbb{Z}^2_1\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}^2_1\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}
\end{aligned}
$$

where $F^{k10}_{[n]}$, $F^{k01}_{[n]}$, $F^{k20}_{[m][n]}$, $F^{k02}_{[m][n]}$ 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

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

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

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

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

if $|i+j|\le K$, $f_{ij}^{k02}=F_{ij}^{k02}$; if $|i+j|>K$, $f_{ij}^{k02}=0$,

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

if $|i-j|\le K$, $f_{ij}^{k11}=F_{ij}^{k11}$; if $|i-j|>K$, $f_{ij}^{k11}=0$, satisfying the equation

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

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*

$$\begin{array}{lcl}(\langle k,\omega\rangle)F_{kl00}&=&\mathrm{i}P_{kl00},\quad |l|\le 1,0<|k|\le K,\\(\langle k,\omega\rangle I-A_{[n]})F^{k10}_{[n]}&=&\mathrm{i}P^{k10}_{[n]},\quad |k|\le K,n\in\mathbb{Z}^2_1\setminus\mathcal{L}_2,\\(\langle k,\omega\rangle I+A_{[n]})F^{k01}_{[n]}&=&\mathrm{i}R^{k01}_{[n]},\quad |k|\le K,n\in\mathbb{Z}^2_1\setminus\mathcal{L}_2,\\(\langle k,\omega\rangle I-\mathcal{A}_n)(F^{k10}_n,F^{k01}_m)^T&=&i(P^{k10}_n,P^{k01}_m)^T,|k|\le K,n\in\mathcal{L}_2,\\(\langle k,\omega\rangle I+\mathcal{A}_n)(F^{k01}_n,F^{k10}_m)^T&=&i(P^{k01}_n,P^{k10}_m)^T,|k|\le K,n\in\mathcal{L}_2.\end{array}\tag{4.12}$$

The Fourier coefficients of F_2 are defined by the following Lemmas
Case 1: $n,m\in\mathbb{Z}^2_1\setminus\mathcal{L}_2$

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

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

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*

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

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*

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

$$\mathcal{A}_n=\begin{pmatrix}\Omega_n+\langle l^+,\omega\rangle & -C(l,\alpha)\xi^{\frac{l}{2}+\alpha}\\ C(l,\alpha)\xi^{\frac{l}{2}+\alpha} & -(\Omega_m-\langle l^-,\omega\rangle)\end{pmatrix}, n\in\mathcal{L}_2.$$

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 $\mathcal{L}_2$, we have

$$\begin{aligned}
&\sum_{|k|\le K,n\in\mathcal{L}_2}[\langle k,\omega\rangle-(\Omega_n+\langle l^+,\omega\rangle)]F^{k10}_n z_n e^{\mathrm{i}\langle k,\theta\rangle}+C(l,\alpha)\xi^{\frac{l}{2}+\alpha}F^{k01}_m z_n e^{\mathrm{i}\langle k,\theta\rangle}\\
&=i\sum_{|k|\le K,n\in\mathcal{L}_2}P^{k10}_n z_n e^{\mathrm{i}\langle k,\theta\rangle},\\
&\sum_{|k|\le K,n\in\mathcal{L}_2}[\langle k,\omega\rangle+(\Omega_m-\langle l^-,\omega\rangle)]F^{k01}_m z_m e^{\mathrm{i}\langle k,\theta\rangle}-C(l,\alpha)\xi^{\frac{l}{2}+\alpha}F^{k10}_n z_m e^{\mathrm{i}\langle k,\theta\rangle}\\
&=i\sum_{|k|\le K,n\in\mathcal{L}_2}P^{k01}_m z_m e^{\mathrm{i}\langle k,\theta\rangle}.
\end{aligned}$$

We 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

$$(\langle k,\omega\rangle I+\mathcal{A}_n)(F^{k01}_n,F^{k10}_m)^T=i(P^{k01}_n,P^{k10}_m)^T, |k|\le K, n\in\mathcal{L}_2.$$

If (n,m) and (n',m') are resonant pairs in $\mathcal{L}_2$, comparing the Fourier coefficients, we have that $(F^{k11}_{nn'},F^{k20}_{nm'},F^{k02}_{mn'},F^{k11}_{m'm})^T$ satisfy

$$\begin{aligned}
&[\langle k,\omega\rangle-(\Omega_n+\langle l^+,\omega\rangle)+(\Omega_{n'}+\langle l'^+,\omega\rangle)]F^{k11}_{nn'}e^{\mathrm{i}\langle k,\theta\rangle}-C(l',\alpha)\xi^{\frac{l'}{2}+\alpha}F^{k20}_{nm'}e^{\mathrm{i}\langle k,\theta\rangle}\\
&+C(l,\alpha)\xi^{\frac{l}{2}+\alpha}F^{k02}_{mn'}e^{\mathrm{i}\langle k,\theta\rangle}=iP^{k11}_{nn'}e^{\mathrm{i}\langle k,\theta\rangle},
\end{aligned}$$

similarly,

$$
\begin{aligned}
&[\langle k,\omega\rangle-(\Omega_n+\langle l^+,\omega\rangle)-(\Omega_{m'}-\langle l'^-,\omega\rangle)]F^{k20}_{nm'}e^{\mathrm{i}\langle k,\theta\rangle}\ +C(l',\alpha)\xi^{\frac{l'}{2}+\alpha}F^{k11}_{nn'}e^{\mathrm{i}\langle k,\theta\rangle}\\
&+C(l,\alpha)\xi^{\frac{l}{2}+\alpha}F^{k11}_{m'm}e^{\mathrm{i}\langle k,\theta\rangle}\ =i\,P^{k20}_{nm'}e^{\mathrm{i}\langle k,\theta\rangle}\ ,\\
&[\langle k,\omega\rangle+(\Omega_m-\langle l^-,\omega\rangle)+(\Omega_{n'}+\langle l'^+,\omega\rangle)]F^{k02}_{mn'}e^{\mathrm{i}\langle k,\theta\rangle}\ -C(l',\alpha)\xi^{\frac{l'}{2}+\alpha}F^{k11}_{m'm}e^{\mathrm{i}\langle k,\theta\rangle}\\
&-C(l,\alpha)\xi^{\frac{l}{2}+\alpha}F^{k11}_{nn'}e^{\mathrm{i}\langle k,\theta\rangle}\ =i\,P^{k02}_{mn'}e^{\mathrm{i}\langle k,\theta\rangle}\ ,\\
&[\langle k,\omega\rangle+(\Omega_m-\langle l^-,\omega\rangle)-(\Omega_{m'}-\langle l'^-,\omega\rangle)]F^{k11}_{m'm}e^{\mathrm{i}\langle k,\theta\rangle}\ +C(l',\alpha)\xi^{\frac{l'}{2}+\alpha}F^{k02}_{mn'}e^{\mathrm{i}\langle k,\theta\rangle}\\
&-C(l,\alpha)\xi^{\frac{l}{2}+\alpha}F^{k20}_{nm'}e^{\mathrm{i}\langle k,\theta\rangle}\ =i\,P^{k11}_{m'm}e^{\mathrm{i}\langle k,\theta\rangle}\ .
\end{aligned}
$$

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. □

Remark. In the case that (n,m) and (n',m') are resonant pairs in $\mathcal{L}_2$, we have that k, (n,m), (n',m') satisfy

$$
\begin{aligned}
&\sum_{|k|\le K,n\in\mathcal{L}_2,n'\in\mathcal{L}_2,|k|+|n-n'|\neq 0}(F^{k11}_{n'n}z_{n'}\bar{z}_n+F^{k11}_{nn'}z_n\bar{z}_{n'})e^{\mathrm{i}\langle k,\theta\rangle}\\
&+\sum_{|k|\le K,n\in\mathcal{L}_2,n'\in\mathcal{L}_2,|k|+|n-m'|\neq 0(or)|k|+|n'-m|\neq 0}(F^{k20}_{n'n}z_{n'}z_n+F^{k02}_{n'm'}\bar{z}_{n'}\bar{z}_n)e^{\mathrm{i}\langle k,\theta\rangle}\\
&+\cdots\cdots
\end{aligned}
$$

Consider the equations

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

matrix $Q_{[n]}$ is the $A_{[n]}$'s orthogonal matrix,

$$(\langle k,\omega\rangle I - Q_{[n]}^T A_{[n]} Q_{[n]}) Q_{[n]}^T F_{[n]}^{k10} = i\, Q_{[n]}^T R_{[n]}^{k10}, \, |k| \le K,$$

that is,

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

Similarly, form

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

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

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

Now we focus on the following equations:

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

Consider the equations

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

matrix $Q_{[n]}$ is the $A_{[n]}$'s orthogonal matrix,

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

that is

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

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

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

Consider 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

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

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

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

Case 1. When Jordan normal form

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

Now we focus on the following equations

$$(\langle k,\omega\rangle-\widetilde{\lambda}_j-\mu_1)(\widetilde{F}^{k20}_{[n]n'})_j=i(\widetilde{P}^{k20}_{[n]n'})_j,$$
$$(\langle k,\omega\rangle-\widetilde{\lambda}_j-\mu_2)(\widetilde{F}^{k11}_{[n]m'})_j=i(\widetilde{P}^{k11}_{[n]m'})_j.$$

Case 2. 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 [45].

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 $F^{k20}_{[n][n]}$ 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

$$F^{k20}_{[m][n],ij}=i\sum_{m_1,n_1}\frac{Q_{[m],im_1}\hat{R}^{k20}_{[m][n],m_1,n_1}Q^T_{[n],n_1j}}{\langle k,\omega\rangle-\widetilde{\lambda}_i-\widetilde{\lambda}_j}.$$

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

Then

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

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

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

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

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

And

$$\|F^{k20}_{[m][n]}\| \le cK_\nu^{3\varepsilon}\varepsilon_{\nu+1}\frac{K_\nu^{(4p+1)(\tau+1)}}{\gamma^{4p+1}}K_\nu^{3\varepsilon} \le \varepsilon_{\nu+1}^{\frac{1}{3}},$$

$$\|\begin{pmatrix} F^{k20}_{[n]n'} \\ F^{k11}_{[n]m'} \end{pmatrix}\| \le cK_\nu^{3\varepsilon}\varepsilon_{\nu+1}\frac{K_\nu^{(4p+1)(\tau+1)}}{\gamma^{4p+1}} \le \varepsilon_{\nu+1}^{\frac{1}{3}},$$

under the assumption

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

4.3. Estimate on the coordinate transformation

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

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

$$\|X_F\|_{D_3,\mathcal{O}} \le c\gamma^{-(4p+1)}K^{(4p+1)(\tau+1)}(r - r_+)^{-c}\varepsilon. \tag{4.15}$$

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

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

Moreover,

$$\|D\phi_F^t - Id\|_{D_{1\eta}} < c\gamma^{-(4p+1)}K^{(4p+1)(\tau+1)}(r-r_+)^{-c}\varepsilon. \tag{4.17}$$

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

$$\|D^m F\|_{D_2,\mathcal{O}} < c\gamma^{-(4p+1)}K^{(4p+1)(\tau+1)}(r-r_+)^{-c}\varepsilon, \tag{4.18}$$

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

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

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

$$\|D\phi_F^t - Id\| \leq 2\|D^2 F\| < c\gamma^{-(4p+1)}K^{(4p+1)(\tau+1)}(r-r_+)^{-c}\varepsilon. \tag{4.19}$$

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_+

$$\begin{aligned} N_+ &= N + P_{0000} + \langle \hat{\omega}, I\rangle + \sum_{[n]} \langle P^{011}_{[n][n]} z_{[n]}, \bar{z}_{[n]}\rangle + \sum_{n\in\mathcal{L}_2} (P^{011}_{nn} z_n \bar{z}_n + P^{011}_{mm} z_m \bar{z}_m) \\ &= \langle \omega_+, I\rangle + \sum_{[n]} \langle A^+_{[n]} z_{[n]}, \bar{z}_{[n]}\rangle + \sum_{n\in\mathcal{L}_2} [(\Omega^+_n + \langle l^+, \omega\rangle) z_n \bar{z}_n + (\Omega^+_m - \langle l^-, \omega\rangle) z_m \bar{z}_m)] \end{aligned}$$

where

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

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

$$\Omega_n^+ = \Omega_n + P_{nn}^{011}, \Omega_m^+ = \Omega_m + P_{mm}^{011}, n \in \mathcal{L}_2.$$

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

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

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

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

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

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:

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

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

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

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,

$$\|D\phi_F^t - Id\|_{D_{1\eta}} < c\gamma^{-(4p+1)} K^{(4p+1)(\tau+1)} (r - r_+)^{-c} \varepsilon, \quad -1 \le t \le 1,$$

thus

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

Due to Lemma 7.3 in [44],

$$\|X_{\{R(t),F\}}\|_{D_{2\eta}} \le c\gamma^{-(4p+1)} K^{(4p+1)(\tau+1)} (r - r_+)^{-c} \eta^{-2} \varepsilon^2,$$

and

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

we have

$$\|X_{P_+}\|_{D_\rho(r_+,s_+)} \le c\eta\varepsilon + c\gamma^{-(4p+1)} K^{(4p+1)(\tau+1)} (r - r_+)^{-c} \eta^{-2} \varepsilon^2 \le c\varepsilon_+.$$

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

Since

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

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

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

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 [25].

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 [25].

A KAM-step cycle is now completed.

5. Iteration lemma and convergence

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

$$\begin{gathered} r_{\nu+1} = r(1 - \sum_{i=2}^{\nu+2} 2^{-i}), \\ \varepsilon_{\nu+1} = c\gamma^{-(4p+1)}(r_\nu - r_{\nu+1})^{-c} K_\nu^{(4p+1)(\tau+1)} \varepsilon_\nu^{\frac{4}{3}}, \\ \eta_{\nu+1} = \varepsilon_{\nu+1}^{\frac{1}{3}}, L_{\nu+1} = L_\nu + \varepsilon_\nu \\ s_{\nu+1} = 2^{-2}\eta_\nu s_\nu = 2^{-2(\nu+1)}(\prod_{i=0}^{\nu} \varepsilon_i)^{\frac{1}{3}} s_0, \\ K_{\nu+1}^{1+3\varepsilon} \rho_{\nu+1} = 1 \\ K_{\nu+1} = 3K_\nu = 3^{\nu+1} K_0 \\ \Delta_{\nu+1} = K_\nu^3 \end{gathered} \tag{5.1}$$

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

$$\begin{gathered} |\langle k, \omega_\nu \rangle| \geq \frac{\gamma}{K_\nu^\tau}, k \neq 0, \\ |\langle k, \omega_\nu \rangle \pm \widetilde{\lambda}_j^\nu| \geq \frac{\gamma}{K_\nu^\tau}, j \in [n], \\ |\langle k, \omega_\nu \rangle \pm \widetilde{\lambda}_i^\nu \pm \widetilde{\lambda}_j^\nu| \geq \frac{\gamma}{K_\nu^\tau}, i \in [m], j \in [n], \end{gathered}$$

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

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

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

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^{4p} smooth and satisfy

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

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

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

and

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

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

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

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

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

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

where

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

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

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

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

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

with $\omega_{\nu+1} = \omega_\nu + P^{\nu}_{0l00}$, 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

$$\begin{aligned} H_{\nu+1} &= e_{\nu+1} + \langle \omega_{\nu+1}, I\rangle + \sum_{[n]} \langle A^{\nu+1}_{[n]}(\xi) z_{[n]}, \bar{z}_{[n]}\rangle \\ &\quad + \sum_{n\in\mathcal{L}_2} [(\Omega^{\nu+1}_n + \langle l^+, \omega\rangle) z_n \bar{z}_n + (\Omega^{\nu+1}_m - \langle l^-, \omega\rangle) z_m \bar{z}_m)] + \mathcal{B}_{\nu+1} + \bar{\mathcal{B}}_{\nu+1} + P_{\nu+1}, \end{aligned}$$

with

$$\begin{aligned} \sup_{\xi\in\mathcal{O}_\nu} \max_{d\le 4p} \|\partial^d_\xi (A^{\nu+1}_{[n]} - A^{\nu}_{[n]})\| &\le c\varepsilon_\nu, \\ \sup_{\xi\in\mathcal{O}_\nu} \max_{d\le 4p} |\partial^d_\xi (\Omega^{\nu+1}_{n'} - \Omega^{\nu}_{n'})| &\le \varepsilon_\nu, \\ \sup_{\xi\in\mathcal{O}_\nu} \max_{d\le 4p} |\partial^d_\xi (\omega_{\nu+1} - \omega_\nu)| &\le \varepsilon_\nu, \end{aligned}$$

and

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

in the sense of Whitney. And

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

5.2. Convergence

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

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

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

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:

$$\mathcal{O}_{\nu+1} \subset \mathcal{O}_\nu,$$

$$\Psi^\nu = \Phi_0 \circ \Phi_1 \circ \cdots \circ \Phi_\nu : D_{\rho_\nu}(r_{\nu+1}, s_{\nu+1}) \times \mathcal{O}_\nu \to D_{\rho_0}(r_0, s_0),\ \nu \geq 0,$$

$$H \circ \Psi^\nu = H_{\nu+1} = N_{\nu+1} + \mathcal{B}_{\nu+1} + \bar{\mathcal{B}}_{\nu+1} + P_{\nu+1}.$$

Let $\tilde{\mathcal{O}} = \cap_{\nu=0}^\infty \mathcal{O}_\nu$. As in [35,36], thanks to Lemma 4.7, it concludes that N_ν, Ψ^ν, $D\Psi^\nu$, ω_ν converge uniformly on $D_{\frac{1}{2}r}(\frac{1}{2}r, 0) \times \tilde{\mathcal{O}}$ with

$$N_\infty + \mathcal{B}_\infty + \bar{\mathcal{B}}_\infty = e_\infty + \langle \omega_\infty, I \rangle + \sum_{[n]} \langle A_{[n]}^\infty(\xi) z_{[n]}, \bar{z}_{[n]} \rangle$$
$$+ \sum_{n \in \mathcal{L}_2} [(\Omega_n^\infty + \langle l^+, \omega \rangle) z_n \bar{z}_n + (\Omega_m^\infty - \langle l^-, \omega \rangle) z_m \bar{z}_m)] + \mathcal{B}_\infty + \bar{\mathcal{B}}_\infty.$$

Since

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

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 Ψ^ν, $D\Psi^\nu$, ω_ν 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

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

where

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

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

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

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

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

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

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

$$|\Omega_n^{\nu+1}(\xi) - \Omega_n(\xi)|_{\mathcal{O}_\nu} \le \sum_{j=0}^{\nu} |R_{nn}^{011,j}| \le \varepsilon.$$

Remark. From the section 4.4, one has that at $(\nu+1)^{\text{th}}$ step, small divisor conditions are automatically satisfied for $|k| \le K_\nu$. 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 $\{\xi \in \mathcal{O}_\nu : |\langle k, \omega_{\nu+1} \rangle + \widetilde{\lambda}_n^{\nu+1} - \widetilde{\lambda}_{n'}^{\nu+1}| < \frac{\gamma}{K_{\nu+1}^\tau}, n, n' \in \mathcal{L}_1\}$ and $\{\xi \in \mathcal{O}_\nu : |det(\langle k, \omega_{\nu+1} \rangle I + \mathcal{A}_n^{\nu+1} \otimes I_2 - I_2 \otimes \mathcal{A}_{n'}^{\nu+1})| < \frac{\gamma}{K_{\nu+1}^\tau}, n, n' \in \mathcal{L}_2\}$. In other cases, the proof is similar, so we omit it. For simplicity, set $M^{\nu+1} = |\langle k, \omega_{\nu+1} \rangle + \widetilde{\lambda}_n^{\nu+1} - \widetilde{\lambda}_{n'}^{\nu+1}|$ and $Y^{\nu+1} = \langle k, \omega_{\nu+1} \rangle I + \mathcal{A}_n^{\nu+1} \otimes I_2 - I_2 \otimes \mathcal{A}_{n'}^{\nu+1}$, $Y^\nu = \langle k, \omega_\nu \rangle I + \mathcal{A}_n^\nu \otimes I_2 - I_2 \otimes \mathcal{A}_{n'}^\nu$, then for $|k| \le K_\nu$

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

Lemma 6.1. *For any given $n, n' \in \mathbb{Z}_1^2$ with $|n - n'| \le K_{\nu+1}$, either $|\langle k, \omega_{\nu+1}\rangle + \widetilde{\lambda}_n^{\nu+1} - \widetilde{\lambda}_{n'}^{\nu+1}| > 1$ or there are $n_0, n_0', c \in \mathbb{Z}^2$ with $|n_0|, |n_0'|, |c| \le 3K_{\nu+1}^2$ and $t \in \mathbb{Z}$, such that $n = n_0 + tc$, $n' = n_0' + tc$.*

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

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

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

In the case that $|\langle n - n', n'\rangle| \le K_{\nu+1}^2$, clearly $n - n' = 0$ is trivial. Assume $n - n' \ne 0$, 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 $|c| \le |n - n'| \le K_{\nu+1}$. 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 $\cdot$)

$$t = [x_1],$$

then $t \in \mathbb{Z}$ and $|n' - tc| \le 2K_{\nu+1}^2$. Take $n_0' = n' - tc \in \mathbb{Z}^2$ and $n_0 = n_0' + n - n' \in \mathbb{Z}^2$. We have $|n_0'| \le 2K_{\nu+1}^2$ and

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

Lemma 6.2.

$$\cup_{n,n' \in \mathcal{L}_1} \mathcal{R}_{k[n][n']}^{\nu+1} \subset \cup_{n_0, n_0', c \in \mathbb{Z}^2, t \in \mathbb{Z}} \mathcal{R}_{k, n_0 + tc, n_0' + tc}^{\nu+1}$$

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

Proof. If $|\langle n - n', n'\rangle| > K_{\nu+1}^2$, $\mathcal{R}_{k[n][n']}^{\nu+1} = \emptyset$. If $|\langle n - n', n'\rangle| \le K_{\nu+1}^2$, there exist $n_0, n_0', c \in \mathbb{Z}^2, t \in \mathbb{Z}$ with $|n_0|, |n_0'|, |c| \le 3K_{\nu+1}^2$ such that $n = n_0 + tc$, $n' = n_0' + tc$. Hence

$$\cup_{n,n' \in \mathcal{L}_1} \mathcal{R}_{k[n][n']}^{\nu+1} \subset \cup_{n_0, n_0', c \in \mathbb{Z}^2, t \in \mathbb{Z}} \mathcal{R}_{k, n_0 + tc, n_0' + tc}^{\nu+1},$$

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

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

$$\operatorname{meas}(\cup_{t\in\mathbb{Z}}\mathcal{R}^{\nu+1}_{k,n_0+tc,n_0'+tc}) < c\frac{\gamma^{\frac{1}{p}}}{K_{\nu+1}^{\frac{\tau}{p(p+1)}}}.$$

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

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

We define resonant set

$$\mathcal{R}^{\nu+1}_{kn_0n_0'c\infty} = \{\xi \in \mathcal{O}_\nu : |\lim_{t\to\infty} M^{\nu+1}(t))| < \frac{\gamma}{K_{\nu+1}^{\frac{\tau}{p+1}}}\} \tag{6.1}$$

For fixed k, n_0, n_0', c,

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

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

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

Case 1: When $|t| > K_{\nu+1}^{\frac{\tau}{p+1}}$, for $\xi \in \mathcal{O}_\nu \backslash \mathcal{R}^{\nu+1}_{kn_0n_0'c\infty}$, we have

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

Case 2: When $|t_1| \leq K_{\nu+1}^{\frac{\tau}{p+1}}$, we define resonant set

$$\mathcal{R}^{\nu+1}_{kn_0n_0'ct} = \{\xi \in \mathcal{O}_\nu : |M^{\nu+1}(t)| < \frac{\gamma}{K_{\nu+1}^{\tau}}\}. \tag{6.2}$$

For fixed k, n_0, n_0', c, t,

$$\mathrm{meas}(\mathcal{R}^{\nu+1}_{kn_0n_0'ct}) < (\frac{\gamma}{K^{\tau}_{\nu+1}})^{\frac{1}{p}},$$

then

$$\mathrm{meas}\{\cup_{|t|\le K_{\nu+1}^{\frac{\tau}{p+1}}} \mathcal{R}^{\nu+1}_{kn_0n_0'ct}\} < K_{\nu+1}^{\frac{\tau}{p+1}} (\frac{\gamma}{K^{\tau}_{\nu+1}})^{\frac{1}{p}} \le \frac{\gamma^{\frac{1}{p}}}{K_{\nu+1}^{\frac{\tau}{p(p+1)}}}.$$

As a consequence,

$$\mathrm{meas}(\cup_{t\in\mathbb{Z}} \mathcal{R}^{\nu+1}_{k,n_0+tc,n_0'+tc}) < c\frac{\gamma^{\frac{1}{p}}}{K_{\nu+1}^{\frac{\tau}{p(p+1)}}}. \quad \square$$

For $K_\nu < |k| \le K_{\nu+1}$, 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 $|n - n'| \le K_{\nu+1}$, 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 $|n_0|, |n_0'|, |c| \le 3K_{\nu+1}^2$ and $t \in \mathbb{Z}$, such that $n = n_0 + tc$, $n' = n_0' + tc$.*

Lemma 6.5.

$$\cup_{n,n'\in\mathbb{Z}_1^2} \mathcal{C}^{\nu+1}_{knn'} \subset \cup_{n_0,n_0',c\in\mathbb{Z}^2,t\in\mathbb{Z}} \mathcal{C}^{\nu+1}_{k,n_0+tc,n_0'+tc}$$

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

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

$$\mathrm{meas}(\cup_{t\in\mathbb{Z}} \mathcal{C}^{\nu+1}_{k,n_0+tc,n_0'+tc}) < c\frac{\gamma^{\frac{1}{4p}}}{K_{\nu+1}^{\frac{\tau}{4p(4p+1)}}}.$$

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

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

We define resonant set

$$\mathcal{C}^{\nu+1}_{kn_0n_0'c\infty} = \left\{ \xi \in \mathcal{O}_\nu : |det \lim_{t\to\infty} Y^{\nu+1}(t)| < \frac{\gamma}{K_{\nu+1}^{\frac{\tau}{4p+1}}} \right\}.$$

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

$$\|(\lim_{t\to\infty} Y^{\nu+1}(t))^{-1}\| \le \frac{K_{\nu+1}^{\frac{\tau}{4p+1}}}{\gamma}.$$

Since

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

for $|t| > K_{\nu+1}^{\frac{\tau}{4p+1}}$, we have

$$\|(Y^{\nu+1}(t))^{-1}\| \le 2\frac{K_{\nu+1}^{\frac{\tau}{4p+1}}}{\gamma} \le \frac{K_{\nu+1}^{\tau}}{\gamma}.$$

For $|t| \le K_{\nu+1}^{\frac{\tau}{4p+1}}$, we define resonant set

$$\mathcal{C}^{\nu+1}_{kn_0n_0'ct} = \{\xi \in \mathcal{O}_\nu : |detY^{\nu+1}(t)| < \frac{\gamma}{K_{\nu+1}^{\tau}}\}. \tag{6.3}$$

In addition

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

For fixed k, n_0, n_0', c, t,

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

then

$$\text{meas}\{\cup_{|t|\le K_{\nu+1}^{\frac{\tau}{4p+1}}} \mathcal{C}^{\nu+1}_{kn_0n_0'ct}\} < K_{\nu+1}^{\frac{\tau}{4p+1}} (\frac{\gamma}{K_{\nu+1}^{\tau}})^{\frac{1}{4p}} \le \frac{\gamma^{\frac{1}{4p}}}{K_{\nu+1}^{\frac{\tau}{4p(4p+1)}}}.$$

As a consequence,

$$\text{meas}(\cup_{t\in\mathbb{Z}} \mathcal{C}^{\nu+1}_{k,n_0+tc,n_0'+tc}) < c\frac{\gamma^{\frac{1}{4p}}}{K_{\nu+1}^{\frac{\tau}{4p(4p+1)}}}. \quad \square$$

Lemma 6.7.

$$\text{meas}(\bigcup_{K_\nu<|k|\le K_{\nu+1}} \mathcal{R}_k^{\nu+1}) \le cK_{\nu+1}^b \frac{\gamma^{\frac{1}{p}}}{K_{\nu+1}^{\frac{\tau}{p(p+1)}}} = c\frac{\gamma^{\frac{1}{p}}}{K_{\nu+1}^{\frac{\tau}{p(p+1)}-b}},$$

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

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

$$\mathrm{meas}(\bigcup_{K_\nu<|k|\le K_{\nu+1},[n],n'}\mathcal{R}^\nu_{k[n]n'})\le c\frac{\gamma^{\frac{1}{p}}}{K_{\nu+1}^{\frac{\tau}{p(p+1)}-b-2}},$$

$$\mathrm{meas}(\bigcup_{K_\nu<|k|\le K_{\nu+1},n,n'}\mathcal{C}^\nu_{knn'})\le c\frac{\gamma^{\frac{1}{4p}}}{K_{\nu+1}^{\frac{\tau}{4p(4p+1)}-b-12}}.$$

Lemma 6.8. *Let $\tau>4p(4p+1)(12+b+1)$, then the total measure need to exclude along the KAM iteration is*

$$\begin{aligned}
&\mathrm{meas}(\bigcup_{\nu\ge0}\mathcal{R}^{\nu+1})\\
&=\mathrm{meas}[\bigcup_{\nu\ge0}\bigcup_{K_\nu<|k|\le K_{\nu+1},[n],[m],n,n'}(\mathcal{R}^{\nu+1}_k\bigcup\mathcal{R}^{\nu+1}_{k[n]}\bigcup\mathcal{R}^{\nu+1}_{k[n][m]}\bigcup\mathcal{R}^{\nu+1}_{k[n]n'}(\gamma)\bigcup\mathcal{C}^{\nu+1}_{knn'}(\gamma))]\\
&\le c\sum_{\nu\ge0}\frac{\gamma^{\frac{1}{4p}}}{K_{\nu+1}}\le c\gamma^{\frac{1}{4p}}.
\end{aligned}$$

Data availability

No data was used for the research described in the article.

Acknowledgments

We are grateful to the anonymous referees for useful comments and suggestions. Those comments and suggestions are all valuable and very helpful for revising and improving our paper, as well as the important guiding significance to our researches.

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, 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.

[6] M. Berti, M. Procesi, Nonlinear wave and Schrödinger equations on compact Lie groups and homogeneous spaces, Duke Math. J. 159 (2011) 479–538.
[7] M. Berti, P. Bolle, Sobolev quasi-periodic solutions of multidimensional wave equations with a multiplicative potential, Nonlinearity 25 (2012) 2579–2613.
[8] 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.
[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, A. Maspero, Long time dynamics of Schrödinger and wave equations on flat tori, J. Differ. Equ. 267 (2019) 1167–1200.
[11] J. Bourgain, Quasiperiodic solutions of Hamiltonian perturbations of 2D linear Schrödinger equations, Ann. Math. 148 (1998) 363–439.
[12] J. Bourgain, Construction of periodic solutions of nonlinear wave equations in higher dimension, Geom. Funct. Anal. 5 (1995) 629–639.
[13] 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.
[14] J. Bourgain, Green's function estimates for lattice Schrödinger operators and applications, Annals of Mathematics Studies, vol. 158, Princeton University Press, Princeton, NJ, 2005.
[15] J. Bourgain, Nonlinear Schrödinger equations, Park city series, vol. 5, American Mathematical Society(Providence, Rhode Island, 1999.
[16] J. Bourgain, On diffusion in high–dimensional Hamiltonian systems and PDE, J. Anal. Math. 80 (2000) 1–35.
[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, S. Xue, Reducible KAM tori for two-dimensional nonlinear Schrödinger equations with explicit dependence on the spatial variable, J. Funct. Anal. 282 (10) (2022) 109430.
[25] 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.
[26] J. Geng, J. You, A KAM Theorem for one dimensional Schrödinger equation with periodic boundary conditions, J. Diff. Equ. 209 (2005) 1–56.
[27] J. Geng, J. You, A KAM theorem for Hamiltonian partial differential equations in higher dimensional spaces, Commun. Math. Phys. 262 (2006) 343–372.
[28] J. Geng, J. You, KAM tori for higher dimensional beam equations with constant potentials, Nonlinearity 19 (2006) 2405–2423.
[29] J. Geng, Y. Yi, Quasi-periodic solutions in a nonlinear Schrödinger equation, J. Diff. Equ. 233 (2007) 512–542.
[30] J. Geng, J. You, A KAM theorem for higher dimensional nonlinear Schrodinger equation, J. Dyn. Differ. Equ. 25 (2013) 451–476.
[31] T. Kappeler, J. Pöschel, KdV & KAM, Springer, 2003.
[32] S.B. Kuksin, Hamiltonian perturbations of infinite–dimensional linear systems with an imaginary spectrum, Funct. Anal. Appl. 21 (1987) 192–205.
[33] S.B. Kuksin, Nearly integrable infinite dimensional Hamiltonian systems, Lecture Notes in Mathematics, vol. 1556, Springer, Berlin, 1993.
[34] S.B. Kuksin, J. Pöschel, Invariant Cantor manifolds of quasiperiodic oscillations for a nonlinear Schrödinger equation, Ann. Math. 143 (1996) 149–179.
[35] J. Pöschel, Quasi-periodic solutions for a nonlinear wave equation, Comment. Math. Helv. 71 (1996) 269–296.
[36] J. Pöschel, A KAM theorem for some nonlinear partial differential equations, Ann. Sc. Norm. Super. Pisa, Cl. Sci. 23 (1996) 119–148.

[37] J. Pöschel, On the construction of almost periodic solutions for a nonlinear Schrödinger equations, Ergod. Theory Dyn. Syst. 22 (2002) 1537–1549.
[38] M. Procesi, C. Procesi, A normal form for the Schrödinger equation with analytic nonlinearities, Commun. Math. Phys. 312 (2012) 501–557.
[39] C. Procesi, M. Procesi, A KAM Algorithm for the Resonant Nonlinear Schrödinger Equation, Adv. Math. 272 (2015) 399–470.
[40] W.–M. Wang, Energy supercritical nonlinear Schrödinger equations: quasiperiodic solutions, Duke Math. J. 165 (2016) 1129–1192.
[41] C.E. Wayne, Periodic and quasi-periodic solutions for nonlinear wave equations via KAM theory, Commun. Math. Phys. 127 (1990) 479–528.
[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. Xu, J. You, Persistence of lower-dimensional tori under the first Melnikov's non-resonance condition, J. Math. Pures Appl. 80 (2001) 1045–1067.
[45] J. You, Perturbations of lower dimensional tori for Hamiltonian systems, J. Differ. Equ. 152 (1999) 1–29.
[46] X. Yuan, Quasi-periodic solutions of completely resonant nonlinear wave equations, J. Diff. Equ. 230 (2006) 213–274.
[47] M. Zhang, J. Si, Construction of quasi-periodic solutions for the quintic schrodinger equation on the two-dimensional torus t-2, Trans. Am. Math. Soc. 374 (2021) 4711–4780.