Available online at www.sciencedirect.com
ScienceDirect
J. Differential Equations 244 (2008) 2185–2225
Journal of Differential Equations
www.elsevier.com/locate/jde

Quasi-periodic solutions for 1D Schrödinger equation with the nonlinearity $|u|^{2p}u$ ☆

Zhenguo Liang

School of Mathematical Sciences and Key Lab of Mathematics for Nonlinear Science, Fudan University, Shanghai 200433, China

Received 22 July 2006; revised 24 January 2008

Available online 19 March 2008

Abstract

In this paper, one-dimensional (1D) nonlinear Schrödinger equation

$$iu_t - u_{xx} + |u|^{2p}u = 0, \quad p \in \mathbb{N},$$

with periodic boundary conditions is considered. It is proved that the above equation admits small-amplitude quasi-periodic solutions corresponding to 2-dimensional invariant tori of an associated infinite-dimensional dynamical system. The proof is based on infinite-dimensional KAM theory, partial normal form and scaling skills.

MSC: 37K55; 35B10; 35J10; 35Q40; 35Q55

Keywords: Schrödinger equation; Hamiltonian systems; KAM theory; Normal form; Quasi-periodic solution

1. Introduction and main result

In this paper, we will prove that one-dimensional (1D) nonlinear Schrödinger equation

$$iu_t - u_{xx} + |u|^{2p}u = 0 \tag{1.1}$$

under *periodic boundary conditions*

☆ Partially supported by the National Natural Science Foundation of China (Grant No. 10701027).
E-mail addresses: zgliang@fudan.edu.cn, lzgmaths@21cn.com.

0022-0396/$ – see front matter
doi:10.1016/j.jde.2008.02.015

$$u(t,x) = u(t, x + 2\pi) \tag{1.2}$$

admits small-amplitude quasi-periodic solutions corresponding to 2-dimensional invariant tori.

As usual, we study Eq. (1.1) as a Hamiltonian system on $\mathcal{P} = H_0^1(\mathbb{T}) = H_0^1([0, 2\pi])$ with the inner product $(u, v) = Re \int_0^{2\pi} u\bar{v}\,dx$, the Sobolev space of all complex-valued L^2-functions on $\mathbb{T}$ with an L^2-derivative. Let $\phi_j(x) = \sqrt{\frac{1}{2\pi}} e^{ijx}$, $\lambda_j = j^2$, $j \in \mathbb{Z}$, be the basic modes and their frequencies for the linear equation $iu_t = u_{xx}$ with periodic boundary conditions. Then every solution is the superposition of oscillations of the basic modes, with the coefficients moving on circles,

$$u(t,x) = \sum_{j \in \mathbb{Z}} q_j(t)\phi_j(x), \quad q_j(t) = q_j^0 e^{i\lambda_j t}.$$

Together they move on a rotational torus of finite or infinite dimension, depending on how many modes are excited. In particular, for every choice

$$\mathcal{J} = \{j_1 < j_2\} \subset \mathbb{Z}$$

of 2 basic modes there is an invariant linear space $E_{\mathcal{J}}$ of complex dimension 2 which is completely foliated into rotational tori:

$$E_{\mathcal{J}} = \left\{u = q_1\phi_{j_1} + q_2\phi_{j_2}\colon\ q \in C^2\right\} = \bigcup_{I \in \overline{P^2}} \mathcal{T}_{\mathcal{J}}(I),$$

where $P^2 = \{I\colon\ I_j > 0\}$ and

$$\mathcal{T}_{\mathcal{J}}(I) = \left\{u = q_1\phi_{j_1} + q_2\phi_{j_2}\colon\ |q_j|^2 = 2I_j \text{ for } 1 \leqslant j \leqslant 2\right\}.$$

In addition, each such torus is linearly stable, and all solutions have vanishing Lyapunov exponents. This is the linear situation.

Upon restoration of the nonlinearity $|u|^{2p}u$, we show that there exist a Cantor set $\mathcal{C} \subset P^2$, an index set $\mathcal{I} = \{n_1 < n_2\}$, where $n_2 > \sqrt{p}n_1 > 0$, and a family of 2-tori

$$\mathcal{T}_{\mathcal{I}}[\mathcal{C}] = \bigcup_{I \in \mathcal{C}} \mathcal{T}_{\mathcal{I}}(I) \subset E_{\mathcal{I}}$$

over $\mathcal{C}$, and a Whitney smooth embedding

$$\Phi : \mathcal{T}_{\mathcal{I}}[\mathcal{C}] \hookrightarrow \mathcal{P},$$

such that the restriction of Φ to each $\mathcal{T}_{\mathcal{I}}(I)$ in the family is an embedding of a rotational 2-torus for the nonlinear equation. In [12], the image $\mathcal{E}_{\mathcal{I}}$ of $\mathcal{T}_{\mathcal{I}}[\mathcal{C}]$ is called a Cantor manifold of rotational 2-tori given by the embedding $\Phi : \mathcal{T}_{\mathcal{I}}[\mathcal{C}] \to \mathcal{E}_{\mathcal{I}}$.

Theorem 1 *(Main Theorem). Consider* 1D *nonlinear Schrödinger equation* (1.1) *with* (1.2). *Then for any index set $\mathcal{I} = \{n_1 < n_2\}$, which satisfies $n_2 > \sqrt{p}n_1 > 0$, there exists a positive-measure Cantor manifold $\mathcal{E}_{\mathcal{I}}$ of real analytic, linearly stable, Diophantine* 2*-tori for the nonlinear Schrödinger equation given by a Whitney smooth embedding $\Phi : \mathcal{T}_{\mathcal{I}}[\mathcal{C}] \to \mathcal{E}_{\mathcal{I}}$.*

Remark 1.1. For 1D nonlinear Schrödinger equations of higher order nonlinearities such as

$$iv_t - v_{xx} + mv + |v|^{2p} v = 0 \tag{1.3}$$

under periodic boundary conditions

$$v(t, x) = v(t, x + 2\pi), \tag{1.4}$$

there exists a well-known transformation $v = e^{imt}u$, the above equation and boundary condition are transformed to Eqs. (1.1) and (1.2).

Remark 1.2. Generally, one cannot prove that Φ is a higher order perturbation of the inclusion map $\Phi_0 : E_{\mathcal{I}} \hookrightarrow \mathcal{P}$ restricted to $\mathcal{T}_{\mathcal{I}}[\mathcal{C}]$. The reason lies in the symplectic transformations Ψ_1, Ψ_2. See Section 2 for details.

There are some known works about Eq. (1.1). For $p = 1$ under *Dirichlet boundary conditions*, see the well-known work of Kuksin and Pöschel [12]. For $p = 2$ under *Dirichlet boundary conditions*, Liang and You (see [13]) also got the similar conclusions as [12]. For the Schrödinger equation under *periodic boundary conditions*, Bourgain obtained the existences of quasi-periodic solutions for the Schrödinger equation including 1D and nD ($n \geqslant 2$) in [2,4,5]. His method, called Craig–Wayne–Bourgain's scheme (see [2–5,7]) is very powerful and different with KAM. It avoids the, sometimes, cumbersome and famous "the second Melnikov conditions" but to a high cost: the approximate linear equations are not of constant coefficients. It results in giving no information on the linear stability of constructed quasi-periodic solutions.

The first work using KAM to construct quasi-periodic solutions of 1D nonlinear PDEs under *periodic boundary conditions* is due to Chierchia and You (see [6]). They obtain the linearly stable quasi-periodic solutions for 1D wave equation. For the Schrödinger equation (1.3) + (1.4) when $p = 1$, it was included in the work of Geng and You [9]. Combing with the methods of [13] and [10], Geng and Yi (see [11]) obtained the similar result for $p = 2$. But all known methods are failed in $p \geqslant 3$.

In the following, we will give a heuristic discussion about our method which works for any p. Our discussion will be confined in 1D Schrödinger equation (1.1) + (1.2). As before, the KAM method for this equation is detached into two steps.

The first is to use some symplectic transformations to the original Hamiltonian. This is the familiar normal form step. When $p \geqslant 3$ or $p = 2$ and $n_2 - n_1 \in 2\mathbb{N}$, this step is more difficult than before. The reason lies in that there exist many terms, which we cannot kill. For the common views, the ones must be killed since they cannot be put into the higher order terms. Therefore, one has to remain them. But we must know what they are. In fact, after some subtle analysis, we can write out all the terms which cannot be killed in the normal form. Except that, we note that the remained are highly symmetric. This is also very important for the following transformations.

More clearly, after introducing the parameters ξ_1, ξ_2, we have the Hamiltonian

$$H = \langle \omega(\xi), y \rangle + \langle G(x)w, \bar{w} \rangle + \text{h.o.t.},$$

where the infinite-dimensional normal matrix

$$G(x)=\begin{pmatrix}\ddots &&&&&&&& \\ & \tilde{\Omega}_{i_0} &&&&&& \tilde{a}_0 e^{-ip(x_1-x_2)} & \\ && \cdots &&&& \cdots && \\ &&& \tilde{\Omega}_{i_{p-2}} && \tilde{a}_{p-2}e^{-2i(x_1-x_2)} &&& \\ &&&& \ddots &&&& \\ &&& \tilde{a}_{p-2}e^{2i(x_1-x_2)} && \tilde{\Omega}_{j_{p-2}} &&& \\ && \cdots &&&& \cdots && \\ & \tilde{a}_0 e^{ip(x_1-x_2)} &&&&&& \tilde{\Omega}_{j_0} & \\ &&&&&&&& \ddots\end{pmatrix},$$

where x_1, x_2 is the angle coordinates and for $\tilde{a}_t$, i_t and j_t, see (2.12), (2.7) and (2.8). Note the symmetry of the normal matrix, we introduce a nonlinear symplectic transformation (see Lemma 2.5) to the above Hamiltonian and re-scale the coordinates and parameters including t and then use another symplectic transformation to diagonalize the normal infinite-dimensional matrix. After all the transformations, one gets the following Hamiltonian

$$H=\langle\omega, y\rangle+\sum_j \tilde{\Omega}_j w_n\bar{w}_n+\text{h.o.t.}$$

This is a standard form for our applying the infinite KAM theorem.

The second step is to estimate the thrown measure. In order to obtain the measure estimates under *periodic boundary conditions*, an easy way is to prove that the perturbation terms always satisfy some special properties. We remark that even though the properties as [10] and [11] do not pertain after the nonlinear symplectic transformation Ψ_1 (see Lemma 2.5), a similar property still holds, which we call generalized compact form. One easily proves that this property holds even after infinite *KAM* steps. The remained difficulty is the measure estimate in the first step, while measure estimates of the remaining steps are standard as [13] and [14]. It is hard to get the so-called twist inequality such as

$$\left|\frac{\partial^{2p} f}{\partial \xi^{2p}}\right| \geqslant c>0,$$

where $f=\langle k,\omega\rangle+\tilde{\Omega}_n-\tilde{\Omega}_m$. It needs some complex computations.

Before ending this section, we give a brief introduction in the recent development of infinite-dimensional KAM theory of higher spatial dimensions. In [10], Geng and You proved the existence and the linear stability of KAM tori for some semilinear beam equation and some non-local Schrödinger equations. Very recently, Eliasson and Kuksin have derived the exciting results of both the existence and the linear stability of KAM tori for nD nonlinear Schrödinger equations in [8]. In [20], Yuan proved that there exist many invariant tori and thus quasi-periodic solutions for nonlinear wave equations, Schrödinger equations and other equations of any spatial dimension. The second Melnikov's conditions are totally eliminated in his method. But till now, the linear stability of the obtained quasi-periodic solutions and invariant tori for NLW equations is still an open problem.

The rest of the paper is organized as follows: In Section 2 the Hamiltonian function is written in infinitely many coordinates, which is then put into partial normal form. In Section 3, we give

KAM steps and Theorem 2. Measure estimates are given in Section 4. In Appendix A, we explain what are the compact form and generalized compact form. Some important lemmas are proved there.

2. Normal form

Using the Hamiltonian formulation, we rewrite Eq. (1.1) with the periodic boundary condition (1.2) as the Hamiltonian system $u_t = \mathrm{i}\frac{\partial H}{\partial \bar{u}}$, where

$$H = \int_0^{2\pi} \left(|u_x|^2\right) dx + \frac{1}{p+1} \int_0^{2\pi} |u|^{2p+2}\, dx.$$

Note that the operator $A = -\partial_{xx}$ with the periodic boundary conditions has an orthonormal basis $\{\phi_n(x) = \sqrt{\frac{1}{2\pi}} e^{\mathrm{i}nx}\}$ and corresponding eigenvalues $\mu_n = n^2$. Let $u(x,t) = \sum_{n\in\mathbb{Z}} q_n(t)\phi_n(x)$. The coordinates are taken from the Hilbert spaces l^ρ of all complex-valued sequences $q = (q_i)_{i\in\mathbb{Z}}$ with

$$\|q\|_\rho^2 = \sum_{j\in\mathbb{Z}} |q_j|^2 e^{2|j|\rho} < \infty.$$

Fix $\rho > 0$ later. Then associated with the symplectic structure $\mathrm{i}\sum_{n\in\mathbb{Z}} dq_n \wedge d\bar{q}_n$, $\{q_n\}_{n\in\mathbb{Z}}$ satisfies the Hamiltonian equations

$$\dot{q}_n = \mathrm{i}\frac{\partial H}{\partial \bar{q}_n}, \quad n \in \mathbb{Z}, \tag{2.1}$$

where

$$H = \Lambda + G \tag{2.2}$$

with

$$\Lambda = \sum_{n\in\mathbb{Z}} \mu_n |q_n|^2, \qquad G = \frac{1}{p+1} \int_0^{2\pi} \left|\sum_{n\in\mathbb{Z}} q_n\phi_n\right|^{2p+2} dx.$$

Lemma 2.1. *The gradient G_q is real analytic map from a neighbourhood of the origin of l^ρ into l^ρ, with*

$$\|G_q\|_\rho = O\left(\|q\|_\rho^{2p+1}\right).$$

The proof is similar as Lemma 3 in [12].

Note that

$$G = \frac{1}{p+1} \sum_{i_1,\dots,i_{p+1},j_1,\dots,j_{p+1}} \left(\int_0^{2\pi} \phi_{i_1} \cdots \phi_{i_{p+1}} \bar{\phi}_{j_1} \cdots \bar{\phi}_{j_{p+1}} \, dx \right) q_{i_1} \cdots q_{i_{p+1}} \bar{q}_{j_{i_1}} \cdots \bar{q}_{j_{p+1}}$$

$$= \frac{1}{p+1} \sum_{i_1,\dots,i_{p+1},j_1,\dots,j_{p+1}} G_{i_1 \cdots i_{p+1} j_1 \cdots j_{p+1}} q_{i_1} \cdots q_{i_{p+1}} \bar{q}_{j_{i_1}} \cdots \bar{q}_{j_{p+1}},$$

where

$$G_{i_1 \cdots i_{p+1} j_1 \cdots j_{p+1}} = \int_0^{2\pi} \phi_{i_1} \cdots \phi_{i_{p+1}} \bar{\phi}_{j_1} \cdots \bar{\phi}_{j_{p+1}} \, dx.$$

It is not difficult to verify that $G_{i_1 \cdots i_{p+1} j_1 \cdots j_{p+1}} = 0$ unless $i_1 + \cdots + i_{p+1} = j_1 + \cdots + j_{p+1}$. Moreover, when $i_1 + \cdots + i_{p+1} = j_1 + \cdots + j_{p+1}$, we have $G_{i_1 \cdots i_{p+1} j_1 \cdots j_{p+1}} = (\frac{1}{2\pi})^{p+1}$.

To transform the Hamiltonian (2.2) into a partial Birkhoff normal form, we fix n_1, n_2 ($n_1 \neq n_2$) and define the index sets Δ_*, $* = 0, 1, 2, 3$, as follows. For each $* = 0, 1, 2$, Δ_* is the set of indices $(i_1, \dots, i_{p+1}, j_1, \dots, j_{p+1})$ which have exactly "$*$" components not in $\{n_1, n_2\}$. Δ_3 is the set of the indices $(i_1, \dots, i_{p+1}, j_1, \dots, j_{p+1})$ which have at least three components not in $\{n_1, n_2\}$. We also consider the resonance sets $\mathcal{N} = \{i_1, \dots, i_{p+1}, i_1, \dots, i_{p+1}\} \cap \Delta_0$, $\mathcal{M} = \{i_1, \dots, i_{p+1}, i_1, \dots, i_{p+1}\} \cap \Delta_2$. For our convenience, denote the sets $\mathcal{T}_1$, $\mathcal{T}_2$,

$$\mathcal{T}_1 = \left\{ (i_1, \dots, i_{p+1}, j_1, \dots, j_{p+1}) \in \Delta_2 \setminus \mathcal{M} \;\middle|\; i_1^2 + \cdots + i_{p+1}^2 = j_1^2 + \cdots + j_{p+1}^2 \right\},$$

$$\mathcal{T}_2 = \left\{ (i_1, \dots, i_{p+1}, j_1, \dots, j_{p+1}) \in \Delta_2 \setminus \mathcal{M} \;\middle|\; i_1^2 + \cdots + i_{p+1}^2 \neq j_1^2 + \cdots + j_{p+1}^2 \right\}.$$

Lemma 2.2. *Let* $(i_1, \dots, i_{p+1}, j_1, \dots, j_{p+1}) \in (\Delta_0 \setminus \mathcal{N}) \cup \Delta_1 \cup \mathcal{T}_2$. *If* $i_1 + \cdots + i_{p+1} = j_1 + \cdots + j_{p+1}$, *then*

$$\mu_{i_1} + \cdots + \mu_{i_{p+1}} - \mu_{j_1} - \cdots - \mu_{j_{p+1}} = i_1^2 + \cdots + i_{p+1}^2 - j_1^2 - \cdots - j_{p+1}^2 \neq 0.$$

Proof. If $(i_1, \dots, i_{p+1}, j_1, \dots, j_{p+1}) \in (\Delta_0 \setminus \mathcal{N})$, without losing generality, suppose there are exactly x's n_1 in $\{i_1, \dots, i_{p+1}\}$ and y's n_1 in $\{j_1, \dots, j_{p+1}\}$. It is obvious that $x \neq y$. Therefore, from $i_1 + \cdots + i_{p+1} = j_1 + \cdots + j_{p+1}$, we have $(x - y)n_1 = (x - y)n_2$. Since $n_1 \neq n_2$ and $x \neq y$, it is impossible. This means that if $i_1 + \cdots + i_{p+1} = j_1 + \cdots + j_{p+1}$, there are no elements in $\Delta_0 \setminus \mathcal{N}$.

If $(i_1, \dots, i_{p+1}, j_1, \dots, j_{p+1}) \in \Delta_1$, without losing generality, suppose x_1's n_1 in $\{i_1, \dots, i_{p+1}\}$ and y_1's n_1 in $\{j_1, \dots, j_{p+1}\}$. And the unique index in $\{j_1, \dots, j_{p+1}\}$ different with n_1, n_2 is denoted by z_1. Similarly, from $i_1 + \cdots + i_{p+1} = j_1 + \cdots + j_{p+1}$, one gets

$$(x_1 - y_1)n_1 + (y_1 + 1 - x_1)n_2 = z_1. \tag{2.3}$$

It is easy to see that

$$\begin{aligned} i_1^2+\cdots+i_{p+1}^2-j_1^2-\cdots-j_{p+1}^2 &= (x_1-y_1)n_1^2+(y_1+1-x_1)n_2^2-z_1^2 \\ &= a_1n_1^2+(1-a_1)n_2^2-\big(a_1n_1+(1-a_1)n_2\big)^2 \\ &= a_1(1-a_1)(n_1-n_2)^2, \end{aligned}$$

where $a_1=x_1-y_1$. Since $z_1\neq n_1,n_2$, this means $a_1\neq 0,1$ from (2.3). Therefore, $a_1(1-a_1)(n_1-n_2)^2\neq 0$. $\square$

Lemma 2.3. *Given $n_1<n_2$, $n_1,n_2\in\mathbb{Z}$, there exists a real analytic, symplectic change of coordinates Γ in a neighborhood of the origin of l^ρ which transforms the Hamiltonian* (2.2) *into a partial Birkhoff normal form*

$$H\circ\Gamma=\Lambda+\bar{G}+\tilde{G}+\hat{G}+K, \tag{2.4}$$

such that the corresponding Hamiltonian vector fields $X_{\bar{G}}$, $X_{\tilde{G}}$, $X_{\hat{G}}$ and X_K are real analytic in a neighborhood of the origin in l^ρ, where

$$\begin{aligned} \bar{G} &= c_p\sum_{k=-1}^{p}\big(C_{p+1}^{k+1}\big)^2|q_{n_1}|^{2(p-k)}|q_{n_2}|^{2(k+1)} \\ &\quad + c_p\big(C_{p+1}^{1}\big)^2\sum_{n\neq n_1,n_2}\sum_{k=0}^{p}\big(C_p^k\big)^2|q_{n_1}|^{2(p-k)}|q_{n_2}|^{2k}|q_n|^2, \\ \tilde{G} &= c_p\sum_{\substack{i_1+\cdots+i_{p+1}=j_1+\cdots+j_{p+1}\\ \{i_1,\dots,i_{p+1},j_1,\dots,j_{p+1}\in\mathcal{T}_1\}}} q_{i_1}\cdots q_{i_{p+1}}\bar{q}_{j_1}\cdots\bar{q}_{j_{p+1}}, \\ \hat{G} &= c_p\sum_{\substack{i_1+\cdots+i_{p+1}=j_1+\cdots+j_{p+1}\\ \{i_1,\dots,i_{p+1},j_1,\dots,j_{p+1}\in\Delta_3\}}} q_{i_1}\cdots q_{i_{p+1}}\bar{q}_{j_1}\cdots\bar{q}_{j_{p+1}}, \\ |K| &= \mathcal{O}\big(\|q\|_\rho^{4p+2}\big), \end{aligned}$$

where $c_p=\frac{1}{(2\pi)^p(p+1)}$. Moreover, $K(q,\bar{q})$ has a special form.

We give an explanation for which K has a special form. If $K=\sum_{\alpha,\beta}K_{\alpha\beta}q^\alpha\bar{q}^\beta$, then

$$K_{\alpha\beta}\neq 0 \quad \text{implies} \quad \sum_{i\in\mathbb{Z}}\alpha_i=\sum_{j\in\mathbb{Z}}\beta_j,$$

where $\alpha=(\alpha_i)_{i\in\mathbb{Z}}$ and $\beta=(\beta_j)_{j\in\mathbb{Z}}$. The proof of Lemma 2.3 is a copy of Proposition 3.1 in [11].

The specific form for $\tilde{G}$ is very important for the following proof. We will give it clearly. For our convenience, we will rewrite the coordinates by a,b, which are different with n_1, n_2 in $\{i_1,\dots,i_{p+1},j_1,\dots,j_{p+1}\}\in\mathcal{T}_1$. It is obvious that $a\neq b$. Otherwise, we have $n_1=n_2$. For

$$\tilde{G} = c_p \sum_{\substack{i_1+\cdots+i_{p+1}=j_1+\cdots+j_{p+1} \\ \{i_1,\ldots,i_{p+1},j_1,\ldots,j_{p+1}\in\mathcal{T}_1\}}} q_{i_1}\cdots q_{i_{p+1}}\bar{q}_{j_1}\cdots\bar{q}_{j_{p+1}},$$

we will suppose there exist k_1's q_{n_1}, k_2's $\bar{q}_{n_1}$, l_1's q_{n_2}, l_2's $\bar{q}_{n_2}$. Before we give the concrete form for $\tilde{G}$, we need a preparation lemma.

Lemma 2.4. *When q_a (or q_b) $\in \{q_{i_1},\ldots,q_{i_{p+1}}\}$, one must have $\bar{q}_b$ (or $\bar{q}_a$) $\in \{\bar{q}_{j_1},\ldots,\bar{q}_{j_{p+1}}\}$.*

Proof. Without losing generality, assume that $q_a, q_b \in \{q_{i_1},\ldots,q_{i_{p+1}}\}$. It is easy to get

$$\begin{cases} k_1 + l_1 = p - 1, \\ k_2 + l_2 = p + 1, \\ k_1 n_1 + l_1 n_2 + a + b = k_2 n_1 + l_2 n_2. \end{cases}$$

We will prove that

$$a^2 + b^2 + k_1 n_1^2 + l_1 n_2^2 \neq k_2 n_1^2 + l_2 n_2^2.$$

If this is not true, one gets

$$\begin{cases} a + b + (k_1 - k_2)n_1 + (l_1 - l_2)n_2 = 0, \\ a^2 + b^2 + (k_1 - k_2)n_1^2 + (l_1 - l_2)n_2^2 = 0. \end{cases}$$

Write $s_1 = k_1 - k_2$. It follows $l_1 - l_2 = -2 - s_1$. Therefore,

$$\begin{cases} a + b + s_1 n_1 + (-2 - s_1)n_2 = 0, \\ a^2 + b^2 + s_1 n_1^2 + (-2 - s_1)n_2^2 = 0. \end{cases}$$

Thus, it follows

$$2a^2 + 2\big(s_1 n_1 - (2 + s_1)n_2\big)a + s_1(s_1 + 1)n_1^2 + (2 + s_1)(1 + s_1)n_2^2 - 2s_1(s_1 + 2)n_1 n_2 = 0. \tag{2.5}$$

Note $\Delta = -4s_1(s_1 + 2)(n_1 - n_2)^2$, one can draw the contradictions from the following three cases.

Case 1. If $s_1 = 0$ or $s_1 = -2$.

If $s_1 = 0$, then $a = n_2$. If $s_1 = -2$, then $a = n_1$. It both contradicts with the choice of a.

Case 2. If $s_1 > 0$ or $s_1 < -2$.

Since $\Delta < 0$ in this case, it is obvious (2.5) cannot hold.

Case 3. If $-2 < s_1 < 0$.

Since $s_1 \in \mathbb{Z}$, it follows $s_1 = -1$ and $\Delta = 4(n_1 - n_2)^2$. From (2.5), it is easy to get $a = n_1, n_2$. It is impossible. □

Thus, from Lemma 2.4, one has

$$\begin{cases} k_1 + l_1 = k_2 + l_2 = p, \\ a + k_1 n_1 + l_1 n_2 = b + k_2 n_1 + l_2 n_2, \\ a^2 + k_1 n_1^2 + l_1 n_2^2 = b^2 + k_2 n_1^2 + l_2 n_2^2, \end{cases}$$

where $k_1, k_2 = 0, 1, \dots, p$, $l_1, l_2 = 0, 1, \dots, p$. If denote $k_1 - k_2 = s$, one has $k_1 - k_2 = s = l_2 - l_1$. Further, we have

$$\begin{cases} sn_1 - sn_2 + a - b = 0, \\ sn_1^2 - sn_2^2 + a^2 - b^2 = 0. \end{cases}$$

From $a \neq b$, we get

$$\begin{cases} a = \dfrac{1}{2}(s+1)(n_2 - n_1) + n_1, \\ b = \dfrac{1}{2}(s+1)(n_1 - n_2) + n_2. \end{cases}$$

It is clear that $s \neq 0, \pm 1$, $s = k_1 - k_2 = l_2 - l_1$, $s \in \{-p, \dots, -1, 0, 1, \dots, p\}$ and $k_1 + l_1 = k_2 + l_2 = p$.

On the contrary, we could clearly write all the terms in $\tilde{G}$. Firstly, give all $s \in \{-p, \dots, -2, 2, \dots, p\}$ satisfying

$$\begin{cases} a = \dfrac{1}{2}(s+1)(n_2 - n_1) + n_1 \in \mathbb{Z}, \\ b = \dfrac{1}{2}(s+1)(n_1 - n_2) + n_2 \in \mathbb{Z}. \end{cases}$$

Denote this set of s by $\mathcal{R}_1$. Corresponding to every $s \in \mathcal{R}_1$ mentioned above, we have many integer pairs (k_1, k_2) satisfying $k_1 - k_2 = s$, $k_1, k_2 \in \{0, 1, \dots, p\}$. Denote this set of (k_1, k_2) by $\mathcal{R}_2^s$. From $(k_1, k_2) \in \mathcal{R}_2^s$ and $k_1 + l_1 = k_2 + l_2 = p$, we can give the corresponding integer pairs (l_1, l_2). In this way, for every $s \in \mathcal{R}_1$, we find many terms in $\tilde{G}$. More concretely, they are all terms made of $c_p q_a q_{n_1}^{k_1} q_{n_2}^{l_1} \bar{q}_b \bar{q}_{n_1}^{k_2} \bar{q}_{n_2}^{l_2}$, where $a = \frac{1}{2}(s+1)(n_2 - n_1) + n_1$, $b = \frac{1}{2}(s+1)(n_1 - n_2) + n_2$ and $(k_1, k_2) \in \mathcal{R}_2^s$. When varying $s \in \mathcal{R}_1$, we have get all terms in $\tilde{G}$.

In this way, suppose that $n_2 - n_1 \in 2\mathbb{N}$, we get

$$\tilde{G} = c_p \sum_{t=0}^{p-2} \sum_{j=0}^{t} q_{j_t} q_{n_1}^{p-j} q_{n_2}^{j} \bar{q}_{i_t} \bar{q}_{n_1}^{t-j} \bar{q}_{n_2}^{p-t+j} + c_p \sum_{t=0}^{p-2} \sum_{j=0}^{t} q_{i_t} q_{n_1}^{t-j} q_{n_2}^{p-t+j} \bar{q}_{j_t} \bar{q}_{n_1}^{p-j} \bar{q}_{n_2}^{j}, \quad (2.6)$$

where

$$i_t = \frac{1}{2}(p - t + 1)(n_1 - n_2) + n_2, \quad (2.7)$$

$$j_t = \frac{1}{2}(p - t + 1)(n_2 - n_1) + n_1, \quad t \in \mathcal{T}. \quad (2.8)$$

When $n_2 - n_1 \in 2\mathbb{N} - 1$ and $p \in 2\mathbb{N}$, we get

$$\tilde{G} = c_p \sum_{\substack{t=0 \\ t\in 2\mathbb{Z}+1}}^{p-2} \sum_{j=0}^{t} q_{j_t} q_{n_1}^{p-j} q_{n_2}^{j} \bar{q}_{i_t} \bar{q}_{n_1}^{t-j} \bar{q}_{n_2}^{p-t+j} + c_p \sum_{\substack{t=0 \\ t\in 2\mathbb{Z}+1}}^{p-2} \sum_{j=0}^{t} q_{i_t} q_{n_1}^{t-j} q_{n_2}^{p-t+j} \bar{q}_{j_t} \bar{q}_{n_1}^{p-j} \bar{q}_{n_2}^{j}. \tag{2.9}$$

When $n_2 - n_1 \in 2\mathbb{N} - 1$ and $p \in 2\mathbb{N} + 1$, we get

$$\tilde{G} = c_p \sum_{\substack{t=0 \\ t\in 2\mathbb{Z}}}^{p-2} \sum_{j=0}^{t} q_{j_t} q_{n_1}^{p-j} q_{n_2}^{j} \bar{q}_{i_t} \bar{q}_{n_1}^{t-j} \bar{q}_{n_2}^{p-t+j} + c_p \sum_{\substack{t=0 \\ t\in 2\mathbb{Z}}}^{p-2} \sum_{j=0}^{t} q_{i_t} q_{n_1}^{t-j} q_{n_2}^{p-t+j} \bar{q}_{j_t} \bar{q}_{n_1}^{p-j} \bar{q}_{n_2}^{j}. \tag{2.10}$$

Remark 2.1. Note the simple case $p = 2$. When $n_2 - n_1 \in 2\mathbb{N} - 1$, (from (2.9)) we know that there is no term in $\tilde{G}$. This responds to the case in [11]. When $n_2 - n_1 \in 2\mathbb{N}$, we have

$$\tilde{G} = c_2 q_a q_{n_1}^2 \bar{q}_b \bar{q}_{n_2}^2 + c_2 q_b q_{n_2}^2 \bar{q}_a \bar{q}_{n_1}^2,$$

where $a = \frac{3}{2}(n_2 - n_1) + n_1$, $b = \frac{3}{2}(n_1 - n_2) + n_2$.

Remark 2.2. The similar phenomenon, as the terms of $\tilde{G}$ do not vanish, exists very popularly. It definitely exists in 1D Schrödinger equation with the nonlinearity $|u|^{2p}u$ $(p \geqslant 2)$ under *Dirichlet boundary conditions*. It is why it is difficult to generalize the conclusions of [13] to any p. We point out that this phenomenon also exists in many other equations such as 1D wave equation and beam equation with the nonlinearity $u^{2\bar{r}+1}$ $(\bar{r} \geqslant 3)$ under different boundary conditions. For example, it exists in 1D wave equation

$$u_{tt} - u_{xx} + mu + u^{2\bar{r}+1} = 0, \quad m > 0,\ \bar{r} \geqslant 3,$$

under *Dirichlet boundary conditions*. If use the same notation as [14], when $\bar{r} = 3$, we will find that the nonresonant term $z_{n_2} z_{n_1}^3 z_i \bar{z}_{n_1} \bar{z}_{n_2} \bar{z}_j$ cannot be killed for some $m > 0$ (depending on i, j), where i, j are normal sites and n_1, n_2 are tangent ones and $\lambda_i = \sqrt{i^2 + m}$, $\lambda_j = \sqrt{j^2 + m}$, $\lambda_{n_1} = \sqrt{n_1^2 + m}$ satisfy

$$\begin{cases} 2\lambda_{n_1} + \lambda_i = \lambda_j, \\ 4n_1 + i = j. \end{cases}$$

This also partly explains why existent KAM results for this equation only hold true for positive measure of $m > 0$. See Bambusi [1] and Liang and You [14] for details.

In the following, we will restrict in the most complex case when $n_2 - n_1 \in 2\mathbb{N}$. When $n_2 - n_1 \in 2\mathbb{N} - 1$, the proof is parallel and the conclusion is the same. We omit it.

Note (2.6), we introduce the symplectic polar and complex coordinates to the Hamiltonian (2.4) by setting

$$q_j = \begin{cases} \sqrt{(\xi_j + y_j)}e^{-\mathrm{i}x_j}, & j = n_1, n_2, \\ w_j, & j \neq n_1, n_2, \end{cases}$$

depending on parameters $\xi \in [0,1]^2$. In order to simplify the expression, we substitute ξ_{n_j}, $j=1,2$, by ξ_j, $j=1,2$. Then one gets

$$\mathrm{i}\sum_{j\in\mathbb{Z}} dq_j \wedge d\bar{q}_j = \sum_{j=n_1,n_2} dx_j \wedge dy_j + \mathrm{i}\sum_{j\neq n_1,n_2} dw_j \wedge d\bar{w}_j.$$

Now the new Hamiltonian is

$$H = \langle \omega(\xi), y\rangle + \sum_{n\neq n_1,n_2} \Omega_n(\xi) w_n \bar{w}_n + \Upsilon_1 + \Upsilon_2 + \Upsilon_3, \tag{2.11}$$

where

$$\begin{aligned}
\omega_1(\xi) &= n_1^2 + c_p \sum_{k=0}^{p} \left(C_{p+1}^k\right)^2 C_{p+1-k}^1 \xi_1^{p-k} \xi_2^k,\\
\omega_2(\xi) &= n_2^2 + c_p \sum_{k=0}^{p} \left(C_{p+1}^{k+1}\right)^2 C_{k+1}^1 \xi_1^{p-k} \xi_2^k,\\
\Omega_n(\xi) &= n^2 + c_p \left(C_{p+1}^1\right)^2 \sum_{k=0}^{p} \left(C_p^k\right)^2 \xi_1^{p-k} \xi_2^k, \quad n\neq n_1,n_2,\\
\Upsilon_1 &= \sum_{t=0}^{p-2} \tilde{a}_t w_{j_t} \bar{w}_{i_t} e^{-\mathrm{i}(p-t)(x_1-x_2)} + \sum_{t=0}^{p-2} \tilde{a}_t \bar{w}_{j_t} w_{i_t} e^{\mathrm{i}(p-t)(x_1-x_2)},\\
\tilde{a}_t &= c_p \sum_{j=0}^{t} \xi_1^{\frac{1}{2}(p+t-2j)} \xi_2^{\frac{1}{2}(p-t+2j)},\\
\Upsilon_2 &= \mathcal{O}\left(|\xi|^{p-1}|y|^2\right) + \mathcal{O}\left(|\xi|^{p-1}|y|\|w\|_\rho^2\right),\\
\Upsilon_3 &= \mathcal{O}\left(|\xi|^{p-\frac{1}{2}}\|w\|_\rho^3\right) + \mathcal{O}\left(|\xi|^{2p+1}\right).
\end{aligned} \tag{2.12}$$

Denote $P = \Upsilon_1 + \Upsilon_2 + \Upsilon_3$. Consider the Taylor–Fourier expansion of P,

$$P = \sum_{k,\alpha,\beta} P_{k\alpha\beta}(y) e^{\mathrm{i}kx} w^\alpha \bar{w}^\beta.$$

We have

$$P_{k\alpha\beta}(y) \neq 0, \quad \text{implies} \quad k_1 n_1 + k_2 n_2 + \sum_{n\in\mathbb{Z}\setminus\{n_1,n_2\}} (-\alpha_n + \beta_n) n = 0.$$

In order to cut our expression, write $\mathcal{N} = \{i_0,\dots,i_{p-2}, j_0,\dots,j_{p-2}\}$ and $\mathcal{J} = \{j_0,\dots,j_{p-2}\}$. It is easy to see that $i_0 < i_1 < \cdots < i_{p-2} < j_{p-2} < \cdots < j_1 < j_0$.

Now we will continue to make a symplectic coordinates transformation for the Hamiltonian (2.11) to obtain the suitable form for our applying the infinite KAM method. Our object is to transform Υ_1 to the terms which do not include the angle variables. The following nonlinear symplectic coordinates transformation works.

Lemma 2.5. *The map* $\Psi_1 : (x, y, w, \bar{w}) \to (x^+, y^+, w^+, \bar{w}^+)$ *defined by*

$$\begin{aligned} x^+ &= x, \\ y^+ &= y + \sum_{t=0}^{p-2} k_t |w_{j_t}|^2, \\ (w_i)^+_{i\in\mathcal{N}} &= E(w_i)_{i\in\mathcal{N}}, \\ w_l^+ &= w_l, \quad l \notin \mathcal{N}, \end{aligned}$$

is symplectic, where

$$\begin{aligned} k_t &= \big(-(p-t), p-t\big)^T, \\ E &= \operatorname{diag}\big(1, \ldots, 1, e^{\mathrm{i}\langle k_{p-2}, x\rangle, \ldots, e^{\mathrm{i}\langle k_0, x\rangle}}\big), \\ (w_i)_{i\in\mathcal{N}} &= (w_{i_0}, \ldots, w_{i_{p-2}}, w_{j_{p-2}}, \ldots, w_{j_0})^T. \end{aligned}$$

Remark 2.3. The similar symplectic transformation as Ψ_1 was used in [18].

Under the above symplectic coordinates transformation Ψ_1, the Hamiltonian (2.11) is changed into the new Hamiltonian (for simplicity, we still use the old coordinates $(x, y, w, \bar{w})$)

$$\begin{aligned} H_+ &= H \circ \Psi_1 \\ &= N_0 + P_0 \\ &= \langle \omega, y\rangle + \sum_{n\notin\mathcal{N}} \langle \Omega_n z_n, \bar{z}_n\rangle + \sum_{t=0}^{p-2} \langle \bar{A}_{i_t} z_{i_t}, \bar{z}_{i_t}\rangle + P_0, \end{aligned} \tag{2.13}$$

where

$$\begin{aligned} z_n &= w_n, \quad n \notin \mathcal{N}, \\ z_{i_t} &= (w_{i_t}, w_{j_t})^T, \qquad \bar{z}_{i_t} = (\bar{w}_{i_t}, \bar{w}_{j_t})^T, \\ \bar{A}_{i_t} &= \begin{pmatrix} \Omega_{i_t} & \tilde{a}_t \\ \tilde{a}_t & \Omega_{i_t} + (p-t)\tilde{A} \end{pmatrix}, \\ \tilde{A} &= \sum_0^p c_p \big[\big(C_{p+1}^k\big)^2 C_{p+1-k}^1 - \big(C_{p+1}^{k+1}\big)^2 C_{k+1}^1\big] \xi_1^{p-k} \xi_2^k, \end{aligned} \tag{2.14}$$

and ω, Ω is the same as those in (2.11). Checking directly, we know that P_0 satisfies a generalized compact form with respect to n_1, n_2 and $\mathcal{J}$ (see Appendix A for the definition). More concretely, consider the Taylor–Fourier expansion of P_0,

$$P_0=\sum_{k,\alpha,\beta}P_{0,k\alpha\beta}(y)e^{\mathrm{i}kx}w^{\alpha}\bar{w}^{\beta},$$

we have that $P_{0,k\alpha\beta}(y)\neq 0$ implies

$$k_1n_1+k_2n_2+\sum_{n\in\mathbb{Z}\setminus\{n_1,n_2\}}(-\alpha_n+\beta_n)n=(n_1-n_2)\sum_{t=0}^{p-2}(\alpha_{j_t}-\beta_{j_t})(p-t). \tag{2.15}$$

For (2.13), rescaling $\xi^{\frac{1}{2}}$ by $\epsilon^6\xi$, w, $\bar{w}$ by $\epsilon^4 w$, $\epsilon^4\bar{w}$, and y by $\epsilon^8 y$, one obtains a new Hamiltonian given by the rescaled Hamiltonian

$$\begin{aligned}\tilde{H}&=\epsilon^{6p+8}H_+\big(x,\epsilon^8y,\epsilon^4w,\epsilon^4\bar{w},\epsilon^6\xi,\epsilon\big)\\&=\langle\tilde{\omega},y\rangle+\sum_{n\notin\mathcal{N}}\langle\tilde{\Omega}_nz_n,\bar{z}_n\rangle+\sum_{t=0}^{p-2}\langle\tilde{\bar{A}}_{i_t}z_{i_t},\bar{z}_{i_t}\rangle+\epsilon\tilde{P}_0,\end{aligned} \tag{2.16}$$

where

$$\begin{aligned}\tilde{\omega}_1(\xi)&=\frac{n_1^2}{\epsilon^{6p}}+c_p\sum_{k=0}^{p}\big(C_{p+1}^k\big)^2C_{p+1-k}^1\xi_1^{2p-2k}\xi_2^{2k},\\ \tilde{\omega}_2(\xi)&=\frac{n_2^2}{\epsilon^{6p}}+c_p\sum_{k=0}^{p}\big(C_{p+1}^{k+1}\big)^2C_{k+1}^1\xi_1^{2p-2k}\xi_2^{2k},\\ \tilde{\Omega}_n(\xi)&=\frac{n^2}{\epsilon^{6p}}+c_p\big(C_{p+1}^1\big)^2\sum_{k=0}^{p}\big(C_p^k\big)^2\xi_1^{2p-2k}\xi_2^{2k},\quad n\neq n_1,n_2,\\ \tilde{\bar{A}}_{i_t}&=\begin{pmatrix}\tilde{\Omega}_{i_t} & a_t\\ a_t & \tilde{\Omega}_{i_t}+(p-t)A\end{pmatrix},\\ a_t&=c_p\sum_{0}^{t}\xi_1^{p+t-2j}\xi_2^{p-t+2j},\end{aligned} \tag{2.17}$$

$$A=\sum_{0}^{p}c_p\big[\big(C_{p+1}^k\big)^2C_{p+1-k}^1-\big(C_{p+1}^{k+1}\big)^2C_{k+1}^1\big]\xi_1^{2(p-k)}\xi_2^{2k}, \tag{2.18}$$

$\xi\in\mathcal{O}=[1,2]^2$. It is obvious that $\tilde{P}_0$ also satisfies a generalized compact form with respect to n_1, n_2 and $\mathcal{J}$. For our convenience, we rewrite $\tilde{H}$ by H, $\tilde{\omega}$ by ω, $\tilde{\Omega}$ by Ω, $\tilde{\bar{A}}$ by $\bar{A}$, $\tilde{\bar{B}}$ by $\bar{B}$ and $\tilde{P}_0$ by P_0.

Now the new Hamiltonian is

$$H=\langle\omega,y\rangle+\sum_{n\notin\mathcal{N}}\langle\Omega_nz_n,\bar{z}_n\rangle+\sum_{t=0}^{p-2}\langle\bar{A}_{i_t}z_{i_t},\bar{z}_{i_t}\rangle+\epsilon P_0. \tag{2.19}$$

It is well known that there exists real orthogonal matrix P_t, $t=0,\dots,p-2$, satisfying

$$P_t^T \bar{A}_{i_t} P_t = P_t^{-1} \bar{A}_{i_t} P_t = A_{i_t} = \operatorname{diag}(\lambda_{1,t}, \lambda_{2,t}), \tag{2.20}$$

where

$$\lambda_{1,t} = \Omega_{i_t} + \frac{1}{2}(p-t)A - \frac{1}{2}\sqrt{4a_t^2 + (p-t)^2 A^2} \tag{2.21}$$

and

$$\lambda_{2,t} = \Omega_{i_t} + \frac{1}{2}(p-t)A + \frac{1}{2}\sqrt{4a_t^2 + (p-t)^2 A^2}. \tag{2.22}$$

Lemma 2.6. *The map $\Psi_2 : (x, y, z, \bar{z}) \to (x^+, y^+, z^+, \bar{z}^+)$ defined by*

$$\begin{aligned}
x^+ &= x, \\
y^+ &= y, \\
z_{i_t}^+ &= P_t^{-1} z_{i_t}, \quad t = 0, \ldots, p-2, \\
z_i^+ &= z_i, \quad i \notin \{i_0, \ldots, i_{p-2}\},
\end{aligned}$$

is symplectic.

Proof. It is easy to check that

$$dx^+ \wedge dy^+ + \mathrm{i}\, dz^+ \wedge d\bar{z}^+ = dx \wedge dy + \mathrm{i}\, dz \wedge d\bar{z}. \qquad \square$$

Under the symplectic coordinates transformation Ψ_2, the Hamiltonian (2.19) is changed into the new Hamiltonian

$$\begin{aligned}
H^+ &= H \circ \Psi_2 \\
&= \langle \omega, y^+ \rangle + \sum_{n \notin \mathcal{N}} \langle \Omega_n z_n^+, \bar{z}_n^+ \rangle + \sum_{t=0}^{p-2} \langle A_{i_t} z_{i_t}^+, \bar{z}_{i_t}^+ \rangle + \epsilon P_0^+,
\end{aligned}$$

where

$$\begin{aligned}
\omega_1(\xi) &= \frac{n_1^2}{\epsilon^{6p}} + c_p \sum_{k=0}^{p} \left(C_{p+1}^k\right)^2 C_{p+1-k}^1 \xi_1^{2p-2k} \xi_2^{2k}, \\
\omega_2(\xi) &= \frac{n_2^2}{\epsilon^{6p}} + c_p \sum_{k=0}^{p} \left(C_{p+1}^{k+1}\right)^2 C_{k+1}^1 \xi_1^{2p-2k} \xi_2^{2k}, \\
\Omega_n(\xi) &= \frac{n^2}{\epsilon^{6p}} + c_p \left(C_{p+1}^1\right)^2 \sum_{k=0}^{p} \left(C_p^k\right)^2 \xi_1^{2p-2k} \xi_2^{2k}, \quad n \neq n_1, n_2, \\
A_{i_t} &= \begin{pmatrix} \lambda_{1,t} & 0 \\ 0 & \lambda_{2,t} \end{pmatrix},
\end{aligned} \tag{2.23}$$

$\xi \in \mathcal{O}$. For $\lambda_{1,t}$, $\lambda_{2,t}$, see (2.21) and (2.22). From Lemma A.3, we know that P_0^+ satisfies the generalized compact form with respect to n_1, n_2 and $\mathcal{J}$. For our convenience, we will rewrite H^+ by H, y^+ by y, z_n^+ by z_n, $\bar{z}_n^+$ by $\bar{z}_n$ and ϵP_0^+ by P. Therefore, the Hamiltonian is

$$\begin{aligned} H &= N + P \\ &= \langle \omega, y\rangle + \sum_{n\notin\mathcal{N}} \langle \Omega_n z_n, \bar{z}_n\rangle + \sum_{t=0}^{p-2} \langle A_{i_t} z_{i_t}, \bar{z}_{i_t}\rangle + P(x, y, z, \bar{z}, \xi, \epsilon) \\ &= \langle \omega, y\rangle + \sum_j \tilde{\Omega}_j w_n \bar{w}_n + P(x, y, w, \bar{w}, \xi, \epsilon), \end{aligned} \tag{2.24}$$

where

$$\tilde{\Omega}_j = \begin{cases} \Omega_j, & j \notin \mathcal{N}, \\ \lambda_{1,t}, & j = i_t,\ t \in \mathcal{T}, \\ \lambda_{2,t}, & j = j_t,\ t \in \mathcal{T}, \end{cases}$$

and P satisfies a generalized compact form (2.15). (The subscript j of $\tilde{\Omega}_j$ certainly satisfies $j \neq n_1, n_2$. We do not mention it again in the following.)

In the following, we will use the KAM iteration which involves infinite many steps of coordinate transformations to prove the existence of the KAM tori. To make this quantitative we introduce the following notations and spaces.

Define the phase space:

$$\mathbb{P} := \left(\mathbb{C}^2/2\pi\mathbb{Z}^2\right) \times \mathbb{C}^2 \times l^\rho \times l^\rho.$$

We endow $\mathbb{P}$ with a symplectic structure $dx \wedge dy + \mathrm{i}\sum_{j\in\mathbb{Z}} dw_j \wedge d\bar{w}_j$, $(x, y, w, \bar{w}) \in \mathbb{P}$. Let

$$\mathcal{T}_0^2 = \left(\mathbb{R}^2/2\pi\mathbb{Z}^2\right) \times \{y = 0\} \times \{w = 0\} \times \{\bar{w} = 0\} \subset \mathbb{P}.$$

Then $\mathcal{T}_0^2$ is a torus in $\mathbb{P}$. Introducing a complex neighborhood of $\mathcal{T}_0^2$ in $\mathbb{P}$:

$$D(s, r) = \left\{(x, y, w, \bar{w}) \in \mathbb{P}\colon |\mathrm{Im}\, x| < s,\ |y| < r^2,\ \|w\|_\rho < r,\ \|\bar{w}\|_\rho < r\right\},$$

where $|\cdot|$ denotes the sup-norm for complex vectors. Define a weighted phase space norms

$$|W|_r = |W|_{r,\rho} = |x| + \frac{1}{r^2}|y| + \frac{1}{r}\|w\|_\rho + \frac{1}{r}\|\bar{w}\|_\rho,$$

for $W = (x, y, w, \bar{w}) \in \mathbb{P}$. Let $\bar{\mathcal{O}} \subset \mathbb{R}^2$ be compact and of positive Lebesgue measure. For a map $W : D(s, r) \times \bar{\mathcal{O}} \to \mathbb{P}$, set

$$|W|_{r,\rho,D(s,r)\times\bar{\mathcal{O}}} := \sup_{(x,\xi)\in D(s,r)\times\bar{\mathcal{O}}} \left|W(x, \xi)\right|_{r,\rho}$$

and

$$|W|^*_{r,\rho,D(s,r)\times\bar{\mathcal{O}}}=\max_{|\alpha|\leqslant 8p}\sup_{(x,\xi)\in D(s,r)\times\bar{\mathcal{O}}}\left|\frac{\partial^\alpha W(x,\xi)}{\partial\xi^\alpha}\right|_{r,\rho}.$$

For an $8p$ order Whitney smooth function $F(\xi)$, define

$$\|F\|^*=\max_{|\alpha|\leqslant 8p}\sup_{\xi\in\bar{\mathcal{O}}}\left|\frac{\partial^\alpha F}{\partial\xi^\alpha}\right|,$$

$$\|F\|_*=\max_{1\leqslant|\alpha|\leqslant 8p}\sup_{\xi\in\bar{\mathcal{O}}}\left|\frac{\partial^\alpha F}{\partial\xi^\alpha}\right|.$$

To functions F, associate a Hamiltonian vector field defined as $X_F=\{-F_y,F_x,-iF_{\bar w},iF_w\}$. Denote the norm for X_F by letting

$$|X_F|^*_{r,D(s,r)}=\max_{|\alpha|\leqslant 8p}\sup_{\substack{\xi\in\bar{\mathcal{O}}\\(x,y,w,\bar w)\in D(r,s)}}\left[\left|\frac{\partial^\alpha F_y}{\partial\xi^\alpha}\right|+\frac{1}{r^2}\left|\frac{\partial^\alpha F_x}{\partial\xi^\alpha}\right|+\frac{1}{r}\left\|\frac{\partial^\alpha F_w}{\partial\xi^\alpha}\right\|_\rho+\frac{1}{r}\left\|\frac{\partial^\alpha F_{\bar w}}{\partial\xi^\alpha}\right\|_\rho\right].$$

In the whole of this paper, by c a universal constant, whose size may be different in different place. If $f\leqslant cg$, we write this inequality as $f\leqslant\cdot g$ when we do not care the size of the constant c. Similarly, if $f\geqslant cg$, we write $f\geqslant\cdot g$.

3. KAM step

Theorem 1 will be proved by a KAM iteration which involves an infinite sequence of changes of variables. Each step of KAM iteration makes the perturbation smaller than the previous step at the cost of excluding a small set of parameters. At the end, the KAM iteration will be convergent and the measure of the total excluding set will remain to be small.

We introduce some notations in the following. Denote the sets

$$\mathcal{S}_0^\nu=\left\{\xi:\ \left|\langle k,\omega_\nu\rangle^{-1}\right|\leqslant\frac{c|k|^{8p\tau+6}}{\varepsilon_\nu^{\beta_0}},\ k\neq 0\right\},\tag{3.1}$$

$$\mathcal{S}_1^\nu=\left\{\xi:\ \left|\left(\langle k,\omega_\nu\rangle+\tilde\Omega_{\nu,n}\right)^{-1}\right|\leqslant\frac{c\max\{|k|^{8p\tau+6},1\}}{\varepsilon_\nu^{\beta_0}}\right\},\tag{3.2}$$

$$\mathcal{S}_{2,1}^\nu=\left\{\xi:\quad\begin{array}{l}|(\langle k,\omega_\nu\rangle+\tilde\Omega_{\nu,n}+\tilde\Omega_{\nu,m})^{-1}|\leqslant\frac{c\max\{|k|^{8p\tau+6},1\}}{\varepsilon_\nu^{\beta_0}(||n|-|m||+1)},\\ \quad\text{where } n,m\notin\mathcal{N}\text{ or } n,m\in\mathcal{N},\\ |(\langle k,\omega_\nu\rangle+\tilde\Omega_{\nu,n}+\tilde\Omega_{\nu,i_t})^{-1}|\leqslant\frac{c\max\{|k|^{8p\tau+6},1\}}{\varepsilon_\nu^{\beta_0}(||i_t|-|n||+1)},\\ \quad\text{where } n\notin\mathcal{N},\ t\in\mathcal{T},\ |k_1n_1+k_2n_2|=|n+i_t|,\\ |(\langle k,\omega_\nu\rangle+\tilde\Omega_{\nu,n}+\tilde\Omega_{\nu,j_t})^{-1}|\leqslant\frac{c\max\{|k|^{8p\tau+6},1\}}{\varepsilon_\nu^{\beta_0}(||j_t|-|n||+1)},\\ \quad\text{where } n\notin\mathcal{N},\ t\in\mathcal{T},\ |k_1n_1+k_2n_2|=|n+j_t+(n_1-n_2)(p-t)|\end{array}\right\},\tag{3.3}$$

and

$$\mathcal{S}_{2,2}^{\nu} = \left\{ \xi: \begin{array}{l} |(\langle k, \omega_\nu\rangle + \tilde{\Omega}_{\nu,n} - \tilde{\Omega}_{\nu,m})^{-1}| \leqslant \frac{c\max\{|k|^{8p\tau+6},1\}}{\varepsilon_\nu^{\beta_0}(||n|-|m||+1)}, \\ \quad \text{where } n, m \notin \mathcal{N},\ |k| + ||n| - |m|| \neq 0,\ |k_1 n_1 + k_2 n_2| = |n-m|, \\ |(\langle k, \omega_\nu\rangle + \tilde{\Omega}_{\nu,n} - \tilde{\Omega}_{\nu,m})^{-1}| \leqslant \frac{c\max\{|k|^{8p\tau+6},1\}}{\varepsilon_\nu^{\beta_0}(||n|-|m||+1)}, \\ \quad \text{where } n, m \in \mathcal{N},\ |k| + |n-m| \neq 0, \\ |(\langle k, \omega_\nu\rangle + \tilde{\Omega}_{\nu,n} - \tilde{\Omega}_{\nu,i_t})^{-1}| \leqslant \frac{c\max\{|k|^{8p\tau+6},1\}}{\varepsilon_\nu^{\beta_0}(||i_t|-|n||+1)}, \\ \quad \text{where } n \notin \mathcal{N},\ t \in \mathcal{T},\ |k_1 n_1 + k_2 n_2| = |n - i_t|, \\ |(\langle k, \omega_\nu\rangle + \tilde{\Omega}_{\nu,n} - \tilde{\Omega}_{\nu,j_t})^{-1}| \leqslant \frac{c\max\{|k|^{8p\tau+6},1\}}{\varepsilon_\nu^{\beta_0}(||j_t|-|n||+1)}, \\ \quad \text{where } n \notin \mathcal{N},\ t \in \mathcal{T},\ |k_1 n_1 + k_2 n_2| = |n - j_t - (n_1 - n_2)(p-t)| \end{array} \right\}, \tag{3.4}$$

where $\omega_0 = \omega$, $\tilde{\Omega}_{0,n} = \tilde{\Omega}_n$, and[1]

$$\varepsilon_\nu = \begin{cases} \epsilon, & \nu = 0, \\ \epsilon_\nu, & \nu \geqslant 1. \end{cases} \tag{3.5}$$

To begin with the KAM iteration, we fix $r, s, \rho > 0$ and restrict the Hamiltonian (2.24) to the domain $D(s, r)$ and restrict the parameters to the set $\mathcal{O}_0 = \mathcal{O} \setminus \mathcal{R}^0$, where

$$\mathcal{O}_0 \subset \mathcal{S}_0^0 \cup \mathcal{S}_1^0 \cup \mathcal{S}_{2,1}^0 \cup \mathcal{S}_{2,2}^0, \tag{3.6}$$

where $0 \leqslant |k| \leqslant K_0$ and

$$\mathcal{R}^0 = \mathcal{R}_0^0 \cup \left(\mathcal{R}_{1,1}^0 \cup \mathcal{R}_{1,2}^0\right) \cup \mathcal{R}_2^0.$$

Please refer to Section 4 and Lemma 4.12 for more. β_0 is a constant and will be chosen later.

Suppose $\|\omega\|_* \leqslant M_1$, $\max_{j\in\mathbb{Z}} |\tilde{\Omega}_j|_* \leqslant M_2$, $M_1 + M_2 \geqslant 1$. Define $M = (M_1 + M_2)^{8p}$. Initially, we set $\omega_0 = \omega$, $\tilde{\Omega}_{0,n} = \tilde{\Omega}_n$, $N_0 = N$, $P_0 = P$, $r_0 = r$, $s_0 = s$, $M_0 = M$ and

$$N_0 = \langle \omega_0, y\rangle + \sum_n \tilde{\Omega}_n w_n \bar{w}_n,$$

$$H_0 = N_0 + P_0.$$

Hence, H_0 is real analytic on $D(r_0, s_0)$ and also depends on $\xi \in \mathcal{O}_0$ Whitney smoothly. It is clear that there is a constant $c_0 > 0$ such that

$$|X_{P_0}|^*_{r_0, D(r_0,s_0),\mathcal{O}_0} \leqslant c_0 \epsilon \equiv \epsilon_0. \tag{3.7}$$

P_0 satisfies a general compact form (2.15).

[1] Note the difference between ϵ_ν and ε_ν.

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

$$H = H_\nu = N_\nu + P_\nu(x, y, w, \bar{w}),$$
$$N = N_\nu = \langle \omega_\nu(\xi), y \rangle + \sum_n \tilde{\Omega}_{\nu,n}(\xi) w_n \bar{w}_n,$$

which is real analytic in $(x, y, w, \bar{w}) \in D_\nu = D(r_\nu, s_\nu)$ and depends on $\xi \in \mathcal{O}_\nu \subset \mathcal{O}$ Whitney smoothly, where

$$\mathcal{O}_\nu \subset \mathcal{S}_0^\nu \cup \mathcal{S}_1^\nu \cup \mathcal{S}_{2,1}^\nu \cup \mathcal{S}_{2,2}^\nu, \tag{3.8}$$

$0 \leqslant |k| \leqslant K_\nu'$[2] for some $r_\nu \leqslant r_0$, $s_\nu \leqslant s_0$ and

$$K_\nu' = \begin{cases} K_0, & \nu = 0, \\ \infty, & \nu \geqslant 1. \end{cases}$$

We also assume that

$$|X_{P_\nu}|^*_{r_\nu, D(r_\nu, s_\nu)} \leqslant \epsilon_\nu \leqslant \epsilon_0$$

and $P_\nu = \sum_{k,\alpha,\beta} P^\nu_{k\alpha\beta}(y) e^{\mathrm{i}\langle k,x\rangle} w^\alpha \bar{w}^\beta$ has a generalized compact form with respect to n_1, n_2 and $\mathcal{J}$.

To simplify notations, in what follows, the quantities without subscripts refer to the ones at the νth step, while the quantities with subscripts "+" denote the corresponding ones at the $(\nu + 1)$th step. We will construct a symplectic transformation $\Phi = \Phi_\nu$, which, in smaller frequency and phase domains, carries the above Hamiltonian into the next KAM cycle.

3.1. Solving the linearized equations

Expand P into the Fourier–Taylor series

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

where[3] $k \in \mathbb{Z}^2$, $l \in \mathbb{N}_0^2$ and the multi-index α, β run over the set $\alpha \equiv (\ldots, \alpha_n, \ldots)$, $\beta \equiv (\ldots, \beta_n, \ldots)$, $\alpha_n, \beta_n \in \mathbb{N}_0$, with finitely many non-vanishing components. We denote by 0 the multi-index whose components are all zeros and by e_n the multi-index whose nth component is 1 and other components are all zeros.

[2] Where "$|k| \leqslant \infty$" means "$|k| < \infty$." We confuse the notation for simplicity.

[3] $\mathbb{N}_0$ means $\mathbb{N} \cup \{0\}$.

Let R be the truncation of P given by

$$R(x,y,w,\bar w)=\sum_{|k|\leqslant K,\,|l|\leqslant 1}P_{kl00}e^{\mathrm{i}\langle k,x\rangle}y^l+\sum_{|k|\leqslant K,n}\bigl(P_n^{k10}w_n+P_n^{k01}\bar w_n\bigr)e^{\mathrm{i}\langle k,x\rangle}$$
$$+\sum_{|k|\leqslant K,n,m}\bigl(P_{nm}^{k20}w_nw_m+P_{nm}^{k02}\bar w_n\bar w_m+P_{nm}^{k11}w_n\bar w_m\bigr)e^{\mathrm{i}\langle k,x\rangle},$$

where $P_n^{k10}=P_{kl\alpha\beta}$ with $\alpha=e_n$, $\beta=0$; $P_n^{k01}=P_{kl\alpha\beta}$ 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 $\alpha=0$, $\beta=e_n+e_m$.

Since P has a generalized compact normal form with respect to n_1, n_2, $\mathcal{J}$, this means

$$\begin{aligned}
&P_{n,i_t}^{k20}=0,\quad \text{if } k_1n_1+k_2n_2-n-i_t\neq 0,\ n\notin\mathcal{N},\ t\in\mathcal{T},\\
&P_{n,j_t}^{k20}=0,\quad \text{if } k_1n_1+k_2n_2-n-j_t\neq (n_1-n_2)(p-t),\ n\notin\mathcal{N},\ t\in\mathcal{T},\\
&P_{nm}^{k11}=0,\quad \text{if } k_1n_1+k_2n_2-n+m\neq 0,\ n,m\notin\mathcal{N},\\
&P_{n,i_t}^{k11}=0,\quad \text{if } k_1n_1+k_2n_2-n+i_t\neq 0,\ n\notin\mathcal{N},\ t\in\mathcal{T},\\
&P_{n,j_t}^{k11}=0,\quad \text{if } k_1n_1+k_2n_2-n+j_t\neq (n_1-n_2)(t-p),\ n\notin\mathcal{N},\ t\in\mathcal{T},\\
&P_{n,i_t}^{k02}=0,\quad \text{if } k_1n_1+k_2n_2+n+i_t\neq 0,\ n\notin\mathcal{N},\ t\in\mathcal{T},\\
&P_{n,j_t}^{k02}=0,\quad \text{if } k_1n_1+k_2n_2+n+j_t\neq (n_1-n_2)(t-p),\ n\notin\mathcal{N},\ t\in\mathcal{T}.
\end{aligned}$$

In particular, $P_{nm}^{k11}=0$ if $|k|=0$ and $n\neq m$, where $n,m\notin\mathcal{N}$.

Below we look for a special F, defined in a domain $D_+=D(r_+,s_+)$ such that the time one map $\Phi=\Phi_F^1$ 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

$$\begin{aligned}
H\circ\Phi_F^1&=(N+R)\circ\Phi_F^1+(P-R)\circ\Phi_F^1\\
&=N+\{N,F\}+R\\
&\quad+\int_0^1(1-t)\bigl\{\{N,F\},F\bigr\}\circ\Phi_F^t\,dt+\int_0^1\{R,F\}\circ\Phi_F^t\,dt+(P-R)\circ\Phi_F^1\\
&=N_++P_++\{N,F\}+R-P_{0000}-\langle\omega',y\rangle-\sum_n R_{nn}^{011}w_n\bar w_n,
\end{aligned}\tag{3.9}$$

where

$$\omega'=\int\frac{\partial P}{\partial y}\,dx\Big|_{w=\bar w=0,\,y=0},$$

$$N_+=N+\hat N=N+P_{0000}+\langle\omega',y\rangle+\sum_n R_{nn}^{011}w_n\bar w_n,\tag{3.10}$$

$$P_+ = \int_0^1 (1-t)\big\{\{N, F\}, F\big\} \circ \Phi_F^t \, dt + \int_0^1 \{R, F\} \circ \Phi_F^t \, dt + (P - R) \circ \Phi_F^1 \tag{3.11}$$

satisfying the homological equation

$$\{N, F\} + R - P_{0000} - \langle \omega', y \rangle - \sum_n R_{nn}^{011} w_n \bar{w}_n = 0. \tag{3.12}$$

Note the term $\sum_n R_{nn}^{011} w_n \bar{w}_n$ has not been eliminated by symplectic change, so we define $F_{nn}^{011} = 0$.

In order to solve the homological equation (3.12), let F has the form

$$\begin{aligned} F(x, y, w, \bar{w}) &= F_0 + F_1 + F_2 \\ &= \sum_{|k| \leqslant K, \, |l| \leqslant 1} F_{kl00} e^{\mathrm{i}\langle k, x\rangle} y^l + \sum_{|k| \leqslant K, n} \big(F_n^{k10} w_n + F_n^{k01} \bar{w}_n\big) e^{\mathrm{i}\langle k, x\rangle} \\ &\quad + \sum_{|k| \leqslant K, n, m} \big(F_{nm}^{k20} w_n w_m + F_{nm}^{k02} \bar{w}_n \bar{w}_m + F_{nm}^{k11} w_n \bar{w}_m\big) e^{\mathrm{i}\langle k, x\rangle}. \end{aligned}$$

By comparing the coefficients, it is easy to see that the homological equation (3.12) is equivalent to

$$\begin{aligned} \langle k, \omega \rangle F_{kl00} &= \mathrm{i} P_{kl00}, \quad k \neq 0, \ |l| \leqslant 1, \\ \big(\langle k, \omega \rangle + \tilde{\Omega}_n\big) F_n^{k10} &= \mathrm{i} P_n^{k10}, \\ \big(\langle k, \omega \rangle - \tilde{\Omega}_n\big) F_n^{k01} &= \mathrm{i} P_n^{k01}, \\ \big(\langle k, \omega \rangle + \tilde{\Omega}_n + \tilde{\Omega}_m\big) F_{nm}^{k20} &= \mathrm{i} P_{nm}^{k20}, \\ \big(\langle k, \omega \rangle + \tilde{\Omega}_n - \tilde{\Omega}_m\big) F_{nm}^{k11} &= \mathrm{i} P_{nm}^{k11}, \quad |k| + \big||n| - |m|\big| \neq 0, \\ \big(\langle k, \omega \rangle - \tilde{\Omega}_n - \tilde{\Omega}_m\big) F_{nm}^{k02} &= \mathrm{i} P_{nm}^{k02}, \end{aligned}$$

where $0 \leqslant |k| \leqslant K'$. Hence the homological equation (3.12) is uniquely solvable on $\mathcal{O}$ to yield the function F which is real analytic in $(x, y, w, \bar{w})$ and Whitney smooth in $\omega \in \mathcal{O}$. Since P has a generalized compact form with respect to n_1, n_2 and $\mathcal{J}$, it is easy to see that F also has the same property. The following lemma is standard, see [15] and [16] for details.

Lemma 3.1. *F satisfies a generalized compact form with respect to n_1, n_2 and $\mathcal{J}$ and*

$$\begin{aligned} |X_{\hat{N}}|_{r, D(s,r)}^* &\leqslant |X_R|_{r, D(s,r)}^*, \\ |X_F|_{r, D(s-\sigma,r)}^* &\leqslant \frac{cM}{\epsilon^{(8p+1)\beta_0} \sigma^\mu} |X_R|_{r, D(s,r)}^*, \end{aligned}$$

where $\mu = 8p(8p+1)\tau + 56p + 8$.

Lemma 3.2. *If $|X_F|^*_{r,D(s-\sigma,r)} \leqslant \sigma$, then for any $\xi \in \mathcal{O}$, the flow $X^t_F(\cdot,\xi)$ exists on $D(s-2\sigma,\frac{r}{2})$ for $|t| \leqslant 1$ and maps $D(s-2\sigma,\frac{r}{2})$ into $D(s-\sigma,r)$. Moreover, for $|t| \leqslant 1$,*

$$\left|X^t_F - \mathrm{id}\right|^*_{r,D(s-2\sigma,\frac{r}{2})}, \sigma\left\|DX^t_F - \mathrm{Id}\right\|^*_{r,r,D(s-3\sigma,\frac{r}{4})} \leqslant c|X_F|^*_{r,D(s-\sigma,r)},$$

where D is the differentiation operator with respect to $(x,y,z,\bar{z})$, id *and* Id *are identity mapping and unit matrix, and the operator norm*

$$\left\|A(\xi,\eta)\right\|_{\bar{r},r,D(s,r)} = \sup_{\eta\in D(s,r)} \sup_{w\neq 0} \frac{\|A(\xi,\eta)w\|_{\rho,\bar{r}}}{\|w\|_{\rho,r}},$$

$$\|A\|^*_{r,r} = \max_{|\alpha|\leqslant 8p}\left\{\left\|\frac{\partial^\alpha A}{\partial\xi^\alpha}\right\|_{r,r}\right\}.$$

For the proof refer to [16].

Below we consider the new perturbation under the symplectic transformation $\Phi = X^t_F|_{t=1}$. Let $|X_P|^*_{r,D(s,r)} \leqslant \epsilon$. From the above, we have

$$R = \sum_{\substack{|k|\leqslant K\\ 2|m|+|q+\bar{q}|\leqslant 2}} R_{kmq\bar{q}}\, y^m w^q \bar{w}^{\bar{q}} e^{\mathrm{i}\langle k,x\rangle}.$$

Thus $|X_R|^*_{r,D(s,r)} \leqslant \cdot|X_P|^*_{r,D(s,r)} \leqslant \cdot\epsilon$, and for $\eta \leqslant \frac{1}{8}$,

$$|X_{P-R}|^*_{\eta r,D(s-\sigma,2\eta r)} \leqslant \cdot\eta\epsilon + e^{-K'\sigma}\epsilon. \tag{3.13}$$

Due to the generalized compact form of P with respect to n_1, n_2 and $\mathcal{J}$, w_n and $\bar{w}_{-n}$ are not coupled in P for any $n \neq 0$ (we check this in Appendix A). This leads to the following simple new normal form

$$\begin{aligned} N_+ &= N + \langle\omega', y\rangle + \sum_n P^{011}_{nn} w_n\bar{w}_n \\ &= \langle\omega_+, y\rangle + \sum_n \tilde{\Omega}_{+,n} w_n\bar{w}_n, \end{aligned}$$

where $\omega_+ = \omega + (\{P_{0l00}\}_{|l|=1})$, $\tilde{\Omega}_{+,n} = \tilde{\Omega}_n + P^{011}_{nn}$. By Lemma 3.1, one has $|X_{\hat{N}}|^*_{r,D(s,r)} \leqslant \cdot\epsilon$. Therefore,

$$\|\omega_+ - \omega\|^*, \|\tilde{\Omega}_+ - \tilde{\Omega}\|^* \leqslant \cdot\epsilon, \tag{3.14}$$

where $\|\tilde{\Omega}\|^* = \max_{j\in\mathbb{Z}}|\tilde{\Omega}_j|^*$. If $\frac{cM\epsilon^{1-(8p+1)\beta_0}}{\sigma^{\mu+1}} \leqslant 1$, by Lemmas 3.1 and 3.2, it follows that for $|t| \leqslant 1$,

$$\frac{1}{\sigma}\left|X^t_F - \mathrm{id}\right|^*_{r,D(s-2\sigma,\frac{r}{2})}, \left\|DX^t_F - \mathrm{Id}\right\|^*_{r,r,D(s-3\sigma,\frac{r}{4})} \leqslant \frac{cM\epsilon^{1-(8p+1)\beta_0}}{\sigma^{\mu+1}}. \tag{3.15}$$

Under the transformation $\Phi = X_F^1$, $(N+R)\circ\Phi = N_+ + R_+$, where $R_+ = \int_0^1\{(1-t)\hat{N} + tR, F\}\circ X_F^t$. Thus, $H\circ\Phi = N_+ + R_+ + (P-R)\circ\Phi = N_+ + P_+$, where the new perturbation

$$P_+ = R_+ + (P-R)\circ\Phi = (P-R)\circ\Phi + \int_0^1 \{\bar{R}(t), F\}\circ X_F^t\,dt,$$

where $\bar{R}(t) = (1-t)\hat{N} + tR$. Hence, the Hamiltonian vector field of the new perturbation is $X_{P_+} = (X_F^1)^*(X_{P-R}) + \int_0^1 (X_F^t)^*[X_{\bar{R}(t)}, X_F]\,dt$. For the estimate of X_{P_+}, we need the following lemma.

Lemma 3.3. *If the Hamiltonian vector field $W(\cdot,\xi)$ on $V = D(s-4\sigma, 2\eta r)$ depends on the parameter $\xi \in \mathcal{O}$ with $\|W\|_{r,V}^* < +\infty$, and $\Phi = X_F^t : U = D(s-5\sigma, \eta r) \to V$, then $\Phi^* W = D\Phi^{-1} W\circ\Phi$ and if $\frac{cM\epsilon^{1-(8p+1)\beta_0}}{\eta^2\sigma^{\mu+1}} \leqslant 1$, we have $\|\Phi^* W\|_{\eta r,U}^* \leqslant c\|W\|_{\eta r,V}^*$.*

For the proof refer to [15].

Now we estimate X_{P_+}. By Lemma 3.3, if $\frac{cM\epsilon^{1-(8p+1)\beta_0}}{\eta^2\sigma^{\mu+1}} \leqslant 1$,

$$|X_{P_+}|_{\eta r, D(s-5\sigma,\eta r)}^* \leqslant \frac{c}{2}|X_{P-R}|_{\eta r, D(s-4\sigma, 2\eta r)}^* + \frac{c}{2}\int_0^1 \big|[X_{\bar{R}(t)}, X_F]\big|_{\eta r, D(s-4\sigma,2\eta r)}^*\,dt.$$

By Cauchy's inequality and Lemma 3.2, one obtains

$$\begin{aligned}\big|[X_{\bar{R}(t)}, X_F]\big|_{\eta r, D(s-4\sigma,2\eta r)}^* &\leqslant \frac{cM\epsilon^{2-(8p+1)\beta_0}}{2\eta^2\sigma^{\mu+1}} \\ &= \frac{c}{2}M\eta\epsilon,\end{aligned}$$

where one chooses $\eta^3 = \frac{\epsilon^{1-(8p+1)\beta_0}}{\sigma^{\mu+1}}$. Combining (3.13) we have

$$|X_{P_+}|_{\eta r, D(s-5\sigma,\eta r)}^* \leqslant \frac{c}{2}M\eta\epsilon + e^{-K'\sigma}\epsilon.$$

If choose $K_0' = K_0 = |\frac{\ln\eta_0}{\sigma_0}|$ and as we know before $K_\nu' = \infty$, $\nu \geqslant 2$, we get

$$|X_{P_+}|_{\eta r, D(s-5\sigma,\eta r)}^* \leqslant cM\eta\epsilon.$$

Lemma 3.4. *P_+ has a generalized compact form with respect to n_1, n_2 and $\mathcal{J}$.*

Proof. Note that

$$P_+ = P - R + \{P, F\} + \frac{1}{2!}\big\{\{N, F\}, F\big\} + \frac{1}{2!}\big\{\{P, F\}, F\big\} + \cdots + \frac{1}{n!}\big\{\dots\{N, \underbrace{F\}\dots, F}_{n's\ F}\big\} + \frac{1}{n!}\big\{\dots\{P, \underbrace{F\}\dots, F}_{n's\ F}\big\} + \cdots.$$

Since P has a generalized compact form with respect to n_1, n_2 and $\mathcal{J}$, so do $P - R$ and $\{N, F\} = P_{0000} + \langle\omega', y\rangle + \sum_n P_{nn}^{011} w_n \bar{w}_n - R$. The lemma then follows from Lemma A.2. □

3.2. Iteration lemma

To iterate the KAM step infinitely we must choose suitable sequences. For $\nu \geqslant 0$ set

$$\epsilon_{\nu+1} = \frac{cM(\nu)\epsilon_\nu^{\frac{4}{3}-\frac{1}{3}(8p+1)\beta_0}}{\sigma_\nu^{\frac{1}{3}(1+\mu)}}, \qquad \sigma_{\nu+1} = \frac{\sigma_\nu}{2}, \qquad \eta_\nu^3 = \frac{\epsilon_\nu^{1-(8p+1)\beta_0}}{\sigma_\nu^{1+\mu}},$$

where $\beta_0 = \frac{1}{2(8p+1)}$. Furthermore,

$$s_{\nu+1} = s_\nu - 5\sigma_\nu, \qquad r_{\nu+1} = \eta_\nu r_\nu, \qquad M(\nu) = \big(M_1 + M_2 + 2c(\epsilon_0 + \cdots + \epsilon_{\nu-1})\big)^{8p},$$

and $D_\nu = D(s_\nu, r_\nu)$. As initial value fix $\sigma_0 = \frac{s_0}{20} \leqslant \frac{1}{2}$. Assume

$$\epsilon_0 \leqslant \gamma_0\sigma_0^{6(\mu+1)}, \qquad \gamma_0 \leqslant \min\left\{\frac{1}{c^6 2^{13(1+\mu)} M^{42}}, \left(\frac{c_0}{8c}\right)^{8p\tau+7}\right\}, \tag{3.16}$$

where $c_0 = \frac{3}{2}c_p(2p)!(p+1)$. Finally, let $K_{\nu+1} = K_0 2^\nu$. We must emphasize that the readers must notice the difference between K_ν and K'_ν.

Lemma 3.5. *Suppose $H_\nu = N_\nu + P_\nu$ ($\nu \geqslant 0$), is given on $D_\nu \times \mathcal{O}_\nu$, where $N_\nu = \langle\omega_\nu(\xi), y\rangle + \langle\tilde{\Omega}_\nu, z\bar{z}\rangle$ is a normal form satisfying*

$$\big|\langle k, \omega_\nu\rangle^{-1}\big| \leqslant \frac{c|k|^{8p\tau+6}}{\epsilon_\nu^{\beta_0}}, \quad k \neq 0,$$

$$\big|\big(\langle k, \omega_\nu\rangle + \tilde{\Omega}_{\nu,n}\big)^{-1}\big| \leqslant \frac{c\max\{|k|^{8p\tau+6}, 1\}}{\epsilon_\nu^{\beta_0}},$$

$$\big|\big(\langle k, \omega_\nu\rangle + \tilde{\Omega}_{\nu,n} + \tilde{\Omega}_{\nu,m}\big)^{-1}\big| \leqslant \frac{c\max\{|k|^{8p\tau+6}, 1\}}{\epsilon_\nu^{\beta_0}(||n| - |m|| + 1)},$$

where $n, m \notin \mathcal{N}$ or $n, m \in \mathcal{N}$,

$$\big|\big(\langle k, \omega_\nu\rangle + \tilde{\Omega}_{\nu,n} + \tilde{\Omega}_{\nu,i_t}\big)^{-1}\big| \leqslant \frac{c\max\{|k|^{8p\tau+6}, 1\}}{\epsilon_\nu^{\beta_0}(||i_t| - |n|| + 1)},$$

where $n \notin \mathcal{N}$, $t \in \mathcal{T}$, $|k_1 n_1 + k_2 n_2| = |n + i_t|$,

$$\big|\big(\langle k,\omega_\nu\rangle+\tilde\Omega_{\nu,n}+\tilde\Omega_{\nu,j_t}\big)^{-1}\big|\leqslant\frac{c\max\{|k|^{8p\tau+6},1\}}{\epsilon_\nu^{\beta_0}(||j_t|-|n||+1)},$$
$$\text{where } n\notin\mathcal{N},\ t\in\mathcal{T},\ |k_1n_1+k_2n_2|=\big|n+j_t+(n_1-n_2)(p-t)\big|,$$
$$\big|\big(\langle k,\omega_\nu\rangle+\tilde\Omega_{\nu,n}-\tilde\Omega_{\nu,m}\big)^{-1}\big|\leqslant\frac{c\max\{|k|^{8p\tau+6},1\}}{\epsilon_\nu^{\beta_0}(||n|-|m||+1)},$$
$$\text{where } n,m\notin\mathcal{N},\ |k|+\big||n|-|m|\big|\neq 0,\ |k_1n_1+k_2n_2|=|n-m|,$$
$$\big|\big(\langle k,\omega_\nu\rangle+\tilde\Omega_{\nu,n}-\tilde\Omega_{\nu,m}\big)^{-1}\big|\leqslant\frac{c\max\{|k|^{8p\tau+6},1\}}{\epsilon_\nu^{\beta_0}(||n|-|m||+1)},$$
$$\text{where } n,m\in\mathcal{N},\ |k|+|n-m|\neq 0,$$
$$\big|\big(\langle k,\omega_\nu\rangle+\tilde\Omega_{\nu,n}-\tilde\Omega_{\nu,i_t}\big)^{-1}\big|\leqslant\frac{c\max\{|k|^{8p\tau+6},1\}}{\epsilon_\nu^{\beta_0}(||i_t|-|n||+1)},$$
$$\text{where } n\notin\mathcal{N},\ t\in\mathcal{T},\ |k_1n_1+k_2n_2|=|n-i_t|,$$
$$\big|\big(\langle k,\omega_\nu\rangle+\tilde\Omega_{\nu,n}-\tilde\Omega_{\nu,j_t}\big)^{-1}\big|\leqslant\frac{c\max\{|k|^{8p\tau+6},1\}}{\epsilon_\nu^{\beta_0}(||j_t|-|n||+1)},$$
$$\text{where } n\notin\mathcal{N},\ t\in\mathcal{T},\ |k_1n_1+k_2n_2|=\big|n-j_t-(n_1-n_2)(p-t)\big|,$$

for above all k satisfying $0\leqslant|k|\leqslant K'_\nu$, P_ν has a generalized compact form with respect to n_1, n_2 and $\mathcal{J}$, and

$$|X_{P_\nu}|^*_{r_\nu,D_\nu}\leqslant\epsilon_\nu.$$

Then there exist a Whitney smooth family of real analytic symplectic coordinate transformations $\Phi_{\nu+1}:D_{\nu+1}\times\mathcal{O}_\nu\to D_\nu$ and a closed subset

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

of $\mathcal{O}_\nu$, where

$$\mathcal{R}^{\nu+1}(\epsilon_{\nu+1})=\mathcal{R}^{\nu+1}_{00}\cup\mathcal{R}^{\nu+1}_{10}\cup\mathcal{R}^{\nu+1}_{20}\cup\mathcal{R}^{\nu+1}_{11},$$
$$\mathcal{R}^{\nu+1}_{20}=\mathcal{R}^{\nu+1}_{20,1}\cup\mathcal{R}^{\nu+1}_{20,2}\cup\mathcal{R}^{\nu+1}_{20,3},$$
$$\mathcal{R}^{\nu+1}_{11}=\mathcal{R}^{\nu+1}_{11,1}\cup\mathcal{R}^{\nu+1}_{11,2}\cup\mathcal{R}^{\nu+1}_{11,3}\cup\mathcal{R}^{\nu+1}_{11,4},$$

and

$$\mathcal{R}^{\nu+1}_{00}=\bigcup_{K'_{\nu+1}\geqslant|k|>K_\nu}\bigg\{\xi\in\mathcal{O}_\nu:\ \big|\langle k,\omega_{\nu+1}\rangle\big|<\frac{\epsilon_{\nu+1}^{\beta_0}}{c|k|^{8p\tau+6}},\ k\neq 0\bigg\},$$

$$\mathcal{R}^{\nu+1}_{10}=\bigcup_{K'_{\nu+1}\geqslant|k|>K_\nu,n}\bigg\{\xi\in\mathcal{O}_\nu:\ \big|\langle k,\omega_{\nu+1}\rangle+\tilde\Omega_{\nu+1,n}\big|<\frac{\epsilon_{\nu+1}^{\beta_0}}{c\max\{|k|^{8p\tau+6},1\}}\bigg\},$$

$$\mathcal{R}_{20,1}^{\nu+1} = \bigcup_{K'_{\nu+1} \geqslant |k| > K_\nu, n, m} \left\{ \xi \in \mathcal{O}_\nu \colon \left| \langle k, \omega_{\nu+1} \rangle + \tilde{\Omega}_{\nu+1,n} + \tilde{\Omega}_{\nu+1,m} \right| < \frac{\epsilon_{\nu+1}^{\beta_0} (||n| - |m|| + 1)}{c \max\{|k|^{8p\tau+6}, 1\}}, \right.$$

$$\left. \textit{where } n, m \notin \mathcal{N} \textit{ or } n, m \in \mathcal{N} \right\},$$

$$\mathcal{R}_{20,2}^{\nu+1} = \bigcup_{K'_{\nu+1} \geqslant |k| > K_\nu, n, t} \left\{ \xi \in \mathcal{O}_\nu \colon \left| \langle k, \omega_{\nu+1} \rangle + \tilde{\Omega}_{\nu+1,n} + \tilde{\Omega}_{\nu+1,i_t} \right| < \frac{\epsilon_{\nu+1}^{\beta_0} (||i_t| - |n|| + 1)}{c \max\{|k|^{8p\tau+6}, 1\}}, \right.$$

$$\left. \textit{where } n \notin \mathcal{N},\ t \in \mathcal{T},\ |k_1 n_1 + k_2 n_2| = |n + i_t| \right\},$$

$$\mathcal{R}_{20,3}^{\nu+1} = \bigcup_{K'_{\nu+1} \geqslant |k| > K_\nu, n, t} \left\{ \xi \in \mathcal{O}_\nu \colon \left| \langle k, \omega_{\nu+1} \rangle + \tilde{\Omega}_{\nu+1,n} + \tilde{\Omega}_{\nu+1,j_t} \right| < \frac{\epsilon_{\nu+1}^{\beta_0} (||j_t| - |n|| + 1)}{c \max\{|k|^{8p\tau+6}, 1\}}, \right.$$

$$\left. \textit{where } n \notin \mathcal{N},\ t \in \mathcal{T},\ |k_1 n_1 + k_2 n_2| = \left| n + j_t + (n_1 - n_2)(p - t) \right| \right\},$$

$$\mathcal{R}_{11,1}^{\nu+1} = \bigcup_{K'_{\nu+1} \geqslant |k| > K_\nu, n, m} \left\{ \xi \in \mathcal{O}_\nu \colon \left| \langle k, \omega_{\nu+1} \rangle + \tilde{\Omega}_{\nu+1,n} - \tilde{\Omega}_{\nu+1,m} \right| < \frac{\epsilon_{\nu+1}^{\beta_0} (||n| - |m|| + 1)}{c \max\{|k|^{8p\tau+6}, 1\}}, \right.$$

$$\left. \textit{where } n, m \notin \mathcal{N},\ |k| + \left| |n| - |m| \right| \neq 0,\ |k_1 n_1 + k_2 n_2| = |n - m| \right\},$$

$$\mathcal{R}_{11,2}^{\nu+1} = \bigcup_{K'_{\nu+1} \geqslant |k| > K_\nu, n, m} \left\{ \xi \in \mathcal{O}_\nu \colon \left| \langle k, \omega_{\nu+1} \rangle + \tilde{\Omega}_{\nu+1,n} - \tilde{\Omega}_{\nu+1,m} \right| < \frac{\epsilon_{\nu+1}^{\beta_0} (||n| - |m|| + 1)}{c \max\{|k|^{8p\tau+6}, 1\}}, \right.$$

$$\left. \textit{where } n, m \in \mathcal{N},\ |k| + |n - m| \neq 0 \right\},$$

$$\mathcal{R}_{11,3}^{\nu+1} = \bigcup_{K'_{\nu+1} \geqslant |k| > K_\nu, n, t} \left\{ \xi \in \mathcal{O}_\nu \colon \left| \langle k, \omega_{\nu+1} \rangle + \tilde{\Omega}_{\nu+1,n} - \tilde{\Omega}_{\nu+1,i_t} \right| < \frac{\epsilon_{\nu+1}^{\beta_0} (||i_t| - |n|| + 1)}{c \max\{|k|^{8p\tau+6}, 1\}}, \right.$$

$$\left. \textit{where } n \notin \mathcal{N},\ t \in \mathcal{T},\ |k_1 n_1 + k_2 n_2| = |n - i_t| \right\},$$

$$\mathcal{R}_{11,4}^{\nu+1} = \bigcup_{K'_{\nu+1} \geqslant |k| > K_\nu, n, t} \left\{ \xi \in \mathcal{O}_\nu \colon \left| \langle k, \omega_{\nu+1} \rangle + \tilde{\Omega}_{\nu+1,n} - \tilde{\Omega}_{\nu+1,j_t} \right| < \frac{\epsilon_{\nu+1}^{\beta_0} (||j_t| - |n|| + 1)}{c \max\{|k|^{8p\tau+6}, 1\}}, \right.$$

$$\left. \textit{where } n \notin \mathcal{N},\ t \in \mathcal{T},\ |k_1 n_1 + k_2 n_2| = \left| n - j_t - (n_1 - n_2)(p - t) \right| \right\},$$

such that for $H_{\nu+1} = H_\nu \circ \Phi_{\nu+1} = N_{\nu+1} + P_{\nu+1}$ *the same assumptions are satisfied with* $\nu + 1$ *in place of* ν.

Proof. Note (3.16), by induction one verifies that

$$\frac{c\epsilon_\nu^{1-(8p+1)\beta_0}}{\eta_\nu^2\sigma_\nu^{1+\mu}} \leqslant 1,$$
$$c\epsilon_\nu K_\nu^{8p\tau+7} \leqslant \epsilon_\nu^{\beta_0} - \epsilon_{\nu+1}^{\beta_0}.$$

It is easy to check that (A.5) holds. From Lemma 3.4, we know $P_{\nu+1}$ has a generalized compact form with respect to n_1, n_2 and $\mathcal{J}$. For the remained proof, see Iterative lemma in [15]. □

With (3.14) and (3.15), we also obtain the following estimate.

Lemma 3.6.

$$\frac{1}{\sigma_\nu}|\Phi_{\nu+1} - \mathrm{id}|^*_{r_\nu, D_{\nu+1}}, \|D\Phi_{\nu+1} - I\|^*_{r_\nu, r_\nu, D_{\nu+1}} \leqslant \frac{cM(\nu)\epsilon_\nu^{1-(8p+1)\beta_0}}{\sigma_\nu^{\mu+1}}, \tag{3.17}$$

$$\|\omega_{\nu+1} - \omega_\nu\|^*_{\mathcal{O}_\nu}, \|\tilde{\Omega}_{\nu+1} - \tilde{\Omega}_\nu\|^*_{\mathcal{O}_{\nu+1}} \leqslant c\epsilon_\nu. \tag{3.18}$$

3.3. Convergence and proof of the existences of tori

Let $\Phi^\nu = \Phi_1 \circ \Phi_2 \circ \cdots \circ \Phi_\nu$, $\nu = 1, 2, \ldots$. Inductively, we have that $\Phi^\nu : D_\nu \times \mathcal{O}_{\nu-1} \to D_0$ and

$$H_0 \circ \Phi^\nu = H_\nu = N_\nu + P_\nu$$

for all $\nu \geqslant 1$.

Let $\tilde{\mathcal{O}}_\epsilon = \bigcap_{\nu=0}^\infty \mathcal{O}_\nu$. We apply Lemmas 3.5, 3.6 and standard arguments (see [15]) to conclude that H_ν, N_ν, P_ν, Φ^ν, $D\Phi^\nu$, ω_ν, $\tilde{\Omega}_{\nu,n}$ converge uniformly on $D(\frac{1}{2}s_0, 0) \times \tilde{\mathcal{O}}_\epsilon$, say to, H_∞, N_∞, P_∞, Φ^∞, $D\Phi^\infty$, ω_∞, $\tilde{\Omega}_{\infty,n}$, respectively. It is clear that

$$N_\infty = \langle \omega_\infty, y\rangle + \sum_n \tilde{\Omega}_{\infty,n} w_n \bar{w}_n.$$

Further, we have

$$|X_{P_\infty}|_{D(\frac{1}{2}s_0,0)\times\tilde{\mathcal{O}}} \equiv 0.$$

Let Φ_H^t denote the flow of any Hamiltonian vector field X_H. Since $H_0 \circ \Phi^\nu = H_\nu$, we have that

$$\Phi_{H_0}^t \circ \Phi^\nu = \Phi^\nu \circ \Phi_{H_\nu}^t.$$

The uniform convergence of Φ^ν, $D\Phi^\nu$, X_{H_ν} imply that one can pass the limit in the above to conclude that

$$\Phi_{H_0}^t \circ \Phi^\infty = \Phi^\infty \circ \Phi_{H_\infty}^t$$

on $D(\frac{1}{2}s_0, 0) \times \tilde{\mathcal{O}}_\epsilon$. It follows that

$$\Phi^t_{H_0}\big(\Phi^\infty\big(\mathbb{T}^2 \times \{\xi\}\big)\big) = \Phi^\infty \Phi^t_{N_\infty}\big(\mathbb{T}^2 \times \{\xi\}\big) = \Phi^\infty\big(\mathbb{T}^2 \times \{\xi\}\big)$$

for all $\xi \in \tilde{\mathcal{O}}_\epsilon$. Hence $\Phi^\infty(\mathbb{T}^2 \times \{\xi\})$ is an embedded invariant torus of the original perturbed Hamiltonian system at $\xi \in \tilde{\mathcal{O}}_\epsilon$. We remark that the frequencies $\omega_\infty(\xi)$ associated with $\Phi^\infty(\mathbb{T}^2 \times \{\xi\})$ are slightly deformed from the unperturbed ones $\omega(\xi)$. The normal behaviors of the invariant tori $\Phi^\infty(\mathbb{T}^2 \times \{\xi\})$ are governed by their respective normal frequencies $\tilde{\Omega}_{\infty,n}(\xi)$.

In fact, combining with Sections 3 and 4 below, we have the following theorem.

Theorem 2. *For the Hamiltonian* (2.24)

$$\begin{aligned} H &= N + P \\ &= \langle \omega, y\rangle + \sum_j \tilde{\Omega}_j w_n \bar{w}_n + P(x, y, w, \bar{w}, \xi, \epsilon), \end{aligned}$$

and P satisfies a generalized compact form with respect to n_1, n_2 and $\mathcal{J}$. Suppose that

$$|X_P|^*_{r,D(s,r)} = \epsilon \leqslant \gamma s^{6(1+\mu)}, \tag{3.19}$$

where γ depends on p, τ and M. Then there exist a Cantor set $\tilde{\mathcal{O}}_\epsilon \subset \mathcal{O} = [1, 2]^2$ with the measure satisfying

$$|\mathcal{O} \setminus \tilde{\mathcal{O}}_\epsilon| \leqslant \cdot \epsilon^{\frac{1}{4p(8p+1)}},$$

a Whitney smooth family of torus embeddings $\Phi : \mathbb{T}^2 \times \tilde{\mathcal{O}}_\epsilon \to \mathbb{P}$, and a Whitney smooth map $\omega_\infty : \tilde{\mathcal{O}}_\epsilon \to \mathbb{R}^2$, such that for each $\xi \in \tilde{\mathcal{O}}_\epsilon$, the map Φ restricted to $\mathbb{T}^2 \times \{\xi\}$ is a real analytic embedding of a rotational torus with frequencies $\omega_\infty(\xi)$ for the Hamiltonian H at ξ.

Each embedding is real analytic on $|\mathrm{Im}\, x| < \frac{s}{2}$, and

$$\begin{aligned} \|\Phi - \Phi_0\|^*_r &\leqslant c\epsilon^{\frac{1}{3}}, \\ \|\omega_* - \omega\|^* &\leqslant c\epsilon \end{aligned}$$

uniformly on that domain and $\tilde{\mathcal{O}}_\epsilon$, where Φ_0 is the trivial embedding $\mathbb{T}^2 \times \mathcal{O} \to \mathcal{T}^2_0$.

Remark 3.1. For the estimates of $\tilde{\mathcal{O}}_\epsilon$, see Section 4 for details.

Remark 3.2. Theorem 1 is a direct result of Theorem 2. For more specific, please refer to the standard proof of [12].

4. Measure estimates

4.1. Measure estimates in the first step

For simplicity, in this section we will denote

$$\lambda_0 = \left(n_1^2, n_2^2\right),$$
$$f_1 = \sum_{k=0}^{p} c_p \left(C_{p+1}^k\right)^2 C_{p+1-k}^1 \xi_1^{2p-2k} \xi_2^{2k},$$
$$f_2 = \sum_{k=0}^{p} c_p \left(C_{p+1}^{k+1}\right)^2 C_{k+1}^1 \xi_1^{2p-2k} \xi_2^{2k},$$
$$f_3 = \sum_{k=0}^{p} c_p \left(C_{p+1}^1\right)^2 \left(C_p^k\right)^2 \xi_1^{2p-2k} \xi_2^{2k}.$$

At the first KAM step, we have to exclude the following resonant set

$$\mathcal{R}^0 = \mathcal{R}_0^0 \cup \left(\mathcal{R}_{1,1}^0 \cup \mathcal{R}_{1,2}^0\right) \cup \mathcal{R}_2^0,$$

where

$$\mathcal{R}_0^0 = \bigcup_{0<|k|\leqslant K_0} \left\{\xi \in \mathcal{O}\colon \left|\langle k, \omega(\xi)\rangle\right| < \frac{\epsilon^{\frac{\beta_0}{4}}}{|k|^{2p\tau}}\right\}, \tag{4.1}$$

$$\mathcal{R}_{1,1}^0 = \bigcup_{\substack{n\notin\mathcal{N}\\ |k|\leqslant K_0}} \left\{\xi \in \mathcal{O}\colon \left|\langle k, \omega(\xi)\rangle + \tilde{\Omega}_n\right| < \frac{\epsilon^{\frac{\beta_0}{4}}}{\max\{1, |k|^{2p\tau}\}}\right\}, \tag{4.2}$$

$$\mathcal{R}_{1,2}^0 = \bigcup_{t\in\mathcal{T},\, |k|\leqslant K_0} \left\{\xi \in \mathcal{O}\colon |g_1| < \frac{\epsilon^{\frac{\beta_0}{2}}}{\max\{1, |k|^{4p\tau}\}}\right\}, \tag{4.3}$$
$$g_1 = \det M_1',$$

and

$$M_1' = \begin{pmatrix} k_1 f_1 + k_2 f_2 + f_3 & a_t \\ a_t & k_1 f_1 + k_2 f_2 + f_3 + (p-t)A \end{pmatrix}.$$

$$\mathcal{R}_{20,1}^0 = \bigcup_{\substack{n,m\notin\mathcal{N}\\ |k|\leqslant K_0}} \left\{\xi \in \mathcal{O}\colon \left|\langle k, \omega(\xi)\rangle + \tilde{\Omega}_n + \tilde{\Omega}_m\right| < \frac{\epsilon^{\frac{\beta_0}{4}}(||n|-|m||+1)}{\max\{1, |k|^{2p\tau}\}}\right\}; \tag{4.4}$$

$$\mathcal{R}_{20,2}^0 = \bigcup_{\substack{t\in\mathcal{T}\\ |k|\leqslant K_0}} \left\{\xi \in \mathcal{O}\colon |g_2| < \frac{\epsilon^{\frac{\beta_0}{2}}}{\max\{1, |k|^{4p\tau}\}}\right\}, \tag{4.5}$$

where

$$g_2 = \det M_2',$$

and

$$M_2' = \begin{pmatrix} k_1 f_1 + k_2 f_2 + 2 f_3 & a_t \\ a_t & k_1 f_1 + k_2 f_2 + 2 f_3 + (p-t)A \end{pmatrix};$$

$$\mathcal{R}^0_{20,3} = \bigcup_{\substack{t_1, t_2 \in \mathcal{T},\, |k| \leqslant K_0 \\ \langle k, \lambda_0\rangle + i_{t_1}^2 + i_{t_2}^2 = 0}} \left\{ \xi \in \mathcal{O}\colon\ |g_3| < \frac{\epsilon^{\beta_0}}{\max\{1, |k|^{8p\tau}\}} \right\}, \tag{4.6}$$

where

$$g_3 = \det M_3', \qquad \Delta_4 = k_1 f_1 + k_2 f_2 + 2 f_3,$$

and

$$M_3' = \begin{pmatrix} \Delta_4 & a_{t_2} & a_{t_1} & 0 \\ a_{t_2} & \Delta_4 + (p-t_2)A & 0 & a_{t_1} \\ a_{t_1} & 0 & \Delta_4 + (p-t_1)A & a_{t_2} \\ 0 & a_{t_1} & a_{t_2} & \Delta_4 + (2p - t_2 - t_1)A \end{pmatrix};$$

$$\mathcal{R}^0_{20,4} = \bigcup_{\substack{|k| + ||n| - |m|| \neq 0 \\ n, m \notin \mathcal{N},\, |k| \leqslant K_0}} \left\{ \xi \in \mathcal{O}\colon\ \left|\left\langle k, \omega(\xi)\right\rangle + \tilde{\Omega}_n - \tilde{\Omega}_m\right| < \frac{\epsilon^{\frac{\beta_0}{4}}(||n| - |m|| + 1)}{\max\{1, |k|^{2p\tau}\}} \right\}, \tag{4.7}$$

$$\mathcal{R}^0_{20,5} = \bigcup_{t \in \mathcal{T}, k} \left\{ \xi \in \mathcal{O}\colon\ |g_4| < \frac{\epsilon^{\frac{\beta_0}{2}}}{\max\{1, |k|^{4p\tau}\}} \right\}, \tag{4.8}$$

where

$$g_4 = \det M_4'$$

and

$$M_4' = \begin{pmatrix} k_1 f_1 + k_2 f_2 & -a_t \\ -a_t & k_1 f_1 + k_2 f_2 - (p-t)A \end{pmatrix};$$

$$\mathcal{R}^0_{20,6} = \bigcup_{\substack{|k| + |t_1 - t_2| \neq 0,\, |k| \leqslant K_0 \\ \langle k, \lambda_0\rangle + i_{t_1}^2 - i_{t_2}^2 = 0,\, t_1, t_2 \in \mathcal{T}}} \left\{ \xi \in \mathcal{O}\colon\ |g_5| < \frac{\epsilon^{\beta_0}}{\max\{1, |k|^{8p\tau}\}} \right\}, \tag{4.9}$$

where

$$g_5 = \det M_5', \qquad \Delta_5 = k_1 f_1 + k_2 f_2,$$

and

$$M_5' = \begin{pmatrix} \Delta_5 & -a_{t_2} & a_{t_1} & 0 \\ -a_{t_2} & \Delta_5 - (p-t_2)A & 0 & a_{t_1} \\ a_{t_1} & 0 & \Delta_5 + (p-t_1)A & -a_{t_2} \\ 0 & a_{t_1} & -a_{t_2} & \Delta_5 + (t_2-t_1)A \end{pmatrix}.$$

The following lemma is used many times in this section. We will not point out it clearly.

Lemma 4.1. *Suppose that $g(x)$ is an mth differentiable function on the closure $\bar{I}$ of I, where $I \subset \mathbb{R}$ is an interval. Let $I_h = \{x \mid |g(x)| < h\}$, $h > 0$. If for some constant $d > 0$, $|g^m(x)| \geqslant d$ for any $x \in I$, then $|I_h| \leqslant ch^{\frac{1}{m}}$, where $|I_h|$ denotes the Lebesgue measure of I_h and $c = 2(2 + 3 + \cdots + m + d^{-1})$.*

For the proof see [16]. The similar method can be found in [17].

Since the proofs for the next nine lemmas are similar, we only give one of them and omit the others.

Lemma 4.2. *If $\tau > 2$, $|\mathcal{R}_0^0| \leqslant \cdot \epsilon^{\frac{\beta_0}{8p}}$.*

Lemma 4.3. *If $\tau > 3$, $|\mathcal{R}_{1,1}^0| \leqslant \cdot \epsilon^{\frac{\beta_0}{8p}}$.*

Lemma 4.4. *If $\tau > 2$, $|\mathcal{R}_{1,2}^0| \leqslant \cdot \epsilon^{\frac{\beta_0}{8p}}$.*

Lemma 4.5. *If $\tau > 5$, $|\mathcal{R}_{20,1}^0| \leqslant \cdot \epsilon^{\frac{\beta_0}{8p}}$.*

Lemma 4.6. *If $\tau > 2$, $|\mathcal{R}_{20,2}^0| \leqslant \cdot \epsilon^{\frac{\beta_0}{8p}}$.*

Lemma 4.7. *If $\tau > 2$, $|\mathcal{R}_{20,3}^0| \leqslant \cdot \epsilon^{\frac{\beta_0}{8p}}$.*

Proof. The difficult point in this proof lies in whether there exist nonzero coefficients in g_3 for any k, $t_1, t_2 \in \mathcal{T}$ and $\langle k, \lambda_0 \rangle + i_{t_1}^2 + i_{t_2}^2 = 0$. We will show this in the following. Write $\Xi_1 = k_1 + (k_2+2)(p+1)$, $\Xi_2 = k_1(p+1) + k_2 + 2$, $\Xi_3 = k_1 + (k_2+2)\frac{p}{2}$, $\Xi_4 = k_1\frac{p}{2} + (k_2+2)$. It is easy to check that

$$\begin{aligned} g_3^{8p,0} = {} & c\Xi_1\big\{\Xi_1^3 - p(4p - 2t_1 - 2t_2)\Xi_1^2 \\ & + p^2\big[(p-t_2)(3p-2t_1-t_2) + (p-t_1)(2p-t_1-t_2)\big]\Xi_1 \\ & - p^3(p-t_1)(p-t_2)(2p-t_1-t_2)\big\}, \end{aligned} \tag{4.10}$$

$$\begin{aligned} g_3^{0,8p} = {} & c\Xi_2\big\{\Xi_2^3 + p(4p - 2t_1 - 2t_2)\Xi_2^2 \\ & + p^2\big[(p-t_2)(3p-2t_1-t_2) + (p-t_1)(2p-t_1-t_2)\big]\Xi_2 \\ & + p^3(p-t_1)(p-t_2)(2p-t_1-t_2)\big\}, \end{aligned} \tag{4.11}$$

$$g_3^{8p-2,2} = c\left\{4\Xi_1^3\Xi_3 - 3p\Xi_1^2\Xi_3(4p-2t_1-2t_2) + (4p-2t_1-2t_2)\left(1-\frac{1}{2}p\right)\Xi_1^3\right.$$
$$+\left[(p-t_1)(2p-t_1-t_2)+(p-t_2)(3p-2t_1-t_2)\right]\left(2p^2\Xi_1\Xi_3 - 2p\left(1-\frac{p}{2}\right)\Xi_1^2\right)$$
$$\left.+(p-t_1)(p-t_2)(2p-t_1-t_2)\left[3p^2\left(1-\frac{p}{2}\right)\Xi_1 - p^3\Xi_3\right]\right\}, \tag{4.12}$$

$$g_3^{2,8p-2} = c\left\{4\Xi_2^3\Xi_4 + 3p\Xi_2^2\Xi_4(4p-2t_1-2t_2) - (4p-2t_1-2t_2)\left(1-\frac{1}{2}p\right)\Xi_2^3\right.$$
$$+\left[(p-t_1)(2p-t_1-t_2)+(p-t_2)(3p-2t_1-t_2)\right]\left(2p^2\Xi_2\Xi_4 - 2p\left(1-\frac{p}{2}\right)\Xi_2^2\right)$$
$$\left.+(p-t_1)(p-t_2)(2p-t_1-t_2)\left[3p^2\left(\frac{p}{2}-1\right)\Xi_2 + p^3\Xi_4\right]\right\}. \tag{4.13}$$

If $g_3^{8p,0}=0$ and $g_3^{0,8p}=0$ for some $k, t_1, t_2 \in \mathcal{T}$ and $\langle k, \lambda_0\rangle + i_{t_1}^2 + i_{t_2}^2 = 0$, one has 16 cases.

Case 1. $\begin{cases}\Xi_1 = 0,\\ \Xi_2 = 0.\end{cases}$ One has $\begin{cases}k_1 = 0,\\ k_2 = -2\end{cases}$ in this case.

Case 2. $\begin{cases}\Xi_1 = 0,\\ \Xi_2 = -p(2p-t_1-t_2).\end{cases}$ One has $k_2 \notin \mathbb{Z}$ or $\begin{cases}k_1 = -p-1,\\ k_2 = -1.\end{cases}$

Cases 3, 4. $\begin{cases}\Xi_1 = 0,\\ \Xi_2 = -p(p-t_1) \text{ or } -p(p-t_2).\end{cases}$ It is easy.

Case 5. $\begin{cases}\Xi_1 = p(2p-t_1-t_2),\\ \Xi_2 = 0.\end{cases}$ One has $k_1 \notin \mathbb{Z}$ or $\begin{cases}k_1 = -1,\\ k_2 = p-1.\end{cases}$

Case 6. $\begin{cases}\Xi_1 = p(2p-t_1-t_2),\\ \Xi_2 = -p(2p-t_1-t_2).\end{cases}$ One has $\begin{cases}k_1 = t_1+t_2-2p,\\ k_2 = -2-t_1-t_2+2p.\end{cases}$ Note $n_2 > \sqrt{p}n_1$, this leads to $\langle k, \lambda_0\rangle + i_{t_1}^2 + i_{t_2}^2 \neq 0$. It is a contradiction.

Cases 7–10. $\begin{cases}\Xi_1 = p(2p-t_1-t_2),\\ \Xi_2 = -p(p-t_1)\end{cases}$ or $\begin{cases}\Xi_1 = p(2p-t_1-t_2),\\ \Xi_2 = -p(p-t_2)\end{cases}$ or $\begin{cases}\Xi_1 = p(p-t_1),\\ \Xi_2 = 0\end{cases}$ or $\begin{cases}\Xi_1 = p(p-t_1),\\ \Xi_2 = -p(2p-t_1-t_2).\end{cases}$ It is easy to get $k_1 \notin \mathbb{Z}$.

Cases 11–14. $\begin{cases}\Xi_1 = p(p-t_1),\\ \Xi_2 = -p(p-t_1)\end{cases}$ or $\begin{cases}\Xi_1 = p(2p-t_1-t_2),\\ \Xi_2 = -p(p-t_2)\end{cases}$ or $\begin{cases}\Xi_1 = p(p-t_2),\\ \Xi_2 = -p(p-t_1)\end{cases}$ or $\begin{cases}\Xi_1 = p(p-t_2),\\ \Xi_2 = -p(p-t_2).\end{cases}$ The four cases are similar as Case 6.

Cases 15, 16. $\begin{cases}\Xi_1 = p(p-t_2),\\ \Xi_2 = 0 \text{ or } -p(2p-t_1-t_2).\end{cases}$ It is easy.

The above proof shows that except the following 3 cases we have

$$\left(g_3^{8p,0}\right)^2 + \left(g_3^{0,8p}\right)^2 \neq 0,$$

which are

$$(1') \quad \begin{cases} k_1 = 0, \\ k_2 = -2, \end{cases} \qquad (2') \quad \begin{cases} k_1 = -1 - p, \\ k_2 = -1, \end{cases} \qquad (3') \quad \begin{cases} k_1 = -1, \\ k_2 = p - 1. \end{cases}$$

Checking directly, it is easy to know for Cases (2′) and (3′), we have

$$\left(g_3^{8p-2,2}\right)^2 + \left(g_3^{2,8p-2}\right)^2 \neq 0.$$

The only remaining case is $\begin{cases} k_1 = 0, \\ k_2 = -2. \end{cases}$ In this case, it is clear

$$g_3(\xi) = \left(a_{t_2}^4 - a_{t_1}^4\right) - (p - t_1)(2p - t_1 - t_2)A^2 a_{t_2}^2.$$

In fact, it is easy to check that

$$g_3^{6p+2t_2,2p-2t_2} \neq 0. \qquad \square$$

Lemma 4.8. *If* $\tau > 5$, $|\mathcal{R}^0_{20,4}| \leqslant \cdot\epsilon^{\frac{\beta_0}{8p}}$.

Lemma 4.9. *If* $\tau > 2$, $|\mathcal{R}^0_{20,5}| \leqslant \cdot\epsilon^{\frac{\beta_0}{8p}}$.

Lemma 4.10. *If* $\tau > 2$, $|\mathcal{R}^0_{20,6}| \leqslant \cdot\epsilon^{\frac{\beta_0}{8p}}$.

Combined with above lemmas, we have the following lemma.

Lemma 4.11. *If* $\tau > 5$, $|\mathcal{R}^0| \leqslant \cdot\epsilon^{\frac{\beta_0}{8p}}$.

In the following, we will give a description lemma about the remaining set $\mathcal{O}_0 = \mathcal{O} \setminus \mathcal{R}^0$.

Lemma 4.12. *For* $|k| \leqslant K_0$ *and all the parameters* $\xi \in \mathcal{O}$*, which belong to the set* $\mathcal{O}_0 = \mathcal{O} \setminus \mathcal{R}^0$*, satisfy*[4] *the following conditions*

$$\left|\langle k, \omega\rangle^{-1}\right| \leqslant \frac{|k|^{2p\tau}}{\epsilon^{\frac{\beta_0}{4}}}, \quad k \neq 0,$$

$$\left|\left(\langle k, \omega\rangle + \Omega_n\right)^{-1}\right| \leqslant \frac{c \max\{|k|^{4p\tau+2}, 1\}}{\epsilon^{\frac{\beta_0}{2}}}, \quad n \notin \mathcal{N},$$

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

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

$\|\cdot\|$ for matrix denotes the operator norm, i.e., $\|M\| = \sup_{|y|=1} |My|$.

$$\left\|\left(\langle k,\omega\rangle I_2 + A_{i_t}\right)^{-1}\right\| \leqslant \frac{c\max\{|k|^{4p\tau+2},1\}}{\epsilon^{\frac{\beta_0}{2}}}, \quad t\in\mathcal{T}, \tag{4.14}$$

$$\left|\left(\langle k,\omega\rangle + \Omega_n + \Omega_m\right)^{-1}\right| \leqslant \frac{c\max\{|k|^{8p\tau+6},1\}}{\epsilon^{\beta_0}(||n|-|m||+1)}, \quad n,m\notin\mathcal{N},$$

$$\left\|\left(\left(\langle k,\omega\rangle + \Omega_n\right)I_2 + A_{i_t}\right)^{-1}\right\| \leqslant \frac{c\max\{|k|^{8p\tau+6},1\}}{\epsilon^{\beta_0}(||i_{t_1}|-|n||+1)}, \tag{4.15}$$

where $|k_1n_1+k_2n_2| = |n+i_t|$ *or* $|k_1n_1+k_2n_2| = |n+j_t+(n_1-n_2)(p-t)|$, $n\notin\mathcal{N}$, $t\in\mathcal{T}$,

$$\left\|\left(I_2\otimes\left(\langle k,\omega\rangle I_2 + A_{i_{t_2}}\right) + A_{i_{t_1}}\otimes I_2\right)^{-1}\right\| \leqslant \frac{c\max\{|k|^{8p\tau+6},1\}}{\epsilon^{\beta_0}(||i_{t_1}|-|i_{t_2}||+1)}, \quad t_1,t_2\in\mathcal{T}, \tag{4.16}$$

$$\left|\left(\langle k,\omega\rangle + \Omega_n - \Omega_m\right)^{-1}\right| \leqslant \frac{c\max\{|k|^{8p\tau+6},1\}}{\epsilon^{\beta_0}(||n|-|m||+1)}, \quad n,m\notin\mathcal{N},\ |k| + \left||n|-|m|\right| \neq 0,$$

where $|k_1n_1+k_2n_2| = |n-m|$,

$$\left\|\left(I_2\otimes\left(\langle k,\omega\rangle I_2 - A_{i_{t_2}}\right) + A_{i_{t_1}}\otimes I_2\right)^{-1}\right\| \leqslant \frac{c\max\{|k|^{8p\tau+6},1\}}{\epsilon^{\beta_0}(||i_{t_1}|-|i_{t_2}||+1)}, \tag{4.17}$$

where $t_1,t_2\in\mathcal{T}$, $|k|+|t_1-t_2|\neq 0$,

$$\left\|\left(\left(\langle k,\omega\rangle + \Omega_n\right)I_2 - A_{i_t}\right)^{-1}\right\| \leqslant \frac{c\max\{|k|^{8p\tau+6},1\}}{\epsilon^{\beta_0}(||i_t|-|n||+1)}, \tag{4.18}$$

where $|k_1n_1+k_2n_2| = |n-i_t|$ *or* $|k_1n_1+k_2n_2| = |n-j_t-(n_1-n_2)(p-t)|$, $n\notin\mathcal{N}$, $t\in\mathcal{T}$.

The proof is given in Appendix A.

Remark 4.1. We must point out that Lemma 4.12 omits one inequality of (3.6), which is

$$|\tilde{\Omega}_{i_t} - \tilde{\Omega}_{j_t}| \leqslant \frac{c}{\epsilon^{\beta_0}(||i_t|-|j_t||+1)}, \quad t\in\mathcal{T}. \tag{4.19}$$

But, from

$$|\tilde{\Omega}_{i_t} - \tilde{\Omega}_{j_t}| = \left|\frac{1}{\sqrt{4a_t^2+(p-t)^2A^2}}\right| \leqslant c,$$

it is easy to know that (4.19) holds naturally.

Remark 4.2. From (4.18) to the corresponding inequalities in (3.6), one inequality is needed. We need the simple inequality as the following:

$$\frac{1}{||i_t|-|n||+1} \leqslant \cdot\frac{1}{||j_t|-|n||+1}. \tag{4.20}$$

Remark 4.3. (3.6) is a direct result from Lemma 4.12 and above two remarks.

4.2. Measure estimates for remaining steps

From Lemma 3.5, we have to exclude the following resonant set

$$\mathcal{R}^{\nu+1} = \mathcal{R}_{00}^{\nu+1} \cup \mathcal{R}_{10}^{\nu+1} \cup \mathcal{R}_{20}^{\nu+1} \cup \mathcal{R}_{11}^{\nu+1},$$
$$\mathcal{R}_{20}^{\nu+1} = \mathcal{R}_{20,1}^{\nu+1} \cup \mathcal{R}_{20,2}^{\nu+1} \cup \mathcal{R}_{20,3}^{\nu+1},$$
$$\mathcal{R}_{11}^{\nu+1} = \mathcal{R}_{11,1}^{\nu+1} \cup \mathcal{R}_{11,2}^{\nu+1} \cup \mathcal{R}_{11,3}^{\nu+1} \cup \mathcal{R}_{11,4}^{\nu+1}$$

(where $\nu \geqslant 0$) at remaining KAM steps. We have the following lemmas which give the corresponding measure estimates. The proofs of the following lemmas are similar with [13] and we omit them.

Lemma 4.13. *If* $\tau > 1$ *and* $K_\nu > \frac{8c}{c_0}$, $|\mathcal{R}_{00}^{\nu+1}| \leqslant \cdot\epsilon_{\nu+1}^{\frac{\beta_0}{2p}}$.

Lemma 4.14. *If* $\tau > 1$ *and* $K_\nu > \frac{8c}{c_0}$, $|\mathcal{R}_{10}^{\nu+1}| \leqslant \cdot\epsilon_{\nu+1}^{\frac{\beta_0}{2p}}$.

Lemma 4.15. *If* $\tau > 2$ *and* $K_\nu > \frac{8c}{c_0}$, $|\mathcal{R}_{20,1}^{\nu+1}| \leqslant \cdot\epsilon_{\nu+1}^{\frac{\beta_0}{2p}}$.

Lemma 4.16. *If* $\tau > 2$ *and* $K_\nu > \frac{8c}{c_0}$, $|\mathcal{R}_{20,2}^{\nu+1}| \leqslant \cdot\epsilon_{\nu+1}^{\frac{\beta_0}{2p}}$.

Lemma 4.17. *If* $\tau > 2$ *and* $K_\nu > \frac{8c}{c_0}$, $|\mathcal{R}_{20,3}^{\nu+1}| \leqslant \cdot\epsilon_{\nu+1}^{\frac{\beta_0}{2p}}$.

Lemma 4.18. *If* $\tau > 2$ *and* $K_\nu > \frac{8c}{c_0}$, $|\mathcal{R}_{11,1}^{\nu+1}| \leqslant \cdot\epsilon_{\nu+1}^{\frac{\beta_0}{2p}}$.

Lemma 4.19. *If* $\tau > 2$ *and* $K_\nu > \frac{8c}{c_0}$, $|\mathcal{R}_{11,2}^{\nu+1}| \leqslant \cdot\epsilon_{\nu+1}^{\frac{\beta_0}{2p}}$.

Lemma 4.20. *If* $\tau > 2$ *and* $K_\nu > \frac{8c}{c_0}$, $|\mathcal{R}_{11,3}^{\nu+1}| \leqslant \cdot\epsilon_{\nu+1}^{\frac{\beta_0}{2p}}$.

Lemma 4.21. *If* $\tau > 2$ *and* $K_\nu > \frac{8c}{c_0}$, $|\mathcal{R}_{11,4}^{\nu+1}| \leqslant \cdot\epsilon_{\nu+1}^{\frac{\beta_0}{2p}}$.

Combining with all the lemmas above, we have

Lemma 4.22. *If* $\tau > 2$ *and* $K_\nu > \frac{8c}{c_0}$, $|\mathcal{R}^{\nu+1}| \leqslant \cdot\epsilon_{\nu+1}^{\frac{\beta_0}{2p}}$ $(\nu \geqslant 0)$.

Note (3.16), this means $K_0 > \frac{8c}{c_0}$. Fix $\tau > 5$. Now we compute the total measure of the parameter sets $\mathcal{R}_\epsilon$ which be thrown in all the steps,

$$
\begin{aligned}
|\mathcal{R}_\epsilon| &\leqslant \cdot\epsilon_0^{\frac{\beta_0}{2p}} + \cdot\epsilon_1^{\frac{\beta_0}{2p}} + \cdots \\
&\leqslant \cdot\epsilon_0^{\frac{\beta_0}{2p}} = \cdot\epsilon_0^{\frac{1}{4p(8p+1)}}.
\end{aligned}
$$

Acknowledgments

The author thanks Professor Xiaoping Yuan and Dr. Jun Yan for their help. The author also thanks Professor Jiangong You and Huawei Niu for several helpful conversations about this paper.

Appendix A

A.1. Compact form and generalized compact form

Given $n_1, n_2 \in \mathbb{Z}$, $n_1 \neq n_2$. A real analytic function

$$F = F(x, y, z, \bar{z}) = \sum_{k,\alpha,\beta} F_{k\alpha\beta}(y) e^{\mathrm{i}\langle k,x\rangle} z^\alpha \bar{z}^\beta$$

on $D(r,s) = \{(x, y, z, \bar{z})\colon\ |\operatorname{Im} x| < s,\ |y| < r^2,\ \|z\|_\rho < r,\ \|\bar{z}\|_\rho < r\}$ is said to admit a compact form with respect to n_1, n_2, if

$$F_{k\alpha\beta} \neq 0 \quad \text{implies} \quad k_1 n_1 + k_2 n_2 + \sum_n (-\alpha_n + \beta_n) n = 0 \quad \text{for any } k, \alpha, \beta,$$

where $k = (k_1, k_2) \in \mathbb{Z}^2$ and $\alpha \equiv (\ldots, \alpha_n, \ldots)$, $\beta \equiv (\ldots, \beta_n, \ldots)$, $\alpha, \beta \in \mathbb{N}_0^\infty$, with finitely many non-vanishing components.

Lemma A.1. *Given $n_1, n_2 \in \mathbb{Z}$ and $n_1 \neq n_2$, consider two real analytic functions*

$$F(x, y, z, \bar{z}), \quad G(x, y, z, \bar{z})$$

on $D(r,s)$. If both F and G have compact forms with respect to n_1, n_2, so does $\{F, G\}$.

For the proof, refer to Lemma 2.4 in [11].

Given n_1, n_2 and specially chosen subscripts set $\mathcal{J} = \{j_0, \ldots, j_{p-2}\}$ and $j_t \notin \{n_1,\ n_2\}$, $t \in \mathcal{T}$. A real analytic function

$$F = F(x, y, z, \bar{z}) = \sum_{k,\alpha,\beta} F_{k\alpha\beta}(y) e^{\mathrm{i}\langle k,x\rangle} z^\alpha \bar{z}^\beta$$

on $D(r,s)$ is said to admit a generalized compact form with respect to n_1, n_2 and $\mathcal{J}$ if

$$F_{k\alpha\beta}(y) \neq 0$$

implies

$$k_1n_1+k_2n_2+\sum_{n\in\mathbb{Z}\setminus\{n_1,n_2\}}(-\alpha_n+\beta_n)n=(n_1-n_2)\sum_{t=0}^{p-2}(\alpha_{j_t}-\beta_{j_t})(p-t) \tag{A.1}$$

for any k, α, β, where $k=(k_1,k_2)\in\mathbb{Z}^2$ and $\alpha\equiv(\dots,\alpha_n,\dots)$, $\beta\equiv(\dots,\beta_n,\dots)$, $\alpha,\beta\in\mathbb{N}_0^\infty$, with finitely many non-vanishing components.

Similar as Lemma A.1, we have the following lemma.

Lemma A.2. *Given $n_1,n_2\in\mathbb{Z}$ and specially chosen subscripts set $\mathcal{J}=\{j_0,\dots,j_{p-2}\}$ and $j_t\notin\{n_1,\ n_2\}$, $t\in\mathcal{T}$. Consider two real analytic functions $F(x,y,z,\bar z)$, $G(x,y,z,\bar z)$ on $D(r,s)$. If both F and G have generalized compact forms with respect to n_1,n_2 and $\mathcal{J}$, so does $\{F,G\}$.*

For the proof, refer to Lemma 2.4 in [11].

The following lemma is needed in Section 2.

Lemma A.3. *P_0^+ satisfies a generalized compact form with respect to n_1, n_2 and $\mathcal{J}$.*

Proof. Write

$$P_t=\begin{pmatrix} p_{11,t} & p_{12,t}\\ p_{21,t} & p_{22,t}\end{pmatrix},$$

where $t\in\mathcal{T}$. As we know,

$$\begin{aligned}P_0^+=\sum_{k,\alpha,\beta}&P_{0,k\alpha\beta}\big(y^+\big)e^{\mathrm{i}\langle k,x^+\rangle}\big(\Pi_{i\notin\mathcal{N}}w_i^{\alpha_i}\bar w_i^{\beta_i}\big)\big(p_{11,0}w_{i_0}^++p_{12,0}w_{j_0}^+\big)^{\alpha_{i_0}}\\
&\cdot\big(p_{21,0}w_{i_0}^++p_{22,0}w_{j_0}^+\big)^{\alpha_{j_0}}\cdots\big(p_{11,p-2}w_{i_{p-2}}^++p_{12,p-2}w_{j_{p-2}}^+\big)^{\alpha_{i_{p-2}}}\\
&\cdot\big(p_{21,p-2}w_{i_{p-2}}^++p_{22,p-2}w_{j_{p-2}}^+\big)^{\alpha_{j_{p-2}}}\big(p_{11,0}\bar w_{i_0}^++p_{12,0}\bar w_{j_0}^+\big)^{\beta_{i_0}}\\
&\cdot\big(p_{21,0}\bar w_{i_0}^++p_{22,0}\bar w_{j_0}^+\big)^{\beta_{j_0}}\cdots\big(p_{11,p-2}\bar w_{i_{p-2}}^++p_{12,p-2}\bar w_{j_{p-2}}^+\big)^{\beta_{i_{p-2}}}\\
&\cdot\big(p_{21,p-2}\bar w_{i_{p-2}}^++p_{22,p-2}\bar w_{j_{p-2}}^+\big)^{\beta_{j_{p-2}}}.\end{aligned} \tag{A.2}$$

If $P_{0,k\alpha\beta}(y^+)=P_{0,k\alpha\beta}(y)\neq0$, then

$$k_1n_1+k_2n_2+\sum_{i\in\mathbb{Z}}(-\alpha_i+\beta_i)i=(n_1-n_2)\sum_{t=0}^{p-2}(\alpha_{j_t}-\beta_{j_t})(p-t). \tag{A.3}$$

We write every term of which its coefficient might be nonzero in (A.2). It is

$$\begin{aligned}&P_{0,k\alpha\beta}\big(y^+\big)e^{\mathrm{i}\langle k,x^+\rangle}\big(\Pi_{i\notin\mathcal{N}}\big(w_i^+\big)^{\alpha_i}\big(\bar w_i^+\big)^{\beta_i}\big)\big(w_{i_0}^+\big)^{k_0^1}\big(w_{j_0}^+\big)^{\alpha_{i_0}-k_0^1}\big(w_{i_0}^+\big)^{k_0^2}\big(w_{j_0}^+\big)^{\alpha_{j_0}-k_0^2}\cdots\\
&\quad\cdot\big(w_{i_{p-2}}^+\big)^{k_{p-2}^1}\big(w_{j_{p-2}}^+\big)^{\alpha_{i_{p-2}}-k_{p-2}^1}\big(w_{i_{p-2}}^+\big)^{k_{p-2}^2}\big(w_{j_{p-2}}^+\big)^{\alpha_{j_{p-2}}-k_{p-2}^2}\\
&\quad\cdot\big(\bar w_{i_0}^+\big)^{l_0^1}\big(\bar w_{j_0}^+\big)^{\beta_{i_0}-l_0^1}\big(\bar w_{i_0}^+\big)^{l_0^2}\big(\bar w_{j_0}^+\big)^{\beta_{j_0}-l_0^2}\cdots\end{aligned}$$

$$
\begin{aligned}
&\cdot \left(\bar{w}^+_{i_{p-2}}\right)^{l^1_{p-2}} \left(\bar{w}^+_{j_{p-2}}\right)^{\beta_{i_{p-2}}-l^1_{p-2}} \left(\bar{w}^+_{i_{p-2}}\right)^{l^2_{p-2}} \left(\bar{w}^+_{j_{p-2}}\right)^{\beta_{j_{p-2}}-l^2_{p-2}} \\
&= P_{0,k\alpha\beta}\left(y^+\right) e^{i\langle k,x^+\rangle} \left(\Pi_{i\notin\mathcal{N}} \left(w^+_i\right)^{\alpha_i} \left(\bar{w}^+_i\right)^{\beta_i}\right) \left(w^+_{i_0}\right)^{k^1_0+k^2_0} \left(w^+_{j_0}\right)^{\alpha_{i_0}+\alpha_{j_0}-k^1_0-k^2_0} \cdots \\
&\quad \cdot \left(w^+_{i_{p-2}}\right)^{k^1_{p-2}+k^2_{p-2}} \left(w^+_{j_{p-2}}\right)^{\alpha_{i_{p-2}}+\alpha_{j_{p-2}}-k^1_{p-2}-k^2_{p-2}} \\
&\quad \cdot \left(\bar{w}^+_{i_0}\right)^{l^1_0+l^2_0} \left(\bar{w}^+_{j_0}\right)^{\beta_{i_0}+\beta_{j_0}-l^1_0-l^2_0} \cdots \left(\bar{w}^+_{i_{p-2}}\right)^{l^1_{p-2}+l^2_{p-2}} \left(\bar{w}^+_{j_{p-2}}\right)^{\beta_{i_{p-2}}+\beta_{j_{p-2}}-l^1_{p-2}-l^2_{p-2}},
\end{aligned}
$$

where k, α, β satisfy (A.3) and

$$0 \leqslant k^1_t \leqslant \alpha_{i_t}, \qquad 0 \leqslant k^2_t \leqslant \alpha_{j_t}, \qquad 0 \leqslant l^1_t \leqslant \beta_{i_t}, \qquad 0 \leqslant l^2_t \leqslant \beta_{j_t}, \quad t \in \mathcal{T}.$$

Then from (A.3), one gets

$$
\begin{aligned}
&k_1 n_1 + k_2 n_2 + \sum_{i\notin\mathcal{N}} i(\beta_i - \alpha_i) \\
&\quad + \sum_{t=0}^{p-2} \left[i_t\left(l^1_t + l^2_t - k^1_t - k^2_t\right) + j_t\left(\beta_{i_t} + \beta_{j_t} - \alpha_{i_t} - \alpha_{j_t} - \left(l^1_t + l^2_t - k^1_t - k^2_t\right)\right)\right] \\
&= k_1 n_1 + k_2 n_2 + \sum_{i\notin\mathcal{N}} i(\beta_i - \alpha_i) + \sum_{t=0}^{p-2} \left[i_t(-\alpha_{i_t} + \beta_{i_t}) + j_t(-\alpha_{j_t} + \beta_{j_t})\right] \\
&\quad + \sum_{t=0}^{p-2} \left[-i_t\left(\beta_{i_t} - \alpha_{i_t} + k^1_t + k^2_t - l^1_t - l^2_t\right) + j_t\left(\beta_{i_t} - \alpha_{i_t} + k^1_t + k^2_t - l^1_t - l^2_t\right)\right] \\
&= (n_1 - n_2) \sum_{t=0}^{p-2} (\alpha_{j_t} - \beta_{j_t})(p-t) + \sum_{t=0}^{p-2} \left(\beta_{i_t} - \alpha_{i_t} + k^1_t + k^2_t - l^1_t - l^2_t\right)(j_t - i_t) \\
&= (n_1 - n_2) \sum_{t=0}^{p-2} (\alpha_{j_t} - \beta_{j_t})(p-t) + \sum_{t=0}^{p-2} \left(\beta_{i_t} - \alpha_{i_t} + k^1_t + k^2_t - l^1_t - l^2_t\right)(n_2 - n_1)(p-t) \\
&= \sum_{t=0}^{p-2} (n_1 - n_2)(p-t)\left[\left(\alpha_{i_t} + \alpha_{j_t} - k^1_t - k^2_t\right) - \left(\beta_{i_t} + \beta_{j_t} - l^1_t - l^2_t\right)\right]. \qquad \square
\end{aligned}
$$

From the generalized compact form of P, we can prove that the coefficient of $w_n \bar{w}_{-n}$ is zero unless $n = 0$ (see Section 3.1 for details).

Proof.

Case 1. $-j_t \neq j_{t'}$, for any $t, t' \in \mathcal{T}$.

Subcase 1. $n \notin \{\pm j_0, \ldots, \pm j_{p-2}\}$. It is easy.

Subcase 2. $n \in \{j_0, \ldots, j_{p-2}\}$. From $n_1 + n_2 \neq 0$, one gets $-2j_t \neq (n_1 - n_2)(p-t)$. The conclusion is easy.

Subcase 3. $n \in \{-j_0, \ldots, -j_{p-2}\}$. This is similar as Subcase 2.

Case 2. For some $t, t' \in \mathcal{T}$, we have $n = j_t \neq 0$, $-n = j_{t'}$. In this case, since $-j_t + j_{t'} = -2j_t \neq 0$, the conclusion is obvious.

Case 3. For some $t, t' \in \mathcal{T}$, we have $n = -j_t = j_{t'} \neq 0, -n = j_t$. This is similar as Case 2. □

A.2. Proof of Lemma 4.12

Proof. We will prove parts of the inequalities in Lemma 4.12. The unproved are similar as the following or obvious.

First, we prove (4.15). Write

$$M_1 = \big(\langle k, \omega\rangle + \Omega_n\big)I_2 + A_{i_t}, \quad t \in \mathcal{T},\ n \notin \mathcal{N}.$$

Obviously,

$$M_1 = P_t^T M_1' P_t = P_t^T \Big(\big(\langle k, \omega\rangle + \Omega_n\big)I_2 + \bar{A}_{i_t}\Big)P_t.$$

In the following we will prove

$$\big\|\big(M_1'\big)^{-1}\big\| \leqslant \frac{c\max\{|k|^{4p\tau+4}, 1\}}{\epsilon^{\frac{\beta_0}{2}}(||i_t| - |n|| + 1)}. \tag{A.4}$$

For our convenience, write $g^1 = \det(M_1')$. We will discuss in two cases.

Case 1. $\langle k, \lambda_0\rangle + n^2 + i_t^2 \neq 0$.

It is obvious that $k \neq 0$. Note the choose of K_0, one has

$$cK_0 \leqslant \frac{1}{\epsilon^{6p}}. \tag{A.5}$$

Therefore,

$$\left|\frac{\langle k, \lambda_0\rangle + n^2 + i_t^2}{\epsilon^{6p}} + k_1 f_1 + k_2 f_2 + 2f_3\right| \geqslant \frac{c}{\epsilon^{6p}},$$
$$\left|\frac{\langle k, \lambda_0\rangle + n^2 + i_t^2}{\epsilon^{6p}} + k_1 f_1 + k_2 f_2 + 2f_3 + (p - t_2)A\right| \geqslant \frac{c}{\epsilon^{6p}}.$$

It follows

$$\big\|\big(M_1'\big)^{-1}\big\| \leqslant c\epsilon^{6p}.$$

Note $|k_1 n_1 + k_2 n_2| = |n + i_t|$ or $|k_1 n_1 + k_2 n_2| = |n + j_t + (n_1 - n_2)(p - t)|$, we have

$$|n| \leqslant c|k|. \tag{A.6}$$

Therefore

$$\frac{c|k|^{4p\tau+2}}{||i_t|-|n||+1} \geqslant c|k|^{4p\tau+1}.$$

Thus, it is easy to get (A.4).

Case 2. $\langle k, \lambda_0\rangle + n^2 + i_t^2 = 0$.

Note we have thrown all the parameters in $\mathcal{R}^0_{20,2}$, this means

$$\left|\frac{1}{g^1}\right| \leqslant \frac{\max\{1, |k|^{4p\tau}\}}{\epsilon^{\frac{\beta_0}{2}}}.$$

From

$$\left(M_1'\right)^{-1} = \frac{1}{g^1}\begin{pmatrix} k_1 f_1 + k_2 f_2 + 2f_3 + (p-t)A & -a_t \\ -a_t & k_1 f_1 + k_2 f_2 + 2f_3 \end{pmatrix}$$

and (A.6), it follows

$$\left\|\left(M_1'\right)^{-1}\right\| \leqslant \frac{c\max\{|k|^{4p\tau+3}, 1\}}{\epsilon^{\frac{\beta_0}{2}}} \leqslant \frac{c\max\{|k|^{4p\tau+4}, 1\}}{\epsilon^{\frac{\beta_0}{2}}(||i_t|-|n||+1)}.$$

Combined with above two cases, the conclusion is clear.

In the following we will prove (4.16). Write

$$M_2 = I_2 \otimes \left(\langle k, \omega\rangle I_2 + A_{i_{t_2}}\right) + A_{i_{t_1}} \otimes I_2.$$

Note

$$M_2' = I_2 \otimes \left(\langle k, \omega\rangle I_2 + \bar{A}_{i_{t_2}}\right) + \bar{A}_{i_{t_1}} \otimes I_2$$

has the same eigenvalues as M_2 (see Lemma 5.3 in [19]), this means that there exists an orthogonal matrix P_{t_1,t_2} so that

$$P^T_{t_1,t_2} M_2' P_{t_1,t_2} = M_2.$$

Denote $g^2 = \det(M_2')$. (4.16) is clear from the equality

$$\left\|\left(M_2'\right)^{-1}\right\| \leqslant \frac{c\max\{1, |k|^{8p\tau+6}\}}{\epsilon^{\beta_0}}. \tag{A.7}$$

We will obtain (A.7) in the following two cases.

Case 1. $\langle k, \lambda_0\rangle + i_{t_1}^2 + i_{t_2}^2 \neq 0$.

As before, we only need discuss it when $cK_0 \leqslant \frac{1}{\epsilon^{6p}}$. It is easy to get

$$\left\|\left(M_2'\right)^{-1}\right\| \leqslant c\epsilon^{6p}.$$

Case 2. $\langle k, \lambda_0\rangle + i_{t_1}^2 + i_{t_2}^2 = 0$.

Note we have thrown out all the parameters in $\mathcal{R}^0_{20,3}$, it follows that

$$|g^2| \geqslant \frac{\epsilon^{\beta_0}}{\max\{1, |k|^{8p\tau}\}}. \tag{A.8}$$

Let $(M_2')^*$ denote the adjoint matrix of M_2'

$$(M_2')^* = \begin{pmatrix} m_{11} & m_{12} & m_{13} & m_{14} \\ m_{21} & m_{22} & m_{23} & m_{24} \\ m_{31} & m_{32} & m_{33} & m_{34} \\ m_{41} & m_{42} & m_{43} & m_{44} \end{pmatrix}.$$

Obviously, we have

$$|m_{ij}| \leqslant c|k|^6. \tag{A.9}$$

Therefore,

$$\left\| (M_2')^{-1} \right\| \leqslant \frac{c \max\{1, |k|^{8p\tau+6}\}}{\epsilon^{\beta_0}}. \qquad \square$$

References

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

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

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

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

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

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

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

[8] H.L. Eliasson, S.B. Kuksin, KAM for the nonlinear Schrödinger equation, Ann. of Math., in press.

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

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

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

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

[13] Z. Liang, J. You, Quasi-periodic solutions for 1D Schrödinger equations with higher order nonlinearity, SIAM J. Math. Anal. 36 (2005) 1965–1990.

[14] Z. Liang, J. You, Quasi-periodic solutions for 1D nonlinear wave equation with a general nonlinearity, http://www.ma.utexas.edu/mp_arc/c/05/05-75.pdf.

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

[16] J. Xu, J. You, Q. Qiu, A KAM theorem of degenerate infinite-dimensional Hamiltonian systems, I, II, Sci. China Ser. A 39 (4) (1996) 372–394.
[17] J. Xu, J. You, Q. Qiu, Invariant tori for nearly integrable Hamiltonian systems with degeneracy, Math. Z. 226 (1997) 375–387.
[18] J. Xu, J. You, Persistence of lower-dimensional tori under the first Melnikov's non-resonance condition, J. Math. Pures Appl. 80 (10) (2001) 1045–1067.
[19] J. You, Perturbations of lower-dimensional tori for Hamiltonian systems, J. Differential Equations 152 (1999) 1–29.
[20] X. Yuan, A KAM theorem with applications to partial differential equations of higher dimension, Comm. Math. Phys. 275 (2007) 97–137.