Small-divisor equation of higher order with large variable coefficient and application to the coupled KdV equation

Meina Gao* Jing Zhang
School of Mathematical Sciences
Fudan University, Shanghai 200433, P R China

April 26, 2010

Abstract: In this paper, we establish an estimate for the solutions of small-divisor equation of higher order with large variable coefficients. Then by formulating an infinite dimensional KAM theorem which allows for multiple normal frequencies and unbounded perturbations, we prove that there are many periodic solutions for the coupled KdV equation subject to small Hamiltonian perturbations.

1 Introduction and main results

Consider a nearly integrable Hamiltonian of infinite dimension

$$H = \sum_{j=1}^{\iota} \omega_j y_j + \sum_{j\geq 1} \Omega_j z_j \bar{z}_j + \varepsilon R(x,y,z,\bar{z}), \quad (x,y,z,\bar{z}) \in (\mathbb{C}^\iota/2\pi\mathbb{Z}^\iota) \times \mathbb{C}^\iota \times \ell^{a,p} \times \ell^{a,p}, \tag{1.1}$$

where $1 \leq \iota < \infty$, and $\ell^{a,p}$ ($a \geq 0$, $p \geq 0$) is the Hilbert space of all complex sequences $z = (z_j : j \geq 1)$ satisfying

$$||z||_{a,p}^2 = \sum_{j\geq 1} |z_j|^2 j^{2p} e^{2aj} < \infty.$$

Let $\mathscr{O}$ be a neighbourhood of $\mathbb{T}^\iota \times \{y=0\} \times \{z=0\} \times \{\bar{z}=0\}$ in $(\mathbb{C}^\iota/2\pi\mathbb{Z}^\iota) \times \mathbb{C}^\iota \times \ell^{a,p} \times \ell^{a,p}$, here $\mathbb{T}^\iota = \mathbb{R}^\iota/2\pi\mathbb{Z}^\iota$ is the usual ι-torus. Assume the vector field X_R of R satisfies:

$$X_R : \mathscr{O} \to (\mathbb{C}^\iota/2\pi\mathbb{Z}^\iota) \times \mathbb{C}^\iota \times \ell^{a,\bar{p}} \times \ell^{a,\bar{p}},$$

we call R a bounded perturbation if $\bar{p} \geq p$, and unbounded one if $\bar{p} < p$. The KAM (Kolmogorov-Arnold-Moser) theory for Hamiltonian H with bounded perturbations has been deeply investigated by Kuksin [10], Wayne [18], Bourgain [4], Eliasson and Kuksin [7] among others. There are too many references to list all of them. Concerning the KAM theory for Hamiltonian H with unbounded perturbations, the only previous result is due to Kuksin [12]. In order to obtain KAM tori for (1.1) with unbounded perturbations, Kuksin [11], [12] initiated the study of the first order differential equation on an ι-dimensional torus

$$-\mathrm{i}\omega \cdot \partial_\phi u(\phi) + \lambda u(\phi) + b(\phi)u(\phi) = f(\phi), \quad \mathrm{i}^2 = -1,\ \phi \in \mathbb{T}^\iota, \tag{1.2}$$

where $u(\phi) : \mathbb{T}^\iota \to \mathbb{C}$ is the unknown function, while $b(\phi),\ f(\phi) : \mathbb{T}^\iota \to \mathbb{C}$ are known functions, being analytic in the complex s-vicinity of $\mathbb{T}^\iota$:

$$\mathbb{T}_s^\iota := \{\phi \in \mathbb{C}^\iota/2\pi\mathbb{Z}^\iota : |\mathrm{Im}\phi| < s\}, \quad s > 0.$$

*E-mail: 071018019@fudan.edu.cn

This is one of the basic homological equations in establishing the KAM theory for infinite dimensional Hamiltonian with unbounded perturbations. Since (1.2) is a scalar equation, it is easy to show that for most values of the vector ω the solution $u(\phi)$ is bounded by

$$\text{Const. } e^{|b|_s}|f|_s, \tag{1.3}$$

where the constant depends on λ and ω, and $|\cdot|_s$ is the sup-norm in $\mathbb{T}_s^\iota$. When KAM theory applies to problems from PDEs (partial differential equations), such as the KdV (Korteweg-de Vries) equation with Hamiltonian perturbations, we often get $|b|_s \gg 1$, so the bound (1.3) becomes non-satisfactory. It was proved by Kuksin in [11], [12] that if $|\lambda| \gg |b|_s$, then the large factor $e^{|b|_s}$ may be removed form (1.3). The basic idea due to Kuksin [11], [12] is as follows:

Approximating ω by a rational vector $\tilde{\omega} \in \mathbb{Q}^\iota$ with periodic T, that is, $\tilde{\omega} = \frac{2\pi}{T}k$ with $k \in \mathbb{Z}^\iota$, and replacing the term $-\mathrm{i}\omega \cdot \partial_\phi u(\phi)$ by $-\mathrm{i}\tilde{\omega} \cdot \partial_\phi u(\phi)$, then the solution of (1.2) can be expressed by an oscillatory integral

$$u(\phi) = \eta_T \int_0^T e^{\mathrm{i}\lambda z + \mathrm{i}B(\phi+z\tilde{\omega}) - \mathrm{i}B(\phi)} f(\phi + z\tilde{\omega})\mathrm{d}z, \tag{1.4}$$

where $\eta_T = \frac{\mathrm{i}}{e^{\mathrm{i}\lambda T}-1}$, and B fulfills $\omega \cdot \partial_\phi B(\phi) = b(\phi)$ and $\frac{1}{(2\pi)^\iota}\int_{\mathbb{T}^\iota} B(\phi)\mathrm{d}\phi = 0$. Using a technique from harmonic analysis to (1.4), Kuksin [11], [12] really removed the factor $e^{|b|_s}$ from (1.3). Nowadays, the result has been called Kuksin's Lemma. See [2], [8] and [13]. The derivation of formula (1.4) profits from the fact that (1.2) is scalar, or, the unknown function $u(\phi)$ is of dimension 1. In order to guarantee that (1.2) is scalar, one needs to assume that the normal frequency Ω_j is simple, that is, the multiplicity of Ω_j is 1. For the KdV equation subject to periodic boundary conditions, the normal frequency $\Omega_j = j^3$ is exactly simple. However, there are many PDEs, for example, the coupled KdV equation subject to periodic boundary conditions, such that Ω_j is not simple. In this case, (1.2) becomes an n-order differential equation where $n > 1$ is the multiplicity of Ω_j. Therefore, it is difficult to express $u(\phi)$ by an oscillatory integral like (1.4). In the present paper, we make an attempt to generalize Kuksin's Lemma to n-order differential equation using a modified method from [20]. More exactly, we have the following theorem:

Theorem 1.1 *Given real constants a, $b > 0$, $0 < \varepsilon \ll 1$ with $b - a > \varepsilon$, and $\lambda \in \mathbb{R}$ with $|\lambda| > 1$. Consider the first order differential equation*

$$-i\omega \frac{d}{d\phi}W + \lambda W + BW = R, \tag{1.5}$$

where W is the unknown function. Assume the following conditions are fulfilled:
(1) The known functions B and R depend on both the variables[1] $\phi \in \mathbb{T}_s$ and the parameter $\omega \in [a,b]$,

$$B(\phi,\omega) : \mathbb{T}_s \times [a,b] \to \mathbb{C}^{n\times n}, \quad R(\phi;\omega) : \mathbb{T}_s \times [a,b] \to \mathbb{C}^n$$

are analytic in $\phi \in \mathbb{T}_s$ and continuously differentiable with respect to ω in Whitney's sense.
(2) $B(\phi;\omega)$ is real for real argument, i.e. $B(\mathbb{T} \times [a,b]) \subset \mathbb{R}^{n\times n}$ and possesses the symmetry:

$$B^T(\phi;\omega) = B(\phi;\omega), \quad \forall\ (\phi;\omega) \in \mathbb{T} \times [a,b], \tag{1.6}$$

here T means the transpose of matrix.
(3) There exists $0 < \theta < 1$ such that[2]

$$|B|_{s,1} := \sum_{k\in\mathbb{Z}} |\widehat{B}(k)|\ |k| e^{|k|s} \le \varepsilon |\lambda|^\theta, \tag{1.7}$$

[1] We shorten $\mathbb{T}^1$, $\mathbb{T}_s^1$ as $\mathbb{T}$, $\mathbb{T}_s$ respectively.

[2] For a vector or a matrix of finite order, by $|\cdot|$ denote Euclidean norm and by $|\cdot|_s$ sup-norm in $\mathbb{T}_s$.

$$|B|^{\mathscr{L}}_{s,1} := \sum_{k\in\mathbb{Z}} |\partial_\omega \widehat{B}(k)|\ |k| e^{|k|s} \le \varepsilon^{\frac{1}{3}} |\lambda|^\theta, \tag{1.8}$$

where $\widehat{B}(k)$ is the k-th Fourier coefficient of B.
(4) $|R|_s = \sup_{\phi\in\mathbb{T}_s} |R(\phi)| < 1$.
Then there is a subset $\underline{\Omega} \subset [a,b]$ with[3]

$$Meas(\underline{\Omega}) > (b-a)(1 - s|\lambda|^{2\theta-1}), \tag{1.9}$$

and a function $W(\phi;\omega) : \mathbb{T}_s \times \underline{\Omega} \to \mathbb{C}^n$ solving (1.5) *approximately in the following sense:*

$$|W|_{s-\sigma} \le \frac{C}{s\sigma^{\frac{1}{2}}|\lambda|^\theta} |R|_s, \tag{1.10}$$

$$|-i\omega \frac{d}{d\phi} W + \lambda W + BW - R|_{s-2\sigma} \le \frac{C\varepsilon}{s\sigma} |R|_s^{\frac{9}{5}}, \tag{1.11}$$

$$|\partial_\omega W(\omega)|_{s-3\sigma} \le \frac{C}{s\sigma^{\frac{3}{2}}|\lambda|^\theta} \big(|\partial_\omega R|_s + \frac{C\varepsilon^{\frac{1}{3}}}{s\sigma^{\frac{3}{2}}} |R|_s\big), \tag{1.12}$$

$$|\partial_\omega(-i\omega \frac{d}{d\phi} W + \lambda W + BW - R)|_{s-4\sigma} \le \big(\frac{8}{\sigma}|\partial_\omega R|_s^2 + \frac{C\varepsilon^{\frac{1}{3}}}{s\sigma}|R|_s^{\frac{9}{5}} + \frac{C\varepsilon}{s\sigma^2}|R|_s^{\frac{9}{10}}\big(|\partial_\omega R|_s + \frac{C\varepsilon^{\frac{1}{3}}}{s\sigma^{\frac{3}{2}}}|R|_s\big)\big), \tag{1.13}$$

for $0 < \sigma = s/10$, where $C(a,b,n) = \frac{160(1+e)^2 n}{(b-a)}$.

Remark 1.1: *Note that in the m-th KAM iteration, $|R|_s = O(\varepsilon_m)$. By* (1.11)*, we have*

$$|-i\omega \frac{d}{d\phi} W + \lambda W + BW - R|_{s-2\sigma} = O(\varepsilon_m^\kappa) \tag{1.14}$$

with some $\kappa > 1$. Therefore, the estimate in (1.11) *is sufficient in KAM iteration, though W is an approximate solution instead of exact one.*

Using Theorem 1.1, we can establish a KAM theorem in an infinite dimensional setting which can be applied to the coupled KdV equation subject to periodic boundary conditions. We start by introducing some notations:
For $j \in \mathbb{Z}_+ := \{1,2,\cdots\}$, let $d_j \in \mathbb{Z}_+$ be positive integers. Given $a \ge 0$, $p \ge 0$, we consider the Hilbert space given by

$$\mathscr{Z}^{a,p} = \{z = (z_j : j \ge 1) : z_j = (z_j^1, \cdots, z_j^{d_j}) \in \mathbb{C}^{d_j},\ ||z||_{a,p} < \infty\},$$

where the norm $||\cdot||_{a,p}$ is defined by [4]

$$||z||^2_{a,p} = \sum_{j\ge1} |z_j|^2 j^{2p} e^{2aj} = \sum_{j\ge1} \big(\sum_{l=1}^{d_j} |z_j^l|^2\big) j^{2p} e^{2aj}.$$

Denote by $\mathscr{L}(\mathscr{Z}^{a,p}, \mathscr{Z}^{a,q})$ the set of all bounded linear operators from $\mathscr{Z}^{a,p}$ to $\mathscr{Z}^{a,q}$ and by $||\cdot||_{p,q}$ the operator norm. Then we introduce the phase space

$$\mathscr{P}^{a,p} = (\mathbb{C}/2\pi\mathbb{Z}) \times \mathbb{C} \times \mathscr{Z}^{a,p} \times \mathscr{Z}^{a,p}.$$

[3] Here 'Meas' means Lebesgue measure.
[4] Of course, we can regard $\mathscr{Z}^{a,p}$ as $\ell^{a,p}$.

We endow $\mathscr{P}^{a,p}$ with a symplectic structure

$$\mathrm{d}x \wedge \mathrm{d}y + \mathrm{i} \sum_{j\geq 1} \mathrm{d}z_j \wedge \mathrm{d}\bar{z}_j = \mathrm{d}x \wedge \mathrm{d}y + \mathrm{i} \sum_{j\geq 1} \sum_{l=1}^{d_j} \mathrm{d}z_j^l \wedge \mathrm{d}\bar{z}_j^l, \quad (x,y,z,\bar{z}) \in \mathscr{P}^{a,p}.$$

Let

$$\mathscr{T}_0 := \mathbb{T} \times \{y=0\} \times \{z=0\} \times \{\bar{z}=0\}$$

be a torus in $\mathscr{P}^{a,p}$. For positive numbers r,s, we introduce complex neighbourhoods of $\mathscr{T}_0$ in $\mathscr{P}^{a,p}$

$$D(s,r) = \{(x,y,z,\bar{z}) \in \mathscr{P}^{a,p} : |\mathrm{Im}x| < s, |y| < r^2, ||z||_{a,p} < r, ||\bar{z}||_{a,p} < r\},$$

$$D_{\Re}(s,r) = \{(x,y,z,\bar{z}) \in D(s,r) : x,y \in \mathbb{R}\}.$$

Note that x,y are real but $z,\bar{z}$ are still complex in $D_{\Re}(s,r)$. For $\tilde{r} > 0$ we define the weighted phase norms for $W = (x,y,z,\bar{z}) \in \mathscr{P}^{a,p}$:

$$||W||_{\tilde{r},p} = |x| + \frac{1}{\tilde{r}^2}|y| + \frac{1}{\tilde{r}}||z||_{a,p} + \frac{1}{\tilde{r}}||\bar{z}||_{a,p}.$$

Finally, we denote by Π a given compact set in $\mathbb{R}$ with positive Lebesgue measure, and $\xi \in \Pi$ the parameter. And for a map $W : D(s,r) \times \Pi \to \mathscr{P}^{a,p}$, let

$$||W||_{\tilde{r},p;D(s,r)\times\Pi} := \sup_{(w,\xi)\in D(s,r)\times\Pi} ||W(w,\xi)||_{\tilde{r},p},$$

$$||W||^{\mathscr{L}}_{\tilde{r},p;D(s,r)\times\Pi} := \sup_{(w,\xi)\in D(s,r)\times\Pi} ||\partial_\xi W(w,\xi)||_{\tilde{r},p},$$

where ∂_ξ is the derivative with respect to ξ in Whitney's sense. Recall that for $\tilde{s} > 0$ the complex $\tilde{s}$-vicinity of $\mathbb{T}$:

$$\mathbb{T}_{\tilde{s}} = \{x \in \mathbb{C}/2\pi\mathbb{Z} : |\mathrm{Im}x| < \tilde{s}\}.$$

Then for a function $f : \mathbb{T}_{\tilde{s}} \times \Pi \to \mathbb{C}^n$ we define

$$|f|_{\mathbb{T}_{\tilde{s}}\times\Pi} := \sup_{(x;\xi)\in\mathbb{T}_{\tilde{s}}\times\Pi} |f(x;\xi)|, \quad |f|^{\mathscr{L}}_{\mathbb{T}_{\tilde{s}}\times\Pi} := \sup_{(x;\xi)\in\mathbb{T}_{\tilde{s}}\times\Pi} |\partial_\xi f(x;\xi)|,$$

for functions that take values in other spaces (in what follows $\mathscr{Z}^{a,p}$ or $\mathscr{L}(\mathscr{Z}^{a,p},\mathscr{Z}^{a,q})$), the notations are defined similarly. For two vectors g,h of order ι with $1 \leq \iota \leq \infty$, we write $\langle g,h\rangle = \sum_{j=1}^{\iota} g_j h_j$.

Consider an infinite dimensional Hamiltonian in the parameter dependent normal form

$$N = \omega(\xi)y + \sum_{j\geq 1} \lambda_j \langle z_j, \bar{z}_j\rangle + \sum_{j\geq 1} \langle B_{jj}(x;\xi) z_j, \bar{z}_j\rangle, \tag{1.15}$$

here $\xi \in \Pi$ is the parameter, λ_j is real and independent of ξ while $B_{jj}(x;\xi)$ may depend on x and ξ, and $B_{jj}(x;\xi)$ is a $d_j \times d_j$ self-adjoint matrix for $(x;\xi) \in \mathbb{T} \times \Pi$. For each $\xi \in \Pi$, the Hamiltonian equations of motion for N, i.e.

$$\frac{dx}{dt} = \omega, \quad \frac{dy}{dt} = 0, \quad \frac{dz_j}{dt} = \mathrm{i}(\lambda_j E_{d_j} + B_{jj})z_j, \quad \frac{d\bar{z}_j}{dt} = -\mathrm{i}(\lambda_j E_{d_j} + B_{jj})\bar{z}_j \tag{1.16}$$

admit a special solution $x = \omega t$, $y = 0$, $z = 0$, $\bar{z} = 0$ corresponding to an invariant torus in $\mathscr{P}^{a,p}$, where E_{d_j} is the unit matrix of order d_j.

Now we study the perturbed Hamiltonian

$$H = N + P = \omega(\xi)y + \sum_{j\geq 1} \lambda_j \langle z_j, \bar{z}_j\rangle + \sum_{j\geq 1} \langle B_{jj}(x;\xi) z_j, \bar{z}_j\rangle + P(x,y,z,\bar{z};\xi). \tag{1.17}$$

Our goal is that, for most values $\xi \in \Pi$ (in Lebesgue measure sense), the Hamiltonian $H = N + P$ still admits an invariant torus provided $||X_P||$ is sufficiently small. To this end we need the following assumptions:

Assumption A: *(Frequency Asymptotics). There exist $\bar{d} \in \mathbb{Z}_+$ and $d > 1$ such that $d_j \leq \bar{d}$ for all j, and the behavior of λ_j's is as follows:*

$$|\lambda_i - \lambda_j| \geq C|i-j|(i^{d-1} + j^{d-1}), \tag{1.18}$$

for all $i \neq j \geq 0$ with some constant $C > 0$. Here $\lambda_0 = 0$.

Assumption B: $B(x;\xi) = diag(B_{jj}(x;\xi) : j \geq 1)$ *satisfies the following conditions:*
(1) Reality condition.

$$B_{jj}^T(x;\xi) = B_{jj}(x;\xi) \in \mathbb{R}^{d_j \times d_j}, \quad (x;\xi) \in \mathbb{T} \times \Pi. \tag{1.19}$$

(2) Finiteness of Fourier modes (we will give K_0 in the following).

$$B_{jj}(x;\xi) = \sum_{|k|\leq K_0} \widehat{B}_{jj}(k;\xi) e^{ikx}. \tag{1.20}$$

(3) Boundedness.

$$||B(x,\xi)||_{p,q;\mathbb{T}_s\times\Pi} \leq \varepsilon, \quad ||B(x,\xi)||^{\mathscr{L}}_{p,q;\mathbb{T}_s\times\Pi} \leq \varepsilon^{\frac{1}{3}}. \tag{1.21}$$

Assumption C *(Non-degeneracy).*

$$0 < C_1 \leq |\frac{d\omega(\xi)}{d\xi}| \leq C_2, \quad \forall\ \xi \in \Pi. \tag{1.22}$$

This assumption enables us to regard ω instead of ξ as a parameter.

Assumption D: *(Regularity). Let s, r be given positive constants. Assume the perturbation term $P(x,y,z,\bar{z};\xi)$ which is defined on the domain $D(s,r)\times\Pi$ is analytic in the space coordinates and continuously differentiable in $\xi \in \Pi$, as well as, for each $\xi \in \Pi$ its Hamiltonian vector field $X_P := (\partial_y P, -\partial_x P,\ i\partial_z P, -i\partial_{\bar{z}} P)$ defines a analytic map*

$$X_P : D(s,r) \subset \mathscr{P}^{a,p} \to \mathscr{P}^{a,q},$$

where $q \geq 0$ satisfies $p - q < d - 1$. Moreover, X_P is continuously differentiable in $\xi \in \Pi$. Without loss of generality, we may also assume that $\delta \geq 0$ is chosen so that

$$p - q \leq \delta < d - 1. \tag{1.23}$$

Assumption E: *(Reality). For any $(x,y,z,\bar{z};\xi) \in D_{\mathfrak{R}}(s,r)\times\Pi$, the perturbation P is real, that is,*

$$\overline{P(x,y,z,\bar{z};\xi)} = P(x,y,z,\bar{z};\xi), \ \ (x,y,z,\bar{z};\xi) \in D_{\mathfrak{R}}(s,r)\times\Pi, \tag{1.24}$$

where the bar means complex conjugate.

Theorem 1.2 (KAM Theory) *Suppose $H = N + P$ satisfies **Assumptions A-E** and smallness assumption:*

$$||X_P||_{r,q;D(s,r)\times\Pi} < \varepsilon, \quad ||X_P||^{\mathscr{L}}_{r,q;D(s,r)\times\Pi} < \varepsilon^{1/3}. \tag{1.25}$$

Then there is a sufficiently small $\varepsilon_* > 0$, such that for $0 < \varepsilon < \varepsilon_*$, there is a subset $\Pi_\varepsilon \subset \Pi$ with

$$Meas\ \Pi_\varepsilon \geq (Meas\ \Pi)(1 - O(\varepsilon)),$$

and there are a family of torus embeddings $\Phi : \mathbb{T} \times \Pi_\varepsilon \to \mathscr{P}^{a,p}$ and a map $\omega_* : \Pi_\varepsilon \to \mathbb{R}$, where $\Phi(\cdot,\xi)$ and $\omega_*(\xi)$ is continuously differentiable in $\xi \in \Pi_\varepsilon$, such that for each $\xi \in \Pi_\varepsilon$ the map Φ restricted to $\mathbb{T} \times \{\xi\}$ is analytic embedding of a rotational torus with frequencies $\omega_*(\xi)$ for the Hamiltonian H at ξ. In other words,

$$t \mapsto \Phi(x + t\omega), \quad t \in \mathbb{R}$$

is a real analytic, periodic solution for the Hamiltonian H for every $x \in \mathbb{T}$ and $\xi \in \Pi_\varepsilon$.
Each embedding is real analytic on $\mathbb{T} \times \{\xi\}$, and the following estimates

$$||\Phi(x;\xi) - \Phi_0(x;\xi)||_{r,p} \leq C\varepsilon, \quad ||\partial_\xi(\Phi(x;\xi) - \Phi_0(x;\xi))||_{r,p} \leq C\varepsilon^{1/3},$$

$$|\omega_*(\xi) - \omega(\xi)| \leq C\varepsilon, \quad |\partial_\xi(\omega_*(\xi) - \omega(\xi))| \leq C\varepsilon^{1/3}$$

hold true uniformly for $x \in \mathbb{T}$ and $\xi \in \Pi_\varepsilon$, where Φ_0 is the trivial embedding $\mathbb{T} \times \Pi \to \mathbb{T} \times \{y = 0\} \times \{z = 0\} \times \{\bar{z} = 0\}$, and $C > 0$ is a constant depending on n, d, p, q.

Consider the coupled KdV equation

$$\begin{cases} u_t = -u_{xxx} + 6uu_x + \varepsilon \dfrac{\partial f(x,u,v)}{\partial u}, \\ v_t = -v_{xxx} + 6vv_x + \varepsilon \dfrac{\partial f(x,u,v)}{\partial v}, \end{cases} \tag{1.26}$$

where $x \in \mathbb{T}$ and $\varepsilon > 0$ is sufficiently small. Denote the Hilbert space $\hbar_p$ of all complex valued sequences $\mathbf{w} = (w_j)_{j\neq 0}$ satisfying

$$||\mathbf{w}||_p^2 = \sum_{j\neq 0} |j|^{2p}|w_j|^2 < \infty, \quad w_{-j} = \bar{w}_j, \quad j \geq 1.$$

Theorem 1.3 *Let Π be a compact subset of positive Lebesgue measure, and $Z \geq 1$. Assume that the Hamiltonian $K = \int_{\mathbb{T}} f(x,u,v)dx$ is real analytic in a complex neighbourhood V of the origin in $\hbar_{Z+\frac{1}{2}} \times \hbar_{Z+\frac{1}{2}}$ and satisfies the regularity condition*

$$X_K : V \to \hbar_{Z+\frac{1}{2}} \times \hbar_{Z+\frac{1}{2}}, \quad ||X_K||_{Z+\frac{1}{2};V} := \sup_{(\boldsymbol{u},\boldsymbol{v})\in V} ||X_K||_{Z+\frac{1}{2}} \leq 1.$$

Then, there exists an $\varepsilon^* > 0$ such that for $|\varepsilon| < \varepsilon^*$ the following holds. There exist
(1) a nonempty Cantor set $\Pi_\varepsilon \subset \Pi$ with $Meas(\Pi - \Pi_\varepsilon) \to 0$ as $\varepsilon \to 0$,
(2) a family of real analytic torus embeddings

$$\Xi : \mathbb{T} \times \Pi_\varepsilon \to V \cap (\hbar_{Z+\frac{1}{2}} \times \hbar_{Z+\frac{1}{2}}),$$

(3) a continuously differentiable map $\chi : \Pi_\varepsilon \to \mathbb{R}$,
such that for each $(\theta,\xi) \in \mathbb{T} \times \Pi_\varepsilon$, the curve $(u(t),v(t)) = \Xi(\theta + \chi(\xi)t, \xi)$ is a periodic solution of (1.26) *winding around the invariant torus $\Xi(\mathbb{T} \times \{\xi\})$.*

Remark 1.2: *For $\varepsilon = 0$, the coupled KdV equation* (1.26) *is integrable. There are many approaches to get periodic or quasi-periodic solutions using the theory for integrable systems. However, for $\varepsilon \neq 0$ and a general $f(x,u,v)$, the equation is not integrable.*

2 Proof of Theorem 1.1

For $K = \frac{10}{s}|\ln(|R|_s)|$, i.e. $e^{-K\sigma} = e^{-Ks/10} = |R|_s$, we introduce a truncation operator Γ_K as follows:

$$(\Gamma_K f)(\phi) := \sum_{|k|\le K} \widehat{f}(k) e^{\mathrm{i}k\phi}, \quad \forall f : \mathbb{T}_s \to \mathbb{C}^n (\text{or } \mathbb{C}^{n\times n}),$$

where $\widehat{f}(k)$ is the k-th Fourier coefficient of f. Then we can write (1.5) as

$$-\mathrm{i}\omega \frac{\mathrm{d}}{\mathrm{d}\phi} W + \lambda W + \Gamma_K(BW) = \Gamma_K R + (R - \Gamma_K R) + (BW - \Gamma_K(BW)). \tag{2.1}$$

We consider the approximate equation of (2.1) with the truncation operator Γ_K:

$$-\mathrm{i}\omega \frac{\mathrm{d}}{\mathrm{d}\phi} W(\phi;\omega) + \lambda W(\phi;\omega) + (\Gamma_K(BW))(\phi;\omega) = (\Gamma_K R)(\phi;\omega), \quad (\phi;\omega) \in \mathbb{T}_s \times [a,b]. \tag{2.2}$$

Main end in this section is to show that there exists unique solution W of (2.2) such that the estimates (1.10)-(1.13) hold.

Suppose (2.2) has a solution W, thus it is easy to see that W fulfills

$$W(\phi;\omega) = (\Gamma_K W)(\phi;\omega) = \sum_{|k|\le K} \widehat{W}(k;\omega) e^{\mathrm{i}k\phi}, \quad \phi \in \mathbb{T}_s, \tag{2.3}$$

where $\widehat{W}(k)$ is the k-th Fourier coefficient of W. For $\phi \in \mathbb{T}_s$, write $\phi = \mathrm{i}t + \nu$, here $t \in \mathbb{R}$, $|t| < s$, $\nu \in \mathbb{T}$. Then

$$W(\phi;\omega) = \sum_{|k|\le K} \widehat{W}(k;\omega) e^{-kt} e^{\mathrm{i}k\nu}. \tag{2.4}$$

Inserting this formula into (2.2) and checking the coefficients of the mode $e^{\mathrm{i}k\nu}$, we get an algebraic equation involving parameter t:

$$(A(\omega) + \widehat{B}(t;\omega))\widehat{W}(t;\omega) = \widehat{R}(t;\omega), \tag{2.5}$$

where

$$A(\omega) = \mathrm{diag}((k\omega+\lambda)E_n : |k| \le K), \quad \widehat{B}(t;\omega) = (\widehat{B}(k-l;\omega)e^{-(k-l)t})_{|k|,|l|\le K},$$
$$\widehat{W}(t;\omega) = (\widehat{W}(k;\omega)e^{-kt})_{|k|\le K}, \quad \widehat{R}(t;\omega) = (\widehat{R}(k;\omega)e^{-kt})_{|k|\le K},$$

here E_n is the unit matrix of order n. Note that $A(\omega), \widehat{B}(t;\omega)$ are matrices of order $\widetilde{K} = (2K+1)n$, while $\widehat{W}(t;\omega), \widehat{R}(t;\omega)$ are vectors in $\mathbb{C}^{\widetilde{K}}$.

Case 1: $|\lambda| \le 2bK$. By the symmetry condition (1.6), we have $\big(\widehat{B}(k-l;\omega)\big)^* = \widehat{B}(l-k;\omega)$ (here $*$ means the transpose and conjugate of a matrix), thus it is easy to check that $\widehat{B}(t;\omega)$ satisfies:

$$\big(\widehat{B}(t;\omega)\big)^* = \widehat{B}(-t;\omega). \tag{2.6}$$

Then we introduce some notations:

$$U(t;\omega) := \frac{\widehat{B}(t;\omega) + \big(\widehat{B}(t;\omega)\big)^*}{2} = \frac{\widehat{B}(t;\omega) + \widehat{B}(-t;\omega)}{2}, \tag{2.7}$$

$$V(t;\omega) := \frac{\widehat{B}(t;\omega) - \big(\widehat{B}(t;\omega)\big)^*}{2} = \frac{\widehat{B}(t;\omega) - \widehat{B}(-t;\omega)}{2}, \tag{2.8}$$

$$A_1(t;\omega) := A(\omega) + U(t;\omega). \tag{2.9}$$

Now all of $U(t;\omega)$, $V(t;\omega)$ and $A_1(t;\omega)$ are self-adjoint for all $|t| < s$,

$$U^*(t;\omega) = U(t;\omega), \quad V^*(t;\omega) = V(t;\omega), \quad A_1^*(t;\omega) = A_1(t;\omega). \tag{2.10}$$

And

$$A(\omega) + \widehat{B}(t;\omega) = A_1(t;\omega) + V(t;\omega). \tag{2.11}$$

Without loss of generality, we may assume that $\lambda > 1$. Using Lemma 5.7, and noting (1.7), (1.8), we have

$$|\widehat{B}(t;\omega)| \le n^{\frac{1}{4}} \max_{|l|\le K} \sum_{|k|\le K} |\widehat{B}(k-l;\omega)| e^{-(k-l)t} \le n^{\frac{1}{4}} \sum_{k\in\mathbb{Z}} |\widehat{B}(k;\omega)| e^{|k|s} \le n^{\frac{1}{4}} |B|_{s,1} \le \varepsilon n^{\frac{1}{4}} \lambda^{\theta}, \tag{2.12}$$

$$|\partial_\omega \widehat{B}(t;\omega)| \le n^{\frac{1}{4}} \max_{|l|\le K} \sum_{|k|\le K} |\partial_\omega B(k-l;\omega)| e^{-(k-l)t} \le n^{\frac{1}{4}} \sum_{k\in\mathbb{Z}} |\partial_\omega B(k;\omega)| e^{|k|s} \le n^{\frac{1}{4}} |B|_{s,1}^{\mathscr{L}} \le \varepsilon^{\frac{1}{3}} n^{\frac{1}{4}} \lambda^{\theta}.$$

So by the definition of $U(t;\omega)$ in (2.7), we obtain that for $|t| < s$,

$$|U(t;\omega)| \le |\widehat{B}(t;\omega)| \le \varepsilon n^{\frac{1}{4}} \lambda^{\theta}, \quad |\partial_\omega U(t;\omega)| \le |\partial_\omega \widehat{B}(t;\omega)| \le \varepsilon^{\frac{1}{3}} n^{\frac{1}{4}} \lambda^{\theta}. \tag{2.13}$$

And by definition of $V(t;\omega)$ in (2.8), using Lemma 5.7, we get that for $|t| < s$,

$$\begin{aligned}
|V(t;\omega)| &\le \frac{n^{\frac{1}{4}}}{2} \max_{|l|\le K} \sum_{|k|\le K} |\widehat{B}(k-l;\omega)|\ |e^{-(k-l)t} - e^{(k-l)t}| \\
&\le n^{\frac{1}{4}} \max_{|l|\le K} \sum_{|k|\le K} |\widehat{B}(k-l;\omega)| (s|k-l| e^{|k-l|\,|s|}) \\
&\le s n^{\frac{1}{4}} \max_{|l|\le K} \sum_{|k|\le K} |\widehat{B}(k-l;\omega)|\ |k-l| e^{|k-l|\,|s|} \\
&\le s n^{\frac{1}{4}} \sum_{k\in\mathbb{Z}} |\widehat{B}(k;\omega)|\ |k| e^{|k|s} \le s n^{\frac{1}{4}} |B|_{s,1} \le \varepsilon s n^{\frac{1}{4}} \lambda^{\theta}.
\end{aligned} \tag{2.14}$$

Similarly, we have

$$|\partial_\omega V(t;\omega)| \le s n^{\frac{1}{4}} |B|_{s,1}^{\mathscr{L}} \le \varepsilon^{\frac{1}{3}} s n^{\frac{1}{4}} \lambda^{\theta}. \tag{2.15}$$

In order to estimate $|W(\omega)|$, we need to estimate $|(A(\omega) + \widehat{B}(\omega))^{-1}| = |(A_1(\omega) + V(\omega))^{-1}|$ (we shorten $A_1(t,\omega) = A_1(\omega)$, and other notations $\widehat{B}(\omega)$, $U(\omega)$, $V(\omega)$ are defined similarly). To this end, we firstly consider the estimate for $|A_1(\omega)^{-1}|$, and then we are able to get the estimate for $|(A_1(\omega) + V(\omega))^{-1}|$ by virtue of $|A_1(\omega)^{-1}|$. By (2.9), we consider

$$A_1(\omega) = \mathrm{diag}((k\omega + \lambda)E_n : |k| \le K) + U(\omega), \tag{2.16}$$

where $A_1(\omega)$ is continuously differentiable in Whitney's sense in $\omega \in [a,b]$. Using Lemma 5.1, there are continuously differentiable functions $\mu_1(\omega), \cdots, \mu_{\widetilde{K}}(\omega)$ which represent the eigenvalues of $A_1(\omega)$ for $\omega \in [a,b]$. Moreover there exists a unitary matrix G of order $\widetilde{K}$, which depends on ω, such that for every $\omega \in [a,b]$, the following equality holds true

$$A_1(\omega) = \mathrm{diag}((k\omega + \lambda)E_n : |k| \le K) + U(\omega) = G(\omega)\mathrm{diag}(\mu_1(\omega), \cdots, \mu_{\widetilde{K}}(\omega))G^*(\omega). \tag{2.17}$$

Then by the perturbation theory of self-adjoint matrix, we have

$$|\mu_j - (k\omega + \lambda)| \le |U(\omega)| \le \varepsilon n^{\frac{1}{4}} \lambda^{\theta}, \quad j = 1, \cdots, \widetilde{K}, \quad |k| \le K. \tag{2.18}$$

Introduce

$$\mathscr{K}_1 = \{k \in \mathbb{Z} : |k| \le K,\ |k\omega + \lambda| \le \lambda^\theta\}, \quad \mathscr{K}_2 = \{k \in \mathbb{Z} : |k| \le K,\ |k\omega + \lambda| > \lambda^\theta\},$$

and by $K_1^\sharp, K_2^\sharp$ we denote for their cardinality, respectively. Thus by (2.18), we have

$$|\mu| > \frac{1}{2}\lambda^\theta, \tag{2.19}$$

for $k \in \mathscr{K}_2$. And for $k \in \mathscr{K}_1$, we have

$$\omega^{-1}(-\lambda^\theta - \lambda) \le k \le \omega^{-1}(\lambda^\theta - \lambda). \tag{2.20}$$

So

$$K_1^\sharp \le 2\omega^{-1}\lambda^\theta. \tag{2.21}$$

Following the above discussion and without loss of generality, we may assume that $|\mu_1| \le \cdots \le |\mu_{K_1^\sharp}| \le \frac{1}{2}\lambda^\theta$. Arbitrarily take $\mu \in \{\mu_1, \cdots, \mu_{K_1^\sharp}\}$, and let ψ be the normalized eigenvector corresponding to μ. Using Lemma 5.2 and noting (2.17), we get

$$\begin{aligned}\partial_\omega \mu &= (\partial_\omega(A_1(\omega))\psi, \psi)\\ &= (\partial_\omega(\mathrm{diag}((k\omega + \lambda)E_n : |k| \le K) + U(\omega))\psi, \psi)\\ &= k + (\partial_\omega U(\omega)\psi, \psi).\end{aligned}$$

Thus by (2.13) and (2.20), we have

$$\frac{1}{2}\omega^{-1}\lambda \le |\partial_\omega \mu| \le 2\omega^{-1}\lambda, \tag{2.22}$$

where in the last inequality, we use $\lambda > 1$, $0 < \theta < 1$ and $0 < \varepsilon \ll 1$.
Let

$$\mathscr{R}_l = \{\omega \in [a,b] : |\mu_l(\omega)| \le \frac{1}{4}(b-a)s\lambda^\theta\}, \quad l = 1, \cdots, K_1^\sharp.$$

By (2.22), $\mathrm{Meas}\mathscr{R}_l \le \frac{(b-a)s\lambda^\theta}{2\omega^{-1}\lambda}$. Thus by (2.21),

$$\mathrm{Meas}(\bigcup_{l=1}^{K_1^\sharp}\mathscr{R}_l) \le \frac{(b-a)s\lambda^\theta}{2\omega^{-1}\lambda}K_1^\sharp \le \frac{(b-a)s\lambda^\theta}{2\omega^{-1}\lambda}(2\omega^{-1}\lambda^\theta) \le (b-a)s\lambda^{2\theta-1}.$$

Let $\underline{\Omega} = [a,b] \backslash (\bigcup_{l=1}^{K_1^\sharp}\mathscr{R}_l)$, therefore we have

$$\mathrm{Meas}\ \underline{\Omega} \ge (b-a)(1 - s\lambda^{2\theta-1}). \tag{2.23}$$

And for any $\omega \in \underline{\Omega}$,

$$|\mu_l(\omega)| > \frac{1}{4}(b-a)s\lambda^\theta, \quad l = 1, \cdots, K_1^\sharp.$$

Furthermore by (2.17), for any $\omega \in \underline{\Omega}$, we have

$$|A_1(\omega)^{-1}| \le |G(\omega)|\ |G(\omega)^*| \max_{1 \le l \le K_1^\sharp} |\mu_l(\omega)^{-1}| \le \frac{4}{(b-a)s\lambda^\theta}. \tag{2.24}$$

So by (2.14), (2.24) and (1), for any $\omega \in \underline{\Omega}$,

$$|A_1(\omega)^{-1}V(\omega)| \leq |A_1(\omega)^{-1}|\,|V(\omega)| \leq \frac{4\varepsilon s n^{\frac{1}{4}}\lambda^{\theta}}{(b-a)s\lambda^{\theta}} = \frac{4\varepsilon n^{\frac{1}{4}}}{b-a} \leq \frac{1}{4}, \tag{2.25}$$

where in the last inequality we use $b-a > \varepsilon$ and $0 < \varepsilon \ll 1$. Now using Neumann series, we see that there exists the inverse of $A_1(\omega)+V(\omega)$ for $\omega \in \underline{\Omega}$, and

$$|(A_1(\omega)+V(\omega))^{-1}| \leq \sum_{j=0}^{\infty} |A_1(\omega)^{-1}V(\omega)|^j\,|A_1(\omega)^{-1}| \leq 2|A_1(\omega)^{-1}|. \tag{2.26}$$

Thus for $\omega \in \underline{\Omega}$ and $|t| < s$, we have

$$|(A(\omega)+\widehat{B}(t;\omega))^{-1}| = |(A_1(t,\omega)+V(t;\omega))^{-1}| \leq \frac{8}{(b-a)s\lambda^{\theta}}. \tag{2.27}$$

Recall that

$$(A(\omega)+\widehat{B}(t;\omega))\widehat{W}(t;\omega) = \widehat{R}(t;\omega). \tag{2.28}$$

Then for $\omega \in \underline{\Omega}$ and $|t| < s$,

$$|\widehat{W}(t;\omega)| \leq |(A(\omega)+\widehat{B}(t;\omega))^{-1}|\;|\widehat{R}(t;\omega)| \leq \frac{8}{(b-a)s\lambda^{\theta}}|\widehat{R}(t;\omega)|. \tag{2.29}$$

Set $\Upsilon = \{1,-1\}$. Arbitrarily take $\kappa \in \{1,-1\}$. Let $t = s'\kappa$ with $s' = s - \sigma/10$. Then by (2.28) we have

$$|\widehat{W}(s'\kappa;\omega)|^2 \leq \frac{64}{(b-a)^2 s^2\lambda^{2\theta}}|\widehat{R}(s'\kappa;\omega)|^2. \tag{2.30}$$

Recall that $\widehat{W}(t;\omega) = (\widehat{W}(k;\omega)e^{-kt})_{|k|\leq K}$ and $\widehat{R}(t;\omega) = (\widehat{R}(k;\omega)e^{-kt})_{|k|\leq K}$. Thus we have

$$\sum_{|k|\leq K} |\widehat{W}(k;\omega)|^2 e^{2|k|s'} \leq \sum_{\kappa\in\{1,-1\}}\sum_{|k|\leq K} |\widehat{W}(k;\omega)e^{-ks'\kappa}|^2 = \sum_{\kappa\in\{1,-1\}} |\widehat{W}(s'\kappa;\omega)|, \tag{2.31}$$

$$|\widehat{R}(s'\kappa;\omega)|^2 \leq \sum_{|k|\leq K} |\widehat{R}(k)|^2 e^{2|k|s'}. \tag{2.32}$$

Consequently, by combining (2.30), (2.31) and (2.32), we have

$$\begin{aligned}\sum_{|k|\leq K} |\widehat{W}(k;\omega)|^2 e^{2|k|s'} &\leq \frac{64}{(b-a)^2 s^2\lambda^{2\theta}} \sum_{\kappa\in\{1,-1\}} |\widehat{R}(s'\kappa;\omega)|^2 \\ &\leq \frac{128}{(b-a)^2 s^2\lambda^{2\theta}} \sum_{|k|\leq K} |\widehat{R}(k;\omega)|^2 e^{2|k|s'}.\end{aligned} \tag{2.33}$$

Using $|\widehat{R}(k)| \leq |R|_s e^{-|k|s}$ and $\sum_{|k|\leq K} e^{-2|k|\sigma/10} \leq \frac{10+10e}{\sigma}$, we have

$$(2.33) \leq \frac{128}{(b-a)^2 s^2\lambda^{2\theta}} \sum_{|k|\leq K} e^{-2|k|(s-s')}|R|_s^2 \leq \frac{128(10+10e)}{(b-a)^2\sigma s^2\lambda^{2\theta}}|R|_s^2. \tag{2.34}$$

Therefore by Cauchy's inequality and (2.32), (2.33), we finally arrive at

$$\begin{aligned}
|W(\omega)|_{s-\sigma} &\leq \sum_{|k|\leq K} |\widehat{W}(k;\omega)| e^{|k|(s-\sigma)} \\
&= \sum_{|k|\leq K} |\widehat{W}(k;\omega)| e^{|k|(s-\sigma/10)} e^{|k|(-9\sigma/10)} \\
&\leq (\sum_{|k|\leq K} e^{-18|k|\sigma/10})^{1/2} (\sum_{|k|\leq K} |\widehat{W}(k;\omega)|^2 e^{2|k|s'})^{1/2} \\
&\leq \frac{80\sqrt{2}(1+e)}{3(b-a)s\sigma^{\frac{1}{2}}\lambda^{\theta}} |R(\omega)|_s, \qquad (2.35)
\end{aligned}$$

where in (2.35) we use $\sum_{|k|\leq K} e^{-18|k|\sigma/10} \leq \frac{10+10e}{9\sigma}$.

Case 2: $|\lambda| > 2bK$. Without loss of generality, we may still assume $\lambda > 1$. In this case, for $|k| \leq K$, we have $|k\omega + \lambda| > \frac{\lambda}{2}$, then $|A^{-1}(\omega)| < \frac{2}{\lambda}$. Using (2.12) and the Neumann series, we get that for $|t| \leq s$,

$$|(A(\omega)+\widehat{B}(t;\omega))^{-1}| \leq \sum_{j=0}^{\infty} |A^{-1}(\omega)\widehat{B}(t;\omega)|^j |A^{-1}(\omega)| \leq \frac{2}{\lambda} \sum_{j=0}^{\infty} (\frac{2\varepsilon n^{\frac{1}{4}}\lambda^{\theta}}{\lambda})^j \leq \frac{4}{\lambda}. \qquad (2.36)$$

Thus, similarly to the proof for (2.35), we have

$$|W(\omega)|_{s-\sigma} \leq \frac{80\sqrt{2}(1+e)}{3\lambda\sigma^{\frac{1}{2}}} |R(\omega)|_s. \qquad (2.37)$$

Now by (2.35) and (2.37), we get (1.10). And the proof of (1.11) is as follows:

$$\begin{aligned}
|(1-\Gamma_K)(BW)|_{s-2\sigma} &\leq \sum_{|k|>K} \Big| \sum_{|l|\leq K} \widehat{B}(k-l)\widehat{W}(l) \Big| e^{|k|(s-2\sigma)} \\
&\leq e^{-9K\sigma/10} \sum_{|k|>K} \Big| \sum_{|l|\leq K} \widehat{B}(k-l)\widehat{W}(l) \Big| e^{|k|(s-11\sigma/10)} \\
&\leq e^{-9K\sigma/10} \sum_{|k|>K} \sum_{|l|\leq K} |\widehat{B}(k-l)| e^{|k-l|(s-11\sigma/10)} |\widehat{W}(l)| e^{|l|(s-11\sigma/10)} \\
&\leq e^{-9K\sigma/10} \Big(\sum_{|k|>K} \max_{|l|\leq K} |\widehat{B}(k-l)| e^{|k-l|(s-11\sigma/10)} \Big) \Big(\sum_{|l|\leq K} |\widehat{W}(l)| e^{|l|(s-11\sigma/10)} \Big) \\
&\leq e^{-9K\sigma/10} (\sum_{k\in\mathbb{Z}} |\widehat{B}(k)| e^{|k|(s-11\sigma/10)}) (\sum_{|k|\leq K} |\widehat{W}(k)| e^{|k|(s-11\sigma/10)}) \\
&\leq e^{-9K\sigma/10} |B|_{s,1} (\sum_{|k|\leq K} e^{-2|k|\sigma/10})^{1/2} (\sum_{|k|\leq K} |\widehat{W}(k)|^2 e^{2|k|(s-\sigma)})^{1/2} \\
&\leq \frac{160\sqrt{5}(1+e)^{\frac{3}{2}}\varepsilon n^{\frac{1}{4}}}{3(b-a)s\sigma} e^{-9K\sigma/10} |R|_s \qquad (2.38) \\
&\leq \frac{160(1+e)^{\frac{3}{2}}\varepsilon n^{\frac{1}{4}}}{(b-a)s\sigma} |R|_s^{\frac{19}{10}} \leq \frac{160(1+e)^{\frac{3}{2}}\varepsilon n^{\frac{1}{4}}}{(b-a)s\sigma} |R|_s^{\frac{9}{5}}, \qquad (2.39)
\end{aligned}$$

where in (2.38) we use (2.35), $|B|_{s,1} \leq \varepsilon n^{\frac{1}{4}}\lambda^{\theta}$, $\sum_{|k|\leq K} e^{-2|k|\sigma/10} \leq \frac{10+10e}{\sigma}$, while in (2.39) we use

$e^{-9K\sigma/10} = |R|_s^{\frac{9}{10}}$ and $|R|_s < 1$. Using the fact $|\widehat{R}(k)| \le e^{-|k|s}|R|_s$ and $e^{-K\sigma} = |R|_s$, we have

$$\begin{aligned} |(1-\Gamma_K)R|_{s-\sigma} &\le \sum_{|k|>K} |\widehat{R}(k)| e^{|k|(s-\sigma)} \\ &\le \sum_{|k|>K} (|R|_s e^{-|k|s}) e^{|k|(s-\sigma)} \\ &\le (\sum_{|k|>K} e^{-|k|\sigma})|R|_s \\ &\le \frac{2(1+e)}{\sigma} e^{-K\sigma}|R|_s \\ &= \frac{2(1+e)}{\sigma}|R|_s^2, \end{aligned} \tag{2.40}$$

where in the last inequality, we use $\sum_{k\in\mathbb{Z}} e^{-k\sigma} \le \frac{2(1+e)}{\sigma}$.

Applying ∂_ω to both sides of (2.2), we have

$$-\mathrm{i}\omega \frac{\mathrm{d}}{\mathrm{d}\phi}(\partial_\omega W) + \lambda(\partial_\omega W) + \Gamma_K(B(\partial_\omega W)) = \Gamma_K(\partial_\omega R) - \Gamma_K((\partial_\omega B)W) + \mathrm{i}\frac{\mathrm{d}}{\mathrm{d}\phi}W := \widetilde{R}. \tag{2.41}$$

By the same proof as (2.35), we have

$$|\partial_\omega W|_{s-3\sigma} \le \frac{C'(a,b)}{s\sigma^{\frac{1}{2}}\lambda^\theta}|\widetilde{R}|_{s-2\sigma}, \tag{2.42}$$

where $C'(a,b) = \frac{80\sqrt{2}(1+e)}{3(b-a)}$. Note that $|\widehat{\partial_\omega R}(k)| \le e^{-|k|s}|\partial_\omega R|_s$, we have

$$\begin{aligned} |\Gamma_K(\partial_\omega R)|_{s-\sigma} &\le \sum_{|k|\le K} |\widehat{\partial_\omega R}(k)| e^{|k|(s-\sigma)} \\ &\le \sum_{|k|\le K} (|\partial_\omega R|_s e^{-|k|s}) e^{|k|(s-\sigma)} \\ &\le (\sum_{|k|\le K} e^{-|k|\sigma})|\partial_\omega R|_s \le \frac{2(1+e)}{\sigma}|\partial_\omega R|_s. \end{aligned} \tag{2.43}$$

Using (2.35) and $|\partial_\omega B|_s \le |\partial_\omega B|_{s,1} \le \varepsilon^{\frac{1}{3}}\lambda^\theta$, we have

$$\begin{aligned} |\Gamma_K((\partial_\omega B)W)|_{s-2\sigma} &\le \sum_{|k|\le K} |(\widehat{(\partial_\omega B)W})(k)| e^{|k|(s-2\sigma)} \\ &\le \sum_{|k|\le K} \big(|(\partial_\omega B)W|_{s-\sigma} e^{-|k|(s-\sigma)}\big) e^{|k|(s-2\sigma)} \\ &\le (\sum_{|k|\le K} e^{-|k|\sigma})|\partial_\omega B|_s |W|_{s-\sigma} \\ &\le \frac{2(1+e)}{\sigma}\varepsilon^{\frac{1}{3}}\lambda^\theta \frac{80\sqrt{2}(1+e)}{3(b-a)s\sigma^{\frac{1}{2}}\lambda^\theta}|R|_s \\ &\le \frac{160\sqrt{2}(1+e)^2\varepsilon^{\frac{1}{3}}}{3(b-a)s\sigma^{\frac{3}{2}}}|R|_s, \end{aligned} \tag{2.44}$$

where in (2.43) and (2.44), we use $\sum_{|k|\le K} e^{-|k|\sigma} \le \frac{2(1+e)}{\sigma}$. Moreover, by (2.35) and $\sum_{|k|\le K}|k|e^{-|k|\sigma} \le \frac{1+e}{e\sigma^2}$,

$$\begin{aligned}
|\mathrm{i}\frac{\mathrm{d}}{\mathrm{d}\phi}W|_{s-2\sigma} &= |\mathrm{i}\sum_{|k|\le K}\widehat{W}(k)(\mathrm{i}k)e^{\mathrm{i}k\phi}|_{s-2\sigma} \\
&\le \sum_{|k|\le K}|\widehat{W}(k)||k|e^{|k|(s-2\sigma)} \\
&\le \sum_{|k|\le K}|W|_{s-\sigma}e^{-|k|(s-\sigma)}|k|e^{|k|(s-2\sigma)} \\
&\le (\sum_{|k|\le K}|k|e^{-|k|\sigma})|W|_{s-\sigma} \\
&\le \frac{1+e}{e\sigma^2}\cdot\frac{80\sqrt{2}(1+e)}{3(b-a)s\sigma^{\frac{1}{2}}\lambda^{\theta}}|R|_s \\
&\le \frac{80\sqrt{2}(1+e)^2}{3e(b-a)s\sigma^{\frac{5}{2}}\lambda^{\theta}}|R|_s. \qquad (2.45)
\end{aligned}$$

Thus, by (2.43), (2.44), (2.45), we have

$$|\widetilde{R}|_{s-2\sigma} \le \big(\frac{2(1+e)}{\sigma}|\partial_\omega R|_s + \frac{C(a,b,n)\varepsilon^{\frac{1}{3}}}{s\sigma^{\frac{5}{2}}}|R|_s\big). \qquad (2.46)$$

And by (2.42), we have

$$\begin{aligned}
|\partial_\omega W|_{s-3\sigma} &\le \frac{C'(a,b)}{s\sigma^{\frac{1}{2}}\lambda^{\theta}}|\widetilde{R}|_{s-2\sigma} \\
&\le \frac{C'(a,b)}{s\sigma^{\frac{1}{2}}\lambda^{\theta}}\big(\frac{2(1+e)}{\sigma}|\partial_\omega R|_s + \frac{C(a,b,n)\varepsilon^{\frac{1}{3}}}{s\sigma^{\frac{5}{2}}}|R|_s\big). \qquad (2.47)
\end{aligned}$$

And the proof for (1.13) is as follows:

$$\begin{aligned}
|(1-\Gamma_K)((\partial_\omega B)W)|_{s-2\sigma} &\le \sum_{|k|>K}\big|\sum_{|l|\le K}(\partial_\omega\widehat{B}(k-l))\widehat{W}(l)\big|e^{|k|(s-2\sigma)} \\
&\le e^{-9K\sigma/10}\sum_{|k|>K}\big|\sum_{|l|\le K}(\partial_\omega\widehat{B}(k-l))\widehat{W}(l)\big|e^{|k|(s-11\sigma/10)} \\
&\le e^{-9K\sigma/10}\sum_{|k|>K}\sum_{|l|\le K}|\partial_\omega\widehat{B}(k-l)|e^{|k-l|(s-11\sigma/10)}|\widehat{W}(l)|e^{|l|(s-11\sigma/10)} \\
&\le e^{-9K\sigma/10}\big(\sum_{|k|>K}\max_{|l|\le K}|\partial_\omega\widehat{B}(k-l)|e^{|k-l|(s-11\sigma/10)}\big)\sum_{|l|\le K}|\widehat{W}(l)|e^{|l|(s-11\sigma/10)} \\
&\le e^{-9K\sigma/10}(\sum_{k\in\mathbb{Z}}|\partial_\omega\widehat{B}(k)|e^{|k|(s-11\sigma/10)})(\sum_{|k|\le K}|\widehat{W}(k)|e^{|k|(s-11\sigma/10)}) \\
&\le e^{-9K\sigma/10}|B|^{\mathscr{L}}_{s,1}(\sum_{|k|\le K}e^{-2|k|\sigma/10})^{1/2}(\sum_{|k|\le K}|\widehat{W}(k)|^2e^{2|k|(s-\sigma)})^{1/2} \\
&\le e^{-9K\sigma/10}\cdot\varepsilon^{\frac{1}{3}}n^{\frac{1}{4}}\lambda^{\theta}\cdot\frac{\sqrt{2(1+e)}}{\sigma^{\frac{1}{2}}}\cdot\frac{80\sqrt{2}(1+e)|R|_s}{3(b-a)s\sigma^{\frac{1}{2}}\lambda^{\theta}} \qquad (2.48) \\
&\le \frac{160(1+e)^{\frac{3}{2}}\varepsilon^{\frac{1}{3}}n^{\frac{1}{4}}}{3(b-a)s\sigma}e^{-9K\sigma/10}|R|_s \le \frac{C(a,b,n)\varepsilon^{\frac{1}{3}}}{s\sigma}|R|_s^{\frac{9}{5}} \qquad (2.49)
\end{aligned}$$

where in (2.48) we use (2.35), $|B|_{s,1}^{\mathscr{L}} \leq \varepsilon^{\frac{1}{3}} n^{\frac{1}{4}} \lambda^{\theta}$, $\sum_{|k|\leq K} e^{-2|k|\sigma/10} \leq \frac{10+10e}{\sigma}$, while in (2.49) we use $e^{-9K\sigma/10} = |R|_s^{\frac{9}{10}}$ and $|R|_s < 1$. Similarly, by (2.47), we have

$$
\begin{aligned}
|(1-\Gamma_K)(B(\partial_\omega W))|_{s-4\sigma} &\leq \sum_{|k|>K} \Big| \sum_{|l|\leq K} \widehat{B}(k-l)(\partial_\omega \widehat{W}(l)) \Big| e^{|k|(s-4\sigma)} \\
&\leq e^{-9K\sigma/10} \Big(\sum_{|k|>K} \max_{|l|\leq K} |\widehat{B}(k-l)| e^{|k-l|s} \Big) \sum_{|l|\leq K} |\partial_\omega \widehat{W}(l)| e^{|l|(s-31\sigma/10)} \\
&\leq e^{-9K\sigma/10} \Big(\sum_{k\in\mathbb{Z}} |\widehat{B}(k)| e^{|k|s} \Big) \Big(\sum_{|k|\leq K} |\partial_\omega \widehat{W}(k)| e^{|k|(s-31\sigma/10)} \Big) \\
&\leq e^{-9K\sigma/10} |B|_{s,1} \Big(\sum_{|k|\leq K} e^{-2|k|\sigma/10} \Big)^{1/2} \Big(\sum_{|k|\leq K} |\partial_\omega \widehat{W}(k)|^2 e^{2|k|(s-3\sigma)} \Big)^{1/2} \\
&\leq |R|^{\frac{9}{10}} \cdot \varepsilon n^{\frac{1}{4}} \lambda^{\theta} \cdot \frac{\sqrt{10}(1+e)}{\sigma^{\frac{1}{2}}} \cdot \frac{C'(a,b)}{s\sigma^{\frac{1}{2}} \lambda^{\theta}} \Big[\frac{2(1+e)}{\sigma} |\partial_\omega R|_s + \frac{C(a,b,n)\varepsilon^{\frac{1}{3}}}{s\sigma^{\frac{5}{2}}} |R|_s \Big] \\
&\leq \frac{C(a,b,n)\varepsilon}{s\sigma} |R|^{\frac{9}{10}} \Big(\frac{2(1+e)}{\sigma} |\partial_\omega R|_s + \frac{C(a,b,n)\varepsilon^{\frac{1}{3}}}{s\sigma^{\frac{5}{2}}} |R|_s \Big).
\end{aligned}
$$

Using the fact $|\widehat{\partial_\omega R}(k)| \leq e^{-|k|s} |\partial_\omega R|_s$ and $e^{-K\sigma} = |R|_s$, we have

$$
\begin{aligned}
|(1-\Gamma_K)(\partial_\omega R)|_{s-\sigma} &\leq \sum_{|k|>K} |\widehat{\partial_\omega R}(k)| e^{|k|(s-\sigma)} \\
&\leq \Big(\sum_{|k|>K} e^{-|k|\sigma} \Big) |\partial_\omega R|_s \\
&\leq \frac{2(1+e)}{\sigma} e^{-K\sigma} |R|_s \\
&= \frac{2(1+e)}{\sigma} |\partial_\omega R|_s^2. \qquad (2.50)
\end{aligned}
$$

Remark 2.1: *By* (1.9) *and the above proof, in the m-th KAM iteration we obtain*

$$
Meas(\underline{\Omega}) \geq (b-a)(1-s|\lambda|^{2\theta-1}) \geq (b-a)(1-s_m(2b)^{2\theta-1}K_m^{2\theta-1}) = (b-a)(1-O(\frac{1}{m^2})). \quad (2.51)
$$

Therefor the estimate in (1.9) *is sufficient in KAM iteration.*

3 The Proof of Theorem 1.2

1. Iterative constants. Theorem 1.2 is proved by a KAM iteration which involves an infinite sequence of coordinate changes. In order to make our iteration procedure run, we need the following iterative constants:

1. $\varepsilon_0 = \varepsilon, \varepsilon_l = \varepsilon^{(4/3)^l}$, $l = 1, 2, \cdots$;
2. $\eta_l = \varepsilon_l^{1/3}$, $l = 0, 1, \cdots$;
3. $K_0 = 1$, $K_l = \frac{10|\ln\varepsilon|}{s_0}(4/3)^{\alpha l}$ with $\alpha > \frac{1}{2(1-\frac{\delta}{d-1})}$, $l = 1, 2, \cdots$;
4. Let $s_0 > 0$ be given (without loss of generality, we can let $s_0 = 1$). And let $s_l = (4/3)^{-(\alpha-1)l} s_0$, $\sigma_l = s_l/10$, $l = 1, 2, \cdots$. Thus $e^{-K_l\sigma_l} = \varepsilon_l$ for $l = 0, 1, 2, \cdots$;
5. $r_0 > 0$ is given, $r_l = \eta_l r_0$, $l = 1, 2, \cdots$;
6. $\mathbb{T}_{s_l} = \{x \in \mathbb{C}/2\pi\mathbb{Z} : |\mathrm{Im}x| < s_l\}$;

7. $D(s_l,r_l)=\{(x,y,z,\bar z)\in\mathscr{P}^{a,p}:|\mathrm{Im}x|<s_l,|y|<r_l^2,||z||_{a,p}<r_l,||\bar z||_{a,p}<r_l\}$.

In the following part of this paper, with C or c a universal constant, whose size may be different in different places, these constants might depend on n, d, p and q. If $f\le Cg$, we write this inequality as $f\lessdot g$, when we do not care about the size of the constant C or c. Similarly, if $f\ge g$ we write $f\gtrdot g$.

2. Iterative Lemma. *Consider a family of Hamiltonian functions H_l $(0\le l\le m)$:*

$$H_l=N_l+P_l=\omega_l(\xi)y+\sum_{j\ge1}\lambda_j\langle z_j,\bar z_j\rangle+\sum_{j\ge1}\langle B_{l,jj}(x;\xi)z_j,\bar z_j\rangle+P_l(x,y,z,\bar z;\xi), \tag{3.1}$$

where $B_{l,jj}(x;\xi)$'s are analytic in $x\in\mathbb{T}_{s_l}$ for any fixed $\xi\in\Pi_l$, and $B_{l,jj}(x;\xi)$'s (for fixed $x\in\mathbb{T}_{s_l}$) are continuously differentiable in $\xi\in\Pi_l$, $P_l(x,y,z,\bar z;\xi)$ is analytic in $(x,y,z,\bar z)\in D(s_l,r_l)$ and continuously differentiable in $\xi\in\Pi_l$. Assume $B_l(x;\xi)=diag(B_{l,jj}(x;\xi):j\ge1)$'s satisfying the following conditions:

(l.1) Reality condition.

$$(B_{l,jj}(x;\xi))^T=B_{l,jj}(x;\xi)\in\mathbb{R}^{d_j\times d_j},\quad\forall\ (x;\xi)\in\mathbb{T}\times\Pi_l. \tag{3.2}$$

(l.2) Finiteness of Fourier modes.

$$B_{l,jj}(x;\xi)=\sum_{|k|\le K_l}\widehat{B}_{l,jj}(k;\xi)e^{ikx}. \tag{3.3}$$

(l.3) Boundedness.

$$||B_l(x;\xi)||_{p,q;\mathbb{T}_{s_l}\times\Pi_l}\lessdot\varepsilon,\ \ ||B_l(x;\xi)||^{\mathscr{L}}_{p,q;\mathbb{T}_{s_l}\times\Pi_l}\lessdot\varepsilon^{\frac13}. \tag{3.4}$$

Moreover, we assume the parameter sets Π_l's satisfy:

(l.4)

$$\Pi_0\supset\cdots\supset\cdots\Pi_l\cdots\supset\Pi_m \tag{3.5}$$

with

$$Meas\ \Pi_l\ge(Meas\ \Pi)(1-O(\frac{1}{l^2})),\quad 1\le l\le m. \tag{3.6}$$

(l.5) The perturbation $P_l(x,y,z,\bar z;\xi)$ is analytic in the space coordinate domain $D(s_l,r_l)$ and continuously differentiable in $\xi\in\Pi_l$, and

$$\overline{P_l(x,y,z,\bar z;\xi)}=P_l(x,y,z,\bar z;\xi),\ \ \forall\ (x,y,z,\bar z;\xi)\in D_{\mathfrak{R}}(s_l,r_l)\times\Pi_l. \tag{3.7}$$

Moreover, its vector field $X_{P_l}:=(\partial_yP_l,-\partial_xP_l,\ i\partial_zP_l,-i\partial_{\bar z}P_l)$ defines on $D(s_l,r_l)$ a analytic map

$$X_{P_l}:D(s_l,r_l)\subset\mathscr{P}^{a,p}\to\mathscr{P}^{a,q}. \tag{3.8}$$

(l.6) In addition, the vector field X_{P_l} is analytic in the domain $D(s_l,r_l)$ with small norms

$$||X_{P_l}||_{r_l,q;D(s_l,r_l)\times\Pi_l}\lessdot\varepsilon_l,\quad ||X_{P_l}||^{\mathscr{L}}_{r_l,q;D(s_l,r_l)\times\Pi_l}\lessdot\varepsilon_l^{1/3}. \tag{3.9}$$

Then there is an absolute positive constant ε_* small enough such that, if $0<\varepsilon<\varepsilon_*$, there is a subset $\Pi_{m+1}\subset\Pi_m$, and a change of variables $\Phi_{m+1}:\mathscr{D}_{m+1}:=D(s_{m+1},r_{m+1})\times\Pi_{m+1}\to D(s_m,r_m)$ being real for real argument, analytic in $(x,y,z,\bar z)\in D(s_{m+1},r_{m+1})$ and continuously differentiable in $\xi\in\Pi_{m+1}$, as well as the following estimates hold true:

$$||\Phi_{m+1}-id||_{r_m,q;\mathscr{D}_{m+1}}\lessdot\varepsilon_m^{1/2}, \tag{3.10}$$

and

$$||\Phi_{m+1}-id||^{\mathscr{L}}_{r_m,q;\mathscr{D}_{m+1}} \lessdot \varepsilon_m^{1/4}, \tag{3.11}$$

$$||D\Phi_{m+1}-I||_{r_m,q,q;\mathscr{D}_{m+1}} \lessdot \varepsilon_m^{1/2}, \tag{3.12}$$

$$||D\Phi_{m+1}-I||^{\mathscr{L}}_{r_m,q,q;\mathscr{D}_{m+1}} \lessdot \varepsilon_m^{1/4}. \tag{3.13}$$

The estimate of $D\Phi_{m+1}-I$ holds also with p for q. Furthermore, the new Hamiltonian $H_{m+1} := H_m \circ \Phi_{m+1}$ is of the form

$$H_{m+1} = N_{m+1}+P_{m+1} = \omega_{m+1}(\xi)y+\sum_{j\geq 1}\lambda_j\langle z_j,\bar{z}_j\rangle+\sum_{j\geq 1}\langle B_{m+1,jj}(x;\xi)z_j,\bar{z}_j\rangle+P_{m+1}(x,y,z,\bar{z};\xi),$$

and it satisfies all the above conditions (l.1-6) with l being replaced by $m+1$.

Proof: The iterative lemma is proved by the usual Newton-type iteration procedure which involves $m+1$ symplectic coordinates changes. At each step of the KAM scheme, we consider a Hamiltonian vector field with

$$H_\nu = N_\nu + P_\nu, \quad \nu = 1,2,\cdots,$$

where N_ν is an "integrable normal form" and P_ν is defined in some set of the form $D(s_\nu,r_\nu)\times\Pi_\nu$. We then construct a map

$$\Phi_\nu : D(s_{\nu+1},r_{\nu+1})\times\Pi_{\nu+1}\subset D(s_\nu,r_\nu)\times\Pi_\nu \to D(s_\nu,r_\nu)\times\Pi_\nu,$$

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

$$||X_{H_\nu\circ\Phi_\nu}-X_{N_{\nu+1}}||_{r_{\nu+1},q;D(s_{\nu+1},r_{\nu+1})\times\Pi_{\nu+1}} \leq \varepsilon_\nu^\kappa$$

with some new normal form $N_{\nu+1}$ and for some fixed ν-independent constant $\kappa>1$. To simplify notations, in what follows, the quantities without subscripts refer to quantities at the ν-th step, while the quantities with subscripts $+$ denote the corresponding quantities at the $\nu+1$-th step.

Expand P into the Fourier-Taylor series

$$P = \sum_{k,l,\alpha_1,\alpha_2}\widehat{P_{l\alpha_1\alpha_2}}(k)e^{\mathrm{i}kx}y^l z^{\alpha_1}\bar{z}^{\alpha_2},$$

where $k\in\mathbb{Z}$, $l\in\mathbb{N}=\{0,1,2,\cdots\}$ and $\alpha_1,\alpha_2\in\otimes_{j\in\mathbb{Z}_+}\mathbb{N}^{d_j}$ with finitely many non-vanishing components, i.e. there exists $j_0>0$ such that $z^{\alpha_1}=\prod_{j=1}^{j_0}z_j^{\alpha_{1,j}}=\prod_{j=1}^{j_0}\prod_{l=1}^{d_j}(z_j^l)^{\alpha_{1,j}^l}$ and similar equalities for $\bar{z}^{\alpha_2}$. Let R be the truncation of P given by

$$R(x,y,z,\bar{z}) = \sum_{2|l|+|\alpha_1|+|\alpha_2|\leq 2}\sum_{k\in\mathbb{Z}}\widehat{P_{l\alpha_1\alpha_2}}(k)e^{\mathrm{i}kx}y^l z^{\alpha_1}\bar{z}^{\alpha_2}, \tag{3.14}$$

here for $\alpha\in\otimes_{j\in\mathbb{Z}_+}\mathbb{N}^{d_j}$, we denote $|\alpha|=\sum_{j\in\mathbb{Z}_+}\sum_{l=1}^{d_j}|\alpha_j^l|$. We will approximate R by P, and we have some estimates of R:

$$||X_R||^*_{r,q,D(s,r)\times\Pi} \lessdot ||X_P||^*_{r,q,D(s,r)\times\Pi} \lessdot \varepsilon_\nu^*, \tag{3.15}$$

$$||X_{P-R}||^*_{r,q,D(s,4\eta r)\times\Pi} \lessdot \eta||X_P||^*_{r,q,D(s,r)\times\Pi} \lessdot \eta\varepsilon_\nu^*, \tag{3.16}$$

for $0<\eta\ll 1$, where *= the blank or $\mathscr{L}$ (that is, $\varepsilon_\nu^*=\varepsilon_\nu$ or $\varepsilon_\nu^{\frac{1}{3}}$). Actually, the proof of (3.15) and (3.16) are similar to that of formula (7) of [15]. It is convenient to rewrite R as $R=R^0+R^1+R^2$, where

$$\begin{aligned} R^0 &= R^x + R^y y, \\ R^1 &= \langle R^z, z\rangle + \langle R^{\bar z}, \bar z\rangle = \sum_{j\geq 1}\langle R^z_j, z_j\rangle + \sum_{j\geq 1}\langle R^{\bar z}_j, \bar z_j\rangle, \\ R^2 &= \frac{1}{2}\langle R^{zz} z, z\rangle + \langle R^{z\bar z} z, \bar z\rangle + \frac{1}{2}\langle R^{\bar z\bar z}\bar z, \bar z\rangle \\ &= \frac{1}{2}\sum_{i,j\geq 1}\langle R^{zz}_{ji} z_i, z_j\rangle + \sum_{i,j\geq 1}\langle R^{z\bar z}_{ji} z_i, \bar z_j\rangle + \frac{1}{2}\sum_{i,j\geq 1}\langle R^{\bar z\bar z}_{ji}\bar z_i, \bar z_j\rangle \end{aligned}$$

with $R^x, R^y : \mathbb{T}_s\times\Pi\to\mathbb{C}$; $R^z = (R^z_j : j\geq 1)$, $R^{\bar z} = (R^{\bar z}_j : j\geq 1) : \mathbb{T}_s\times\Pi\to\mathscr{Z}^{a,q}$; $R^{zz} = (R^{zz}_{ji} : i,j\geq 1)$, $R^{z\bar z} = (R^{z\bar z}_{ji} : i,j\geq 1)$, $R^{\bar z\bar z} = (R^{\bar z\bar z}_{ji} : i,j\geq 1) : \mathbb{T}_s\times\Pi\to\mathscr{L}(\mathscr{Z}^{a,p}, \mathscr{Z}^{a,q})$. By (3.7) and (3.14), we have

$$\overline{R(x,y,z,\bar z)} = R(x,y,z,\bar z), \quad (x,y,z,\bar z)\in D_{\mathfrak{R}}(s,r)\times\Pi. \tag{3.17}$$

Then by (3.17), we have that for real x,

$$\overline{R^x(x)} = R^x(x), \quad \overline{R^y(x)} = R^y(x), \tag{3.18}$$

$$\overline{R^z(x)} = R^{\bar z}(x), \tag{3.19}$$

$$(R^{zz}_{ji}(x))^T = R^{zz}_{ij}(x), \quad \overline{R^{zz}_{ji}(x)} = R^{\bar z\bar z}_{ji}(x), \tag{3.20}$$

$$(R^{z\bar z}_{ji}(x))^T = R^{z\bar z}_{ij}(x), \quad \overline{R^{z\bar z}_{ji}(x)} = R^{z\bar z}_{ji}(x). \tag{3.21}$$

Let

$$\Gamma_K R = \sum_{2|l|+|\alpha_1|+|\alpha_2|\leq 2}\ \sum_{|k|\leq K}\widehat{P_{l\alpha_1\alpha_2}}(k)e^{\mathrm{i}kx}y^l z^{\alpha_1}\bar z^{\alpha_2}, \quad R_K = R - \Gamma_K R.$$

And by the choice of K_+, K, s, $\sigma = s/10$ and η_+ we have $\sigma K_+\geq\sigma K = |\ln\varepsilon_\nu|\geq|\ln\eta_+^2|$, and $K_+\eta_+ < 1$. Then by Lemma 4.2 in [20], we have

$$||X_{\Gamma_{K_+}R}||^*_{r,q,D(s-\sigma,r)\times\Pi}\leq ||X_R||^*_{r,q,D(s-\sigma,r)\times\Pi}\lessdot\varepsilon^*_\nu, \tag{3.22}$$

$$||X_{R_{K_+}}||^*_{r,q,D(s-\sigma,r)\times\Pi}\lessdot\eta_+\varepsilon^*_\nu\leq\varepsilon^*_{\nu+1}. \tag{3.23}$$

where *=the blank or $\mathscr{L}$ (that is, $\varepsilon^*_{\nu+1} = \varepsilon_{\nu+1}$ or $\varepsilon^{\frac{1}{3}}_{\nu+1}$). Now we can rewrite $P = R + (P-R) = \Gamma_{K_+}R + R_{K_+} + (P-R)$. Let $\mathscr{R} = \Gamma_{K_+}R$. Thus we can write

$$\begin{aligned} \mathscr{R} &= \mathscr{R}^x + \mathscr{R}^y y + \langle\mathscr{R}^z, z\rangle + \langle\mathscr{R}^{\bar z}, \bar z\rangle \\ &\quad \frac{1}{2}\langle\mathscr{R}^{zz} z, z\rangle + \langle\mathscr{R}^{z\bar z} z, \bar z\rangle + \frac{1}{2}\langle\mathscr{R}^{\bar z\bar z}\bar z, \bar z\rangle, \end{aligned}$$

here

$$\mathscr{R}^* = \sum_{|k|\leq K_+}\widehat{R^*}(k)e^{\mathrm{i}kx} \quad\text{for}\quad * = x, y, z, \bar z, zz, z\bar z, \bar z\bar z. \tag{3.24}$$

And as $\widehat{\mathscr{R}^x}(0)$ does not affect the equation of motion, we may assume $\widehat{\mathscr{R}^x}(0) = 0$. Furthermore, under the smallness assumption (3.9) on P, we have the following estimates by Lemma 4.3 in [20]:

$$|\partial_x\mathscr{R}^x|_{\mathbb{T}_s\times\Pi}\leq r^2\varepsilon_\nu, \quad |\partial_x\mathscr{R}^x|^{\mathscr{L}}_{\mathbb{T}_s\times\Pi}\leq r^2\varepsilon^{\mathscr{L}}_\nu, \tag{3.25}$$

$$|\mathscr{R}^y|_{\mathbb{T}_s\times\Pi}\leq\varepsilon_\nu, \quad |\mathscr{R}^y|^{\mathscr{L}}_{\mathbb{T}_s\times\Pi}\leq\varepsilon^{\mathscr{L}}_\nu, \tag{3.26}$$

$$||\mathscr{R}^u||_{p,\mathbb{T}_s\times\Pi} \le r\varepsilon_\nu, \quad ||\mathscr{R}^u||^{\mathscr{L}}_{p,\mathbb{T}_s\times\Pi} \le r\varepsilon^{\mathscr{L}}_\nu, \quad u\in\{z,\bar z\}, \tag{3.27}$$

$$||\mathscr{R}^{uv}||_{p,q,\mathbb{T}_s\times\Pi} \le \varepsilon_\nu, \quad ||\mathscr{R}^{uv}||^{\mathscr{L}}_{p,q,\mathbb{T}_s\times\Pi} \le \varepsilon^{\mathscr{L}}_\nu, \quad u,v\in\{z,\bar z\}. \tag{3.28}$$

Then we shall find a function F defined in domain $D_+(s_+,r_+)$, such that the time one map ϕ^1_F of the Hamiltonian vector field X_F defines a map from $D_+\to D$ and transforms H into H_+. To this end, we suppose that F is of the same form as R, that is, $F=F^0+F^1+F^2$, where

$$\begin{aligned} F^0 &= F^x+F^y y,\\ F^1 &= \langle F^z,z\rangle+\langle F^{\bar z},\bar z\rangle = \sum_{j\ge1}\langle F^z_j,z_j\rangle+\sum_{j\ge1}\langle F^{\bar z}_j,\bar z_j\rangle,\\ F^2 &= \frac12\langle F^{zz}z,z\rangle+\langle F^{z\bar z}z,\bar z\rangle+\frac12\langle F^{\bar z\bar z}\bar z,\bar z\rangle\\ &= \frac12\sum_{i,j\ge1}\langle F^{zz}_{ji}z_i,z_j\rangle+\sum_{i,j\ge1}\langle F^{z\bar z}_{ji}z_i,\bar z_j\rangle+\frac12\sum_{i,j\ge1}\langle F^{\bar z\bar z}_{ji}\bar z_i,\bar z_j\rangle. \end{aligned}$$

Thus by second order Taylor formula, we have

$$H\circ\phi^1_F = (N+\mathscr{R})\circ X^1_F+(R_K+P-R)\circ X^1_F = N_+ + P^1_+ + P^2_+,$$

where

$$N_+ = N+\widehat{\mathscr{R}^y}(0)y+\sum_{j\ge1}\langle \mathscr{R}^{z\bar z}_{jj}z_j,\bar z_j\rangle = \omega_+y+\sum_{j\ge1}\lambda_j\langle z_j,\bar z_j\rangle+\sum_{j\ge1}\langle B_{+,jj}(x;\xi)z_j,\bar z_j\rangle,$$

$$\omega_+ = \omega+\widehat{\mathscr{R}^y}(0),$$

$$B_{+,jj}(x;\xi) = B_{jj}(x;\xi)+\mathscr{R}^{z\bar z}_{jj}(x;\xi), \tag{3.29}$$

$$P^1_+ = \frac12\int_0^1 ds\int_0^s\{\{N+\mathscr{R},F\},F\}\circ X^t_F dt+\{\mathscr{R},F\}+(R_K+P-R)\circ X^1_F, \tag{3.30}$$

and

$$P^2_+ = \{N,F\}+\mathscr{R}-\widehat{\mathscr{R}}(0)y-\sum_{j\ge1}\langle\mathscr{R}^{z\bar z}_{jj}z_j,\bar z_j\rangle. \tag{3.31}$$

Actually we make our attempts to find out a function F to make (3.31) vanish or of higher order in the following. Moreover, we can compute the Poisson Bracket $\{F,N\}$ with respect to the symplectic structure $\mathrm{d}x\wedge\mathrm{d}y+\mathrm{i}\sum_{j\ge1}\sum_{l=1}^{d_j}dz^l_j\wedge\mathrm{d}\bar z^l_j$ and derive the homological equations in components. That is, F satisfies the following three kinds of homological equations according to the small divisor condition:

(1) The Kolmogrov condition:

$$\omega\frac{\mathrm{d}}{\mathrm{d}x}F^x = \mathscr{R}^x(x;\xi), \tag{3.32}$$

$$\omega\frac{\mathrm{d}}{\mathrm{d}x}F^y = \mathscr{R}^y(x;\xi)-\widehat{\mathscr{R}^y}(0). \tag{3.33}$$

(2) The first Melnikov condition:

$$-\mathrm{i}\omega\frac{\mathrm{d}}{\mathrm{d}x}F^z_j-\lambda_jF^z_j-B_{jj}(x;\xi)F^z_j = -\mathrm{i}\mathscr{R}^z_j(x;\xi),\ \ j\ge1, \tag{3.34}$$

$$-\mathrm{i}\omega\frac{\mathrm{d}}{\mathrm{d}x}F^{\bar z}_j+\lambda_jF^{\bar z}_j+B_{jj}(x;\xi)F^{\bar z}_j = -\mathrm{i}\mathscr{R}^{\bar z}_j(x;\xi),\ \ j\ge1, \tag{3.35}$$

$$-\mathrm{i}\omega\frac{\mathrm{d}}{\mathrm{d}x}F_{ij}^{zz}-(\lambda_i+\lambda_j)F_{ij}^{zz}-B_{ii}(x;\xi)F_{ij}^{zz}-F_{ij}^{zz}B_{jj}(x;\xi)=-\mathrm{i}\mathscr{R}_{ij}^{zz}(x;\xi),\ i,j\geq 1, \tag{3.36}$$

$$-\mathrm{i}\omega\frac{\mathrm{d}}{\mathrm{d}x}F_{ij}^{\bar z\bar z}+(\lambda_i+\lambda_j)F_{ij}^{\bar z\bar z}+B_{ii}(x;\xi)F_{ij}^{\bar z\bar z}+F_{ij}^{\bar z\bar z}B_{jj}(x;\xi)=-\mathrm{i}\mathscr{R}_{ij}^{\bar z\bar z}(x;\xi),\ i,j\geq 1. \tag{3.37}$$

(3) The second Melnikov condition:

$$-\mathrm{i}\omega\frac{\mathrm{d}}{\mathrm{d}x}F_{ij}^{z\bar z}-(\lambda_i-\lambda_j)F_{ij}^{z\bar z}-B_{ii}(x;\xi)F_{ij}^{z\bar z}+F_{ij}^{z\bar z}B_{jj}(x;\xi)=-\mathrm{i}\mathscr{R}_{ij}^{z\bar z}(x;\xi),\ i\neq j, \tag{3.38}$$

and $F_{ii}^{z\bar z}=0$.

Lemma 3.1 *(Solutions of* (3.32), (3.33)*) There exists a subset $\Pi_+^1\subset\Pi$ with $Meas\Pi_+^1\geq(Meas\Pi)(1-K_+^{-1})$, and on $\mathbb{T}_{s-\sigma}\times\Pi_+^1$,* (3.32) *has a solution $F^x(x;\xi)$ which is analytic in $x\in\mathbb{T}_{s-\sigma}$ for ξ fixed and continuously differentiable in ξ for x fixed, and which is real for real argument, such that*

$$|\frac{d}{dx}F^x|_{\mathbb{T}_s\times\Pi_+^1}\lessdot K_+^C r^2\varepsilon_\nu,\quad |\frac{d}{dx}F^x|^{\mathscr{L}}_{\mathbb{T}_s\times\Pi_+^1}\lessdot K_+^C r^2\varepsilon_\nu^{\mathscr{L}}. \tag{3.39}$$

Similarly, on $\mathbb{T}_{s-\sigma}\times\Pi_+^1$, (3.33) *has a solution $F^y(x;\xi)$ which is analytic in $x\in\mathbb{T}_{s-\sigma}$ for ξ fixed and continuously differentiable in ξ for x fixed, and which is real for real argument, such that*

$$|\frac{d}{dx}F^y|_{\mathbb{T}_s\times\Pi_+^1}\lessdot K_+^C\varepsilon_\nu,\quad |\frac{d}{dx}F^y|^{\mathscr{L}}_{\mathbb{T}_s\times\Pi_+^1}\lessdot K_+^C\varepsilon_\nu^{\mathscr{L}}. \tag{3.40}$$

Furthermore, we have

$$||X_{F^0}||_{r,p;D(s-3\sigma,r)}\lessdot K_+^C\varepsilon_\nu,\quad ||X_{F^0}||^{\mathscr{L}}_{r,p;D(s-3\sigma,r)}\lessdot K_+^C\varepsilon_\nu^{\mathscr{L}}. \tag{3.41}$$

Proof: The proof is standard in KAM theory, see Proposition 1, 2 in [20] for details. □

As to (3.34)-(3.38), we can use Theorem 1.1 to obtain the estimates. In the following we consider the equation for $F_{ij}^{z\bar z}$, which is slightly more complicated than the ones for F_j^z, $F_j^{\bar z}$, F_{ij}^{zz} and $F_{ij}^{\bar z\bar z}$.

Lemma 3.2 (*Solutions of* (3.38)) *There exists a subset $\Pi_+^2\subset\Pi$ with $Meas\Pi_+^2\geq(Meas\Pi)(1-\frac{1}{(\nu+1)^2})$, and on $\mathbb{T}_{s-\sigma}\times\Pi_+^2$,* (3.38) *has a solution $F_{ij}^{z\bar z}(x;\xi)$ which is analytic in $x\in\mathbb{T}_{s-\sigma}$ for ξ fixed and continuously differentiable in ξ for x fixed, and which is real for real argument, such that*

$$||F^{z\bar z}||_{p,p,\mathbb{T}_{s-2\sigma}},\ ||F^{z\bar z}||_{q,q,\mathbb{T}_{s-2\sigma}}\leq K_+^C\varepsilon_\nu, \tag{3.42}$$

$$||F^{z\bar z}||^{\mathscr{L}}_{p,p,\mathbb{T}_{s-2\sigma}},\ ||F^{z\bar z}||^{\mathscr{L}}_{q,q,\mathbb{T}_{s-2\sigma}}\leq K_+^C\varepsilon_\nu^{\mathscr{L}}, \tag{3.43}$$

$$||i\omega\frac{d}{dx}F^{z\bar z}+(\Lambda+B)F^{z\bar z}-F^{z\bar z}(\Lambda+B)-\mathscr{R}^{z\bar z}||_{q,q,\mathbb{T}_{s-3\sigma}}\lessdot\varepsilon_{\nu+1}, \tag{3.44}$$

$$||i\omega\frac{d}{dx}F^{z\bar z}+(\Lambda+B)F^{z\bar z}-F^{z\bar z}(\Lambda+B)-\mathscr{R}^{z\bar z}||^{\mathscr{L}}_{q,q,\mathbb{T}_{s-4\sigma}}\lessdot\varepsilon_{\nu+1}^{\mathscr{L}}. \tag{3.45}$$

Proof: Firstly for $i\neq j$, we can use Lemma 5.6 to change (3.38) into vector equations:

$$-\mathrm{i}\omega\frac{\mathrm{d}}{\mathrm{d}x}(F_{ij}^{z\bar z})'-(\lambda_i-\lambda_j)(F_{ij}^{z\bar z})'-[E_{d_j}\otimes B_{ii}(x;\xi)-B_{jj}(x;\xi)\otimes E_{d_i}](F_{ij}^{z\bar z})'=\mathrm{i}(\mathscr{R}_{ij}^{z\bar z})'(x;\xi). \tag{3.46}$$

Here $\otimes$ means the tensor product of two matrices. And note that $F_{ij}^{z\bar z}$, $\mathscr{R}_{ij}^{z\bar z}$ are $d_i\times d_j$ matrices, if we denote $F_{ij}^{z\bar z}=(F_{ij,1}^{z\bar z},\cdots,F_{ij,d_j}^{z\bar z})$, then $(F_{ij}^{z\bar z})'=((F_{ij,1}^{z\bar z})^T,\cdots,(F_{ij,d_j}^{z\bar z})^T)^T$. Similarly, we have $(\mathscr{R}_{ij}^{z\bar z})'$. By

$(\nu.1)$, $(\nu.2)$ and $(\nu.3)$ and the definition of the tensor product, we can check that (1)-(4) in Theorem 1.1 holds true on $\mathbb{T}_{s-\sigma}$. By the proof of Theorem 1.1 in Section 2, for $|\lambda_i-\lambda_j|\le 2K_+$, there exists a subset $\Pi_+^2\subset\Pi$, with

$$\mathrm{Meas}\Pi_+^2\ge \mathrm{Meas}\Pi(1-\frac{1}{(\nu+1)^2}),\tag{3.47}$$

where in the last inequality, we use the estimate (1.9) and $s|\lambda_i-\lambda_j|^{2\theta-1}\le sK_+^{2\theta-1}\le\frac{1}{(\nu+1)^2}$. And on $\mathbb{T}_{s-\sigma}\times\Pi_+^2$, there exists $(F_{ij}^{z\bar z})'$ which satisfies

$$|(F_{ij}^{z\bar z})'|_{\mathbb{T}_{s-\sigma}\times\Pi_+}\lessdot K_+^C|(\mathscr{R}_{ij}^{z\bar z})'|_{\mathbb{T}_s\times\Pi},\tag{3.48}$$

and

$$|(F_{ij}^{z\bar z})'|_{\mathbb{T}_{s-\sigma}\times\Pi_+}^{\mathscr{L}}\lessdot K_+^C\big(|(\mathscr{R}_{ij}^{z\bar z})'|_{\mathbb{T}_s\times\Pi}+(\mathscr{R}_{ij}^{z\bar z})'|_{\mathbb{T}_s\times\Pi}^{\mathscr{L}}\big).\tag{3.49}$$

And for $|\lambda_i-\lambda_j|>2K_+$, we can get (3.48) and (3.49) or even better estimates. Thus by Lemma 5.4, we can get (3.42) and (3.43). Furthermore, by similar proof, we obtain (3.44) and (3.45) hold. □

The same, and even better estimates hold for F^{zz} and $F^{\bar z\bar z}$. Multiplying with z and $\bar z$ we then get

$$\frac{1}{r^2}|F^2|_{D(s-2\sigma,r)}\lessdot K_+^C\varepsilon_\nu,$$

and finally with Cauchy's estimate

$$||X_{F^2}||_{r,p;D(s-3\sigma,r)}\lessdot K_+^C\varepsilon_\nu.$$

Applying ∂_ξ to both sides of (3.38) and using the method similar to the above, we get

$$||X_{F^2}||_{r,p;D(s-3\sigma,r)}^{\mathscr{L}}\lessdot K_+^C\varepsilon_\nu^{\mathscr{L}}.$$

And similarly, we have

$$\frac{1}{r}|F^1|_{D(s-2\sigma,r)}\lessdot K_+^C\varepsilon_\nu,\quad ||X_{F^1}||_{r,p;D(s-3\sigma,r)}\lessdot K_+^C\varepsilon_\nu,\quad ||X_{F^1}||_{r,p;D(s-3\sigma,r)}^{\mathscr{L}}\lessdot K_+^C\varepsilon_\nu^{\mathscr{L}}.$$

Thus

$$||X_F||_{r,p;D(s-3\sigma,r)}\lessdot K_+^C\varepsilon_\nu,\quad ||X_F||_{r,p;D(s-3\sigma,r)}^{\mathscr{L}}\lessdot K_+^C\varepsilon_\nu^{\mathscr{L}}.$$

By (3.21), we have

$$(\mathscr{R}_{jj}^{z\bar z}(x;\xi))^T=\mathscr{R}_{jj}^{z\bar z}(x;\xi)=\overline{\mathscr{R}_{jj}^{z\bar z}(x;\xi)},\quad (x;\xi)\in\mathbb{T}\times\Pi.\tag{3.50}$$

Then $(\nu+1.1)$ holds true. By (3.24), we have $(\nu+1.2)$ fulfills. And by (3.28), we get

$$||\mathrm{diag}(\mathscr{R}_{jj}^{z\bar z}(x;\xi):j\ge 1)||_{p,q,\mathbb{T}_s\times\Pi}^{*}\le||\mathscr{R}^{z\bar z}||_{p,q,\mathbb{T}_s\times\Pi}^{*}\le\varepsilon_\nu^{*},\tag{3.51}$$

where $^*=$ the blank or $\mathscr{L}$. As $\varepsilon_\nu\ll\varepsilon$, we also can get $(\nu+1.3)$. Let $\Pi_+=\Pi_+^1\bigcap\Pi_+^2$, then we have

$$\mathrm{Meas}\Pi_+\ge\mathrm{Meas}\Pi(1-\frac{1}{(\nu+1)^2}).\tag{3.52}$$

Thus $(\nu+1.4)$ is true. The estimate for P_+^1 is well-known in KAM theory, we have

$$||X_{P_+^1}||_{r,q;D(s_+,r)}^{*}\lessdot\varepsilon_{\nu+1}^{*}.$$

Moreover, by (3.44) and (3.45) and similar formula for other terms, we have

$$||X_{P_+^2}||^*_{r,q;D(s_+,r)} \lessdot \varepsilon^*_{\nu+1}. \tag{3.53}$$

Then we can get

$$||X_{P_+}||^*_{r,q;D(s_+,r)} \lessdot \varepsilon^*_{\nu+1}, \tag{3.54}$$

where $*=$ the blank or $\mathscr{L}$ in the above three inequalities. As F^x, F^y, F^z, $F^{\bar{z}}$, F^{zz}, $F^{z\bar{z}}$, $F^{\bar{z}\bar{z}}$ are the unique solutions of the homological equations (3.32)-(3.38) respectively, by (3.18)-(3.21) and $\overline{B_{ii}(x;\xi)} = B_{ii}(x;\xi)$ for $(x;\xi) \in \mathbb{T}\times\Pi$. Thus by the uniqueness of the solutions, for $(x;\xi) \in \mathbb{T}\times\Pi_+$, we get

$$\overline{F^x(x)} = F^x, \quad \overline{F^y(x)} = F^y, \quad \overline{F^z(x)} = F^{\bar{z}}(x), \tag{3.55}$$

$$(F^{zz}_{ij}(x))^T = F^{zz}_{ji}(x), \quad \overline{F^{zz}_{ij}(x)} = F^{\bar{z}\bar{z}}_{ij}(x), \tag{3.56}$$

$$(F^{z\bar{z}}_{ij}(x))^T = F^{z\bar{z}}_{ji}(x), \quad \overline{F^{z\bar{z}}_{ij}(x)} = F^{z\bar{z}}_{ij}(x). \tag{3.57}$$

To sum up, by (3.55), (3.56), (3.57), we have

$$\overline{F(x,y,z,\bar{z};\xi)} = F(x,y,z,\bar{z};\xi),\ \forall\,(x,y,z,\bar{z};\xi) \in D_{\mathfrak{R}}(s_+,r_+)\times\Pi_+. \tag{3.58}$$

So $(\nu+1.5)$ also holds true. The estimates (3.10)-(3.13) can be similarly obtained as in [20]. □

3. Proof of Theorem 1.2. The proof is similar to that of [10]. Here we give the outline. By assumption in Theorem 1.2, the conditions (l.1-6) in the iterative lemma are fulfilled with $l=0$. Hence the iterative lemma applies to H which is defined in Theorem 4.1 Inductively, we get what as follows:
(i) Domains: for $m = 0,1,2,\cdots$

$$\mathscr{D}_m := D(s_m,r_m)\times\Pi_m, \quad \mathscr{D}_{m+1}\subset\mathscr{D}_m;$$

(ii) Coordinate changes:

$$\Psi^m = \Phi_1\circ\cdots\circ\Phi_{m+1} : \mathscr{D}_{m+1} \to D(s_0,r_0);$$

(iii) Hamiltonian functions $H_m (m=0,1,\cdots)$ satisfy the conditions $(l.1-7)$ with l replaced by m. Let $\Pi_\infty = \bigcap_{m=0}^\infty \Pi_m, \mathscr{D}_\infty = \bigcap_{m=0}^\infty \mathscr{D}_m$. By the same argument as in [10, p.134], we conclude that $\Psi^m, D\Psi^m, H_m, X_{H_m}, \omega_m$ converges uniformly on the domain $\mathscr{D}_\infty$, and $X_{\widetilde{H}_\infty}\circ\Psi^\infty = D\Psi^\infty\cdot X_{\omega_\infty}$ where

$$\widetilde{H}_\infty := \lim_{m\to\infty}\widetilde{H}_m = \omega_\infty y + \sum_{n\geq1}\lambda_n\langle z_n,\bar{z}_n\rangle + \sum_{n\geq1}\langle B^\infty_{nn}z_n,\bar{z}_n\rangle,$$

here $B^\infty_{nn} = \lim_{m\to\infty}B_{m,nn}$, and X_{ω_∞} is the constant vector field ω_∞ on the torus $\mathbb{T}$. Thus, $\mathbb{T}\times\{y=0\}\times\{z=0\}\times\{\bar{z}=0\}$ is an embedding torus with rotational frequencies $\omega_\infty\in\omega_\infty(\Pi_\infty)$ of the Hamiltonian H_∞. Returning to the original Hamiltonian H_∞, it has an embedding torus $\Phi^\infty(\mathbb{T}\times\{y=0\}\times\{z=0\}\times\{\bar{z}=0\})$ with frequencies ω_∞. This proves the Theorem. □

4 Preserving of finite-gap solutions for coupled KdV Equation

Consider the coupled Korteweg-de Vries equation

$$\begin{cases} u_t = -u_{xxx} + 6uu_x + \varepsilon[\dfrac{\partial f(x,u,v)}{\partial u}]_x, \\ v_t = -v_{xxx} + 6vv_x + \varepsilon[\dfrac{\partial f(x,u,v)}{\partial v}]_x, \end{cases} \tag{4.1}$$

where $x\in\mathbb{T}$ and $\varepsilon>0$ is sufficiently small. Introduce for any integer $Z\geq 0$ the phase space

$$\mathscr{H}_0^Z\times\mathscr{H}_0^Z=\{u\in L^2(\mathbb{T};\mathbb{R}):\hat{u}(0)=0,\ ||u||_Z<\infty\}\times\{v\in L^2(\mathbb{T};\mathbb{R}):\hat{v}(0)=0,\ ||v||_Z<\infty\}$$

of real valued functions on $\mathbb{T}=\mathbb{R}/2\pi\mathbb{Z}$, where $||u||_Z^2=|\hat{u}(0)|^2+\sum_{k\in\mathbb{Z}}|k|^{2Z}|\hat{u}(k)|^2$ is defined in terms of the Fourier transform $\hat{u}$ of u, $u(x)=\sum_{k\in\mathbb{Z}}\hat{u}(k)e^{\mathrm{i}kx}$. We endow $\mathscr{H}_0^Z\times\mathscr{H}_0^Z$ with the Poisson structure proposed by Gardner

$$\{F,G\}=\int_{\mathbb{T}}\{\frac{\partial F}{\partial u(x)}\frac{\mathrm{d}}{\mathrm{d}x}\frac{\partial G}{\partial u(x)}+\frac{\partial F}{\partial v(x)}\frac{\mathrm{d}}{\mathrm{d}x}\frac{\partial G}{\partial v(x)}\}\mathrm{d}x,$$

where F,G are differentiable functions on $\mathscr{H}_0^Z\times\mathscr{H}_0^Z$ with L^2-gradients in $\mathscr{H}_0^1\times\mathscr{H}_0^1$. The Hamiltonian corresponding to (4.1) is then given by

$$\begin{aligned}H(u,v)&=\int_{\mathbb{T}}(\frac{1}{2}u_x^2+\frac{1}{2}v_x^2+u^3+v^3)\mathrm{d}x+\varepsilon\int_{\mathbb{T}}f(x,u,v)\mathrm{d}x\\&=H_0(u,v)+\varepsilon K(u,v),\end{aligned}\tag{4.2}$$

where $H_0(u,v)=\int_{\mathbb{T}}(\frac{1}{2}u_x^2+\frac{1}{2}v_x^2+u^3+v^3)\mathrm{d}x$, $K(u,v)=\int_{\mathbb{T}}f(x,u,v)\mathrm{d}x$. And

$$\frac{\partial}{\partial t}\begin{pmatrix}u\\v\end{pmatrix}=\frac{\mathrm{d}}{\mathrm{d}x}\begin{pmatrix}\frac{\partial H}{\partial u}\\\frac{\partial H}{\partial v}\end{pmatrix}$$

is (4.1) written in Hamiltonian form.

To write the Hamiltonian system more explicitly as an infinite dimensional system we introduce infinitely many coordinates $\mathbf{u}=(u_j)_{j\neq0},\mathbf{v}=(v_j)_{j\neq0}$ by writing

$$u=\mathscr{F}(\mathbf{u})\triangleq\sum_{j\neq0}\gamma_ju_je^{\mathrm{i}jx},\quad v=\mathscr{F}(\mathbf{v})\triangleq\sum_{j\neq0}\gamma_jv_je^{\mathrm{i}jx},\tag{4.3}$$

where

$$\gamma_j=\sqrt{|j|}$$

are fixed positive weights. The sequence $(\mathbf{u},\mathbf{v})$ is an element of $\hbar_p\times\hbar_p$ (see Section 1 for the definition of $\hbar_p$). Due to the choice of the weights, we have an isomorphism $\mathscr{F}:\hbar_{Z+1/2}\to\mathscr{H}_0^Z$ for each $Z\geq1$. The complex space $\hbar_p$ is canonically identified with the real space h_p by setting

$$w_j=(x_j-\mathrm{i}y_j)/\sqrt{2},\quad w_{-j}=\bar{w}_j,\quad j\geq1.$$

The minus sign in the definition of w_j is chosen so that $\mathrm{d}w_j\wedge\mathrm{d}w_{-j}=\mathrm{i}\mathrm{d}x_j\wedge\mathrm{d}y_j$. A function on $\hbar_p$ is said to be real analytic, if with this identification it is real analytic in x_j and y_j in the usual sense. The complexification of $\hbar_p$ is the same space of sequences, but with the condition $w_{-j}=\bar{w}_j$ dropped.

The Hamiltonian expressed in the new coordinates $\mathbf{u},\mathbf{v}$ is determined by inserting the expansion (4.3) of u,v into the definition (4.2) of H. Using for simplicity the same symbol for the Hamiltonian as a function of u,v we obtain

$$H(\mathbf{u},\mathbf{v})=\Lambda(\mathbf{u},\mathbf{v})+L(\mathbf{u},\mathbf{v})+\varepsilon K(\mathbf{u},\mathbf{v})\tag{4.4}$$

with

$$\Lambda(\mathbf{u},\mathbf{v})=2\pi\sum_{j\geq1}\gamma_j^6(|u_j|^2+|v_j|^2),\quad L(\mathbf{u},\mathbf{v})=2\pi\sum_{k,l,m\neq0,k+l+m=0}\gamma_k\gamma_l\gamma_m(u_ku_lu_m+v_kv_lv_m),\tag{4.5}$$

and

$$K(\mathbf{u},\mathbf{v}) = \int_{\mathbb{T}} f(x, \sum_{j\neq 0} \gamma_j u_j e^{\mathrm{i}jx}, \sum_{j\neq 0} \gamma_j v_j e^{\mathrm{i}jx})\mathrm{d}x. \tag{4.6}$$

The phase space $\hbar_p \times \hbar_p$ is endowed with the Poisson structure

$$\{F,G\} = \mathrm{i}\sum_{j\neq 0} \sigma_j (\frac{\partial F}{\partial u_j}\frac{\partial G}{\partial u_{-j}} + \frac{\partial F}{\partial v_j}\frac{\partial G}{\partial v_{-j}}),$$

where σ_j is the sign of j, and the equations of motion in the new coordinates are given by

$$\begin{cases} \dot{u}_j = \mathrm{i}\sigma_j \dfrac{\partial H}{\partial u_{-j}}, & j \neq 0, \\ \dot{v}_j = \mathrm{i}\sigma_j \dfrac{\partial H}{\partial v_{-j}}, & j \neq 0. \end{cases}$$

Since the transformed Poisson structure is nondegenerate, it also defines a symplectic structure

$$-\mathrm{i}\sum_{j\geq 1} (\mathrm{d}u_j \wedge \mathrm{d}u_{-j} + \mathrm{d}v_j \wedge \mathrm{d}v_{-j})$$

on $\hbar_p \times \hbar_p$, according to which the above equations of motion are the usual Hamiltonian equations with Hamiltonian. The associated Hamiltonian vector field with Hamiltonian H is given by

$$X_H = \mathrm{i}\sum_{j\neq 0} \sigma_j (\frac{\partial H}{\partial u_{-j}}\frac{\partial}{\partial u_j} + \frac{\partial H}{\partial v_{-j}}\frac{\partial}{\partial v_j}).$$

We study (4.1) as a perturbed system of the integrable Hamiltonian system

$$\begin{cases} u_t = -u_{xxx} + 6uu_x, \\ v_t = -v_{xxx} + 6vv_x, \end{cases} \tag{4.7}$$

and the Hamiltonian corresponding to (4.3) is

$$H_0(u,v) = \int_{\mathbb{T}} (\frac{1}{2}u_x^2 + \frac{1}{2}v_x^2 + u^3 + v^3)\mathrm{d}x. \tag{4.8}$$

Actually, we can regard the two KdV equations $u_t = -u_{xxx} + 6uu_x$ and $v_t = -v_{xxx} + 6vv_x$ in (4.1) as independent with each other. Thus similar to the discussion for the KdV equation, we have the following observations. The vector field of the quadratic Hamiltonian Λ takes values in $\hbar_{p-3} \times \hbar_{p-3}$ for $(\mathbf{u},\mathbf{v})$ in $\hbar_p \times \hbar_p$, hence it is unbounded of order 3. Strictly speaking, it is not a genuine vector field. The vector field of the cubic Hamiltonian L is also unbounded, but only of order 1. More precisely, we have the following regularity property of X_L.

Lemma 4.1 *The Hamiltonian vector field X_L is real analytic as a map from $\hbar_p \times \hbar_p$ into $\hbar_{p-1} \times \hbar_{p-1}$ for each $p \geq \frac{3}{2}$. Moreover,*

$$||X_L||_{p-1} = O(||\boldsymbol{u}||_p^2) + O(||\boldsymbol{v}||_p^2).$$

Proof: It is similar to the proof of Lemma 14.1 in [8]. We omit it here. □

Theorem 4.2 (**The Birkhoff normal form**) *There exists a real analytic coordinate transformation*

$$\Psi : \begin{cases} \boldsymbol{u} = \Phi(\boldsymbol{w}_1) \\ \boldsymbol{v} = \Phi(\boldsymbol{w}_2) \end{cases} \tag{4.9}$$

defined in a neighborhood of the origin in $\hbar_{3/2}\times\hbar_{3/2}$ which transforms Hamiltonian H, into its Birkhoff normal form up to order four. More precisely,

$$H_0\circ\Psi=\Lambda-B+M \tag{4.10}$$

with

$$B=6\pi\sum_{j\geq 1}(|w_{1j}|^4+|w_{2j}|^4),\quad ||X_M||_{1/2}=O(||\boldsymbol{w}_1||_{3/2}^4)+O(||\boldsymbol{w}_2||_{3/2}^4). \tag{4.11}$$

Moreover, for each $r\geq 3/2$, the restriction of Φ to some neighborhood of the origin in $\hbar_p$ defines a similar coordinate transformation in $\hbar_p$, so that

$$||X_M||_{p-1}=O(||\boldsymbol{w}_1||_p^4)+O(||\boldsymbol{w}_2||_p^4). \tag{4.12}$$

Proof: Choose Φ just be the real analytic symplectic coordinate transformation in Theorem 14.2, then we can get the result. □

Now we consider the transformed Hamiltonian

$$\begin{aligned}H\circ\Psi &= H_0\circ\Psi+\varepsilon K\circ\Psi\\ &= \Lambda-B+M+\varepsilon R\\ &= \Lambda(|\mathbf{w}_1|^2,|\mathbf{w}_2|^2)-B(|\mathbf{w}_1|^2,|\mathbf{w}_2|^2)+M(|\mathbf{w}_1|^2,|\mathbf{w}_2|^2)+\varepsilon R(\mathbf{w}_1,\mathbf{w}_2),\end{aligned}$$

which is real analytic on $U\subset\hbar_p\times\hbar_p$.

The perturbing Hamiltonian vector field in the origin phase space is

$$X_K=(\frac{\mathrm{d}}{\mathrm{d}x}\frac{\partial K}{\partial u},\frac{\mathrm{d}}{\mathrm{d}x}\frac{\partial K}{\partial v}).$$

It is defined on V and of of order 1. Since Ψ is a symplectic diffeomorphism of the two Hilbert scales $(\hbar_{Z+\frac{1}{2}}\times\hbar_{Z+\frac{1}{2}})_{Z\geq 1}$ and $(\mathscr{H}_0^Z\times\mathscr{H}_0^Z)_{Z\geq 1}$, the vector field of the transformed Hamiltonian $R=K\circ\Psi$ is

$$||X_R||_{Z-\frac{1}{2};U}\leq C||X_K||_{Z-1;V}\leq C||(\frac{\partial K}{\partial u},\frac{\partial K}{\partial v})||_{Z;V}\leq C.$$

As a second step, We introduce symplectic polar and real coordinates by setting

$$\begin{cases}w_{11}=\sqrt{I_1+y}e^{-\mathrm{i}x},\\ w_{21}=z_1,\\ (w_{1j},w_{2j})=z_j,\quad j\geq 2,\\ (\bar{w}_{1j},\bar{w}_{2j})=\bar{z}_j,\quad j\geq 2.\end{cases} \tag{4.13}$$

Setting $\mathbf{I}_0=(\xi,0)$ the integrable Hamiltonian in (4.10) in the new coordinates is, up to a constant depending only on ξ, given by

$$H=N+Q=N(y,z,\bar{z};\xi)+Q(y,z,\bar{z};\xi), \tag{4.14}$$

where

$$N=\omega_1(\xi)y+\Omega_1(\xi)z_1+\sum_{n\geq 2}\Omega_n(\xi)\langle z_n,\bar{z}_n\rangle,$$

$$\Omega_j(\xi)=\omega_j(\xi),\quad j\geq 2.$$

And with $I_1 = y$ and $I_j = \langle z_j, \bar{z}_j \rangle$, $j \geq 2$, $Q = \sum_{i,j\geq 1} Q_{ij}(\xi, I) I_i I_j$ with $|\sum_{j\geq 1} Q_{ij}(\mathbf{I}_0, \mathbf{I}) I_j| \leq ci||\mathbf{I}||_{\ell^1_{2Z+1}}$ (see (17.3) in [4,pp140] for details). We check Assumptions A-E of the KAM Theorem 1.2 for this normal form. Its external frequencies Ω_j may be written as

$$\Omega_j(\xi) = \bar{\Omega}_j + \tilde{\Omega}_j(\xi),$$

with $\bar{\Omega}_j = \Omega_j(0)$ and $\tilde{\Omega}_j(\xi) = \Omega_j(\xi) - \Omega_j(0)$. We have

$$\bar{\Omega}_j = 8\pi^3 j^3,$$

and a map

$$\tilde{\Omega} = (\tilde{\Omega}_j)_{j\geq 1} : \Pi \to \ell^\infty_{-1},$$

where ℓ^∞_β is the Banach space of all real sequences $z = (z_1, z_2, \cdots)$ with $|z|_\beta = \sup_{j\geq 1} j^\beta |z_j| < \infty$. So Assumption A and B are satisfied with $d = 3$, $\bar{d} = 2$ and $\delta = 1$. Assumptions C is satisfied as we have $\omega(\xi) = 8\pi^3 - 6\xi + \cdots$ where the dots stand for the higher terms in ξ.

This coordinates transformation is real analytic and symplectic on the phase space

$$\begin{aligned} D(s,r) &= \{|\mathrm{Im}x| < s\} \times \{|y| < r^2\} \times \{||z||_p + ||\bar{z}||_p < r\} \\ &\subset \mathbb{C} \times \mathbb{C} \times (\hbar_p \times \hbar_p) \times (\hbar_p \times \hbar_p) \\ &= \mathscr{P}_{p,\mathbb{C}}, \end{aligned}$$

for all $s > 0$, and $r > 0$ sufficiently small, where $p = Z + \frac{1}{2}$. In the following we may fix such an s arbitrarily, while we keep the freedom to choose r smaller. And the parameter domain

$$\Xi_r^- = U_{-4r^2}\Xi_r, \quad \Xi_r = \{\xi : 0 < \xi < r\},$$

where $U_{-\rho}\Xi$ is the subset of all points in Ξ with boundary distance greater than ρ. The total Hamiltonian H is well-defined on these domains, and we have $||X_P||_{r,p-1;D(s,r)\times\Pi} = O(r^2)$. Thus Assumption D is satisfied with $q = p - 1$. Furthermore, by we have $\overline{K(\mathbf{u},\mathbf{v})} = K(\bar{\mathbf{u}}, \bar{\mathbf{v}}) = K(\mathbf{u},\mathbf{v})$, also as Φ is real analytic, we have

$$\begin{cases} \bar{\mathbf{u}} = \overline{\Phi(\mathbf{w}_1)} = \Phi(\bar{\mathbf{w}}_1), \\ \bar{\mathbf{v}} = \overline{\Phi(\mathbf{w}_2)} = \Phi(\bar{\mathbf{w}}_2), \end{cases} \tag{4.15}$$

Thus

$$\overline{P(\mathbf{w}_1, \mathbf{w}_2)} = P(\mathbf{w}_1, \mathbf{w}_2). \tag{4.16}$$

Denote $D_{\mathfrak{R}}(s,r) = \{(x,y,z,\bar{z}) \in D(s,r) : x, y \in \mathbb{R}\}$. By the coordinates (4.13) and (4.16), we have for

$$\overline{P(x,y,z,\bar{z})} = P(x,y,z,\bar{z}), \quad (x,y,z,\bar{z}) \in D_{\mathfrak{R}}(s,r). \tag{4.17}$$

In other words, Assumption E is also satisfied. Then we can use Theorem 1.2 to get Theorem 4.1 and thus Theorem 1.3. □

5 Technical Lemmas

Lemma 5.1 *Consider a $n \times n$ complex matrix function $Y(\xi)$ which depends on the real parameter $\xi \in \mathbb{R}$, and $Y(\xi)$ satisfies the following conditions:*
(1) $Y(\xi)$ is self-adjoint for every $\xi \in \mathbb{R}$, i.e. $Y(\xi) = (Y(\xi))^*$, where star denotes the conjugate transpose of a matrix;
(2) $Y(\xi)$ is continuously differentiable in an interval I of the real variable ξ.

Then there exist n continuously differentiable functions $\mu_1(\xi),\cdots,\mu_n(\xi)$ on I that represent the repeated eigenvalues of $Y(\xi)$.

Proof: See page122-124 of [9]. $\square$

Lemma 5.2 *Assume $Y=Y(\xi)$ satisfies the conditions in Lemma 5.1. Let $\mu=\mu(\xi)$ be any eigenvalue of Y and ϕ be the normalized eigenfunction corresponding to μ. Then*

$$\partial_\xi \mu = ((\partial_\xi Y)\phi,\phi).$$

Proof. See page 125 of [9]. $\square$

Lemma 5.3 *For $\sigma>0$, $\nu>0$, the following inequalities hold true:*

$$\sum_{k\in\mathbb{Z}^n} e^{-2|k|\sigma} \le \frac{(1+e)^n}{\sigma^n}, \tag{5.1}$$

$$\sum_{k\in\mathbb{Z}^n} e^{-2|k|\sigma}|k|^\nu \le (\frac{\nu}{e})^\nu \frac{1}{\sigma^{\nu+n}}(1+e)^n. \tag{5.2}$$

Proof: See page 22 of [1]. $\square$

Lemma 5.4 *Let $U_j, j\ge 1$ be complex n-matrix functions on $\mathbb{T}$ that their elements are real analytic on $\mathbb{T}_s=\{x\in\mathbb{C}/2\pi\mathbb{Z}: |Imx|<s\}$. Then*

$$(\sum_{j\ge1}\sup_{x\in\mathbb{T}_{s-\sigma}}|U_j(x)|^2_{\ell^\infty\to\ell^\infty})^{\frac12} \le \frac{4n}{\sigma}\sup_{x\in\mathbb{T}_s}(\sum_{j\ge1}|U_j(x)|^2_{\ell^\infty\to\ell^\infty})^{\frac12}. \tag{5.3}$$

Proof: Note that for complex functions $u_j, j\ge1$ defined on $\mathbb{T}$ that are real analytic on $\mathbb{T}_s$, we have

$$(\sum_{j\ge1}\sup_{x\in\mathbb{T}_{s-\sigma}}|u_j(x)|^2)^{\frac12} \le \frac{4}{\sigma}\sup_{x\in\mathbb{T}_s}(\sum_{j\ge1}|u_j(x)|^2)^{\frac12}. \tag{5.4}$$

Let $U_j(x)=(U_j^{kl}(x): 1\le k,l\le n)$, then

$$\begin{aligned}
(\sum_{j\ge1}\sup_{x\in\mathbb{T}_{s-\sigma}}|U_j(x)|^2_{\ell^\infty\to\ell^\infty})^{\frac12} &= (\sum_{j\ge1}\sup_{x\in\mathbb{T}_{s-\sigma}}(\max_{1\le k\le n}\sum_{l=1}^n|U_j^{kl}(x)|)^2)^{\frac12}\\
&\le (\sum_{j\ge1}\sup_{x\in\mathbb{T}_{s-\sigma}} n^2\max_{1\le k\le n}\sum_{l=1}^n|U_j^{kl}(x)|^2)^{\frac12}\\
&= n(\sum_{j\ge1}\sup_{x\in\mathbb{T}_{s-\sigma}}\max_{1\le k\le n}\sum_{l=1}^n|U_j^{kl}(x)|^2)^{\frac12}\\
&= n(\sum_{j\ge1}\sup_{x\in\mathbb{T}_{s-\sigma}}\sum_{l=1}^n|U_j^{k_jl}(x)|^2)^{\frac12} \quad k_j\in\{1,\cdots,n\}\\
&\le n(\sum_{j\ge1}\sum_{l=1}^n\sup_{x\in\mathbb{T}_{s-\sigma}}|U_j^{k_jl}(x)|^2)^{\frac12}\\
&= (\sum_{j\ge1,1\le l\le n}\sup_{x\in\mathbb{T}_{s-\sigma}}|U_j^{k_jl}(x)|^2)^{\frac12}.
\end{aligned} \tag{5.5}$$

By (5.4),we have the above

$$
\begin{aligned}
(5.5)\quad &\leq \frac{4n}{\sigma}\sup_{x\in\mathbb{T}_s}\Big(\sum_{j\geq 1,1\leq l\leq n}|U_j^{k_jl}(x)|\Big)^{\frac{1}{2}}\\
&\leq \frac{4n}{\sigma}\sup_{x\in\mathbb{T}_s}\Big(\sum_{j\geq 1}\sum_{l=1}^{n}|U_j^{k_jl}(x)|^2\Big)^{\frac{1}{2}}\\
&\leq \frac{4n}{\sigma}\sup_{x\in\mathbb{T}_s}\Big(\sum_{j=1}\max_{1\leq k\leq n}\sum_{l=1}^{n}|U_j^{k_jl}(x)|^2\Big)^{\frac{1}{2}}\\
&\leq \frac{4n}{\sigma}\sup_{x\in\mathbb{T}_s}\Big(\sum_{j\geq 1}|U_j(x)|^2_{\ell^\infty\to\ell^\infty}\Big)^{\frac{1}{2}}.
\end{aligned}
$$

Lemma 5.5 *Let $A=(A_{ij})_{i,j\geq 1}$ be a bounded operator on ℓ^2 which depends on $x\in\mathbb{T}$ such that all coefficients are analytic on $\mathbb{T}_s$, here A_{ij} is a $n\times n$-matrix. Suppose $B=(B_{ij})_{i,j\geq 1}$ is another operator on ℓ^2 depending on x whose coefficients satisfy*

$$
\sup_{x\in\mathbb{T}_s}|B_{ij}(x)|_{\ell^2\to\ell^2}\leq\frac{1}{|i-j|}\sup_{x\in\mathbb{T}_s}|A_{ij}(x)|_{\ell^2\to\ell^2},\qquad i\neq j,
$$

and $B_{jj}=0$ for $j\geq 1$. Then B is a bounded operator on ℓ^2 for every $x\in\mathbb{T}_s$, and

$$
\sup_{x\in\mathbb{T}_{s-\sigma}}||B(x)||_{\ell^2\to\ell^2}\leq\frac{16n^{\frac{3}{2}}}{\sigma}\sup_{x\in\mathbb{T}_s}||A(x)||_{\ell^2\to\ell^2},
$$

for $0<\sigma\leq s\leq 1$.

Proof: For $x\in\mathbb{T}_{s-\sigma}$, we have by Schawarz inequality and Lemma 5.3

$$
\begin{aligned}
\sum_{j\geq 1}|B_{ij}(x)|_{\ell^2\to\ell^2} &\leq \sum_{j\geq 1}\sup_{x\in\mathbb{T}_{s-\sigma}}|B_{ij}(x)|_{\ell^2\to\ell^2}\\
&\leq \Big(\sum_{j\geq 1}\sup_{x\in\mathbb{T}_{s-\sigma}}|A_{ij}(x)|^2_{\ell^2\to\ell^2}\Big)^{\frac{1}{2}}\Big(\sum_{j\neq i}\frac{1}{|i-j|^2}\Big)^{\frac{1}{2}}\\
&\leq \sqrt{n}\sup_{x\in\mathbb{T}_{s-\sigma}}|A_{ij}(x)|^2_{\ell^\infty\to\ell^\infty}\Big)^{\frac{1}{2}}\Big(\sum_{j\neq i}\frac{1}{|i-j|^2}\Big)^{\frac{1}{2}}\\
&\leq \frac{16n^{\frac{3}{2}}}{\sigma}\sup_{x\in\mathbb{T}_s}\Big(\sum_{j\geq 1}|A_{ij}(x)|^2_{\ell^\infty\to\ell^\infty}\Big)^{\frac{1}{2}}\\
&\leq \frac{16n^{\frac{3}{2}}}{\sigma}\sup_{x\in\mathbb{T}_s}\Big(\sum_{j\geq 1}|A_{ij}(x)|^2_{\ell^2\to\ell^2}\Big)^{\frac{1}{2}}\\
&\leq \frac{16n^{\frac{3}{2}}}{\sigma}\sup_{x\in\mathbb{T}_s}||A(x)||_{\ell^2\to\ell^2}.
\end{aligned}
$$

The same estimates applies to $\sum_{i\geq 1}|B_{ij}(x)|_{\ell^2\to\ell^2}$. Hence, for $x\in\mathbb{T}_{s-\sigma}, v=(v_j)_{j\geq 1}\in\ell^2$, here $v_j\in\mathbb{C}^n$,

$$\begin{aligned}
||B(x)v||_{\ell^2}^2 &\leq \sum_{i\geq 1}(\sum_{j\geq 1}|B_{ij}(x)|_{\ell^2\to\ell^2}|v_j|_{\ell^2})^2\\
&\leq \sum_{i\geq 1}(\sum_{j\geq 1}|B_{ij}(x)|_{\ell^2\to\ell^2})(\sum_{j\geq 1}|B_{ij}(x)|_{\ell^2\to\ell^2}|v_j|_{\ell^2}^2)\\
&\leq (\sup_i\sum_{j\geq 1}|B_{ij}(x)|_{\ell^2\to\ell^2})(\sup_j\sum_{i\geq 1}|B_{ij}(x)|_{\ell^2\to\ell^2})(\sum_{j\geq 1}|v_j|_{\ell^2}^2)\\
&\leq (\frac{16n^{\frac{3}{2}}}{\sigma}\sup_{x\in\mathbb{T}_s}||A(x)||_{\ell^2\to\ell^2})^2||v||_{\ell^2}^2.
\end{aligned}$$

From this the final estimate follows. □

Lemma 5.6 *Let A, B, C be respectively $n\times n$, $m\times m$, $n\times m$ matrices, and let X be an $n\times m$ unknown matrix. The matrix equation*

$$AX-XB=C \tag{5.6}$$

is solvable if and only if

$$(E_m\otimes A-B\otimes E_n)X'=C'$$

is solvable. Here if we denote $A=(A_1,\cdots,A_n)$, $C=(C_1,\cdots,C_m)$, $X=(X_1,\cdots,X_m)$, then $X'=(X_1^T,\cdots,X_m^T)^T$, $C'=(C_1^T,\cdots,C_m^T)^T$.

Proof: We can find the proof of this lemma on page 526 in [14]. □

Lemma 5.7 *Let $A=(A(k-l):k,l\in\mathbb{Z})$, where $A(k-l)=(A_{ij}(k-l):1\leq i,j\leq n)$. Then*

$$||A||_{\ell^2\to\ell^2}\leq n^{\frac{1}{4}}\max_{l\in\mathbb{Z}}\sum_{k\in\mathbb{Z}}|A(k-l)|_{\ell^2\to\ell^2}\leq n^{\frac{1}{4}}\sum_{k\in\mathbb{Z}}|A(k)|_{\ell^2\to\ell^2}. \tag{5.7}$$

Proof: By Riesz interpolation theorem, we have

$$\begin{aligned}
||A||_{\ell^2\to\ell^2} &\leq \sqrt{||A||_{\ell^1\to\ell^1}}\sqrt{||A||_{\ell^\infty\to\ell^\infty}}\\
&= \sqrt{\max_{1\leq j\leq n,l\in\mathbb{Z}}\sum_{1\leq i\leq n,k\in\mathbb{Z}}|A_{ij}(k-l)|}\sqrt{\max_{1\leq i\leq n,k\in\mathbb{Z}}\sum_{1\leq j\leq n,l\in\mathbb{Z}}|A_{ij}(k-l)|}\\
&\leq \sqrt{\max_{l\in\mathbb{Z}}\sum_{k\in\mathbb{Z}}(\max_{1\leq j\leq n}\sum_{1\leq i\leq n}|A_{ij}(k-l)|)}\sqrt{\max_{k\in\mathbb{Z}}\sum_{l\in\mathbb{Z}}(\max_{1\leq i\leq n}\sum_{1\leq j\leq n}|A_{ij}(k-l)|)}\\
&= \sqrt{\max_{l\in\mathbb{Z}}\sum_{k\in\mathbb{Z}}|A(k-l)|_{\ell^1\to\ell^1}}\sqrt{\max_{k\in\mathbb{Z}}\sum_{l\in\mathbb{Z}}|A(k-l)|_{\ell^\infty\to\ell^\infty}}\\
&\leq \sqrt{\max_{l\in\mathbb{Z}}\sum_{k\in\mathbb{Z}}\sqrt{n}|A(k-l)|_{\ell^2\to\ell^2}}\sqrt{\max_{k\in\mathbb{Z}}\sum_{l\in\mathbb{Z}}|A(k-l)|_{\ell^2\to\ell^2}}\\
&= n^{\frac{1}{4}}\sqrt{\max_{l\in\mathbb{Z}}\sum_{k\in\mathbb{Z}}|A(k-l)|_{\ell^2\to\ell^2}}\sqrt{\max_{l\in\mathbb{Z}}\sum_{k\in\mathbb{Z}}|A(k-l)|_{\ell^2\to\ell^2}}\\
&\leq n^{\frac{1}{4}}\max_{l\in\mathbb{Z}}\sum_{k\in\mathbb{Z}}|A(k-l)|_{\ell^2\to\ell^2}\leq n^{\frac{1}{4}}\sum_{k\in\mathbb{Z}}|A(k)|_{\ell^2\to\ell^2}.
\end{aligned}$$

References

[1] N. N. Bogolyubov, Yu. A. Mitropolskii, A. M. Samoilemko, Methods of Accelerated Convergence in Nonlinear Mechanics, Springer-Verlag, New York (1976) (Russian Original: Naukova Dumka, Kiev, 1969).

[2] D. Bambusi, S. Graffi, Time quasi-periodic unbounded perturbations of Schrödinger operators and KAM methods. Commun. Math. Phys. 219, 465-480 (2001).

[3] D. Bambusi, S. Paleari, Families of periodic oribits for resonant PDE's. J. Nonlinear Science 11, 69-87 (2001).

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

[5] J. Bourgain, Construction of quasi-periodic solutions for Hamiltonian perturbatins of linear equations and application to nonlinear pde. Int. Math. Research Notices 11, 475-497 (1994).

[6] L. Chierchia, J. You, KAM tori for 1D nonlinear wave equations with periodic boundary conditions, Comm. Math. Phys. 211, 497-525 (2000) .

[7] L. H. Eliasson and S. B. Kuksin, KAM for non-linear Schrödinger equation. Preprint, 2006.

[8] T. Kappeler, J. Pöschel, KdV&KAM, Springer-Verlag Berlin Heidelberg 2003.

[9] T. Kato, Perturbation Theory for Linear Operators, (Corrected printing of the second edition) Springer, Berlin Heidelberg New York (1980).

[10] S. B. Kuksin, Nearly integrable infinite-dimensional Hamiltonian systems. Lecture Notes in Mathematics 1556, Springer, 1993.

[11] S. B. Kuksin, On small-denominators equations with large variable coefficients, J. Appl. Math. Phys. (ZAMP) 48, 262-271 (1997).

[12] S. B. Kusin, Analysis of Hamiltonian PDEs. Oxford University Press. Oxford, 2000.

[13] J. Liu, X. Yuan, Spectrum for quantum Duffing oscillator and small-divisor equation with large variable coefficient, Communcuations on Pure and Applied Mathematics, to appear.

[14] P. Lancaster, Theory of Matrices, New York and London: Academic Press LTD, 1969.

[15] J. Pöschel, A KAM-Theorem for some Nonlinear Partial Differential Equations, Ann. Sc. Norm. Sup. Pisa 23, 119-148 (1996).

[16] H. Rüssmann, Invariant tori in non-degenerate nearly integrable Hamiltonian systems. Regul. Chaotic Dyn. 6, 119-204 (2001).

[17] P.R., Halmos, V.S., Sunder, Bounded Integral Operators on L^2 Space. Springer-Verlag Berlin Heidelberg New York 1978.

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

[19] K. Yosida, Functional Analysis. Berlin Heidelberg New York: Springer-Verlag (1980); reprinted in China by Beijing World Punlishing corporation, 1999.

[20] X. Yuan, A KAM Theorem with Applications to Partial Differential Equations of Higher Dimensions, Commun. Math. Phys. 275, 97-137 (2007) .

[21] X. Yuan, Quasi-periodic solutions of completely resonant nonlinear wave equations. J. Diff. Eq. 230, 213-274 (2006).

[22] X. Yuan, Invariant tori of nonlinear wave equations with a given potential. Discrete and continuous dynamical systems 16 (3), 615-634 (2006).

[23] X. Yuan, Invariant manifold of hyperbolic-type for nonlinear wave equation. Int, J. Math. Science 18, 1111-1136 (2003).

[24] W. P. Ziemer, Weakly differentiable functions, (Graduate Texts in Math. No.120) Springer, Berlin Heidelberg New York (1989), reprinted in China by Beijing World Publishing Corporation (1999).