A VARIANT OF KAM THEOREM WITH APPLICATIONS TO NONLINEAR WAVE EQUATIONS OF HIGHER DIMENSION

Xiaoping Yuan

Department of Mathematics, Fudan University, Shanghai 200433, China
Email: xpyuan@fudan.edu.cn; yuanxiaoping@hotmail.com

Abstract. *The existence of lower dimensional KAM tori is shown for a class of nearly integrable Hamiltonian systems where the second Melnikov's conditions are relaxed, at the cost of the stronger regularity of the perturbed nonlinear term. As a consequence, it is proved that there exist many linearly stable invariant tori and thus quasi-periodic solutions for nonlinear wave equations of non-local nonlinearity and of higher spatial dimension.*

1. Introduction and main results.

Let us begin with the non-linear wave (NLW) equation

$$u_{tt} - u_{xx} + V(x)u + h(x,u) = 0 \tag{1.1}$$

subject to Dirichlet boundary conditions. The existence of solutions, periodic in time, for NLW equations has been studied by many authors. See [B-P, Br, L-S] and the references theirin, for example. While finding quasi-periodic solutions, the so-called small divisor difficulty arises. The KAM (Kolmogorov-Arnold-Moser) theory is a very powerful tool to overcome the difficulty. This theory deals with the existence of invariant tori for nearly integrable Hamiltonian systems. In order to obtain the quasi-periodic solutions of a partial differential equation, we may show the existence of the lower (finite) dimensional invariant tori for the infinitely dimensional Hamiltonian system defined by the equation. Assume the hamiltonian is of the form:

$$H = (\omega, y) + \sum_{j=1}^{\infty} \Omega_j z_j \bar{z}_j + R(x, y, z, \bar{z})$$

with tangential frequencies $\omega = (\omega_1, ..., \omega_n)$ and normal frequencies $\Omega = (\Omega_1, ...,)$. When $R \equiv 0$, there is a trivial invariant torus $x = \omega t, y = 0, z = \bar{z} = 0$. The KAM theory guarantees the persistence of the trivial invariant torus for sufficiently small perturbation R, provided that the well-known Melnikov conditions are fulfilled:

$$(k, \omega) - \Omega_j \neq 0 \qquad \text{(the first Melnikov's)}$$

for all $k \in \mathbb{Z}^n$ and $1 \le j < \infty$, and

$$(k, \omega) + \Omega_{j_1} - \Omega_{j_2} \neq 0 \qquad \text{(the second Melnikov's)}$$

Supported by NNSFC and in part by the grant of STCSM: 03JC14013.

Typeset by $\mathcal{A}\mathcal{M}\mathcal{S}$-TEX for all $k \in \mathbb{Z}^n$ and $1 \leq j_1, j_2 < \infty, j_1 \neq j_2$. See [E,K1,P1,W] for the details. This KAM theorem can be applied to a wide array of Hamiltonian partial differential equations of 1-dimensional spatial variable, including (1.1). Kuksin[K1,2] shows that there are many quasi-periodic solutions of (1.1), assuming that the potential V depends on an n-dimensional external parameter in some non-degenerate way. Wayne[W] obtains also the existence of the quasi-periodic solutions of (1.1), when the potential V is lying on the outside of the set of some "bad" potentials. In [W], the set of all potentials is given some Gaussian measure and then the set of "bad" potentials is of small measure. Bobenko & Kuksin[Bo-K] and Pöschel[P2] investigate the case $V(x) \equiv m \in (0, \infty)$. By the remark in [P2], the same result holds also true for the parameter values $-1 < m < 0$. When $m \in (-\infty, -1) \setminus \mathbb{Z}$, it is shown in [Y1] that there are many hyperbolic-elliptic invariant tori. More recently, the existence of invariant tori (thus quasi-periodic solutions) of (1.1) are shown for any prescribed potential[1] $V(x) \neq 0$ in [Y2] and for $V(x) \equiv 0$ in [Y3]. In [C-Y] and [Br-K-S], the equation (1.1) subject to periodic boundary conditions is investigated.

For NLW equation (1.1) of spatial dimension 1, the multiplicity of normal frequency Ω_j is 1 in Dirichlet boundary condition or 2 in periodic boundary condition. Considering PDE's with spatial dimension> 1, a significant new problem arises due to the presence of clusters of normal frequencies of the Hamiltonian system defined by the PDEs. In this case, the multiplicity of Ω_j goes to ∞ as $|j| \to \infty$; consequently, the second Melnikov's conditions is destroyed seriously, preventing the application of the KAM theorems mentioned above to Hamiltonian partial differential equations of higher spatial dimension. Bourgain[Bo1-4] developed another profound approach, originally proposed by Craig-Wayne in [C-W], and successfully obtained the existence of quasi-periodic solutions of the nonlinear Schrödinger (NLS) equations and NLW equations of higher dimension in space. This method is called C-W-B method in some references. The techniques used in [C-W] and [Bo1-4] are based on not KAM theory, but rather on a generalization of Lyapunov-Schmidt procedure and on techniques by Fröhlich and Spencer[F-S].

The advantage of the KAM approach is, from one hand, to possibly simplify the proof and, on the other hand, to allow the construction of local normal forms closed to the considered torus, which could be useful for the better understanding of the dynamics. For example, in generally, one can easy check the linear stability and the vanishing Lyapunov exponents.

Naturally, we should ask that *whether or not one can establish a new KAM theorem for some nonlinear partial differential equations, such as NLW and NLS, of higher spatial dimension.*

In a private talk, the present author was told that Eliasson and Kuksin got a new KAM theorem which could be applied to NLS equations. This is an excited news! In the present paper,we will prove a variant of the KAM theorem due to Kuksin[K1] and Pöschel[P1]. In the variant, the requirements of the normal frequencies Ω_j's are weaker than those in [K1,P1], at the expense of stronger regularity of nonlinearity. Consequently, we can show that there are many invariant tori which are linearly stable, for the NLW equations of non-local nonlinear term and higher spatial dimension:

$$u_{tt} - \triangle u + V(x,\xi)u + \Psi((\Psi u)^3) = 0, \quad \text{in } \mathbb{R} \times (0, 2\pi)^d \tag{1.2}$$

[1]This potential V contains no parameter.

and

$$u_{tt} - \triangle u + M_\xi u + \Psi((\Psi u)^3) = 0, \quad \text{in } \mathbb{R} \times (0, 2\pi)^d \tag{1.3}$$

subject to Dirichlet boundary condition

$$u(t, x)|_{x \in \partial[0,2\pi]^d} = 0 \tag{1.4}$$

where $\triangle$ is d-Laplacian, the potential V depends on parameter ξ in some kind of non-degenerate way, and M_ξ is a Fourier multiplier, i.e.,

$$M_\xi e^{\sqrt{-1}(j,x)} = \xi_j e^{\sqrt{-1}(j,x)}, \quad \xi_j \in \mathbb{R}, j \in \mathbb{Z}^d \tag{1.5}$$

and $\Psi : u \mapsto \psi \star u$ is a convolution operator with a function ψ, even in each entry of $x \in \mathbb{R}^d$. Assume the operator Ψ is smoothing of order[2] $\kappa = 577d/2$:

$$\begin{aligned} \Psi :&H_0^p([0, 2\pi]) \to H_0^{\bar{p}}([0, 2\pi]), \ \bar{p} = p + \kappa, \\ &||\Psi u||_{H^{\bar{p}}} \lessdot ||\Psi u||_{H^p}, \ \forall \ p > d/2. \end{aligned} \tag{1.6}$$

The variant of the KAM theorem also applies to nonlinear Schrödinger equations of higher spatial dimension:

$$\sqrt{-1} u_t + \mathcal{A} u + \Psi((\Psi u)^3) = 0, \quad \text{in } \mathbb{R} \times (0, 2\pi)^d \tag{1.7}$$

subject to b. c. (1.4) where $\mathcal{A} = -\triangle + M_\xi$ or $\mathcal{A} = -\triangle + V(x, \xi)$. Geng and You[G-Y] also show the existence of stable invariant tori of (1.7) with the regularity $\kappa > 0$. The requirement of the regularity in [G-Y] is weaker than ours, but our result can apply to nonlinear wave equations (1.2) and (1.3). In addition, Pöschel[P3] shows that there are many almost periodic solutions of (1.7) when $d = 1$. The non-local condition is not satisfactory. It is an intereating problem that whether or not the non-local condition can be removed.

The paper is organized as follows: In §.2, we formulate a general infinitely dimensional KAM theorem designed to deal with the presence of clusters of normal frequencies of the Hamiltonian system. In §.3, we show how to apply the preceding KAM theorem to NLW equation (1.3) with b. c. (1.4). Sect.4-8 are devoted to the proof of the KAM theorem. In §.4, the homological equations are reduced and solved; in §.5, the symplectic transform X_F^1 is given out and the new perturbed term R_+ is estimated; in §.6, the iterative lemma is given out; in §.7, The KAM theorem is proven by using the iterative lemma;in §.8, the measure estimates for the parameter sets is given out. Some technical lemmas are provided in §.9 − 10.

2. A variant of the KAM theorem due to Kuksin and Pöschel.

2.1. Some notations. Denote by $(\ell_2, ||\cdot||)$ the usual space of the square summable sequences, and by $(L^2, ||\cdot||)$ the space of the square integrable functions. By $|\cdot|$ the Euclidian norm. Let $a \geq 0$ and $p \geq d/2$. For a sequence $u = (u_j \in \mathbb{C}^* : j \in \mathbb{Z}^d)$ with $* = 1$ or 2, we define its norm as follows:

$$||u||_{a,p}^2 = \sum_{j \in \mathbb{Z}^d} |j|^{2p} e^{2a|j|} |u_j|^2. \tag{2.1}$$

[2]Here $\kappa = 577d/2$ is not optimal. Since the variant of KAM theorem does not hold true for $\kappa = 0$, we do not pursue the optimal $\kappa > 0$.

Let $\ell^{a,p}$ be the set of all sequences satisfying (2.1). It is easy to see that $\ell^{a,p}$ is a Hilbert space with an inner product corresponding to (2.1). When $a = 0$, we sometimes write $\ell^{0,p} = \ell^p$ and $||\cdot||_{0,p} = ||\cdot||_p$. Introduce the phase space:

$$\mathcal{P} := (\mathbb{C}^n/2\pi\mathbb{Z}^n) \times \mathbb{C}^n \times \ell^p, \tag{2.2}$$

where n is a given positive integer. We endow $\mathcal{P}$ with a symplectic structure

$$dx \wedge dy + \sum_{j\in\mathbb{Z}^d} du_j^1 \wedge du_j^2, \quad (x, y, u) \in \mathcal{P},$$

where $u = (u_j)_{j\in\mathbb{Z}^d}$ with $u_j = (u_j^1, u_j^2) \in \mathbb{C}^2$. Let

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

Then $\mathcal{T}_0^n$ is an torus in $\mathcal{P}$. Introduce a complex neighborhoods of $\mathcal{T}_0^n$ in $\mathcal{P}$:

$$D(s, r) := \{(x, y, u) \in \mathcal{P} : |\mathrm{Im} x| < s, |y| < r^2, ||u||_p < r\}$$

where $r, s > 0$ are constants.

Recall $\bar{p} = p + \kappa$ in (1.6). For $\tilde{p} = p$ or $\tilde{p} = \bar{p}$, let

$$\mathcal{P}^{a,\tilde{p}} := \mathbb{C}^n \times \mathbb{C}^n \times \ell^{a,\tilde{p}}, \quad \forall a \geq 0.$$

Then for $\tilde{r} > 0$ we define the weighted phase norms

$$_{\tilde{r}}|W|_{a,\tilde{p}} = |X| + \frac{1}{\tilde{r}^2}|Y| + \frac{1}{\tilde{r}}||U||_{a,\tilde{p}}$$

for $W = (X, Y, U) \in \mathcal{P}^{a,\tilde{p}}$. Let $\Pi \subset \mathbb{R}^n$ be compact and of positive Lebesgue measure. For a map $W : D(s, r) \times \Pi \to \mathcal{P}^{a,\tilde{p}}$, set

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

and

$$_{\tilde{r}}|W|^{\mathcal{L}}_{a,\tilde{p},D(s,r)\times\Pi} := \max_{1\leq j\leq n} \sup_{(x,\xi)\in D(s,r)\times\Pi} {}_{\tilde{r}}|\partial_{\xi_j} W(x, \xi)|_{a,\tilde{p}}, \quad \xi = (\xi_1, ..., \xi_n).$$

Denote by $\mathcal{B}(\ell^{a_*,p_*}, \ell^{a_*,p^*})$ the set of all bounded linear operators from ℓ^{a_*,p_*} to ℓ^{a^*,p^*} and by $|||\cdot|||_{a_*,a^*,p_*,p^*}$ the operator norm.

In the whole of this paper, by C or c a universal constant, whose size may be different in different place. If $f \leq Cg$, we write this inequality as $f \lessdot g$ when we dot not care the size of the constant C. Similarly, if $f \geq Cg$ we write $f \gtrdot g$.

2.2. The statement of the KAM theorem. For two vectors $b, c \in \mathbb{C}^k$ or $\mathbb{R}^k$, we write $(b, c) = \sum_{j=1}^k b_j c_j$. Consider an infinitely dimensional Hamiltonian in the parameter dependent normal form

$$N_0 = (\omega^0(\xi), y) + \sum_{j\in\mathbb{Z}^d} \Omega_j^0(\xi) u_j^2, \quad (x, y, u) \in \mathcal{P}$$

where $u_j^2 = u_{j_1}^2 + u_{j_2}^2$ with $u_j = (u_{j_1}, u_{j_2})$ and the phase space $\mathcal{P}$ is endowed with the symplectic form

$$dx \wedge dy + \sum_{j\in\mathbb{Z}^d} du_{j_1} \wedge du_{j_2}.$$

The tangent frequencies $\omega^0 = (\omega_1^0, \cdots, \omega_n^0)$ and the normal frequencies $\Omega^0 = (\Omega_j^0 : j \in \mathbb{Z}^d)$ depend on n parameters $\xi \in \Pi_0 \subset \mathbb{R}^n$, Π_0 a given compact set of positive Lebesgue measure. Let $J = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$ and

$$J_m = \mathrm{diag}(\underbrace{J, ..., J}_{m}), \quad J_\jmath = \mathrm{diag}(\underbrace{J, ..., J}_{\jmath^\sharp}), \quad J_\infty = \mathrm{diag}(\underbrace{..., J, ..., J, ...}_{\infty}).$$

The Hamiltonian equation of motion of N_0 are

$$\dot{x} = \omega^0(\xi), \quad \dot{y} = 0, \quad \dot{u} = J_\infty \Omega^0(\xi) u.$$

Hence, for each $\xi \in \Pi_0$, there is an invariant n-dimensional torus $\mathcal{T}_0^n = \mathbb{T}^n \times \{0\} \times \{0\}$ with frequencies $\omega(\xi)$. The aim is to prove the persistence of the torus $\mathcal{T}_0^n$, for "most" (in the sense of Lebesgue measure) values of parameter $\xi \in \Pi_0$, under small perturbations R of the Hamiltonian N_0. To this end the following assumptions are required.

Assumption A: *(Multiplicity.) Give $\nu_0 > 0$. Let $\mathcal{O}_0$ be the ν_0-neighborhood of Π_0 in $\mathbb{R}^n$. Assume that*[3] *for all $\xi \in \mathcal{O}_0$,*

$$\Omega_i^0(\xi) = \Omega_j^0(\xi) \quad \textit{if} \quad |i| = |j|.$$

Set

$$\mathcal{N} = \{|j| : j \in \mathbb{Z}^d\} \subset \mathbb{R}_+.$$

It is easy to see that the set $\mathcal{N}$ is countable. For $\jmath \in \mathcal{N}$, let

$$S_\jmath = \{j \in \mathbb{Z}^d : |j| = \jmath\}$$

and denote by $\jmath^\sharp$ the cardinality of the set $S_\jmath$. it is well known that for $d \geq 2$ we have $\jmath^\sharp \leq \jmath^{d-2+\underline{\varepsilon}}$ where $\underline{\varepsilon} > 0$ is a small constant, and it can be removed if $d \geq 5$. By Assumption A, we can let $\tilde{\Omega}_\jmath^0 = \Omega_j^0$ if $j \in \mathbb{Z}^d$ and $|j| = \jmath$. And let $\tilde{\Omega}^0 = (\tilde{\Omega}_\jmath^0 : \jmath \in \mathcal{N})$ and $\Lambda_\jmath^0 =$ diag $(\tilde{\Omega}_\jmath^0 : |j| = \jmath)$. Notice that $\Lambda_\jmath^0 = \tilde{\Omega}_\jmath^0 E_{\jmath^\sharp}$ where $E_{\jmath^\sharp}$ is the unit matrix of order $\jmath^\sharp$. Let $u_\jmath$ be the vector consisting in the entries of u_j with $|j| = \jmath$. Thus $u_\jmath$ is a vector of order $\jmath^\sharp$. Then we can rewrite N_0 as

$$N_0 = (\omega^0(\xi), y) + \sum_{\jmath\in\mathcal{N}} (\Lambda_\jmath^0 u_\jmath, u_\jmath).$$

Assumption B: *(Non-degeneracy.) There are two absolute constant $c_1, c_2 > 0$ such that*

$$\sup_{\mathcal{O}_0} |det\ \partial_\xi \omega^0(\xi)| \geq c_1, \quad \sup_{\mathcal{O}_0} |\partial_\xi^j \omega| \leq c_2, \ j = 0, 1.$$

[3] This assumption can be relaxed to that the multiplicity of Ω_j^0 is bounded by $\underline{c}|j|^{\bar{c}}$ with constant $\underline{c}, \bar{c} > 0$. However, this general assumption is not necessary in finding quasi-periodic solutions of NLS and NLW equations.

Moreover, assume that both $\omega^0(\xi)$ and $\tilde\Omega^0(\xi)$ are real analytic in each entry ξ_l $(l = 1, ..., n)$ of the variable vector $\xi \in \mathcal{O}_0$.

Assumption C. *(Bounded conditions of Normal frequencies.) Assume that there exists constants $c_3, c_4 > 0$ such that*

$$\inf_{\mathcal{O}_0} \tilde\Omega^0_\jmath \geq c_3, \quad \sup_{\mathcal{O}_0} |\partial_\xi \tilde\Omega^0_\jmath| \leq c_4 \ll 1$$

uniformly for all $\jmath$. In addition, assume there is a constant $c_5 > 0$ such that the following spectra gap conditions hold true:

$$|\tilde\Omega^0_\imath(\xi) - \tilde\Omega^0_\jmath(\xi)| \geq c_5 \imath^{-d} \jmath^{-d}, \quad \imath > \jmath, \ \forall\ \xi \in \mathcal{O}_0.$$

Assumption D: *(Regularity.) Give s_0, r_0, and $0 < \epsilon_0 \ll 1$. Let $\epsilon_m = \epsilon_0^{\wedge}(4/3)^m$ and $\varsigma_m = \epsilon_m^{4/(2\kappa-d)}$. Assume the perturbation $R^0(x, y, u; \xi)$ can be decomposed into*

$$R^0 = \sum_{m=0}^{\infty} R^{0m}(x, y, u; \xi),$$

and each term R^{0m} is defined on the domain $D(s_0, r_0) \times \mathcal{O}_0$ is analytic in the space coordinates and also analytic in each entry ξ_l $(l = 1, ..., n)$ of the parameter vector $\xi \in \mathcal{O}_0$, and is real for real argument, as well as, for each $\xi \in \mathcal{O}_0$ its Hamiltonian vector field $X_{R^{0m}} := (R^{0m}_y, -R^{0m}_x, J_\infty R^{0m}_u)^T$ defines a analytic map

$$X_{R^{0m}} : D(s_0, r_0) \subset \mathcal{P} \to \mathcal{P}^{\varsigma_m, p}.$$

Also assume that $X_{R^{0m}}$ is analytic in each entry of $\xi \in \mathcal{O}_0$.

Theorem 2.1. *Suppose $H = N_0 + R^0$ satisfies assumptions A, B, C and D, and smallness assumption:*

$${}_{r_0}|X_{R^{0m}}|_{\varsigma_m, p, D(s_0,r_0)\times\mathcal{O}_0} < \epsilon_m, \quad {}_{r_0}|X_{R^{0m}}|^{\mathcal{L}}_{\varsigma_m, p, D(s_0,r_0)\times\mathcal{O}_0} < \epsilon_m^{1/3}, \ m = 0, 1, 2, ...$$

Then, for given $\alpha \ll 1$, there is a Cantor set $\Pi_\alpha \subset \Pi_0$ with

$$\text{Meas}\ \Pi_\alpha \geq (\ \text{Meas}\ \Pi_0)(1 - O(\alpha)),$$

a family of torus embedding $\Phi : \mathbb{T}^n \times \Pi_\alpha \to \mathcal{P}$ and a map $\omega_* : \Pi_\alpha \to \mathbb{R}^n$ where $\Phi(\cdot, \xi)$ and $\omega_*(\xi)$ is analytic in each entry ξ_j of the parameter vector $\xi = (\xi_1, ..., \xi_n)$ for other arguments fixed, such that for each $\xi \in \Pi_\alpha$ the map Φ restricted to $\mathbb{T}^n \times \{\xi\}$ is a analytic embedding of a rational torus with frequencies $\omega_*(\xi)$ for the Hamiltonian H at ξ.

Each embedding is analytic on $D(s_0/2) := \{x \in \mathbb{C}^n : |\Im x| < s_0/2\}$, and

$${}_{r_0}|\Phi - \Phi_0|_{0,p,D(s_0/2)\times\Pi_\alpha} \leq c\epsilon_0, \quad {}_{r_0}|\Phi - \Phi_0|^{\mathcal{L}}_{0,p,D(s_0/2)\times\Pi_\alpha} \leq c\epsilon_0^{1/3},$$

$$|\omega_* - \omega| \leq c\epsilon_0, \quad |\omega_* - \omega|^{\mathcal{L}} \leq c\epsilon_0^{1/3},$$

where Φ_0 is the trivial embedding $\mathbb{T}^n \times \Pi_0 \to \mathbb{T}^n \times \{0\} \times \{0\}$, and $c > 0$ is a constant depending on n, α, and ${}_{r_0}|\cdot|_{0,p,D(s_0/2)\times\Pi_\alpha}$ is defined in the way similar to ${}_r|\cdot|_{a,p,D(s,r)\times\Pi}$.

3. Application to nonlinear wave equations of higher dimension.

For technical simplicity, we consider (1.3) instead of (1.2). Essentially, our results hold true for (1.4). We study equation (1.3) as an infinitely dimensional Hamiltonian system. Since the quasi-periodic solutions to be constructed are of small amplitude, (1.3) may be considered as the linear equation $u_{tt} = \mathcal{A}u$ with a small nonlinear perturbation $\Psi((\Psi u)^3)$ where $\mathcal{A} = -\triangle + M_\xi$. Let $\phi_j(x)$ and μ_j^0 $(j \in \mathbb{Z}^d)$ be the eigenfunctions and eigenvalues of the operator $\mathcal{A}$, respectively. By a simple computation,

$$\phi_j(x) = \frac{\sqrt{2}}{(2\pi)^{d/2}} \sin(j, x),$$

and

$$\mu_j^0 = |j|^2 + \xi_j,\ |j|^2 = j_1^2 + \cdots + j_d^2.$$

Then every solution of the linear system is the superposition of their harmonic oscillations and of the form

$$u(t,x) = \sum_{j \in \mathbb{Z}^d} q_j(t)\phi_j(x), \qquad q_j(t) = y_j^0 \cos(\sqrt{\mu_j^0} t)$$

with amplitude $y_j^0 \geq 0$. The solution $u(t,x)$ is time-periodic, quasi-periodic or almost periodic of the linear equation, depending on whether one, finitely many or infinitely many modes are excited, respectively. In particular,

$$N_n = \{j \in \mathbb{Z}^d : 0 \leq |j| \leq n_0\},$$

where $n_0 \in \cup_\jmath S_\jmath = \cup_\jmath \{\jmath = |j| : j \in \mathbb{Z}^d\}$ is given and $n = \sum_{0 \leq \jmath \leq n_0} \jmath^\sharp$. The reason why we choose this N_n is just for convenience. Essentially, we can choose any finite subset N_n of $\mathbb{Z}^d$. Consider the Fourier multiplier M_ξ in (1.5). Let

$$\begin{cases} \xi_j \in [1,2], j \in N_n \\ \xi_j = 0, \ \text{otherwise}\ . \end{cases}$$

Observe that $(\xi_j : j \in N_n)$ is a vector of dimension n. For convenience, we write $\xi = (\xi_j : j \in N_n) = (\xi_1, ..., \xi_l, ..., \xi_n)$. Note that the eigenvalues μ_j^0 depends on ξ. Let $\Pi_0 = [1,2]^n$ and $\mathcal{O}_0$ be the ν_0-neighborhood of Π_0 in $\mathbb{R}^n$. Write

$$\{\sqrt{\mu_j^0(\xi)} : j \in N_n\} = \{\omega_l^0(\xi) : 1 \leq l \leq n\}$$

and

$$\{\phi_j(x) : j \in N_n\} = \{\phi_l^0 : 1 \leq l \leq n\}.$$

Let $\omega^0 = (\omega_1^0, ..., \omega_n^0)$. Then

$$u_0(t,x) = \sum_{l=1}^n y_l^0 \cos \omega_l t \cdot \phi_l^0(x)$$

is a quasi-periodic solution of the linear equation $u_{tt} = -\mathcal{A}u$ for any $\xi \in \Pi_0$ and $y^0 = (y_1^0, ..., y_n^0) \in \mathbb{R}_+^n$. Upon restoring the nonlinearity $\Psi((\Psi u)^3)$ the quasi-periodic solutions will not persist in their entirety due to resonance among the modes and the strong perturbing effect of $\Psi((\Psi u)^3)$ for large amplitudes. In a sufficiently small neighborhood of $u=0$ in the space $H^p([0,2\pi])$, however, it will be shown that there does persist the quasi-periodic solutions $u_0(t,x)$'s which are only slightly deformed for "most" $\xi\in\Pi_0$.

We study the nonlinear NLW equation (1.3) as an infinite dimensional Hamiltonian system. Since the solutions to be constructed are of small amplitude, we can rewrite (1.3) as

$$u_{tt}-\triangle u+M_\xi u+\epsilon\Psi((\Psi u)^3)=0,\quad \text{in } \mathbb{R}\times(0,2\pi)^d \tag{1.3*}$$

by re-scaling $u=\sqrt{\epsilon}u$. To apply Theorem 2.1, we let $\epsilon=\epsilon_0$.

As the phase space one may take, for example, the product of the usual Sobolev space $H_0^1([0,2\pi]^d)\times L^2([0,2\pi]^d)$ with coordinates u and $v=u_t$. The Hamiltonian is then

$$H=\frac{1}{2}\langle v,v\rangle+\frac{1}{2}\langle\mathcal{A}u,u\rangle+\frac{\epsilon}{4}\int_0^{2\pi}(\Psi u)^4\,dx \tag{3.0}$$

where $\langle\cdot,\cdot\rangle$ denotes the usual scalar product in L^2. Here the Hamiltonian structure is $du\wedge dv$. Note that the Dirichlet boundary condition (1.4) is equivalent to

$$x\in\mathbb{T}^d=\mathbb{R}^d/(2\pi\mathbb{Z})^d \quad\text{and}\quad u(-x)=-u(x).$$

Let $L_0^2(\mathbb{T}^d)$ be the subspace of $L^2(\mathbb{T}^d)$ satisfying $u(x)=-u(-x)$, and ℓ_{20} be the subspace of ℓ_2 satisfies $q_j=-q_{-j}$. Let

$$\mathcal{F}:\ \ell_{20}\to L_0^2,\quad q\mapsto\mathcal{F}q=\sum_{j\in\mathbb{Z}^d}q_je^{\sqrt{-1}(j,x)},\ q_{-j}=-q_j$$

be the inverse discrete Fourier transform, which defines an isometry between the two space, and $\mathcal{F}$ can be extended into a isometry from ℓ_2 to L^2. It is obvious that $q\in\ell^{0,p}\subset\ell_2$ if and only if $\mathcal{F}q\in H^p([0,2\pi]^d)\subset L^2([0,2\pi]^d)$. Let

$$\tilde{\Psi}q=\mathcal{F}^{-1}\Psi(\mathcal{F}q),\ \forall\ q\in\ell^{0,p}.$$

Since ψ is even, we have $\Psi(\mathcal{F}q)\in L_0^2$ if $\mathcal{F}(q)\in L_0^2$.

Formally, letting

$$u=\sum_{j\in\mathbb{Z}^d}\tilde{q}_j(t)\phi_j(x) \tag{3.1}$$

and inserting it into (1.3*) and noting $\{\phi_j:j\in\mathbb{Z}^d\}$ is a real basis of L_0^2 we get

$$\frac{d^2\tilde{q}_j}{dt^2}+\mu_j^0\tilde{q}_j+\epsilon\langle\Psi((\Psi u)^3),\phi_j\rangle=0,\quad j\in\mathbb{Z}^d. \tag{3.2}$$

Let

$$q_j=\sqrt{\mu_j^0}\tilde{q}_j,\quad p_j=\frac{1}{\sqrt{\mu_j^0}}\frac{d\tilde{q}_j}{dt},\quad j\in\mathbb{Z}^d. \tag{3.3}$$

Then we get a Hamiltonian system

$$\dot{p}_j=-\frac{\partial H}{\partial q_j},\quad \dot{q}_j=\frac{\partial H}{\partial p_j},\quad j\in\mathbb{Z}^d \tag{3.4}$$

where

$$H(p,q)=\frac{1}{2}\sum_{j\in\mathbb{Z}^d}\sqrt{\mu_j^0}(p_j^2+q_j^2)+G(\tilde{\Psi}q) \tag{3.5}$$

with

$$G(q)=\sum_{i,j,k,l}G_{ijkl}q_iq_jq_kq_l \tag{3.6}$$

$$G_{ijkl}=\frac{\epsilon}{4}(\mu_i^0\mu_j^0\mu_k^0\mu_l^0)^{-1/2}\int_{[0,2\pi]^d}\phi_i\phi_j\phi_k\phi_l dx. \tag{3.7}$$

Since $\phi_j(x)=\sin(j,x)$, it is not difficult to verify that $G_{ijkl}=0$ unless $i\pm j\pm k\pm l=0$ for some combination of plus and minus signs. Hence the sum in (3.6) is restricted to indices i,j,k,l such that $i\pm j\pm k\pm l=0$. Let $\partial_q G$ and $\partial_q^2 G$ are the first and second derivatives of G, respectively. Then, obviously,

$$\partial_q G(q)=(\partial_{q_l}G)_{l\in\mathbb{Z}^d},\ \partial_{q_l}G=4\sum_{\pm i\pm j\pm k=l}G_{ijkl}q_iq_jq_k, \tag{3.8}$$

and

$$\partial_q^2 G(q)=\left(\frac{\partial^2 G}{\partial_{q_k}\partial_{q_l}}\right)_{k,l\in\mathbb{Z}^d},\ \frac{\partial^2 G}{\partial_{q_k}\partial_{q_l}}=12\sum_{\pm i\pm j=\pm k+l}G_{ijkl}q_iq_j, \tag{3.9}$$

Lemma 3.1. *For any $a\geq 0$, $p>d/2$ and $q\in\ell^{a,p}$, we have $\partial_q G(q)\in\ell^{a,p}$ with*

$$||\partial_q G(q)||_{a,p}\lessdot\epsilon||q||_{a,p}^3; \tag{3.10}$$

moreover, $\partial_q^2 G(q)$ is a bounded linear operator from $\ell^{a,p}$ to $\ell^{a,p}$ with

$$|||\partial_q^2 G(q)|||_{a,a,p,p}\lessdot\epsilon||q||_{a,p}^2. \tag{3.11}$$

Proof. Without loss of generality, we assume the sum in (3.8) is restricted to $i+j-k=l$ and the sum in (3.9) is restricted to $i-j=k-l$. For convenience, let $|k|=1$ if k is a zero vector. Let

$$\eta_{ik}(l)=\frac{|i+k-l||i||k|}{|l|}e^{(a/p)(|i+k-l|+|i|+|k|-|l|)} \tag{3.12}$$

It is easy to verify that for any $l\in\mathbb{Z}^d$ and $p>d/2$,

$$\sum_{i,k\in\mathbb{Z}^d}\frac{1}{\eta_{ik}(l)^{2p}}\lessdot 1. \tag{3.13}$$

For any $u,v,w\in\ell^{a,p}$, let $S(u,v,w)=(S_l)_{l\in\mathbb{Z}^d}$ with

$$S_l=\sum_{i-j+k=l}G_{ijkl}u_iu_jw_k. \tag{3.14}$$

By the Schwarz inequality,

$$\begin{aligned}
&||S(u,v,w)||_{a,p}^2 \\
&= \sum_l |l|^{2p} e^{2a|l|} \left| \sum_{i-j+k=l} G_{ijkl} u_i v_j w_k \right|^2 \\
&= \sum_l G_{ijkl} |l|^{2p} e^{2a|l|} \left| \sum_{i,k\in\mathbb{Z}^d} \frac{\eta_{ij}(l)^p u_i v_{i+k-l} w_k}{\eta_{ij}(l)^p} \right|^2 \qquad (3.15)\\
&\le C\epsilon \sum_l \sum_{i,k\in\mathbb{Z}^d} |i+k-l|^{2p} e^{2a|i+k-l|} |v_{i+k-l}|^2 |i|^{2p} e^{2a|i|} |u_i|^2 |k|^{2p} e^{2a|k|} |w_k|^2 \\
&\le C\epsilon ||u||_{a,p}^2 ||v||_{a,p}^2 ||w||_{a,p}^2.
\end{aligned}$$

Note $\partial_q G(q) = S(q,q,q)$. Thus, the proof of (3.10) is completed by (3.15). For any $u \in \ell^{a,p}$, observing that

$$\sum_l \sum_{i-j=k-l} G_{ijkl} q_i q_j u_l = \sum_{i-j+l=k} G_{ijkl} q_i q_j u_l,$$

we get

$$(\partial_q^2 G(q))u = S(q,q,u).$$

By (3.15),

$$||(\partial_q^2 G(q))u||_{a,p} \le C\epsilon ||q||_{a,p}^2 ||u||_{a,p}.$$

This implies that (3.11) holds true. □

In Sect. 10, we will construct a family of operators $T_m : \ell_2 \supset \ell^{0,\bar{p}} \to \ell^{a,p} \subset \ell_2$ which satisfy Lemma B.3. Now we introduce a Hamiltonian R:

$$R(q) = R_0(q) + \sum_{m=1}^{\infty} R_m(q) := G(T_0\tilde{\Psi}(q)) + \sum_{m=1}^{\infty} (G(T_m\tilde{\Psi}(q)) - G(T_{m-1}\tilde{\Psi}(q))) \tag{3.16}$$

Lemma 3.2. *For $q \in \ell^{0,p}$ with $||q||_{0,p} \lessdot 1$, we have that*

$$G(\tilde{\Psi}(q)) = R(q) \tag{3.17}$$

$$||\partial_q R_m(q)||_{\varsigma_m,p} \lessdot \epsilon_m, \tag{3.18}$$

where $\epsilon_m = \epsilon^\wedge(4/3)^m$ and $\varsigma_m = \epsilon_m^{4/(2\kappa-d)}$, $(m = 0,1,2,...)$

Proof. For $\theta \in [0,1]$, let

$$q^* := \tilde{\Psi}(q) + \theta(T_m\tilde{\Psi}(q) - \tilde{\Psi}(q)).$$

Note that for any $a \ge 0$ and $\bar{p} \ge p > d/2$, we have

$$||q||_{a,p} \ge ||q||_{0,p} \ge ||q||_{\ell_2}, \quad q \in \ell^{a,p},$$

and
$$||q||_{0,p} \le ||q||_{0,\bar p}, \quad q \in \ell^{0,\bar p}.$$
By the definition of Ψ, we have
$$||\tilde\Psi(q)||_{0,p} \le ||\tilde\Psi(q)||_{0,\bar p} \lessdot ||q||_{0,p}.$$
In view of Lemma B.3 (10.13),
$$||\theta(T_m\tilde\Psi(q) - \tilde\Psi(q))||_{0,p} \le ||\tilde\Psi(q)||_{0,\bar p} \lessdot ||q||_{0,p}.$$
Thus,
$$||q^*||_{0,p} \lessdot ||q||_{0,p}.$$
Using Talylor's formula, we get
$$\begin{aligned}
&|G(T_m\tilde\Psi(q)) - G(\tilde\Psi(q))|\\
=&|\langle \partial_q G(q^*), (T_m - 1)\tilde\Psi(q)\rangle_{\ell_2}|\\
\le&||\partial_q G(q^*)||_{\ell_2}||(T_m - 1)\tilde\Psi(q)||_{\ell_2}\\
\le&||\partial_q G(q^*)||_{0,p}||(T_m - 1)\tilde\Psi(q)||_{0,p}\\
\lessdot&||q^*||^3_{0,p}\epsilon_{m+1}||\tilde\Psi(q)||_{0,\bar p} \Leftarrow \text{ Lemma 3.1, B.3(10.13)}\\
\le&||q||^4_{0,p}\epsilon_{m+1} \lessdot \epsilon_{m+1} \to 0, \quad \text{as } m \to \infty.
\end{aligned}$$

This proves (3.17). We are position to show (3.18). Let $q_m = T_m\tilde\Psi(q)$ and $n = m-1$. By Lemma B.2, $||q_m||_{0,p} \le ||q_m||_{\varsigma_m,p} \le ||q||_{0,p}$. It will be shown in Lemma B.4 in Sect. 10 that $T_m\tilde\Psi$ is self-adjoint in ℓ_2. Then
$$\partial_q R_0 = T_0\tilde\Psi(\partial_{q_0}G(q_0),$$
$$\partial_q R_m = T_n\tilde\Psi\partial_{q_n}G(q_n) - T_m\tilde\Psi\partial_{q_m}G(q_m).$$
Furthermore, by (3.10) and Lemma B.2(10.14),
$$\begin{aligned}
&||\partial_q R_0||_{\varsigma_0,p} \le ||\tilde\Psi\partial_{q_0}G(q_0)||_{0,p} \le ||\tilde\Psi(\partial_{q_0}G(q_0)||_{0,\bar p}\\
&\le ||\partial_{q_0}G(q_0)||_{0,p} \le \epsilon||q_0||^3_{0,p} \lessdot \epsilon = \epsilon_0,
\end{aligned}$$

and
$$\begin{aligned}
||\partial_q R_m||_{\varsigma_m,p} &= ||\partial_q(G(T_m\tilde\Psi(q) - G(T_n\tilde\Psi(q))||_{\varsigma_m,p}\\
&= ||T_m\tilde\Psi\partial_{q_m}G(q_m) - T_n\tilde\Psi\partial_{q_n}G(q_n)||_{\varsigma_m,p}\\
&\le ||(T_m - T_n)\tilde\Psi\partial_{q_m}G(q_m)||_{\varsigma_m,p} \qquad (3.19)\\
&\quad + ||T_n\tilde\Psi\left((\partial_{q_m}G(q_m) - \partial_{q_n}G(q_n)\right)||_{\varsigma_m,p}. \qquad (3.20)
\end{aligned}$$

Thus,
$$\begin{aligned}
(3.19) &\lessdot \epsilon_m||\tilde\Psi\partial_{q_m}G(q_m)||_{0,\bar p} \ (\Leftarrow \text{ Lemma B.(10.12)})\\
&\lessdot \epsilon_m||\partial_{q_m}G(q_m)||_{0,p} \ (\Leftarrow \text{ Definition of } \Psi)\\
&\lessdot \epsilon_m||q_m||^3_{0,p} \ (\Leftarrow \ (3.10))\\
&\le \epsilon_m||q||^3_{0,p}
\end{aligned}$$

Let $q^* := q_n + \theta(q_m - q_n)$ with $\theta \in [0,1]$. We have

$$
\begin{aligned}
(3.20) &\le ||T_n\tilde{\Psi}\left((\partial_{q_m}G(q_m) - \partial_{q_n}G(q_n)\right)||_{\varsigma_n,p}, \quad \varsigma_m < \varsigma_n \\
&\lessdot ||\tilde{\Psi}(\partial_{q_m}G(q_m) - \partial_{q_n}G(q_n))||_{0,p} \ (\Leftarrow \text{ Lemma B.3(10.14)}) \\
&\le ||\tilde{\Psi}(\partial_{q_m}G(q_m) - \partial_{q_n}G(q_n))||_{0,\bar{p}} \\
&\lessdot ||\partial_{q_m}G(q_m) - \partial_{q_n}G(q_n)||_{0,p} \ (\Leftarrow \text{ Definition of } \Psi) \\
&= ||\partial^2_{q^*}G(q^*)(q_m - q_n)||_{0,p} \ (\Leftarrow \text{ Taylor's formula }) \\
&\le |||\partial^2_{q^*}G(q^*)|||_{0,0,p,p}||q_m - q_n||_{0,p} \\
&\le \epsilon||q^*||^2_{0,p}||(T_m - T_n)\tilde{\Psi}(q)||_{0,p} \\
&\le \epsilon||q||^2_{0,p}||(T_m - T_n)\tilde{\Psi}(q)||_{\varsigma_m,p} \\
&\le \epsilon\epsilon_m||q||^2_{0,p}||\tilde{\Psi}(q)||_{0,\bar{p}} \ (\Leftarrow \text{ Lemma B.(10.12)}) \\
&\le \epsilon\epsilon_m||q||^3_{0,p}.
\end{aligned}
$$

Consequently, if $||q||_{0,p} \lessdot 1$

$$||\partial_q R_m||_{\varsigma_m,p} \le \epsilon_m||q||^3_{0,p} \lessdot \epsilon_m.$$

This completes the proof of this lemma. $\square$

Observe that for $1 \le l \le n$, there is a $j \in N_n$ such that

$$\omega^0_l(\xi) = \sqrt{|j|^2 + \xi_l}, \tag{3.21}$$

and for $j \in \mathbb{Z}^d \setminus N_n$,

$$\Omega^0_j = \sqrt{\mu^0_j} = \jmath. \tag{3.22}$$

It follows from (3.21,22) that Assumptions A, B, C of Theorem 2.1 are fulfilled. Now let us check Assumption D of Theorem 2.1. Write $q = (\underline{q}, \bar{q})$ with $\underline{q} = (q_j)_{j\in N_n}$ and $\bar{q} = (q_j)_{j\notin N_n}$. Let

$$q_j = \sqrt{2(y^0_j + y_j)}\cos x_j, \ p_j = \sqrt{2(y^0_j + y_j)}\sin x_j, \ j \in N_n, y \in [0,1]^n. \tag{3.23}$$

Then, in view of (3.17), Hamiltonian (3.5) is transformed into

$$H = (\omega(\xi), y) + \frac{1}{2}\sum_{j\notin N_n}\Omega^0_j(p^2_j + q^2_j) + R^0(x, y, \bar{q}) \tag{3.24}$$

where

$$R^0(x, y, \bar{q}) = R(\underline{q}(x, y), \bar{q}) = \sum_{m=0}^{\infty} R_m(\underline{q}(x, y), \bar{q})$$

with $\underline{q}(x, y)$ defined by (3.23). Observe that for $|\Im x| \le s_0, |y| \le r_0 < 1$,

$$|\partial_x R_m(x, y, \bar{q})|, |\partial_y R_m(x, y, \bar{q})|, ||\partial_{\bar{q}} R_m(x, y, \bar{q})||_{a,p} \lessdot ||\partial_q R(q)||_{a,p}.$$

Let $u = (u_j : j \notin N_n)$ with $u_j = (p_j, q_j)$ and

$$X_{R_m} = (\partial_x R_m, -\partial_y R_m, J_\infty \partial_u R_m), \quad \partial_{u_j} = (\partial_{p_j}, \partial_{q_j}).$$

It follows from Lemma 3.2 that the Assumption D is fulfilled and

$${}_{r_0}|X_{R_m}|_{\varsigma_m, p, D(s_0,r_0)\times\mathcal{O}_0} \lessdot \epsilon_m.$$

Using the fact $\mu_j^0 = |j|^2 + \xi_j$ and (3.6) and (3.7), we get that the vector field X_{R_m} is analytic in each entry of $\xi \in \mathcal{O}_0$ and

$${}_{r_0}|X_{R_m}|^{\mathcal{L}}_{\varsigma_m, p, D(s_0,r_0)\times\mathcal{O}_0} \lessdot \epsilon_m.$$

By invoking Theorem 2.1, we get the invariant torus and thus quasi-periodic solutions for (1.3).

Theorem 3.3. *For any $0 < \alpha \ll 1$, there ia a set $\Pi_\alpha \subset \Pi_0$ with*

$$\textit{Meas}\ (\Pi \setminus \Pi_\alpha) \le \tilde{c}\alpha$$

(here $\tilde{c} > 0$ is an absolute constant) such that for any $\xi \in \Pi_\alpha$, the NLW equation (1.3)$_{\xi\in\Pi_\alpha}$ possesses a smooth quasi-periodic solution $u(t,x)$ of frequencies ω_* which satisfies

$$|u(t,x) - u_0(t,x)| \le \sqrt{\epsilon}$$

and

$$|\omega_* - \omega_0| \le \epsilon.$$

Besides, the solution $u(t,x)$ is linearly stable.

4. The linearized equation.

4.1. split and estimate for small perturbation. Recall that for $\jmath \in \mathcal{N}$, the notation $\jmath^\sharp$ denotes the number of the elements of the set $\{j \in \mathbb{Z}^d : |j| = \jmath\}$, and $u_\jmath$ is the vector consisting of u_j with $j \in \mathbb{Z}^d$ and $|j| = \jmath$. Let $E_\jmath$ be the unit matrix of order $\jmath^\sharp$. Let $\mathcal{O}$ be an open set in $\mathbb{R}^n$. Consider two infinitely dimensional vectors $u = (u_j)_{j\in\mathbb{Z}^d}$ and $v = (v_j)_{j\in\mathbb{Z}^d}$ where both u_j and v_j are in $\mathbb{C}^2$. Define $\langle u, v\rangle = \sum_{j\in\mathbb{Z}^d}(u_j, v_j)$. Therefore, if write $u = (u_\jmath)_{\jmath\in\mathcal{N}}$ and $v = (v_\jmath)_{\jmath\in\mathcal{N}}$ where both $u_\jmath$ and $v_\jmath$ are $\jmath^\sharp$-dimensional vectors, then $\langle u, v\rangle = \sum_{\jmath\in\mathcal{N}}(u_\jmath, v_\jmath)$. Let N be an integrable Hamiltonian:

$$N = (\omega(\xi), y) + \sum_{\jmath\in\mathcal{N}} \langle \tilde{\Omega}_\jmath(\xi) E_\jmath u_\jmath, u_\jmath\rangle + \sum_{\jmath\in\mathcal{N}} \langle B_{\jmath\jmath}(\xi) u_\jmath, u_\jmath\rangle$$

where $B_{\jmath\jmath}(\omega)$ is a real symmetric matrix of order $\jmath^\sharp$ for any $\omega \in \mathcal{O}$, and all of the coefficients $\omega(\xi), \tilde{\Omega}_\jmath(\xi)$ and $B_{\jmath\jmath}(\xi)$ are analytic in each entry ξ_j $(j = 1, ..., n)$ of $\xi \in \mathcal{O}$. Moreover, we assume $|\det \frac{\partial\omega(\xi)}{\partial_\xi}| > c > 0$ for all $\xi \in \mathcal{O}$. If we write $\Lambda = \mathrm{diag}(\tilde{\Omega}_\jmath(\xi) E_\jmath : \jmath \in \mathcal{N})$ and $B = \mathrm{diag}(B_{\jmath\jmath} : \jmath \in \mathcal{N})$, then

$$N = (\omega(\xi), y) + \langle \Lambda u, u\rangle + \langle Bu, u\rangle.$$

We now consider a perturbation $H = N + \grave{R}$ where $\grave{R} = \grave{R}(x, y, u; \xi)$ is a Hamiltonian defined on $D(s, r)$ and depends on the parameter $\xi \in \mathcal{O}$. We assume that there are quantities $\varepsilon = \varepsilon(r, s, \mathcal{O})$ and $\varepsilon^{\mathcal{L}} = \varepsilon^{\mathcal{L}}(r, s, \mathcal{O})$ which are dependent on $r, s, \mathcal{O}$ such that

$$ {}_r|X_{\grave{R}}|_{\varsigma,p,D(s,r)\times\mathcal{O}} \lessdot \varepsilon, \quad {}_r|X_{\grave{R}}|^{\mathcal{L}}_{\varsigma,p,D(s,r)\times\mathcal{O}} \lessdot \varepsilon^{\mathcal{L}}, \quad \varepsilon < \varepsilon^{\mathcal{L}} \ll 1. \tag{4.1} $$

For $u = (u_j)_{j\in\mathbb{Z}^d}$ with $u_j = (u_j^1, u_j^2)$, write $u^1 = (u_j^1)_{j\in\mathbb{Z}^d}$ and $u^2 = (u_j^2)_{j\in\mathbb{Z}^d}$. Let

$$ R = \sum_{2|m|+|q_1+q_2|\le 2} \sum_{k\in\mathbb{Z}^d,} R_{kmq_1q_2} e^{\sqrt{-1}(k,x)} y^m (u^1)^{q_1} (u^2)^{q_2}, $$

with the Taylor-Fourier coefficients $R_{kmq_1q_2}$ of $\grave{R}$ depending on $\xi \in \mathcal{O}$, and being analytic in each entry ξ_j of ξ, such that the vector field $X_R : \quad \mathcal{P} \to \mathcal{P}^{\varsigma,p}$ is real, analytic in $(x, y, u) \in D(s, r)$ and in each entry of $\xi \in \mathcal{O}$. We will approximate $\grave{R}$ by its partial Taylor-Fourier expansion R. For convenience we decompose $R = R^0 + R^1 + R^2$, where R^j's $(j = 0, 1, 2)$ comprises all terms with $|q + \bar{q}| = j$, and furthermore,

$$ \begin{aligned} R^0 &= R^x + (R^y, y), \\ R^1 &= \langle R^u, u\rangle, \\ R^2 &= \langle R^{uu}u, u\rangle, \end{aligned} $$

where R^x, R^y, R^{uu} depend on x, ξ. Let $D(s) = \{x \in \mathbb{C}^n/2\pi\mathbb{Z}^n : |\Im x| < s\}$. In order to derive the linearized equation, we need some notations. For any operator $Y : \ell^p \to \ell^{\varsigma,p} \subset \ell^p$, we regard it as a matrix of infinite dimension. Denote by Y^{ij}'s the elements of this matrix. For any $\imath, \jmath \in \mathcal{N}$, let $Y_{\imath\jmath}$ is the sub-matrix of Y with $Y_{\imath\jmath} = (Y^{ij})_{|i|=\imath,|j|=\jmath}$. Denote by $Y^{ij}_{\imath\jmath}$ the elements of the sub-matrix $Y_{\imath\jmath}$. We split the matrix Y as follows: $Y = Y_g + Y_{ng}$ where R_g is a quasi-diagonal matrix with $Y_g = (Y_{\jmath\jmath} : \jmath \in \mathcal{N})$ and Y_{ng} is a non-diagonal matrix with $Y_{ng} = Y - Y_g$. Denote by Y^{ij}_{ng} the elements of matrix Y_{ng}. Thus, $Y^{ij}_{ng} = 0$ if $|i| = |j| = \jmath$. For any vector or matrix Y dependent on $x \in D(s)$, let

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

Besides, we suppose that $[R^x] = 0$ without loss of generality, since the Hamiltonian dynamics will not be changed by adding (or removing) a constant to (or from) the Hamiltonian function. Now we give some estimates of R.

Lemma 4.0.

$$ {}_r|X_R|^*_{\varsigma,p,D(s,r)\times\mathcal{O}} \lessdot {}_r|X_{\grave{R}}|^*_{\varsigma,p,D(s,r)\times\mathcal{O}} \le \varepsilon^*, \tag{4.2} $$

$$ {}_{\eta r}|X_R - X_{\grave{R}}|^*_{\varsigma,p,D(s,4\eta r)\times\mathcal{O}} \lessdot \eta \cdot {}_r|X_{\grave{R}}|^*_{\varsigma,p,D(s,r)\times\mathcal{O}} \lessdot \eta\varepsilon^*, \tag{4.3} $$

for any $0 < \eta \ll 1$, where * = the blank or $\mathcal{L}$, for example, $\varepsilon^* = \varepsilon$ or $\varepsilon^{\mathcal{L}}$.

Proof. The proof is similar to that of formula (7) of [P1,129].

Lemma 4.1. *Under the smallness assumption on $\grave{R}$, the following estimates hold true:*

$$ |\partial_x R^x|_{D(s)\times\mathcal{O}} \le r^2\varepsilon, \quad |\partial_x R^x|^{\mathcal{L}}_{D(s)\times\mathcal{O}} \le r^2\varepsilon^{\mathcal{L}} \tag{4.4} $$

$$ |R^y|_{D(s)\times\mathcal{O}} \le \varepsilon, \quad |R^y|^{\mathcal{L}}_{D(s)\times\mathcal{O}} \le \varepsilon^{\mathcal{L}} \tag{4.5} $$

$$||R^u||_{\varsigma,p;D(s)\times\mathcal{O}} \le r\varepsilon, \quad ||R^u||^{\mathcal{L}}_{\varsigma,p;D(s)\times\mathcal{O}} \le r\varepsilon^{\mathcal{L}} \tag{4.6}$$

$$|||R^{uu}|||_{0,\varsigma,p,p;D(s)\times\mathcal{O}} \le \varepsilon, \quad |||R^{uu}|||^{\mathcal{L}}_{0,\varsigma,p,p;D(s)\times\mathcal{O}} \le \varepsilon^{\mathcal{L}} \tag{4.7}$$

Proof. Consider R^{uu}. Observe that $R^{uu} = \partial_u\partial_u R|_{u=0}$ with $\partial_{u_j} = (\partial_{u_j^1}, \partial_{u_j^2})$ and $u_j = (u_j^1, u_j^2)$. By the generalized Cauchy inequality (See Lemma A.3 in [P1]),

$$|||R^{uu}|||_{0,\varsigma,p,p;D(s)\times\mathcal{O}} \le \frac{1}{r}||\partial_u R||_{\varsigma,p,D(s,r)\times\mathcal{O}} \le {}_r|X_R|_{\varsigma,p,D(s,r)\times\mathcal{O}} < \varepsilon.$$

The remaining proof is simple. We omit the details. $\square$

It follows from Lemma 4.2 that R^{uu} is a bounded linear operator from ℓ^p to $\ell^{\varsigma,p}$ for any $x \in D(s)$. Write $R^{uu} = (R_{ij} : i,j \in \mathbb{Z}^d)$ where R_{ij}'s are 2×2 complex matrix. In fact,

$$R_{ij} = \begin{pmatrix} \partial_{u_i^1}\partial_{u_j^1}R & \partial_{u_i^2}\partial_{u_j^1}R \\ \partial_{u_j^2}\partial_{u_i^1}R & \partial_{u_i^2}\partial_{u_j^2}R \end{pmatrix}.$$

We see that $R^t_{ij} = R_{ji}$ where t means the transpose of matrix. Therefore, R^{uu} is a symmetric operator. Besides, R^{uu} is real for real x. Recall R_{ij} is a 2×2 matrix. Denote by $|R_{ij}|$ the maximum norm of matrix.

Lemma 4.2. *For $|i| \ge |j|$, we have*

$$|R_{ij}|_{D(s)\times\mathcal{O}}, |R_{ji}|_{D(s)\times\mathcal{O}} \lessdot e^{-\varsigma|i|}|i|^{-p}|j|^p\varepsilon. \tag{4.8}$$

$$|R_{ij}|^{\mathcal{L}}_{D(s)\times\mathcal{O}}, |R_{ji}|^{\mathcal{L}}_{D(s)\times\mathcal{O}} \lessdot e^{-\varsigma|i|}|i|^{-p}|j|^p\varepsilon^{\mathcal{L}}. \tag{4.9}$$

Proof. Let $u = (u_k)_{k\in\mathbb{Z}^d}$ with $u_j = |j|^{-p}(1/\sqrt{2}, 1\sqrt{2})$ and $u_k = 0$ for $k \ne j$. Then $||u||_p = 1$. In terms of the definition of the operator norm $|||\cdot|||_{0,\varsigma,,p,p}$ and (4.7), we have

$$\sum_{l\in\mathbb{Z}^d} e^{2\varsigma|l|}|l|^{2p}|\sum_{k\in\mathbb{Z}^d} R_{lk}u_k|^2 \le |||R^{uu}|||^2_{0,\varsigma,p,p} \le \varepsilon^2,$$

that is,

$$\sum_{l\in\mathbb{Z}^d} e^{2\varsigma|l|}|l|^{2p}|R_{lj}|^2|j|^{-2p} \lessdot \varepsilon^2,$$

in particular, for $|i| > |j|$,

$$|R_{ij}| \lessdot e^{-\varsigma|i|}|i|^{-p}|j|^p\varepsilon.$$

The remaining proof is similar. This completes the proof. $\square$

Lemma 4.3. *For $\jmath \in \mathcal{N}$,*

$$\sup_{D(s)\times\mathcal{O}} ||R_{\jmath\jmath}||_2 \lessdot \jmath^{(d-1)/2}e^{-\varsigma\jmath}\varepsilon. \tag{4.10}$$

$$\sup_{D(s)\times\mathcal{O}} ||R_{\jmath\jmath}||^{\mathcal{L}}_2 \lessdot \jmath^{(d-1)/2}e^{-\varsigma\jmath}\varepsilon^{\mathcal{L}}. \tag{4.11}$$

Proof. Observe a well-known fact in matrix theory:

$$||R_{\jmath\jmath}||_2^2 \le \max_{|j|=\jmath} \sum_{|i|=\jmath} |R_{ij}| \cdot \max_{|i|=|j|=\jmath} |R_{ij}|$$

Note that the cardinality of the set $\{j : |j| = \jmath\}$ is bounded by $\jmath^{d-2+\underline{\epsilon}} \le \jmath^{d-1}$. By Lemma 4.2, we have

$$||R_{\jmath\jmath}||_2 \le \sqrt{\sum_{|j|=\jmath} 1} \cdot \max_{|i|=|j|=\jmath} |R_{ij}| \lessdot \jmath^{(d-1)/2} e^{-\varsigma\jmath}\varepsilon$$

The proof of the another estimate is the same.

Let $Y : \ell_2 \to \ell_2$ be a matrix of infinity order. Write $Y = (Y_{ij} : i, j\mathbb{Z}^d)$. For $M > 0$, let $\Upsilon_M Y : \ell_2 \to \ell_2$ be a matrix of infinity order whose matrix elements are defined by

$$(\Upsilon_M Y)_{ij} = \begin{cases} Y_{ij}, & |i| \le M \text{ and } |j| \le M \\ 0, & |i| > M \text{ or } |j| > M. \end{cases}$$

If $Y = (Y_j : j \in \mathbb{Z}^d)$ be a vector in ℓ_2. Let $\Upsilon_M Y = ((\Upsilon Y)_j : j \in \mathbb{Z}^d)$ is a vector in ℓ_2 defined by

$$(\Upsilon_M Y)_j = \begin{cases} Y_j, & |j| \le M \\ 0, & |j| > M. \end{cases}$$

Lemma 4.4. *For $0 < \eta < 1$ and $0 < \sigma < 1$, let $M = 4|\ln \eta\sigma^n|/\varsigma$. Then*

$$||(1 - \Upsilon_M)R^u||_{\varsigma/2,p,D(s)\times\mathcal{O}} \le r\varepsilon\sigma^n\eta^2, \ ||(1 - \Upsilon_M)R^u||^{\mathcal{L}}_{\varsigma/2,p,D(s)\times\mathcal{O}} \le \eta^2 r\sigma^n\varepsilon^{\mathcal{L}}.$$

Proof. By (4.6) and the definition of Υ,

$$||(1 - \Upsilon_M)R^u||_{\varsigma/2,p,D(s)\times\mathcal{O}} \le e^{\varsigma M/2}||R^u||_{\varsigma,p,D(s)\times\mathcal{O}} \le r\varepsilon\eta^2\sigma^n.$$

The remaining inequality is proven similarly. □

Lemma 4.5. *For $0 < \eta < 1$, let $M = 4|\ln \eta\sigma^n|/\varsigma$. Then*

$$|||(1 - \Upsilon_M)R^{uu}|||_{0,\varsigma/2,p,p,D(s)\times\mathcal{O}} \lessdot \sigma^n\varepsilon\eta^2,$$

$$|||(1 - \Upsilon_M)R^{uu}|||^{\mathcal{L}}_{0,\varsigma/2,p,p,D(s)\times\mathcal{O}} \lessdot \sigma^n\eta^2\varepsilon^{\mathcal{L}}.$$

Proof. By (4.8) and the definition of Υ,

$$\begin{aligned} &|((1 - \Upsilon_M)R^{uu})_{ij}|_{D(s)\times\mathcal{O}}, \ |((1 - \Upsilon_M)R^{uu})_{ji}|_{D(s)\times\mathcal{O}} \\ &\le e^{\varsigma M/4} e^{-(3/4)\varsigma|i|}|i|^{-p}|j|^p\varepsilon \le e^{-(3/4)\varsigma|i|}|i|^{-p}|j|^p(\varepsilon\eta^2\sigma^n). \end{aligned}$$

By the same argument as in the proof of Lemma A.2 in Appendix A, we complete the proof of the first inequality of this lemma. The remaining proof is similar. □

Lemma 4.6. *Suppose that $r = r_0\eta$ with a absolute constant $r_0 > 0$ defined in Theorem 2.1. Let*

$$R_M = \langle (1-\Upsilon_M)R^u, u\rangle + \langle (1-\Upsilon_M)R^{uu}u, u\rangle.$$

Then

$$_{\eta r}|X_{R_M}|_{\varsigma/2,p,D(s-\sigma,\eta r)\times\mathcal{O}} \lessdot \eta\varepsilon, \quad _{\eta r}|X_{R_M}|^{\mathcal{L}}_{\varsigma/2,p,D(s-\sigma,\eta r)\times\mathcal{O}} \lessdot \eta\varepsilon^{\mathcal{L}}.$$

Proof. This is a easy corollary of Lemmas 4.4 and 4.5. $\square$

For $K > 0$ and a function $f(x) = \sum_{k\in\mathbb{Z}^n} \hat f(k)e^{\sqrt{-1}(k,x)}$, define a function $\Gamma_K f$ by

$$(\Gamma_K f)(x) = \sum_{|k|\le K} \hat f(k)e^{\sqrt{-1}(k,x)}.$$

Lemma 4.7. *Let $K = |\ln\eta|/\sigma$ and $R_K = (R - R_M) - \Gamma_K(R - R_M)$. We have*

$$_r|X_{R_K}|_{\varsigma,p,D(s-\sigma,r)\times\mathcal{O}} \lessdot \eta\varepsilon,$$

$$_r|X_{R_K}|^{\mathcal{L}}_{\varsigma,p,D(s-\sigma,r)\times\mathcal{O}} \lessdot \eta\varepsilon^{\mathcal{L}}$$

Proof. Write $R_K = R_K^x + (R_K^y, y) + \langle R_K^u, u\rangle + \langle R_K^{uu}u, u\rangle$. Note that the terms R_K^x, R_K^y, and so on, are analytic in $x \in D(s)$. Then by Cauchy's formula, we have $|\widehat{R_K^x}(k)| \le e^{-s|k|}\sup_{D(s)}|R^x|$, and so on. Observe that $|k| > K$ in those Fourier coefficients $\widehat{R_K^x}(k)$'s. We can get

$$_r|X_{R_K}|^{*}_{\varsigma,p,D(s-\sigma,r)\times\mathcal{O}} \le e^{-\sigma K}\ _r|X_R|_{\varsigma,p,D(s,r)\times\mathcal{O}} \le \eta\varepsilon^*$$

where $*=$ the blank or $\mathcal{L}$. $\square$

Finally, let $\mathcal{R} = \Gamma_K(R - R_M)$. Then

$$R = \mathcal{R} + R_M + R_K.$$

Also write

$$\mathcal{R} = \mathcal{R}^x + (\mathcal{R}^y, y) + \langle \mathcal{R}^u, u\rangle + \langle \mathcal{R}^{uu}u, u\rangle.$$

By Lemma 4.6 and 4.7 we see that Lemmas 4.0, 4.1,4.2,4.3 hold still true after replacing R by $\mathcal{R}$.

4.2. Derivation of homological equations. The KAM theorem is proven by the usual Newton-type iteration procedure which involves an infinite sequence of coordinate changes. Each coordinate change is obtained as the time-1 map $X_F^t|_{t=1}$ of a Hamiltonian vector field X_F. Its generating Hamiltonian F solves the linearized equation

$$\{F, N\} = \mathcal{R} - [[\mathcal{R}]]$$

where $\{\cdot,\cdot\}$ is Poisson bracket with respect to the symplectic structure $dx\wedge dy + du^1\wedge du^2$ and $[[\mathcal{R}]]$ is defined as

$$[[\mathcal{R}]] = ([\mathcal{R}^y], y) + \langle [\mathcal{R}^{uu}_g]u, u\rangle.$$

It is easy to see that

$$[[\mathcal{R}]] = \sum_{0\le |m|\le 1, |i|=|j|} R_{0mij} y^m u_i^1 u_j^2$$

which is of the same form as N. We are now in position to find a solution of this equation and give some estimates for the solution. To this end, we suppose that F is of the same form as $\mathcal{R} - [[\mathcal{R}]]$, that is, $F = F^0 + F^1 + F^2$, where

$$\begin{aligned} F^0 &= F^x + \langle F^y, y\rangle, \\ F^1 &= \langle F^u, u\rangle, \\ F^2 &= \langle F^{uu} u, u\rangle, \end{aligned}$$

with F^x, F^y, F^{uu} depending on x, ξ. We furthermore suppose that $[[F]] = 0$ where the definition of $[[F]]$ is the same as $[[\mathcal{R}]]$. Set $A_{\jmath\jmath}(\xi) = \tilde{\Omega}_\jmath(\xi) E_\jmath + B_{\jmath\jmath}(\xi), \quad A = \operatorname{diag}(A_{\jmath\jmath} : \jmath \in \mathcal{N})$.

As in [K1, p.62], now the linearized equation is reduced to the following equations:

$$\partial F^x/\partial\omega = \mathcal{R}^x(x,\xi), \tag{4.12}$$

$$\partial F^y/\partial\omega = \mathcal{R}^y(x,\xi) - [\mathcal{R}^y](\xi), \tag{4.13}$$

$$\partial F^u/\partial\omega - AJ_\infty F^u = \mathcal{R}^u(x,\xi), \tag{4.14}$$

$$\partial F^{uu}/\partial\omega + F^{uu} J_\infty A - AJ_\infty F^{uu} = \mathcal{R}^{uu}(x,\xi) - [\mathcal{R}^{uu}_g](\xi) \tag{4.15*}$$

where $\partial/\partial\omega = (\omega, \frac{\partial}{\partial x})$, for example, $\partial F^y/\partial\omega = (\omega, \frac{\partial F^y}{\partial x})$. Observe that both A and J_∞ are quasi-diagonal. We can split (4.15*) into the following systems:

$$\partial F^{uu}_{\imath\jmath}/\partial\omega + F^{uu}_{\imath\jmath} J_\jmath A_{\jmath\jmath} - A_{\imath\imath} J_\imath F^{uu}_{\imath\jmath} = \mathcal{R}^{uu}_{\imath\jmath}, \quad [\mathcal{R}^{uu}_{\jmath\jmath}] = 0, \imath, \jmath \le M. \tag{4.15}$$

Recall $X_{\mathcal{R}} : D(s,r) \subset \mathcal{P} \to \mathcal{P}^{\varsigma,p}$ is real analytic in $(x,y,u) \in D(s,r)$ and each entry of $\xi \in \mathcal{O}$.

4.3. Solutions of the homological equations.

Proposition 1. *(Solution of (4.12).) Assume that uniformly on $\xi \in \mathcal{O}$,*

$$|(k,\omega(\xi))| \gtrdot \frac{\alpha}{|k|^\tau}, \text{ for all } 0 \ne k \in \mathbb{Z}^n, |k| \le K \tag{4.16}$$

where $\alpha > 0$ and $\tau > n$. Then on $D(s-\sigma)\times\mathcal{O}$ with $0 < \sigma < s$, the equation (4.12) has a solution $F^x(x,\xi)$ which is analytic in $x \in D(s-\sigma)$ for ξ fixed and analytic in each $\xi_j, (j = 1, ..., n)$ for other variables fixed, and which is real for real argument, such that

$$|\partial_x F^x|_{D(s-\sigma)\times\mathcal{O}} \lessdot \frac{r^2\varepsilon}{\alpha\sigma^{\tau+n}}, \quad |\partial_x F^x|^{\mathcal{L}}_{D(s-\sigma)\times\mathcal{O}} \lessdot \frac{r^2\varepsilon^{\mathcal{L}}}{\alpha\sigma^{2(\tau+n)}}. \tag{4.17}$$

Proof. Expanding $\partial_x\mathcal{R}^x$ into Fourier series

$$\partial_x\mathcal{R}^x = \sum_{0\ne k\in\mathbb{Z}^n, |k|\le K} \widehat{\partial_x\mathcal{R}^x}(k) e^{\sqrt{-1}(k,x)}.$$

Since $\partial_x \mathcal{R}^x$ is analytic in $x \in D(s)$, we get that the Fourier coefficients $\widehat{\partial_x \mathcal{R}^x}(k)$'s decay exponentially in k, that is,

$$|\widehat{\partial_x \mathcal{R}^x}(k)| \lessdot |\partial_x \mathcal{R}^x|_{D(s)\times\mathcal{O}} e^{-s|k|} \lessdot e^{-s|k|} r^2 \varepsilon, \tag{4.18}$$

where we have used (4.4), since Lemma 4.1 holds true not only for R but also for $\mathcal{R}$. Expanding $\partial_x F^x$ into Fourier series as that of $\partial_x \mathcal{R}^x$ and putting them into (4.12), we get

$$\partial_x F^x(x,\xi) = \sum_{0\neq k\in\mathbb{Z}^n, |k|\leq K} \frac{\widehat{\partial_x \mathcal{R}^x}(k)}{\sqrt{-1}(k,\omega)} e^{\sqrt{-}(k,x)}.$$

By (4.16), (4.18) as well as Lemma A.1, we get that for $x \in D(s-\sigma)$,

$$|\partial_x F^x(x,\xi)| \lessdot \frac{r^2\varepsilon}{\alpha} \sum_{k\in\mathbb{Z}^n} |k|^\tau e^{-|k|\sigma} \lessdot \frac{r^2\varepsilon}{\alpha}\sigma^{-\tau-n}.$$

Applying ∂_{ξ_j} to both sides of (4.12) and using a method similar to the above, we can get

$$|\partial_x F^x(x,\xi)|^{\mathcal{L}} \lessdot \frac{r^2\varepsilon^{\mathcal{L}}}{\alpha\sigma^{2(\tau+n)}}. \quad \square$$

Proposition 2. *(Solution of (4.13).) Assume (4.16) holds true. Then on $D(s-\sigma)\times\mathcal{O}$ with $0<\sigma<s$, the equation (4.13) has a solution $F^y(x,\xi)$ which is analytic in $x \in D(s-\sigma)$ for ξ fixed and analytic in each $\xi_j, (j=1,...,n)$ for other variables fixed, and which is real for real argument, such that*

$$|F^y|_{D(s-\sigma)\times\mathcal{O}} \lessdot \frac{\varepsilon}{\alpha\sigma^{\tau+n}}, \quad |F^y|^{\mathcal{L}}_{D(s-\sigma)\times\mathcal{O}} \lessdot \frac{\varepsilon^{\mathcal{L}}}{\alpha\sigma^{2(\tau+n)}}$$

Proof. The proof is the same as that of Prop. 1. We omit it. $\square$

We are now in position to find the solution of (4.14). Recall that $A = \mathrm{diag}(A_{\jmath\jmath} : \jmath \in \mathcal{N})$ with $A_{\jmath\jmath} = \tilde{\Omega}_\jmath E_\jmath + B_{\jmath\jmath}$. We assume that $B_{\jmath\jmath}$ is symmetric and $\min_\jmath \tilde{\Omega}_\jmath \geq c > 0$ and $||B_{\jmath\jmath}||_2 \lessdot \epsilon_0^{1/3}$. Let

$$\tilde{A}_{\jmath\jmath} = E_\jmath + \tilde{\Omega}_\jmath^{-1} B_{\jmath\jmath} := E_\jmath + \tilde{B}_{\jmath\jmath}.$$

Then $\tilde{A}_{\jmath\jmath}$ and $\tilde{B}_{\jmath\jmath}$ are symmetric and $A_{\jmath\jmath} = \tilde{\Omega}_\jmath \tilde{A}_{\jmath\jmath}$. In addition, we will assume that $||\tilde{B}_{\jmath\jmath}||_2 \lessdot \epsilon_0 \jmath^{-d}$. It is easy to see that $\tilde{A}_{\jmath\jmath}$ is positively definite and $||\tilde{A}_{\jmath\jmath}||_2 \leq 1 + C\epsilon_0$. Write $\mathcal{R}^u = (\mathcal{R}_\jmath)_{\jmath\in\mathcal{N}}$ with $\mathcal{R}_\jmath = (\mathcal{R}^u_j)_{|j|=\jmath}$; Similarly, write $F^u = (F_\jmath)_{\jmath\in\mathcal{N}}$. Then (4.14) can be written as a system of equations:

$$\partial F_\jmath/\partial\omega - A_{\jmath\jmath} J_\jmath F_\jmath = \mathcal{R}_\jmath, \quad \jmath \in \mathcal{N}, \jmath \leq M. \tag{4.19}$$

Multiplying both sides of the above equation by a scalar $e^{\varsigma\jmath/2}\jmath^p$ and letting

$$e^{\varsigma\jmath/2}\jmath^p F_\jmath = \tilde{F}_\jmath,\ e^{\varsigma\jmath/2}\jmath^p \mathcal{R}_\jmath = \tilde{\mathcal{R}}_\jmath. \tag{4.20}$$

Then

$$\partial\tilde{F}_\jmath/\partial\omega - \tilde{\Omega}_\jmath \tilde{A}_{\jmath\jmath} J_\jmath \tilde{F}_\jmath = \tilde{\mathcal{R}}_\jmath, \quad \jmath \leq M. \tag{4.21}$$

Let $\tilde{\mathcal{R}} = (\tilde{\mathcal{R}}_\jmath)_{\jmath\in\mathbb{N}}$ and $\tilde{F} = (\tilde{F}_\jmath)_{\jmath\in\mathcal{N}}$.

Letting $\bar{F}_\jmath = \tilde{A}_{\jmath\jmath}^{-1/2}\tilde{F}_\jmath$ and $\bar{\mathcal{R}}_\jmath = \tilde{A}_{\jmath\jmath}^{-1/2}\tilde{\mathcal{R}}_\jmath$ and $\bar{A}_{\jmath\jmath} = \sqrt{-1}\tilde{A}_{\jmath\jmath}^{1/2}J_\jmath\tilde{A}_\jmath^{1/2}$, then, by noting $\tilde{\Omega}_\jmath$ is a scalar, we get

$$\partial\bar{F}_\jmath/\partial\omega + \sqrt{-1}\tilde{\Omega}_\jmath\bar{A}_{\jmath\jmath}\bar{F}_\jmath = \bar{\mathcal{R}}_\jmath, \tag{4.22}$$

Since $\tilde{A}_{\jmath\jmath}$ is (real) symmetric and $J_\jmath$ is skew-symmetric, we get $\bar{A}_{\jmath\jmath}$ is hermitian. Since $B_{\jmath\jmath}(\xi)$ is analytic in $\xi_j (j = 1, ..., n)$, it is easy to see that $\bar{A}_{\jmath\jmath} = \bar{A}_{\jmath\jmath}(\xi)$ is analytic in ξ_j. Therefore we have the following lemma which will be used in estimating the measure of some non-resonance sets.

Lemma 4.8. *Assume $B_{\jmath\jmath}(\xi)$ is real symmetric and*

$$\sup_{\mathcal{O}} ||B_{\jmath\jmath}(\xi)||_2 \lessdot \epsilon_0\jmath^{-d},\ \sup_{\mathcal{O}} ||\partial_\xi B_{\jmath\jmath}(\xi)||_2 \lessdot \jmath^{-d}\epsilon_0^{1/3},\ \sup_{\mathcal{O}} |\partial_\xi\tilde{\Omega}_\jmath| \le C\jmath^{-d},\ C \ll 1.$$

Let $\Lambda_\jmath = \{\lambda_j : |j| = \jmath\}$ be the collection of all eigenvalues of $\tilde{\Omega}_\jmath\bar{A}_{\jmath\jmath}$ and $\Lambda = \cup_{\jmath\in\mathcal{N}}\Lambda_\jmath$. Then for any $\lambda_l \in \Lambda$, it is a function of $\xi \in \mathcal{O}$ and is analytic in each entry[4] *ξ_l's $(l = 1, ..., n)$of $\xi \in \mathcal{O}$. Moreover,*

$$|\lambda_j(\xi) \pm \tilde{\Omega}_\jmath| \lessdot \jmath^{-d}\epsilon_0, \quad |j| = \jmath$$

and

$$\sup_{\mathcal{O}} |\partial_\xi\lambda_j(\xi)| \le C\jmath^{-d}, \quad |j| = \jmath, 0 < C \ll 1.$$

Proof. By the assumption, we can write

$$\tilde{\Omega}_\jmath\bar{A}_{\jmath\jmath} = \sqrt{-1}\tilde{\Omega}_\jmath J_\jmath + \bar{B}_{\jmath\jmath},$$

with

$$\sup_{\mathcal{O}} ||\bar{B}_{\jmath\jmath}||_2 \le \epsilon_0\jmath^{-d},\ \ \sup_{\mathcal{O}} ||\partial_\xi\bar{B}_{\jmath\jmath}||_2 \lessdot \epsilon_0^{1/3}\jmath^{-d}.$$

Note that the eigenvalues of $\sqrt{-1}\tilde{\Omega}_\jmath J_\jmath$ are $\pm\tilde{\Omega}_\jmath$'s. The proof is finished by the combination of Lemmas A.3,4,5 in Section 9. $\square$

Proposition 3. *Suppose that for any $k \in \mathbb{Z}^n$ and $\jmath \ge 0$, $\lambda_j \in \Lambda_\jmath$ and $\xi \in \mathcal{O}$, the following inequality holds true:*

$$|(k, \omega(\xi)) \pm \lambda_j| \gtrdot \alpha/(\jmath^d|k|^\tau),\ |k| \le K, \jmath \le M,\ (\ 1^{st}\ \textit{Melnikov's}\ ) \tag{4.23}$$

where we take $|0|$ as 1 for convenience. Then equation (4.14) has a solution $F^u(x, \xi)$ which is analytic in $x \in D(s - \sigma)$ for ξ fixed and analytic in each $\xi_j, (j = 1, ..., n)$ for other variables fixed, and which is real for real argument, such that

$$||F^u||_{\varsigma/2, D(s-\sigma)\times\mathcal{O}} \le \varsigma^{-d}\alpha^{-1}\sigma^{-\tau-n}r\varepsilon, \tag{4.24}$$

$$||F^u||^{\mathcal{L}}_{\varsigma/2, D(s-\sigma)\times\mathcal{O}} \le \varsigma^{-d}\alpha^{-1}\sigma^{-2(\tau+n)}r\varepsilon^{\mathcal{L}}. \tag{4.25}$$

[4]The function λ_l is not necessarily analytic in the whole of ξ. See [Ka] for a example.

Proof. By (4.23), we get

$$||((k,\omega)E_\jmath+\tilde{\Omega}_\jmath\bar{A}_\jmath)^{-1}||_2 \lessdot \jmath^d|k|^\tau/\alpha,\ \jmath\le M, |k|\le K.$$

Expand $\bar{\mathcal{R}}_\jmath$ and $\bar{F}_\jmath$ into Fourier series and putting then into (4.22), we get

$$\bar{F}_\jmath=\sum_{|k|\le K}((k,\omega)E_\jmath+\tilde{\Omega}_\jmath\bar{A}_\jmath)^{-1}\widehat{\bar{\mathcal{R}}_\jmath}(k)e^{\sqrt{-1}(k,x)},\ \jmath\le M. \tag{4.26}$$

Since $\bar{\mathcal{R}}_\jmath$ is analytic in $x\in D(s)$, we have $||\widehat{\bar{\mathcal{R}}_\jmath}(k)||_2 \lessdot e^{-s|k|}\sup_{D(s)\times\mathcal{O}}||\bar{\mathcal{R}}_\jmath||_2$. Therefore, for $x\in D(s-\sigma)$,

$$||\bar{F}_\jmath(x)||_2 \lessdot \alpha^{-1}\jmath^d\sup_{D(s)\times\mathcal{O}}||\bar{\mathcal{R}}_\jmath||_2\sum_{k\in\mathbb{Z}^n}|k|^\tau e^{-\sigma|k|}\lessdot \jmath^d\alpha^{-1}\sigma^{-\tau-n}\sup_{D(s)\times\mathcal{O}}||\bar{\mathcal{R}}_\jmath||_2.$$

Notice that $||\tilde{A}_{\jmath\jmath}^{-1/2}||_2=1+o(1)$. It follows that

$$||\tilde{F}_\jmath(x)||_2\le||\bar{F}_\jmath(x)||_2,\ ||\bar{\mathcal{R}}_\jmath(x)||_2\le||\tilde{\mathcal{R}}_\jmath(x)||_2.$$

We get

$$||\tilde{F}_\jmath(x)||_2\lessdot \jmath^d\alpha^{-1}\sigma^{-\tau-n}\sup_{D(s)\times\mathcal{O}}||\tilde{\mathcal{R}}_\jmath||_2.$$

Recalling that

$$e^{\varsigma\jmath/2}\jmath^pF_\jmath=\tilde{F}_\jmath,\ e^{\varsigma\jmath/2}\jmath^p\mathcal{R}_\jmath=\tilde{\mathcal{R}}_\jmath.$$

Hence,

$$||F_\jmath(x)||_2\lessdot \jmath^d\alpha^{-1}\sigma^{-\tau-n}\sup_{D(s)\times\mathcal{O}}||\mathcal{R}_\jmath||_2.$$

Finally, noting a simple fact

$$\sup_{0\le t}t^\beta e^{-\alpha t}=(\beta/\alpha)^\beta e^{-\beta},\ \text{ for any } \beta,\alpha>0,$$

we have

$$\begin{aligned}||F^u(x)||_{\varsigma/2,p}=&\sqrt{\sum_{\jmath\le M}e^{\varsigma\jmath}\jmath^{2p}||F_\jmath(x)||_2^2}\\ \lessdot&\alpha^{-1}\sigma^{-\tau-n}\sup_{D(s)\times\mathcal{O}}\sqrt{\sum_{\jmath\le M}(\jmath^{2d}e^{-\varsigma\jmath})(e^{2\varsigma\jmath}\jmath^{2p}||\mathcal{R}_\jmath||_2^2)}\\ \lessdot&\varsigma^{-d}\alpha^{-1}\sigma^{-\tau-n}\sup_{D(s)\times\mathcal{O}}\sqrt{\sum_{\jmath\le M}e^{2\varsigma\jmath}\jmath^{2p}||\mathcal{R}_\jmath||_2^2}\\ =&\varsigma^{-d}\alpha^{-1}\sigma^{-\tau-n}\sup_{D(s)\times\mathcal{O}}||\mathcal{R}^u||_{\varsigma,p}\\ \le&\varsigma^{-d}\alpha^{-1}\sigma^{-\tau-n}r\varepsilon,\end{aligned}$$

where (4.6) is used in the last inequality, since Lemma 4.1 holds true not only for R but also for $\mathcal{R}$. Differentiating (4.22) with respect to ξ and repeating the procedure above, we can prove (4.25). $\square$

Finally we turn to the solutions of the homological equation (4.15). Recall that, we have let

$$\tilde{A}_{\jmath\jmath} = E_\jmath + \tilde{\Omega}_\jmath^{-1} B_{\jmath\jmath} := E_\jmath + \tilde{B}_{\jmath\jmath}.$$

Thus $A_{\jmath\jmath} = \tilde{\Omega}_\jmath \tilde{A}_{\jmath\jmath}$. Note that $\tilde{\Omega}_\jmath$ is a scalar. Therefore, the equation (4.15) can be written as

$$\partial F_{\imath\jmath}/\partial\omega + \tilde{\Omega}_\jmath F_{\imath\jmath} J_\jmath \tilde{A}_{\jmath\jmath} - \tilde{\Omega}_\imath \tilde{A}_{\imath\imath} J_\imath F_{\imath\jmath} = \mathcal{R}_{\imath\jmath}(x,\xi), \tag{4.27}$$

where we omit the superscript ^{uu} of F^{uu} and $\mathcal{R}^{uu}$. Let

$$\bar{A}_{\jmath\jmath} = \sqrt{-1}\tilde{A}_{\jmath\jmath}^{1/2} J_\jmath \tilde{A}_{\jmath\jmath}^{1/2}. \tag{4.28}$$

Notice that $\tilde{B}_{\jmath\jmath}$ is real symmetric and $||\tilde{B}_{\jmath\jmath}||_2 < \epsilon_0$. It is easy to see that both $\tilde{A}_{\jmath\jmath}$ and $\bar{A}_{\jmath\jmath}$ are hermitian and positive and

$$||\tilde{A}_{\jmath\jmath}||_2, ||\tilde{A}_{\jmath\jmath}^{-1}||_2, ||\bar{A}_{\jmath\jmath}||_2, ||\bar{A}_{\jmath\jmath}^{-1}||_2 = 1 + o(1). \tag{4.29}$$

Lemma 4.9. *Let M and N be $m \times n$ and $n \times l$ matrices, respectively. Denote by $||\cdot||_\infty$ the maximum norm of matrix. Then*

$$||M||_\infty \le ||M||_2, \ ||MN||_\infty \le n||M||_\infty ||N||_\infty.$$

Proof. The proof is rather simple.

Let

$$\bar{F}_{\imath\jmath} = \tilde{A}_{\imath\imath}^{-1/2} F_{\imath\jmath} \tilde{A}_{\jmath\jmath}^{-1/2}, \quad \bar{\mathcal{R}}_{\imath\jmath} = \tilde{A}_{\imath\imath}^{-1/2} \mathcal{R}_{\imath\jmath} \tilde{A}_{\jmath\jmath}^{-1/2}. \tag{4.30}$$

Then (4.27) is changed into

$$\partial \bar{F}_{\imath\jmath}/\partial\omega - \sqrt{-1}\tilde{\Omega}_\jmath \bar{F}_{\imath\jmath} \bar{A}_{\jmath\jmath} + \sqrt{-1}\tilde{\Omega}_\imath \bar{A}_{\imath\imath} \bar{F}_{\imath\jmath} = \bar{\mathcal{R}}_{\imath\jmath}(x,\xi), \ [\mathcal{R}_{\jmath\jmath}] = 0, \ \imath, \jmath \le M \tag{4.31}$$

Recall that we have denoted by $\Lambda_\jmath$ the collection of the eigenvalues of $\tilde{\Omega}_\jmath \bar{A}_{\jmath\jmath}$. Write $\Lambda_\jmath = \{\lambda_j : |j| = \jmath\}$. By abuse use of notation, we also by $\Lambda_\jmath$ the diagonal matrix $\mathrm{diag}(\lambda_\jmath : |j| = \jmath)$. Then there is a unitary matrix $Q_{\jmath\jmath}$ such that

$$\tilde{\Omega}_\jmath \bar{A}_{\jmath\jmath} = Q_{\jmath\jmath}^* \Lambda_\jmath Q_{\jmath\jmath}. \tag{4.32}$$

Let

$$\underline{F}_{\imath\jmath} = Q_{\imath\imath} \bar{F}_{\imath\jmath} Q_{\jmath\jmath}^*, \quad \underline{\mathcal{R}}_{\imath\jmath} = Q_{\imath\imath} \bar{\mathcal{R}}_{\imath\jmath} Q_{\jmath\jmath}^*. \tag{4.33}$$

Then (4.31) is changed into

$$\partial \underline{F}_{\imath\jmath}/\partial\omega - \sqrt{-1}(\underline{F}_{\imath\jmath}\Lambda_\jmath - \Lambda_\imath \underline{F}_{\imath\jmath}) = \underline{\mathcal{R}}_{\imath\jmath}(x,\xi), \ [\underline{\mathcal{R}}_{\jmath\jmath}] = 0, \imath, \jmath \le M \tag{4.34}$$

Recall that both $\underline{F}_{\imath\jmath}$ and $\underline{\mathcal{R}}_{\imath\jmath}$ are $\imath^\sharp \times \jmath^\sharp$-matrices. Denote by $\underline{F}_{\imath\jmath}^{ij}$ and $\underline{\mathcal{R}}_{\imath\jmath}^{ij}$ the elements of the matrix $\underline{F}_{\imath\jmath}$ and $\underline{\mathcal{R}}_{\imath\jmath}$, respectively. Expanding $\underline{F}_{\imath\jmath}^{ij}$ and $\underline{\mathcal{R}}_{\imath\jmath}^{ij}$ into Fourier series and putting them into (4.31) we get

$$\widehat{\underline{F}_{\imath\jmath}^{ij}}(k) = -\sqrt{-1}\frac{\widehat{\underline{\mathcal{R}}_{\imath\jmath}^{ij}}(k)}{(k,\omega) \pm \lambda_i \pm \lambda_j}, \quad \begin{cases} |i| = |j| = \jmath \le M, 0 \ne |k| \le K \\ |i| \ne |j|, \imath, \jmath \le M, |k| \le K. \end{cases} \tag{4.35}$$

In order to that (4.35) is solvable, we need the *2^{nd} Melnikov's conditions:*[5]
Assume that for any $\xi \in \mathcal{O}$, $\lambda_i, \lambda_j \in \Lambda_\jmath$, we have

$$|(k,\omega) \pm \lambda_i \pm \lambda_j| \gg \alpha/(\imath^d \jmath^d |k|^\tau), \quad \begin{cases} |i| = |j| = \jmath \leq M, 0 \neq |k| \leq K \\ |i| \neq |j|, \imath, \jmath \leq M, |k| \leq K. \end{cases} \tag{4.36}$$

We assume $|i| \geq |j|$ without loss of generality. By Lemma 4.2 and Cauchy's theorem, we get

$$|\widehat{\mathcal{R}_{\imath\jmath}^{ij}}(k)| \lessdot e^{-|k|s} e^{-\varsigma\imath} \imath^{-p} \jmath^p \varepsilon, \ |i| \geq |j|.$$

Note that (4.30,33) and the fact $Q_{\imath\imath}$ is unitary and of order $\imath^\sharp \leq \imath^{d-2+\underline{\varepsilon}}$ with some constant $0 < \underline{\varepsilon} \ll 1$. Using Lemma 4.9, we have

$$|\widehat{\underline{\mathcal{R}}_{\imath\jmath}^{ij}}(k)| \leq |\widehat{\mathcal{R}_{\imath\jmath}^{ij}}(k)| (\imath^\sharp \jmath^\sharp)^2 \leq e^{-|k|s} e^{-\varsigma\imath} \imath^{4(d-1)} \imath^{-p} \jmath^p \varepsilon, \ |i| \geq |j| \tag{4.37}$$

Using (4.32,33,34) we get

$$|\widehat{\underline{F}_{\imath\jmath}^{ij}}(k)| \leq \frac{e^{-|k|s}|k|^\tau \varepsilon}{\alpha} e^{-\varsigma\imath} \imath^{6d-4} \imath^{-p} \jmath^p, \ |i| \geq |j|; |i|, |j| \leq M; |k| \leq K \tag{4.38}$$

Moreover, the function

$$\underline{F}_{\imath\jmath}^{ij}(x,\xi) = \sum_{k \in \mathbb{Z}^n, |k| \leq K} \widehat{\underline{F}_{\imath\jmath}^{ij}}(k) e^{\sqrt{-1}(k,x)}$$

is well-defined on a small domain $D(s-\sigma) \times \mathcal{O}$ and on this domain

$$|\underline{F}_{\imath\jmath}^{ij}(x,\xi)| \leq \frac{\varepsilon}{\alpha\sigma^{n+\tau}} e^{-\varsigma\imath} \imath^{6d-4} \imath^{-p} \jmath^p, \ |i| \geq |j|; |i|, |j| \leq M \tag{4.39}$$

where Lemma A.1 is used. Using Lemma 4.9 and (4.30,33), we get

$$|F_{\imath\jmath}^{ij}(x,\xi)| \leq \frac{\varepsilon}{\alpha\sigma^{n+\tau}} e^{-\varsigma\imath} \imath^{10d-8} \imath^{-p} \jmath^p, \ |i| \geq |j|; |i|, |j| \leq M \tag{4.40}$$

Note that $\widehat{\underline{F}_{\jmath\jmath}}(k)$, ω, λ_i and λ_j are analytically dependent on $\xi_j (j = 1, ..., n)$. Applying ∂_{ξ_j} to (4.34) and using the same method as the above, we get

$$|F_{\imath\jmath}^{ij}(x,\xi)|^{\mathcal{L}} \leq \frac{\varepsilon^{\mathcal{L}}}{\alpha\sigma^{2(n+\tau)}} e^{-\varsigma\imath} \imath^{10d-8} \imath^{-p} \jmath^p, \ |i| \geq |j|; |i|, |j| \leq M. \tag{4.41}$$

Using Lemma A.2 in Appendix A, we have the follow lemma.

Proposition 4. *Assume that the non-resonant conditions (4.5) and (5.12) hold true. Then there is an operator $F^{uu}(x,\xi)$ defined on $D(s-\sigma) \times \mathcal{O}$ solves (4.4*) and*

$$|||F^{uu}(x,\xi)|||_{0,\varsigma/2,p,p} \leq \frac{\varepsilon}{\varsigma^{12d-8}\sigma^{n+\tau}\alpha}, \tag{4.42}$$

and

$$|||F^{uu}(x,\xi)|||_{0,\varsigma/2,p,p}^{\mathcal{L}} \leq \frac{\varepsilon^{\mathcal{L}}}{\varsigma^{12d-8}\sigma^{2(n+\tau)}\alpha}. \tag{4.43}$$

[5]These conditions are weaker than the usual second Melnikov's ones as in [P1], but similar to ones in [B,B-B,B-G].

5. Symplectic change of variables.

In this section, our procedure is standard and almost the same as that of Section 3 in [P1,p.128-132]. Here we give out the outline of the procedure. See [P1] for the details.

Coordinate transformation. By Propositions 1-4, we get a Hamiltonian F on $D(s-\sigma,r)$ where

$$F = F^x + (F^y, y) + \langle F^u, u\rangle + \langle F^{uu}u, u\rangle$$

and give estimates of F^x, F^y, F^u and F^{uu}. Let X_F be the vector field corresponding to the Hamiltonian F, that is,

$$X_F = (-\partial_y F, \partial_x F, J_\infty \partial_u F),$$

here $\partial_u F$ is the usual ℓ^2-gradient. It follows from Prop.1,2,3 and 4 that for $(x, y, u; \xi) \in D(s-\sigma, r) \times \xi \in \mathcal{O}$,

$$\begin{aligned}
&{}_r|X_F|_{\varsigma/2,p} = |\partial_y F| + \frac{1}{r^2}|\partial_x F| + \frac{1}{r}||J_\infty \partial_u F||_{\varsigma/2,p}\\
=&|F^y| + \frac{1}{r^2}|\partial_x F^x| + \frac{1}{r}||J_\infty F^u||_{\varsigma/2,p} + \frac{1}{r}||J_\infty F^{uu}u||_{\varsigma/2,p}\\
\le&|F^y| + \frac{1}{r^2}|\partial_x F^x| + \frac{1}{r}||F^u||_{\varsigma/2,p} + \frac{1}{r}|||F^{uu}|||_{0,\varsigma/2,p,p}||u||_p\\
\lessdot&\frac{1}{\alpha\sigma^{n+\tau}\varsigma^{12d-8}}\cdot\varepsilon,
\end{aligned}$$

where we have used $0 < \varsigma < 1$ and $||u||_p < r$. That is,

$${}_r|X_F|_{\varsigma/2,p,D(s-\sigma,r)\times\mathcal{O}} \lessdot Q\varepsilon, \tag{5.1}$$

where

$$Q = \frac{1}{\alpha\sigma^{2(n+\tau)}\varsigma^{12d-8}}. \tag{5.2}$$

Similarly, we have

$${}_r|X_F|^{\mathcal{L}}_{\varsigma/2,p,D(s-\sigma,r)\times\mathcal{O}} \lessdot Q\varepsilon^{\mathcal{L}}. \tag{5.3}$$

As in [P1,p.129], we introduce the operator norm

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

Using (5.1), (5.3) and the generalized Cauchy's inequality (See Lemma A.3 of [P1,p.147]) and the observation that every point in $D(s-2\sigma, r/2)$ has at least $|\cdot|_{p,r}$-distance $\sigma/2$ to the boundary of $D(s-\sigma,r)$, we get

$$\sup_{D(s-2\sigma,r/2,\varsigma/2)\times\mathcal{O}} {}_r||DX_F||_{\varsigma/2,p} \lessdot \sigma^{-1}\ {}_r|X_F|_{\varsigma/2,p,D(s-\sigma,r)\times\mathcal{O}} \le \sigma^{-1}Q\varepsilon. \tag{5.4}$$

$$\sup_{D(s-2\sigma,r/2,\varsigma/2)\times\mathcal{O}} {}_r||DX_F||^{\mathcal{L}}_{\varsigma/2,p} \lessdot \sigma^{-1}\ {}_r|X_F|^{\mathcal{L}}_{\varsigma/2,p,D(s-\sigma,r)\times\mathcal{O}} \le \sigma^{-1}Q\varepsilon^{\mathcal{L}}, \tag{5.5}$$

where DX_F is the differential of X_F. Assume that $\sigma^{-1}Q\varepsilon$ and $\sigma^{-1}Q\varepsilon^{\mathcal{L}}$ are small enough. (These assumptions will be fulfilled in the following KAM iterations. Also see (5.12).) Arbitrarily fix $\xi \in \mathcal{O}$. By (5.1), the flow X_F^t of the vector field X_F exists on $D(s-3\sigma, r/4)$ for $t \in [-1,1]$ and takes the domain into $D(s-2\sigma, r/2)$, and by Lemma A.4 of [P1, p.147], we have

$$ {}_r|X_F^t - id|_{\varsigma/2,p,D(s-3\sigma,r/4)\times\mathcal{O}} \lessdot {}_r|X_F|_{\varsigma/2,p,D(s-\sigma,r)\times\mathcal{O}} \lessdot Q\varepsilon \tag{5.6}$$

and

$$\begin{aligned} &{}_r|X_F^t - id|^{\mathcal{L}}_{\varsigma/2,p,D(s-3\sigma,r/4)\times\mathcal{O}} \\ &\lessdot \exp({}_r\|DX_F\|_{\varsigma/2,p,D(s-2\sigma,r/2)\times\mathcal{O}}) \cdot {}_r|X_F|^{\mathcal{L}}_{\varsigma/2,p,D(s-\sigma,r)\times\mathcal{O}} \\ &\lessdot \exp(\sigma^{-1}Q\varepsilon)Q\varepsilon^{\mathcal{L}} \lessdot Q\varepsilon^{\mathcal{L}}, \end{aligned}\tag{5.7}$$

for $t \in [-1,1]$. Furthermore, by the generalized Cauchy's inequality,

$$ {}_r\|DX_F^t - I\|_{\varsigma/2,p,D(s-4\sigma,r/8)\times\mathcal{O}} \lessdot \sigma^{-1}Q\varepsilon, \tag{5.8}$$

and

$$ {}_r\|DX_F^t - I\|^{\mathcal{L}}_{\varsigma/2,p,D(s-4\sigma,r/8)\times\mathcal{O}} \lessdot \sigma^{-1}Q\varepsilon^{\mathcal{L}}, \tag{5.9}$$

The new error term. Subjecting $H = N + \grave{R}$ to the symplectic transformation $\Phi = X_F^t|_{t=1}$ we get the new Hamiltonian scale $H_+ := H\circ\Phi = H\circ X_F^1$ on $D(s-5\sigma, \eta r$ where $0<\eta<1/8$. By Taylor's formula

$$\begin{aligned} H_+ &= (N + R + (\grave{R} - R)) \circ X_F^1 \\ &= (N + \mathcal{R} + R_M + R_K + (\grave{R} - R)) \circ X_F^1 \\ &= N - \{F, N\} + \int_0^1 \{t\{F,N\}, F\} \circ X_F^t \, dt \\ &\quad + \mathcal{R} + \int_0^1 \{\mathcal{R}, F\} \circ X_F^t \, dt + (R_M + R_K + (\grave{R} - R)) \circ X_F^1. \end{aligned}$$

Recall that F solves the linearized equation

$$\{F, N\} = \mathcal{R} - [[\mathcal{R}]].$$

Thus,

$$H_+ = N_+ + \grave{R}_+$$

where

$$N_+ = N + [[\mathcal{R}]]$$

$$\grave{R}_+ = R_M \circ X_F^1 + R_K \circ X_F^1 + (\grave{R} - R) \circ X_F^1 + R_K \circ X_F^1 + \int_0^t \{\mathcal{R}(t), F\} \circ X_F^t \, dt$$

with

$$\mathcal{R}(t) = \mathcal{R} + t(\mathcal{R} - [[\mathcal{R}]]).$$

Hence, the new perturbing vector field is

$$X_{\grave{R}_+} = (X_F^1)^*(X_{\grave{R}} - X_R + R_M + R_K) + \int_0^t (X_F^t)^* [X_{\mathcal{R}(t)}, X_F] \, dt,$$

where $(X_F^t)^*$ is the pull-back of X_F^t and $[\cdot,\cdot]$ is the commutator of vector fields. We are now in position to estimate the new perturbing vector field $X_{\grave{R}_+}$. Let $Y : D(s-\sigma,r) \subset \mathcal{P} \to \mathcal{P}^{a,p}$ be a vector field on $D(s-\sigma,r)$, depending on the parameter $\xi \in \mathcal{O}$. Let $U = D(s-5\sigma,\eta r)\times\mathcal{O}$ and $V = D(s-4\sigma,2\eta r)\times\mathcal{O}$ and $W = D(s-2\sigma,4\eta r)\times\mathcal{O}$. By (5.9) and the "proof of estimate (12)" of [P1, p.131-132], we have that for any $a > 0$,

$$ {}_{\eta r}|(X_F^t)^* Y|_{a,p,U} \lessdot {}_{\eta r}|Y|_{a,p,V} \tag{5.10}$$

and

$$ {}_{\eta r}|(X_F^t)^* Y|^{\mathcal{L}}_{a,p,U} \lessdot {}_{\eta r}|Y|^{\mathcal{L}}_{a,p,V} + \frac{1}{\sigma\eta^2}\ {}_{\eta r}|Y|_{a,p,W}\cdot {}_{\eta r}|X_F|^{\mathcal{L}}_{a,p,V}. \tag{5.11}$$

We assume that

$$\varepsilon Q/\sigma\eta^2 \le \varepsilon Q/\sigma^2\eta^2 \lessdot 1. \tag{5.12}$$

These assumptions will be fulfilled in the KAM iterative lemma later. By (4.3) and (5.10,11),

$$ {}_{\eta r}|(X_F^1)^*(X_{\grave{R}} - X_R)|_{\varsigma/2,p,U} \lessdot {}_{\eta r}|X_{\grave{R}} - X_R|_{\varsigma/2,p,V} \lessdot \eta\varepsilon \tag{5.13}$$

and

$$ {}_{\eta r}|(X_F^1)^*(X_{\grave{R}} - X_R)|^{\mathcal{L}}_{a,p,U} \lessdot \eta\varepsilon^{\mathcal{L}} + \frac{\varepsilon\eta}{\sigma\eta^2} Q\varepsilon^{\mathcal{L}} \lessdot \eta\varepsilon^{\mathcal{L}}. \tag{5.14}$$

Let $D_l = D(s-l\sigma, r/l)\times\mathcal{O}$ $(l = 1,2,...)$. Recall that (4.2) holds still true after replacing R by $\mathcal{R}$. By (4.2) and (5.4,5) and using the generalized Cauchy estimate, following [P1,p.130-131] we get

$$\begin{aligned} {}_r|[X_{\mathcal{R}(t)}, X_F]|_{\varsigma/2,p,D_2} &\lessdot \sigma^{-1}\ {}_r|X_{\mathcal{R}}|_{\varsigma/2,p,D_1}\cdot {}_r|X_F|_{\varsigma/2,p,D_1} \\ &\lessdot \sigma^{-1}\ {}_r|X_{\grave{R}}|_{\varsigma/2,p,D_1}\cdot {}_r|X_F|_{\varsigma/2,p,D_1} \\ &\lessdot \sigma^{-1} Q\varepsilon^2 < \eta\varepsilon \end{aligned} \tag{5.15}$$

and

$$\begin{aligned} &{}_r|[X_{\mathcal{R}(t)}, X_F]|^{\mathcal{L}}_{\varsigma/2,p,D_2} \\ &\lessdot \sigma^{-1}\ {}_r|X_{\grave{R}}|^{\mathcal{L}}_{\varsigma/2,p,D_1}\ {}_r|X_F|_{\varsigma/2,p,D_1} + \sigma^{-1}\ {}_r|X_{\grave{R}}|_{\varsigma/2,p,D_1}\ {}_r|X_F|^{\mathcal{L}}_{\varsigma/2,p,D_1} \\ &\lessdot \sigma^{-1}\varepsilon^{\mathcal{L}} Q\varepsilon + \sigma^{-1}\varepsilon Q\varepsilon^{\mathcal{L}} < \eta\varepsilon^{\mathcal{L}} \end{aligned} \tag{5.16}$$

Finally, we have

$$ {}_{\eta r}|Y|_{\varsigma/2,p,D_l} \lessdot \eta^{-2}\ {}_r|Y|_{\varsigma/2,p,D_l}, \quad {}_{\eta r}|Y|^{\mathcal{L}}_{\varsigma/2,p,D_l} \lessdot \eta^{-2}\ {}_r|Y|^{\mathcal{L}}_{\varsigma/2,p,D_l}, \tag{5.17}$$

for any vector field Y on D_l's $(l = 1,2,...)$. Collecting all terms above and Lemma 4.6 and 4.7, we then arrive ate the estimates

$$ {}_{\eta r}|X_{\grave{R}_+}|_{\varsigma/2,p,D(s-5\sigma,\eta r)\times\mathcal{O}} \lessdot \eta\varepsilon, \quad {}_{\eta r}|X_{\grave{R}_+}|^{\mathcal{L}}_{\varsigma/2,p,D(s-5\sigma,\eta r)\times\mathcal{O}} \lessdot \eta\varepsilon^{\mathcal{L}}. \tag{5.18}$$

The new normal form. This is $N_+ = N + [[\mathcal{R}]]$. Recall

$$N = (\omega(\xi), y) + \langle \Lambda u, u\rangle + \langle Bu, u\rangle$$

and

$$[[\mathcal{R}]] = ([\mathcal{R}^y], y) + \langle [\mathcal{R}_g^{uu}]u, u\rangle.$$

Let

$$\omega_+ = \omega + [\mathcal{R}^y] \tag{5.19}$$

and

$$B_+ = B + [\mathcal{R}_g^{uu}]. \tag{5.20}$$

Then

$$N_+ = (\omega_+, y) + \langle \Lambda u, u\rangle + \langle B_+ u, u\rangle. \tag{5.21}$$

6. Iterative lemma.

6.1. Iterative constants. As usual, the KAM theorem is proven by the Newton-type iteration procedure which involves an infinite sequence of coordinate changes. In order to make our iteration procedure run, we need the following iterative constants:

1. $\epsilon_0 = \epsilon$, $\epsilon_m = \epsilon_0^{\wedge}(4/3)^m$, $m = 1, 2, ...$;
2. $\alpha_0 = \alpha$, $\alpha_m = \alpha/m^2$, $m = 1, 2, ...$;
3. $\eta_m = \epsilon_m^{1/3}$, $m = 0, 1, 2, ..$;
4. $e_0 = 0$, $e_m = (1^{-2} + \cdots + m^{-2})/2\sum_{j=1}^{\infty} j^{-2}$, (thus, $e_m < 1/2$ for all $m \in \mathbb{N}$);
5. $s_0 = s$, $s_m = s_0(1 - e_m)$, $m = 1, 2, ...$, (thus, $s_m > s_0/2$ for all $m \in \mathbb{N}$);
6. $\sigma_m = (s_m - s_{m+1})/10$, $m = 1, 2, ...$, (thus, $s_m - l\sigma_m > s_{m+1}$ for $1 \le l \le 6$ and $\sigma_m^{-1} = O(m^2)$);
7. $\varsigma_m = \epsilon_m^{4/(2\kappa - d)} = \epsilon_m^{1/144d}$, (Recall $\kappa = 577d/2$);
8. $r_0 = r$, $r_m = \eta_m r_0$, $m = 1, 2, ...$;
9. $M_m = 2|\ln(\sigma_m^n \eta_m)|/\varsigma_m$;
10. $K_m = |\ln \eta_m|/\sigma_m$;
11. $\nu_m = \alpha_m/(2M_m^{2d}K_m^{\tau+1})$;
12. $\Pi_0 = [1, 2]^d$, and Π_m ($m = 1, 2, ...$) are defined in Section 8. $\mathcal{O}_m$'s are the ν_m-neighborhood of Π_m in $\mathbb{R}^n$.

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

$$H_l = (\omega_l(\xi), y) + \sum_{\jmath\in\mathcal{N}} (\Omega_\jmath^0 u_\jmath, u_\jmath) + \frac{1}{2}\sum_{\jmath\in\mathcal{N}} (B_{\jmath\jmath}^l(\xi)u_\jmath, u_\jmath) + \sum_{\wp \ge l}^{\infty} \grave{R}^{l\wp}(x, y, u; \xi), \tag{6.1}$$

where the following conditions are imposed:

(l.1). the parameter sets $\Pi_0 \supset \cdots \supset \cdots \Pi_l \supset \cdots \Pi_m$ with

$$\textit{Meas}\ \Pi_l \ge (\ \textit{Meas}\ \Pi_0)(1 - \alpha_l/(1 + l)^2); \tag{6.2}$$

The map $\xi \mapsto \omega_l(\xi)$ is analytic in each entry of $\xi \in \mathcal{O}_l$, ($\mathcal{O}_l$ is the ν_l-neighborhood of Π_l in $\mathbb{R}^n$.) and

$$\inf_{\mathcal{O}_l}\left|det\ \frac{\partial\omega_l}{\partial\xi}\right| \ge (1 - e_l)c_1,\ \sup_{\mathcal{O}_l}|\partial_\xi^j \omega_l| \le e_l c_2,\ j = 0, 1. \tag{6.3}$$

(l.2). $B_{\jmath\jmath}^l$ is real symmetric matrix of order $\jmath^\sharp$ and analytic in each entry ξ_k ($k = 1, ..., n$) of $\xi \in \mathcal{O}_l$, and

$$\sup_{\mathcal{O}_l}||B_{\jmath\jmath}^l||_2 \le \jmath^{-d}e_l\epsilon_0,\ \sup_{\mathcal{O}_l}||B_{\jmath\jmath}^l||_2^{\mathcal{L}} \le \jmath^{-d}e_l\epsilon_0^{1/3},\ \ \textit{for any}\ \jmath \in \mathcal{N}. \tag{6.4}$$

In addition, $B^0_{\jmath\jmath} \equiv 0$.

$(l.3)$. *For* $\wp \geq l$ *and* $0 \leq l \leq m$, *the perturbation* $\grave{R}^{l\wp}(x, y, u; \xi)$ *is analytic in the space coordinate domain* $D(s_\wp, r_\wp)$ *and also analytic in each entry* ξ_k $(k = 1, ..., n)$ *of the parameter vector* $\xi \in \mathcal{O}_l$, *and is real for real argument; moreover, its Hamiltonian vector field* $X_{\grave{R}^{l\wp}} := (\grave{R}^{l\wp}_y, -\grave{R}^{l\wp}_x, J_\infty \grave{R}^{l\wp}_u)^T$ *defines on* $D(s_\wp, r_\wp)$ *a analytic map*

$$X_{\grave{R}^{l\wp}} : D(s_l, r_l) \subset \mathcal{P} \to \mathcal{P}^{\varsigma_\wp, p}. \tag{6.5}$$

In addition, the vector field $X_{\grave{R}^{l\wp}}$ *is analytic in the domain* $D(s_\wp, r_\wp)$ *with small norms*

$$_{r_\wp}|X_{\grave{R}^{l\wp}}|_{\varsigma_\wp, p, D(s_l, r_l)\times\mathcal{O}_l} \lessdot \epsilon_\wp, \quad _{r_\wp}|X_{\grave{R}^{l}}|^{\mathcal{L}}_{\varsigma_\wp, p, D(s_l, r_l)\times\mathcal{O}_l} \lessdot \epsilon_\wp^{1/3}. \tag{6.6}$$

Then there is is an absolute positive constant ϵ^* *enough small such that, if* $0 < \epsilon_0 < \epsilon^*$, *there is a set* $\Pi_{m+1} \subset \Pi_m$, *and a change of variables* $\Phi_{m+1} : \mathcal{D}_{m+1} := D(s_{m+1}, r_{m+1}) \times \mathcal{O}_{m+1} \to D(s_m, r_m)$ *being real*[6], *analytic in* $(x, y, u) \in D(s_{m+1}, r_{m+1})$ *and each entry* $\xi \in \mathcal{O}_{m+1}$, *as well as following estimates holds true:*

$$_{r_m}|\Phi_{m+1} - id|_{\varsigma_m/2, p, \mathcal{D}_{m+1}} \lessdot \epsilon_m^{1/2} \tag{6.7}$$

and

$$_{r_m}|\Phi_{m+1} - id|^{\mathcal{L}}_{\varsigma_m/2, p, \mathcal{D}_{m+1}} \lessdot \epsilon_m^{1/4}. \tag{6.8}$$

Furthermore, the new Hamiltonian $H_{m+1} := H_m \circ \Phi_{m+1}$ *of the form*

$$H_{m+1} = (\omega_{m+1}, y) + \sum_{\jmath\in\mathcal{N}} (\Omega^0_\jmath u_\jmath, u_\jmath) + \frac{1}{2}\sum_{\jmath\in\mathcal{N}} (B^{m+1}_{\jmath\jmath} u_\jmath, u_\jmath) + \sum_{\wp\geq m+1}^{\infty} \grave{R}^{m+1\wp} \tag{6.9}$$

satisfies all the above conditions $(l.1, 2, 3)$ *with* l *being replaced by* $m + 1$.

6.3. Proof of The Iterative Lemma.

As stated as in the iterative lemma, we have got a family of Hamiltonian functions H_l's $(l = 0, 1, ..., m)$ which satisfy the conditions $(l.1, 2, 3)$. We now consider the Hamiltonian H_m.That is,

$$H_m = (\omega_m, y) + \sum_{\jmath\in\mathcal{N}} (\Omega^0_\jmath u_\jmath, u_\jmath) + \frac{1}{2}\sum_{\jmath\in\mathcal{N}} (B^m_{\jmath\jmath} u_\jmath, u_\jmath) + \sum_{\wp\geq m}^{\infty} \grave{R}^{m\wp} \tag{6.10}$$

which satisfy the conditions $(m.1, 2, 3)$. First, let us consider

$$\breve{H}_m := (\omega_m, y) + \sum_{\jmath\in\mathcal{N}} (\Omega^0_\jmath u_\jmath, u_\jmath) + \frac{1}{2}\sum_{\jmath\in\mathcal{N}} (B^m_{\jmath\jmath} u_\jmath, u_\jmath) + \grave{R}^{mm} \tag{6.11}$$

instead of H_m. Let $s = s_m$, $\eta = \eta_m$, $r = r_m = \eta_m r_0$, $\varepsilon = \epsilon_m$, $\varepsilon^{\mathcal{L}} = \epsilon_m^{1/3}$, $\sigma = \sigma_m$, $\varsigma = \varsigma_m$, $\omega = \omega_m$, $\Lambda =$ diag $(\Omega^0_\jmath : \jmath \in \mathbb{N})$, $B =$ diag $(B^m_{\jmath\jmath} : \jmath \in \mathbb{N})$, and $\grave{R} = \grave{R}^{mm} := R_m$. Clearly, $\varepsilon < \varepsilon^{\mathcal{L}}$. Let

$$\Pi = \Pi_{m+1} := \{\xi \in \Pi_m : \text{non-resonant conditions (4.16,23,36) hold. }\} \tag{6.12}$$

---

[6]The word "real" means $\overline{\Phi_{m+1}(z, \xi)} = \Phi_{m+1}(\bar{z}, \xi)$ for any $(z, \xi) \in \mathcal{D}_{m+1}$.

Set

$$Q = Q_m = \frac{1}{\alpha \sigma_m^{2(n+\tau)} \varsigma_m^{12d-8}}.$$

Then

$$\epsilon_m Q_m / (\sigma_m^2 \eta_m^2) \lessdot m^{2(n+\tau+1)} \epsilon_m^{\frac{1}{3} - \frac{12d-8}{144d}} / \alpha \lessdot \epsilon_m^{1/4} \le 1,$$

if $\alpha \ll \epsilon_0$. This implies that (5.12) holds true.

By means of the arguments in Section 5, we got that there is a Hamiltonian F_m defined on[7] $D(s_m - 4\sigma_m, r_m/8) \times \mathcal{O}_{m+1}$ and a symplectic change of variables $\Phi_{m+1} = X_{F_m}^t|_{t=1}$ with[8]

$$_{r_m}|\Phi_{m+1} - id|_{\varsigma_m/2, p, \mathcal{D}_{m+1}} \lessdot Q_m \epsilon_m < \epsilon_m^{1/2} \tag{6.13}$$

and

$$_{r_m}|\Phi_{m+1} - I|^{\mathcal{L}}_{\varsigma_m/2, p, \mathcal{D}_{m+1}} \lessdot \sigma_m^{-1} Q_m \epsilon_m^{1/3} < \epsilon_m^{1/4} \tag{6.14}$$

$$_{r_m}\|D\Phi_{m+1} - id\|_{\varsigma_m/2, p, \mathcal{D}_{m+1}} \lessdot Q_m \epsilon_m < \epsilon_m^{1/2} \tag{6.15}$$

$$_{r_m}\|D\Phi_{m+1} - id\|^{\mathcal{L}}_{\varsigma_m/2, p, \mathcal{D}_{m+1}} \lessdot \sigma_m^{-1} Q_m \epsilon_m^{1/3} < \epsilon_m^{1/4} \tag{6.16}$$

[9]such that

$$H_+ := \breve{H}_m \circ \Phi_{m+1} = N_{m+1} + \grave{R}_{m+1} \tag{6.17}$$

where[10]

$$N_{m+1} := N_+ = (\omega_{m+1}, y) + \langle \Lambda u, u \rangle + \frac{1}{2} \langle B^{m+1} u, u \rangle \tag{6.18}$$

$$\omega_{m+1} = \omega_m + [\mathcal{R}_m^y] \tag{6.19}$$

$$B^{m+1} = B^m + [(\mathcal{R}_m^{uu})_g], \quad \text{or} \quad B_{\jmath\jmath}^{m+1} = B_{\jmath\jmath}^m + [(\mathcal{R}_m)_{\jmath\jmath}] \tag{6.20}$$

and[11]

$$_{r_{m+1}}|X_{\grave{R}_{m+1}}|_{\varsigma_{m+1}, p, \mathcal{D}_{m+1}} \lessdot \eta_m \epsilon_m = \epsilon_{m+1}, \tag{6.21}$$

$$_{r_{m+1}}|X_{\grave{R}_{m+1}}|^{\mathcal{L}}_{\varsigma_{m+1}, p, \mathcal{D}_{m+1}} \lessdot \eta_m \epsilon_m^{1/3} < \epsilon_{m+1}^{1/3}. \tag{6.22}$$

Verification of the condition $((m+1).1)$. According the condition $(m.1)$ we have $\inf_{\mathcal{O}_m} |\det \partial_\xi \omega_m| > (1 - e_m) c_1$. Using $|R_m^y|^{\mathcal{L}}_{D(s_m) \times \mathcal{O}_m} \le \epsilon_m^{1/3}$ (See (4.5)), we have that $|\partial_\xi [R_m^y]| \le \epsilon_m^{1/3}$. Thus,

$$|\det \partial_\xi \omega_{m+1}| \ge (1 - e_m) c_1 - C \epsilon_m^{1/3} \ge (1 - e_{m+1}) c_1. \tag{6.23}$$

In addition, we will verify in Section 9 that

$$\text{Meas } \Pi_{m+1} \ge \text{ Meas } \Pi_0 (1 - \alpha/(m+1)^2). \tag{6.24}$$

[7]Note $\mathcal{D}_{m+1} := D(s_{m+1}, r_{m+1}) \times \Pi_{m+1} \subset D(s_m - 4\sigma_m, r_m/8) \times \mathcal{O}_{m+1}$.

[8]See (5.6-9).

[9](6.13,14) has fulfilled (6.7,8).

[10]See (5.19,20,21).

[11]Note $\varsigma_{m+1} < \varsigma_m/2$ and $\eta r = \eta_m r_m < r_{m+1}$ and $|X|_{a,p} \le |X|_{b,p}$ if $a \le b$

Hence, the condition $((m+1).1)$ is verified for the tangential frequency ω_{m+1}.

Verification of the condition $((m+1).2)$. By the condition $(m.2)$ and Lemma 4.3, it follows from (6.20) that

$$\begin{aligned} &||B_{\jmath\jmath}^{m+1}||_2 \leq ||B_{\jmath\jmath}^{m}||_2 + ||[R_{\jmath\jmath}]||_2 \\ &\leq e_m\epsilon_0 \jmath^{-d} + \varsigma_m^{-2d}\epsilon_m \jmath^{-d} < \jmath^{-d} e_{m+1}\epsilon_0. \end{aligned} \tag{6.25}$$

Similarly, we have

$$||B_{\jmath\jmath}^{m+1}||_2^{\mathcal{L}} \leq \jmath^{-d} e_{m+1}\epsilon_0^{1/3}. \tag{6.26}$$

The combination of (6.25) and (6.26) verifies the condition $((m+1).2)$. This also fulfills the assumptions in Lemma 4.8.

Verification of the condition $((m+1).3)$. Let us consider

$$H_{m+1} := H_m \circ \Phi_{m+1}. \tag{6.27}$$

According to (6.10,11) and (6.17) we get

$$H_{m+1} = N_{m+1} + \grave{R}_{m+1} + \sum_{\wp \geq m+1} \grave{R}^{m\wp} \circ \Phi_{m+1}. \tag{6.28}$$

Observe that

$$X_{\grave{R}^{m\wp}\circ\Phi_{m+1}} = (\Phi_{m+1})^* X_{\grave{R}^{m\wp}} = D\Phi_{m+1}^{-1} X_{\grave{R}^{m\wp}} \circ \Phi_{m+1}. \tag{6.29}$$

By (6.15), we have

$${}_{r_m}||D\Phi_{m+1}^{-1}||_{\varsigma_m,p,\mathcal{D}_{m+1}} \lessdot 1. \tag{6.30}$$

Furthermore, by means of (6.6) with $l=m$, we get that for $\wp \geq m+1$

$$\begin{aligned} &{}_{r_\wp}|X_{\grave{R}^{m\wp}\circ\Phi_{m+1}}|_{\varsigma_\wp,p,D(s_{m+1},r_{m+1})} \leq {}_{r_\wp}|X_{\grave{R}^{m\wp}\circ\Phi_{m+1}}|_{\varsigma_\wp,p,D(s_m,r_m)} \\ &\leq {}_{r_m}||D\Phi_{m+1}^{-1}||_{\varsigma_m,p,\mathcal{D}_{m+1}} \cdot {}_{r_\wp}|X_{\grave{R}^{m\wp}}|_{\varsigma_\wp,p,D(s_m,r_m)} \lessdot \epsilon_\wp. \end{aligned} \tag{6.31}$$

Similarly, by means of (6.8)with with $l=m$ and (6.16), we get

$${}_{r_\wp}|X_{\grave{R}^{m\wp}\circ\Phi_{m+1}}|^{\mathcal{L}}_{\varsigma_\wp,p,D(s_{m+1},r_{m+1})} \lessdot \epsilon_\wp^{1/3}, \ \wp \geq m+1. \tag{6.32}$$

Let

$$\grave{R}^{(m+1)(m+1)} = \grave{R}_{m+1} + \grave{R}^{m(m+1)} \circ \Phi_{m+1}, \tag{6.33}$$

$$\grave{R}^{(m+1)\wp} = \grave{R}^{m\wp} \circ \Phi_{m+1}, \wp \geq m+2. \tag{6.34}$$

By (6.28),

$$H_{m+1} = N_{m+1} + \sum_{\wp \geq m+1} \grave{R}^{(m+1)\wp}. \tag{6.35}$$

By combination of (6.21,22) and (6.31,32), we conclude that (6.6) holds true with $l = m+1$. It is plain that (6.5) holds true with $l = m+1$. This complete the verification of $(l.3)$ with $l = m+1$. Therefore, the proof of the iterative lemma is complete. $\square$

7. Proof of the theorem 2.1.

The proof is similar to that of [P1]. Here we give an outline. By Assumptions A,B, C, D and the smallness assumption in Theorem 2.1, the conditions $(l.1, 2, 3)$ in the iterative lemma in Section 6.2 are fulfilled with $l = 0$. Hence the iterative lemma applies to $\tilde{H}$. Inductively, we get what as follows:

(i). Domains: for $m = 0, 1, 2, \ldots,$

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

(ii). Coordinate changes:

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

(iii). Hamiltonian functions $\tilde{H}_m$ $(m = 0, 1, \ldots)$ satisfy the conditions $(l.1, 2, 3)$ with l replaced by m;

Let $\Pi_\infty = \cap_{m=0}^{\infty} \Pi_m$, $\mathcal{D}_\infty = \cap \mathcal{D}_m$. By the same argument as in [P1, pp.134], we conclude that $\Psi^m, D\Psi^m, \tilde{H}_m, X_{H_m}$ converges uniformly on the domain $\mathcal{D}_\infty$, and $X_{\tilde{H}_\infty} \circ \Psi^\infty = D\Psi^\infty \cdot X_{\omega_*}$ where

$$\tilde{H}_\infty := \lim_{m\to\infty} \tilde{H}_m = (\omega_*(\xi), y) + \sum_{\jmath\in\mathcal{N}} (\Lambda_\jmath^0 u_\jmath, u_\jmath) + \frac{1}{2} \sum_{\jmath\in\mathcal{N}} (B_{\jmath\jmath}^\infty(\xi) u_\jmath, u_\jmath)$$

here $B_{\jmath\jmath}^\infty = \lim_{m\to\infty} B_{\jmath\jmath}^m$ and X_{ω_*} is the constant vector field ω_* on the torus $\mathbb{T}^n$. Thus, $\mathbb{T}^N \times \{0\} \times \{0\}$ is an embedding torus with rotational frequencies $\omega_*(\xi) \in \omega_*(\Pi_\infty)$ of the Hamiltonian $\tilde{H}_\infty$. Returning the original Hamiltonian $\tilde{H}$, it has an embedding torus $\Phi^\infty(\mathbb{T}^n \times \{0\} \times \{0\})$ with frequencies $\omega_*(\xi)$. This proves the Theorem. □

8. Verification of Non-resonant conditions–estimate of measure.

In estimating the measure of the resonant zones it is not necessary to distinguish between the various perturbations ω_l and Ω_l of the frequencies, since only the size of the perturbation matters. Therefore, now we write ω and Ω for all of them, and by Assumptions B and C as well as (6.3) and Lemma 4.8 we have that the map $\xi \mapsto \omega(\xi)$ is analytic in each entry of $\xi \in \mathcal{O}$ here $\mathcal{O}$ is a ν-neighborhood of Π, and there are two absolute constants $c_1, c_2, c_3, c_4, c > 0$ such that

$$\inf_{\mathcal{O}} \left| \det \frac{\partial \omega}{\partial \xi} \right| \geq c_1, \ \sup_{\mathcal{O}} |\partial_\xi^j \omega_l| \leq c_2, \ j = 0, 1, \tag{8.1}$$

$$\inf_{\mathcal{O}} \lambda_j \geq c_3 > 0, \ \sup_{\mathcal{O}} |\partial_\xi \lambda_j| \leq c_4 \ll 1. \tag{8.2}$$

and

$$\inf_{\mathcal{O}} |\lambda_i - \lambda_j| \geq c|i|^{-d}|j|^{-d}, \ c > 0, |i| > |j|. \tag{8.3}$$

Lemma 8.1. *Under the condition (8.1), there is a subset Ξ^1 of Π with Lebesgue measure Meas$\Xi^1 \lessdot \alpha$ such that for any $\xi \in \Pi \setminus \Xi^1$, the non-resonant condition (4.16) is fulfilled, i.e.,*

$$|(k, \omega(\xi))| \gtrdot \frac{\alpha}{|k|^\tau}, \ \textit{for all } k \in \mathbb{Z}^n \textit{ with } 0 \neq |k| \leq K$$

where $\alpha > 0$ and $\tau > n$.

Proof. The proof is standard in KAM theory. We omit it. □

Lemma 8.2. *Under the condition (8.1) and (8.2), there is a subset Ξ^2 of Π with Meas $\Xi^2 \lessdot \alpha$ such that for any $\xi \in \Pi \setminus \Xi^2$, the non-resonant condition (4.23) is fulfilled, i.e.,*

$$|(k,\omega(\xi)) \pm \lambda_j(\xi)| \gtrdot \frac{\alpha}{\jmath^d|k|^\tau}, \quad |k| \le K, j \le M$$

where $\alpha > 0$ and $\tau > n$ and $\lambda_j \in \Lambda$[12]*.*

Proof. Let

$$\Xi^2_{k,j} = \{\xi \in \Pi : |(k,\omega(\xi)) \pm \lambda_j(\xi)| < \frac{\alpha}{\jmath^d|k|^\tau}\}, \quad \Xi^2 = \bigcup_{|k|\le K, |j|\le M} \Xi^2_{k,j},$$

and

$$\tilde{\Xi}^2_{k,j} = \{\eta \in \omega(\Pi) : |(k,\eta) \pm \lambda_j(\omega^{-1}(\eta))| < \frac{\alpha}{\jmath^d|k|^\tau}\}, \quad \tilde{\Xi}^2 = \bigcup_{|k|\le K, j\in\mathbb{Z}^d} \tilde{\Xi}^2_{k,j}.$$

By (8.1) and (8.2),

$$\sup_{\omega(\Pi)} |\partial_\eta \lambda_j(\omega^{-1}(\eta)| \ll 1 \tag{8.4}$$

and

$$\inf_{\omega(\Pi)} \lambda_j(\omega^{-1}(\eta)) > c > 0. \tag{8.5}$$

Again by (8.1), we have Meas$\Xi^2 \lessdot$ Meas$\tilde{\Xi}^2$. Observe that the set $\tilde{\Xi}^2_{k,j}$ is empty when $k = 0$ and $0 < \alpha \ll 1$, in view of (8.5). In the following argument, suppose that $k \ne 0$. Write $k = (k_1, ..., k_n)$. Suppose $k_1 \ne 0$ without loss of generality. By Lemma (8.4),

$$|\partial_{\eta_1}(k,\eta)) \pm \lambda_j(\omega^{-1}(\eta))| = |k_1 \pm \partial_{\eta_1}\lambda_j(\omega^{-1}(\eta))| \ge 1/2.$$

It follows that

$$\text{Meas } \tilde{\Xi}^2_{k,j} \lessdot \frac{\alpha}{\jmath^d|k|^\tau}.$$

Hence,

$$\text{Meas } \Xi^2 \le \text{ Meas } \tilde{\Xi}^2 \le \text{ Meas } \bigcup_{|k|\le K, |j|\le M} \tilde{\Xi}^2_{k,j} \lessdot \sum_{k\in\mathbb{Z}^n, j\in\mathbb{Z}^d} \frac{\alpha}{\jmath^d|k|^\tau} \lessdot \alpha. \quad \square$$

Lemma 8.3. *Under the condition (8.1) and (8.2), there is a subset Ξ^3 of Π with Meas $\Xi^3 \lessdot \alpha$ such that for any $\xi \in \Pi \setminus \Xi^3$, the non-resonant condition (4.36) is fulfilled, i.e.,*

$$|(k,\omega(\xi)) \pm \lambda_i(\xi) \pm \lambda_j(\xi)| \gtrdot \alpha/(\imath^d\jmath^d|k|^\tau), \quad \begin{cases} |i| = |j| = \jmath \le M, 0 \ne |k| \le K \\ |i| \ne |j|, |i|, |j| \le M, |k| \le K \end{cases}$$

Proof. Let

$$\Xi^3_{k,i,j} = \left\{\xi \in \Pi : |(k,\omega(\xi)) \pm \lambda_i(\xi) \pm \lambda_j(\xi)| < \frac{\alpha}{1+|k|^\tau}\frac{1}{\imath^d\jmath^d}\right\},$$

[12]See Lemma 4.8 for the definition of Λ.

and

$$\Xi^3 = \bigcup_{|k|\le K, |i|,|j|\le M} \Xi^3_{k,i,j},$$

where $k \neq 0$ when $|i| = |j|$. It follows from (8.3) that the set $\Xi^3_{k,i,j}$ is empty when $k = 0$ and $|i| \neq |j|$. Therefore assume $k = (k_1, ...k_n) \neq 0$. Suppose $k_1 \neq 0$ without loss of generality. Let

$$\tilde{\Xi}^3_{k,i,j} = \left\{ \eta \in \omega(\Pi) : |(k,\eta) \pm \lambda_i(\omega^{-1}(\eta)) \pm \lambda_j(\omega^{-1}(\eta))| < \frac{\alpha}{1+|k|^\tau} \frac{1}{\imath^d \jmath^d} \right\},$$

and

$$\tilde{\Xi}^3 = \bigcup_{|k|\le K, |i|,|j|\le M} \tilde{\Xi}^3_{k,i,j},$$

where $k \neq 0$ when $|i| = |j|$. By (8.1), we have $\mathrm{Meas}\Xi^3 \lessdot \mathrm{Meas}\tilde{\Xi}^3$. By 8.4, we have

$$|\partial_{\eta_1}(k,\eta) \pm \lambda_i(\omega^{-1}(\eta)) \pm \lambda_j(\omega^{-1}(\eta))| = |k_1 \pm \partial_{\eta_1}\lambda_i \pm \partial_{\eta_1}\lambda_j| \geq 1/2.$$

It follows that

$$\mathrm{Meas}\ \tilde{\Xi}^3_{k,i,j} \lessdot \frac{\alpha}{|k|^\tau} \frac{1}{\imath^d \jmath^d}.$$

Hence

$$\mathrm{Meas}\ \Xi^3 \leq\ \mathrm{Meas} \bigcup_{0<|k|\le K, |i|,|j|\le M} \tilde{\Xi}^3_{k,j} \lessdot \sum_{k\in\mathbb{Z}^d} \frac{\alpha}{|k|^\tau} \sum_{i\in\mathbb{Z}^d} \frac{1}{|i|^d} \sum_{j\in\mathbb{Z}^d} \frac{1}{|j|^d} \lessdot \alpha. \quad \square$$

Lemma 8.4. *There is a subset $\Pi_+ \subset \Pi$ with*

$$\textit{Meas}\ \Pi_+ \geq (\ \textit{Meas}\ \Pi)(1 - C\alpha). \tag{8.6}$$

And there is a positive ν_+ such that for any $\xi \in \mathcal{O}_+$, a ν_+-neighborhood of Π_+, all resonant conditions in Lemmas 8.1,2,3 hold true.

Proof. Let $\Pi_+ = \Pi \setminus (\Xi^1 \cup \Xi^2 \cup \Xi^3)$. Then (8,6) holds true clearly. Let

$$\nu_+ = \alpha/(2M^{2d}K^{\tau+1}).$$

Since $\nu_+ < \nu$, we get $\mathcal{O}_+ \subset \mathcal{O}$. Let

$$f(\xi) = (k, \omega(\xi)) \pm \lambda_i(\xi) \pm \lambda_j(\xi), |k| \leq K, |i|, |j| \leq M.$$

Then, by (8.1) and (8.2),

$$\sup_{\mathcal{O}} |\partial_\xi f(\xi)| \leq |k|c_2 + 2c_4 \lessdot K.$$

Since $\mathcal{O}_+$ is the ν_+-neighborhood of Π_+, we get that for any $\xi \in \mathcal{O}_+$, there is a $\xi_0 \in \Pi_+$ such that $|\xi - \xi_0| < \nu_+$. Thus,

$$|f(\xi) - f(\xi_0)| \leq \sup_{\mathcal{O}} |\partial_\xi f(\xi)||\xi - \xi_0| \leq K\nu_+ \leq \frac{\alpha}{2M^{2d}K^\tau}.$$

Consequently, for $\xi \in \mathcal{O}_+$ and $|k| \leq K, |i|, |j| \leq M$, we have

$$|f(\xi)| \geq |f(\xi_0)| - |f(\xi) - f(\xi_0)| \gtrdot \frac{\alpha}{|i|^d|j|^d|k|^\tau} - \frac{\alpha}{2M^{2d}K^\tau} \geq \frac{\alpha}{2|i|^d|j|^d|k|^\tau}.$$

This implies that the non-resonant conditions in Lemma 8.3 hold true for $\xi \in \mathcal{O}_+$. The remaining proof is similar to that above. $\square$

9. Appendix A. Some Technical lemmas.

Lemma A.1. *For $\mu > 0, \nu > 0$, the following inequality holds true:*

$$\sum_{k\in\mathbb{Z}^d} e^{-2|k|\mu}|k|^\nu \le (\frac{\nu}{e})^\nu \frac{1}{\mu^{\nu+d}}(1+e)^d.$$

Proof. This Lemma can be found in [B-M-S].

Lemma A.2. *Suppose that an operator $Y = Y(x,\xi) : \ell^{a,p} \to \ell^{a,p}$ is analytically dependent on $x \in D(s_*) = \{x \in \mathbb{C}^n : |\Im x| < s_*\}$ and each entry $\xi^l\,(l = 1, ..., n)$ of $\xi \in \Pi$, and suppose that following estimates hold true*

$$|Y_{ij}|_{D(s_*)\times\Pi}, |Y_{ji}|_{D(s)\times\Pi} \lessdot e^{-\varsigma|i|}|i|^\beta|i|^{-p}|j|^p\varepsilon_*, \quad |i| \ge |j|, \tag{*}$$

and

$$|Y_{ij}|^{\mathcal{L}}_{D(s_*)\times\Pi}, |Y_{ji}|^{\mathcal{L}}_{D(s_*)\times\Pi} \lessdot e^{-\varsigma|i|}|i|^\beta|i|^{-p}|j|^p\varepsilon^{\mathcal{L}}_*, \quad |i| \ge |j|, \tag{**}$$

where constants ς, β, ε_* and $\varepsilon^{\mathcal{L}}_*$ are positive. Then we have

$$\sup_{D(s_*)\times\Pi}|||Y|||_{0,\varsigma/2,p,\bar{p}} \lessdot \varsigma^{-2d-\beta}\varepsilon_*, \qquad \sup_{D(s_*)\times\Pi}|||Y|||^{\mathcal{L}}_{0,\varsigma/2,p,p} \lessdot \varsigma^{-2d-\beta}\varepsilon^{\mathcal{L}}_*$$

Proof. Let $w_i = e^{\varsigma/2|i|}|i|^p$. Let $J_i = \{j \in \mathbb{Z}^d : |j| \le |i|\}$ and $J^i = \{j \in \mathbb{Z}^d : |j| > |i|\}$. Let $I^j = \{i \in \mathbb{Z}^d : |i| \ge |j|\}$ and $I_j = \{i \in \mathbb{Z}^d : |i| < |j|\}$. For any $u \in \ell_p$ with $||u||_p = 1$, we have

$$\begin{aligned}
||Yu||^2_{\varsigma/2,p} &= \sum_{i\in\mathbb{Z}^d} w_i^2|\sum_{j\in\mathbb{Z}^d} Y_{ij}u_j|^2 \\
&\lessdot \sum_{i\in\mathbb{Z}^d} w_i^2\left(\sum_{J_i}|Y_{ij}||u_j| + \sum_{J^i}|Y_{ij}||u_j|\right)^2 \\
&\lessdot \sum_{i\in\mathbb{Z}^d} w_i^2\left(\sum_{J_i}|Y_{ij}||u_j|\right)^2 + \sum_{i\in\mathbb{Z}^d} w_i^2\left(\sum_{J^i}|Y_{ij}||u_j|\right)^2 \\
&:= (1) + (2),
\end{aligned}$$

where Y_{ij}'s are the matrix elements of Y. Note that $\{j \in \mathbb{Z}^d : |j| = \jmath\}^\sharp \le \jmath^{d-1}$ where $\sharp =$ is the cardinality of the set. By assumption (*), we have

$$\sum_{J_i}|Y_{ij}|e^{\varsigma|i|/2}|i|^p|j|^{-p} \le \sum_{J_i} e^{-\varsigma|i|/2}|i|^\beta\varepsilon_* \lessdot \sum_{l\in\mathbb{N}} l^{d+\beta-1}e^{-\varsigma l/2}\varepsilon_* \lessdot \varsigma^{-2d-\beta}\varepsilon_*,$$

where Lemma A.1 is used in the last inequality. Similarly,

$$\sum_{I^j}|Y_{ij}|e^{\varsigma|i|/2}|i|^p|j|^{-p} \lessdot \varsigma^{-2d-\beta}\varepsilon_*,$$

$$\sum_{I_j}|Y_{ij}|e^{\varsigma|i|/2}|i|^p|j|^{-p} \lessdot \varsigma^{-2d-\beta}\varepsilon_*,$$

and

$$\sum_{J^i} |Y_{ij}| e^{\varsigma|j|/2} |j|^p |i|^{-p} \lessdot \varsigma^{-2d-\beta} \varepsilon_*.$$

By Hölder inequality, we have

$$\begin{aligned}
(1) &\leq \sum_i e^{\varsigma|i|} |i|^p \left(\sum_{J_i} |Y_{ij}| e^{\varsigma|i|/2} |i|^p |j|^{-p} \right) \left(\sum_{J_i} |Y_{ij}| e^{-\varsigma|i|/2} |j|^p |u_j|^2 \right) \\
&\lessdot \varsigma^{-2d-\beta} \varepsilon_* \sum_i e^{\varsigma|i|} |i|^p \sum_{J_i} |Y_{ij}| e^{-\varsigma|i|/2} |j|^p |u_j|^2 \\
&\lessdot \varsigma^{-2d-\beta} \varepsilon_* \sum_j (\sum_{I^j} |Y_{ij}| e^{\varsigma|i|/2} |i|^p |j|^{-p})(|j|^{2p} |u_j|^2) \\
&\lessdot (\varsigma^{-2d-\beta} \varepsilon_*)^2 \sum_j |j|^{2p} |u_j|^2 = (\varsigma^{-2d-\beta} \varepsilon_*)^2 ||u||_p^2 = \varsigma^{-4d-2\beta} \varepsilon_*^2.
\end{aligned}$$

We are now in position to estimate (2). Again by Hölder inequality we have

$$\begin{aligned}
(2) &\leq \sum_i w_i^2 \sum_{J^i} |Y_{ij}| e^{\varsigma|j|/2} |j|^p |i|^{-p} \sum_{J^i} |Y_{ij}| e^{-\varsigma|j|/2} |j|^{-p} |i|^p |u_j|^2 \\
&\lessdot \varsigma^{-2d-\beta} \varepsilon_* \sum_i w_i^2 \sum_{J^i} |Y_{ij}| e^{-\varsigma|j|/2} |j|^{-p} |i|^p |u_j|^2 \\
&= \varsigma^{-2d-\beta} \varepsilon_* \sum_j \left(\sum_{I_j} |Y_{ij}| e^{\varsigma|i| - \varsigma|j|/2} |i|^{2p+p} |j|^{-\bar{p}-2p} \right) |j|^{2p} |u_j|^2 \\
&\leq \varsigma^{-2d-\beta} \varepsilon_* \sum_j \left(\sum_{I_j} |Y_{ij}| e^{\varsigma|i|/2} |i|^p |j|^{-p} \right) |j|^{2p} |u_j|^2 \\
&\leq \varsigma^{-4d-2\beta} \varepsilon_*^2 \sum_j |j|^{2p} |u_j|^2 = \varsigma^{-4d-2\beta} \varepsilon_*^2.
\end{aligned}$$

Consequently, we get that for any $(x.\xi) \in D(s_*) \times \Pi$,

$$||Y(x,\xi)u||_{\varsigma/2,p} \lessdot \varsigma^{-2d-\beta} \varepsilon_*, \quad \text{for} \quad ||u||_p = 1.$$

That is,

$$\sup_{D(s_*)\times\Pi} |||Y|||_{0,\varsigma/2,p,p} \lessdot \varsigma^{-2d-\beta} \varepsilon_*.$$

Using the same arguments as above for $\partial_{\xi_j} Y_{ng}$ with $j = 1, ..., n$ and using (**), we can get the estimates for $|||Y|||^{\mathcal{L}}_{0,\varsigma/2,p,\bar{p}}$. This completes the proof. $\square$

Lemma A.3. *Suppose that $Y = X + Z$ where both X and Z are hermitian matrices of order m, and the eigenvalues of Y and X are $\lambda_1 \geq ... \geq \lambda_m$ and $\mu_1 \geq ... \geq \mu_m$, respectively. Then*

$$|\lambda_l - \mu_l| \leq ||Z||_2, \quad l = 1, ..., m.$$

Proof. The proof can be found in most text books on matrix theory.

Lemma A.4. *Consider an $n \times n$ complex matrix function $Y(\xi)$ which depends on the real parameter $\xi \in \mathbb{R}$. Let $Y(\xi)$ be a matrix function satisfying conditions:*

(i) $Y(\xi)$ is self-adjoint for every $\xi \in \mathbb{R}$; i.e., $Y(\xi) = (Y(\xi))^*$, where star denotes the conjugate transpose matrix;

(ii) $Y(\xi)$ is an analytic function of the real variable ξ.

Then there exist scalar functions $\mu_1(\xi), \cdots, \mu_n(\xi)$ and a matrix-valued function $U(\xi)$, which are analytic for real ξ and possess the following properties for every $\xi \in \mathbb{R}$:

$$Y(\xi) = U(\xi) diag(\mu_1(\xi), \cdots, \mu_n(\xi)) U^*(\xi), \quad U(\xi)(U(\xi))^* = E.$$

Proof. See [pp.394-396, G-L-R]. □

It is worth to point out that this lemma does not hold true for $\xi \in \mathbb{R}^k$ with $k > 1$. See [Ka].

Lemma A.5. *Under the same assumptions as in Lemma A.4, we have*

$$|\mu_l'(\xi)| \le ||Y'(\xi)||_2, \quad l = 1, ..., m, \ \textit{here} \ ' = \frac{d}{d\xi}.$$

Proof. See [Ka, p.125]. □

10. Appendix B. Theorem on regularity of linear operator.

In order to overcome the difficulty arising from the delicate small divisors of the form $\langle k, \omega \rangle + \lambda_i - \lambda_j$ with $|i| \neq |j|$, we have to raise up the regularity of the linear operator R^{uu} coming from the second term $\langle R^{uu} u, u \rangle$ of the perturbed Hamiltonian. We start with some natation and definitions. For $x = (x_1, ..., x_d)$, we denote $D_j = \partial/\partial x_j$, $D^k = D_1^{k_1} \circ D_2^{k_2} \circ \cdots \circ D_d^{k_d}$, $|k| = \sum_{j=1}^d k_j$. We define the complex strips U_a for all $a > 0$ as follows:

$$U_a = \{x \in \mathbb{C}^d/(2\pi\mathbb{Z})^d : |\Im x_j| < a, j = 1, ..., d\}.$$

For a function $u : U_a \to \mathbb{C}$ and integers $p^* \ge 0$, we introduce the seminorms

$$|u|_{a,p^*} = \sup_{x \in U_a, |k| = p} |D^k u(x)|.$$

When $a = 0$, we write $|u|_{0,p^*}$ as $|u|_{p^*}$. Let $C^{p^*}(T)$ be the set of all functions defined on T with $\sup_{x \in T, |k| = p^*} |D^k u(x)| < \infty$. For $p^* \ge 0$, the Banach spaces $A(a, C^{p^*})$ are then defined as spaces of real holomorphic functions u on U_a (u being real means $\overline{u(x)} = u(\bar{x})$), with period 2π in each variable and such that $|u|_{a,p} < \infty$. Take a function $\tilde{s} \in C_0^\infty(\mathbb{R})$, vanishing outside a compact set and identically equal to 1 in a neighborhood of 0, and let s be its Fourier transform. Moreover, we can require $s(x)$ is even function. When $x = (x_1, ..., x_d) \in \mathbb{R}^d$, by a slight abuse of notation, we denote $s(x) = s(x_1) \cdots s(x_d)$. For $a > 0$ we introduce the families of linear operators $S_a : C^p(\mathbb{T}^d) \to A(a, C^p)$, by means of the convolution $S_a u = s_a \bar{\star} u$, $s_a(z) = a^{-d} s(a^{-1} z)$:

$$s_a \bar{\star} u(z) = a^{-d} \int_{\mathbb{R}^d} s\left(\frac{y - z}{a}\right) u(y) \ dy, \quad u \in C^p(\mathbb{T}^d).$$

It is clear that $S_a u$ is an entire real holomorphic function on $\mathbb{C}^d$ and has period 2π since u has period 2π.

Lemma B.1. *There exists a constant $C = C(p,d) \geq 1$ depending only on positive integers p and d such that, for all $0 \leq \sigma \leq a$,*

$$|(S_a - S_\sigma)u|_{\sigma,p_*} \leq C|u|_{p^*}a^{p^*-p_*}, \quad 0 \leq p_* \leq p^*$$

$$|(S_a - 1)u|_{p_*} \leq C|u|_{p^*}a^{p^*-p_*}$$

and for $\alpha \in \mathbb{N}^d$ with $|\alpha| \leq p_*$,

$$\sup_{|\Im x| \leq a} |D^\alpha S_a u(x) - \sum_{|\beta| \leq p_* - |\alpha|} \frac{D^{\alpha+\beta}u(\Re x)}{\beta!}(\sqrt{-1}\Im x)^\beta| \leq C|u|_{p_*}a^{p_*-|\alpha|},$$

in particular, for $|\alpha| = p_*$,

$$|S_a u|_{a,p_*} = \sup_{|\Im x| \leq a} |D^\alpha S_a u(x)| \leq (1+C)|u|_{p_*}.$$

Proof. This lemma is the so-called Jackson's analytic approximation theorem. The proof consists in a direct check based on standard tools from calculus and some simple properties of Fourier transform. Refer to [Z].

Remark. If u depends on some parameter $\xi \in \Pi \subset \mathbb{R}^n$ and if the Lipschitz semi-norm of u and its $x-$derivatives are uniformly bounded by $|u|_p^{\mathcal{L}}$, then all estimates in Lemma 8.1 hold true with $|\cdot|$ replaced by $|\cdot|^{\mathcal{L}}$. The proof in [Z] is still valid here only if $|\cdot|$ is replaced by $|\cdot|^{\mathcal{L}}$.

Let

$$H^{p^*}(\mathbb{T}^d) = \{u \in L^2(\mathbb{T}^d) : ||u||_{p^*}^2 = \sum_j |j|^{2p^*}|\widehat{u}(k)|^2 < \infty\}.$$

Define $\mathcal{F} : \ell^{p^*} \to H^{p^*}(\mathbb{T}^d)$ by

$$\mathcal{F}(q) = \sum_j q_j e^{\sqrt{-1}\langle j,x\rangle}, \quad q \in \ell^{p^*}.$$

By means of Parseval equality, $\mathcal{F} : \ell^{p^*} \to H^{p^*}(\mathbb{T}^d)$ is isometric.

Lemma B.2. *If $u \in H^{p^*}(\mathbb{T}^d)$ with $p^* > d/2$, then $u \in C^{p^*-d/2}(\mathbb{T}^d)$ and there is an absolute constant C such that $|u|_{p^*-d/2} \leq C||u||_{p^*}$.*

Proof. Formally, for $k \in \mathbb{R}_+^d$ with $|k| = p^* - d/2$ and $u \in H^p(\mathbb{T}^d)$,

$$D^k u = \sum_j \hat{u}(j)(\sqrt{-1}j)^k e^{\sqrt{-1}\langle j,x\rangle}.$$

Then

$$\begin{aligned}
&\sum_j \sup_{x \in \mathbb{T}^d} |\hat{u}(j)(\sqrt{-1}j)^k e^{\sqrt{-1}\langle j,x\rangle}| \leq \sum_j |\hat{u}(j)||j|^{p^*-d/2} \\
&\leq (\sum_j |\hat{u}(j)|^2|j|^{2p^*})^{1/2} \cdot (\sum_j 1/|j|^d)^{1/2} \leq C||u||_{p^*},
\end{aligned}$$

so $u \in C^{p^*-d/2}(\mathbb{T}^d)$ and $|u|_{p^*-d/2} \leq C||u||_{p^*}$.

Let us now take $q \in \ell^{p^*}$. Then $u(x) = \mathcal{F}(q) \in H^{p^*}$. It is plain that $(\hat{u}(j))_{j\in\mathbb{Z}^d} = q$. By Lemma B.2, we have $u \in C^{p^*-d}(\mathbb{T}^d)$ and $|u|_{p^*-d} \le C||u||_{p^*}$. By Lemma 7.1, for any $0 < \tau < \sigma$, the functions $S_\tau u, S_\sigma u$ are entire real holomorphic functions on $\mathbb{C}^d$ and has period 2π; moreover, letting $p^* = \bar{p} - d/2$ and $p_* = p$ and recalling $\bar{p} - p = \kappa$ and using Lemmas B.1 and B.2 we have the following estimates hold:

$$|(S_\sigma - S_\tau)u|_{\tau,p} \lessdot |u|_{\bar{p}-d/2}\sigma^{\kappa-(d/2)} \lessdot ||u||_{\bar{p}}\sigma^{\kappa-(d/2)}, \tag{10.1}$$

and

$$|(S_\tau - 1)u|_p \lessdot |u|_{\bar{p}-d/2}\tau^{\kappa-(d/2)} \lessdot ||u||_{\bar{p}}\sigma^{\kappa-(d/2)}, \tag{10.2}$$

$$|S_\sigma u|_{\sigma,p} \lessdot |u|_p. \tag{10.3}$$

Note that $(S_\sigma - S_\tau)u$ is analytic in the strip $|\Im x| \le \tau$. By means of Cauchy's formula and (10.1) we get

$$|j|^p|\widehat{S_\tau u}(j) - \widehat{S_\sigma u}(j)| \lessdot e^{-\tau|j|} \cdot ||u||_{\bar{p}}\sigma^{\kappa-(d/2)}.$$

It follows that

$$\sum_{j\in\mathbb{Z}^d} |j|^{2p}e^{2\tau|j|}|\widehat{S_\tau u}(j) - \widehat{S_\sigma u}(j)|^2 \lessdot ||u||_{\bar{p}}^2\sigma^{2\kappa-d}. \tag{10.4}$$

Note that $||u||_{\bar{p}} = ||\mathcal{F}(q)||_{\bar{p}} = ||q||_{\bar{p}}$, since $\mathcal{F} : \ell^{\bar{p}} \to H^{\bar{p}}(\mathbb{T}^d)$ is isometric. Let

$$q_\sigma(j) = \widehat{S_\sigma u}(j), \quad q_\sigma = (q_\sigma(j))_{j\in\mathbb{Z}^d}.$$

Then (10.4) implies that $q_\sigma - q_\tau \in \ell^{\tau,p}$ and

$$||q_\sigma - q_\tau||_{\tau,p} \lessdot ||u||_{\bar{p}}\sigma^{\kappa-(d/2)} = ||q||_{\bar{p}}\sigma^{\kappa-(d/2)}. \tag{10.5}$$

By (10.2), we have

$$||q_\sigma - q||_p \lessdot ||q||_{\bar{p}}\sigma^{\kappa-(d/2)}. \tag{10.6}$$

Using (10.3), we get

$$||q_\sigma||_{\sigma,p} \lessdot ||q||_p. \tag{10.7}$$

For any $0 < \sigma$, we define an operator $T_\sigma : \ell^{\bar{p}} \to \ell^{a,p}$ by means of

$$T_\sigma q = q_\sigma, \quad q \in \ell^{\bar{p}}.$$

In view of (10.7), the operator is well defined and bounded. It is plain that $T_\sigma = \mathcal{F}^{-1} \circ S_\sigma \circ \mathcal{F}$, and it is linear since S_σ and $\mathcal{F}$ are linear. We can now rewrite (10.5-7) as

$$|||T_\sigma - T_\tau|||_{0,\tau,\bar{p},p} \lessdot \sigma^{\kappa-(d/2)}. \tag{10.8}$$

$$|||T_\sigma - 1|||_{0,0,\bar{p},p} \lessdot \sigma^{\kappa-(d/2)}. \tag{10.9}$$

$$|||T_\sigma|||_{0,\sigma,p,p} \lessdot 1. \tag{10.10}$$

Now, given a decreasing sequence $\varsigma_m = \epsilon_m^\wedge(\frac{4}{2\kappa-d}) \downarrow 0$ $(m = 0, 1, 2, ...)$[13], we get a family of bounded linear operators $T_m := T_{\varsigma_m}$ from ℓ^p to $\ell^{\varsigma_m,p}$.

[13]See Section 6.1 for ς_m and ϵ_m.

Lemma B.3. *There are a family of operators $T_m : \ell^{0,p} \to \ell^{\varsigma_m,p}$ $(m = 0, 1, ...)$ such that*

$$|||T_m - T_{m+1}|||_{0,\varsigma_{m+1},\bar{p},p} \lessdot \epsilon_m^2 < \epsilon_{m+1}, \tag{10.12}$$

$$|||T_m - 1|||_{0,0,\bar{p},p} \lessdot \epsilon_{m+1}, \tag{10.13}$$

$$|||T_m|||_{0,\sigma,p,p} \lessdot 1, \quad \forall\ 0 \le \sigma \le \varsigma_m. \tag{10.14}$$

Proof. This lemma is a direct result of (10.8,9,10). □

Lemma B.4. *The composition $T_m \circ \tilde{\Psi}$ of T_m and $\tilde{\Psi}$ is self-adjoint in ℓ_{20}.*[14]

Proof. Let $S_m := S_{\varsigma_m}$. Then the operator $T_m \circ \tilde{\Psi}$ is self-adjoint in ℓ_{20} if and only if the operator $\mathcal{S}_m := S_m \circ \Psi$ is is self-adjoint in L_0^2. It is easy to verify that

$$S_m \circ \Psi(u) = (s_a \bar{\star} \psi) \star u, \quad a = \varsigma_m.$$

For any $u, v \in L_0^2$, (We can assume u, v are real without loss of generality.), then

$$\begin{aligned}
\langle \mathcal{S}_m u, v\rangle &= \int_0^{2\pi} v(z)(s_a \bar{\star} \psi) \star u(z)\ dz \\
&= \int_0^{2\pi} \int_0^{2\pi} v(z)(s_a \bar{\star} \psi)(z - t)u(t)\ dtdz \\
&= \int_0^{2\pi} u(t) \int_0^{2\pi} (s_a \bar{\star} \psi)(z - t)v(z)\ dzdt \\
&= \int_0^{2\pi} u(t) \int_0^{2\pi} (s_a \bar{\star} \psi)(t - z)v(z)\ dzdt \\
&= \int_0^{2\pi} u(t)(s_a \bar{\star} \psi) \star v(z)\ dz \\
&= \langle u, \mathcal{S}_m v\rangle,
\end{aligned}$$

where the fact $s_a(-x) = s(x)$ and $\psi(-x) = \psi(x)$ are used in the fourth equality. Note the operator $\mathcal{S}$ is bounded. The proof is complete. □

References.

[B] Bambusi, D., Birkhoff normal form for some nonlinear PDEs, Commun. Math. Phys., **234**, 253–285 (2003)

[B-B] Berti, M. and Bambusi, D., A Birkhoff-Lewis type theorem for some Hamiltonian PDEs, preprint, (2003)

[B-G] Bambusi, D. and Grébert, Birkhoff normal form for PDEs with TAME modulus, Duke Math. *to appear*

[Bo-K] Bobenko, A. I., and Kuksin, S., The nonlinear Klein-Gordon equation on an interval as perturbed sine-Gordon equation, Comm. Math. Helv. **70**, 63–112 (1995)

[Br-K-S] Bricmont, J., Kupiainen, A. and Schenkel, A., Renormalization group and the Melnikov Problems for PDE's, Commun. Math. Phys., **221**, 101–140 (2001)

[14] See Sect.3 for the definitions of ℓ_{20} and L_0^2.

[B-M-S] Bogolyubov, N. N., Mitropolskij, Yu. A. and Samojlenko, A. M.: Methods of Accelerated Convergence in Nonlinear Mechanics. Springer-Verlag, New York (1976) [Russian Original: Naukova Dumka, Kiev, 1969]

[Bo1] Bourgain, J., Construction of quasi-periodic solutions for Hamiltonian perturbations of linear equations and application to nonlinear pde, Int. Math. Research Notices **11** 475–497(1994)

[Bo2] Bourgain, J., Periodic solutions of nonlinear wave equations, Harmonic analysis and partial equations, Chicago Univ. Press, 1999, pp.69-97

[Bo3] Bourgain, J., Quasi-periodic solutions of Hamiltonian perturbations for 2D linear Schrödinger equation, Ann. Math. **148** (1998) 363–439

[Bo4] Bourgain, J., Green function estimates for lattice Schrödinger operators and applications, preprint

[B-P] Bambusi, D., and Paleari, S., Families of periodic orbits for resonant PDE's, J. Nonlinear Science, **11** (2001), 69-87

[Br] Brézis, H., Periodic solutions of nonlinear vibrating strings and duality principles, Bull. AMS **8** 409–426(1983)

[C-W] Craig, W. and Wayne, C., Newton's method and periodic solutions of nonlinear wave equation, Commun. Pure. Appl. Math.,**46** 1409–1501 (1993)

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

[E] Eliasson, L. H. Perturbations of stable invariant tori for Hamiltonian systems, Ann. Scula Norm. Sup. Pisa CL Sci. **15** 115-147 (1988)

[F-S] Frohlich,J., Spencer,T. Absence of diffusion in the Anderson tight binding model for large disorder or lower energy. Commun. Math. Phys.**88** 151-184 (1983)

[G-Y] Geng, J. and You, J., a talk in the conference on dynamical systems in Huang Shan, 2005

[G-L-R] Gohberg, I., Lancaster, P., and Rodman, L.: Matrix Polynomials. *(Computer Science and Applied Math.,)* Academic Press, New York London (1982)

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

[K1] Kuksin, S. B.: Nearly integrable infinite-dimensional Hamiltonian systems.(Lecture Notes in Math. 1556). Springer-Verlag, New York (1993)

[K2] Kuksin, S. B.: Elements of a qualitative theory of Hamiltonian PDEs, Doc. Math. J. DMV (Extra Volume ICM) **II** 819–829 (1998)

[L-S] Lidskij, B. V. and Shulman, E., Periodic solutions of the equation $u_{tt}-u_{xx}+u^3=0$, Funct. Anal. Appl. **22** 332-333 (1988)

[P1] Pöschel, J., A KAM-theorem for some nonlinear PDEs, Ann. Scuola Norm. Sup. Pisa, Cl. Sci., IV Ser. 15, **23** 119–148 (1996)

[P2] Pöschel, J., Quasi-periodic solutions for nonlinear wave equation, Commun. Math. Helvetici **71** 269–296 (1996)

[P3] Pöschel, J., On the construction of almost periodic solutions for a nonlinear Schrödinger equation, Ergod. Th. & Dynam. Syst. **22** 1537–1549 (2002)

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

[Y1] Yuan, X., Invariant Manifold of Hyperbolic-Elliptic Type for Nonlinear Wave Equation, Int. J. Math. Math. Science, **18** 1111-1136 (2003)

[Y2] Yuan, X., Invariant tori of nonlinear wave equations with a given potential. Preprint. FDIM 2004-11, Fudan University (2004), *(to appear in DCDS).*

[Y3] Yuan, X., *Quasi-periodic solutions of completely resonant nonlinear wave equations. Preprint. FDIM 2004-12, Fudan University (2004).*

Department of Mathematics
Fudan University
Shanghai 200433
China
E-mail address: yuanxiaoping@hotmail.com; xpyuan@fudan.edu.cn