# Quasi-Periodic Solutions for 1D Schrödinger Equation with the Nonlinearity $|u|^{2p}u^*$

Zhenguo Liang

School of Mathematical Sciences, Fudan University, Shanghai 200433, China;

Key Lab of Mathematics for Nonlinear Science

(Fudan University), Ministry of Education, China.

**Abstract**

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

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

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.

## 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*

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

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

As usual, we study the equation (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,

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

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:

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

*The work was supported by the National Natural Science Foundation of China (10671035).

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

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

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 exists 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 $\Phi$ to each $\mathcal{T}_{\mathcal{I}}(I)$ in the family is an embedding of a rotational 2-torus for the nonlinear equation. In [14], 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 the equation (1.1) and (1.2).*

**Remark 1.2** *Generally, one can't prove that* $\Phi$ *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* $\Psi_1,\ \Psi_2$*. See Section 2 for details.*

There are some known woks about the equation (1.1). For $p=1$ under *Dirichlet boundary conditions*, see the well-known work of Kuksin and Pöschel [14]. For $p=2$ under *Dirichlet boundary conditions*, Liang and You(see [15]) also got the similar conclusions as [14]. But their method is hard to be generalized to $p\ge 3$. The reason will be given in the following. Before that, we will turn to some works about the Schrödinger equation under *periodic boundary conditions*. In [2, 4, 5], Bourgain obtained the existences of quasi–periodic solutions for the Schrödinger equation including 1D and nD($n\ge 2$). His method, called Craig-Wayne-Bourgain's scheme(see [7, 2, 3, 4, 5]) 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 equation 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. But their method is hard to deal with the Schrödinger equation. For the Schrödinger equation (1.3)+(1.4) when $p=1$, it was included in the work of Geng and You [11]. Combing with the methods of [15] and [12], Geng and Yi(see [13]) obtained the similar result for $p=2$. But their methods failed in $p\ge 3$. What is the problem?

Before we turn to the problem mentioned above and explain our method, we want to give a fast introduction in the recent development in KAM of higher dimension. In [12], Geng and You proved a KAM type of theorem which is applicable to certain Hamiltonian partial differential equations in higher space dimension including beam equations and Schrödinger equations with nonlocal nonlinearity. The important point in their proof is that they find the perturbation terms of the iterative Hamiltonian pertain some special form. Unfortunately, they expel the most interesting cases such as the higher dimensional Schrödinger equation with the general nonlinearities and higher dimensional wave equation. Very recently, there is new important development towards two problems. In [8, 9, 10], Eliasson and Kuksin give a KAM for higher dimensional Schrödinger equation with the general nonlinearities. The constructed quasi-periodic solutions have all Lyapounov exponents equal to zero. In [22], Yuan obtains a KAM theorem which can be applied to both the nonlinear wave equations and Schrödinger equations of higher dimension. The second Melnikov's conditions are totally eliminated in his method.

Now, we come back to the Schrödinger equation (1.1)+(1.2). We give a short discussion and the reason why the existent results only restricted in $p\leq 2$. In the end, we will give the sketch of our proof and point out the main difficulties in our proof. After the standard way as [15] and [13], we have the Hamiltonian

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

where

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

$$\begin{aligned}\widehat{G}&=c_p\sum_{\substack{i_1+\cdots+i_{p+1}=j_1+\cdots+j_{p+1}\\ \{i_1,\cdots,i_{p+1},j_1,\cdots,j_{p+1}\in\Delta_3\}}}q_{i_1}\cdots q_{i_{p+1}}\bar{q}_{j_1}\cdots\bar{q}_{j_{p+1}},\\ |K|&=\mathcal{O}(||q||_\rho^{4p+2}),\end{aligned}$$

where $n_1,\ n_2$ are the different tangent sites and the set

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

Please see section 2 for other relative notations. The existence of $\tilde{G}$ is the trouble-maker. For $p=1$, $\tilde{G}$ doesn't exist. For $p=2$, if choose $n_2-n_1\in 2\mathbb{N}-1$, the terms of $\tilde{G}$ also don't exist. This is the reason why Geng and Yi(see [13]) restrict the tangent site in the case $n_2-n_1\in 2\mathbb{N}-1$(also see Remark 2.1). For either $p=2$ and $n_2-n_1\in 2\mathbb{N}$ or $p\geq 3$, the terms of $\tilde{G}$ aren't empty.

This similar phenomenon, as the terms of $\tilde{G}$ don't vanish, exists very popularly. It definitely exists in 1D Schrödinger equation with the nonlinearity $|u|^{2p}u(p\geq 2)$ under *Dirichlet boundary conditions*. It is why it is difficult to generalize the conclusions of [15] 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 $u^{2\bar{r}+1}(\bar{r}\geq 3)$ under different boundary conditions. For example, it exists in 1D wave equation

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

under *Dirichlet boundary conditions*. If use the same notation as [16], when $\bar{r}=3$, we will find that the nonresonant term $z_{n_2}z_{n_1}^3z_i\bar{z}_{n_1}\bar{z}_{n_2}\bar{z}_j$ can't 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 [16] for details.

In the following, we give a sketch of our proof. Firstly, we give the concrete form of $\tilde{G}$(see (2.6), (2.9) and (2.10)). The proof is only restricted in two different tangent sites $n_1,\ n_2$. After the standard way of introducing the parameters $\xi_1,\ \xi_2$, we have the Hamiltonian

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

See (2.11) for details. In some sense, $\Upsilon_2+\Upsilon_3=h.o.t.$. We absorb the term of $\Upsilon_1$ into the term of $\sum_{n\neq n_1,n_2}\tilde{\Omega}_n(\xi)w_n\bar{w}_n$. And rewrite the two terms as $\langle G(x)w,\bar{w}\rangle$, where the infinite dimensional normal matrix

$$G(x)=\begin{pmatrix}\ddots &&&&&&&& \\ & \tilde{\Omega}_{i_0} &&&&&& \tilde{a}_0e^{-ip(x_1-x_2)} & \\ && \cdots &&&& \cdots && \\ &&& \tilde{\Omega}_{i_{p-2}} && \tilde{a}_{p-2}e^{-2i(x_1-x_2)} &&& \\ &&&& \ddots &&&& \\ &&& \tilde{a}_{p-2}e^{2i(x_1-x_2)} && \tilde{\Omega}_{j_{p-2}} &&& \\ && \cdots &&&& \cdots && \\ & \tilde{a}_0e^{ip(x_1-x_2)} &&&&&& \tilde{\Omega}_{j_0} & \\ &&&&&&&& \ddots\end{pmatrix},$$

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). In the following, 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

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

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

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

For $\Omega_j$, $j\neq n_1,\ n_2$, $A$ and $a_t$, see (2.23), (2.14) and (2.17). 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 common properties as ([12]) and ([13]) don't pertain after the nonlinear symplectic transformation $\Psi_1$(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. Another main difficulty is the measure estimate in the first step, while measure estimates of the remaining steps are standard as [15] and [16]. The difficulty lie in $\lambda_{1,t}$ and $\lambda_{2,t}$, $t\in\mathcal{T}$. It is hard to get the inequality such as

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

where $f=\langle k,\omega\rangle+\tilde{\Omega}_n-\tilde{\Omega}_m$. Our method is technical. See Lemma 4.12 for details.

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 the Appendix, we explain what is the compact form and generalized compact form. Some important lemmata are proved there.

# 2 Normal Form

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

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

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 $u(x,t)=\sum\limits_{n\in\mathbb{Z}}q_n(t)\phi_n(x)$. The coordinates are taken from the Hilbert spaces $l^\rho$ of all complex-valued sequences $q=(q_i)_{i\in\mathbb{Z}}$ with

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

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

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

where

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

with

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

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

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

The proof is similar as Lemma 3 in [14].

Note that

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

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 $\Delta_*$, $*=0,1,2,3$, as follows. For each $*=0,1,2$, $\Delta_*$ is the set of indices

$(i_1,\cdots,i_{p+1},j_1,\cdots,j_{p+1})$ which have exactly "$*$" components not in $\{n_1,n_2\}$. $\Delta_3$ is the set of the indices $(i_1,\cdots,i_{p+1},j_1,\cdots,j_{p+1})$ which have at least three components not in $\{n_1,n_2\}$. We also consider the resonance sets $\mathcal{N}=\{i_1,\cdots,i_{p+1},i_1,\cdots,i_{p+1}\}\bigcap\Delta_0$, $\mathcal{M}=\{i_1,\cdots,i_{p+1},i_1,\cdots,i_{p+1}\}\bigcap\Delta_2$. For our convenience, denote the sets $\mathcal{T}_1$, $\mathcal{T}_2$,

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

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

**Lemma 2.2** *Let $(i_1,\cdots,i_{p+1},j_1,\cdots,j_{p+1})\in(\Delta_0\setminus\mathcal{N})\bigcup\Delta_1\bigcup\mathcal{T}_2$. 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 $(i_1,\cdots,i_{p+1},j_1,\cdots,j_{p+1})\in(\Delta_0\setminus\mathcal{N})$, without losing generality, suppose there are exactly $x$'s $n_1$ in $\{i_1,\cdots,i_{p+1}\}$ and $y$'s $n_1$ in $\{j_1,\cdots,j_{p+1}\}$. 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 $(i_1,\cdots,i_{p+1},j_1,\cdots,j_{p+1})\in\Delta_1$, without losing generality, suppose $x_1$'s $n_1$ in $\{i_1,\cdots,i_{p+1}\}$ and $y_1$'s $n_1$ in $\{j_1,\cdots,j_{p+1}\}$. And the unique index in $\{j_1,\cdots,j_{p+1}\}$ 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

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

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, $a_1(1-a_1)(n_1-n_2)^2\neq 0$. ∎

**Lemma 2.3** *Given $n_1<n_2$, $n_1,n_2\in\mathbb{Z}$, there exists a real analytic, symplectic change of coordinates $\Gamma$ in a neighborhood of the origin of $l^\rho$ which transforms the Hamiltonian (2.24) into a partial Birkhoff normal form*

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

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

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

*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

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

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 [13].

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 $\{i_1,\cdots,i_{p+1},j_1,\cdots,j_{p+1}\}\in\mathcal{T}_1$. It is obvious that $a\neq b$. Otherwise, we have $n_1=n_2$. For

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

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

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

$$\begin{cases} k_1+l_1=p-1\\ k_2+l_2=p+1\\ k_1n_1+l_1n_2+a+b=k_2n_1+l_2n_2.\end{cases}$$

We will prove that

$$a^2+b^2+k_1n_1^2+l_1n_2^2\neq k_2n_1^2+l_2n_2^2.$$

If this isn't 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,

$$\begin{cases} a+b+s_1n_1+(-2-s_1)n_2=0\\ a^2+b^2+s_1n_1^2+(-2-s_1)n_2^2=0.\end{cases}$$

Thus, it follows

$$\begin{aligned} &2a^2+2(s_1n_1-(2+s_1)n_2)a\\ &\quad+s_1(s_1+1)n_1^2+(2+s_1)(1+s_1)n_2^2-2s_1(s_1+2)n_1n_2=0. \end{aligned}\tag{2.5}$$

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) can't 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

$$\begin{cases} k_1+l_1=k_2+l_2=p\\ a+k_1n_1+l_1n_2=b+k_2n_1+l_2n_2\\ a^2+k_1n_1^2+l_1n_2^2=b^2+k_2n_1^2+l_2n_2^2,\end{cases}$$

where $k_1,k_2=0,1,\cdots,p$, $l_1,l_2=0,1,\cdots,p$. 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

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

It is clear that $s\neq 0,\pm 1$, $s=k_1-k_2=l_2-l_1$, $s\in\{-p,\cdots,-1,0,1,\cdots,p\}$ 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 $s\in\{-p,\cdots,-2,2,\cdots,p\}$ satisfying

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

Denote this set of $s$ by $\mathcal{R}_1$. Corresponding to every $s\in\mathcal{R}_1$ mentioned above, we have many integer pair $(k_1,k_2)$ satisfying $k_1-k_2=s$, $k_1,k_2\in\{0,1,\cdots,p\}$. 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 $c_pq_aq_{n_1}^{k_1}q_{n_2}^{l_1}\bar{q}_b\bar{q}_{n_1}^{k_2}\bar{q}_{n_2}^{l_2}$, 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 the all terms in $\tilde{G}$.

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

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

where

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

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

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 [13]. When $n_2-n_1\in 2\mathbb{N}$, we have*

$$\tilde{G}=c_2q_aq_{n_1}^2\bar{q}_b\bar{q}_{n_2}^2+c_2q_bq_{n_2}^2\bar{q}_a\bar{q}_{n_1}^2,$$

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

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 $\xi_{n_j},\ j=1,2$ by $\xi_j,\ j=1,2$. 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

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

where

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

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

$$P_{k\alpha\beta}(y)\neq 0,\ \text{implies}\ k_1n_1+k_2n_2+\sum_{n\in\mathbb{Z}\setminus\{n_1,n_2\}}(-\alpha_n+\beta_n)n=0.$$

In order to cut our expression, write $\mathcal{N}=\{i_0,\cdots,i_{p-2},j_0,\cdots,j_{p-2}\}$ and $\mathcal{J}=\{j_0,\cdots,j_{p-2}\}$. 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 the infinite KAM theorem. Our object is to transform $\Upsilon_1$ to the terms which don't 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:*

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

*is symplectic, where*

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

**Remark 2.2** *The similar sympletic transformation as $\Psi_1$ was used in [20].*

Under the above symplectic coordinates transformation $\Psi_1$, 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

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

and $\omega$, $\Omega$ 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 the Appendix 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

$$\tilde{H}=\epsilon^{6p+8}H_+(x,\epsilon^8y,\epsilon^4w,\epsilon^4\bar{w},\epsilon^6\xi,\epsilon) \tag{2.16}$$

$$=\langle\tilde{\omega},y\rangle+\sum_{n\notin\mathcal{N}}\langle\tilde{\Omega}_nz_n,\bar{z}_n\rangle+\sum_{t=0}^{p-2}\langle\tilde{\bar{A}}_{i_t}z_{i_t},\bar{z}_{i_t}\rangle+\epsilon\tilde{P}_0,$$

where

$$\tilde{\omega}_1(\xi)=\frac{n_1^2}{\epsilon^{6p}}+c_p\sum_{k=0}^p(C_{p+1}^k)^2C_{p+1-k}^1\xi_1^{2p-2k}\xi_2^{2k},$$

$$\tilde{\omega}_2(\xi)=\frac{n_2^2}{\epsilon^{6p}}+c_p\sum_{k=0}^p(C_{p+1}^{k+1})^2C_{k+1}^1\xi_1^{2p-2k}\xi_2^{2k},$$

$$\tilde{\Omega}_n(\xi)=\frac{n^2}{\epsilon^{6p}}+c_p(C_{p+1}^1)^2\sum_{k=0}^p(C_p^k)^2\xi_1^{2p-2k}\xi_2^{2k},\ n\neq n_1,\ n_2,$$

$$\tilde{\bar{A}}_{i_t}=\begin{pmatrix} \tilde{\Omega}_{i_t} & a_t \\ a_t & \tilde{\Omega}_{i_t}+(p-t)A \end{pmatrix},$$

$$a_t=c_p\sum_0^t\xi_1^{p+t-2j}\xi_2^{p-t+2j}, \tag{2.17}$$

$$A=\sum_0^p c_p[(C_{p+1}^k)^2C_{p+1-k}^1-(C_{p+1}^{k+1})^2C_{k+1}^1]\xi_1^{2(p-k)}\xi_2^{2k}, \tag{2.18}$$

$\xi\in\mathcal{O}=[1,2]^2$. 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 $\omega$, $\tilde{\Omega}$ by $\Omega$, $\tilde{\bar{A}}$ by $\bar{A}$, $\tilde{\bar{B}}$ by $\bar{B}$ and $\tilde{P}_0$ by $P_0$.

Now the new Hamiltonian is

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

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

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

where

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

and

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

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

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

*is symplectic.*

*Proof:* It is easy to check that

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

Under the above symplectic coordinates transformation $\Psi_2$, 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

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

$$A_{i_t}=\begin{pmatrix}\lambda_{1,t}&0\\0&\lambda_{2,t}\end{pmatrix},$$

$\xi\in\mathcal{O}$. For $\lambda_{1,t}$, $\lambda_{2,t}$, see (2.21) and (2.22). From Lemma 5.3, 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 $\epsilon P_0^+$ by $P$. Therefore, the Hamiltonian is

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

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 don't 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:

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

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

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

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

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

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

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

and

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

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

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

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

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

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

In the whole of this paper, by $c$ a universal constant, whose size may be different in different place. If $f\le cg$, we write this inequality as $f\le\cdot g$ when we don't care the size of the constant $c$. Similarly, if $f\ge cg$, we write $f\ge\cdot g$.

# 3 KAM Step

Theorem 1 will be proved by a KAM iteration which involves an infinite sequence of change of variables. Each step of KAM iteration makes the perturbation smaller than 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.

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

$$\mathcal{O}_0\subset\left\{\xi:\begin{array}{l}|\langle k,\omega\rangle^{-1}|\le\frac{c|k|^{8p\tau+6}}{\epsilon^{\beta_0}},\quad k\neq0,\\ |(\langle k,\omega\rangle+\tilde\Omega_n)^{-1}|\le\frac{c\max\{|k|^{8p\tau+6},1\}}{\epsilon^{\beta_0}},\\ |(\langle k,\omega\rangle+\tilde\Omega_n+\tilde\Omega_m)^{-1}|\le\frac{c\max\{|k|^{8p\tau+6},1\}}{\epsilon^{\beta_0}(||n|-|m||+1)},\\ \text{where } n,\ m\notin\mathcal{N}\text{ or } n,\ m\in\mathcal{N},\\ |(\langle k,\omega\rangle+\tilde\Omega_n+\tilde\Omega_{i_t})^{-1}|\le\frac{c\max\{|k|^{8p\tau+6},1\}}{\epsilon^{\beta_0}(||i_t|-|n||+1)},\\ \text{where } n\notin\mathcal{N},\ t\in\mathcal{T},\ |k_1n_1+k_2n_2|=|n+i_t|,\\ |(\langle k,\omega\rangle+\tilde\Omega_n+\tilde\Omega_{j_t})^{-1}|\le\frac{c\max\{|k|^{8p\tau+6},1\}}{\epsilon^{\beta_0}(||j_t|-|n||+1)},\\ \text{where } n\notin\mathcal{N},\ t\in\mathcal{T},\ |k_1n_1+k_2n_2|=|n+j_t+(n_1-n_2)(p-t)|,\\ |(\langle k,\omega\rangle+\tilde\Omega_n-\tilde\Omega_m)^{-1}|\le\frac{c\max\{|k|^{8p\tau+6},1\}}{\epsilon^{\beta_0}(||n|-|m||+1)},\\ \text{where } n,\ m\notin\mathcal{N},\ |k|+||n|-|m||\neq0,\ |k_1n_1+k_2n_2|=|n-m|,\\ |(\langle k,\omega\rangle+\tilde\Omega_n-\tilde\Omega_m)^{-1}|\le\frac{c\max\{|k|^{8p\tau+6},1\}}{\epsilon^{\beta_0}(||n|-|m||+1)},\\ \text{where } n,\ m\in\mathcal{N},\ |k|+|n-m|\neq0,\\ |(\langle k,\omega\rangle+\tilde\Omega_n-\tilde\Omega_{i_t})^{-1}|\le\frac{c\max\{|k|^{8p\tau+6},1\}}{\epsilon^{\beta_0}(||i_t|-|n||+1)},\\ \text{where } n\notin\mathcal{N},\ t\in\mathcal{T},\ |k_1n_1+k_2n_2|=|n-i_t|,\\ |(\langle k,\omega\rangle+\tilde\Omega_n-\tilde\Omega_{j_t})^{-1}|\le\frac{c\max\{|k|^{8p\tau+6},1\}}{\epsilon^{\beta_0}(||j_t|-|n||+1)},\\ \text{where } n\notin\mathcal{N},\ t\in\mathcal{T},\ |k_1n_1+k_2n_2|=|n-j_t-(n_1-n_2)(p-t)|\end{array}\right\},\tag{3.1}$$

and $0\le|k|\le K_0$ and

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

For more concretely, please refer to section 4 and Lemma 4.12. $\beta_0$ is a constant and will be chosen later.

Suppose $\|\omega\|_*\le M_1,\ \max_{j\in\mathbb{Z}}|\tilde\Omega_j|_*\le M_2,\ M_1+M_2\ge1$. 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

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

$P_0$ satisfies a general compact form (2.15).

Suppose that after a $\nu$th 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

$$\mathcal{O}_\nu\subset\left\{\xi:\begin{array}{l}|\langle k,\omega_\nu\rangle^{-1}|\le\frac{c|k|^{8p\tau+6}}{\epsilon_\nu^{\beta_0}},\quad k\neq 0,\\ |(\langle k,\omega_\nu\rangle+\tilde\Omega_{\nu,n})^{-1}|\le\frac{c\max\{|k|^{8p\tau+6},1\}}{\epsilon_\nu^{\beta_0}},\\ |(\langle k,\omega_\nu\rangle+\tilde\Omega_{\nu,n}+\tilde\Omega_{\nu,m})^{-1}|\le\frac{c\max\{|k|^{8p\tau+6},1\}}{\epsilon_\nu^{\beta_0}(||n|-|m||+1)},\\ \text{where } n,\ m\notin\mathcal{N} \text{ or } n,\ m\in\mathcal{N},\\ |(\langle k,\omega_\nu\rangle+\tilde\Omega_{\nu,n}+\tilde\Omega_{\nu,i_t})^{-1}|\le\frac{c\max\{|k|^{8p\tau+6},1\}}{\epsilon_\nu^{\beta_0}(||i_t|-|n||+1)},\\ \text{where } n\notin\mathcal{N},\ t\in\mathcal{T},\ |k_1n_1+k_2n_2|=|n+i_t|,\\ |(\langle k,\omega_\nu\rangle+\tilde\Omega_{\nu,n}+\tilde\Omega_{\nu,j_t})^{-1}|\le\frac{c\max\{|k|^{8p\tau+6},1\}}{\epsilon_\nu^{\beta_0}(||j_t|-|n||+1)},\\ \text{where } n\notin\mathcal{N},\ t\in\mathcal{T},\ |k_1n_1+k_2n_2|=|n+j_t+(n_1-n_2)(p-t)|,\\ |(\langle k,\omega_\nu\rangle+\tilde\Omega_{\nu,n}-\tilde\Omega_{\nu,m})^{-1}|\le\frac{c\max\{|k|^{8p\tau+6},1\}}{\epsilon_\nu^{\beta_0}(||n|-|m||+1)},\\ \text{where } n,\ m\notin\mathcal{N},\ |k|+||n|-|m||\neq 0,\ |k_1n_1+k_2n_2|=|n-m|,\\ |(\langle k,\omega_\nu\rangle+\tilde\Omega_{\nu,n}-\tilde\Omega_{\nu,m})^{-1}|\le\frac{c\max\{|k|^{8p\tau+6},1\}}{\epsilon_\nu^{\beta_0}(||n|-|m||+1)},\\ \text{where } n,\ m\in\mathcal{N},\ |k|+|n-m|\neq 0,\\ |(\langle k,\omega_\nu\rangle+\tilde\Omega_{\nu,n}-\tilde\Omega_{\nu,i_t})^{-1}|\le\frac{c\max\{|k|^{8p\tau+6},1\}}{\epsilon_\nu^{\beta_0}(||i_t|-|n||+1)},\\ \text{where } n\notin\mathcal{N},\ t\in\mathcal{T},\ |k_1n_1+k_2n_2|=|n-i_t|,\\ |(\langle k,\omega_\nu\rangle+\tilde\Omega_{\nu,n}-\tilde\Omega_{\nu,j_t})^{-1}|\le\frac{c\max\{|k|^{8p\tau+6},1\}}{\epsilon_\nu^{\beta_0}(||j_t|-|n||+1)},\\ \text{where } n\notin\mathcal{N},\ t\in\mathcal{T},\ |k_1n_1+k_2n_2|=|n-j_t-(n_1-n_2)(p-t)|\end{array}\right\},\tag{3.2}$$

$0\le|k|\le K'_\nu$ [1] for some $r_\nu\le r_0,\ s_\nu\le s_0$ and

$$K'_\nu=\begin{cases}K_0, & \nu=0,\\ \infty, & \nu\ge 1.\end{cases}$$

We also assume that

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

and $P_\nu=\sum\limits_{k,\alpha,\beta}P^\nu_{k\alpha\beta}(y)e^{i\langle k,x\rangle}w^\alpha\bar w^\beta$ 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 $\nu$th 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

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

where [2] $k\in\mathbb{Z}^2,l\in\mathbb{N}_0^2$ and the multi–index $\alpha$, $\beta$ run over the set $\alpha\equiv(\cdots,\alpha_n,\cdots)$, $\beta\equiv(\cdots,\beta_n,\cdots)$, $\alpha_n,\beta_n\in\mathbb{N}_0$ 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 $n$th components is 1 and other components are all zeros.

Let $R$ be the truncation of $P$ given by

$$\begin{aligned}R(x,y,w,\bar w)=&\sum_{|k|\le K,|l|\le 1}P_{kl00}e^{\mathrm{i}\langle k,x\rangle}\,y^l+\sum_{|k|\le K,n}(P^{k10}_nw_n+P^{k01}_n\bar w_n)e^{\mathrm{i}\langle k,x\rangle}\\&+\sum_{|k|\le K,n,m}(P^{k20}_{nm}w_nw_m+P^{k02}_{nm}\bar w_n\bar w_m+P^{k11}_{nm}w_n\bar w_m)e^{\mathrm{i}\langle k,x\rangle}\end{aligned}$$

[1] Where "$|k|\le\infty$" means "$|k|<\infty$". We confuse the notation for simplicity.
[2] $\mathbb{N}_0$ means $\mathbb{N}\cup\{0\}$.

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

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

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

$$\begin{aligned}
H\circ\Phi_F^1&=(N+R)\circ\Phi_F^1+(P-R)\circ\Phi_F^1\\
&=N+\{N,F\}+R\\
&\quad+\int_0^1(1-t)\{\{N,F\},F\}\circ\Phi_F^t dt+\int_0^1\{R,F\}\circ\Phi_F^t dt+(P-R)\circ\Phi_F^1 \qquad (3.3)\\
&=N_++P_++\{N,F\}+R-P_{0000}-\langle\omega',y\rangle-\sum_n R_{nn}^{011}w_n\bar{w}_n,
\end{aligned}$$

where

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

$$N_+=N+\widehat{N}=N+P_{0000}+\langle\omega',y\rangle+\sum_n R_{nn}^{011}w_n\bar{w}_n, \qquad (3.4)$$

$$P_+=\int_0^1(1-t)\{\{N,F\},F\}\circ\Phi_F^t dt+\int_0^1\{R,F\}\circ\Phi_F^t dt+(P-R)\circ\Phi_F^1. \qquad (3.5)$$

satisfying the homological equation

$$\{N,F\}+R-P_{0000}-\langle\omega',y\rangle-\sum_n R_{nn}^{011}w_n\bar{w}_n=0. \qquad (3.6)$$

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 (3.6), let $F$ has the form

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

By comparing the coefficients, it is easy to see that the homological equation (3.6) is equivalent to

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

where $0\le |k|\le K'$. Hence the homological equation (3.6) 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 [17] and [18] for details.

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

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

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

**Lemma 3.2** *If $|X_F|^*_{r,D(s-\sigma,r)}\le\sigma$, then for any $\xi\in\mathcal{O}$, the flow $X_F^t(\cdot,\ \xi)$ exists on $D(s-2\sigma,\ \frac{r}{2})$ for $|t|\le 1$ and maps $D(s-2\sigma,\ \frac{r}{2})$ into $D(s-\sigma,r)$. Moreover, for $|t|\le1$,*

$$|X_F^t-id|^*_{r,D(s-2\sigma,\frac{r}{2})},\sigma||DX_F^t-Id||^*_{r,r,D(s-3\sigma,\frac{r}{4})}\le c|X_F|^*_{r,D(s-\sigma,r)},$$

*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*

$$||A(\xi,\eta)||_{\bar r,r,D(s,r)}=\sup_{\eta\in D(s,r)}\sup_{w\neq0}\frac{||A(\xi,\eta)w||_{\rho,\bar r}}{||w||_{\rho,r}},$$

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

*For the proof refer to [18].*

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

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

Thus $|X_R|^*_{r,D(s,r)}\le\cdot|X_P|^*_{r,D(s,r)}\le\cdot\epsilon$, and for $\eta\le\frac18$,

$$|X_{P-R}|^*_{\eta r,D(s-\sigma,2\eta r)}\le\cdot\eta\epsilon+e^{-K'\sigma}\epsilon. \tag{3.7}$$

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 $n\neq0$(we check this in the appendix). This leads to the following simple new normal form

$$\begin{aligned} N_+ &= N+\langle\omega',y\rangle+\sum_n P_{nn}^{011}w_n\bar w_n\\ &=\langle\omega_+,y\rangle+\sum_n\tilde\Omega_{+,n}w_n\bar w_n, \end{aligned}$$

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

$$\|\omega_+ - \omega\|^*, \|\tilde{\Omega}_+ - \tilde{\Omega}\|^* \leq \cdot\epsilon, \tag{3.8}$$

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

$$\frac{1}{\sigma}|X_F^t - id|^*_{r,D(s-2\sigma,\frac{r}{2})}, \|DX_F^t - Id\|^*_{r,r,D(s-3\sigma,\frac{r}{4})} \leq \frac{cM\epsilon^{1-(8p+1)\beta_0}}{\sigma^{\mu+1}}. \tag{3.9}$$

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 $\bar{R}(t) = (1-t)\widehat{N} + tR$. 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 $\frac{cM\epsilon^{1-(8p+1)\beta_0}}{\eta^2\sigma^{\mu+1}} \leq 1$, we have $\|\Phi^*W\|^*_{\eta r,\ U} \leq c\|W\|^*_{\eta r,\ V}$.*

For the proof refer to [17].

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

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

By Cauchy's inequality and Lemma 3.2, one obtains

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

where one chooses $\eta^3 = \frac{\epsilon^{1-(8p+1)\beta_0}}{\sigma^{\mu+1}}$. Combining (3.7) we have

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

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

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

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

*Proof.* Note that

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

Since $P$ has a generalized compact form with respect to $n_1$, $n_2$ and $\mathcal{J}$, so do $P-R$ and $\{N,F\} = P_{0000} + \langle\omega', y\rangle + \sum_n P^{011}_{nn} w_n\bar{w}_n - R$. The lemma then follows from Lemma 5.2.

### 3.2 Iteration Lemma

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

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

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

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

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

$$\epsilon_0\le\gamma_0\sigma_0^{6(\mu+1)},\ \gamma_0\le\min\{\frac{1}{c^6 2^{13(1+\mu)}M^{42}},\ (\frac{c_0}{8c})^{8p\tau+7}\}, \tag{3.10}$$

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 $K_\nu$ and $K'_\nu$.

**Lemma 3.5** *Suppose $H_\nu=N_\nu+P_\nu(\nu\ge 0)$, 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*

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

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

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

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

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

where $n\notin\mathcal{N},\ t\in\mathcal{T},\ |k_1n_1+k_2n_2|=|n+i_t|$,

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

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

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

where $n,\ m\notin\mathcal{N},\ |k|+||n|-|m||\ne 0,\ |k_1n_1+k_2n_2|=|n-m|$,

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

where $n,\ m\in\mathcal{N},\ |k|+|n-m|\ne 0$,

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

where $n\notin\mathcal{N},\ t\in\mathcal{T},\ |k_1n_1+k_2n_2|=|n-i_t|$,

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

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

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

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

*Then there exists 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*

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

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

$$\begin{aligned}\mathcal{R}^{\nu+1}(\epsilon_{\nu+1})&=\mathcal{R}^{\nu+1}_{00}\bigcup\mathcal{R}^{\nu+1}_{10}\bigcup\mathcal{R}^{\nu+1}_{20}\bigcup\mathcal{R}^{\nu+1}_{11},\\ \mathcal{R}^{\nu+1}_{20}&=\mathcal{R}^{\nu+1}_{20,1}\bigcup\mathcal{R}^{\nu+1}_{20,2}\bigcup\mathcal{R}^{\nu+1}_{20,3},\\ \mathcal{R}^{\nu+1}_{11}&=\mathcal{R}^{\nu+1}_{11,1}\bigcup\mathcal{R}^{\nu+1}_{11,2}\bigcup\mathcal{R}^{\nu+1}_{11,3}\bigcup\mathcal{R}^{\nu+1}_{11,4},\end{aligned}$$

*and*

$$\mathcal{R}^{\nu+1}_{00}=\bigcup_{K'_{\nu+1}\ge|k|>K_\nu}\{\xi\in\mathcal{O}_\nu:\ |\langle k,\omega_{\nu+1}\rangle|<\frac{\epsilon_{\nu+1}^{\beta_0}}{c|k|^{8p\tau+6}},\quad k\ne0\},$$

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

$$\mathcal{R}^{\nu+1}_{20,1}=\bigcup_{K'_{\nu+1}\ge|k|>K_\nu,n,m}\{\xi\in\mathcal{O}_\nu:|\langle k,\omega_{\nu+1}\rangle+\tilde\Omega_{\nu+1,n}+\tilde\Omega_{\nu+1,m}|<\frac{\epsilon_{\nu+1}^{\beta_0}(||n|-|m||+1)}{c\max\{|k|^{8p\tau+6},1\}},\ \text{where } n,\ m\notin\mathcal{N}\text{ or } n,\ m\in\mathcal{N}\},$$

$$\mathcal{R}^{\nu+1}_{20,2}=\bigcup_{K'_{\nu+1}\ge|k|>K_\nu,n,t}\{\xi\in\mathcal{O}_\nu:|\langle k,\omega_{\nu+1}\rangle+\tilde\Omega_{\nu+1,n}+\tilde\Omega_{\nu+1,i_t}|<\frac{\epsilon_{\nu+1}^{\beta_0}(||i_t|-|n||+1)}{c\max\{|k|^{8p\tau+6},1\}},\ \text{where } n\notin\mathcal{N},\ t\in\mathcal{T},\ |k_1n_1+k_2n_2|=|n+i_t|\},$$

$$\mathcal{R}^{\nu+1}_{20,3}=\bigcup_{K'_{\nu+1}\ge|k|>K_\nu,n,t}\{\xi\in\mathcal{O}_\nu:|\langle k,\omega_{\nu+1}\rangle+\tilde\Omega_{\nu+1,n}+\tilde\Omega_{\nu+1,j_t}|<\frac{\epsilon_{\nu+1}^{\beta_0}(||j_t|-|n||+1)}{c\max\{|k|^{8p\tau+6},1\}},\ \text{where } n\notin\mathcal{N},\ t\in\mathcal{T},\ |k_1n_1+k_2n_2|=|n+j_t+(n_1-n_2)(p-t)|\},$$

$$\mathcal{R}^{\nu+1}_{11,1}=\bigcup_{K'_{\nu+1}\ge|k|>K_\nu,n,m}\{\xi\in\mathcal{O}_\nu:|\langle k,\omega_{\nu+1}\rangle+\tilde\Omega_{\nu+1,n}-\tilde\Omega_{\nu+1,m}|<\frac{\epsilon_{\nu+1}^{\beta_0}(||n|-|m||+1)}{c\max\{|k|^{8p\tau+6},1\}},\ \text{where } n,\ m\notin\mathcal{N},\ |k|+||n|-|m||\ne0,\ |k_1n_1+k_2n_2|=|n-m|\},$$

$$\mathcal{R}^{\nu+1}_{11,2}=\bigcup_{K'_{\nu+1}\ge|k|>K_\nu,n,m}\{\xi\in\mathcal{O}_\nu:|\langle k,\omega_{\nu+1}\rangle+\tilde\Omega_{\nu+1,n}-\tilde\Omega_{\nu+1,m}|<\frac{\epsilon_{\nu+1}^{\beta_0}(||n|-|m||+1)}{c\max\{|k|^{8p\tau+6},1\}},\ \text{where } n,\ m\in\mathcal{N},\ |k|+|n-m|\ne0\},$$

$$\mathcal{R}^{\nu+1}_{11,3}=\bigcup_{K'_{\nu+1}\ge|k|>K_\nu,n,t}\{\xi\in\mathcal{O}_\nu:|\langle k,\omega_{\nu+1}\rangle+\tilde\Omega_{\nu+1,n}-\tilde\Omega_{\nu+1,i_t}|<\frac{\epsilon_{\nu+1}^{\beta_0}(||i_t|-|n||+1)}{c\max\{|k|^{8p\tau+6},1\}},\ \text{where } n\notin\mathcal{N},\ t\in\mathcal{T},\ |k_1n_1+k_2n_2|=|n-i_t|\},$$

$$\mathcal{R}^{\nu+1}_{11,4}=\bigcup_{K'_{\nu+1}\ge|k|>K_\nu,n,t}\{\xi\in\mathcal{O}_\nu:|\langle k,\omega_{\nu+1}\rangle+\tilde\Omega_{\nu+1,n}-\tilde\Omega_{\nu+1,j_t}|<\frac{\epsilon_{\nu+1}^{\beta_0}(||j_t|-|n||+1)}{c\max\{|k|^{8p\tau+6},1\}},\ \text{where } n\notin\mathcal{N},\ t\in\mathcal{T},\ |k_1n_1+k_2n_2|=|n-j_t-(n_1-n_2)(p-t)|\},$$

*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* $\nu$.

*Proof:* Note (3.10), by induction one verifies that

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

It is easy to check that (5.5) 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 [17].

With (3.8) and (3.9), we also obtain the following estimate.

**Lemma 3.6**

$$\frac{1}{\sigma_\nu}|\Phi_{\nu+1} - id|^*_{r_\nu, D_{\nu+1}}, \|D\Phi_{\nu+1} - I\|^*_{r_\nu, r_\nu, D_{\nu+1}} \le \frac{cM(\nu)\epsilon_\nu^{1-(8p+1)\beta_0}}{\sigma_\nu^{\mu+1}} \tag{3.11}$$
$$\|\omega_{\nu+1} - \omega_\nu\|^*_{\mathcal{O}_\nu}, \|\tilde{\Omega}_{\nu+1} - \tilde{\Omega}_\nu\|^*_{\mathcal{O}_{\nu+1}} \le c\epsilon_\nu. \tag{3.12}$$

## 3.3 Convergence and Proof of the Existences of Tori

Let $\Phi^\nu = \Phi_1 \circ \Phi_2 \circ \cdots \circ \Phi_\nu$, $\nu = 1, 2, \cdots, \cdots$. 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 $\nu \ge 1$.

Let $\tilde{\mathcal{O}}_\epsilon = \bigcap_{\nu=0}^{\infty} \mathcal{O}_\nu$. We apply Lemma 3.5, Lemma 3.6 and standard arguments (see [17]) to conclude that $H_\nu$, $N_\nu$, $P_\nu$, $\Phi^\nu$, $D\Phi^\nu$, $\omega_\nu$, $\tilde{\Omega}_{\nu,n}$ converge uniformly on $D(\frac{1}{2}s_0, 0) \times \tilde{\mathcal{O}}_\epsilon$, say to, $H_\infty$, $N_\infty$, $P_\infty$, $\Phi^\infty$, $D\Phi^\infty$, $\omega_\infty$, $\tilde{\Omega}_{\infty,n}$ 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 $\Phi_H^t$ 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 $\Phi^\nu$, $D\Phi^\nu$, $X_{H_\nu}$ 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

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

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 section 3 and section 4 below, we have the following theorem.

**Theorem 2** *For the Hamiltonian (2.24)*

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

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

$$|X_P|^*_{r,D(s,r)} = \epsilon \leq \gamma s^{6(1+\mu)}, \tag{3.13}$$

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

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

*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 $\Phi$ 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 $\xi$.*

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

$$\|\Phi - \Phi_0\|^*_r \leq c\epsilon^{\frac{1}{3}},$$
$$||\omega_* - \omega||^* \leq c\epsilon,$$

*uniformly on that domain and $\tilde{\mathcal{O}}_\epsilon$, where $\Phi_0$ 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 [14].*

# 4 Measure estimates

## 4.1 Measure estimates in the first step

For simplicity, in this section we will denote

$$\lambda_0 = (n_1^2, n_2^2),$$
$$f_1 = \sum_{k=0}^{p} c_p (C^k_{p+1})^2 C^1_{p+1-k} \xi_1^{2p-2k} \xi_2^{2k},$$
$$f_2 = \sum_{k=0}^{p} c_p (C^{k+1}_{p+1})^2 C^1_{k+1} \xi_1^{2p-2k} \xi_2^{2k},$$
$$f_3 = \sum_{k=0}^{p} c_p (C^1_{p+1})^2 (C^k_p)^2 \xi_1^{2p-2k} \xi_2^{2k}.$$

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

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

where

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

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

$$\mathcal{R}^0_{1,2} = \bigcup_{t\in\mathcal{T}, |k|\leq K_0} \{\xi \in \mathcal{O} : |g_1| < \frac{\epsilon^{\frac{\beta_0}{2}}}{\max\{1, |k|^{4p\tau}\}}\}, \tag{4.3}$$

$$g_1 = detM_1',$$

and

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

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

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

where

$$g_2 = detM_2',$$

and

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

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

where

$$g_3 = detM_3', \ \Delta_4 = k_1 f_1 + k_2 f_2 + 2f_3,$$

and

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

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

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

where

$$g_4 = detM_4',$$

and

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

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

where

$$g_5 = detM_5',\ \Delta_5 = k_1 f_1 + k_2 f_2,$$

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 won't 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 $I_h = \{x || g(x)| < h\}, h > 0$. If for some constant $d > 0$, $|g^m(x)| \geq d$ for any $x \in I$, then $|I_h| \leq ch^{\frac{1}{m}}$, where $|I_h|$ denotes the Lebesgue measure of $I_h$ and $c = 2(2+3+\cdots+m+d^{-1})$.*

For the proof see [18]. The similar method can be found in [19].

**Lemma 4.2** *If $\tau > 2$,*

$$|\mathcal{R}_0^0| \leq \cdot \epsilon^{\frac{\beta_0}{8p}}.$$

**Lemma 4.3** *If $\tau > 3$,*

$$|\mathcal{R}_{1,1}^0| \leq \cdot \epsilon^{\frac{\beta_0}{8p}}.$$

**Lemma 4.4** *If $\tau > 2$,*

$$|\mathcal{R}_{1,2}^0| \leq \cdot \epsilon^{\frac{\beta_0}{8p}}.$$

*Proof.* If $k \neq 0$, it is easy to check that

$$\begin{aligned} g_1(\xi) &= c(k_1 + k_2(p+1) + (p+1))(k_1 + k_2(p+1) + (p+1) - p(p-t))\xi_1^{4p} \\ &\quad + c(k_1(p+1) + k_2 + (p+1))(k_1(p+1) + k_2 + (p+1) + p(p-t))\xi_2^{4p} + \mathcal{O}_1(\xi^{4p}), \end{aligned}$$

where $\mathcal{O}_1(\xi^{4p})$ mean other $4p$th order terms in $\xi$, $\xi_2$ which are different with $\xi_1^{4p}$ and $\xi_2^{4p}$(the similar notations will appear many times in following and be in the similar sense). It follows that

$$\begin{aligned} \frac{\partial^{4p} g_1}{\partial \xi_1^{4p}} &= c(k_1 + k_2(p+1) + (p+1))(k_1 + k_2(p+1) + (p+1) - p(p-t)) \\ \frac{\partial^{4p} g_1}{\partial \xi_2^{4p}} &= c(k_1(p+1) + k_2 + (p+1))(k_1(p+1) + k_2 + (p+1) + p(p-t)). \end{aligned}$$

If $\frac{\partial^{4p} g_1}{\partial \xi_1^{4p}} = 0$ and $\frac{\partial^{4p} g_1}{\partial \xi_2^{4p}} = 0$ hold for the same $k \neq 0$, then at least one of the following four cases is true.

Case 1.

$$\begin{cases} k_1 + k_2(p+1) + (p+1) = 0 \\ k_1(p+1) + k_2 + (p+1) = 0. \end{cases}$$

It is easy to get $k_1 \notin \mathbb{Z}$. It is impossible.

Case 2.

$$\begin{cases} k_1 + k_2(p+1) + (p+1) = p(p-t) \\ k_1(p+1) + k_2 + (p+1) = -p(p-t). \end{cases}$$

It is easy to get $k_1 + k_2 \notin \mathbb{Z}$. It is impossible.

Case 3.

$$\begin{cases} k_1 + k_2(p+1) + (p+1) = p(p-t) \\ k_1(p+1) + k_2 + (p+1) = 0. \end{cases}$$

It is easy to get either $k_1+k_2\notin\mathbb{Z}$ or $k_1\notin\mathbb{Z}$. It is impossible.

Case 4.

$$\begin{cases} k_1+k_2(p+1)+(p+1)=0\\ k_1(p+1)+k_2+(p+1)=-p(p-t).\end{cases}$$

It is easy to obtain $k_1+k_2=p-t+\frac{2t+2}{p+2}$. From $k_1+k_2\in\mathbb{Z}$, one has $t=\frac{p}{2}$. It follows $k_2=-\frac{3}{2}+\frac{2}{p}$. From $k_2\in\mathbb{Z}$, we have $p=4,\ k_1=4,\ k_2=-1, t=2$. For this special case, we compute the coefficients of $\xi_1^{4p-2}\xi_2^2$ of $g_1$ denoted by $g_1^{4p-2,2}$. In fact

$$\begin{aligned} g_1^{4p-2,2}=&c\{(p-t)(k_1(1-\frac{3}{2}p)+k_2(\frac{p}{2}+1-p^2)+(\frac{p}{2}+1-\frac{3}{2}p^2))\\ &+2(k_1+k_2(p+1)+(p+1))(k_1+k_2\frac{p}{2}+p)\}.\end{aligned}$$

It is easy to check that $g_1^{4p-2,2}\neq 0$.

In fact from the above discussions, we have proved that there exists $|\alpha_1|=4p$ so that $|\frac{\partial^{\alpha_1}g_1}{\partial\xi^{\alpha_1}}|\geq c>0$ for $k\neq 0$. For $k=0$, it is more easy to get the same conclusion. The following proof is standard. Please refer to [15] and [16] for details. ■

**Lemma 4.5** *If $\tau>5$,*

$$|\mathcal{R}^0_{20,1}|\leq\cdot\epsilon^{\frac{\beta_0}{8p}}.$$

For the proof, refer to Lemma 4.8.

**Lemma 4.6** *If $\tau>2$,*

$$|\mathcal{R}^0_{20,2}|\leq\cdot\epsilon^{\frac{\beta_0}{8p}}.$$

*Proof.* We only give a sketch. As Lemma 4.4, one can prove $g_2^{4p,0}$ or $g_2^{0,4p}$ is not equal to zero except one case which is $t=0,\ k_1=-1-p,\ k_2=-1$. For the remaining case, compute $g_2^{4p-2,2}=c(p-2)\neq 0$ when $p>2$. For $p=2$, we turn to compute $g_2^{4p-4,4}=g_2^{4,4}\neq 0$. Now the conclusion is clear.

**Lemma 4.7** *If $\tau>2$,*

$$|\mathcal{R}^0_{20,3}|\leq\cdot\epsilon^{\frac{\beta_0}{8p}}.$$

*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

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

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

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

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

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 the following 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 $\begin{cases} k_1 = -p-1 \\ k_2 = -1. \end{cases}$

Case 3.

$$\begin{cases} \Xi_1 = 0 \\ \Xi_2 = -p(p-t_1). \end{cases}$$

It is easy to get $k_2 \notin \mathbb{Z}$.

Case 4.

$$\begin{cases} \Xi_1 = 0 \\ \Xi_2 = -p(p-t_2). \end{cases}$$

It is easy to get $k_2 \notin \mathbb{Z}$.

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.

Case 7.

$$\begin{cases} \Xi_1 = p(2p-t_1-t_2) \\ \Xi_2 = -p(p-t_1). \end{cases}$$

It is easy to get $k_1 \notin \mathbb{Z}$.

Case 8.

$$\begin{cases} \Xi_1 = p(2p-t_1-t_2) \\ \Xi_2 = -p(p-t_2). \end{cases}$$

It is easy to get $k_1 \notin \mathbb{Z}$.

Case 9.

$$\begin{cases} \Xi_1 = p(p-t_1) \\ \Xi_2 = 0. \end{cases}$$

It is easy to get $k_1 \notin \mathbb{Z}$.

Case 10.

$$\begin{cases} \Xi_1 = p(p-t_1) \\ \Xi_2 = -p(2p-t_1-t_2). \end{cases}$$

It is easy to get $k_1 \notin \mathbb{Z}$.

Case 11.

$$\begin{cases} \Xi_1 = p(p-t_1) \\ \Xi_2 = -p(p-t_1). \end{cases}$$

One has $\begin{cases} k_1 = t_1 - p \\ k_2 = -2 - t_1 + p. \end{cases}$ As case 6, it is impossible.

Case 12.

$$\begin{cases} \Xi_1 = p(2p-t_1-t_2) \\ \Xi_2 = -p(p-t_2). \end{cases}$$

One has $k_1 \notin \mathbb{Z}$ or $\begin{cases} k_1 = t_2 - p = t_1 - p \\ k_2 = -2 - t_2 + p = -2 - t_1 + p. \end{cases}$ As Case 6, it is impossible.

Case 13.

$$\begin{cases} \Xi_1 = p(p-t_2) \\ \Xi_2 = 0. \end{cases}$$

It is easy to get $k_1 \notin \mathbb{Z}$.

Case 14.

$$\begin{cases} \Xi_1 = p(p-t_2) \\ \Xi_2 = -p(2p-t_1-t_2). \end{cases}$$

It is easy to get $k_1 \notin \mathbb{Z}$.

Case 15.

$$\begin{cases} \Xi_1 = p(p-t_2) \\ \Xi_2 = -p(p-t_1). \end{cases}$$

One has $k_1 \notin \mathbb{Z}$ or $\begin{cases} k_1 = t_2 - p = t_1 - p \\ k_2 = -2 - t_2 + p = -2 - t_1 + p. \end{cases}$ As case 6, it is impossible.

Case 16.

$$\begin{cases} \Xi_1 = p(p-t_2) \\ \Xi_2 = -p(p-t_2). \end{cases}$$

One has $k_1 \notin \mathbb{Z}$ or $\begin{cases} k_1 = t_2 - p \\ k_2 = -2 - t_2 + p. \end{cases}$ As case 6, it is impossible.

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

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

which are

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

Checking directly, it is easy to know for Case $2'$ and $3'$, we have

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

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

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

In fact, it is easy to check that

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

■

**Lemma 4.8** *If $\tau > 5$,*

$$|\mathcal{R}^0_{20,4}| \leq \cdot \epsilon^{\frac{\beta_0}{8p}}.$$

*Proof.* Suppose $k \neq 0$. Without losing generality, also suppose $|n| \leq |m|$. If $|n| < |m|$, we have the following discussion.

If $|m| > c|k|$ and $c$ is large enough, we have

$$\frac{m^2 - n^2}{|m| - |n| + 1} \geq \frac{c}{2}|k| > |\langle k, \lambda_0 \rangle|.$$

This means

$$|\langle k, \omega \rangle + \Omega_n - \Omega_m| > \frac{\epsilon^{\frac{\beta_0}{4}}(||n| - |m|| + 1)}{|k|^{2p\tau}}.$$

Therefore, $|n| < |m| \leq c|k|$. It follows that

$$|\bigcup_{\substack{0<|k|\leq K \\ |n|\neq|m|}} \{\xi : |\langle k, \omega \rangle + \Omega_n - \Omega_m| < \frac{\epsilon^{\frac{\beta_0}{4}}(||n| - |m|| + 1)}{|k|^{2p\tau}}\}| \leq \sum_{0<|k|\leq K} \frac{c\epsilon^{\frac{\beta_0}{8p}}}{|k|^{\tau-3}}.$$

If $|n| = |m|$, it is obvious to obtain

$$|\bigcup_{\substack{0<|k|\leq K \\ |n|=|m|}} \{\xi : |\langle k, \omega \rangle + \Omega_n - \Omega_m| < \frac{\epsilon^{\frac{\beta_0}{4}}(||n| - |m|| + 1)}{|k|^{2p\tau}}\}| \leq \sum_{0<|k|\leq K} \frac{c\epsilon^{\frac{\beta_0}{8p}}}{|k|^{\tau}}.$$

If $k = 0$, it is easy to get $|m| \neq |n|$. It follows that

$$|\bigcup_{\substack{k=0 \\ |n|\neq|m|}} \{\xi : |\langle k, \omega \rangle + \Omega_n - \Omega_m| < \epsilon^{\frac{\beta_0}{4}}(||n| - |m|| + 1)\}| = 0.$$

Now the conclusion is clear. ∎

**Lemma 4.9** *If $\tau > 2$,*

$$|\mathcal{R}^0_{20,5}| \leq \cdot \epsilon^{\frac{\beta_0}{8p}}.$$

*Proof.* We only give a sketch. Firstly, $k = 0$, $g_4(\xi) = -a_t^2$. The conclusion is clear. If $k \neq 0$, we will compute $g_4^{4p,0}$ and $g_4^{0,4p}$ as Lemma 4.4. It is easy to show that $(g_4^{4p,0})^2 + (g_4^{0,4p})^2 \neq 0$ except when $k_1 = p - t$ and $k_2 = t - p$. In fact, it is easy to check that $g_4(\xi) = -a_t^2$ when $k_1 = p - t$ and $k_2 = t - p$. ∎

**Lemma 4.10** *If $\tau > 2$,*

$$|\mathcal{R}^0_{20,6}| \leq \cdot \epsilon^{\frac{\beta_0}{8p}}.$$

*Proof.* We will show there exist nonzero coefficients in $g_5$ for any $k$, $|k| + |t_1 - t_2| \neq 0$ and $\langle k, \lambda_0 \rangle + i^2_{t_1} - i^2_{t_2} = 0$, where $t_1, t_2 \in \mathcal{T}$. If $t_1 \neq t_2$, since $\langle k, \lambda_0 \rangle + i^2_{t_1} - i^2_{t_2} = 0$, one gets $k \neq 0$. If $t_1 = t_2$, from $|k| + |t_1 - t_2| \neq 0$, we also have $k \neq 0$.

Write $x = k_1 + k_2(p+1)$, $y = k_1(p+1) + k_2, x' = k_1 + \frac{1}{2}k_2 p$, $y' = \frac{1}{2}k_1 p + k_2$. Denote

$$\chi_1 = (p - t_1)(t_2 - t_1) + (p - 2t_1 + t_2)(t_2 - p), \ \chi_2 = (t_2 - p)(p - t_1)(t_2 - t_1).$$

It is easy to check that

$$g_5^{8p,0} = c\{x^4 + p(2t_1 - 2t_2)x^3 + \chi_1 p^2 x^2 - \chi_2 p^3 x\}, \tag{4.14}$$

$$g_5^{0,8p} = c\{y^4 - p(2t_1 - 2t_2)y^3 + \chi_1 p^2 y^2 + \chi_2 p^3 y\}, \tag{4.15}$$

$$\begin{aligned} g_5^{8p-2,2} &= c\{4x^3x' + (2t_2 - 2t_1)[-3px' + (1 - \frac{p}{2})x]x^2 \\ &+ \chi_1 p(2px' + (p-2)x)x + \chi_2 p^2[-px' + 3(1 - \frac{p}{2})x]\}, \end{aligned} \tag{4.16}$$

$$\begin{aligned} g_5^{2,8p-2} &= c\{4y^3y' - (2t_2 - 2t_1)[-3py' + (1 - \frac{p}{2})y]y^2 \\ &+ \chi_1 p(2py' + (p-2)y)y - \chi_2 p^2[-py' + 3(1 - \frac{p}{2})y]\}. \end{aligned} \tag{4.17}$$

From $g_5^{8p,0} = 0$ and $g_5^{0,8p} = 0$ 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 the following 16 cases.

Case 1.

$$\begin{cases} x = 0 \\ y = 0. \end{cases}$$

One has $k = 0$. It contradicts with $k \neq 0$.

Case 2.

$$\begin{cases} x = 0 \\ y = (t_1 - t_2)p. \end{cases}$$

One has $k = 0$ or $t_2 \geq p$. It is impossible.

Case 3.

$$\begin{cases} x = 0 \\ y = p(t_1 - p). \end{cases}$$

It is easy to get $k_2 \notin \mathbb{Z}$.

Case 4.

$$\begin{cases} x = 0 \\ y = p(p - t_2). \end{cases}$$

It is easy to get $k_2 \notin \mathbb{Z}$.

Case 5.

$$\begin{cases} x = (t_2 - t_1)p \\ y = 0. \end{cases}$$

One has $k_1 \notin \mathbb{Z}$ or $k = 0$.

Case 6.

$$\begin{cases} x = p(t_2 - t_1)p \\ y = (t_1 - t_2)p. \end{cases}$$

One has $\begin{cases} k_1 = t_1 - t_2 \\ k_2 = t_2 - t_1. \end{cases}$

Case 7.

$$\begin{cases} x = (t_2 - t_1)p \\ y = p(t_1 - p). \end{cases}$$

It is easy to get $k_2 \notin \mathbb{Z}$.

Case 8.

$$\begin{cases} x = (t_2 - t_1)p \\ y = -p(p - t_2). \end{cases}$$

It is easy to get $k_2 \notin \mathbb{Z}$.

Case 9.

$$\begin{cases} x = p(p - t_1) \\ y = 0. \end{cases}$$

It is easy to get $k_1 \notin \mathbb{Z}$.

Case 10.

$$\begin{cases} x = p(p - t_1) \\ y = (t_1 - t_2)p. \end{cases}$$

It is easy to get $k_1 \notin \mathbb{Z}$.

Case 11.

$$\begin{cases} x = p(p-t_1) \\ y = p(p-t_2). \end{cases}$$

One has $k_2 \notin \mathbb{Z}$ or $\begin{cases} k_1 = t_1+1 \\ k_2 = t_2+1 \end{cases}$ and $t_1+t_2 = p-2$.

Case 12.

$$\begin{cases} x = p(p-t_1) \\ y = p(t_1-p). \end{cases}$$

One has $\begin{cases} k_1 = t_1 - p \\ k_2 = p - t_1. \end{cases}$

Case 13.

$$\begin{cases} x = p(t_2-p) \\ y = 0. \end{cases}$$

It is easy to get $k_1 \notin \mathbb{Z}$.

Case 14.

$$\begin{cases} x = p(t_2-p) \\ y = p(t_1-t_2). \end{cases}$$

It is easy to get $k_2 \notin \mathbb{Z}$.

Case 15.

$$\begin{cases} x = p(t_2-p) \\ y = p(t_1-p). \end{cases}$$

One has $k_2 \notin \mathbb{Z}$ or $\begin{cases} k_1 = -t_2-1 \\ k_2 = -t_1-1 \end{cases}$ and $t_1+t_2 = p-2$.

Case 16.

$$\begin{cases} x = p(t_2-p) \\ y = p(p-t_2). \end{cases}$$

One has $\begin{cases} k_1 = p-t_2 \\ k_2 = t_2-p. \end{cases}$

The above proof shows that except the following 5 cases we have

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

if $p > 2$, which are

$$(1')\begin{cases} k_1 = t_1 - t_2 \\ k_2 = t_2 - t_1, \end{cases} (2')\begin{cases} k_1 = t_1 - p \\ k_2 = p - t_1, \end{cases} (3')\begin{cases} k_1 = p - t_2 \\ k_2 = t_2 - p. \end{cases}$$

$$(4')\begin{cases} k_1 = t_1+1 \\ k_2 = t_2+1 \\ t_1+t_2 = p-2, \end{cases} (5')\begin{cases} k_1 = -t_2-1 \\ k_2 = -t_1-1 \\ t_1+t_2 = p-2. \end{cases}$$

Checking directly, it is easy to know for Case $1'$, $2'$ and $3'$, we have $g_5^{8p-2,2} \neq 0$(note $p > 2$). For the remaining Case $4'$ and $5'$, we clearly prove Case $4'$ since the latter case is totally similar. In fact, a simple transformation can turn Case $5'$ into Case $4'$. The following, we will prove $g_5^{2,8p-2} \neq 0$ under Case $4'$.

First, we must have $t_1 \geq t_2$ under Case $4'$. If not, one can easily prove

$$k_1 n_1^2 + k_2 n_2^2 \neq i_{t_2}^2 - i_{t_1}^2$$

from the direct computation.

For simplicity, we write $x_1 = p - t_1,\ y_1 = p - t_2,\ 2 \leq x_1 \leq y_1 \leq p$. It follows that

$$y' = (x_1 - 1) + \frac{1}{2}(y_1 - 1)p,\ \chi_1 = x_1^2 + y_1^2 - 3x_1 y_1,\ \chi_2 = x_1 y_1 (y_1 - x_1).$$

Under the above notations, one has

$$\begin{aligned} g_5^{2,8p-2} &= c\{4(k_1+1)^2y' + (t_2-t_1)(k_1+1)^2(p-2) + 6y'(t_2-t_1)(k_1+1) \\ &\quad + \chi_1[(p-2)(k_1+1)+2y'] + \frac{3}{2}(p-2)\chi_2 + (p-t_1)(t_1-t_2)y'\} \\ &= c(I-II), \end{aligned}$$

Where

$$\begin{aligned} I &= 4y_1[(x_1-1)+\frac{1}{2}(y_1-1)p] + \frac{3}{2}(p-2)x_1y_1(y_1-x_1) + x_1(y_1-x_1)[(x_1-1)+\frac{1}{2}(y_1-1)p] \\ &\quad + (p-2)y_1(x_1^2+y_1^2) + 2[(x_1-1)+\frac{1}{2}(y_1-1)p](x_1^2+y_1^2), \end{aligned}$$

and

$$\begin{aligned} II &= (p-2)(y_1-x_1)y_1^2 + 6[(x_1-1)+\frac{1}{2}(y_1-1)p](y_1-x_1)y_1 + 3x_1y_1^2(p-2) \\ &\quad + 6x_1y_1[(x_1-1)+\frac{1}{2}(y_1-1)p]. \end{aligned}$$

It is easy to check that

$$I-II = \frac{p}{2}(x_1y_1-x_1^2) + 2(x_1^3-x_1^2) + 2x_1y_1 > 0.$$

This shows that $g_5^{2,8p-2} \neq 0$.

If $p=2$, we have $t=0$ and $a_t = \xi_1^2\xi_2^2$ in this case. We only need discuss Case $2'$ and Case $3'$(Case $1'$ doesn't exist and Case $4'$ and $5'$ are the same). Computing directly in the two cases, we both have

$$g_5 = -288\xi_1^4\xi_2^4(\xi_1^4-\xi_2^4)^2.$$

The conclusion is obvious. ■

Combined with above lemmata, we have the following lemma.

**Lemma 4.11** *If $\tau > 5$,*

$$|\mathcal{R}^0| \leq \cdot\epsilon^{\frac{\beta_0}{8p}}$$

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 $|k| \leq K_0$ and all the parameters $\xi \in \mathcal{O}$, which belong to the set $\mathcal{O}_0 = \mathcal{O}\setminus\mathcal{R}^0$, satisfy* [3] *the following conditions*

$$\begin{aligned} |\langle k,\omega\rangle^{-1}| &\leq \frac{|k|^{2p\tau}}{\epsilon^{\frac{\beta_0}{4}}}, \quad k\neq 0, \\ |(\langle k,\omega\rangle + \Omega_n)^{-1}| &\leq \frac{c\max\{|k|^{4p\tau+2},1\}}{\epsilon^{\frac{\beta_0}{2}}}, \quad n\notin\mathcal{N}, \\ \|(\langle k,\omega\rangle I_2 + A_{i_t})^{-1}\| &\leq \frac{c\max\{|k|^{4p\tau+2},1\}}{\epsilon^{\frac{\beta_0}{2}}}, \quad t\in\mathcal{T}, \end{aligned} \tag{4.18}$$

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

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

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

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

$$\|((\langle k,\omega\rangle+\Omega_n)I_2+A_{i_t})^{-1}\|\le\frac{c\max\{|k|^{8p\tau+6},1\}}{\epsilon^{\beta_0}(||i_{t_1}|-|n||+1)},\tag{4.19}$$

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

$$\|(I_2\otimes(\langle k,\omega\rangle I_2+A_{i_{t_2}})+A_{i_{t_1}}\otimes I_2)^{-1}\|\le\frac{c\max\{|k|^{8p\tau+6},1\}}{\epsilon^{\beta_0}(||i_{t_1}|-|i_{t_2}||+1)},\quad t_1,t_2\in\mathcal{T},\tag{4.20}$$

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

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

$$\|(I_2\otimes(\langle k,\omega\rangle I_2-A_{i_{t_2}})+A_{i_{t_1}}\otimes I_2)^{-1}\|\le\frac{c\max\{|k|^{8p\tau+6},1\}}{\epsilon^{\beta_0}(||i_{t_1}|-|i_{t_2}||+1)},\tag{4.21}$$

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

$$\|(((\langle k,\omega\rangle+\Omega_n)I_2-A_{i_t})^{-1}\|\le\frac{c\max\{|k|^{8p\tau+6},1\}}{\epsilon^{\beta_0}(||i_t|-|n||+1)},\tag{4.22}$$

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 the Appendix.

**Remark 4.1** *We must point out that Lemma 4.12 omits one inequality of (3.1), which is*

$$|\tilde\Omega_{i_t}-\tilde\Omega_{j_t}|\le\frac{c}{\epsilon^{\beta_0}(||i_t|-|j_t||+1)},\ t\in\mathcal{T}.\tag{4.23}$$

*But, from*

$$|\tilde\Omega_{i_t}-\tilde\Omega_{j_t}|=|\frac{1}{\sqrt{4a_t^2+(p-t)^2A^2}}|\le c,$$

*it is easy to know that (4.23) holds naturally.*

**Remark 4.2** *From (4.22) to the corresponding inequalities in (3.1), one inequality is needed. We need the simple inequality as the following:*

$$\frac{1}{||i_t|-|n||+1}\le\cdot\frac{1}{||j_t|-|n||+1}.\tag{4.24}$$

*(4.24) can be easily proved by discussing two cases. One is $|i_t|\ne|n|$ and the other is $|i_t|=|n|$.*

**Remark 4.3** *(3.1) 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

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

(where $\nu\ge 0$) at remaining KAM steps. We have the following lemmata which give the corresponding measure estimates. The proofs of the following lemmata are similar, we only give one and omit the others.

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

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

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

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

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

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

*Proof.* Define $v_1=(2p,0)^T$, and $v_2=(0,2p)^T$. It is easy to get

$$\langle\frac{k}{|k|},\omega\rangle=const.+c_p(\frac{k_1}{|k|}(p+1)+\frac{k_2}{|k|}(p+1)^2)\xi_1^{2p}+\cdots+c_p(\frac{k_1}{|k|}(p+1)^2+\frac{k_2}{|k|}(p+1))\xi_2^{2p}.$$

Denote $f=\frac{1}{c_p(2p)!(p+1)}\langle\frac{k}{|k|},\omega\rangle$ and $F=\langle k,\omega_{\nu+1}\rangle+\tilde{\Omega}_{\nu+1,n}-\tilde{\Omega}_{\nu+1,m}$. It follows that

$$\begin{cases}\frac{\partial^{v_1}f}{\partial\xi^{v_1}}=\frac{k_1}{|k|}+(p+1)\frac{k_2}{|k|},\\ \frac{\partial^{v_2}f}{\partial\xi^{v_2}}=(p+1)\frac{k_1}{|k|}+\frac{k_2}{|k|}.\end{cases}\tag{4.25}$$

Write $\beta=(\beta_0,\beta_1)=(\frac{k_1}{|k|},\frac{k_2}{|k|})$. Obviously, $|\beta|_1=1$. Denote

$$D=\begin{pmatrix}1 & p+1\\ p+1 & 1\end{pmatrix}.$$

Computing directly, it is easy to obtain $|D\beta|_1\geq 3$. It follows that for any $k\neq 0$, there exists $v_{i_0}$, $i_0=1$ or 2 so that $|\frac{\partial^{v_{i_0}}f}{\partial\xi^{v_{i_0}}}|\geq\frac{3}{2}$. This means that

$$|\frac{\partial^{v_{i_0}}\langle k,\omega\rangle}{\partial\xi^{v_{i_0}}}|\geq c_0|k|.$$

Note

$$|\frac{\partial^{v_{i_0}}\langle k,\omega_{\nu+1}-\omega\rangle}{\partial\xi^{v_{i_0}}}|\leq c\epsilon_0|k|,$$

$$|\frac{\partial^{v_{i_0}}\tilde{\Omega}_n}{\partial\xi^{v_{i_0}}}|\leq c,$$

$$|\frac{\partial^{v_{i_0}}(\tilde{\Omega}_{\nu+1,n}-\tilde{\Omega}_n)}{\partial\xi^{v_{i_0}}}|\leq c\epsilon_0,$$

$$|\frac{\partial^{v_{i_0}}\tilde{\Omega}_m}{\partial\xi^{v_{i_0}}}|\leq c,$$

$$|\frac{\partial^{v_{i_0}}(\tilde{\Omega}_{\nu+1,m}-\tilde{\Omega}_m)}{\partial\xi^{v_{i_0}}}|\leq c\epsilon_0,$$

it follows that

$$|\langle k,\omega_{\nu+1}-\omega\rangle+\tilde{\Omega}_{\nu+1,n}-\tilde{\Omega}_{\nu+1,m}|\leq 2c+1+c\epsilon_0|k|.$$

Therefore, for $|k|>\frac{8c}{c_0}$,

$$|\frac{\partial^{v_{i_0}}F}{\partial\xi^{v_{i_0}}}|\geq 1.$$

The following we estimate $\mathcal{R}^{\nu+1}_{11,1}$. Write $\mathcal{R}^{\nu+1}_{11,1}=\mathcal{R}_{+}+\mathcal{R}_{-}$, where $\mathcal{R}_{+}=\mathcal{R}^{\nu+1}_{11,1}\cap\{|n|\neq|m|\}$ and $\mathcal{R}_{-}=\mathcal{R}^{\nu+1}_{11,1}\cap\{|n|=|m|\}$. For the measure of $\mathcal{R}_{+}$, it is standard(See [15]). It is easy to get

$$|\mathcal{R}_{+}|\leq\cdot\epsilon_{\nu+1}^{\frac{\beta_0}{2p}}$$

when $\tau>2$ and $|k|>\frac{8c}{c_0}$. Write $\mathcal{R}_{-}=\mathcal{R}^{1}_{-}\cup\mathcal{R}^{2}_{-}$, where $\mathcal{R}^{1}_{-}=\mathcal{R}^{\nu+1}_{11,1}\cap\{n=m\}$ and $\mathcal{R}^{1}_{-}=\mathcal{R}^{\nu+1}_{11,1}\cap\{n=-m\}$. It is obvious that $|\mathcal{R}_{+}|\leq\cdot\epsilon_{\nu+1}^{\frac{\beta_0}{2p}}$ when $\tau>2$ and $|k|>\frac{8c}{c_0}$. When $n=-m$, we know that $|k_1n_1+k_2n_2|=|2n|$. This means $|n|\leq c|k|$. Therefore, when $\tau>2$ and $|k|>\frac{8c}{c_0}$ we have

$$\begin{aligned}|\mathcal{R}^{1}_{-}|&\leq\sum_{|k|>K_\nu,|n|\leq c|k|}\cdot\left(\frac{\epsilon^{\beta_0}_{\nu+1}}{|k|^{8p\tau+6}}\right)^{\frac{1}{2p}}\\&\leq\cdot\epsilon_{\nu+1}^{\frac{\beta_0}{2p}}.\end{aligned}$$

Combining with all the estimates before, one has $|\mathcal{R}^{\nu+1}_{11,1}|\leq\cdot\epsilon_{\nu+1}^{\frac{\beta_0}{2p}}$.

**Remark 4.4** *In fact, there is a standard proof of Lemma 4.18(see Lemma 5.7 in [15]). Here we give a minor different proof.*

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

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

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

Combining with all the lemmata before, we have

**Lemma 4.22** *If $\tau>2$ and $K_\nu>\frac{8c}{c_0}$, $|\mathcal{R}^{\nu+1}|\leq\cdot\epsilon_{\nu+1}^{\frac{\beta_0}{2p}}(\nu\geq0)$.*

Note (3.10), 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.

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

∎

# 5 Appendix

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

$$F_{k\alpha\beta}\neq0,\ \text{ implies } k_1n_1+k_2n_2+\sum_n(-\alpha_n+\beta_n)n=0 \text{ for any k, } \alpha,\ \beta,$$

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

**Lemma 5.1** *Given $n_1,\ n_2\in\mathbb{Z}$ and $n_1\neq n_2$, consider two real analytic functions*

$$F(x,y,z,\bar z),\ G(x,y,z,\bar z)$$

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

For the proof, refer to Lemma 2.4 in [13].

Given $n_1,\ n_2$ and specially chosen subscripts set $\mathcal{J}=\{j_0,\cdots,j_{p-2}\}$ 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

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

imply

$$k_1n_1+k_2n_2+\sum_{n\in\mathbb{Z}\setminus\{n_1,n_2\}}(-\alpha_n+\beta_n)n=(n_1-n_2)\sum_{t=0}^{p-2}(\alpha_{j_t}-\beta_{j_t})(p-t) \tag{5.1}$$

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

Similar as Lemma 5.1, we have the following lemma.

**Lemma 5.2** *Given $n_1,n_2\in\mathbb{Z}$ and specially chosen subscripts set $\mathcal{J}=\{j_0,\cdots,j_{p-2}\}$ 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 $\mathcal{J}$, then so does $\{F,G\}$.*

*For the proof, refer to Lemma 2.4 in [13].*

The following lemma is needed in Section 2.

**Lemma 5.3** *$P_0^+$ 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,

$$\begin{aligned}P_0^+\ =&\sum_{k,\alpha,\beta}P_{0,k\alpha\beta}(y^+)e^{i\langle k,x^+\rangle}(\Pi_{i\notin\mathcal{N}}w_i^{\alpha_i}\bar w_i^{\beta_i})(p_{11,0}w_{i_0}^++p_{12,0}w_{j_0}^+)^{\alpha_{i_0}}(p_{21,0}w_{i_0}^++p_{22,0}w_{j_0}^+)^{\alpha_{j_0}}\\ &\cdots(p_{11,p-2}w_{i_{p-2}}^++p_{12,p-2}w_{j_{p-2}}^+)^{\alpha_{i_{p-2}}}(p_{21,p-2}w_{i_{p-2}}^++p_{22,p-2}w_{j_{p-2}}^+)^{\alpha_{j_{p-2}}}\\ &(p_{11,0}\bar w_{i_0}^++p_{12,0}\bar w_{j_0}^+)^{\beta_{i_0}}(p_{21,0}\bar w_{i_0}^++p_{22,0}\bar w_{j_0}^+)^{\beta_{j_0}}\\ &\cdots(p_{11,p-2}\bar w_{i_{p-2}}^++p_{12,p-2}\bar w_{j_{p-2}}^+)^{\beta_{i_{p-2}}}(p_{21,p-2}\bar w_{i_{p-2}}^++p_{22,p-2}\bar w_{j_{p-2}}^+)^{\beta_{j_{p-2}}}.\end{aligned} \tag{5.2}$$

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

$$k_1n_1+k_2n_2+\sum_{i\in\mathbb{Z}}(-\alpha_i+\beta_i)i=(n_1-n_2)\sum_{t=0}^{p-2}(\alpha_{j_t}-\beta_{j_t})(p-t). \tag{5.3}$$

We write every term of which its coefficient might be nonzero in (5.2). It is

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

where $k$, $\alpha$, $\beta$ satisfy (5.3) and

$$
\begin{aligned}
&0\le k_t^1\le\alpha_{i_t},\ 0\le k_t^2\le\alpha_{j_t},\\
&0\le l_t^1\le\beta_{i_t},\ 0\le l_t^2\le\beta_{j_t},\ t\in\mathcal{T}.
\end{aligned}
$$

Then from (5.3), one gets

$$
\begin{aligned}
&k_1n_1+k_2n_2+\sum_{i\notin\mathcal{N}}i(\beta_i-\alpha_i)\\
&+\sum_{t=0}^{p-2}[i_t(l_t^1+l_t^2-k_t^1-k_t^2)+j_t(\beta_{i_t}+\beta_{j_t}-\alpha_{i_t}-\alpha_{j_t}-(l_t^1+l_t^2-k_t^1-k_t^2))]\\
&=k_1n_1+k_2n_2+\sum_{i\notin\mathcal{N}}i(\beta_i-\alpha_i)+\sum_{t=0}^{p-2}[i_t(-\alpha_{i_t}+\beta_{i_t})+j_t(-\alpha_{j_t}+\beta_{j_t})]\\
&+\sum_{t=0}^{p-2}[-i_t(\beta_{i_t}-\alpha_{i_t}+k_t^1+k_t^2-l_t^1-l_t^2)+j_t(\beta_{i_t}-\alpha_{i_t}+k_t^1+k_t^2-l_t^1-l_t^2)]\\
&=(n_1-n_2)\sum_{t=0}^{p-2}(\alpha_{j_t}-\beta_{j_t})(p-t)+\sum_{t=0}^{p-2}(\beta_{i_t}-\alpha_{i_t}+k_t^1+k_t^2-l_t^1-l_t^2)(j_t-i_t)\\
&=(n_1-n_2)\sum_{t=0}^{p-2}(\alpha_{j_t}-\beta_{j_t})(p-t)+\sum_{t=0}^{p-2}(\beta_{i_t}-\alpha_{i_t}+k_t^1+k_t^2-l_t^1-l_t^2)(n_2-n_1)(p-t)\\
&=\sum_{t=0}^{p-2}(n_1-n_2)(p-t)[(\alpha_{i_t}+\alpha_{j_t}-k_t^1-k_t^2)-(\beta_{i_t}+\beta_{j_t}-l_t^1-l_t^2)].
\end{aligned}
$$

■

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

*Proof.*

Case 1.

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

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

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

Subcase 3. $n\in\{-j_0,\ \cdots,-j_{p-2}\}$. 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.

## 5.2 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 (4.19). Write

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

Obviously,

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

The following we will prove

$$\|(M_1')^{-1}\| \leq \frac{c\max\{|k|^{4p\tau+4},1\}}{\epsilon^{\frac{\beta_0}{2}}(||i_t|-|n||+1)}. \tag{5.4}$$

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

$$cK_0 \leq \frac{1}{\epsilon^{6p}}. \tag{5.5}$$

Therefore,

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

It follows

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

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

$$|n| \leq c|k|. \tag{5.6}$$

Therefore

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

Thus, it is easy to get (5.4).

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

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

From

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

and (5.6), it follows

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

Combined with above two cases, the conclusion is clear.

In the following we will prove (4.20). Write

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

Note

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

has the same eigenvalues as $M_2$(see Lemma 5.3 in [21]), this means that there exists an orthogonal matrix $P_{t_1,t_2}$ so that

$$P_{t_1,t_2}^T M_2' P_{t_1,t_2} = M_2.$$

Denote $g^2 = det(M_2')$. (4.20) is clear from the equality

$$\|(M_2')^{-1}\| \le \frac{c\max\{1, |k|^{8p\tau+6}\}}{\epsilon^{\beta_0}}. \tag{5.7}$$

We will obtain (5.7) in the following two cases.

Case 1.

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

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

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

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 $\mathcal{R}_{20,3}^0$, it follows that

$$|g^2| \ge \frac{\epsilon^{\beta_0}}{\max\{1, |k|^{8p\tau}\}}. \tag{5.8}$$

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

$$|m_{ij}| \le c|k|^6. \tag{5.9}$$

Therefore,

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

## Acknowledgements

The author thanks Professor Xiaoping Yuan and Jun Yan for their help. The author would also like to thank Professor Jiangong You for his encouragement. Part of this paper was discussed with Huawei Niu.

# References


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

[2] Bourgain, J., Construction of quasi-periodic solutions for Hamiltonian perturbations of linear equations and applications to nonlinear PDE, International Mathematics Research Notices, 1994, 475-497.

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

[4] Bourgain, J., Quasiperiodic solutions of Hamiltonian perturbations of 2D linear Schrödinger equations, Annals of Mathematics, 148(1998), 363-439.

[5] Bourgain, J., Green's function estimates for lattice Schrödinger operators and applications, Annals of Mathematics Studies, 158, Princeton University Press, Princeton, NJ, 2005.

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

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

[8] H.L. Eliasson and S.B. Kuksin, Homological equations for the non-linear Schrödinger equation, Preprint(2005).

[9] H.L. Eliasson and S.B. Kuksin, Infinite Töplitz-Lipschitz matrices and operators, Preprint(2005).

[10] H.L. Eliasson and S.B. Kuksin, KAM for the nonlinear Schrödinger equation, Preprint(2006).

[11] J. Geng and J. You, A KAM theorem for one dimensional Schrödinger equation with periodic boundary conditions, J. Diff. Eqs., 209(2005), 1-56.

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

[13] J. Geng and Y. Yi, Quasi-periodic solutions in a nonlinear Schrödinger equation, J. Diff. Eqs., 233(2007), 512-542.

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

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

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

[17] Pöschel, J., A KAM Theorem for some nonlinear partial differential equations, Ann. Sc. Norm. sup. Pisa CI. sci., 23(1996), 119-148.

[18] J. Xu, J. You, and Q. Qiu, A KAM theorem of degenerate infinite dimensional Hamiltonian systems(I)(II), Science in China(Series A), Vol.39 No.4.(1996), 372-394.

[19] J. Xu, J. You, and Q. Qiu, Invariant tori for nearly integrable Hamiltonian systems with degenaracy, Math. Z., 226(1997), 375-387.

[20] J. Xu and J. You, Persistence of lower-dimensional tori under the first Melnikov's non-resonance condition, J. Math. Pure Appl. 80, 10(2001), 1045-1067.

[21] J. You, Perturbations of lower dimensional tori for Hamiltonian systems, J. Diff. Eqs., 152(1999), 1-29.

[22] Yuan, X., A KAM Theorem with applications to partial differential equations of higher dimension, Preprint(2006).