A KAM Theorem for Hamiltonian Partial Differential Equations with Unbounded Perturbations *

Jianjun Liu,   Xiaoping Yuan

School of Mathematical Sciences and Key Lab of Math. for Nonlinear Science,
Fudan University, Shanghai 200433, P R China

September 30, 2010

Abstract: We establish an abstract infinite dimensional KAM theorem dealing with unbounded perturbation vector-field, which could be applied to a large class of Hamiltonian PDEs containing the derivative ∂_x in the perturbation. Especially, in this range of application lie a class of derivative nonlinear Schrödinger equations with Dirichlet boundary conditions and perturbed Benjamin-Ono equation with periodic boundary conditions, so KAM tori and thus quasi-periodic solutions are obtained for them.

1   Introduction and Main Results

Consider a Hamiltonian partial differential equation (HPDE)

$$\dot{w} = Aw + F(w),$$

where Aw is linear Hamiltonian vector-field with $d := \mathrm{ord}\, A > 0$ and $F(w)$ is nonlinear Hamiltonian vector-field with $\tilde{d} := \mathrm{ord}\, F$ and is analytic in the neighborhood of the origin $w = 0$.

When $\tilde{d} = \mathrm{ord}\, F \leq 0$, the vector-field F is called bounded perturbation. For example, in this case lie a class of nonlinear Schrödinger equations (with $\tilde{d} = 0$)

$$\mathbf{i}u_t + u_{xx} + V(x)u + |u|^2 u = 0$$

and a class of nonlinear wave equations (with $\tilde{d} = -1$)

$$u_{tt} - u_{xx} + V(x)u + u^3 = 0.$$

For the existence of KAM tori of the PDEs with bounded perturbations has been deeply and widely investigated by many authors. In this field of study there are too many references to list here. We give just two survey papers by Kuksin [8] and Bourgain [4].

When $\tilde{d} = \mathrm{ord}\, F > 0$, the vector-field F is called *unbounded* perturbation. According to a well-known example, due to Lax [11] and Klainerman [5] (See also [7]), it is reasonable to assume

$$\tilde{d} \leq d - 1$$

in order to guarantee the existence of KAM tori for the PDE. The quantity $d - \tilde{d}$ measures the strength of nonlinearity of the PDE. The smaller the $d - \tilde{d}$ is, the stronger the nonlinearity is. When $d - \tilde{d} = 1$, the nonlinearity of the PDE is the strongest.

1

For the PDE with unbounded Hamiltonian perturbation, the only previous KAM theorem is due to Kuksin [7] where it is assumed that $d - \tilde{d} > 1$. Kuksin's theorem is in [7] used to prove the persistence of the finite-gap solutions of KdV equation, as well as its hierarchy, subject to periodic boundary conditions. See also Kappeler-Pöschel [10]. Another KAM theorem with unbounded linear Hamiltonian perturbation is due to Bambusi-Graffi [2] where the spectrum property is investigated for the time dependent linear Schrödinger equation

$$\mathbf{i}\dot{y}(t) = (A + \varepsilon F(t))y(t)$$

with $d - \tilde{d} > 1$ also.

The assumption $d - \tilde{d} > 1$ excludes a large class of interesting partial differential equations such as a class of derivative nonlinear Schrödinger (DNLS) equations

$$\mathbf{i}u_t + u_{xx} - M_\sigma u + \mathbf{i}f(u, \bar{u})u_x = 0$$

subject to Dirichlet boundary conditions, and perturbed Benjamin-Ono (BO) equation

$$u_t + \mathscr{H}u_{xx} - uu_x + \text{perturbation} = 0$$

subject to periodic boundary conditions, where $\mathscr{H}$ is the Hilbert transform. In both cases, $d - \tilde{d} = 1$. In the present paper, we will construct a KAM theorem including $d - \tilde{d} > 1$ and limiting case $d - \tilde{d} = 1$ and show the existence of KAM tori for DNLS and perturbed BO equations.

If the Hamiltonian operator A has continuous spectra, there is usually no KAM tori for the HPDE. Therefore, we should assume that the Hamiltonian operator A is of pure point spectra before stating our theorems. Taking the eigenfunctions of A as a basis and changing partial coordinates into action-angle variables, from the HPDE one can usually obtain a small perturbation $H = N + P$ of an infinite dimensional Hamiltonian in the parameter dependent normal form

$$N = \sum_{1 \le j \le n} \omega_j(\xi)y_j + \frac{1}{2}\sum_{j \ge 1} \Omega_j(\xi)(u_j^2 + v_j^2) \tag{1.1}$$

on a phase space

$$\mathscr{P}^{a,p} = \mathbb{T}^n \times \mathbb{R}^n \times \ell^{a,p} \times \ell^{a,p} \ni (x, y, u, v)$$

with symplectic structure $\sum_{1 \le j \le n} dy_j \wedge dx_j + \sum_{j \ge 1} du_j \wedge dv_j$, where $\ell^{a,p}$ is the Hilbert space of all real sequences $w = (w_1, w_2, \cdots)$ with

$$\|w\|_{a,p}^2 = \sum_{j \ge 1} e^{2aj} j^{2p} |w_j|^2 < \infty,$$

where $a \ge 0$ and $p \ge 0$. The tangential frequencies $\omega = (\omega_1, \cdots, \omega_n)$ and normal frequencies $\Omega = (\Omega_1, \Omega_2, \cdots)$ are real vectors depending on parameters $\xi \in \Pi \subset \mathbb{R}^n$, Π a closed bounded set of positive Lebesgue measure, and roughly

$$\Omega_j(\xi) = j^d + \cdots.$$

The perturbation term P is real analytic in the space coordinates and Lipschitz in the parameters, and for each $\xi \in \Pi$ its Hamiltonian vector field $X_P = (P_y, -P_x, -P_v, P_u)^T$ defines near $\mathscr{T}_0 := \mathbb{T}^n \times \{y = 0\} \times \{u = 0\} \times \{v = 0\}$ a real analytic map

$$X_P : \quad \mathscr{P}^{a,p} \to \mathscr{P}^{a,q},$$

where

$$p - q = \tilde{d}.$$

2

In the whole of this paper the parameter a is fixed. Moreover, throughout this paper, for convenience, we will adopt lots of notations and definitions from [10].

We denote by $\mathscr{P}_{\mathbb{C}}^{a,p}$ the complexification of $\mathscr{P}^{a,p}$. For $s, r > 0$, We introduce the complex $\mathscr{T}_0$-neighborhoods in $\mathscr{P}_{\mathbb{C}}^{a,p}$

$$D(s,r): \quad |\mathrm{Im}x| < s, \quad |y| < r^2, \quad \|u\|_{a,p} + \|v\|_{a,p} < r, \tag{1.2}$$

and weighted norm for $W = (X, Y, U, V) \in \mathscr{P}_{\mathbb{C}}^{a,q}$

$$\|W\|_{r,a,q} = |X| + \frac{|Y|}{r^2} + \frac{\|U\|_{a,q}}{r} + \frac{\|V\|_{a,q}}{r},$$

where $|\cdot|$ denotes the sup-norm for complex vectors. Furthermore, for a map $W : D(s,r) \times \Pi \to \mathscr{P}_{\mathbb{C}}^{a,q}$, for example, the Hamiltonian vector field X_P, we define the norms

$$\|W\|_{r,a,q,D(s,r)\times\Pi} = \sup_{D(s,r)\times\Pi} \|W\|_{r,a,q},$$

$$\|W\|_{r,a,q,D(s,r)\times\Pi}^{\mathrm{lip}} = \sup_{\xi,\zeta\in\Pi,\xi\neq\zeta} \sup_{D(s,r)} \frac{\|\Delta_{\xi\zeta}W\|_{r,a,q}}{|\xi - \zeta|},$$

where $\Delta_{\xi\zeta}W = W(\cdot;\xi) - W(\cdot;\zeta)$. In a completely analogous manner, the Lipschitz semi-norm of the frequencies ω and Ω are defined as

$$|\omega|_{\Pi}^{\mathrm{lip}} = \sup_{\xi,\zeta\in\Pi,\xi\neq\zeta} \frac{|\Delta_{\xi\zeta}\omega|}{|\xi - \zeta|}, \quad |\Omega|_{-\delta,\Pi}^{\mathrm{lip}} = \sup_{\xi,\zeta\in\Pi,\xi\neq\zeta} \sup_{j\geq 1} \frac{j^{-\delta}|\Delta_{\xi\zeta}\Omega_j|}{|\xi - \zeta|} \tag{1.3}$$

for any real number δ.

Theorem 1.1. *Suppose the normal form N described above satisfies the following assumptions:*

(A) *The map $\xi \mapsto \omega(\xi)$ between Π and its image is a homeomorphism which is Lipschitz continuous in both directions, i.e. there exist positive constants M_1 and L such that $|\omega|_{\Pi}^{\mathrm{lip}} \leq M_1$ and $|\omega^{-1}|_{\omega(\Pi)}^{\mathrm{lip}} \leq L$;*

(B) *There exists $d > 1$ such that*
$$|\Omega_i - \Omega_j| \geq m|i^d - j^d| \tag{1.4}$$
for all $i \neq j \geq 0$ uniformly on Π with some constant $m > 0$. Here $\Omega_0 = 0$;

(C) *There exists $\delta \leq d - 1$ such that the functions $\xi \mapsto \frac{\Omega_j(\xi)}{j^\delta}$ are uniformly Lipschitz on Π for $j \geq 1$, i.e. there exist a positive constant M_2 such that $|\Omega|_{-\delta,\Pi}^{\mathrm{lip}} \leq M_2$;*

(D) *We additionally assume*
$$4ELM_2 \leq m, \tag{1.5}$$
where $E = |\omega|_{\Pi} := \sup_{\xi\in\Pi} |\omega(\xi)|$.

Set $M = M_1 + M_2$. Then for every $\beta > 0$, there exists a positive constant γ, depending only on n, d, δ, m, the frequencies ω and Ω, $s > 0$ and β, such that for every perturbation term P described above with

$$\tilde{d} = p - q \leq d - 1 \tag{1.6}$$

and

$$\varepsilon := \|X_P\|_{r,a,q,D(s,r)\times\Pi} + \frac{\alpha}{M}\|X_P\|_{r,a,q,D(s,r)\times\Pi}^{\mathrm{lip}} \leq (\alpha\gamma)^{1+\beta} \tag{1.7}$$

for some $r > 0$ and $0 < \alpha < 1$, there exist

(1) a Cantor set $\Pi_\alpha \subset \Pi$ with

$$|\Pi \setminus \Pi_\alpha| \le c_1 \alpha, \tag{1.8}$$

where $|\cdot|$ denotes Lebesgue measure and $c_1 > 0$ is a constant depends on n, ω and Ω;

(2) a Lipschitz family of smooth torus embeddings $\Phi : \mathbb{T}^n \times \Pi_\alpha \to \mathscr{P}^{a,p}$ satisfying: for every non-negative integer multi-index $k = (k_1, \ldots, k_n)$,

$$\|\partial_x^k (\Phi - \Phi_0)\|_{r,a,p,\mathbb{T}^n \times \Pi_\alpha} + \frac{\alpha}{M} \|\partial_x^k (\Phi - \Phi_0)\|_{r,a,p,\mathbb{T}^n \times \Pi_\alpha}^{\text{lip}} \le c_2 \varepsilon^{\frac{1}{1+\beta}} / \alpha, \tag{1.9}$$

where $\partial_x^k := \frac{\partial^{|k|}}{\partial x_1^{k_1} \cdots \partial x_n^{k_n}}$ with $|k| := |k_1| + \cdots + |k_n|$,

$$\Phi_0 : \mathbb{T}^n \times \Pi \to \mathscr{T}_0, \quad (x, \xi) \mapsto (x, 0, 0, 0)$$

is the trivial embedding for each ξ, and c_2 is a positive constant which depends on k and the same parameters as γ;

(3) a Lipschitz map $\phi : \Pi_\alpha \to \mathbb{R}^n$ with

$$|\phi - \omega|_{\Pi_\alpha} + \frac{\alpha}{M} |\phi - \omega|_{\Pi_\alpha}^{\text{lip}} \le c_3 \varepsilon, \tag{1.10}$$

where c_3 is a positive constant which depends on the same parameters as γ,

such that for each $\xi \in \Pi_\alpha$ the map Φ restricted to $\mathbb{T}^n \times \{\xi\}$ is a smooth embedding of a rotational torus with frequencies $\phi(\xi)$ for the perturbed Hamiltonian H at ξ. In other words,

$$t \mapsto \Phi(\theta + t\phi(\xi), \xi), \quad t \in \mathbb{R}$$

is a smooth quasi-periodic solution for the Hamiltonian H evaluated at ξ for every $\theta \in \mathbb{T}^n$ and $\xi \in \Pi_\alpha$.

This theorem can be applied to the class of derivative nonlinear Schrödinger equations mentioned above. However, the assumption (1.5) excludes the perturbed Benjamin-Ono equation mentioned above. Thus, in the following, we give a modified version of the above theorem:

Theorem 1.2. *The above theorem also holds true with, respectively, replacing the assumption (D) and conclusion (1) by the assumption (D*) and conclusion (1*) below:*

(D*) For every $k \in \mathbb{Z}^n$ and $l \in \mathbb{Z}^\infty$ with $1 \le |l| \le 2$ (here $|l| = \sum_{j \ge 1} |l_j|$), the resonance set

$$\{\xi \in \Pi : \langle k, \omega(\xi) \rangle + \langle l, \Omega(\xi) \rangle = 0\}$$

has Lebesgue measure zero; Moreover, if $\delta = d - 1$, we additively assume that there exist $\delta_0 < d - 1$, a partition $\Omega = \Omega^1 + \Omega^2$, positive constant M_3 and M_4 with

$$8ELM_4 \le m, \tag{1.11}$$

such that $|\Omega^1|_{-\delta_0,\Pi}^{\text{lip}} \le M_3$, $|\Omega^2|_{-\delta,\Pi}^{\text{lip}} \le M_4$;

(1*) a Cantor set $\Pi_\alpha \subset \Pi$ with $|\Pi \setminus \Pi_\alpha| \to 0$ as $\alpha \to 0$, where $|\cdot|$ denotes Lebesgue measure.

This paper is organized as follows: In section 2 we give an outline of the proof of the above theorems, and some new difficult and ideas compared with [7] [10] are exhibited. In sections 3-6 the above theorems are proved in detail. The proof of Theorem 1.2 is the same as that of Theorem 1.1 except the measure estimate. Thus, sections 3-5 and subsection 6.1 are devoted to the proof of Theorem 1.1, while the measure estimate for Theorem 1.2 is given in subsection 6.2. In sections 7-8 Theorem 1.1 and Theorem 1.2 are applied to derivative nonlinear Schrödinger equations and perturbed Benjamin-Ono equation, respectively. Finally, a technical lemma is listed in section 9.

2 Outline of The Proof and more Remarks

The above theorems generalize Kukisn's theorem from $\tilde{d} < d-1$ to $\tilde{d} \leq d-1$ such that the range of application is extended to a class of derivative nonlinear Schrödinger equations and perturbed Benjamin-Ono equation. Here we would like to compare the proof of our theorems with that of Kuksin's theorem. By and large, as any KAM theorem, in both cases Newton iteration is used to overcome the notorious small divisor difficult. Therefore, our proof is mainly based on Kuksin's approach in [7]. (Also see [10]). There is, however, some essential differences between the proof of our theorems and that of Kuksin's theorem. In order to see clearly the differences, let us give the basic procedure of the proof of KAM theorem from [7] and [10], which consists of the following steps.

2.1. Derivation of the homological equations. For convenience, introduce complex variables $z = (u - \mathbf{i}v)/\sqrt{2}$ and $\bar{z} = (u + \mathbf{i}v)/\sqrt{2}$. Assume we are now in the v-th KAM iterative step. Write the integrable part N_v of the Hamiltonian H_v

$$N_v = \langle \omega, y \rangle + \sum_{j=1}^{\infty} \Omega_j z_j \bar{z}_j$$

and develop the perturbation P_v into Taylor series in $(y, z, \bar{z})$:

$$P_v = \varepsilon_v R_v + O(|y|^2 + |y| ||z||_{a,p} + ||z||_{a,p}^3),$$

where ε_v goes to zero very fast, for example, taking $\varepsilon_v \approx \varepsilon^{(5/4)^v}$, and

$$R_v = R^x(x) + \langle R^y, y \rangle + \langle R^z(x), z \rangle + \langle R^{\bar{z}}(x), \bar{z} \rangle + \langle R^{zz}(x)z, z \rangle + \langle R^{\bar{z}\bar{z}}(x)\bar{z}, \bar{z} \rangle + \langle R^{z\bar{z}}(x)z, \bar{z} \rangle = O(1).$$

A key point is, very roughly speaking, to search for a Hamiltonian function of the same form as R_v:

$$F_v = F^x(x) + \langle F^y, y \rangle + \langle F^z(x), z \rangle + \langle F^{\bar{z}}(x), \bar{z} \rangle + \langle F^{zz}(x)z, z \rangle + \langle F^{\bar{z}\bar{z}}(x)\bar{z}, \bar{z} \rangle + \langle F^{z\bar{z}}(x)z, \bar{z} \rangle$$

which satisfies

$$\{N_v, F_v\} + R_v = 0, \tag{2.1}$$

where $\{\cdot, \cdot\}$ is the Poisson bracket with respect to the symplectic structure

$$\sum_{1 \leq j \leq n} dy_j \wedge dx_j - \mathbf{i} \sum_{j \geq 1} dz_j \wedge d\bar{z}_j.$$

By $\Phi_v := X_{\varepsilon_v F_v}^1$ denote the time-1 map of the Hamiltonian vector field $X_{\varepsilon_v F_v}$. It is a symplectic transformation. A simple calculation shows that Φ_v changes $H_v = N_v + P_v$ into

$$H_{v+1} = H_v \circ \Phi_v = N_v + \tilde{R}_{v+1} + O(|y|^2 + |y| ||z||_{a,p} + ||z||_{a,p}^3) \tag{2.2}$$

with

$$\tilde{R}_{v+1} = \frac{1}{2} \varepsilon_v^2 \{\{N_v, F_v\}, F_v\} + \varepsilon_v^2 \{R_v, F_v\} + \cdots.$$

Our task is now to search for F_v satisfying $\{N_v, F_v\} + R_v = 0$ which is a set of the first order partial differential equations:

$$\omega \cdot \partial_x F^x(x) = R^x(x), \quad \omega \cdot \partial_x F^y(x) = R^y(x),$$

$$\cdots$$

$$\omega \cdot \partial_x F^{z\bar{z}}(x) + \mathbf{i}\Lambda F^{z\bar{z}}(x) - \mathbf{i}F^{z\bar{z}}(x)\Lambda = R^{z\bar{z}}(x),$$

5

where $\Lambda = \mathrm{diag}\,(\Omega_j : j = 1, 2, \ldots)$. Let us consider the last equation which is the most difficulty one. By $F_{ij}(x)$ and $R_{ij}(x)$, denote the matrix elements of the operators $F^{z\bar{z}}$ and $R^{z\bar{z}}$, respectively. Then the last equation becomes

$$-\mathbf{i}\omega \cdot \partial_x F_{ij}(x) + (\Omega_i - \Omega_j)F_{ij}(x) = -\mathbf{i}R_{ij}(x), \ i, j = 1, 2, \ldots \tag{2.3}$$

or

$$(\langle k, \omega \rangle + \Omega_i - \Omega_j)\widehat{F}_{ij}(k) = -\mathbf{i}\widehat{R}_{ij}(k),$$

where $\widehat{F}_{ij}(k)$ and $\widehat{R}_{ij}(k)$ are the k-th Fourier coefficients of $F_{ij}(x)$ and $R_{ij}(x)$, respectively. One can assume $\widehat{R}_{jj}(0) = 0$, otherwise $\widehat{R}_{jj}(0)$ is put into Ω_j as a modification of the normal form N_v. Thus, (2.3) can be solved by

$$\widehat{F}_{ij}(k) = \frac{\widehat{R}_{ij}(k)}{\langle k, \omega \rangle + \Omega_i - \Omega_j}$$

under the non-resonant conditions

$$\langle k, \omega \rangle + \Omega_i - \Omega_j \neq 0 \quad \text{unless} \quad k = 0, i = j.$$

This is actually the KAM iterative procedure for bounded perturbation P. However, the thing is not so simple when the perturbation P is unbounded, i.e., $\tilde{d} = p - q > 0$. When

$$X_P: \ \mathscr{P}^{a,p} \to \mathscr{P}^{a,q}, \quad \tilde{d} = p - q > 0,$$

one has, very roughly,

$$|\widehat{R}_{ij}(k)| \approx |i|^{\tilde{d}} + |j|^{\tilde{d}} \to \infty, \quad \text{as } |i| + |j| \to \infty.$$

This leads usually to $\widehat{F}_{ij}(k) \to \infty$. In other words, the solution F_v or the transformation $\Phi_v = X^1_{\varepsilon_v F_v}$ would be unbounded. One should note that the coordinate transformations $\Phi_v = X^1_{\varepsilon_v F_v}$ *must be bounded* even if the the perturbation P is unbounded in order that the domains of the KAM iteration are always in the same phase space $\mathscr{P}^{a,p}$ and that the KAM iterative procedure can work. In order to guarantee the bounded-ness of F_v, it is required in [7] and [10] that

$$|\Omega_i - \Omega_j| \approx |i|^{d-1} + |j|^{d-1}, \quad i \neq j$$

or rather roughly,

$$|\langle k, \omega \rangle + \Omega_i - \Omega_j| \approx |i|^{d-1} + |j|^{d-1}.$$

Together with $\tilde{d} \leq d - 1$, one has that for $i \neq j$,

$$|\widehat{F}_{ij}(k)| \approx \frac{|i|^{\tilde{d}} + |j|^{\tilde{d}}}{|i|^{d-1} + |j|^{d-1}} = O(1).$$

It is clear that this estimate fails when $i = j$. To avoid this plight, Kuksin[7] smartly put the whole $R_{jj}(x)$ rather than $\widehat{R}_{jj}(0)$ into Ω_j as a modification of the normal form N_v so that it is not necessary to solve the equation for $F_{jj}(x)$. In doing so, the term N_v becomes into the generalized normal form (i.e. normal frequencies depend on the angle variable x)

$$\tilde{N}_v := N_v + \sum_{j=1}^{\infty} R_{jj}(x)z_j\bar{z}_j = \langle \omega, y \rangle + \sum_{j=1}^{\infty} (\Omega_j + R_{jj}(x))z_j\bar{z}_j,$$

the homological equation (2.1) is modified into

$$\{\tilde{N}_v, F_v\} + R_v = 0, \tag{2.4}$$

6

and the remaining term $\tilde{R}_{\nu+1}$ is changed into

$$\tilde{R}_{\nu+1} = \frac{1}{2}\varepsilon_\nu^2\{\{\tilde{N}_\nu, F_\nu\}, F_\nu\} + \varepsilon_\nu^2\{R_\nu, F_\nu\} + \cdots. \tag{2.5}$$

Accordingly, the homological equation (2.3) becomes into

$$-\mathbf{i}\omega \cdot \partial_x F_{ij}(x) + (\Omega_i - \Omega_j + R_{ii}(x) - R_{jj}(x))F_{ij}(x) = -\mathbf{i}R_{ij}(x), \quad i \neq j. \tag{2.6}$$

Let $u = F_{ij}(x)$, $\lambda = \Omega_i - \Omega_j$, $\mu(x) = R_{ii}(x) - R_{jj}(x)$ and $r(x) = -\mathbf{i}R_{ij}(x)$. Then this equation can be abbreviated as an abstract equation

$$-\mathbf{i}\omega \cdot \partial_x u + (\lambda + \mu(x))u = r(x). \tag{2.7}$$

Since R_{ii} and R_{jj} are large, $\mu(x)$ is usually large. And the coefficient $\mu(x)$ involves the angle variable. The equation of this type is called "small-denominators equations with large variable coefficients" by Kuksin [6]. Remark that, for simplicity, the modification of ω is omitted here.

2.2. Solving the homological equations. In order to make KAM iterative procedure work, the existence domain of the solution u should be the strip-type neighborhood of $\mathbb{T}^n$ with some width $s > 0$: $D(s) := \{x \in \mathbb{C}^n/2\pi\mathbb{Z}^n : |\mathrm{Im}\, x| \leq s\}$. Assume

$$\mu(x) \approx C\tilde{\gamma},$$

where C is some small constant and where $\tilde{\gamma}$ should be usually a large magnitude. Since (2.7) is scalar, it can be solved directly and estimated by

$$\sup_{x \in D(s-\sigma)} |u(x)| \lesssim e^{2C\tilde{\gamma}} \|r\|_s, \quad 0 \leq \sigma < s, \tag{2.8}$$

where $\|r\|_s := \sup_{x \in D(s)} |r(x)|$. This estimate is, however, not good enough to support the KAM iteration procedure, since the solution u becomes too large as μ is large. In fact, in the ν-th KAM iteration step, $\tilde{\gamma} \approx 2^\nu$. Thus, $u \approx \exp(2^\nu)$ goes to infinity very rapidly, which makes the coordinate transformations essentially unbounded. When $\tilde{d} < d - 1$, Kuksin's lemma [6] can solve this problem. Following Kuksin [6], we assume there are constant $C > 0$ and $0 < \theta < 1$ such that

$$|\lambda|^\theta \geq C\|\mu\|_s, \tag{2.9}$$

Kuksin's Lemma states that under suitable non-resonant conditions on ω, the solution u satisfies the estimate:

$$\|u\|_{s-\sigma} \leq C_1 \exp(C_2 C_3^{\frac{1}{1-\theta}})\|r\|_s, \tag{2.10}$$

where C_1 and C_2 are positive constants depending on only n and σ, and $C_3 > 0$ is a constant depending on the non-resonant conditions. When this estimate is applied to (2.6), one needs to take

$$\lambda = \Omega_i - \Omega_j \approx i^d - j^d \approx i^{d-1} + j^{d-1}, \quad \tilde{\gamma} \approx i^{\tilde{d}} + j^{\tilde{d}}, \quad i, j = 1, 2, 3, \ldots.$$

From this, we see that there is indeed a constant $0 < \theta < 1$ such that (2.9) holds true if $\tilde{d} < d - 1$. Therefore, the solution u of the homological equation (2.7) has a uniform bound independent of the size of μ. This makes the coordinate transformation bounded. We see also that $\tilde{d} = d - 1$ leads to $\theta = 1$ in (2.9). In this case, the estimate (2.10) is invalid. (The right hand side of (2.10) is equal to ∞.) Now it is clear that we need some new estimate for the solution u covering not only the case $\tilde{d} < d - 1$ but also the limit case $\tilde{d} = d - 1$. The new estimate has been obtained in our recent paper [12]:

$$\|u\|_{s-\sigma} \leq C_4 e^{C\tilde{\gamma}s}\|r\|_s, \tag{2.11}$$

7

where C_4 is a constant depending the non-resonant conditions on ω and where C is a positive constant small enough. Since the parameter $\tilde{\gamma}$ (which measures the magnitude of the perturbation $\mu(x)$) goes into the exponential in the right hand side of (2.11), in this sense, the upper bound of this new estimate looks weaker than that of the original Kuksin's lemma. However, compared with (2.8), there is some essential improvement in (2.11): there is an s in the exponential in (2.11). This number s will be used crucially in the following manner: In the ν-th KAM iteration, we let $s = 2^{-\nu}$. One will find that there is no the small divisor problem in the homological equation (2.7) when $\tilde{\gamma} >$ some large constant $K \approx 2^{\nu}(5/4)^{\nu}|\ln \varepsilon|$. In this case, the homological equation (2.7) can be solved by the implicit function theorem. So the non-trivial case is when $\tilde{\gamma} \leq K$. At this time, we find that

$$e^{C\tilde{\gamma}s} \lesssim \varepsilon_\nu^{-C}$$

with $\varepsilon_\nu := \varepsilon^{(5/4)^\nu}$ and the constant $C \ll 1$. Thus, by the new estimate (2.11) and noting $||r||_s \leq \varepsilon_\nu$, we have

$$||u||_{s-\sigma} \lesssim \varepsilon_\nu^{-C}||r||_s \leq \varepsilon_\nu^{-C}\varepsilon_\nu \lesssim \varepsilon_\nu. \tag{2.12}$$

In the usual KAM iteration, one would have obtained $||u||_{s-\sigma} \leq 2^{\nu}\varepsilon_\nu \lesssim \varepsilon_\nu$. Although here $\varepsilon_\nu^{-C}\varepsilon_\nu \gg 2^{\nu}\varepsilon_\nu$, inequality (2.12) can guarantee the KAM procedure to be iterated. Therefore, although the new estimate (2.11) is weaker than that of the original Kuksin's Lemma, it can covers both $\tilde{d} < d-1$ and the limit case $\tilde{d} = d-1$, and it is sufficient for the proof of the KAM theorems of the present paper.

2.3. Estimate of the remaining terms $\tilde{R}_{\nu+1}$, etc. Recall $R_\nu = O(1)$. By (2.11), $F_\nu = O(\varepsilon_\nu^{-C})$ with $0 < C \ll 1$. Thus,

$$\{F_\nu, R_\nu\} = O(\varepsilon_\nu^{-C}).$$

By (2.4),

$$\{\tilde{N}_\nu, F_\nu\} = -R_\nu = O(1).$$

Thus,

$$\{\{\tilde{N}_\nu, F_\nu\}, F_\nu\} = \{O(1), O(\varepsilon_\nu^{-C})\} = O(\varepsilon_\nu^{-C}).$$

Consequently, by (2.5), very roughly,

$$\tilde{R}_{\nu+1} = \frac{1}{2}\varepsilon_\nu^2\{\{\tilde{N}, F_\nu\}, F_\nu\} + \varepsilon_\nu^2\{R_\nu, F_\nu\} + \cdots = O(\varepsilon_\nu^{2-C}) = O(\varepsilon_{\nu+1}) := \varepsilon_{\nu+1}R_{\nu+1},$$

where $R_{\nu+1} = O(1)$. Therefore, we can rewrite $H_{\nu+1}$ as

$$H_{\nu+1} = \tilde{N}_{\nu+1} + \varepsilon_{\nu+1}R_{\nu+1} + O(|y|^2 + |y|||z||_{a,p} + ||z||_{a,p}^3). \tag{2.13}$$

2.4. Convergence of the iterative procedure. Repeating the above procedure and letting $\nu \to \infty$ and noting $\varepsilon_\nu \to 0$ very fast, one can finally get

$$H_\infty := \lim_{\nu \to \infty} H \circ \Phi_1 \circ \cdots \circ \Phi_\nu = \tilde{N}_\infty + O(|y|^2 + |y|||z||_{a,p} + ||z||_{a,p}^3),$$

where $\tilde{N}_\infty = \lim_{\nu \to \infty} \tilde{N}_\nu = \langle \omega_\infty, y \rangle + \sum_{j \geq 1} \Omega_{\infty,j}(x)z_j\bar{z}_j$. The system corresponding to H_∞ is

$$\begin{cases} \dot{x} = \frac{\partial H_\infty}{\partial y} = \omega_\infty + O(|y| + ||z||_{a,p}) \\ \dot{y} = -\frac{\partial H_\infty}{\partial x} = O(|y|^2 + |y|||z||_{a,p} + ||z||_{a,p}^2) \\ \dot{z}_j = -\mathbf{i}\frac{\partial H_\infty}{\partial \bar{z}_j} = -\mathbf{i}\Omega_{\infty,j}z_j + O(|y| + ||z||_{a,p}^2), \quad j \geq 1. \end{cases}$$

It is clearly seen that

$$\mathscr{T}_\infty := \{(x, y, z, \bar{z}) \in \mathscr{P}^{a,p} : x = \omega_\infty t, y = 0, z = \bar{z} = 0, t \in \mathbb{R}\}$$

8

forms an invariant torus of the hamiltonian vector field X_{H_∞}. Going back to the original vector field X_H, then $(\lim_{\nu \to \infty} \Phi_1 \circ \cdots \circ \Phi_\nu) \mathscr{T}_\infty$ is an invariant torus of the Hamiltonian system defined by H.

2.5. Estimate of the measure of the parameters. When $\tilde{d} = d - 1$, another new difficult also arises in search for F. It is under suitable non-resonant conditions that either (2.11) or (2.10) can hold true. In other words, one has to remove some resonant sets consisting of "bad" parameters ξ, equivalently, to remove the "bad" parameters ω when $\omega = \omega(\xi)$ depends on ξ in some non-degenerate way. For example, we need to eliminate the resonant set

$$\{\xi \in \Pi : \langle k, \omega(\xi) \rangle + \Omega_i(\xi) - \Omega_j(\xi) \text{ is small}\}$$

where $i \neq j \in \mathbb{N}$ and $k \in \mathbb{Z}^n$. Clearly we hope that the Lebesgue measure of the set is small. To that end we need to verify that $\langle k, \omega(\xi) \rangle + \Omega_i(\xi) - \Omega_j(\xi)$ is twisted with respect to $\xi \in \Pi$, equivalently, twisted with respect to $\omega \in \omega(\Pi)$, that is, we need to show

$$(\star) := |\langle k, \omega \rangle + \Omega_i(\xi(\omega)) - \Omega_j(\xi(\omega))|_{\omega(\Pi)}^{\text{lip}} > 0,$$

where $\omega(\xi(\omega)) = \omega$.

Recall $|\Omega_i - \Omega_j| \geq m|i^d - j^d|$. Thus, $\langle k, \omega \rangle + \Omega_i - \Omega_j$ is not small if $|k| \leq \tilde{C}_0|i^d - j^d|$ with some constant $\tilde{C}_0$ depending on m and $|\omega|_\Pi$. Now we assume $|k| \geq \tilde{C}_0|i^d - j^d|$. At the ν-th KAM step, because of the modification of frequencies from unbounded perturbation, we have

$$|\Omega_j|_{\omega(\Pi)}^{\text{lip}} = O(j^\delta) + O(j^{\tilde{d}}).$$

By a small trick (See §3.2 below.), we can let $\delta = \tilde{d}$. Therefore, there exists a constant $\tilde{C}_1$ such that

$$(\star) \quad \geq \quad |k| - \tilde{C}_1(i^{\tilde{d}} + j^{\tilde{d}}) \tag{2.14}$$

$$\geq \quad \frac{1}{2}\tilde{C}_0(i^{d-1} + j^{d-1}) - \tilde{C}_1(i^{\tilde{d}} + j^{\tilde{d}}) \tag{2.15}$$

$$\geq \quad \left(\frac{1}{4}\tilde{C}_0(i^{d-1-\tilde{d}} + j^{d-1-\tilde{d}}) - \tilde{C}_1\right)(i^{\tilde{d}} + j^{\tilde{d}}), \tag{2.16}$$

which is similar to (10.50) in [7] (See also page 174 in [10]). When $\tilde{d} < d - 1$, it follows from (2.16) that $(\star) > 0$ if

$$\max(i, j) > \left(\frac{4\tilde{C}_1}{\tilde{C}_0}\right)^{\frac{1}{d-1-\tilde{d}}} := \mathscr{K}. \tag{2.17}$$

Therefore, if $\tilde{d} < d - 1$ it remains to verify the twist condition $(\star) > 0$ just for only finite number of cases

$$(k, i, j): \quad \tilde{C}_0|i^d - j^d| \leq |k| \leq \tilde{C}_1(i^{\tilde{d}} + j^{\tilde{d}}) \quad \text{and} \quad i, j \leq \mathscr{K}. \tag{2.18}$$

Note that the constant $\mathscr{K}$ is independent of i, j and k, so is the number of the cases. For these cases, the measure estimate of the resonant sets can be dealt with by some initial assumptions (For example, see Proposition 22.2 in [10]). When $\tilde{d} = d - 1$, $\mathscr{K} = +\infty$. However, it follows directly from (2.15) that $(\star) > (\frac{1}{2}\tilde{C}_0 - \tilde{C}_1)(i^{d-1} + j^{d-1}) > 0$ if $\tilde{C}_0 > 2\tilde{C}_1$. This completes the measure estimate of the resonant sets for Theorem 1.1. One can verify that the condition $\tilde{C}_0 > 2\tilde{C}_1$ is indeed satisfied by DNLS equations.

However, the inequality $\tilde{C}_0 > 2\tilde{C}_1$ is not satisfied by BO equation. The procedure of the measure estimate of the resonant set is modified as follows:

Assume that Ω can be split into two part: $\Omega_j = \Omega_j^1 + \Omega_j^2$ such that $|\Omega_j^1|_{\omega(\Pi)}^{\text{lip}} \leq \tilde{C}_2 j^{\delta_0}$ with $\delta_0 < d - 1$, and $|\Omega_j^2|_{\omega(\Pi)}^{\text{lip}} \leq \tilde{C}_3 j^{d-1}$ with $\tilde{C}_3$ suitably small. Thus,

$$(\star) \geq \left(|k| - \tilde{C}_2(i^{\delta_0} + j^{\delta_0}) - \tilde{C}_3(i^{d-1} + j^{d-1})\right)$$

$$\geq (\frac{1}{2}\tilde{C}_0 - \tilde{C}_3)(i^{d-1} + j^{d-1}) - \tilde{C}_2(i^{\delta_0} + j^{\delta_0}).$$

9

It follows from $\delta_0 < d-1$ that there is a constant $\tilde{\mathscr{K}} = \tilde{\mathscr{K}}(d,\delta_0,\tilde{C}_0,\tilde{C}_2,\tilde{C}_3) > 0$ such that $(\star) > 0$ if $\max\{i,j\} > \tilde{\mathscr{K}}$ and $\frac{1}{2}\tilde{C}_0 > \tilde{C}_3$. The latter inequality is satisfied by BO equation with $d = 2$ and $\delta_0 = 0$. We also mention that the partition of Ω into $\Omega^1 + \Omega^2$ is rather natural. In fact, Ω^1 is usually regarded as the initial frequency vector, while Ω^2 corresponds to the modification in KAM iteration steps.

Remarks. 1. In [12] we mentioned Theorem 1.1 above and its application to DNLS equation without proof. In this paper we give the proof and add Theorem 1.2 and a new application to Benjamin-Ono equation.

2. The proofs of both (2.10) and (2.11) depends heavily on the fact that the homological equation (2.7) is 1-dimensional so that the solution can be expressed explicitly. This dimensional restriction requires that the normal frequency Ω_j's must be *simple*, i.e., $\Omega_j^\sharp = 1$. Therefore, the range of application of Theorems 1.1 and 1.2 and Kuksin's KAM theorem ([7]) lies in those PDEs with simple frequencies such as DNLS equation subject to Dirichlet boundary conditions, KdV and BO equations with periodic boundary conditions. For a class of DNLS equation

$$\mathbf{i}u_t + u_{xx} - M_\sigma u + \mathbf{i}(|u|^2 u)_x = 0$$

subject to periodic boundary conditions, the multiplicity $\Omega_j^\sharp = 2$. And for Kadomtsev-Petviashvili (KP) equation

$$(u_t + u_{xxx} + uu_x)_x \pm u_{yy} = 0, \quad u = u(t,x,y), \ (x,y) \in \mathbb{T}^2,$$

its frequency multiplicity $\Omega_j^\sharp \to \infty$ as $|j| \to \infty$. There is nothing to know about the existence of KAM tori for these two classes of PDEs with perturbations. In particular, the existence of KAM tori for KP equation is a well-known open problem by Kuksin. See [8],[4].

3 The Homological Equations

3.1 Derivation of Homological Equations

The proof of Theorem 1.1 employs the rapidly converging iteration scheme of Newton type to deal with small divisor problems introduced by Kolmogorov, involving infinite sequence of coordinate transformations. At the ν-th step of the scheme, a Hamiltonian

$$H_\nu = N_\nu + P_\nu$$

is considered, as a small perturbation of some normal form N_ν. A transformation Φ_ν is set up so that

$$H_\nu \circ \Phi_\nu = N_{\nu+1} + P_{\nu+1}$$

with another normal form $N_{\nu+1}$ and a much smaller perturbation $P_{\nu+1}$. We drop the index ν of H_ν, N_ν, P_ν, Φ_ν and shorten the index $\nu+1$ as $+$.

Using convenient complex notation $z = (u - \mathbf{i}v)/\sqrt{2}$ and $\bar{z} = (u + \mathbf{i}v)/\sqrt{2}$, the generalized normal form reads

$$N = \langle \omega(\xi), y \rangle + \sum_{j \geq 1} \Omega_j(x;\xi) z_j \bar{z}_j. \tag{3.1}$$

Let R be 2-order Taylor polynomial truncation of P, that is,

$$R = R^x + \langle R^y, y \rangle + \langle R^z, z \rangle + \langle R^{\bar{z}}, \bar{z} \rangle + \langle R^{zz}z, z \rangle + \langle R^{\bar{z}\bar{z}}\bar{z}, \bar{z} \rangle + \langle R^{z\bar{z}}z, \bar{z} \rangle, \tag{3.2}$$

where $\langle \cdot, \cdot \rangle$ is formal product for two column vectors and $R^x, R^y, R^z, R^{\bar{z}}, R^{zz}, R^{\bar{z}\bar{z}}, R^{z\bar{z}}$ depend on x and ξ. For a function u on $\mathbb{T}^n$, let

$$[u] = \frac{1}{(2\pi)^n} \int_{\mathbb{T}^n} u(x)dx.$$

By $\llbracket R \rrbracket$ denote the part of R in generalized normal form as follows

$$\llbracket R \rrbracket = [R^x] + \langle [R^y], y \rangle + \langle \mathrm{diag}(R^{z\bar{z}})z, \bar{z} \rangle,$$

where $\mathrm{diag}(R^{z\bar{z}})$ is the diagonal of $R^{z\bar{z}}$. Note that $[R^x]$ and $[R^y]$ are independent of x. In the following, the term $[R^x]$ will be omitted since it does not affect the dynamics.

The coordinate transformation Φ is obtained as the time-1-map $X_F^t|_{t=1}$ of a Hamiltonian vector field X_F, where F is of the same form as R:

$$F = F^x + \langle F^y, y \rangle + \langle F^z, z \rangle + \langle F^{\bar{z}}, \bar{z} \rangle + \langle F^{zz}z, z \rangle + \langle F^{\bar{z}\bar{z}}\bar{z}, \bar{z} \rangle + \langle F^{z\bar{z}}z, \bar{z} \rangle, \tag{3.3}$$

and $\llbracket F \rrbracket = 0$. Denote $\partial_\omega = \sum_{1 \le b \le n} \omega_b \frac{\partial}{\partial x_b}$, $\Lambda = \mathrm{diag}(\Omega_j : j \ge 1)$. Then we have

$$
\begin{aligned}
H \circ \Phi &= (N+R) \circ X_F^1 + (P-R) \circ X_F^1 \\
&= N + \{N,F\} + R + \int_0^1 \{(1-t)\{N,F\} + R, F\} \circ X_F^t dt + (P-R) \circ X_F^1 \\
&= N + \sum_{j \ge 1} \langle \partial_x \Omega_j, F^y \rangle z_j \bar{z}_j + \langle [R^y], y \rangle + \langle \mathrm{diag}(R^{z\bar{z}})z, \bar{z} \rangle
\end{aligned}
\tag{3.4}
$$

$$+ (-\partial_\omega F^x + R^x) \tag{3.5}$$
$$+ \langle -\partial_\omega F^y + R^y - [R^y], y \rangle \tag{3.6}$$
$$+ \langle -\partial_\omega F^z + \mathbf{i}\Lambda F^z + R^z, z \rangle \tag{3.7}$$
$$+ \langle -\partial_\omega F^{\bar{z}} - \mathbf{i}\Lambda F^{\bar{z}} + R^{\bar{z}}, \bar{z} \rangle \tag{3.8}$$
$$+ \langle (-\partial_\omega F^{zz} + \mathbf{i}\Lambda F^{zz} + \mathbf{i}F^{zz}\Lambda + R^{zz})z, z \rangle \tag{3.9}$$
$$+ \langle (-\partial_\omega F^{\bar{z}\bar{z}} - \mathbf{i}\Lambda F^{\bar{z}\bar{z}} - \mathbf{i}F^{\bar{z}\bar{z}}\Lambda + R^{\bar{z}\bar{z}})\bar{z}, \bar{z} \rangle \tag{3.10}$$
$$+ \langle (-\partial_\omega F^{z\bar{z}} - \mathbf{i}\Lambda F^{z\bar{z}} + \mathbf{i}F^{z\bar{z}}\Lambda + R^{z\bar{z}} - \mathrm{diag}(R^{z\bar{z}}))z, \bar{z} \rangle \tag{3.11}$$

$$+ \int_0^1 \{(1-t)\{N,F\} + R, F\} \circ X_F^t dt + (P-R) \circ X_F^1. \tag{3.12}$$

We wish to find the function F such that (3.5)-(3.11) vanish. To this end, F^x, F^y, F^z, $F^{\bar{z}}$, F^{zz}, $F^{\bar{z}\bar{z}}$ and $F^{z\bar{z}}$ should satisfy the homological equations:

$$\partial_\omega F^x = R^x, \tag{3.13}$$
$$\partial_\omega F^y = R^y - [R^y], \tag{3.14}$$
$$\partial_\omega F_j^z - \mathbf{i}\Omega_j F_j^z = R_j^z, \quad j \ge 1, \tag{3.15}$$
$$\partial_\omega F_j^{\bar{z}} + \mathbf{i}\Omega_j F_j^{\bar{z}} = R_j^{\bar{z}}, \quad j \ge 1, \tag{3.16}$$
$$\partial_\omega F_{ij}^{zz} - \mathbf{i}(\Omega_i + \Omega_j)F_{ij}^{zz} = R_{ij}^{zz}, \quad i, j \ge 1, \tag{3.17}$$
$$\partial_\omega F_{ij}^{\bar{z}\bar{z}} + \mathbf{i}(\Omega_i + \Omega_j)F_{ij}^{\bar{z}\bar{z}} = R_{ij}^{\bar{z}\bar{z}}, \quad i, j \ge 1, \tag{3.18}$$
$$\partial_\omega F_{ij}^{z\bar{z}} + \mathbf{i}(\Omega_i - \Omega_j)F_{ij}^{z\bar{z}} = R_{ij}^{z\bar{z}}, \quad i, j \ge 1, \ i \ne j. \tag{3.19}$$

3.2 Solving the Homological Equations

Let $\Omega = (\Omega_j : j \ge 1)$, $\bar{\Omega} = [\Omega]$ and $\tilde{\Omega} = \Omega - \bar{\Omega}$. Define

$$\langle k \rangle = \max\{1, |k|\}, \quad \langle l \rangle_d = \max\{1, |\sum_{j \ge 1} j^d l_j|\}.$$

Moreover, for an analytic function u on $D(s)$, we define

$$|u|_{s,\tau} := \sum_{k \in \mathbb{Z}^n} |\hat{u}_k| |k|^\tau e^{|k|s},$$

where $\hat{u}_k := (2\pi)^{-n} \int_{\mathbb{T}^n} u(x) e^{-\mathbf{i}k \cdot x} dx$ is the k-Fourier coefficient of u.

Consider the conditions $\delta \leq d-1$ and $\tilde{d} = p-q \leq d-1$. If $\delta > \tilde{d}$, decreasing q such that $\delta = \tilde{d}$, then for the new q, the inequality (1.7) still holds true; if $\delta < \tilde{d}$, increasing δ such that $\delta = \tilde{d}$, then for the new δ, the assumption (C) still holds true, and if $\tilde{d} = d-1$, the assumption (D*) is satisfied with $\Omega_2 = 0$. Thus, without loss of generality we assume $\delta = \tilde{d} \leq d-1$ in the following.

Equations (3.13)-(3.19) will be solved under the following conditions: uniformly on Π,

$$|\langle k, \omega(\xi) \rangle + \langle l, \bar{\Omega}(\xi) \rangle| \geq \alpha \frac{\langle l \rangle_d}{\langle k \rangle^\tau}, \quad k \neq 0, \ |l| \leq 2, \tag{3.20}$$

$$|\langle l, \bar{\Omega}(\xi) \rangle| \geq m \langle l \rangle_d, \quad 0 < |l| \leq 2, \tag{3.21}$$

$$|\tilde{\Omega}_j|_{s,\tau+1} \leq \alpha \gamma_0 j^\delta, \quad j \geq 1, \tag{3.22}$$

with constants $\tau \geq n$, $d > 1$, $0 < \gamma_0 \leq 1/8$, $m > 0$, and a parameter $0 < \alpha \leq m$. We mention that d is the same as in Theorem 1.1 and α, m will be the iteration parameters α_ν, m_ν in the ν-th KAM step.

Equations (3.13) (3.14) can be easily solved by a standard approach in classical, finite dimensional KAM theory, so we only give the related results at the end of this subsection. Equations (3.15)-(3.18) are easier than (3.19) and can be solved in the same way as (3.19) done, so we only give the details of solving (3.19) in the following.

For any positive number K, we introduce a truncation operator Γ_K as follows:

$$(\Gamma_K f)(x) := \sum_{|k| \leq K} \hat{f}_k e^{\mathbf{i}k \cdot x}, \quad \forall f : \mathbb{T}^n \to \mathbb{C},$$

where $\hat{f}_k$ is the k-Fourier coefficient of f.

Set $C_0 = 2|\omega|_\Pi / m$ and K being a positive number which will be the iteration parameter K_ν in the ν-th KAM step.

(1) For (i,j) with $0 < |i^d - j^d| < C_0 K$, we solve exactly (3.19):

$$\partial_\omega F_{ij}^{z\bar{z}} + \mathbf{i}(\Omega_i - \Omega_j) F_{ij}^{z\bar{z}} = R_{ij}^{z\bar{z}}; \tag{3.23}$$

(2) for (i,j) with $|i^d - j^d| \geq C_0 K$, we solve the truncated equation of (3.19):

$$\partial_\omega F_{ij}^{z\bar{z}} + \mathbf{i}\Gamma_K\left((\Omega_i - \Omega_j) F_{ij}^{z\bar{z}}\right) = \Gamma_K R_{ij}^{z\bar{z}}, \quad \Gamma_K F_{ij}^{z\bar{z}} = F_{ij}^{z\bar{z}}. \tag{3.24}$$

Comparing (3.24) with (3.19), we find that (3.11) doesn't vanish. Actually, at this time, (3.11) is equivalent to $\langle \hat{R}^{z\bar{z}} z, \bar{z} \rangle$ with the matrix elements of $\hat{R}^{z\bar{z}}$ being defined by

$$\hat{R}_{ij}^{z\bar{z}} = \begin{cases} 0, & |i^d - j^d| < C_0 K, \\ (1 - \Gamma_K)\left(-\mathbf{i}(\Omega_i - \Omega_j) F_{ij}^{z\bar{z}} + R_{ij}^{z\bar{z}}\right), & |i^d - j^d| \geq C_0 K. \end{cases} \tag{3.25}$$

Letting $\Omega_{ij} = \Omega_i - \Omega_j = \bar{\Omega}_{ij} + \tilde{\Omega}_{ij}$ and dropping the superscript '$z\bar{z}$' for brevity, (3.23) (3.24) (3.25) become

$$-\mathbf{i}\partial_\omega F_{ij} + \bar{\Omega}_{ij} F_{ij} + \tilde{\Omega}_{ij} F_{ij} = -\mathbf{i}R_{ij}, \tag{3.26}$$

12

$$-\mathbf{i}\partial_\omega F_{ij} + \bar{\Omega}_{ij}F_{ij} + \Gamma_K(\tilde{\Omega}_{ij}F_{ij}) = -\mathbf{i}\Gamma_K R_{ij}, \tag{3.27}$$

$$\hat{R}_{ij} = \begin{cases} 0, & 0 < |i^d - j^d| < C_0 K, \\ (1-\Gamma_K)(-\mathbf{i}\tilde{\Omega}_{ij}F_{ij} + R_{ij}), & |i^d - j^d| \geq C_0 K. \end{cases} \tag{3.28}$$

We are now in position to solve the homological equations (3.26) (3.27) by using the following two lemmas, which have been proved in [12] as Theorem 1.4 and Lemma 2.6 respectively:

Lemma 3.1 ([12]). *Consider the first order partial differential equation*

$$-\mathbf{i}\partial_\omega u + \lambda u + \mu(x)u = p(x), \quad x \in \mathbb{T}^n, \tag{3.29}$$

for the unknown function u defined on the torus $\mathbb{T}^n$, where $\omega = (\omega_1, \cdots, \omega_n) \in \mathbb{R}^n$ and $\lambda \in \mathbb{C}$. Assume

(1) There are constants $\alpha, \tilde{\gamma} > 0$ and $\tau > n$ such that

$$|k \cdot \omega| \geq \frac{\alpha}{|k|^\tau}, \quad k \in \mathbb{Z}^n \setminus \{0\}, \tag{3.30}$$

$$|k \cdot \omega + \lambda| \geq \frac{\alpha\tilde{\gamma}}{1 + |k|^\tau}, \quad k \in \mathbb{Z}^n. \tag{3.31}$$

(2) $\mu : D(s) \to \mathbb{C}$ is real analytic (here 'real' means $\mu(\mathbb{T}^n) \subset \mathbb{R}$) and is of zero average: $[\mu] = 0$. Moreover, assume there is constant $C > 0$ such that

$$|\mu|_{s,\tau+1} \leq C\tilde{\gamma}. \tag{3.32}$$

(3) $p(x)$ is analytic in $x \in D(s)$.

Then (3.29) has a unique solution $u(x)$ which is defined in a narrower domain $D(s-\sigma)$ with $0 < \sigma < s$, and which satisfies

$$\sup_{x \in D(s-\sigma)} |u(x)| \leq \frac{c(n,\tau)}{\alpha\gamma\sigma^{n+\tau}} e^{2C\tilde{\gamma}s/\alpha} \sup_{x \in D(s)} |p(x)| \tag{3.33}$$

for $0 < \sigma < \min\{1,s\}$, where the constant $c(n,\tau) = (6e+6)^n[1 + (\frac{3\tau}{e})^\tau]$.

Lemma 3.2 ([12]). *Consider the first order partial differential equation with the truncation operator Γ_K*

$$-\mathbf{i}\partial_\omega u + \lambda u + \Gamma_K(\mu u) = \Gamma_K p, \quad x \in \mathbb{T}^n, \tag{3.34}$$

for the unknown function u defined on the torus $\mathbb{T}^n$, where $\omega \in \mathbb{R}^n$, $0 \neq \lambda \in \mathbb{C}$, and $0 < 2K|\omega| \leq |\lambda|$. Assume that μ is real analytic in $x \in D(s)$ with

$$\sum_{k \in \mathbb{Z}^n} |\hat{\mu}_k| e^{|k|s} \leq \frac{|\lambda|}{4\iota} \tag{3.35}$$

for some constant $\iota \geq 1$, and assume $p(x)$ is analytic in $x \in D(s)$. Then (3.34) has a unique solution $u(x)$ with $u = \Gamma_K u$ and

$$\sup_{x \in D(s-\sigma)} |u(x)| \leq \frac{c(n)}{|\lambda|\sigma^n} \sup_{x \in D(s)} |p(x)|, \tag{3.36}$$

$$\sup_{x \in D(s-\sigma)} |(1-\Gamma_K)(\mu u)(x)| \leq \frac{c(n)}{\iota\sigma^n} e^{-9K\sigma/10} \sup_{x \in D(s)} |p(x)| \tag{3.37}$$

for $0 < \sigma < s$, where the constant $c(n) = 4(20e+20)^n$.

Set $0 < \sigma < \min\{1, s/5\}$. In what follows the notation $a \lessdot b$ stands for "there exists a positive constant c such that $a \le cb$, where c can only depend on n, τ."

First, let us consider (3.26) for (i, j) with $0 < |i^d - j^d| < C_0 K$. From (3.20) (3.21) we get

$$|\langle k, \omega(\xi) \rangle| \ge \frac{\alpha}{|k|^\tau}, \quad k \in \mathbb{Z}^n \setminus \{0\}, \tag{3.38}$$

$$|\langle k, \omega(\xi) \rangle + \bar{\Omega}_{ij}| \ge \frac{\alpha |i^d - j^d|}{1 + |k|^\tau}, \quad k \in \mathbb{Z}^n. \tag{3.39}$$

From (3.22) we get

$$|\tilde{\Omega}_{ij}|_{s, \tau+1} \le \alpha \gamma_0 (i^\delta + j^\delta) \le 2\alpha \gamma_0 |i^d - j^d|. \tag{3.40}$$

Applying Lemma 3.1 to (3.26), we have

$$|F_{ij}|_{D(s-2\sigma)} \lessdot \frac{1}{\alpha |i^d - j^d| \sigma^{n+\tau}} e^{4\gamma_0 |i^d - j^d| s} |R_{ij}|_{D(s-\sigma)}. \tag{3.41}$$

In view of

$$|i^d - j^d| < C_0 K,$$

we get

$$|F_{ij}|_{D(s-2\sigma)} \lessdot \frac{1}{\alpha |i^d - j^d| \sigma^{n+\tau}} e^{4 C_0 \gamma_0 K s} |R_{ij}|_{D(s-\sigma)}. \tag{3.42}$$

Then, let us consider (3.27) for (i, j) with $|i^d - j^d| \ge C_0 K$. From (3.22) (3.21) we get

$$|\tilde{\Omega}_{ij}|_{s, 0} \le |\tilde{\Omega}_{ij}|_{s, \tau+1} \le \alpha \gamma_0 (i^\delta + j^\delta) \le \frac{2\alpha \gamma_0 |\bar{\Omega}_{ij}|}{m |i - j|} \le \frac{|\bar{\Omega}_{ij}|}{4 |i - j|}. \tag{3.43}$$

Now applying Lemma 3.2 to (3.27), we have

$$|F_{ij}|_{D(s-2\sigma)} \lessdot \frac{1}{m |i^d - j^d| \sigma^n} |R_{ij}|_{D(s-\sigma)}, \tag{3.44}$$

$$|(1 - \Gamma_K)(\tilde{\Omega}_{ij} F_{ij})|_{D(s-2\sigma)} \lessdot \frac{1}{|i - j| \sigma^n} e^{-9K\sigma/10} |R_{ij}|_{D(s-\sigma)}. \tag{3.45}$$

For a bounded linear operator from $\ell^{a,p}$ to $\ell^{a,q}$, define its operator norm by $\| \cdot \|_{a,q,p}$. As in lemma 19.1 of [10], in view of (3.42) and (3.44), using Lemma 9.1 below, we get the estimates of $F^{z\bar{z}}$:

$$
\begin{aligned}
\|F^{z\bar{z}}\|_{a,p,p,D(s-2\sigma)}, \|F^{z\bar{z}}\|_{a,q,q,D(s-2\sigma)} \quad &\lessdot \quad \frac{1}{\alpha \sigma^{2n+\tau}} e^{4C_0 \gamma_0 K s} \|R^{z\bar{z}}\|_{a,q,p,D(s)} \\
&\lessdot \quad \frac{1}{\alpha \sigma^{2n+\tau}} e^{4C_0 \gamma_0 K s} \|X_R\|_{r,a,q,D(s,r)}.
\end{aligned}
\tag{3.46}
$$

Multiplying by $z, \bar{z}$ we then get

$$\frac{1}{r^2} |\langle F^{z\bar{z}} z, \bar{z} \rangle|_{D(s-2\sigma, r)} \le \|F^{z\bar{z}}\|_{a,p,p,D(s-2\sigma)}, \tag{3.47}$$

and finally by Cauchy's estimate we have

$$\|X_{\langle F^{z\bar{z}} z, \bar{z} \rangle}\|_{r,a,p,D(s-3\sigma, r)} \lessdot \frac{1}{\alpha \sigma^{2n+\tau+1}} e^{4C_0 \gamma_0 K s} \|X_R\|_{r,a,q,D(s,r)}. \tag{3.48}$$

To obtain the estimate of the Lipschitz semi-norm, we proceed as follows. Shortening $\Delta_{\xi\zeta}$ as Δ and applying it to (3.26) and (3.27), one gets that, for (i,j) with $0 < |i^d - j^d| < C_0 K$,

$$\mathbf{i}\partial_\omega(\Delta F_{ij}) + \bar{\Omega}_{ij}\Delta F_{ij} + \tilde{\Omega}_{ij}\Delta F_{ij} = -\mathbf{i}\partial_{\Delta\omega}F_{ij} - (\Delta\Omega_{ij})F_{ij} + \mathbf{i}\Delta R_{ij} := Q_{ij}, \tag{3.49}$$

and that, for (i,j) with $|i^d - j^d| \geq C_0 K$,

$$\mathbf{i}\partial_\omega(\Delta F_{ij}) + \bar{\Omega}_{ij}\Delta F_{ij} + \Gamma_K(\tilde{\Omega}_{ij}\Delta F_{ij}) = -\mathbf{i}\partial_{\Delta\omega}F_{ij} - \Gamma_k((\Delta\Omega_{ij})F_{ij} - \mathbf{i}\Delta R_{ij}) := Q_{ij}. \tag{3.50}$$

For $0 < |i^d - j^d| < C_0 K$, we have

$$
\begin{aligned}
|Q_{ij}|_{D(s-3\sigma)} &\leq \frac{|\Delta\omega|}{\sigma}|F_{ij}|_{D(s-2\sigma)} + (i^\delta + j^\delta)|\Delta\Omega|_{-\delta,D(s)}|F_{ij}|_{D(s-3\sigma)} + |\Delta R_{ij}|_{D(s-3\sigma)} \\
&\lessdot \frac{e^{4C_0\gamma_0 Ks}}{\alpha\sigma^{n+\tau+1}}(|\Delta\omega| + |\Delta\Omega|_{-\delta,D(s)})|R_{ij}|_{D(s-\sigma)} + |\Delta R_{ij}|_{D(s-\sigma)} \\
&\lessdot \frac{e^{4C_0\gamma_0 Ks}}{\sigma^{n+\tau+1}}\left(\frac{|\Delta\omega| + |\Delta\Omega|_{-\delta,D(s)}}{\alpha}|R_{ij}|_{D(s-\sigma)} + |\Delta R_{ij}|_{D(s-\sigma)}\right).
\end{aligned}
\tag{3.51}
$$

Again applying Lemma 3.1 to (3.49), we have

$$|\Delta F_{ij}|_{D(s-4\sigma)} \lessdot \frac{e^{8C_0\gamma_0 Ks}}{\alpha|i^d - j^d|\sigma^{2n+2\tau+1}}\left(\frac{|\Delta\omega| + |\Delta\Omega|_{-\delta,D(s)}}{\alpha}|R_{ij}|_{D(s-\sigma)} + |\Delta R_{ij}|_{D(s-\sigma)}\right). \tag{3.52}$$

For $|i^d - j^d| \geq C_0 K$, we have

$$|Q_{ij}|_{D(s-3\sigma)} \lessdot \frac{1}{\sigma^{3n}}\left(\frac{|\Delta\omega| + |\Delta\Omega|_{-\delta,D(s)}}{m}|R_{ij}|_{D(s-\sigma)} + |\Delta R_{ij}|_{D(s-\sigma)}\right). \tag{3.53}$$

Again applying Lemma 3.2 to (3.50), we have,

$$|\Delta F_{ij}|_{D(s-4\sigma)} \lessdot \frac{1}{m|i^d - j^d|\sigma^{4n}}\left(\frac{|\Delta\omega| + |\Delta\Omega|_{-\delta,D(s)}}{m}|R_{ij}|_{D(s-\sigma)} + |\Delta R_{ij}|_{D(s-\sigma)}\right), \tag{3.54}$$

$$|(1 - \Gamma_K)(\tilde{\Omega}_{ij}\Delta F_{ij})|_{D(s-4\sigma)} \lessdot \frac{e^{-9K\sigma/10}}{|i - j|\sigma^{4n}}\left(\frac{|\Delta\omega| + |\Delta\Omega|_{-\delta,D(s)}}{m}|R_{ij}|_{D(s-\sigma)} + |\Delta R_{ij}|_{D(s-\sigma)}\right). \tag{3.55}$$

In view of (3.52) and (3.54), applying Lemma 9.1 below again, we get the estimates of $\Delta F^{z\bar{z}}$:

$$
\begin{aligned}
&\|\Delta F^{z\bar{z}}\|_{a,p,p,D(s-4\sigma)}, \|\Delta F^{z\bar{z}}\|_{a,q,q,D(s-4\sigma)} \\
&\lessdot \frac{e^{8C_0\gamma_0 Ks}}{\alpha\sigma^{3n+2\tau+1}}\left(\frac{|\Delta\omega| + |\Delta\Omega|_{-\delta,D(s)}}{\alpha}\|R^{z\bar{z}}\|_{a,q,p,D(s)} + \|\Delta R^{z\bar{z}}\|_{a,q,p,D(s)}\right).
\end{aligned}
\tag{3.56}
$$

Dividing by $|\xi - \zeta| \neq 0$ and taking the supremum over Π, we get

$$
\begin{aligned}
&\|F^{z\bar{z}}\|^{\text{lip}}_{a,p,p,D(s-4\sigma)\times\Pi}, \|F^{z\bar{z}}\|^{\text{lip}}_{a,q,q,D(s-4\sigma)\times\Pi} \\
&\lessdot \frac{e^{8C_0\gamma_0 Ks}}{\alpha\sigma^{3n+2\tau+1}}\left(\frac{M}{\alpha}\|R^{z\bar{z}}\|_{a,q,p,D(s)\times\Pi} + \|R^{z\bar{z}}\|^{\text{lip}}_{a,q,p,D(s)\times\Pi}\right),
\end{aligned}
\tag{3.57}
$$

where $M := |\omega|^{\text{lip}}_\Pi + |\Omega|^{\text{lip}}_{-\delta,D(s)\times\Pi}$. Thus, in the same way as (3.48), we get

$$\|X_{\langle F^{z\bar{z}}z,\bar{z}\rangle}\|^{\text{lip}}_{r,a,p,D(s-5\sigma,r)\times\Pi} \lessdot \frac{e^{8C_0\gamma_0 Ks}}{\alpha\sigma^{3n+2\tau+2}}\left(\frac{M}{\alpha}\|X_R\|_{r,a,q,D(s,r)\times\Pi} + \|X_R\|^{\text{lip}}_{r,a,q,D(s,r)\times\Pi}\right). \tag{3.58}$$

For $\lambda \geq 0$, define

$$\| \cdot \|_{\Pi}^{\lambda} = \| \cdot \|_{\Pi} + \lambda \| \cdot \|_{\Pi}^{\mathrm{lip}}.$$

The symbol 'λ' in $\| \cdot \|_{\Pi}^{\lambda}$ will always be used in this role and never has the meaning of exponentiation. Set $0 \leq \lambda \leq \alpha/M$. From (3.48) (3.58) we get

$$\|X_{\langle F^{z\bar{z}}z,\bar{z}\rangle}\|_{r,a,p,D(s-5\sigma,r)\times\Pi}^{\lambda} \lessdot \frac{e^{8C_0\gamma_0 Ks}}{\alpha\sigma^{3n+2\tau+2}} \|X_R\|_{r,a,q,D(s,r)\times\Pi}^{\lambda}. \tag{3.59}$$

Now considering the homological equations (3.13) (3.14), by a standard approach in finite dimensional KAM theory, we can easily get

$$\|X_{F^x}\|_{r,a,p,D(s-\sigma,r)}, \quad \|X_{\langle F^y,y\rangle}\|_{r,a,p,D(s-\sigma,r)} \lessdot \frac{1}{\alpha\sigma^{\tau+n}} \|X_R\|_{r,a,q,D(s,r)}, \tag{3.60}$$

$$\|X_{F^x}\|_{r,a,p,D(s-2\sigma,r)\times\Pi}^{\mathrm{lip}}, \quad \|X_{\langle F^y,y\rangle}\|_{r,a,p,D(s-2\sigma,r)\times\Pi}^{\mathrm{lip}}$$
$$\lessdot \frac{1}{\alpha\sigma^{2\tau+2n+1}} \left(\frac{M}{\alpha}\|X_R\|_{r,a,q,D(s,r)\times\Pi} + \|X_R\|_{r,a,q,D(s,r)\times\Pi}^{\mathrm{lip}}\right). \tag{3.61}$$

From (3.60) (3.61) we get

$$\|X_{F^x}\|_{r,a,p,D(s-2\sigma,r)\times\Pi}^{\lambda}, \quad \|X_{\langle F^y,y\rangle}\|_{r,a,p,D(s-2\sigma,r)\times\Pi}^{\lambda} \lessdot \frac{1}{\alpha\sigma^{2\tau+2n+1}} \|X_R\|_{r,a,q,D(s,r)\times\Pi}^{\lambda}. \tag{3.62}$$

For the other terms of F, i.e. $\langle F^z,z\rangle$, $\langle F^{\bar{z}},\bar{z}\rangle$, $\langle F^{zz}z,z\rangle$, $\langle F^{\bar{z}\bar{z}}\bar{z},\bar{z}\rangle$, the same results - even better - than (3.59) can be obtained. Thus, we finally get the estimate for F:

$$\|X_F\|_{r,a,p,D(s-5\sigma,r)\times\Pi}^{\lambda} \lessdot \frac{e^{8C_0\gamma_0 Ks}}{\alpha\sigma^{3n+2\tau+2}} \|X_R\|_{r,a,q,D(s,r)\times\Pi}^{\lambda}. \tag{3.63}$$

4 The New Hamiltonian

From (3.4)-(3.12) we get the new Hamiltonian

$$H \circ \Phi = N_+ + P_+, \tag{4.1}$$

where $N_+ = (3.4)$ and

$$P_+ = \hat{R} + \int_0^1 \{(1-t)(\hat{N}+\hat{R})+tR, F\} \circ X_F^t dt + (P-R)\circ X_F^1, \tag{4.2}$$

where $\hat{R} = (3.7) + \cdots + (3.11) := \langle \hat{R}^z,z\rangle + \langle \hat{R}^{\bar{z}},\bar{z}\rangle + \langle \hat{R}^{zz}z,z\rangle + \langle \hat{R}^{\bar{z}\bar{z}}\bar{z},\bar{z}\rangle + \langle \hat{R}^{z\bar{z}}z,\bar{z}\rangle$. The aim of this section is to estimate the new normal form N_+ and the new perturbation P_+.

4.1 The New Normal Form

In view of (3.4), denote $N_+ = N + \hat{N}$ with

$$\hat{N} = \langle \hat{\omega},y\rangle + \sum_{j\geq 1} \hat{\Omega}_j z_j \bar{z}_j,$$

where

$$\hat{\omega} := [R^y], \tag{4.3}$$

16

$$\hat{\Omega}_j := R_{jj} + \langle \partial_x \Omega_j, F^y \rangle = R_{jj} + \langle \partial_x \tilde{\Omega}_j, F^y \rangle. \tag{4.4}$$

From (4.3) we easily get

$$|\hat{\omega}|_\Pi^\lambda \lessdot \|X_R\|_{r,a,q,D(s,r) \times \Pi}^\lambda. \tag{4.5}$$

In the following, we estimate $\hat{\Omega} = (\hat{\Omega}_j : j \geq 1)$. In view of the second estimate of (3.60),

$$|\langle \partial_x \tilde{\Omega}_j, F^y \rangle|_{D(s-\sigma)} \leq |\tilde{\Omega}_j|_{s,\tau+1} \|X_{\langle F^y, y \rangle}\|_{r,a,p,D(s-\sigma,r)} \lessdot \frac{\gamma_0 j^\delta}{\sigma^{\tau+n}} \|X_R\|_{r,a,q,D(s,r)}.$$

Thus, together with

$$|R_{jj}|_{D(s-\sigma)} \leq j^\delta \|X_R\|_{r,a,q,D(s,r)},$$

we get

$$|\hat{\Omega}|_{-\delta,D(s-\sigma)} \lessdot \frac{1}{\sigma^{\tau+n}} \|X_R\|_{r,a,q,D(s,r)}. \tag{4.6}$$

Applying Δ to $\hat{\Omega}_j$, we have

$$\Delta \hat{\Omega}_j = \Delta R_{jj} - \langle \partial_x \Delta \tilde{\Omega}_j, F^y \rangle - \langle \partial_x \tilde{\Omega}_j, \Delta F^y \rangle. \tag{4.7}$$

Since

$$|\Delta R_{jj}|_{D(s-2\sigma)} \leq j^\delta \|\Delta X_R\|_{r,a,q,D(s,r)},$$

$$|\langle \partial_x \Delta \tilde{\Omega}_j, F^y \rangle|_{D(s-2\sigma)} \leq \frac{1}{\sigma} |\Delta \Omega_j|_{D(s-\sigma)} \|X_{\langle F^y, y \rangle}\|_{r,a,p,D(s-2\sigma,r)} \lessdot \frac{|\Delta \Omega|_{-\delta,D(s)} j^\delta}{\alpha \sigma^{\tau+n+1}} \|X_R\|_{r,a,q,D(s,r)},$$

$$|\langle \partial_x \tilde{\Omega}_j, \Delta F^y \rangle|_{D(s-2\sigma)} \leq |\tilde{\Omega}_j|_{s,\tau+1} \|\Delta X_{\langle F^y, y \rangle}\|_{r,a,p,D(s-2\sigma,r)} \leq \alpha \gamma_0 j^\delta \|\Delta X_{\langle F^y, y \rangle}\|_{r,a,p,D(s-2\sigma,r)},$$

we get

$$|\hat{\Omega}|_{-\delta,D(s-2\sigma) \times \Pi}^{\mathrm{lip}} \lessdot \frac{1}{\sigma^{2n+2\tau+1}} \left(\frac{M}{\alpha} \|X_R\|_{r,a,q,D(s,r) \times \Pi} + \|X_R\|_{r,a,q,D(s,r) \times \Pi}^{\mathrm{lip}} \right). \tag{4.8}$$

Therefore, from (4.6) (4.8) we get

$$|\hat{\Omega}|_{-\delta,D(s-2\sigma) \times \Pi}^\lambda \lessdot \frac{1}{\sigma^{2n+2\tau+1}} \|X_R\|_{r,a,q,D(s,r) \times \Pi}^\lambda. \tag{4.9}$$

4.2 The New Perturbation

We firstly estimate the error term $\hat{R}^{z\bar{z}}$ with its matrix elements $\hat{R}_{ij}$ in (3.28). Split $\hat{R}^{z\bar{z}}$ into three parts: $\hat{R}^{z\bar{z}} = S^1 + S^2 + S^3$, such that S^1, S^2 have their matrix elements as follows:

$$S_{ij}^1 = \begin{cases} 0, & if \ |i^d - j^d| < C_0 K, \\ (1-\Gamma_K)(-\mathbf{i}\tilde{\Omega}_{ij} F_{ij}), & if \ |i^d - j^d| \geq C_0 K, \end{cases} \tag{4.10}$$

$$S_{ij}^2 = \begin{cases} -(1-\Gamma_K)(R_{ij}), & if \ 0 < |i^d - j^d| < C_0 K, \\ 0, & if \ i = j \ or \ |i^d - j^d| \geq C_0 K, \end{cases} \tag{4.11}$$

and S^3 is the cut-off of the perturbation $R^{z\bar{z}}$, that is,

$$S^3 = (1-\Gamma_K)(R^{z\bar{z}} - \mathrm{diag}(R^{z\bar{z}})). \tag{4.12}$$

In view of (3.45), and using Lemma 9.1 below, we get

$$\|S^1\|_{a,q,p,D(s-2\sigma)} \lessdot \frac{e^{-9K\sigma/10}}{\sigma^{2n}} \|R^{z\bar{z}}\|_{a,q,p,D(s)}. \tag{4.13}$$

17

From (3.55) and

$$|(1-\Gamma_K)((\Delta\tilde{\Omega}_{ij})F_{ij})|_{D(s-3\sigma)} \quad \lessdot \quad \frac{e^{-9K\sigma/10}}{\sigma^{2n}}(i^\delta+j^\delta)|\Delta\Omega|_{-\delta,D(s-2\sigma)}|F_{ij}|_{D(s-2\sigma)}$$

$$\lessdot \quad \frac{e^{-9K\sigma/10}}{m|i-j|\sigma^{3n}}|\Delta\Omega|_{-\delta,D(s-2\sigma)}|R_{ij}|_{D(s-\sigma)}, \qquad (4.14)$$

by using Lemma 9.1 below, we get

$$\|S^1\|^{\mathrm{lip}}_{a,q,p,D(s-4\sigma)\times\Pi} \lessdot \frac{e^{-9K\sigma/10}}{\sigma^{5n}}\left(\frac{M}{m}\|R^{z\bar{z}}\|_{a,q,p,D(s)\times\Pi}+\|R^{z\bar{z}}\|^{\mathrm{lip}}_{a,q,p,D(s)\times\Pi}\right). \qquad (4.15)$$

Since

$$\|S^2_{ij}\|_{D(s-2\sigma)} \lessdot \frac{e^{-9K\sigma/10}}{\sigma^n}\|R_{ij}\|_{D(s-\sigma)}, \qquad (4.16)$$

by Lemma 9.1 below, we get

$$\|S^2\|_{a,q,p,D(s-2\sigma)} \quad \lessdot \quad \frac{e^{-9K\sigma/10}}{\sigma^{2n}}\max\{|i-j| : |i^2-j^2| < C_0 K\}\|R^{z\bar{z}}\|_{a,q,p,D(s)}$$

$$\lessdot \quad \frac{C_0 K e^{-9K\sigma/10}}{\sigma^{2n}}\|R^{z\bar{z}}\|_{a,q,p,D(s)}$$

$$\lessdot \quad \frac{C_0 e^{-4K\sigma/5}}{\sigma^{2n+1}}\|R^{z\bar{z}}\|_{a,q,p,D(s)}. \qquad (4.17)$$

Again applying Lemma 9.1 to ΔS^2, we can get

$$\|S^2\|^{\mathrm{lip}}_{a,q,p,D(s-4\sigma)\times\Pi} \lessdot \frac{C_0 e^{-4K\sigma/5}}{\sigma^{2n+1}}\|R^{z\bar{z}}\|^{\mathrm{lip}}_{a,q,p,D(s)\times\Pi}. \qquad (4.18)$$

It is obvious that

$$\|S^3\|_{a,q,p,D(s-\sigma)} \lessdot \frac{e^{-9K\sigma/10}}{\sigma^n}\|R^{z\bar{z}}\|_{a,q,p,D(s)}, \qquad (4.19)$$

$$\|S^3\|^{\mathrm{lip}}_{a,q,p,D(s-\sigma)\times\Pi} \lessdot \frac{e^{-9K\sigma/10}}{\sigma^n}\|R^{z\bar{z}}\|^{\mathrm{lip}}_{a,q,p,D(s)\times\Pi}. \qquad (4.20)$$

Thus

$$\|X_{\langle\hat{R}^{z\bar{z}}z,\bar{z}\rangle}\|_{r,a,q,D(s-3\sigma,r)} \lessdot \frac{(1+C_0)e^{-4K\sigma/5}}{\sigma^{2n+2}}\|X_R\|_{r,a,q,D(s,r)}, \qquad (4.21)$$

$$\|X_{\langle\hat{R}^{z\bar{z}}z,\bar{z}\rangle}\|^{\mathrm{lip}}_{r,a,q,D(s-5\sigma,r)\times\Pi} \lessdot \frac{(1+C_0)e^{-4K\sigma/5}}{\sigma^{5n+1}}\left(\frac{M}{\alpha}\|X_R\|_{r,a,q,D(s,r)\times\Pi}+\|X_R\|^{\mathrm{lip}}_{r,a,q,D(s,r)\times\Pi}\right). \quad (4.22)$$

Therefore, from (4.21) (4.22) we get

$$\|X_{\langle\hat{R}^{z\bar{z}}z,\bar{z}\rangle}\|^{\lambda}_{r,a,q,D(s-5\sigma,r)\times\Pi} \lessdot \frac{(1+C_0)e^{-4K\sigma/5}}{\sigma^{5n+1}}\|X_R\|^{\lambda}_{r,a,q,D(s,r)\times\Pi}. \qquad (4.23)$$

For the other terms of $\hat{R}$, i.e. $\langle\hat{R}^z,z\rangle$, $\langle\hat{R}^{\bar{z}},\bar{z}\rangle$, $\langle\hat{R}^{zz}z,z\rangle$, $\langle\hat{R}^{\bar{z}\bar{z}}\bar{z},\bar{z}\rangle$, the same results - even better - than (4.23) can be obtained. Thus, we finally get the estimate for the error term $\hat{R}$:

$$\|X_{\hat{R}}\|^{\lambda}_{r,a,q,D(s-5\sigma,r)\times\Pi} \lessdot \frac{(1+C_0)e^{-4K\sigma/5}}{\sigma^{5n+1}}\|X_R\|^{\lambda}_{r,a,q,D(s,r)\times\Pi}. \qquad (4.24)$$

18

Now consider the new perturbation (4.2). By setting $R(t) = (1-t)(\hat{N} + \hat{R}) + tR$, we have

$$X_{P_+} = X_{\hat{R}} + \int_0^1 (X_F^t)^* [X_{R(t)}, X_F] dt + (X_F^1)^* (X_P - X_R). \tag{4.25}$$

We assume that

$$\|X_P\|_{r,a,q,D(s,r) \times \Pi}^{\lambda} \leq \frac{\alpha \eta^2}{B_\sigma} e^{-8C_0 \gamma_0 K s} \tag{4.26}$$

for $0 \leq \lambda \leq \alpha/M$ with some $0 < \eta < 1/16$ and $0 < \sigma < \min\{s, 1\}$, where $B_\sigma = c\sigma^{-9(n+\tau+1)}$ with c being a sufficiently large constant depending only on n, τ and $|\omega|_\Pi$. Since R is 2-order Taylor polynomial truncation in $y, z, \bar{z}$ of P, we can obtain

$$\|X_R\|_{r,a,q,D(s,r) \times \Pi}^{\lambda} \lessdot \|X_P\|_{r,a,q,D(s,r) \times \Pi}^{\lambda}, \tag{4.27}$$

$$\|X_P - X_R\|_{\eta r,a,q,D(s,4\eta r) \times \Pi}^{\lambda} \lessdot \eta \|X_P\|_{r,a,q,D(s,r) \times \Pi}^{\lambda}. \tag{4.28}$$

As in lemma 19.3 of [10], we obtain

$$\|DX_F\|_{r,a,p,p,D(s-6\sigma,r) \times \Pi}^{\lambda}, \|DX_F\|_{r,a,q,q,D(s-6\sigma,r) \times \Pi}^{\lambda} \lessdot \frac{e^{8C_0 \gamma_0 K s}}{\alpha \sigma^{3n+2\tau+3}} \|X_R\|_{r,a,q,D(s,r) \times \Pi}^{\lambda}. \tag{4.29}$$

From (3.63), (4.5), (4.9), (4.24), (4.29) and (4.27) we get

$$\|X_F\|_{r,a,p,D(s-5\sigma,r) \times \Pi}^{\lambda} \lessdot \frac{e^{8C_0 \gamma_0 K s}}{\alpha \sigma^{3n+2\tau+2}} \|X_P\|_{r,a,q,D(s,r) \times \Pi}^{\lambda}, \tag{4.30}$$

$$\|X_{\hat{N}}\|_{r,a,q,D(s-2\sigma,r) \times \Pi}^{\lambda} \lessdot \frac{1}{\sigma^{2n+2\tau+1}} \|X_P\|_{r,a,q,D(s,r) \times \Pi}^{\lambda}, \tag{4.31}$$

$$\|X_{\hat{R}}\|_{r,a,q,D(s-5\sigma,r) \times \Pi}^{\lambda} \lessdot \frac{(1+C_0)e^{-4K\sigma/5}}{\sigma^{5n+1}} \|X_P\|_{r,a,q,D(s,r) \times \Pi}^{\lambda}, \tag{4.32}$$

$$\|DX_F\|_{r,a,p,p,D(s-6\sigma,r) \times \Pi}^{\lambda}, \|DX_F\|_{r,a,q,q,D(s-6\sigma,r) \times \Pi}^{\lambda} \lessdot \frac{e^{8C_0 \gamma_0 K s}}{\alpha \sigma^{3n+2\tau+3}} \|X_P\|_{r,a,q,D(s,r) \times \Pi}^{\lambda}. \tag{4.33}$$

Moreover, together with the smallness assumptions (4.26), by properly choosing c, we get

$$\|X_F\|_{r,a,p,D(s-5\sigma,r) \times \Pi}^{\lambda}, \|DX_F\|_{r,a,p,p,D(s-6\sigma,r) \times \Pi}^{\lambda}, \|DX_F\|_{r,a,q,q,D(s-6\sigma,r) \times \Pi}^{\lambda} \leq \frac{\eta^2 \sigma}{c_0} \tag{4.34}$$

with some suitable constant $c_0 \geq 1$. Then the flow X_F^t of the vector field X_F exists on $D(s-7\sigma, r/2)$ for $-1 \leq t \leq 1$ and takes this domain into $D(s-6\sigma, r)$. Similarly, it takes $D(s-8\sigma, r/4)$ into $D(s-7\sigma, r/2)$. In the same way as (20.6) in [10], we obtain

$$\|X_F^t - \mathrm{id}\|_{r,a,p,D(s-7\sigma,r/2) \times \Pi}^{\lambda} \lessdot \|X_F\|_{r,a,p,D(s-6\sigma,r) \times \Pi}^{\lambda}, \tag{4.35}$$

$$\|DX_F^t - I\|_{r,a,p,p,D(s-8\sigma,r/4) \times \Pi}^{\lambda} \lessdot \|DX_F\|_{r,a,p,p,D(s-6\sigma,r) \times \Pi}^{\lambda}, \tag{4.36}$$

$$\|DX_F^t - I\|_{r,a,q,q,D(s-8\sigma,r/4) \times \Pi}^{\lambda} \lessdot \|DX_F\|_{r,a,q,q,D(s-6\sigma,r) \times \Pi}^{\lambda}. \tag{4.37}$$

Also in the same way as (20.7) in [10], we obtain that for any vector field Y,

$$\|(DX_F^t)^* Y\|_{\eta r,a,q,D(s-9\sigma,\eta r) \times \Pi}^{\lambda} \lessdot \|Y\|_{\eta r,a,q,D(s-7\sigma,4\eta r) \times \Pi}^{\lambda}. \tag{4.38}$$

From (4.27) (4.31) (4.32) and the assumption $(1+C_0)e^{-4K\sigma/5} \ll 1$, we get

$$\|X_{R(t)}\|^{\lambda}_{r,a,q,D(s-5\sigma,r)\times\Pi} \ll \frac{1}{\sigma^{3n+2\tau+1}} \|X_P\|^{\lambda}_{r,a,q,D(s,r)\times\Pi}. \tag{4.39}$$

Moreover, we have

$$
\begin{aligned}
&\|[X_{R(t)},X_F]\|^{\lambda}_{r,a,q,D(s-6\sigma,r/2)\times\Pi} \\
\ll\quad &\|DX_{R(t)}\|^{\lambda}_{r,a,q,p,D(s-6\sigma,r/2)\times\Pi}\|X_F\|^{\lambda}_{r,a,p,D(s-6\sigma,r/2)\times\Pi} \\
&+\|DX_F\|^{\lambda}_{r,a,q,q,D(s-6\sigma,r/2)\times\Pi}\|X_{R(t)}\|^{\lambda}_{r,a,q,D(s-6\sigma,r/2)\times\Pi} \\
\ll\quad &\frac{e^{8C_0\gamma_0 Ks}}{\alpha\sigma^{6n+4\tau+4}}\left(\|X_P\|^{\lambda}_{r,a,q,D(s,r)\times\Pi}\right)^2.
\end{aligned} \tag{4.40}
$$

Hence, also

$$\|[X_{R(t)},X_F]\|^{\lambda}_{\eta r,a,q,D(s-6\sigma,r/2)\times\Pi} \ll \frac{e^{8C_0\gamma_0 Ks}}{\alpha\eta^2\sigma^{6n+4\tau+4}}\left(\|X_P\|^{\lambda}_{r,a,q,D(s,r)\times\Pi}\right)^2. \tag{4.41}$$

Together with the estimates of $\hat{R}$ in (4.32) and $X_P - X_R$ in (4.28), we finally arrive at the estimate

$$
\begin{aligned}
&\|X_{P_+}\|^{\lambda}_{\eta r,a,q,D(s-9\sigma,\eta r)\times\Pi} \\
\leq\quad &\frac{1}{3}\left(\frac{B_\sigma e^{8C_0\gamma_0 Ks}}{\alpha\eta^2}\|X_P\|^{\lambda}_{r,a,q,D(s,r)\times\Pi} + \frac{B_\sigma e^{-4K\sigma/5}}{\alpha\eta^2} + \eta\right)\|X_P\|^{\lambda}_{r,a,q,D(s,r)\times\Pi}.
\end{aligned} \tag{4.42}
$$

This is the bound for the new perturbation.

5 Iteration and Convergence

Set $\beta' = \min\{\frac{\beta}{1+\beta},\frac{1}{4}\}$ and $\kappa = \frac{4}{3} - \frac{\beta'}{3}$. Now we give the precise set-up of iteration parameters. Let $\nu \geq 0$ be the ν-th KAM step.

$\alpha_\nu = \frac{\alpha_0}{10}(9+2^{-\nu})$, which is used to dominate the measure of removed parameters,

$m_\nu = \frac{m_0}{10}(9+2^{-\nu})$, which is used for describing the growth of external frequencies,

$E_\nu = \frac{E_0}{9}(10-2^{-\nu})$, which is used to dominate the norm of internal frequencies,

$M_{1,\nu} = \frac{M_{1,0}}{9}(10-2^{-\nu})$, $M_{2,\nu} = \frac{M_{2,0}}{9}(10-2^{-\nu})$, $M_\nu = M_{1,\nu} + M_{2,\nu}$, which are used to dominate the Lipschitz semi-norm of frequencies,

$L_\nu = \frac{L_0}{9}(10-2^{-\nu})$, which is used to dominate the inverse Lipschitz semi-norm of internal frequencies,

$J_0 = \gamma_0^{-\frac{1}{\tau+1}}$, $J_\nu = J_0^{\kappa^\nu}$, which is used for estimate of measure,

$s_\nu = s_0/2^\nu$, which dominates the width of the angle variable x,

$\sigma_\nu = s_\nu/20$, which serves as a bridge from s_ν to $s_{\nu+1}$,

$B_\nu = B_{\sigma_\nu} := c\sigma_\nu^{-9(n+\tau+1)}$, here c is a large constant only depending on n, τ and E_0,

$\varepsilon_\nu = \left(\varepsilon_0 \prod_{\mu=0}^{\nu-1} \left(\frac{B_\mu}{\alpha_\mu}\right)^{\frac{1}{3\kappa^{\mu+1}}}\right)^{\kappa^\nu}$, which dominates the size of the perturbation P_ν in ν-the KAM iteration,

$K_\nu = 5|\ln \varepsilon_\nu|/(4\sigma_\nu)$, which is the length of the truncation of Fourier series,

$\eta_\nu^3 = \varepsilon_\nu^{1-\beta'} \alpha_\nu^{-1} B_\nu, \quad r_{\nu+1} = \eta_\nu r_\nu, \quad D_\nu = D(s_\nu, r_\nu), \quad \lambda_\nu = \frac{\alpha_\nu}{M_\nu}.$

5.1 Iterative Lemma

Lemma 5.1. *Suppose that*

$$\varepsilon_0 \leq \left(\frac{\alpha_0 \gamma_0}{80}\right)^{\frac{1}{1-\beta'}} \prod_{\mu=0}^{\infty} B_\mu^{-\frac{1}{3\kappa^{\mu+1}}}, \quad \alpha_0 \leq \frac{m_0}{10}. \tag{5.1}$$

Suppose $H_\nu = N_\nu + P_\nu$ is regular on $D_\nu \times \Pi_\nu$, where N_ν is a generalized normal form with coefficients satisfying

$$|\langle k, \omega_\nu(\xi)\rangle + \langle l, \bar{\Omega}_\nu(\xi)\rangle| \geq \alpha_\nu \frac{\langle l \rangle_d}{\langle k \rangle^\tau}, \quad k \neq 0, \ |l| \leq 2, \tag{5.2}$$

$$|\langle l, \bar{\Omega}_\nu(\xi)\rangle| \geq m_\nu \langle l \rangle_d, \quad 0 < |l| \leq 2, \tag{5.3}$$

$$|\tilde{\Omega}_{\nu,j}|_{s_\nu,\tau+1} \leq (\alpha_0 - \alpha_\nu)\gamma_0 j^\delta, \quad j \geq 1, \tag{5.4}$$

$$|\omega_\nu|_{\Pi_\nu} \leq E_\nu, \quad |\omega_\nu|_{\Pi_\nu}^{\text{lip}} \leq M_{1,\nu}, \quad |\omega_\nu^{-1}|_{\omega_\nu(\Pi_\nu)}^{\text{lip}} \leq L_\nu, \tag{5.5}$$

$$|\Omega_\nu|_{-\delta,D_\nu \times \Pi_\nu}^{\text{lip}} \leq M_{2,\nu} \tag{5.6}$$

on Π_ν, and P_ν satisfies

$$\|X_{P_\nu}\|_{r_\nu,a,q,D_\nu \times \Pi_\nu}^{\lambda_\nu} \leq \varepsilon_\nu. \tag{5.7}$$

Then there exists a Lipschitz family of real analytic symplectic coordinate transformations $\Phi_{\nu+1} : D_{\nu+1} \times \Pi_\nu \to D_\nu$ satisfying

$$\|\Phi_{\nu+1} - \text{id}\|_{r_\nu,a,p,D_{\nu+1} \times \Pi_\nu}^{\lambda_\nu}, \|D\Phi_{\nu+1} - I\|_{r_\nu,a,p,p,D_{\nu+1} \times \Pi_\nu}^{\lambda_\nu}, \|D\Phi_{\nu+1} - I\|_{r_\nu,a,q,q,D_{\nu+1} \times \Pi_\nu}^{\lambda_\nu} \leq \frac{B_\nu}{\alpha_\nu} \varepsilon_\nu^{1-\beta'}, \tag{5.8}$$

and a closed subset

$$\Pi_{\nu+1} = \Pi_\nu \setminus \bigcup_{|k|>J_\nu, |l| \leq 2} \mathfrak{R}_{kl}^{\nu+1}(\alpha_{\nu+1}), \tag{5.9}$$

where

$$\mathfrak{R}_{kl}^{\nu+1}(\alpha) = \left\{ \xi \in \Pi_\nu : |\langle k, \omega_{\nu+1}(\xi)\rangle + \langle l, \bar{\Omega}_{\nu+1}(\xi)\rangle| < \alpha \frac{\langle l \rangle_d}{\langle k \rangle^\tau} \right\}, \tag{5.10}$$

such that for $H_{\nu+1} = H_\nu \circ \Phi_{\nu+1} = N_{\nu+1} + P_{\nu+1}$, the estimate

$$|\omega_{\nu+1} - \omega_\nu|_{\Pi_\nu}^{\lambda_\nu}, \quad |\Omega_{\nu+1} - \Omega_\nu|_{-\delta,D_{\nu+1} \times \Pi_\nu}^{\lambda_\nu} \leq B_\nu \varepsilon_\nu \tag{5.11}$$

holds and the same assumptions as above are satisfied with '$\nu+1$' in place of 'ν'.

Proof. Setting $C_{0,\nu} = 2E_\nu/m_\nu$, then it's obvious $C_{0,\nu} \leq 4C_{0,0}$. Thus we have

$$e^{8C_{0,\nu}\gamma_0 K_\nu s_\nu} \leq \varepsilon_\nu^{-\beta'} \tag{5.12}$$

by $K_\nu s_\nu = 20 K_\nu \sigma_\nu = 25|\ln \varepsilon_\nu|$ and choosing γ_0 small enough such that $800 C_{0,0} \gamma_0 \leq \beta'$. In view of the definition of η_ν, namely $\eta_\nu^3 = \varepsilon_\nu^{1-\beta'} \alpha_\nu^{-1} B_\nu$, the smallness condition (4.26), namely $\varepsilon_\nu \leq \frac{\alpha_\nu \eta_\nu^2}{B_\nu} e^{-8 C_{0,\nu} \gamma_0 K_\nu s_\nu}$, is satisfied if

$$\varepsilon_\nu^{1-\beta'} \leq \frac{\alpha_\nu}{B_\nu}. \tag{5.13}$$

To verify the last inequality we argue as follows. As B_ν and α_ν^{-1} are increasing with ν,

$$\left(\frac{B_\nu}{\alpha_\nu}\right)^{\frac{1}{1-\beta'}} = \left(\frac{B_\nu}{\alpha_\nu}\right)^{\frac{1}{3(\kappa-1)}} = \left(\prod_{\mu=\nu}^{\infty} \left(\frac{B_\nu}{\alpha_\nu}\right)^{\frac{1}{3\kappa^{\mu+1}}}\right)^{\kappa^\nu} \leq \left(\prod_{\mu=\nu}^{\infty} \left(\frac{B_\mu}{\alpha_\mu}\right)^{\frac{1}{3\kappa^{\mu+1}}}\right)^{\kappa^\nu}. \tag{5.14}$$

By the definition of ε_ν above, the bound $\alpha_\nu \geq 9\alpha_0/10$ and the smallness condition on ε_0 in (5.1),

$$\varepsilon_\nu^{1-\beta'} \frac{B_\nu}{\alpha_\nu} \leq \left(\varepsilon_0 \prod_{\mu=0}^{\infty} \left(\frac{B_\mu}{\alpha_\mu}\right)^{\frac{1}{3\kappa^{\mu+1}}}\right)^{\kappa^\nu(1-\beta')} \leq \left(\frac{\gamma_0}{72}\right)^{\kappa^\nu} \leq 1. \tag{5.15}$$

So the smallness condition (4.26) is satisfied for each $\nu \geq 0$. In particular, noticing $\kappa \geq 5/4$, we have

$$\varepsilon_\nu^{1-\beta'} \frac{B_\nu}{\alpha_\nu} \leq \frac{\gamma_0}{2^{\nu+6}}. \tag{5.16}$$

Now there exists a coordinate transformation $\Phi_{\nu+1} : D_{\nu+1} \times \Pi_\nu \to D_\nu$ taking H_ν into $H_{\nu+1}$. Moreover, (5.8) is obtained by (4.30) (4.33) (4.35)-(4.37), and (5.11) is obtained by (4.31). More explicitly, (5.11) is written as

$$|\omega_{\nu+1} - \omega_\nu|_{\Pi_\nu}, \quad |\Omega_{\nu+1} - \Omega_\nu|_{-\delta, D_{\nu+1} \times \Pi_\nu} \leq B_\nu \varepsilon_\nu, \tag{5.17}$$

$$|\omega_{\nu+1} - \omega_\nu|_{\Pi_\nu}^{\mathrm{lip}}, \quad |\Omega_{\nu+1} - \Omega_\nu|_{-\delta, D_{\nu+1} \times \Pi_\nu}^{\mathrm{lip}} \leq \frac{M_\nu}{\alpha_\nu} B_\nu \varepsilon_\nu. \tag{5.18}$$

In view of (5.16)-(5.18), by choosing γ_0 properly small, (5.3)-(5.6) are satisfied with '$\nu+1$' in place of 'ν'. As to the diophantine conditions, in view of the definition of $\Pi_{\nu+1}$ in (5.9), only the case $\langle k \rangle \leq J_\nu$ remains to verify. In this case, by the definition of J_ν, we have

$$B_\nu \varepsilon_\nu \leq \frac{\alpha_\nu \gamma_0^{\kappa^\nu}}{2^{\nu+6}} \leq \frac{\alpha_\nu - \alpha_{\nu+1}}{3 J_\nu^{\tau+1}}. \tag{5.19}$$

Hence, for $k \neq 0$ and $\langle k \rangle \leq J_\nu$,

$$
\begin{aligned}
|\langle k, \omega_{\nu+1} - \omega_\nu \rangle + \langle l, \bar{\Omega}_{\nu+1} - \bar{\Omega}_\nu \rangle| &\leq |k||\omega_{\nu+1} - \omega_\nu| + 2\langle l \rangle_d |\bar{\Omega}_{\nu+1} - \bar{\Omega}_\nu|_{-\delta} \\
&\leq 3|k|\langle l \rangle_d B_\nu \varepsilon_\nu \\
&\leq (\alpha_\nu - \alpha_{\nu+1}) \frac{|k|\langle l \rangle_d}{J_\nu^{\tau+1}} \\
&\leq (\alpha_\nu - \alpha_{\nu+1}) \frac{\langle l \rangle_d}{\langle k \rangle^\tau}
\end{aligned}
\tag{5.20}
$$

on $D_{\nu+1} \times \Pi_\nu$. This completes the proof of (5.2) with '$\nu+1$' in place of 'ν'. On the other hand, from (4.42) we get

$$
\begin{aligned}
\|X_{P_{\nu+1}}\|_{r_{\nu+1}, a, q, D_{\nu+1} \times \Pi_\nu}^{\lambda_{\nu+1}} &\leq \frac{1}{3}\left(\frac{B_\nu e^{8 C_{0,\nu} \gamma_0 K_\nu s_\nu}}{\alpha_\nu \eta_\nu^2} \varepsilon_\nu + \frac{B_\nu e^{-4 K_\nu \sigma_\nu/5}}{\alpha_\nu \eta_\nu^2} + \eta_\nu\right) \varepsilon_\nu \\
&\leq \frac{1}{3}\left(\frac{B_\nu}{\alpha_\nu \eta_\nu^2} \varepsilon_\nu^{1-\beta'} + \frac{B_\nu}{\alpha_\nu \eta_\nu^2} \varepsilon_\nu^{1-\beta'} + \eta_\nu\right) \varepsilon_\nu \\
&= \left(\frac{B_\nu}{\alpha_\nu}\right)^{\frac{1}{3}} \varepsilon_\nu^\kappa \\
&= \varepsilon_{\nu+1}.
\end{aligned}
\tag{5.21}
$$

22

This completes the proof of iterative lemma. □

5.2 Convergence

We are now in a position to prove the KAM theorem. To apply iterative lemma with $\nu = 0$, set

$$N_0 = N, \quad P_0 = P, \quad s_0 = s, \quad r_0 = r$$

and similarly $E_0 = E$, $L_0 = L$, $M_{1,0} = M_1$, $M_{2,0} = M_2$, $m_0 = m$, $\alpha_0 = \alpha$ and $\lambda_0 = \lambda = \alpha/M$. Define γ in the KAM theorem by setting

$$\gamma = \gamma_0 \gamma_s, \quad \gamma_s = \frac{1}{80} \Big(\prod_{\mu=0}^{\infty} B_\mu^{-\frac{1}{3\kappa^{\mu+1}}} \Big)^{1-\beta'}, \tag{5.22}$$

where γ_0 is the same parameter as before and γ_s only depends on n, τ, E, s, β. The smallness condition (5.1) of the iterative lemma is then satisfied by the assumption of the KAM theorem:

$$\varepsilon_0 := \|X_{P_0}\|_{r_0,a,q,D_0 \times \Pi_0}^{\lambda_0} \leq (\alpha \gamma)^{1+\beta} \leq (\alpha_0 \gamma_0 \gamma_s)^{\frac{1}{1-\beta'}}. \tag{5.23}$$

The small divisor conditions (5.2) are satisfied by setting

$$\Pi_0 = \Pi \setminus \bigcup_{(k,l) \neq (0,0), |l| \leq 2} \mathfrak{R}_{kl}^0(\alpha_0), \tag{5.24}$$

and the other conditions (5.3)-(5.6) about the unperturbed frequencies are obviously true.

Hence, the iterative lemma applies, and we obtain a decreasing sequence of domains $D_\nu \times \Pi_\nu$ and a sequence of transformations

$$\Phi^\nu = \Phi_1 \circ \cdots \circ \Phi_\nu : \quad D_\nu \times \Pi_{\nu-1} \to D_0,$$

such that $H \circ \Phi^\nu = N_\nu + P_\nu$ for $\nu \geq 1$. Moreover, the estimates (5.8) and (5.11) hold.

Shorten $\|\cdot\|_{r,a,p}$ as $\|\cdot\|_r$ and consider the operator norm

$$\|L\|_{r,\tilde{r}} = \sup_{W \neq 0} \frac{\|LW\|_r}{\|W\|_{\tilde{r}}}.$$

For $r \geq \tilde{r}$, these norms satisfy $\|AB\|_{r,\tilde{r}} \leq \|A\|_{r,r} \|B\|_{\tilde{r},\tilde{r}}$ since $\|W\|_r \leq \|W\|_{\tilde{r}}$. For $\nu \geq 1$, by the chain rule, using (5.8) (5.16), we get

$$\|D\Phi^\nu\|_{r_0,r_\nu,D_\nu \times \Pi_{\nu-1}} \leq \prod_{\mu=1}^{\nu} \|D\Phi_\mu\|_{r_\mu,r_\mu,D_\mu \times \Pi_{\mu-1}} \leq \prod_{\mu=1}^{\infty} \Big(1 + \frac{1}{2^{\mu+6}} \Big) \leq 2, \tag{5.25}$$

$$\begin{aligned}
\|D\Phi^\nu\|_{r_0,r_\nu,D_\nu \times \Pi_{\nu-1}}^{\text{lip}} &\leq \sum_{\mu=1}^{\nu} \|D\Phi_\mu\|_{r_\mu,r_\mu,D_\mu \times \Pi_{\mu-1}}^{\text{lip}} \prod_{1 \leq \rho \leq \nu, \rho \neq \mu} \|D\Phi_\rho\|_{r_\rho,r_\rho,D_\rho \times \Pi_{\rho-1}} \\
&\leq 2 \sum_{\mu=1}^{\nu} \|D\Phi_\mu - I\|_{r_\mu,r_\mu,D_\mu \times \Pi_{\mu-1}}^{\text{lip}} \leq 2 \prod_{\mu=1}^{\infty} \frac{M_\mu}{\alpha_\mu} \frac{1}{2^{\mu+6}} \leq \frac{M_0}{\alpha_0}. \quad (5.26)
\end{aligned}$$

Thus, with the mean value theorem we obtain

$$\begin{aligned}
&\|\Phi^{\nu+1} - \Phi^\nu\|_{r_0,D_{\nu+1} \times \Pi_\nu} \\
&\leq \|D\Phi^\nu\|_{r_0,r_\nu,D_\nu \times \Pi_{\nu-1}} \|\Phi_{\nu+1} - \text{id}\|_{r_\nu,D_{\nu+1} \times \Pi_\nu} \leq 2\|\Phi_{\nu+1} - \text{id}\|_{r_\nu,D_{\nu+1} \times \Pi_\nu}, \quad (5.27)
\end{aligned}$$

$$\|\Phi^{v+1}-\Phi^v\|_{r_0,D_{v+1}\times\Pi_v}^{\text{lip}}$$
$$\leq \|D\Phi^v\|_{r_0,r_v,D_v\times\Pi_{v-1}}^{\text{lip}}\|\Phi_{v+1}-\text{id}\|_{r_v,D_{v+1}\times\Pi_v}+\|D\Phi^v\|_{r_0,r_v,D_v\times\Pi_{v-1}}\|\Phi_{v+1}-\text{id}\|_{r_v,D_{v+1}\times\Pi_v}^{\text{lip}}$$
$$\leq \frac{M_0}{\alpha_0}\|\Phi_{v+1}-\text{id}\|_{r_v,D_{v+1}\times\Pi_v}+2\|\Phi_{v+1}-\text{id}\|_{r_v,D_{v+1}\times\Pi_v}^{\text{lip}}. \tag{5.28}$$

It follows that
$$\|\Phi^{v+1}-\Phi^v\|_{r_0,D_{v+1}\times\Pi_v}^{\lambda_0}\leq 3\|\Phi_{v+1}-\text{id}\|_{r_v,D_{v+1}\times\Pi_v}^{\lambda_v}. \tag{5.29}$$

From (5.29) (5.8), we get
$$\|\Phi^{v+1}-\Phi^v\|_{r_0,D_{v+1}\times\Pi_v}^{\lambda_0}\leq 3\frac{B_v}{\alpha_v}\varepsilon_v^{1-\beta'}. \tag{5.30}$$

For every non-negative integer multi-index $k=(k_1,\ldots,k_n)$, by Cauchy's estimate we have
$$\|\partial_x^k(\Phi^{v+1}-\Phi^v)\|_{r_0,D_{v+2}\times\Pi_v}^{\lambda_0}\leq 3\frac{B_v}{\alpha_v}\varepsilon_v^{1-\beta'}\frac{k_1!\cdots k_n!}{(\frac{s_0}{2^{v+2}})^{|k|}}, \tag{5.31}$$

the right side of which super-exponentially decay with v. This shows that Φ^v converge uniformly on $D_*\times\Pi_\alpha$, where $D_*=\mathbb{T}^n\times\{0\}\times\{0\}\times\{0\}$ and $\Pi_\alpha=\bigcap_{v\geq 0}\Pi_v$, to a Lipschitz continuous family of smooth torus embeddings
$$\Phi:\mathbb{T}^n\times\Pi_\alpha\to\mathscr{P}^{a,p},$$

for which the estimate (1.9) holds. Similarly, the frequencies ω_v converge uniformly on Π_α to a Lipschitz continuous limit ω_*, and the frequencies Ω_v converge uniformly on $D_*\times\Pi_\alpha$ to a regular limit Ω_*, with estimate (1.10) holding. Moreover, $X_H\circ\Phi=D\Phi\cdot X_{N_*}$ on D_* for each $\xi\in\Pi_\alpha$, where N_* is the generalized normal form with frequencies ω_* and Ω_*. Thus, the embedded tori are invariant under the perturbed Hamiltonian flow, and the flow on them is linear. Now it only remains to prove the claim about the set $\Pi\setminus\Pi_\alpha$, which is the subject of the next section.

6 Measure Estimate

6.1 Proof of (1.8)

We know
$$\Pi\setminus\Pi_\alpha=\bigcup_{v\geq 0}\bigcup_{|k|>J_{v-1},|l|\leq 2}\mathfrak{R}_{kl}^v, \tag{6.1}$$

where $J_{-1}=0$, $J_v=\gamma_0^{-\kappa^v/(\tau+1)}$ and
$$\mathfrak{R}_{kl}^v=\left\{\xi\in\Pi_{v-1}:|\langle k,\omega_v(\xi)\rangle+\langle l,\bar{\Omega}_v(\xi)\rangle|<\alpha_v\frac{\langle l\rangle_d}{\langle k\rangle^\tau}\right\} \tag{6.2}$$

with $\Pi_{-1}=\Pi$. Here, ω_v and $\bar{\Omega}_v$ are defined and Lipschitz continuous on Π_{v-1}, and $\omega_0=\omega$, $\bar{\Omega}_0=\Omega$ are the frequencies of the unperturbed system.

Lemma 6.1. *If γ_0 is sufficiently small and $\tau\geq n+1+2/(d-1)$, then*
$$|\Pi\setminus\Pi_\alpha|\leq c\alpha, \tag{6.3}$$

where $c>0$ is a constant depends on n,d,E,L and m.

Proof. We only need to give the proof of the most difficult case that l has two non-zero components of opposite sign. In this case, rewriting $\mathfrak{R}^v_{kl}$ as

$$\mathfrak{R}^v_{kij} = \left\{ \xi \in \Pi_{v-1} : |\langle k, \omega_v(\xi) \rangle + \bar{\Omega}_{v,i}(\xi) - \bar{\Omega}_{v,j}(\xi)| < \alpha_v \frac{|i^d - j^d|}{\langle k \rangle^\tau} \right\}, \quad i \neq j, \qquad (6.4)$$

we only need to estimate the measure of

$$\Xi_\alpha := \bigcup_{v \geq 0} \bigcup_{|k| > J_{v-1}, i \neq j} \mathfrak{R}^v_{kij}. \qquad (6.5)$$

To estimate the measure of $\bigcup_{|k| > J_{v-1}, i \neq j} \mathfrak{R}^v_{kij}$, we introduce the perturbed frequencies $\zeta = \omega_v(\xi)$ as parameters over the domain $Z := \omega_v(\Pi_v)$ and consider the resonance zones $\dot{\mathfrak{R}}^v_{kij} = \omega_v(\mathfrak{R}^v_{kij})$ in Z. Regarding $\bar{\Omega}_v$ as function of ζ, then from the iterative lemma above, we know

$$|\zeta| \leq E_v \leq \frac{10}{9} E_0, \quad |\xi|^{\text{lip}}_Z \leq L_v \leq \frac{10}{9} L_0, \qquad (6.6)$$

$$|\bar{\Omega}_{v,i} - \bar{\Omega}_{v,j}|_Z \geq m_v |i^d - j^d| \geq \frac{9}{10} m_0 |i^d - j^d|, \qquad (6.7)$$

$$|\bar{\Omega}_{v,j}|^{\text{lip}}_Z \leq L_v M_{2,v} j^\delta \leq (\frac{10}{9})^2 L_0 M_{2,0} j^\delta, \qquad (6.8)$$

where $E_0, L_0, m_0, M_{2,0}$ are just E, L, m, M_2 in theorem 1.1.

Now we consider a fixed $\dot{\mathfrak{R}}^v_{kij}$.

If $|k| < \frac{9 m_v}{10 E_v} |i^d - j^d|$, we get $|\langle k, \zeta \rangle| < \frac{9 m_v}{10} |i^d - j^d|$. In view of (6.7) and $\alpha_v \leq \frac{m_v}{10}$, we know $\dot{\mathfrak{R}}^v_{kij}$ is empty.

If $|k| \geq \frac{9 m_v}{10 E_v} |i^d - j^d|$, we have $|k| \geq \frac{1}{2} (\frac{9}{10})^3 \frac{m}{E} (i^\delta + j^\delta)$. Fix $w_1 \in \{-1, 1\}^n$ such that $|k| = k \cdot w_1$ and write $\zeta = a w_1 + w_2$ with $w_1 \perp w_2$. As a function of a, for $t > s$,

$$\langle k, \zeta \rangle |^t_s = |k|(t - s), \qquad (6.9)$$

$$(|\bar{\Omega}_{v,i} - \bar{\Omega}_{v,j}|)|^t_s \leq (\frac{10}{9})^2 L M_2 (i^\delta + j^\delta)(t - s). \qquad (6.10)$$

Thus

$$
\begin{aligned}
(\langle k, \zeta \rangle + \bar{\Omega}_{v,i} - \bar{\Omega}_{v,j})|^t_s &\geq |k|(t - s)\left(1 - (\frac{9}{10})^2 L M_2 (i^\delta + j^\delta)/|k|\right) \\
&\geq |k|(t - s)\left(1 - 2(\frac{10}{9})^5 \frac{E L M_2}{m}\right) \\
&\geq \frac{1}{10} |k|(t - s),
\end{aligned}
\qquad (6.11)
$$

by using the assumption $4 E L M_2 \leq m$ in theorem 1.1 and the fact $(9/10)^6 > 1/2$. Therefore, we get

$$|\dot{\mathfrak{R}}^v_{kij}| \leq 10 (\text{diam} Z)^{n-1} \alpha_v \frac{|i^d - j^d|}{\langle k \rangle^{\tau+1}} \leq 10 (\frac{20}{9} E)^{n-1} \alpha \frac{(\frac{10}{9})^3 \frac{E}{m}}{\langle k \rangle^\tau}. \qquad (6.12)$$

Going back to the original parameter domain Π_v by the inverse frequency map ω^{-1}, we get

$$|\mathfrak{R}^v_{kij}| \leq c_1 \frac{\alpha}{\langle k \rangle^\tau}, \qquad (6.13)$$

25

where $c_1 = 5(10/9)^{2n+2}(2EL)^n/m$. Consequently, we have that for any $|k| > J_{\nu-1}$,

$$|\bigcup_{i \neq j} \mathfrak{R}_{kij}^\nu| \leq \bigcup_{|i^d - j^d| \leq (10/9)^3(E/m)|k|} |\mathfrak{R}_{kij}^\nu| \leq c_2 \frac{\alpha}{\langle k \rangle^{\tau - \frac{2}{d-1}}}, \tag{6.14}$$

where $c_2 = c_1 \big(2(10/9)^3(E/m)\big)^{2/(d-1)}$. Moreover, since $\tau \geq n+1+2/(d-1)$, we have

$$| \bigcup_{|k| > J_{\nu-1}, i \neq j} \mathfrak{R}_{kij}^\nu| \leq c_2 c_3 \frac{\alpha}{1 + J_{\nu-1}}, \tag{6.15}$$

where $c_3 > 0$ depends only on n. The sum of the latter inequality over all ν converges, and we finally obtain the estimate of lemma 6.1. $\qquad \square$

6.2 Proof of Theorem 1.2

Since the proof of the case $\delta = \tilde{d} < d - 1$ can be found in [7] or [10], we assume $\delta = \tilde{d} = d - 1$ in the following.

From the assumption (D*) we know $\bar{\Omega}_0 = \Omega^1 + \Omega^2$ with $|\Omega^1|^{\text{lip}}_{-\delta_0,\Pi} \leq M_3$ and $|\Omega^2|^{\text{lip}}_{-\delta,\Pi} \leq M_4$. Thus in ν-th KAM step, setting $\bar{\Omega}_\nu = \Omega^1 + \Omega_\nu^2$, we know $\Omega_0^2 = \Omega^2$ and

$$|\Omega_\nu^2|^{\text{lip}}_{-\delta,\Pi} \leq M_{4,\nu}, \tag{6.16}$$

where the iteration parameter $M_{4,\nu} = \frac{M_4}{9}(10 - 2^{-\nu})$.

In the following we consider the excluded set of parameters under the assumption (D*) instead of (D). Our aim is to verify the conclusion (1*) in Theorem 1.2.

In the same way as [10], we write $\Pi \setminus \Pi_\alpha = \Xi_\alpha^1 + \Xi_\alpha^2$, where

$$\Xi_\alpha^1 = \bigcup_{0 < |k| \leq J_0, |l| \leq 2} \mathfrak{R}_{kl}^0, \tag{6.17}$$

$$\Xi_\alpha^2 = \bigcup_{\nu \geq 0} \bigcup_{|k| > \max(J_0, J_{\nu-1}), |l| \leq 2} \mathfrak{R}_{kl}^\nu. \tag{6.18}$$

Since $|k| \leq J_0$ and for each k there are only finitely many l for which $\mathfrak{R}_{kl}^0$ is not empty, the set Ξ_α^1 is a finite union of resonance zones. For each of its members we know that $|\mathfrak{R}_{kl}^0| \to 0$ as $\alpha \to 0$ for $l \neq 0$ by the first part of assumption (D*), and for $l = 0$ by elementary volume estimates. Thus $|\Xi_\alpha^1| \to 0$ as $\alpha \to 0$. In the remainder of this section we estimate the measure of Ξ_α^2.

Lemma 6.2. *If γ_0 is sufficiently small and $\tau \geq n+1+2/(d-1)$, then*

$$|\Xi_\alpha^2| \leq c\alpha, \tag{6.19}$$

where $c > 0$ is a constant depends on n, d, E, L and m.

Proof. As in Lemma 6.1, we only need to give the proof of the most difficult case that l has two non-zero components of opposite sign. Seeing (6.4) for the definition of $\mathfrak{R}_{kij}^\nu$, we only need to estimate the measure of

$$\bigcup_{\nu \geq 0} \bigcup_{|k| > \max(J_0, J_{\nu-1}), i \neq j} \mathfrak{R}_{kij}^\nu. \tag{6.20}$$

26

For the measure of $\bigcup_{|k|>\max(J_0,J_{v-1}),i\neq j}\mathfrak{R}^v_{kij}$, we introduce the perturbed frequencies $\zeta = \omega_v(\xi)$ as parameters over the domain $Z = \omega_v(\Pi_v)$ and consider the resonance zones $\dot{\mathfrak{R}}^v_{kij} = \omega_v(\mathfrak{R}^v_{kij})$ in Z. The estimate (6.6) (6.7) still hold true. Moreover, for $\bar{\Omega}_{v,j} = \Omega^1_j + \Omega^2_{v,j}$, we have the estimate

$$|\Omega^1_j|^{\text{lip}}_Z \leq L_v M_3 j^{\delta_0} \leq \frac{10}{9} L M_3 j^{\delta_0}, \tag{6.21}$$

$$|\Omega^2_{v,j}|^{\text{lip}}_Z \leq L_v M_{4,v} j^{\delta} \leq (\frac{10}{9})^2 L M_4 j^{\delta}. \tag{6.22}$$

Let $\delta_1 = \min(\delta - \delta_0, \delta)$ and $\delta_* = \max(1, \delta_0/\delta_1)$,

$$L_* = (\frac{10}{9})^4 \frac{2ELM_3}{m}, \quad J_* = \frac{20}{9} L M_3 L^{\delta_*}_*.$$

Choose γ_0 sufficiently small so that

$$J_0 = \gamma_0^{-\frac{1}{\tau+1}} \geq J_*, \tag{6.23}$$

and thus $J_v \geq J_0 \geq J_*$ for all $v \geq 0$. Now we consider a fixed $\dot{\mathfrak{R}}^v_{kij}$ with $|k| > J_*$.

If $|k| < \frac{9m_v}{10E_v}|i^d - j^d|$, then $\dot{\mathfrak{R}}^v_{kij}$ is empty.

If $|k| \geq \frac{9m_v}{10E_v}|i^d - j^d|$, we have

$$|k| \geq \frac{1}{2}(\frac{9}{10})^3 \frac{m}{E}(i^{\delta} + j^{\delta}). \tag{6.24}$$

Fix $w_1 \in \{-1,1\}^n$ such that $|k| = k \cdot w_1$ and write $\zeta = aw_1 + w_2$ with $w_1 \perp w_2$. As a function of a, for $t > s$,

$$\langle k, \zeta \rangle|^t_s = |k|(t - s), \tag{6.25}$$

$$(|\Omega^1_i - \Omega^1_j|)|^t_s \leq \frac{10}{9} L M_3 (i^{\delta_0} + j^{\delta_0})(t - s), \tag{6.26}$$

$$(|\Omega^2_{v,i} - \Omega^2_{v,j}|)|^t_s \leq (\frac{10}{9})^2 L M_4 (i^{\delta} + j^{\delta})(t - s). \tag{6.27}$$

We claim

$$\frac{10}{9} L M_3 (i^{\delta_0} + j^{\delta_0}) \leq \frac{|k|}{2}. \tag{6.28}$$

In fact,

(1) If $i^{\delta_1} + j^{\delta_1} \geq L_*$, in view of (6.24) and

$$2(i^{\delta_0} + j^{\delta_0})(i^{\delta_1} + j^{\delta_1}) \leq i^{\delta} + j^{\delta}, \tag{6.29}$$

then (6.28) follows from

$$|k| \geq (\frac{9}{10})^3 \frac{m}{E}(i^{\delta_0} + j^{\delta_0})(i^{\delta_1} + j^{\delta_1}) \geq (\frac{9}{10})^3 \frac{m}{E}(i^{\delta_0} + j^{\delta_0})L_* = \frac{20}{9} L M_3 (i^{\delta_0} + j^{\delta_0}); \tag{6.30}$$

(2) If $i^{\delta_1} + j^{\delta_1} < L_*$, then

$$\frac{10}{9} L M_3 (i^{\delta_0} + j^{\delta_0}) \leq \frac{10}{9} L M_3 (i^{\delta_1} + j^{\delta_1})^{\delta_*} \leq \frac{10}{9} L M_3 L^{\delta_*}_* = \frac{J_*}{2} \leq \frac{|k|}{2}. \tag{6.31}$$

27

This completes the proof of (6.28).

In view of (1.11) in assumption (D*), (6.24) and the fact $(9/10)^6 > 1/2$, we get

$$(\frac{10}{9})^2 LM_4(i^\delta + j^\delta) \leq \frac{9}{20}|k|. \tag{6.32}$$

Thus, from (6.25)-(6.28) and (6.32),

$$(\langle k, \zeta \rangle + \bar{\Omega}_{v,i} - \bar{\Omega}_{v,j})|_s^t \geq \frac{1}{20}|k|(t-s). \tag{6.33}$$

In the same process as (6.12)-(6.15) in Lemma 6.1, we finally obtain the estimate of lemma 6.2. $\quad\square$

7   Application to Derivative Nonlinear Schrödinger Equation

In this section, using Theorem 1.1, we show the existence of quasi-periodic solutions for a class of derivative nonlinear Schrödinger equations subject to Dirichlet boundary conditions

$$\begin{cases} \mathbf{i}u_t + u_{xx} - M_\sigma u + \mathbf{i}f(u,\bar{u})u_x = 0, & (t,x) \in \mathbb{R} \times [0,\pi], \\ u(t,0) = 0 = u(t,\pi), \end{cases} \tag{7.1}$$

where M_σ is a real Fourier multiplier,

$$M_\sigma \sin jx = \sigma_j \sin jx, \quad \sigma_j \in \mathbb{R}, \quad j \geq 1, \tag{7.2}$$

and f is analytic in some neighborhood of the origin in $\mathbb{C}^2$ with

$$\overline{f(u,\bar{u})} = f(u,\bar{u}), \quad f(-u,-\bar{u}) = -f(u,\bar{u}).$$

We study this equation as a Hamiltonian system on some suitable phase space $\mathscr{P}$. As the phase space one may take, for example, the usual Sobolev space $\mathscr{H}_0^2([0,\pi])$. The same as in [9], introducing the inner product

$$\langle u, v \rangle = \text{Re} \int_0^\pi u\bar{v}dx, \tag{7.3}$$

then (7.1) can be written in Hamiltonian form

$$\frac{\partial u}{\partial t} = -\mathbf{i}\nabla H(u), \tag{7.4}$$

$$H(u) = \frac{1}{2}\int_0^\pi |u_x|^2 dx + \frac{1}{2}\int_0^\pi (M_\sigma u)\bar{u}dx + \frac{1}{2}\int_0^\pi g(u,\bar{u})u_x dx, \tag{7.5}$$

where the gradient ∇ is defined with respect to $\langle \cdot, \cdot \rangle$, and $g(z_1, z_2) = -\mathbf{i}\int_0^{z_2} f(z_1, \zeta)d\zeta$. To write it in infinitely many coordinates, we make the ansatz

$$u = \mathscr{S}q = \sum_{j \geq 1} q_j \phi_j, \quad \phi_j = \sqrt{\frac{2}{\pi}}\sin jx, \quad j \geq 1. \tag{7.6}$$

The coordinates are taken from the Hilbert space $\ell_{\mathbb{C}}^{a,p}$ of all complex-valued sequences $q = (q_1, q_2, \cdots)$ with

$$\|q\|_{a,p}^2 = \sum_{j \geq 1} |q_j|^2 j^{2p} e^{2aj} < \infty. \tag{7.7}$$

We fix $a \geq 0$ and $p > \frac{3}{2}$ in the following. Then (7.4) can be rewritten as

$$\dot{q}_j = -2\mathbf{i}\frac{\partial H}{\partial \bar{q}_j}, \quad j \geq 1 \tag{7.8}$$

with the Hamiltonian

$$\begin{aligned} H(q) &= \Lambda + G \\ &= \frac{1}{2}\sum_{j\geq 1}(j^2 + \sigma_j)|q_j|^2 + \frac{1}{2}\int_0^\pi g(\mathscr{S}q, \mathscr{S}\bar{q})(\mathscr{S}q)_x dx. \end{aligned} \tag{7.9}$$

The perturbation term G has the following properties:

Lemma 7.1. *For $a \geq 0$ and $p > \frac{3}{2}$, the function G is analytic in some neighborhood of the origin in $\ell_{\mathbb{C}}^{a,p}$ with real value, and the Hamiltonian vector field X_G is an analytic map from some neighborhood of the origin in $\ell_{\mathbb{C}}^{a,p}$ into $\ell_{\mathbb{C}}^{a,p-1}$ with*

$$\|X_G\|_{a,p-1} = O(\|q\|_{a,p}^2). \tag{7.10}$$

Proof. Set $h(z_1, z_2) = \int_0^{z_2}\int_0^{z_1} f(\eta, \zeta)d\eta d\zeta$. From $\overline{f(u,\bar{u})} = f(u,\bar{u})$, we know $h(u,\bar{u})$ is real. In view of the definition of $g(z_1, z_2)$ above,

$$\overline{g(u,\bar{u})} = -\mathbf{i}\frac{\overline{\partial h(u,\bar{u})}}{\partial u} = \mathbf{i}\frac{\partial h(u,\bar{u})}{\partial \bar{u}}. \tag{7.11}$$

Therefore,

$$0 = -\mathbf{i}\int_0^\pi \frac{dh}{dx}dx = -\mathbf{i}\int_0^\pi\left(\frac{\partial h}{\partial u}u_x + \frac{\partial h}{\partial \bar{u}}\bar{u}_x\right)dx = \int_0^\pi\left(g(u,\bar{u})u_x - \overline{g(u,\bar{u})}\bar{u}_x\right)dx = 2(G - \bar{G}). \tag{7.12}$$

This illustrates that G is real valued. In view of G in (7.9), we have

$$\frac{\partial G}{\partial \bar{q}_j} = -\frac{\mathbf{i}}{2}\int_0^\pi f(u,\bar{u})u_x\phi_j dx, \quad u = \mathscr{S}q. \tag{7.13}$$

From $f(-u,-\bar{u}) = -f(u,\bar{u})$, we know that $f(0,0) = 0$ and $f(u,\bar{u})u_x$ can be expanded as Fourier sine series. Thus the components of the gradient $G_{\bar{q}}$ are the Fourier sine coefficients of $f(u,\bar{u})u_x$. Now the estimate (7.10) can be obtained in the same way as that of Lemma 3 in [9]. □

Pick a set

$$J = \{j_1 < j_2 < \cdots < j_n\} \subset \mathbb{N}$$

as n basic modes. We assume

$$\begin{cases} \sigma_{j_b} = \xi_b, & b = 1, \cdots, n, \\ \sigma_j = 0, & j \notin J, \end{cases} \tag{7.14}$$

and take $\xi := (\xi_1, \cdots, \xi_n) \in \Pi \subset \mathbb{R}^n$ as parameters, where Π is a closed bounded set of positive Lebesgue measure. We introduce symplectic polar and real coordinates (x, y, u, v) by setting

$$\begin{cases} q_{j_b} = \sqrt{2(I_b + y_b)}e^{-\mathbf{i}x_b}, & b = 1, \cdots, n, \\ q_j = u_j - \mathbf{i}v_j, & j \notin J, \end{cases} \tag{7.15}$$

where $I = (I_1, \cdots, I_n)$ is fixed. Then we have

$$-\frac{\mathbf{i}}{2}\sum_{j\geq 1}dq_j \wedge d\bar{q}_j = \sum_{1\leq b\leq n}dy_b \wedge dx_b + \sum_{j\notin J}du_j \wedge dv_j. \tag{7.16}$$

Therefore, up to a constant term, the Hamiltonian (7.9) can be rewritten as

$$
\begin{aligned}
H &= N + P \\
&= \sum_{1 \leq b \leq n} \omega_b y_b + \frac{1}{2} \sum_{j \notin J} \Omega_j (u_j^2 + v_j^2) + G(q(x,y,u,v))
\end{aligned}
\tag{7.17}
$$

with symplectic structure $\sum_{1 \leq b \leq n} dy_b \wedge dx_b + \sum_{j \notin J} du_j \wedge dv_j$, where

$$
\omega_b = j_b^2 + \xi_b, \quad b = 1, \cdots, n,
\tag{7.18}
$$

$$
\Omega_j = j^2, \quad j \notin J.
\tag{7.19}
$$

It's obvious that the tangential frequencies $\omega = (\omega_1, \cdots, \omega_n)$ and normal frequencies $\Omega = (\Omega_j : j \notin J)$ satisfy the assumptions (A) (B) (C) (D) in Theorem 1.1. From Lemma 7.1 above, we know there exists $r > 0$, such that, for every fixed I satisfying $|I| = O(r^2)$, the Hamiltonian vector field X_P is real analytic from $D(1,r)$ to $\mathscr{P}^{a,p-1}$ with

$$
\|X_P\|_{r,a,p-1,D(1,r)} = O(r).
\tag{7.20}
$$

Moreover, since X_P is independent of ξ, we know

$$
\|X_P\|_{r,a,p-1,D(1,r) \times \Pi}^{\text{lip}} = 0.
\tag{7.21}
$$

Therefore, defining ε as (1.7), we have

$$
\varepsilon = O(r).
\tag{7.22}
$$

We simply assume $|\Pi| = O(1)$ and fix $\beta = 1/5$. Then by letting $\alpha = O(r^{2/3})$, the inequality in (1.7) is satisfied when r is small enough. Now Theorem 1.1 yields the following

Theorem 7.2. *Consider a family of derivative nonlinear Schrödinger equation (7.1) parameterized by the Fourier multiplier M_σ with $\sigma = \sigma(\xi)$ defined by (7.14). Then for any $0 < \varepsilon \ll 1$, there is a subset $\Pi_\varepsilon \subset \Pi$ with*

$$
|\Pi \setminus \Pi_\varepsilon| = O(\varepsilon^{2/3}),
\tag{7.23}
$$

such that for every $\xi \in \Pi_\varepsilon$, the equation has a smooth quasi-periodic solution of the form

$$
u(t,x) = \sum_{j \geq 1} \tilde{q}_j(t) \phi_j(x),
\tag{7.24}
$$

where $\{\tilde{q}_j\}_{j \geq 1}$ are quasi-periodic functions with frequencies $\omega' := (\omega_1', \cdots, \omega_n')$. Moreover,

$$
|\omega' - \omega| = O(\varepsilon),
\tag{7.25}
$$

and for every non-negative integer ν, there exists a positive constant c depending on ν such that

$$
|\frac{d^\nu \tilde{q}_j}{dt^\nu}| \leq c\varepsilon, \quad j \in J,
\tag{7.26}
$$

$$
\sum_{j \notin J} e^{2aj} j^{2p} |\frac{d^\nu \tilde{q}_j}{dt^\nu}|^2 \leq c\varepsilon^{7/3}.
\tag{7.27}
$$

Remarks. 1. The estimate (7.25) follows from (1.10) in Theorem 1.1. From (1.9) in Theorem 1.1, we know

$$\left(\max_{1 \leq b \leq n} |\frac{d^\nu (\tilde{x}_b - \theta_b - \omega_b' t)}{dt^\nu}| \right) + \left(\max_{1 \leq b \leq n} |\frac{d^\nu \tilde{y}_b}{dt^\nu}| \right) r^{-2} + \left(\sum_{j \notin J} e^{2aj} j^{2p} |\frac{d^\nu \tilde{q}_j}{dt^\nu}|^2 \right)^{1/2} r^{-1}$$

$$= O(\varepsilon^{\frac{1}{1+\beta}}/\alpha) = O(\varepsilon^{1/6}),$$

where $\tilde{q}_{j_b} = \sqrt{2(I_b + \tilde{y}_b)} e^{\mathbf{i} \tilde{x}_b}$, $1 \leq b \leq n$ and $\theta = (\theta_1, \ldots, \theta_n) \in \mathbb{T}^n$. Furthermore,

$$\frac{d^\nu (\tilde{x}_b - \theta_b - \omega_b' t)}{dt^\nu} = O(\varepsilon^{1/6}), \quad \frac{d^\nu \tilde{y}_b}{dt^\nu} = O(\varepsilon^{13/6}), \tag{7.28}$$

$$\left(\sum_{j \notin J} e^{2aj} j^{2p} |\frac{d^\nu \tilde{q}_j}{dt^\nu}|^2 \right)^{1/2} = O(\varepsilon^{7/6}). \tag{7.29}$$

The estimate (7.26) follows from (7.28) and $I = O(\varepsilon^2)$. The estimate (7.27) follows from (7.29).

2. Under Dirichlet boundary conditions, in order to get the Hamiltonian (7.9) of discrete form, one needs to develop $u = u(t,x)$ into Fourier sine series (7.6). Once it is done, the nonlinearity must be developed into Fourier sine series in proving the regularity of nonlinear Hamiltonian vector field (see Lemma 7.1). This excludes the usual nonlinearity $(|u|^2 u)_x$. Naturally, one expects to investigate the DNLS equation with nonlinearity $(|u|^2 u)_x$ subject to periodic boundary conditions. In this case, one must extend Lemma 3.1 to higher dimension. In present time, this is an open problem.

8 Application to Perturbed Benjamin-Ono Equation

The Benjamin-Ono equation describes the evolution of the interface between two inviscid fluids under some physical conditions (see [3]). Under periodic boundary conditions it reads

$$u_t + \mathcal{H} u_{xx} - u u_x = 0, \quad (t,x) \in \mathbb{R} \times \mathbb{T}, \tag{8.1}$$

where u is real-valued and $\mathcal{H}$ is the Hilbert transform defined for 2π-periodic functions with mean value zero by

$$\widehat{\mathcal{H}(f)}(0) = 0, \quad \widehat{\mathcal{H}(f)}(j) = -\mathbf{i} \operatorname{sgn}(j) \hat{f}(j), \quad j \in \mathbb{Z} \setminus \{0\}. \tag{8.2}$$

The Benjamin-Ono equation is an integrable system (see [1]). For the global well-posedness of the Cauchy problem of the above equation, see [13], [14]. In the first subsection, we transform the Benjamin-Ono Hamiltonian into its Birkhoff normal form up to order four. In the second subsection, by using Theorem 1.2, we prove there are many KAM tori and thus quasi-periodic solutions for the above equation with small Hamiltonian perturbations.

8.1 Birkhoff Normal Form

We introduce for any $N > 3/2$ the phase space

$$\mathcal{H}_0^N = \{u \in L^2(\mathbb{T}; \mathbb{R}): \ \hat{u}(0) = 0, \ \|u\|_N^2 = \sum_{j \in \mathbb{Z} \setminus \{0\}} |j|^{2N} |\hat{u}(j)|^2 < \infty\}$$

of real valued functions on $\mathbb{T}$, where

$$\hat{u}(j) = \int_0^{2\pi} u(x) e_{-j}(x) dx, \quad e_j(x) = \frac{1}{\sqrt{2\pi}} e^{\mathbf{i} j x}.$$

Under the standard inner product on $L^2(\mathbb{T};\mathbb{R})$, equation (8.1) can be written in Hamiltonian form

$$\frac{\partial u}{\partial t} = -\frac{d}{dx}\frac{\partial H}{\partial u} \tag{8.3}$$

with Hamiltonian

$$H(u) = \frac{1}{2}\int_{\mathbb{T}} u(\mathscr{H}u_x)dx - \frac{1}{6}\int_{\mathbb{T}} u^3 dx. \tag{8.4}$$

To write it in infinitely many coordinates, we make the ansatz

$$u(t,x) = \sum_{j\neq 0}\gamma_j q_j(t)e_j(x), \tag{8.5}$$

where $\gamma_j = \sqrt{|j|}$. The coordinates are taken from the Hilbert space $\ell_{N+1/2}$ of all complex-valued sequences $q = (q_j)_{j\neq 0}$ with

$$\|q\|^2_{N+1/2} = \sum_{j\neq 0}|q_j|^2 j^{2N+1} < \infty, \quad q_{-j} = \bar{q}_j. \tag{8.6}$$

Then (8.3) can be rewritten as

$$\dot{q}_j = -\mathbf{i}\sigma_j\frac{\partial H}{\partial q_{-j}}, \quad \sigma_j = \begin{cases} 1, & j \geq 1 \\ -1, & j \leq -1 \end{cases} \tag{8.7}$$

with the Hamiltonian

$$H(q) = \Lambda + G = \sum_{j\geq 1} j^2|q_j|^2 - \frac{1}{6\sqrt{2\pi}}\sum_{j+k+l=0}\gamma_j\gamma_k\gamma_l q_j q_k q_l. \tag{8.8}$$

The function G is analytic in $\ell_{N+1/2}$ with real value, and the Hamiltonian vector field X_G is an analytic map from $\ell_{N+1/2}$ into $\ell_{N-1/2}$ with

$$\|X_G\|_{N-1/2} = O(\|q\|^2_{N+1/2}). \tag{8.9}$$

In the following theorem, we transform the above Hamiltonian into its Birkhoff normal form up to order four.

Theorem 8.1. *There exists a real analytic symplectic coordinate transformation Φ defined in a neighborhood of the origin of $\ell_{N+1/2}$ which transforms the above Hamiltonian H into its Birkhoff normal form up to order four. More precisely,*

$$H \circ \Phi = \Lambda + B + R \tag{8.10}$$

with

$$B = -\frac{1}{8\pi}\sum_{j,k\geq 1}\min(j,k)|q_j|^2|q_k|^2, \tag{8.11}$$

$$\|X_R\|_{N-1/2} = O(\|q\|^4_{N+1/2}). \tag{8.12}$$

Proof. (1). The first step is to eliminate the three order term G. Define $F^3 = \sum_{j,k,l\neq 0}F^3_{jkl}q_j q_k q_l$ by

$$iF^3_{jkl} = \begin{cases} \frac{1}{6\sqrt{2\pi}}\frac{\gamma_j\gamma_k\gamma_l}{\lambda_j+\lambda_k+\lambda_l}, & \text{for } j+k+l=0, \\ 0, & \text{otherwise,} \end{cases} \tag{8.13}$$

32

where $\lambda_j = \sigma_j j^2$. Then we have

$$\{\Lambda, F^3\} + G = 0, \tag{8.14}$$

where $\{\cdot, \cdot\}$ is Poisson bracket with respect to the symplectic structure $-\mathbf{i}\sum_{j\geq 1} dq_j \wedge dq_{-j}$. Letting $\Phi_1 = X_{F^3}^1$, then

$$
\begin{aligned}
H \circ \Phi_1 &= H \circ X_{F^3}^t|_{t=1} \\
&= \Lambda + \{\Lambda, F^3\} + \int_0^1 (1-t)\{\{\Lambda, F^3\}, F^3\} \circ X_{F^3}^t \, dt + G + \int_0^1 \{G, F^3\} \circ X_{F^3}^t \, dt \\
&= \Lambda + \int_0^1 t\{G, F^3\} \circ X_{F^3}^t \, dt \\
&= \Lambda + \frac{1}{2}\{G, F^3\} + \frac{1}{2}\int_0^1 (1-t^2)\{\{G, F^3\}, F^3\} \circ X_{F^3}^t \, dt.
\end{aligned} \tag{8.15}
$$

The j-th element of vector field X_{F^3} reads explicitly

$$-\mathbf{i}\sigma_j \frac{\partial F^3}{\partial q_{-j}} = -\mathbf{i}\sigma_j \sum_{k+l=j} 3F^3_{(-j)kl} q_k q_l = \frac{-\sigma_j}{2\sqrt{2\pi}} \sum_{k+l=j} \frac{\gamma_j \gamma_k \gamma_l}{\lambda_{-j} + \lambda_k + \lambda_l} q_k q_l. \tag{8.16}$$

For any $j, k, l \neq 0$ with $j + k + l = 0$,

$$\lambda_j + \lambda_k + \lambda_l = 2jkl / \max\{|j|, |k|, |l|\}. \tag{8.17}$$

Thus,

$$\left| \frac{\gamma_{k+l} \gamma_k \gamma_l}{\lambda_{-k-l} + \lambda_k + \lambda_l} \right| = \frac{\max\{|k+l|, |k|, |l|\}}{2\sqrt{|(k+l)kl|}} \leq \frac{\sqrt{2}}{2}. \tag{8.18}$$

Hence we have the estimate

$$\left| -\mathbf{i}\sigma_j \frac{\partial F^3}{\partial q_{-j}} \right| \leq \frac{1}{4\sqrt{\pi}} \sum_{k+l=j} |q_k||q_l| = \frac{1}{4\sqrt{\pi}} g_{-j}, \tag{8.19}$$

where g_{-j} stands for the sum $\sum_{k+l=j} |q_k||q_l|$. Obviously, $g = (g_j)_{j\neq 0}$ is the two-fold convolution of $w = (|q_j|)_{j\neq 0}$, that is, $g = w * w$. Thus, for any $r > 1/2$,

$$\|X_{F^3}\|_r \leq \frac{1}{4\sqrt{\pi}} \|g\|_r \leq c\|w\|_r^2 = c\|q\|_r^2, \tag{8.20}$$

where $c > 0$ depends only on r. This establishes the regularity of the vector field X_{F^3}.

(2). The second step is to normalize the four order term $\frac{1}{2}\{G, F^3\}$ in (8.15). By a simple calculation, we have

$$
\begin{aligned}
\frac{1}{2}\{G, F^3\} &= -\frac{\mathbf{i}}{2} \sum_{j\neq 0} \sigma_j \frac{\partial G}{\partial q_j} \frac{\partial F^3}{\partial q_{-j}} \\
&= \frac{1}{16\pi} \sum_{j\neq 0} \sigma_j \left(\sum_{k+l=-j} \gamma_j \gamma_k \gamma_l q_k q_l \right) \left(\sum_{m+n=j} \frac{\gamma_{-j} \gamma_m \gamma_n}{\lambda_{-j} + \lambda_m + \lambda_n} q_m q_n \right) \\
&= \frac{1}{16\pi} \sum_{\substack{k,l,m,n \neq 0 \\ k+l+m+n=0 \\ k+l \neq 0}} \frac{(m+n)\gamma_k \gamma_l \gamma_m \gamma_n}{-\lambda_{m+n} + \lambda_m + \lambda_n} q_k q_l q_m q_n.
\end{aligned} \tag{8.21}
$$

33

Let B consist of all terms with $k+m=0$ or $k+n=0$ in (8.21). Then write B explicitly,

$$
\begin{aligned}
B &= \frac{1}{16\pi}\sum_{k\neq 0}\frac{-2k^3}{\lambda_{2k}-2\lambda_k}|q_k|^4 + \frac{1}{8\pi}\sum_{k,l,k\pm l\neq 0}\frac{-(k+l)|kl|}{\lambda_{k+l}-\lambda_k-\lambda_l}|q_k|^2|q_l|^2 \\
&= -\frac{1}{16\pi}\sum_{k\neq 0}|k||q_k|^4 - \frac{1}{16\pi}\sum_{k,l,k\pm l\neq 0}\sigma_{kl}\max\{|k|,|l|,|k+l|\}|q_k|^2|q_l|^2 \\
&= -\frac{1}{8\pi}\sum_{k\geq 1}k|q_k|^4 - \frac{1}{4\pi}\sum_{k>l\geq 1}l|q_k|^2|q_l|^2,
\end{aligned}
\tag{8.22}
$$

which is (8.11). Let $Q=\frac{1}{2}\{G,F^3\}-B$. We will find a coordinate transformation $\Phi_2 = X_{F^4}^1$ to eliminate Q. In complete analogy to the first step we let

$$
F^4 = \sum_{k,l,m,n\neq 0}F^4_{klmn}q_kq_lq_mq_n,
\tag{8.23}
$$

$$
iF^4_{klmn} = \begin{cases} -\frac{1}{16\pi}\frac{(m+n)\gamma_k\gamma_l\gamma_m\gamma_n}{(-\lambda_{m+n}+\lambda_m+\lambda_n)(\lambda_k+\lambda_l+\lambda_m+\lambda_n)}, & \text{for } \begin{array}{l} k+l+m+n=0 \\ k+l,k+m,k+n\neq 0 \end{array}, \\ 0, & \text{otherwise.} \end{cases}
\tag{8.24}
$$

Then we have

$$
\{\Lambda,F^4\}+Q=0.
\tag{8.25}
$$

In order to complete the second step, we only need to establish the regularity of the vector field X_{F^4}. We claim that for $f_{klmn}:=\frac{-\lambda_{m+n}+\lambda_m+\lambda_n}{m+n}(\lambda_k+\lambda_l+\lambda_m+\lambda_n)$ with $k+l+m+n=0$ and $k,l,m,n,k+l,k+m,k+n\neq 0$,

$$
|f_{klmn}|\geq \frac{1}{2}\max\{|k|,|l|,|m|,|n|\}.
\tag{8.26}
$$

We prove this claim in four cases. Ahead of the proof, we give a simple inequality for two positive integers a, b:

$$
2ab\geq a+b,
$$

which will be frequently used. Without loss of generality, we assume $|k|\geq |l|$ and $|m|\geq |n|$ because of their symmetry in f_{klmn}.

(1) l,m,n have the opposite sign of k. Without loss of generality, we assume $k<0$ and $l,m,n>0$. In this case, we have

$$
|-\lambda_{m+n}+\lambda_m+\lambda_n| = 2|mn|\geq |m+n|,
$$

$$
|\lambda_k+\lambda_l+\lambda_m+\lambda_n| = 2(lm+ln+mn) \geq 2(lm+ln) = 2l(|k|-l) \geq |k|,
$$

which leads to (8.26).

(2) k,l,n have the opposite sign of m. Without loss of generality, we assume $m<0$ and $k,l,n>0$. In this case, we have

$$
|-\lambda_{m+n}+\lambda_m+\lambda_n| = 2|(m+n)n| \geq |m+n|,
$$

$$
|\lambda_k+\lambda_l+\lambda_m+\lambda_n| = 2(kl+kn+ln) \geq 2(kn+ln) = 2(|m|-n)n \geq |m|,
$$

which leads to (8.26).

(3) k, l have the same sign while m, n have the same sign. Without loss of generality, we assume $k, l > 0$ and $m, n < 0$. Here (8.26) follows from the following inequalities:

$$\left|\frac{-\lambda_{m+n} + \lambda_m + \lambda_n}{m+n}\right| = \left|\frac{2mn}{m+n}\right| \geq |n|,$$

$$|n||\lambda_k + \lambda_l + \lambda_m + \lambda_n| = 2|n|(k - |n|)|k - |m|| \geq k|k - |m|| \geq \frac{1}{2}\max\{k, |m|\}.$$

(4) k, n have the same sign while l, m have the same sign. Without loss of generality, we assume $k, n > 0$ and $l, m < 0$. Here (8.26) follows from the following inequalities:

$$\left|\frac{-\lambda_{m+n} + \lambda_m + \lambda_n}{m+n}\right| = 2n,$$

$$2n|\lambda_k + \lambda_l + \lambda_m + \lambda_n| = 4n(|m| - n)||m| - k| \geq 2|m||k - |m|| \geq \max\{k, |m|\}.$$

Finally, the proof of (8.26) is finished. From (8.26) we get

$$|F_{klmn}^4| \leq \frac{\gamma_k \gamma_l \gamma_m \gamma_n}{8\pi \max\{|k|, |l|, |m|, |n|\}}. \tag{8.27}$$

The j-th element of vector field X_{F^4} explicitly reads

$$-\mathbf{i}\sigma_j \frac{\partial F^4}{\partial q_{-j}} = -\mathbf{i}\sigma_j \sum_{l+m+n=j} \left(F_{(-j)lmn}^4 + F_{l(-j)mn}^4 + F_{lm(-j)n}^4 + F_{lmn(-j)}^4\right) q_l q_m q_n. \tag{8.28}$$

Hence we have the estimate

$$\left|-\mathbf{i}\sigma_j \frac{\partial F^4}{\partial q_{-j}}\right| \leq \frac{1}{2\pi\gamma_j} \sum_{l+m+n=j} \gamma_l \gamma_m \gamma_n |q_l q_m q_n| \leq \frac{1}{2\pi\gamma_j} g_{-j}, \tag{8.29}$$

where g_{-j} stands for the sum $\sum_{l+m+n=j} \gamma_l \gamma_m \gamma_n |q_l||q_m||q_n|$. Obviously, $g = (g_j)_{j \neq 0}$ is the three-fold convolution of $w = (\gamma_j |q_j|)_{j \neq 0}$, that is $g = w * w * w$. Thus, for any $r > 1$,

$$\|X_{F^4}\|_r \leq \frac{1}{2\pi} \|g\|_{r-1/2} \leq c\|w\|_{r-1/2}^3 = c\|q\|_r^3, \tag{8.30}$$

where $c > 0$ depends only on r. This establishes the regularity of the vector field X_{F^4}, and finishes the proof of Theorem 8.1. $\qquad \square$

8.2 Quasi-periodic Solutions for Perturbed Benjamin-Ono Equation

Here is our main result in this section:

Theorem 8.2. *Consider the Benjamin-Ono equation (8.1) with a small Hamiltonian perturbation, written in the Hamiltonian form*

$$\frac{\partial u}{\partial t} = -\frac{d}{dx}\left(\frac{\partial H}{\partial u} + \varepsilon \frac{\partial K}{\partial u}\right), \tag{8.31}$$

where the Hamiltonian H is defined by (8.4), and the Hamiltonian K is real analytic in a complex neighborhood V of the origin in $\mathscr{H}_{0,\mathbb{C}}^N$, which is the complexification of $\mathscr{H}_0^N$, $N > 3/2$. Moreover, K satisfies the regularity condition

$$\frac{\partial K}{\partial u} : \quad V \to \mathscr{H}_{0,\mathbb{C}}^N, \quad \left\|\frac{\partial K}{\partial u}\right\|_{N,V} \leq 1. \tag{8.32}$$

Then, for each index set $J = \{j_1 < j_2 < \cdots < j_n\} \subset \mathbb{N}$, there exists an $\varepsilon_0 > 0$ depending only on J, N and the size of V, such that for $\varepsilon < \varepsilon_0$, the equation has uncountable quasi-periodic solutions with frequency vector close to $(j_1^2, \cdots, j_n^2)$.

Proof. Set the perturbed Hamiltonian $\tilde{H} = H + \varepsilon K$. Then, by the transformation Φ in Theorem 8.1, we get the new Hamiltonian, still denoted by $\tilde{H}$,

$$\tilde{H} = \Lambda + B + R + \varepsilon K \circ \Phi, \tag{8.33}$$

which is analytic in some neighborhood U of the origin of $\ell_{N+1/2}$ with Λ in (8.8), B in (8.11), R satisfying (8.12) and the last term satisfying

$$\|X_{K \circ \Phi}\|_{N-1/2,U} \leq 2. \tag{8.34}$$

We introduce symplectic polar and real coordinates (x, y, u, v) by setting

$$\begin{cases} q_{j_b} = \sqrt{\xi_b + y_b}e^{-\mathbf{i}x_b}, & q_{-j_b} = \sqrt{\xi_b + y_b}e^{\mathbf{i}x_b}, & b = 1, \cdots, n, \\ q_j = \frac{1}{\sqrt{2}}(u_j - \mathbf{i}v_j), & q_{-j} = \frac{1}{\sqrt{2}}(u_j + \mathbf{i}v_j), & j \in \mathbb{N}_* := \mathbb{N} \setminus J, \end{cases} \tag{8.35}$$

where $\xi = (\xi_1, \cdots, \xi_n) \in \mathbb{R}_+^n$. Then

$$\Lambda = \sum_{1 \leq b \leq n} j_b^2(\xi_b + y_b) + \frac{1}{2} \sum_{j \in \mathbb{N}_*} j^2(u_j^2 + v_j^2), \tag{8.36}$$

$$\begin{aligned} -8\pi B &= \sum_{1 \leq b, b' \leq n} \min(j_b, j_{b'})(\xi_b + y_b)(\xi_{b'} + y_{b'}) \\ &+ \sum_{1 \leq b \leq n, j \in \mathbb{N}_*} \min(j_b, j)(\xi_b + y_b)(u_j^2 + v_j^2) \\ &+ \frac{1}{4} \sum_{j, j' \in \mathbb{N}_*} \min(j, j')(u_j^2 + v_j^2)(u_{j'}^2 + v_{j'}^2). \end{aligned} \tag{8.37}$$

Thus the new Hamiltonian, still denoted by $\tilde{H}$, up to a constant depending only on ξ, is given by

$$\tilde{H} = N + P = \sum_{1 \leq b \leq n} \omega_b y_b + \frac{1}{2} \sum_{j \in \mathbb{N}_*} \Omega_j(u_j^2 + v_j^2) + Q + R + \varepsilon K, \tag{8.38}$$

with symplectic structure $\sum_{1 \leq b \leq n} dy_b \wedge dx_b + \sum_{j \in \mathbb{N}_*} du_j \wedge dv_j$, where

$$\omega_b = j_b^2 - \frac{1}{4\pi} \sum_{1 \leq b' \leq n} \min(j_b, j_{b'})\xi_{b'}, \tag{8.39}$$

$$\Omega_j = j^2 - \frac{1}{4\pi} \sum_{1 \leq b \leq n} \min(j_b, j)\xi_b, \tag{8.40}$$

$$\begin{aligned} Q = &- \frac{1}{8\pi} \sum_{1 \leq b, b' \leq n} \min(j_b, j_{b'})y_b y_{b'} - \frac{1}{8\pi} \sum_{1 \leq b \leq n, j \in \mathbb{N}_*} \min(j_b, j)y_b(u_j^2 + v_j^2) \\ &- \frac{1}{32\pi} \sum_{j, j' \in \mathbb{N}_*} \min(j, j')(u_j^2 + v_j^2)(u_{j'}^2 + v_{j'}^2). \end{aligned} \tag{8.41}$$

Set

$$\Pi = \{\xi \in \mathbb{R}_+^n : |\xi| \leq \varepsilon^{2/5}\}. \tag{8.42}$$

In the following we check the assumptions (A) (B) (C) and (D*).

36

In view of (8.39), we know that $\xi \mapsto \omega$ is an affine transformation from Π to $\mathbb{R}^n$. Regarding ω as n-dimensional column vector,

$$
\omega = \breve{\omega} - \frac{1}{4\pi}A\xi = \begin{pmatrix} j_1^2 \\ j_2^2 \\ j_3^2 \\ \vdots \\ j_n^2 \end{pmatrix} - \frac{1}{4\pi}\begin{pmatrix} j_1 & j_1 & j_1 & \cdots & j_1 \\ j_1 & j_2 & j_2 & \cdots & j_2 \\ j_1 & j_2 & j_3 & \cdots & j_3 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ j_1 & j_2 & j_3 & \cdots & j_n \end{pmatrix}\begin{pmatrix} \xi_1 \\ \xi_2 \\ \xi_3 \\ \vdots \\ \xi_n \end{pmatrix}, \tag{8.43}
$$

where $\breve{\omega}$ is a n-dimensional column vector with its b-th element $\breve{\omega}_b = j_b^2$ and A is a $n \times n$ matrix with its (b,m)-th element $A_{bm} = \min(j_b, j_m)$. Let

$$
T = \begin{pmatrix} 1 & -1 & & & \\ & 1 & -1 & & \\ & & 1 & \ddots & \\ & & & \ddots & -1 \\ & & & & 1 \end{pmatrix}\begin{pmatrix} 1 & & & & \\ -\frac{j_1}{j_2-j_1} & 1 & & & \\ & -\frac{j_2-j_1}{j_3-j_2} & 1 & & \\ & & \ddots & \ddots & \\ & & & -\frac{j_{n-1}-j_{n-2}}{j_n-j_{n-1}} & 1 \end{pmatrix}. \tag{8.44}
$$

Then by a simple calculation, we have $AT = \tilde{A} := \mathrm{diag}(j_b - j_{b-1} : 1 \leq b \leq n)$ (here $j_0 = 0$). Thus,

$$
\det\left(\frac{\partial\omega}{\partial\xi}\right) = (-\frac{1}{4\pi})^n j_1(j_2 - j_1)(j_3 - j_2)\cdots(j_n - j_{n-1}) \neq 0. \tag{8.45}
$$

This implies that the assumption (A) is fulfilled with positive constant M_1 and L only depending on the set J.

In view of (8.40), we know

$$
\Omega_j = j^2 - \frac{1}{4\pi}B_j\xi, \tag{8.46}
$$

where B_j is a n-dimensional row vector with its m-th element $B_{jm} = \min(j, j_m)$. It's obvious that $\Omega_j \approx j^2$, thus, the assumption (B) is fulfilled with $d = 2$ and $m = 1/2$.

From (8.46), we get

$$
|\Omega_j|_\Pi^{\mathrm{lip}} = \frac{1}{4\pi}\sum_{1\leq m\leq n}B_{jm} = \frac{1}{4\pi}\sum_{1\leq m\leq n}\min(j, j_m). \tag{8.47}
$$

Thus,

$$
|\Omega|_{-1,\Pi}^{\mathrm{lip}} = \sup_{j\in\mathbb{N}_*}j^{-1}|\Omega_j|_\Pi^{\mathrm{lip}} \leq \frac{n}{4\pi}, \tag{8.48}
$$

that is, assumption (C) is fulfilled with $\delta = 1$ and $M_2 = \frac{n}{4\pi}$.

Moreover, taking $\delta_0 = 0$ and letting $\Omega_j^1 = \Omega_j$, $\Omega_j^2 = 0$ in the assumption (D*), then from (8.47) we have

$$
|\Omega^1|_{-\delta_0,\Pi}^{\mathrm{lip}} = \sup_{j\in\mathbb{N}_*}|\Omega_j|_\Pi^{\mathrm{lip}} = \frac{1}{4\pi}\sum_{1\leq m\leq n}j_m := M_3, \tag{8.49}
$$

which only depends on J. On the other hand, since $\Omega^2 = 0$, the positive constant M_4 can be chosen arbitrarily small. Therefore, we can choose it small enough such that (1.11) is fulfilled. Thus the second part of the assumption (D*) holds true.

In view of (8.46), regarding Ω as an infinite dimensional column vector with its index $j \in \mathbb{N}_*$, we have

$$
\Omega = \breve{\Omega} - \frac{1}{4\pi}B\xi, \tag{8.50}
$$

37

where $\check{\Omega}$ is an infinite dimensional column vector with its j-th element $\check{\Omega}_j = j^2$ and B is a $\infty \times n$ matrix with its j-th row B_j. For the first part of the assumption (D*), regarding k and l as n-dimensional and infinite dimensional row vectors respectively, we have to check for every $k \in \mathbb{Z}^n$ and $l \in \mathbb{Z}^\infty$ with $1 \leq |l| \leq 2$,

$$k\check{\omega} + l\check{\Omega} \neq 0 \quad or \quad kA + lB \neq 0. \tag{8.51}$$

By a simple calculation, we have $BA^{-1} = BT\tilde{A}^{-1} = \tilde{B}$, which is a $\infty \times n$ matrix with its j-th row $\tilde{B}_j$:

(1) $\tilde{B}_j = (\frac{j}{j_1}, 0, \cdots, 0)$ for $j < j_1$, only the first element nonzero;

(2) $\tilde{B}_j = (0, \cdots, 0, \frac{j_{m+1}-j}{j_{m+1}-j_m}, \frac{j-j_m}{j_{m+1}-j_m}, 0, \cdots, 0)$ for $j_m < j < j_{m+1}$ with $1 \leq m < n$, only the m-th and $(m+1)$-th elements nonzero;

(3) $\tilde{B}_j = (0, \cdots, 0, 1)$ for $j > j_n$, only the last element nonzero.

Then the second inequality of (8.51) is equivalent to

$$k + l\tilde{B} \neq 0. \tag{8.52}$$

Set $\mathbb{N}_* = \mathbb{N}_*^1 \bigcup \mathbb{N}_*^2 \bigcup \mathbb{N}_*^3$, where $\mathbb{N}_*^1 = \{1 \leq j < j_1\}$, $\mathbb{N}_*^2 = \{j_1 \leq j \leq j_n\} \setminus J$ and $\mathbb{N}_*^3 = \{j > j_n\}$. Notice that for $j \in \mathbb{N}_*^1 \bigcup \mathbb{N}_*^2$ the nonzero elements of $\tilde{B}_j$ are positive and less than 1. Thus the equality (8.52) holds true for $k \in \mathbb{Z}^n$ and $1 \leq |l| \leq 2$ except the following three cases:

(1)
$$l_j = \begin{cases} \pm 2, & j = j_1/2 \\ 0, & \text{otherwise} \end{cases}, \qquad k_m = \begin{cases} \mp 1, & m = 1 \\ 0, & \text{otherwise} \end{cases};$$

(2) For a fixed m' with $1 \leq m' < n$,

$$l_j = \begin{cases} \pm 2, & j = (j_{m'} + j_{m'+1})/2 \\ 0, & \text{otherwise} \end{cases}, \qquad k_m = \begin{cases} \mp 1, & m = m' \\ \mp 1, & m = m' + 1 \\ 0, & \text{otherwise} \end{cases};$$

(3) $l_j = 0$ for $j \in \mathbb{N}_*^1 \bigcup \mathbb{N}_*^2$, and

$$k_m = \begin{cases} 0, & 1 \leq m < n \\ -\sum_{j \in \mathbb{N}_*} l_j, & m = n \end{cases}.$$

It's easy to check that the first inequality of (8.51) holds true for all of the above three cases. Therefore, the first part of the assumption (D*) holds true.

Now we consider sup norm and Lipschitz semi-norm of the perturbation

$$P = Q + R + \varepsilon K \tag{8.53}$$

on $D(s,r) \times \Pi$, where $D(s,r)$ is defined in (1.2) by letting $p = N + 1/2$ and Π is defined in (8.42). We choose $s > 0$ a constant, and

$$r = \varepsilon^{1/4}. \tag{8.54}$$

In view of (8.41), we have

$$\|X_Q\|_{r,a,p-1,D(s,r)\times\Pi} = O(r^2) = O(\varepsilon^{1/2}). \tag{8.55}$$

In view of (8.12), we know R is at least five order of q. Thus,

$$\|X_R\|_{r,a,p-1,D(s,r)\times\Pi} = O\left((\varepsilon^{1/5})^5 r^{-2}\right) = O(\varepsilon^{1/2}). \tag{8.56}$$

In view of (8.34), we have

$$\|X_K\|_{r,a,p-1,D(s,r)\times\Pi} = O(r^{-2}) = O(\varepsilon^{-1/2}). \tag{8.57}$$

From (8.55) (8.56) (8.57) we get

$$\|X_P\|_{r,a,p-1,D(s,r)\times\Pi} = O(\varepsilon^{1/2}). \tag{8.58}$$

Since X_P is real analytic in ξ, we have

$$\|X_P\|_{r,a,p-1,D(s,r)\times\Pi}^{\mathrm{lip}} = O(\varepsilon^{1/2}\varepsilon^{-2/5}). \tag{8.59}$$

We choose

$$\alpha = \varepsilon^{9/20}\gamma^{-1}, \quad \beta = 1/18, \tag{8.60}$$

where γ is taken from the KAM theorem. Set $M := M_1 + M_2$, which only depends on the set J. It's obvious that when ε is small enough,

$$\tilde{\varepsilon} := \|X_P\|_{r,a,p-1,D(s,r)\times\Pi} + \frac{\alpha}{M}\|X_P\|_{r,a,p-1,D(s,r)\times\Pi}^{\mathrm{lip}} = O(\varepsilon^{1/2}) \le (\alpha\gamma)^{1+\beta}, \tag{8.61}$$

which is just the smallness condition (1.7). Now the conclusion of Theorem 8.2 follows from Theorem 1.2. $\qquad\square$

9 A Technical Lemma

Lemma 9.1. *Let $F = (F_{ij})_{i,j\ge 1}$ be a bounded operator on ℓ^2 which depends on $x \in \mathbb{T}^n$ such that all elements (F_{ij}) are analytic on $D(s)$. Suppose $R = (R_{ij})_{i,j\ge 1}$ is another operator on ℓ^2 depending on x whose elements satisfy*

$$\sup_{x\in D(s')} |R_{ij}(x)| \le \frac{1}{|i-j|} \sup_{x\in D(s-\sigma)} |F_{ij}(x)|, \quad i \ne j, \tag{9.1}$$

and $R_{ii} = 0$. Then R is a bounded operator on ℓ^2 for every $x \in D(s')$, and

$$\sup_{x\in D(s')} \|R(x)\| \le \frac{4^{n+1}}{\sigma^n} \sup_{x\in D(s)} \|F(x)\|. \tag{9.2}$$

Proof. This is lemma M.3 in [10]. $\qquad\square$

References

[1] Ablowitz, M. J., Fokas, A. S.: The inverse scattering transform for the Benjamin-Ono equation, a pivot for multidimensional problems, Stud. Appl. Math. 68, 1-10 (1983)

[2] Bambusi, D.; Graffi, S. Time Quasi-Periodic Unbounded Perturbations of Schrödinger Operators and KAM Methods. Commun. Math. Phys. 219, 465-480 (2001)

[3] Benjamin, T. B.: Internal waves of permanent form in fluids of great depth. J. Fluid Mech. 29, 559-592 (1967)

[4] Bourgain, J.: Recent progress on quasi-periodic lattice Schrödinger operators and Hamiltonian PDEs. Russian Math. Surveys 59(2), 231-246 (2004)

[5] Klainerman, S.: Long-time behaviour of solutions to nonlinear wave equations. In Proceedings of International Congress of Mathematics (Warsaw), 1209-1215 (1983)

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

[7] Kuksin, S. B.: Analysis of Hamiltonian PDEs, Oxford Univ. Press, Oxford, 2000

[8] Kuksin, S. B.: Fifteen years of KAM in PDE, "Geometry, topology, and mathematical physics": S. P. Novikov's seminar 2002-2003 edited by V. M. Buchstaber, I. M. Krichever. Amer. Math. Soc. Transl. Series 2, Vol.212, 2004, pp.237-257

[9] Kuksin, S. B., Pöschel, J: Invariant Cantor manifolds of quasi-periodic oscillations for a nonlinear Schrödinger equation. Ann. Math. 143, 147-179 (1996)

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

[11] Lax, P. D.: Development of singularities of solutions of nonlinear hyperbolic partial differential equations. J. Math. Phys. 5, 611-613 (1964)

[12] Liu, J., Yuan, X.: Spectrum for quantum Duffing oscillator and small-divisor equation with large variable coefficient. Commun. Pure. Appl. Math. 63, 1145-1172 (2010)

[13] Molinet, L.: Global well-posedness in the energy space for the Benjamin-Ono equation on the circle. Math. Ann. 337, 353-383 (2007)

[14] Molinet, L.: Global well-posedness in L^2 for the periodic Benjamin-Ono equation. Amer. J. Math. 130, 635-683 (2008)