Available online at www.sciencedirect.com
SCIENCE DIRECT®
J. Differential Equations 209 (2005) 1–56
Journal of Differential Equations
www.elsevier.com/locate/jde

A KAM theorem for one dimensional Schrödinger equation with periodic boundary conditions☆

Jiansheng Geng, Jiangong You*

Department of Mathematics, Nanjing University, Nanjing 210093, P.R. China

Received 15 January 2003; revised 30 August 2004

Available online 13 November 2004

Abstract

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

$$iu_t - u_{xx} + mu + \frac{\partial g(u,\bar{u})}{\partial \bar{u}} = 0,$$

with *Periodic Boundary Conditions* is considered; $m \notin \frac{1}{12}\mathbb{Z}$ is a real parameter and the nonlinearity

$$g(u,\bar{u}) = \sum_{j,l,j+l\geqslant 4} a_{jl} u^j \bar{u}^l, \quad a_{jl} = a_{lj} \in \mathbb{R}, \quad a_{22} \neq 0$$

is a real analytic function in a neighborhood of the origin. The KAM machinery is adapted to fit the above equation so as to construct small-amplitude periodic or quasi-periodic solutions corresponding to finite dimensional invariant tori for an associated infinite dimensional dynamical system.

Keywords: Schrödinger equation; KAM tori; Decay property; Birkhoff normal form

☆ The work was supported by the National Natural Science Foundation of China (19925107) and the Special Funds for Major State Basic Research Projects (973 projects).

* Corresponding author.

E-mail address: jyou@nju.edu.cn (J. You).

0022-0396/$ - see front matter © 2004 Elsevier Inc. All rights reserved.
doi:10.1016/j.jde.2004.09.013

1. Introduction and main result

The dynamics of linear Hamiltonian partial differential equations is quite clear: in many cases, the equation has families of periodic solutions, quasi-periodic solutions and almost-periodic solutions. The stability of the solutions is also obvious. One would like to know if those solutions and the related dynamics continue to the nonlinear equations in the neighborhood of equilibrium. There are plenty of works along this line. The existence of a Cantor family of quasi-periodic solutions and its linear stability for $1D$ Hamiltonian PDEs under *Dirichlet Boundary Conditions* by KAM theory are obtained by Wayne [15], Kuksin [10], Pöschel [12,13], Kuksin and Pöschel [11]. To construct the quasi-periodic solutions of Hamiltonian PDEs under *Periodic Boundary Conditions*, Craig and Wayne [8] developed a new method based on Lyapunov–Schmidt procedure and techniques by Fröhlich and Spencer [9], and later improved by Bourgain [3–6]. More remarkably, this method works for higher space dimension case [5,6]. We point out that, following Craig–Wayne–Bourgain's method, one is not able to obtain linear stability and zero Lyapunov exponents of quasi-periodic solutions comparing with the approach of Wayne [15], Kuksin [10], Pöschel [12,13], Kuksin and Pöschel [11].

The difference between Dirichlet boundary conditions and periodic boundary conditions is the multiplicity of eigenvalues in latter case. This fact would bring a lot of trouble for KAM approach. A natural problem is whether the infinite dimensional KAM theory can be generalized to deal with *Periodic Boundary Condition* case as well as higher space dimension case. Essentially, KAM machinery includes two parts: analytic part which deals with the iteration and proves convergence under some small divisor conditions, and geometric part which proves that the parameter set left after infinitely many times iteration has positive Lebesgue measure. In [7], Chierchia and You improved the analytic part of the KAM machinery so that it applies to multiple normal frequency case encountered in 1D Hamiltonian PDEs with periodic boundary conditions. For the geometric part, one has to assume a kind of regularity property, i.e., the vector field generated by nonlinear terms of the PDEs sends a sequence with decay to a sequence with faster decay to guarantee that there are essentially finitely many resonances at each KAM step (for truncated K). 1D wave equations with periodic boundary conditions meet such kind of regularity requirement. Recently, the authors of this paper further generalized the KAM theorem so that it applies to beam equations of higher space dimension.

Unfortunately, Schrödinger equation does not have such kind of regularity. In this paper, by digging out some "decay property" from nonlinearity of equation, we will prove that the KAM machinery does apply to non-linear Schrödinger equation with periodic boundary conditions. It concludes that the equation admits small-amplitude periodic or quasi-periodic solutions corresponding to finite dimensional invariant tori. We then give a new proof for Bourgain's result on the existence of quasi-periodic solutions of 1D Schrödinger equation with periodic boundary conditions. What's more, besides the existence result, the solutions we obtained have linear stability and zero Lyapunov exponents.

Let us firstly mention the works existed for this equation concerning the existence of quasi-periodic solutions of 1D nonlinear Schrödinger equations. In [11], Kuksin and

Pöschel considered the equation with "special" nonlinearity

$$iu_t - u_{xx} + mu + f(|u|^2)u = 0, \quad m \in \mathbb{R} \tag{1.1}$$

under *Dirichlet Boundary Conditions* and with f real analytic in a neighborhood of the origin. It is proved that there exists a Cantor manifold of quasi-periodic oscillations in a sufficiently small neighborhood of the origin. The existence of quasi-periodic solutions for Eq. (1.1) under *Periodic Boundary Conditions* is given by Bourgain [5]. In [3], $1D$ Schrödinger equation with variable "potential"

$$iu_t - u_{xx} + V(x,\xi)u + \varepsilon\frac{\partial g}{\partial \bar{u}} = 0, \tag{1.2}$$

with *Periodic Boundary Conditions* and g a real analytic Hamiltonian on a neighborhood of the origin is also considered. It is proved that for most ξ (in the sense of Lebesgue measure), there exist perturbed quasi-periodic solutions with perturbed frequencies. Remarkably, in [6], Bourgain constructed the quasi-periodic solutions of nonlinear Schrödinger equation and nonlinear wave equation in arbitrary space dimension by controlling Green's functions relying on more powerful methods than those in [5].

In this paper, we consider $1D$ Schrödinger equation with constant potential and more general nonlinearity

$$iu_t - u_{xx} + mu + \frac{\partial g(u,\bar{u})}{\partial \bar{u}} = 0,$$

with periodic boundary conditions

$$u(t,x) = u(t, x+2\pi),$$

m is a real parameter and the nonlinearity

$$g(u,\bar{u}) = \sum_{j,l,\, j+l \geqslant 4} a_{jl} u^j \bar{u}^l, \quad a_{jl} = a_{lj} \in \mathbb{R}, \quad a_{22} \neq 0$$

is a real analytic function in a neighborhood of the origin.

The following theorem is the main result of this paper.

Theorem 1. *Consider $1D$ nonlinear Schrödinger equation*

$$iu_t - u_{xx} + mu + \frac{\partial g(u,\bar{u})}{\partial \bar{u}} = 0, \tag{1.3}$$

with Periodic Boundary Conditions

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

and $m \notin \frac{1}{12}\mathbb{Z}$[1] *a real parameter, g is a real analytic function of the form*

$$g(u,\bar{u}) = \sum_{j,l,j+l\geqslant 4} a_{jl}u^j\bar{u}^l, \quad a_{jl} = a_{lj} \in \mathbb{R}, \quad a_{22} \neq 0. \tag{1.5}$$

Fix finite number of integers $n_1,\ldots,n_d$ *satisfying* $|n_i| \neq |n_j|$ *for* $i \neq j$, *then linearized equation has solutions*

$$u(t,x) = \sum_{i=1}^{d} \sqrt{\xi_i} e^{i(\omega_i t + n_i x)}, \quad \omega_i = n_i^2 + m, \quad \xi_i > 0.$$

Take $\xi = (\xi_1,\ldots,\xi_d) \in \mathcal{O} \subset \mathbb{R}_+^d$ *as parameters, then there exists a positive-measure Cantor subset* $\tilde{\mathcal{O}} \subset \mathcal{O}$, *such that for any* $\xi \in \tilde{\mathcal{O}}$, *there is a real analytic quasi-periodic solution of* (1.3)

$$u(t,x) = \sum_{i=1}^{d} \sqrt{\xi_i} e^{i(\tilde{\omega}_i t + n_i x)} + O(|\xi|^{\frac{3}{2}})$$

with

$$\tilde{\omega}_i = \omega_i + O(|\xi|), \qquad 1 \leqslant i \leqslant d.$$

Since the result itself is in principle not new we will emphasize how KAM machinery can be applied to the nonlinear Schrödinger equations. We hope the underlying idea might be helpful for more general problems, such as higher dimensional case.

KAM machinery is built up with infinite many KAM iteration steps. Roughly speaking, each KAM step is a change of variables which transforms the Hamiltonian into a nice normal form plus a *smaller* perturbation. For this purpose we have to solve some homological equations which forces us to assume that the tangential frequencies and normal frequencies (infinitely many) together satisfy some non-resonant relations. For example, to get rid of the term as $e^{i\langle k,\theta\rangle}w_n\bar{w}_m$ (where w_n, w_m are normal variables, θ is the angle variable), one need to assume the non-resonance condition $|\langle k,\omega(\xi)\rangle + \lambda_n(\xi) - \lambda_m(\xi)| \geqslant \frac{\gamma}{|k|^\tau}$. In order to satisfy this condition, one must discard some parameters. In general, the measure of such set of ξ for each fixed k,n,m is of $\frac{\gamma}{|k|^\tau}$. Since n,m range from 1 to infinity, there might be no ξ satisfying all the

[1] We use the notations $\mathbb{Z} = \{\ldots,-1,0,1,\ldots\}$, $\mathbb{N} = \{0,1,2,\ldots\}$, $\mathbb{Z}_+ = \{1,2,\ldots\}$.

above relations if we insist to get rid of all terms of the form $e^{i\langle k,\theta\rangle}w_n\bar{w}_m$. Such problem does not appear in case $\lambda_n(\xi)-\lambda_m(\xi)$ are integers as in the first KAM iteration step. However, the problem appears from the second step since both the tangential frequencies and the normal frequencies will drift a little bit. In the multiple normal frequency case, say $\lambda_n=\lambda_{-n}$ at the first KAM step, the drifted frequencies are $\tilde{\lambda}_n=\lambda_n+O(\varepsilon)$ and $\tilde{\lambda}_{-n}=\lambda_n+O(\varepsilon)$ respectively after iterating. Then $|\langle k,\omega(\xi)\rangle+\tilde{\lambda}_n-\tilde{\lambda}_{-n}|=|\langle k,\omega(\xi)\rangle+O_{n,-n}(\varepsilon)|\geqslant\frac{\gamma}{|k|^\tau}$ is required if we want to get rid of the terms $e^{i\langle k,\theta\rangle}w_n\bar{w}_{-n}$. There are infinitely many such terms. If the tails $O_{n,-n}(\varepsilon)$ are random, we might arrive at an empty parameter set even after one step of KAM iteration. In the case of 1D wave equations with periodic boundary condition, the Hamiltonian has a kind of regularity property which guarantees that the drift of frequencies is asymptotically zero, i.e., $\tilde{\lambda}_n-\lambda_n=O(n^{-1})$. With this property, one finds that $|\langle k,\omega(\xi)\rangle+\lambda_n-\lambda_m+O_{mn}(\varepsilon)|\geqslant\frac{\gamma}{|k|^\tau}$ is asymptotically $|\langle k,\omega(\xi)\rangle+\lambda_n-\lambda_m|\geqslant\frac{\gamma}{|k|^\tau}$ when $n,m\to\infty$. The latter are small divisor conditions assumed to be satisfied before doing KAM iteration. This key property will lead to a positive measure estimate of the remained parameter set after infinitely many KAM steps. However, the Schrödinger equations do not have such regularity property. Instead, in this paper, we use a kind of decay property of the vector field generated by partial differential equations. Roughly speaking, the decay property implies that the coefficient of the term $a_{n(-n)}e^{i\langle k,\theta\rangle}w_n\bar{w}_{-n}$ is of exponential small $O(\varepsilon e^{-2|n|\rho})$. Therefore, the terms with $|n|>K\sim|\log\varepsilon|$ need not to be killed in the present KAM step since they are already sufficiently small.[2] As a result, we need not to assume $|\langle k,\omega(\xi)\rangle+\lambda_n(\xi)-\lambda_{-n}(\xi)|\geqslant\frac{\gamma}{|k|^\tau}$ for $|n|>K$ (K being the truncation of Fourier modes at each KAM step). It is this fact that makes the positive measure estimate available! In fact, all the other non-resonance conditions are easy to control as in the Dirichlet boundary condition case:

1. The fact $\lambda_n\sim n^2$ implies that the *non-trivial* non-resonance assumptions in

$$|\langle k,\omega(\xi)\rangle\pm\lambda_n(\xi)|\geqslant\frac{\gamma}{|k|^\tau}$$

$$|\langle k,\omega(\xi)\rangle+\lambda_n(\xi)+\lambda_m(\xi)|\geqslant\frac{\gamma}{|k|^\tau}$$

are of finite many for fixed k.

2. The fact $|\lambda_n-\lambda_m|\sim|n|+|m|$ for $|n|\neq|m|$ implies that the non-trivial non-resonance assumptions in

$$|\langle k,\omega(\xi)\rangle+\lambda_n(\xi)-\lambda_m(\xi)|\geqslant\frac{\gamma}{|k|^\tau}$$

are of finite many for fixed k if $|n|\neq|m|$.

[2] One of the key ideas in KAM theory is to remove those “big” bad terms at each KAM step, while keep the smaller ones to the following KAM steps.

3. While

$$|\langle k, \omega(\xi)\rangle + \lambda_n(\xi) - \lambda_{-n}(\xi)| \geqslant \frac{\gamma}{|k|^\tau}$$

for $|n| > K$ need not to be assumed since we don't solve the corresponding terms.

As a result, at each KAM iteration step, we only need to deal with finite (about $O(K)$) many non-resonance assumptions. Then the parameter set excluded at each KAM step can be well controlled.

In summary, we use the decay property to the reduce the number of the non-resonance assumptions, especially those might cause problem for measure estimate. The persistence of the decay property of the perturbation along the iteration has to be proved since one has to run KAM iteration infinite many steps.

Further remarks: 1. The restriction $m \notin \frac{1}{12}\mathbb{Z}$ is required when doing the normal form for the general nonlinearity. If the nonlinearity is of the form $f(|u|^2)u$ as considered in Bourgain [5], the restriction is not necessary.

2. Our methods are also valid when applied to the equations of the form (1.2) considered by Bourgain [3], at that time, ξ are outer-parameters instead of amplitudes.

3. Besides the existence of solutions, we also obtain its linear stability and zero Lyapunov exponents, this could be useful for a better understanding of the dynamics.

As for the proof, we try to follow the steps of the previous KAM approach. Meanwhile, we shall always emphasize the decay property which is the only difference between this paper and the previous papers. The rest of the paper is organized as follows: In Section 2 the Hamiltonian function is written in infinitely many coordinates and its regularity is established. In Section 3 it is transformed into Birkhoff normal form of order four so that the transformed Hamiltonian is a small perturbation of some nonlinear integrable system in a sufficiently small neighborhood of the origin. Sections 4–6 dedicate to the proof of Theorem 1. Some technical lemmata are proved in the Appendix.

2. The Hamiltonian

Eq. (1.3) with *Periodic Boundary Conditions* (1.4) is Hamiltonian system

$$u_t = i\frac{\partial H}{\partial \bar{u}}, \qquad H = \int_0^{2\pi} |u_x|^2 + m|u|^2\, dx + \int_0^{2\pi} g(u, \bar{u})\, dx, \tag{2.1}$$

The operator $A = -\partial_{xx} + m$ with *Periodic Boundary Conditions* has an orthonormal basis $\phi_n(x) = \sqrt{\frac{1}{2\pi}}e^{inx}$ and corresponding eigenvalues

$$\mu_n = n^2 + m. \tag{2.2}$$

Let

$$u(x) = \sum_{n\in\mathbb{Z}} q_n \phi_n(x),$$

with finite weighted norm

$$\|q\|_\rho = \sum_{n\in\mathbb{Z}} |q_n| e^{|n|\rho} < \infty,$$

for some $\rho > 0$. Then corresponding Hamiltonian is

$$H = \Lambda + G = \sum_{n\in\mathbb{Z}} \mu_n |q_n|^2 + \int_0^{2\pi} g\left(\sum_{n\in\mathbb{Z}} q_n\phi_n, \sum_{n\in\mathbb{Z}} \bar{q}_n\bar{\phi}_n\right) dx, \tag{2.3}$$

as a result, Eq. (1.3) is equivalent to the following Hamiltonian equations (symplectic structure $i\sum_n dq_n \wedge d\bar{q}_n$)

$$\dot{q}_n = i\frac{\partial H}{\partial \bar{q}_n}, \qquad n \in \mathbb{Z}. \tag{2.4}$$

Next we consider the regularity of the gradient of G. To this end, let ℓ^ρ be the Banach spaces of all bi-infinite, complex valued sequences $q = (\ldots, q_{-1}, q_0, q_1, \ldots)$ with finite weighted norm

$$\|q\|_\rho = \sum_{n\in\mathbb{Z}} |q_n| e^{|n|\rho}.$$

The convolution $q * p$ of two such sequences is defined by $(q * p)_n = \sum_m q_{n-m} p_m$.

Lemma 2.1. *For $\rho > 0$, the space ℓ^ρ is a Banach algebra with respect to convolution of sequences, and*

$$\|q * p\|_\rho \leqslant \|q\|_\rho \|p\|_\rho.$$

Proof.

$$\begin{aligned}\|q * p\|_\rho &= \sum_n |(q * p)_n| e^{|n|\rho} \\ &= \sum_n \left|\sum_m q_{n-m} p_m\right| e^{|n|\rho}\end{aligned}$$

$$\leqslant \sum_{n,m} |q_{n-m}| e^{|n-m|\rho} |p_m| e^{|m|\rho}$$
$$\leqslant \|q\|_\rho \|p\|_\rho.$$

Lemma 2.2. *For $\rho > 0$, the gradient $G_{\bar q}$ is real analytic as a map from some neighborhood of the origin in ℓ^ρ into ℓ^ρ, with*

$$\|G_{\bar q}\|_\rho = O(\|q\|_\rho^3).$$

Proof. We have

$$\frac{\partial G}{\partial \bar q_n} = \int_0^{2\pi} \frac{\partial g(u,\bar u)}{\partial \bar u} \bar\phi_n \, dx, \qquad u = \sum_{n\in\mathbb{Z}} q_n \phi_n.$$

Let q be in ℓ^ρ, then $\|u\|_\rho = \|q\|_\rho$. By the Banach algebra property and the analyticity of g, then

$$\left\| \frac{\partial g(u,\bar u)}{\partial \bar u} \right\|_\rho \leqslant c\|u\|_\rho^3,$$

in a sufficiently small neighborhood of the origin. The components of the gradient $G_{\bar q}$ are its Fourier coefficients, therefore $G_{\bar q}$ belongs to ℓ^ρ with

$$\|G_{\bar q}\|_\rho \leqslant \left\| \frac{\partial g(u,\bar u)}{\partial \bar u} \right\|_\rho \leqslant c\|u\|_\rho^3 \leqslant c\|q\|_\rho^3.$$

The regularity of $G_{\bar q}$ follows from the regularity of its components and its local boundedness ([14], Appendix A). □

In the following Lemma, we prove the coefficients of nonlinearity G have decay properties (similar proofs see Bambusi [1,2]). This fact is crucial for this paper, and we will prove it to be preserved along KAM iteration, consequently, in each KAM step, we essentially handle finite small-denominator conditions.

Lemma 2.3. *For $\rho > 0$,*

$$G(q,\bar q) = \int_0^{2\pi} g\left(\sum_{n\in\mathbb{Z}} q_n\phi_n, \sum_{n\in\mathbb{Z}} \bar q_n \bar\phi_n\right) dx$$
$$\triangleq \sum_{N=4}^{\infty} \sum_{n_1,n_2,\ldots,n_N,\iota_1,\iota_2,\ldots,\iota_N} G^{\iota_1\iota_2\cdots\iota_N}_{n_1 n_2\cdots n_N} q^{\iota_1}_{n_1} q^{\iota_2}_{n_2} \cdots q^{\iota_N}_{n_N},$$

where $\iota = \pm 1$, $q^{+1} \equiv q$, $q^{-1} \equiv \bar{q}$. *then* $G^{\iota_1\iota_2\cdots\iota_N}_{n_1 n_2\cdots n_N}$ *satisfies*

$$|G^{\iota_1\iota_2\cdots\iota_N}_{n_1 n_2\cdots n_N}| \leqslant c e^{-\rho|\iota_1 n_1+\iota_2 n_2+\cdots+\iota_N n_N|}.$$

Proof.

$$\begin{aligned} |G^{\iota_1\iota_2\cdots\iota_N}_{n_1 n_2\cdots n_N}| &= \left| a_{jl}\int_0^{2\pi} \phi^{\iota_1}_{n_1}\phi^{\iota_2}_{n_2}\cdots\phi^{\iota_N}_{n_N}\,dx \right| \\ &\leqslant c e^{-\rho|\iota_1 n_1+\iota_2 n_2+\cdots+\iota_N n_N|}, \end{aligned}$$

where $j+l=N$ and j, l are the number of $+1$, -1 among $\iota_1,\ldots,\iota_N$ respectively. □

For the convenience of notations, we introduce vectors $\vec{n} = (n_1, n_2, \ldots)$, $\vec{\iota} = (\iota_1, \iota_2, \ldots)$ with finitely many non-vanishing components, then if $\iota_j = 0$ for some j, this means that there is no factor $q_{n_j}^{\iota_j}$ in monomial $q_{\vec{n}}^{\vec{\iota}}$. Thus $G(q,\bar{q})$ can be denoted as $G(q,\bar{q}) = \sum_{\vec{n},\vec{\iota}} G^{\vec{\iota}}_{\vec{n}} q^{\vec{\iota}}_{\vec{n}}$ with $|G^{\vec{\iota}}_{\vec{n}}| \leqslant c e^{-\rho|\langle\vec{\iota},\vec{n}\rangle|}$ ($\langle\cdot,\cdot\rangle$ being standard inner product). For this end, we introduce the definition of decay property:

Definition 2.1. If $G(q,\bar{q}) = \sum_{\vec{n},\vec{\iota}} G^{\vec{\iota}}_{\vec{n}} q^{\vec{\iota}}_{\vec{n}}$, and $|G^{\vec{\iota}}_{\vec{n}}| \leqslant c e^{-\rho|\langle\vec{\iota},\vec{n}\rangle|}$, then G is called to have decay property.

When the nonlinearity $g(u,\bar{u})$ has only terms of fourth order, we have

$$\begin{aligned} G &= \int_0^{2\pi} (a_{40}u(x)^4 + a_{31}u(x)^3\bar{u}(x) + a_{22}u(x)^2\bar{u}(x)^2 \\ &\quad + a_{13}u(x)\bar{u}(x)^3 + a_{04}\bar{u}(x)^4)\,dx \\ &= \sum_{\vec{n},\vec{\iota}} G^{\vec{\iota}}_{\vec{n}} q^{\vec{\iota}}_{\vec{n}} \end{aligned} \tag{2.5}$$

with

$$G^{\vec{\iota}}_{\vec{n}} = \int_0^{2\pi} a_{jl}\phi^{\iota_1}_{n_1}\phi^{\iota_2}_{n_2}\phi^{\iota_3}_{n_3}\phi^{\iota_4}_{n_4}\,dx, \tag{2.6}$$

where $j+l=4$, j and l are the number of $+1$ and -1 among $\iota_1, \iota_2, \iota_3, \iota_4$, respectively. It is not difficult to verify that $G^{\vec{\iota}}_{\vec{n}} = 0$ if $\langle\vec{\iota},\vec{n}\rangle \neq 0$. This will play an important role later on. In particular, if $\vec{n} = (n_1, n_1, n_2, n_2)$, $\vec{\iota} = (1, (-1), 1, (-1))$, we have

$$G^{\vec{\iota}}_{\vec{n}} = \frac{a_{22}}{2\pi}. \tag{2.7}$$

From now on we focus our attention on the nonlinearity of fourth order, since higher order terms will not make any difference. Moreover, we take $a_{40}=a_{31}=a_{22}=a_{13}=a_{04}=1$ for convenience.

3. The Birkhoff normal form

Next we transform the Hamiltonian (2.3) into some Birkhoff normal form of order four so that the Hamiltonian (2.3) may serve as a small perturbation of some nonlinear integrable system in a sufficiently small neighborhood of the origin.

Proposition 3.1. *For the Hamiltonian $H=\Lambda+G$ with the nonlinearity* (2.5), *there exists a real analytic, symplectic change of coordinates Γ in some neighborhood of the origin in ℓ^ρ that takes the Hamiltonian* (2.3) *into*

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

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

$$\bar{G}=\sum_{n_1,n_2}\bar{G}_{n_1n_2}|q_{n_1}|^2|q_{n_2}|^2,\qquad |K|=O(\|q\|_\rho^6), \tag{3.1}$$

with uniquely determined coefficients $\bar{G}_{n_1n_2}=\frac{4-3\delta_{n_1n_2}}{2\pi}$. In addition, $K(q,\bar{q})=\sum_{\vec{n},\vec{\imath}}K_{\vec{n}}^{\vec{\imath}}q_{\vec{n}}^{\vec{\imath}}$ has property: $K_{\vec{n}}^{\vec{\imath}}=0$ if $\langle\vec{\imath},\vec{n}\rangle\neq 0$.

Remark. The last property of K in above proposition is a stronger version of *decay property*. We will prove the *decay property* is preserved along KAM iteration. This is crucial in this paper. With *decay property*, we actually consider finite small-divisors at each KAM step, which makes measure estimates available.

Proof. Since for $m\notin\frac{1}{12}\mathbb{Z}$, $\imath_1\mu_{n_1}+\imath_2\mu_{n_2}+\imath_3\mu_{n_3}+\imath_4\mu_{n_4}\neq 0$ except for $\{\imath_1,\imath_2,\imath_3,\imath_4,n_1,n_2,n_3,n_4\}=\{\imath_1,(-\imath_1),\imath_2,(-\imath_2),e,e,f,f\}$. Then this allows to eliminate all terms in $G(q,\bar{q})$ that are not of the form $|q_{n_1}|^2|q_{n_2}|^2$ by symplectic transformation. More precisely, we will find symplectic change $\Gamma=X_F^t|_{t=1}$ given by

$$F=\sum_{\vec{n},\vec{\imath}}F_{\vec{n}}^{\vec{\imath}}q_{\vec{n}}^{\vec{\imath}},\qquad \vec{n}=(n_1,n_2,n_3,n_4),\quad \vec{\imath}=(\imath_1,\imath_2,\imath_3,\imath_4)$$

to satisfy the desired consequence. By second order Taylor formula, we formally obtain

$$\begin{aligned} H \circ \Gamma &= \Lambda + (G + \{\Lambda, F\}) + O(\|q\|_\rho^6) \\ &= \Lambda + \bar{G} + K \end{aligned}$$

with

$$\{\Lambda, F\} = -i \sum_{\vec{n}, \vec{\imath}} (\imath_1 \mu_{n_1} + \imath_2 \mu_{n_2} + \imath_3 \mu_{n_3} + \imath_4 \mu_{n_4}) F_{\vec{n}}^{\vec{\imath}} q_{\vec{n}}^{\vec{\imath}}$$

Let

$$i F_{\vec{n}}^{\vec{\imath}} = \begin{cases} \dfrac{G_{\vec{n}}^{\vec{\imath}}}{\imath_1 \mu_{n_1} + \imath_2 \mu_{n_2} + \imath_3 \mu_{n_3} + \imath_4 \mu_{n_4}} & \text{for} \quad \{\imath_1, \imath_2, \imath_3, \imath_4, n_1, n_2, n_3, n_4\} \\ & \qquad \neq \{\imath_1, (-\imath_1), \imath_2, (-\imath_2), e, e, f, f\}, \\ 0 & \text{otherwise.} \end{cases} \tag{3.2}$$

Thus, we have

$$\bar{G} = \sum_{n_1, n_2} \bar{G}_{n_1 n_2} |q_{n_1}|^2 |q_{n_2}|^2$$

with

$$\bar{G}_{n_1 n_2} = \frac{4 - 3\delta_{n_1 n_2}}{2\pi}.$$

The next is to prove analyticity of the preceding transformation,

$$\begin{aligned} \left| \frac{\partial F}{\partial q_{n_4}^{\imath_4}} \right| &\leqslant \sum_{\imath_1 n_1 + \imath_2 n_2 + \imath_3 n_3 + \imath_4 n_4 = 0} |F_{\vec{n}}^{\vec{\imath}}| |q_{n_1}^{\imath_1} q_{n_2}^{\imath_2} q_{n_3}^{\imath_3}| \\ &\leqslant c \sum_{\imath_1 n_1 + \imath_2 n_2 + \imath_3 n_3 = -\imath_4 n_4} |G_{\vec{n}}^{\vec{\imath}}| |q_{n_1}^{\imath_1} q_{n_2}^{\imath_2} q_{n_3}^{\imath_3}| \\ &\leqslant c \sum_{\imath_1 n_1 + \imath_2 n_2 + \imath_3 n_3 = -\imath_4 n_4} |q_{n_1}^{\imath_1} q_{n_2}^{\imath_2} q_{n_3}^{\imath_3}| \\ &= c(q * q * q)_{n_4}. \end{aligned}$$

Hence by Lemma 2.1

$$\|F_q\|_\rho \leqslant c\|q * q * q\|_\rho \leqslant c\|q\|_\rho^3. \tag{3.3}$$

The analyticity of F_q follows from the analyticity of each component function and its local boundedness ([14], Appendix A).

Next we prove $K = \sum_{\langle \vec{\imath}, \vec{n} \rangle = 0} K_{\vec{n}}^{\vec{\imath}} q_{\vec{n}}^{\vec{\imath}}$ which obviously implies the decay property of K. Since

$$\begin{aligned} K &= \{G, F\} + \frac{1}{2!}\{\{\Lambda, F\}, F\} + \frac{1}{2!}\{\{G, F\}, F\} \\ &\quad + \cdots + \frac{1}{n!}\{\cdots\{\Lambda, \underbrace{F\} \ldots, F}_{n}\} + \frac{1}{n!}\{\cdots\{G, \underbrace{F\} \ldots, F}_{n}\} + \cdots \end{aligned}$$

We first consider $\{G, F\}$, due to

$$\begin{aligned} G &= \sum_{\imath_1 n_1 + \imath_2 n_2 + \imath_3 n_3 + \imath_4 n_4 = 0} G_{n_1 n_2 n_3 n_4}^{\imath_1 \imath_2 \imath_3 \imath_4} q_{n_1}^{\imath_1} q_{n_2}^{\imath_2} q_{n_3}^{\imath_3} q_{n_4}^{\imath_4}, \\ F &= \sum_{\imath_1' m_1 + \imath_2' m_2 + \imath_3' m_3 + \imath_4' m_4 = 0} F_{m_1 m_2 m_3 m_4}^{\imath_1' \imath_2' \imath_3' \imath_4'} q_{m_1}^{\imath_1'} q_{m_2}^{\imath_2'} q_{m_3}^{\imath_3'} q_{m_4}^{\imath_4'}, \end{aligned}$$

then

$$\begin{aligned} \{G, F\} &= i \sum_n \left(\frac{\partial G}{\partial q_n} \frac{\partial F}{\partial \bar{q}_n} - \frac{\partial G}{\partial \bar{q}_n} \frac{\partial F}{\partial q_n} \right) \\ &= i \sum_n \Bigg(\sum_{\imath_2 n_2 + \imath_3 n_3 + \imath_4 n_4 = -n} G_{n n_2 n_3 n_4}^{(+1) \imath_2 \imath_3 \imath_4} q_{n_2}^{\imath_2} q_{n_3}^{\imath_3} q_{n_4}^{\imath_4} \\ &\qquad \times \sum_{\imath_2' m_2 + \imath_3' m_3 + \imath_4' m_4 = n} F_{n m_2 m_3 m_4}^{(-1) \imath_2' \imath_3' \imath_4'} q_{m_2}^{\imath_2'} q_{m_3}^{\imath_3'} q_{m_4}^{\imath_4'} \\ &\qquad - \sum_{\imath_2 n_2 + \imath_3 n_3 + \imath_4 n_4 = n} G_{n n_2 n_3 n_4}^{(-1) \imath_2 \imath_3 \imath_4} q_{n_2}^{\imath_2} q_{n_3}^{\imath_3} q_{n_4}^{\imath_4} \\ &\qquad \times \sum_{\imath_2' m_2 + \imath_3' m_3 + \imath_4' m_4 = -n} F_{n m_2 m_3 m_4}^{(+1) \imath_2' \imath_3' \imath_4'} q_{m_2}^{\imath_2'} q_{m_3}^{\imath_3'} q_{m_4}^{\imath_4'} \Bigg) \\ &= i \sum_{\imath_2 n_2 + \imath_3 n_3 + \imath_4 n_4 + \imath_2' m_2 + \imath_3' m_3 + \imath_4' m_4 = 0} K_{n_2 n_3 n_4 m_2 m_3 m_4}^{\imath_2 \imath_3 \imath_4 \imath_2' \imath_3' \imath_4'} q_{n_2}^{\imath_2} q_{n_3}^{\imath_3} q_{n_4}^{\imath_4} q_{m_2}^{\imath_2'} q_{m_3}^{\imath_3'} q_{m_4}^{\imath_4'} \end{aligned} \tag{3.4}$$

i.e., $K^{\iota_2\iota_3\iota_4\iota'_2\iota'_3\iota'_4}_{n_2n_3n_4m_2m_3m_4} = 0$ if $\iota_2 n_2 + \iota_3 n_3 + \iota_4 n_4 + \iota'_2 m_2 + \iota'_3 m_3 + \iota'_4 m_4 \neq 0$. Analogously, $\frac{1}{n!}\{\cdots\{\Lambda, \underbrace{F\}\ldots, F\}}_{n}$ and $\frac{1}{n!}\{\cdots\{G, \underbrace{F\}\ldots, F\}}_{n}$ have also this property, therefore, K has also this property. □

Note that the definition (3.2) is reasonable in view of the following lemma.

Lemma 3.1. *If $\vec{n} = (n_1, n_2, n_3, n_4)$, $\vec{\iota} = (\iota_1, \iota_2, \iota_3, \iota_4)$, such that $\langle\vec{\iota}, \vec{n}\rangle = 0$, but $\{\iota_1, \iota_2, \iota_3, \iota_4, n_1, n_2, n_3, n_4\} \neq \{\iota_1, (-\iota_1), \iota_2, (-\iota_2), e, e, f, f\}$, then for m lying in the compact set Π of $\{x | x \in \mathbb{R}$ and $x \neq \frac{n}{12}, n \in \mathbb{Z}\}$, one has*

$$|\iota_1\mu_{n_1} + \iota_2\mu_{n_2} + \iota_3\mu_{n_3} + \iota_4\mu_{n_4}| \geqslant c,$$

with c dependent only on m. Hence the denominators in (3.2) are uniformly bounded away from zero.

Proof. We have to consider divisors of the form $\delta = \pm\mu_{n_1} \pm \mu_{n_2} \pm \mu_{n_3} \pm \mu_{n_4}$ for all possible combinations of plus and minus signs. To this end, we distinguish them according to their number of minus signs. To shorten notation we let for example $\delta_{+-++} = \mu_{n_1} - \mu_{n_2} + \mu_{n_3} + \mu_{n_4}$. Similarly, for all other combinations of plus and minus signs.

Case 0: *No minus sign.*

$$|\delta_{++++}| = |n_1^2 + n_2^2 + n_3^2 + n_4^2 + 4m|,$$

hence if $m \in \Pi$, then $|\delta_{++++}| \geqslant c$.

Case 1: *One minus sign.*

$$|\delta_{-+++}| = |-n_1^2 + n_2^2 + n_3^2 + n_4^2 + 2m|,$$

hence if $m \in \Pi$, then $|\delta_{-+++}| \geqslant c$. Analogously, if $m \in \Pi$, $|\delta_{+-++}| \geqslant c$, $|\delta_{++-+}| \geqslant c$, $|\delta_{+++-}| \geqslant c$.

Case 2: *Two minus signs*. Owing to assumption $\iota_1 n_1 + \iota_2 n_2 + \iota_3 n_3 + \iota_4 n_4 = 0$, there are two "+1" and two "−1" among $\iota_1, \iota_2, \iota_3, \iota_4$; and $\{n_1, n_2, n_3, n_4\} \neq \{e, e, f, f\}$, without loss of generality, we assume $\iota_1 = \iota_2 = 1$ and $\iota_3 = \iota_4 = -1$, then

$$|\delta_{++--}| = |n_1^2 + n_2^2 - n_3^2 - n_4^2| = 2|(n_2 - n_3)(n_2 - n_4)| \geqslant c.$$

Analogously, $|\delta_{--++}| \geqslant c$, $|\delta_{-+-+}| \geqslant c$, $|\delta_{-++-}| \geqslant c$, $|\delta_{+--+}| \geqslant c$, $|\delta_{+-+-}| \geqslant c$.

Cases 3 *and* 4: *Three and four minus signs*. These ones reduce to *Case* 1 and *Case* 0, respectively. □

Now our Hamiltonian is

$$H = \Lambda + \bar{G} + K = \sum_n \mu_n |q_n|^2 + \sum_{n_1,n_2} \bar{G}_{n_1 n_2} |q_{n_1}|^2 |q_{n_2}|^2 + O(\|q\|_\rho^6),$$

where

$$\bar{G}_{n_1 n_2} = \frac{4 - 3\delta_{n_1 n_2}}{2\pi}.$$

We choose distinguished finite number d of modes $\phi_{n_1}, \ldots, \phi_{n_d}$ satisfying $|n_i| \neq |n_j|$ for $i \neq j$. Introduce the standard action-angle variables $(I, \theta) \in \mathbb{R}^d \times \mathbb{T}^d$ and linearize H around a given value for the action variable, namely, for some $\xi = (\xi_1, \ldots, \xi_d) \in \mathbb{R}^d_+$,

$$q_{n_i} \bar{q}_{n_i} = I_{n_i} + \xi_i, \quad i = 1, \ldots, d$$

and set $q_n = w_n$ for $n \notin \{n_1, \ldots, n_d\}$. One finally obtains

$$H = \langle \omega(\xi), I \rangle + \sum_n \Omega_n(\xi) w_n \bar{w}_n + P(\theta, I, w, \bar{w}, \xi), \tag{3.5}$$

where $\omega(\xi) = (\omega_1(\xi), \ldots, \omega_d(\xi))$, $\Omega(\xi) = (\Omega_n(\xi))_{n \neq n_1,\ldots,n_d}$ are given by

$$\omega(\xi) = \alpha + A\xi, \tag{3.6}$$

$$\Omega(\xi) = \beta + B\xi \tag{3.7}$$

and $\alpha = (\mu_{n_1}, \ldots, \mu_{n_d})$, $\beta = (\mu_n)_{n \neq n_1,\ldots,n_d}$, $A = (\bar{G}_{n_i n_j})_{1 \leqslant i,j \leqslant d}$, $B = (\bar{G}_{n n_i})_{\substack{1 \leqslant i \leqslant d \\ n \neq n_1,\ldots,n_d}}$. P is just $K + O(|I|^2) + O(|I| \|w\|_\rho^2) + O(\|w\|_\rho^4)$ with the variables $q_i, \bar{q}_i, i = 1, \ldots, d$ expressed in terms of I, θ. In addition, P is real analytic in θ, real analytic in $I, w, \bar{w}$ in a sufficiently small neighborhood of the origin, and C^4 in ξ lying on the compact set of $\mathbb{R}^d_+$ in the sense of Whitney. What's more, due to Proposition 3.1, we obtain $K = \sum_{\langle \vec{\imath}, \vec{n} \rangle = 0} K_{\vec{n}}^{\vec{\imath}} q_{\vec{n}}^{\vec{\imath}}$. After the above action-angle transformation, $P = \sum_{\vec{n}', \vec{\imath}'} P_{\vec{n}'}^{\vec{\imath}'}(\theta, I) w_{\vec{n}'}^{\vec{\imath}'}$. Note that each non-vanishing term comes from $K_{\vec{n}}^{\vec{\imath}} q_{n_1}^{k_1} \bar{q}_{n_1}^{\bar{k}_1} \cdots q_{n_d}^{k_d} \bar{q}_{n_d}^{\bar{k}_d} w_{\vec{n}'}^{\vec{\imath}'}$ with $\langle \vec{\imath}', \vec{n}' \rangle = -((k_1 - \bar{k}_1) n_1 + \cdots + (k_d - \bar{k}_d) n_d)$. Thus we has still the decay estimates,[3]

$$\begin{aligned} |P_{\vec{n}'}^{\vec{\imath}'}(\theta, I)| &= |K_{\vec{n}}^{\vec{\imath}} q_{n_1}^{k_1} \bar{q}_{n_1}^{\bar{k}_1} \cdots q_{n_d}^{k_d} \bar{q}_{n_d}^{\bar{k}_d}| \leqslant c |q_{n_1}|^{k_1} |\bar{q}_{n_1}|^{\bar{k}_1} \cdots |q_{n_d}|^{k_d} |\bar{q}_{n_d}|^{\bar{k}_d} \\ &\leqslant ce^{-\rho(|k_1 n_1| + |\bar{k}_1 n_1| + \cdots + |k_d n_d| + |\bar{k}_d n_d|)} \leqslant ce^{-\rho|(k_1 - \bar{k}_1) n_1 + \cdots + (k_d - \bar{k}_d) n_d|} \\ &= ce^{-\rho |\langle \vec{\imath}', \vec{n}' \rangle|}, \end{aligned} \tag{3.8}$$

[3] Later, we will see $|q_{n_i}|, |\bar{q}_{n_i}| \sim e^{-|n_i|\rho}$, $1 \leqslant i \leqslant d$.

where $\vec{n}' = (n'_1, n'_2, \ldots)$, $\vec{\imath}' = (\imath'_1, \imath'_2, \ldots)$ with finitely many non-vanishing components.

To this end, we proved the decay property of the perturbation P. Below we will find that decay estimates (3.8) can be preserved during the KAM iteration, so that smaller terms may be delayed to the following KAM steps by shrinking the weight of vector fields, as a consequence, in each KAM step, we practically dispose of finite small-divisors. This is the motivation we can settle Schrödinger equation with *Periodic Boundary Conditions*.

Scaling ξ by ε^6, $w, \bar{w}$ by ε^4, and I by ε^8, one obtains a Hamiltonian system given by the rescaled Hamiltonian

$$\begin{aligned}\tilde{H}(I,\theta,w,\bar{w},\xi) &= \varepsilon^{-14}H(\varepsilon^8 I,\theta,\varepsilon^4 w,\varepsilon^4\bar{w},\varepsilon^6\xi)\\ &= \langle\tilde{\omega}(\xi), I\rangle + \sum_n \tilde{\Omega}_n(\xi)w_n\bar{w}_n + \tilde{P}(I,\theta,w,\bar{w},\xi,\varepsilon), \qquad (3.9)\end{aligned}$$

where

$$\tilde{\omega}(\xi) = \varepsilon^{-6}\alpha + A\xi, \quad \tilde{\Omega}_n(\xi) = \varepsilon^{-6}\beta + B\xi, \tag{3.10}$$

and

$$\|X_{\tilde{P}}\|_\rho \leqslant \varepsilon. \tag{3.11}$$

For simplicity, we still denote $\tilde{H}$ by H, $\tilde{\omega}$ by ω, $\tilde{\Omega}$ by Ω and $\tilde{P}$ by P.

Below we check non-resonance conditions:

$$\langle k, \omega(\xi)\rangle \not\equiv 0, \qquad k \neq 0,$$

$$\langle k, \omega(\xi)\rangle + \Omega_n(\xi) \not\equiv 0,$$

$$\langle k, \omega(\xi)\rangle + \Omega_n(\xi) + \Omega_m(\xi) \not\equiv 0,$$

$$\langle k, \omega(\xi)\rangle + \Omega_n(\xi) - \Omega_m(\xi) \not\equiv 0, \qquad |k| + ||n| - m|| \neq 0.$$

Since

$$A = (\bar{G}_{ij})_{1\leqslant i,j\leqslant d} = \frac{1}{2\pi}\begin{pmatrix} 1 & 4 & \cdots & 4\\ 4 & 1 & \cdots & 4\\ \vdots & \vdots & \ddots & \vdots\\ 4 & 4 & \cdots & 1\end{pmatrix}$$

then

$$\det(A) = \frac{(-3)^{d-1}(4d-3)}{(2\pi)^d} \neq 0.$$

Hence $\langle k, \omega(\xi)\rangle \not\equiv 0$ for $k \neq 0$.

For the last three non-resonance conditions, one has to check that

$$\langle \alpha, k\rangle + \langle \beta, l\rangle \neq 0 \text{ or } Ak + B^T l \neq 0$$

for $1 \leqslant |l| \leqslant 2$. Suppose $Ak + B^T l = 0$. Multiplying by 2π we have

$$2\pi Ak + 2\pi B^T l = 0,$$

and all coefficients of $2\pi B^T$ are 4. Thus all components of k are equal, say p, and $(4d-3)p + 4q = 0$, where q is the sum of at most two nonzero components of l. Then integer solutions to this equation are

$$\begin{aligned} &\text{(i)} && q = \pm 1,\ p = \mp 4,\ d = 1;\\ &\text{(ii)} && q = \pm 2,\ p = \mp 8,\ d = 1;\\ &\text{(iii)} && q = p = 0. \end{aligned}$$

then in (i), $\mp 4\mu_{n_i} \pm \mu_n = \mp 4n_i^2 \pm n^2 \mp 3m \neq 0$. In (ii), $\mp 8\mu_{n_i} \pm (\mu_n + \mu_m) \neq 0$. In (iii), So $k = 0$ and l has one "1" and one "-1", but then $\langle \alpha, k\rangle + \langle \beta, l\rangle = \mu_n - \mu_m = n^2 - m^2$ for some $|n| \neq m$, this expression does not vanish. Consequently non-resonance conditions are satisfied.

In the following, we will use KAM iteration which involves infinite many steps of coordinate transformation to eliminate lower order θ-dependent terms in P. To make this quantitative we introduce

$$D_\rho(r,s) = \{(\theta, I, w, \bar{w}) : |\mathrm{Im}\,\theta| < r, |I| < s^2, \|w\|_\rho < s, \|\bar{w}\|_\rho < s\},$$

a complex neighborhood of $\mathbb{T}^d \times \{I = 0\} \times \{w = 0\} \times \{\bar{w} = 0\}$. Where $|\cdot|$ denotes the sup-norm for complex vectors, and $\|w\|_\rho, \|\bar{w}\|_\rho$ denote a family of finite weighted norm

$$\|w\|_\rho = \sum_{\substack{n\in\mathbb{Z}\\ n\neq n_1,\ldots,n_d}} |w_n| e^{|n|\rho},$$

$$\|\bar{w}\|_\rho = \sum_{\substack{n\in\mathbb{Z}\\ n\neq n_1,\ldots,n_d}} |\bar{w}_n| e^{|n|\rho}.$$

For any given function

$$F=\sum_{\alpha,\beta}F_{\alpha\beta}w^{\alpha}\bar{w}^{\beta}, \tag{3.12}$$

where the multi-index α, β run over the set $\alpha\equiv(\ldots,\alpha_n,\ldots)$, $\beta\equiv(\ldots,\beta_n,\ldots)$, $\alpha_n,\beta_n\in\mathbb{N}$ with finitely many non-vanishing components, we define its finite weighted norm

$$\|F\|_{D_\rho(r,s),\mathcal{O}}\equiv\sup_{\substack{\|w\|_\rho<s\\ \|\bar{w}\|_\rho<s}}\sum_{\alpha,\beta}\|F_{\alpha\beta}\|\,|w^{\alpha}||\bar{w}^{\beta}|. \tag{3.13}$$

Where, if $F_{\alpha\beta}=\sum_{k\in\mathbb{Z}^d,I\in\mathbb{N}^d}F_{kl\alpha\beta}(\xi)I^le^{i\langle k,\theta\rangle}$, ($\langle\cdot,\cdot\rangle$ being the standard inner product in $\mathbb{C}^d$), $\|F_{\alpha\beta}\|$ is short for

$$\|F_{\alpha\beta}\|\equiv\sum_{k,l}|F_{kl\alpha\beta}|_{\mathcal{O}}s^{2|l|}e^{|k|r},\quad |F_{kl\alpha\beta}|_{\mathcal{O}}\equiv\sup_{\xi\in\mathcal{O}}\max_{|p|\leqslant 4}\left(\left|\frac{\partial^pF_{kl\alpha\beta}}{\partial\xi^p}\right|\right). \tag{3.14}$$

(the derivatives with respect to ξ are in the sense of Whitney).

To function F, we associate a Hamiltonian vector field defined as

$$X_F=(F_I,-F_\theta,\{iF_{w_n}\}_{n\neq n_1,\ldots,n_d},\{-iF_{\bar{w}_n}\}_{n\neq n_1,\ldots,n_d}).$$

Define its weighted norm by [4]

$$\begin{aligned}\|X_F\|_{D_\rho(r,s),\mathcal{O}}\equiv\ &\|F_I\|_{D_\rho(r,s),\mathcal{O}}+\frac{1}{s^2}\|F_\theta\|_{D_\rho(r,s),\mathcal{O}}\\ &+\frac{1}{s}\left(\sum_{n\neq n_1,\ldots,n_d}\|F_{w_n}\|_{D_\rho(r,s),\mathcal{O}}e^{|n|\rho}+\sum_{n\neq n_1,\ldots,n_d}\|F_{\bar{w}_n}\|_{D_\rho(r,s),\mathcal{O}}e^{|n|\rho}\right)\end{aligned} \tag{3.15}$$

Having the above preparation, below we will outline the mechanism of KAM iteration. At the first step, the unperturbed Hamiltonian will be a family of integrable

[4] The norm $\|\cdot\|_{D_\rho(r,s),\mathcal{O}}$ for scalar functions is defined in (3.13). For vector (or matrix-valued) functions $G:D_\rho(r,s)\times\mathcal{O}\to\mathbb{C}^m$, ($m<\infty$) is similarly defined as $\|G\|_{D_\rho(r,s),\mathcal{O}}=\sum_{i=1}^m\|G_i\|_{D_\rho(r,s),\mathcal{O}}$ (for the matrix-valued case the sum will run over all entries).

Hamiltonian of the form

$$N = \langle \omega(\xi), I\rangle + \sum_n \Omega_n(\xi) w_n \bar{w}_n,$$

where all the variables $w_n, \bar{w}_n$ are separated. Due to the multiple normal frequencies, the terms $P^{(1,(-1))}_{(n,(-n))} w_n \bar{w}_{-n}$ in perturbation cannot be solved in the homological equations so that they have to be moved to the normal form part if they are not small enough. The decay property of the perturbation P (see Lemma 2.3, Proposition 3.1 and (3.8)) implies that the coefficient $|P^{(1,(-1))}_{(n,(-n))}| \leqslant c\varepsilon e^{-2|n|r}$ with $\rho \leqslant r$. It follows that the smaller-weighted norm of the vector field of $\sum_{2|n|>K} P^{(1,(-1))}_{(n,(-n))} w_n \bar{w}_{-n}$ is of size $\varepsilon_+ \ll \varepsilon$ (here $K \sim -\log \varepsilon$), so that it can be postponed to the next KAM step. But the terms $P^{(1,(-1))}_{(n,(-n))} w_n \bar{w}_{-n}$ with $2|n| \leqslant K$ has to be moved to the normal form part since it is not small enough. The normal form part thus takes the following more complicated form

$$N = \langle \omega(\xi), I\rangle + \sum_n \Omega_n(\xi) w_n \bar{w}_n + \sum_n P^{(1,(-1))}_{(n,n)} w_n \bar{w}_n + \sum_{2|n|\leqslant K} P^{(1,(-1))}_{(n,(-n))} w_n \bar{w}_{-n} \tag{3.16}$$

Therefore, we couple the variables w_n, w_{-n} corresponding to multiple eigenvalues, More precisely, we let $z_n = (w_n, w_{-n})$, $\bar{z}_n = (\bar{w}_n, \bar{w}_{-n})$, $n \geqslant 0$. Correspondingly, the Hamiltonian with normal form (3.16) after one KAM step, takes the form

$$H = \langle \omega(\xi), I\rangle + \sum_{n\geqslant 0} \langle A_n z_n, \bar{z}_n\rangle + P(\theta, I, z, \bar{z}, \xi, \varepsilon), \tag{3.17}$$

where, $\langle \cdot, \cdot \rangle$ is standard inner product,

$$A_n = \begin{pmatrix} \Omega_n + P^{(1,(-1))}_{(n,n)} & P^{(1,(-1))}_{((-n),n)} \\ P^{(1,(-1))}_{(n,(-n))} & \Omega_{-n} + P^{(1,(-1))}_{((-n),(-n))} \end{pmatrix} \tag{3.18}$$

$$\stackrel{\Delta}{=} \begin{pmatrix} \tilde{\Omega}_n & b_{((-n),n)} \\ b_{(n,(-n))} & \tilde{\Omega}_{-n} \end{pmatrix}, \text{ for } 2n \leqslant K,$$

$$A_n = \begin{pmatrix} \Omega_n + P^{(1,(-1))}_{(n,n)} & 0 \\ 0 & \Omega_{-n} + P^{(1,(-1))}_{((-n),(-n))} \end{pmatrix}$$

$$\stackrel{\Delta}{=} \begin{pmatrix} \tilde{\Omega}_n & 0 \\ 0 & \tilde{\Omega}_{-n} \end{pmatrix}, \text{ for } 2n > K$$

and

$$|P^{(1,(-1))}_{(n,n)}|_{\mathcal{O}},\, |P^{(1,(-1))}_{((-n),(-n))}|_{\mathcal{O}} \leqslant \varepsilon, \tag{3.19}$$

$$|P^{(1,(-1))}_{((-n),n)}|_{\mathcal{O}} = |b_{((-n),n)}|_{\mathcal{O}},\, |P^{(1,(-1))}_{(n,(-n))}|_{\mathcal{O}} = |b_{(n,(-n))}|_{\mathcal{O}} \leqslant \varepsilon e^{-2|n|r}.$$

Remark. Different from wave equation and beam equation where $|P^{(1,(-1))}_{(n,n)}|_{\mathcal{O}}$, $|P^{(1,(-1))}_{((-n),(-n))}|_{\mathcal{O}} = O(\varepsilon|n|^{-1})$ due to the regularity, Schrödinger equation has only the smallness of the drift (3.19). This might cause trouble for the measure estimates. To overcome this difficulty, we prove that the perturbation from PDEs has a kind of decay property (3.8) persisting along the KAM iteration. Especially, the coefficient of $w_n\bar{w}_{-n}$ is of the size $\varepsilon e^{-2|n|\rho}$, which means that in this step, the terms of the form $\sum_{2n>K} P^{(1,(-1))}_{(n,(-n))}(\theta)w_n\bar{w}_{-n}$ need not to be solved since it is small enough to be postponed to the next KAM step. Due to this fact, all of A_n with $2n > K$ are diagonal matrices. Consequently, in this step, small-divisor $\langle k,\omega\rangle + \Omega_n - \Omega_{-n}$ when $2n > K$ need not to be controlled. As a result, we essentially deal with finite small-denominator conditions. But in the next step, there will be some terms of $\sum_{2|n|>K} P^{(1,(-1))}_{(n,(-n))}(\theta)w_n\bar{w}_{-n}$ to be handled, hence, some of A_n with $2n > K$ will become non-diagonal matrices of the form (3.18), so eventually, we will handle all small-denominator conditions. This fact is very important for the measure estimate.

4. KAM step

Theorem 1 will be proved by a KAM iteration which involves an infinite sequence of change of variables. Each step of KAM iteration makes the perturbation smaller than the previous step at the cost of excluding a small set of parameters. What we need to worry about is the convergence and the measure of the excluding set after infinite KAM steps.

At each step of the KAM scheme, we consider a Hamiltonian vector field with

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

where N_ν is an "integrable normal form" and P_ν with decay property is defined in $D_\rho(r_\nu, s_\nu) \times \mathcal{O}_{\nu-1}$ for any fixed $\frac{1}{2}r_0 < \rho \leqslant r_\nu$.

We then construct a map

$$\Phi_\nu : D_\rho(r_{\nu+1}, s_{\nu+1}) \times \mathcal{O}_\nu \to D_\rho(r_\nu, s_\nu) \times \mathcal{O}_{\nu-1}$$

for any $\frac{1}{2}r_0 < \rho \leqslant r_{\nu+1}$, so that the vector field $X_{H_\nu\circ\Phi_\nu}$ defined on $D_\rho(r_{\nu+1}, s_{\nu+1})$ with $\frac{1}{2}r_0 < \rho \leqslant r_{\nu+1}$ satisfies

$$\|X_{H_\nu\circ\Phi_\nu} - X_{N_{\nu+1}}\|_{D_\rho(r_{\nu+1},s_{\nu+1}),\mathcal{O}_\nu} \leqslant \varepsilon_\nu^\kappa$$

with some new normal form $N_{\nu+1}$ for some fixed ν-independent constant $\kappa > 1$.

To simplify notations, in what follows, the quantities without subscripts refer to quantities at the vth step, while the quantities with subscripts $+$ denote the corresponding quantities at the $(v+1)$th step. Let us then consider the Hamiltonian

$$H = N + P \equiv e + \langle \omega, I\rangle + \sum_{n\geqslant 0}\langle A_n z_n, \bar{z}_n\rangle + P(\theta, I, z, \bar{z}, \xi, \varepsilon),$$

defined in $D_\rho(r,s)\times\mathcal{O}$. We assume that $\xi\in\mathcal{O}$ satisfies[5] (for a suitable $\tau>0$ to be specified later)

$$|\langle k,\omega\rangle^{-1}|\leqslant\frac{K^\tau}{\gamma},\ \|(\langle k,\omega\rangle I + A_n)^{-1}\|\leqslant\left(\frac{K^\tau}{\gamma}\right)^2,$$

$$\|(\langle k,\omega\rangle I \pm A_n\otimes I \pm I\otimes A_m)^{-1}\|\leqslant\left(\frac{K^\tau}{\gamma}\right)^4,\quad n\neq m,$$

$$\|(\langle k,\omega\rangle I + A_n\otimes I + I\otimes A_n)^{-1}\|\leqslant\left(\frac{K^\tau}{\gamma}\right)^4, \tag{4.1}$$

$$|(\langle k,\omega\rangle I + A_n\otimes I - I\otimes A_n)^{-1}\|\leqslant\left(\frac{K^\tau}{\gamma}\right)^4,\quad 2n\leqslant K,$$

for $0<|k|\leqslant K$, and (See (3.18) for the notation $b_{((-n),n)}, b_{(n,(-n))}$)

$$\sup_{\xi\in\mathcal{O}}\max_{0<|p|\leqslant 4}\left\|\frac{\partial^p A_n}{\partial\xi^p}\right\|\leqslant L, \tag{4.2}$$

$$|b_{((-n),n)}|_{\mathcal{O}}, |b_{(n,(-n))}|_{\mathcal{O}}\leqslant\varepsilon_0 e^{-2|n|\rho}, \tag{4.3}$$

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

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

$\|\cdot\|$ for matrix denotes the operator norm, i.e., $\|M\| = \sup_{|y|=1}|My|$. Recall that ω and A_n depend on ξ.

moreover,

$$\|X_P\|_{D_\rho(r,s),\mathcal{O}} \leqslant \varepsilon \tag{4.4}$$

and $P = \sum_{\vec{n},\vec{l}} P_{\vec{n}}^{\vec{l}} w_{\vec{n}}^{\vec{l}}$ satisfies decay estimates,

$$||P_{\vec{n}}^{\vec{l}}||_{\mathcal{O}} < \varepsilon e^{-|\langle \vec{l},\vec{n}\rangle|\rho}, \tag{4.5}$$

for any ρ with $\frac{1}{2}r_0 < \rho \leqslant r$. We now let $0 < r_+ < r$ and define

$$s_+ = \frac{1}{8} s\varepsilon^{\frac{1}{3}}, \quad K^d e^{-K(r-r_+)} = \varepsilon^{\frac{1}{6}}, \quad \gamma = \varepsilon_0^{\frac{1}{50}}, \quad \varepsilon_+ = c\gamma^{-8}(r-r_+)^{-c}\varepsilon^{\frac{7}{6}}. \tag{4.6}$$

Here and later, the letter c denotes suitable (possibly different) constants that do not depend on the iteration step.

We now describe how to construct a set $\mathcal{O}_+ \subset \mathcal{O}$ and a change of variables $\Phi : D_+ \times \mathcal{O}_+ = D_\rho(r_+, s_+) \times \mathcal{O}_+ \to D_\rho(r, s) \times \mathcal{O}$, with $\frac{1}{2}r_0 < \rho \leqslant r_+$, such that the transformed Hamiltonian $H_+ = N_+ + P_+ \equiv H \circ \Phi$ satisfies all the above iterative assumptions with new parameters $s_+, \varepsilon_+, r_+, L_+$ and with $\xi \in \mathcal{O}_+$.

4.1. Solving the linearized equations

Expand P into the Fourier–Taylor series

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

where $k \in \mathbb{Z}^d, l \in \mathbb{N}^d$ and the multi-index α, β run over the set $\alpha \equiv (\ldots, \alpha_n, \ldots)$, $\beta \equiv (\ldots, \beta_n, \ldots)$, $\alpha_n, \beta_n \in \mathbb{N}$ with finitely many non-vanishing components.

Let R be the truncation of P given by

$$\begin{aligned} R(\theta, I, w, \bar{w}) = & \sum_{|k|\leqslant K, |l|\leqslant 1} P_{kl00} e^{i\langle k,\theta\rangle} I^l \\ & + \sum_{|k|\leqslant K, |n|\leqslant K} (P_n^{k10} w_n + P_n^{k01} \bar{w}_n) e^{i\langle k,\theta\rangle} \\ & + \sum_{|k|\leqslant K, |n+m|\leqslant K} (P_{nm}^{k20} w_n w_m + P_{nm}^{k02} \bar{w}_n \bar{w}_m) e^{i\langle k,\theta\rangle} \\ & + \sum_{|k|\leqslant K, |n-m|\leqslant K} P_{nm}^{k11} w_n \bar{w}_m e^{i\langle k,\theta\rangle} \end{aligned} \tag{4.7}$$

where $P_n^{k10} = P_{kl\alpha\beta}$ with $\alpha = e_n, \beta = 0$; $P_n^{k01} = P_{kl\alpha\beta}$ with $\alpha = 0, \beta = e_n$; $P_{nm}^{k20} = P_{kl\alpha\beta}$ with $\alpha = e_n + e_m, \beta = 0$; $P_{nm}^{k11} = P_{kl\alpha\beta}$ with $\alpha = e_n, \beta = e_m$; $P_{nm}^{k02} = P_{kl\alpha\beta}$ with $\alpha = 0, \beta = e_n + e_m$.

Remark. Due to the decay property, the terms $\sum_{|k|\leqslant K, 2n>K} P_{n(-n)}^{k11} w_n \bar{w}_{-n} e^{i\langle k,\theta\rangle}$ is small enough to be delayed to the next KAM step. This fact makes it possible to save infinite many non-resonance assumptions $|\langle k, \omega(\xi)\rangle + \lambda_n(\xi) - \lambda_m(\xi)| \geqslant \frac{\gamma}{|k|^\tau}$, $|n| > K$, which is different from the previous KAM approach. As a result, we essentially deal with finite small-divisors.

In order to have a compact formulation when solving homological equations, we rewrite R in matrix form. Let $z_n = (w_n, w_{-n})$, $\bar{z}_n = (\bar{w}_n, \bar{w}_{-n})$, $n \geqslant 0$. R can be re-written as follows:

$$\begin{aligned} R(\theta, I, z, \bar{z}) \;=\; R_0 + R_1 + R_2 = & \sum_{|k|\leqslant K, |l|\leqslant 1} P_{kl00} e^{i\langle k,\theta\rangle}\, I^l \\ & + \sum_{|k|\leqslant K, n} (\langle R_n^{k10}, z_n\rangle + \langle R_n^{k01}, \bar{z}_n\rangle) e^{i\langle k,\theta\rangle} \\ & + \sum_{|k|\leqslant K, n, m} (\langle R_{mn}^{k20} z_n, z_m\rangle + \langle R_{mn}^{k02} \bar{z}_n, \bar{z}_m\rangle) e^{i\langle k,\theta\rangle} \\ & + \sum_{|k|\leqslant K, n, m} \langle R_{mn}^{k11} z_n, \bar{z}_m\rangle e^{i\langle k,\theta\rangle}\ , \end{aligned} \tag{4.8}$$

where, R_n^{k10}, R_n^{k01}, R_{mn}^{k20}, R_{mn}^{k02} and R_{mn}^{k11} are, respectively,

$$R_n^{k10} = \begin{pmatrix} P_n^{k10} \\ P_{-n}^{k10} \end{pmatrix}, \quad R_n^{k01} = \begin{pmatrix} P_n^{k01} \\ P_{-n}^{k01} \end{pmatrix}, \quad n \leqslant K,$$

$$R_{mn}^{k20} = \begin{pmatrix} P_{nm}^{k20} & P_{(-n)m}^{k20} \\ P_{n(-m)}^{k20} & P_{(-n)(-m)}^{k20} \end{pmatrix}, \quad n + m \leqslant K,$$

$$R_{mn}^{k20} = \begin{pmatrix} 0 & P_{(-n)m}^{k20} \\ P_{n(-m)}^{k20} & 0 \end{pmatrix}, \quad n + m > K, |n - m| \leqslant K,$$

$$R_{mn}^{k02} = \begin{pmatrix} P_{nm}^{k02} & P_{(-n)m}^{k02} \\ P_{n(-m)}^{k02} & P_{(-n)(-m)}^{k02} \end{pmatrix} \quad n + m \leqslant K,$$

$$R^{k02}_{mn} = \begin{pmatrix} 0 & P^{k02}_{(-n)m} \\ P^{k02}_{n(-m)} & 0 \end{pmatrix} \quad n+m > K, |n-m| \leqslant K,$$

$$R^{k11}_{mn} = \begin{pmatrix} P^{k11}_{nm} & P^{k11}_{(-n)m} \\ P^{k11}_{n(-m)} & P^{k11}_{(-n)(-m)} \end{pmatrix} \qquad n+m \leqslant K, \tag{4.9}$$

$$R^{k11}_{mn} = \begin{pmatrix} P^{k11}_{nm} & 0 \\ 0 & P^{k11}_{(-n)(-m)} \end{pmatrix} \qquad n+m > K, |n-m| \leqslant K.$$

Note that all of R^{k11}_{nn} with $2n > K$ are diagonal matrices, as a result, we need not small-divisor conditions $(\langle k, \omega\rangle I + A_n \otimes I - I \otimes A_n)$, $2n > K$, which is different from the previous KAM machinery, consequently, Schrödinger equation with periodic boundary conditions can be handled by this technique. In addition, we have $R^{k20}_{nm} = (R^{k20}_{mn})^T$, $R^{k11}_{nm} = (R^{k11}_{mn})^T$ and $R^{k02}_{nm} = (R^{k02}_{mn})^T$.

Rewrite H as $H = N + R + (P - R)$. By the choice of s_+ in (4.6) and by the definition of the norms, it follows immediately that

$$\|X_R\|_{D_\rho(r,s),\mathcal{O}} \leqslant \|X_P\|_{D_\rho(r,s),\mathcal{O}} \leqslant \varepsilon. \tag{4.10}$$

for any $\frac{1}{2} r_0 < \rho \leqslant r$. In the next, we prove that, $\frac{1}{2} r_0 < \rho \leqslant r_+$,

$$\|X_{(P-R)}\|_{D_\rho(r,s_+)} \; < \; \mathrm{c}\, \varepsilon_+. \tag{4.11}$$

In fact, due to (4.7), $P - R = P^* + \text{h.o.t.}$, where $P^* = \sum_{n\in\mathbb{Z}, n>K} P^{\imath_1}_n(\theta) w^{\imath_1}_n + \sum_{|\imath_2 n + \imath_3 m| > K} P^{(\imath_2,\imath_3)}_{(n,m)}(\theta) w^{\imath_2}_n w^{\imath_3}_m$ be the linear and quadratic terms in the perturbation. By virtue of (4.5), (4.6), $\|X_P\|_{D_\rho(r,s),\mathcal{O}} \leqslant \varepsilon$, and Cauchy estimates, one has that

$$\begin{aligned} \|X_{P^*}\|_{D_\rho(r_+,s),\mathcal{O}} \;&\leqslant\; (r-r_+)^{-1} \sum_{|n|>K} \varepsilon e^{-|n|r} e^{|n|\rho} + (r-r_+)^{-1} \\ &\quad \times \sum_{|\imath_2 n + \imath_3 m|>K} \varepsilon e^{-|\imath_2 n + \imath_3 m| r} |w^{\imath_3}_m| e^{|n|\rho} \\ &\leqslant\; (r-r_+)^{-1} \Bigg(\sum_{|n|>K} \varepsilon e^{-|n|r} e^{|n|\rho} \\ &\quad + \sum_{|\imath_2 n + \imath_3 m|>K} \varepsilon e^{-|\imath_2 n + \imath_3 m| r} |w^{\imath_3}_m| e^{m\rho} e^{|\imath_2 n + \imath_3 m|\rho} \Bigg) \end{aligned}$$

$$\begin{aligned} &\leqslant (r-r_+)^{-1}\left(\sum_{|n|>K}\varepsilon e^{-|n|r}e^{|n|\rho}+\sum_{m,|n|>K}\varepsilon e^{-|n|r}|w_m^{l_3}|e^{m\rho}e^{|n|\rho}\right)\\ &\leqslant (r-r_+)^{-1}\sum_{|n|>K}\varepsilon e^{-|n|(r-\rho)}\\ &\leqslant (r-r_+)^{-1}\varepsilon e^{-K(r-\rho)}\leqslant\varepsilon_+, \end{aligned} \tag{4.12}$$

provided $\rho\leqslant r_+$.

Moreover, we take $s_+\ll s$ such that in a domain $D_\rho(r,s_+)$, the norm of the vector field of the higher order terms in the perturbation is bounded by ε_+. In conclusion, we have (4.11).

Below we look for a special F, defined in a domain $D_+=D_\rho(r_+,s_+)$ with $\rho\leqslant r_+$, such that the time one map ϕ_F^1 of the Hamiltonian vector field X_F defines a map from $D_+\to D$ and transforms H into H_+. More precisely, by second order Taylor formula, we have

$$\begin{aligned} H\circ\phi_F^1 &= (N+R)\circ\phi_F^1+(P-R)\circ\phi_F^1\\ &= N+\{N,F\}+R\\ &\quad+\int_0^1(1-t)\{\{N,F\},F\}\circ\phi_F^t\,dt\\ &\quad+\int_0^1\{R,F\}\circ\phi_F^t\,dt+(P-R)\circ\phi_F^1\\ &= N_++P_++\{N,F\}+R-P_{0000}-\langle\omega',I\rangle-\sum_n\langle R_{nn}^{011}z_n,\bar z_n\rangle, \end{aligned} \tag{4.13}$$

where

$$\begin{aligned} \omega' &= \int\frac{\partial P}{\partial I}\,d\theta|_{z=\bar z=0,I=0},\\ N_+ &= N+P_{0000}+\langle\omega',I\rangle+\sum_n\langle R_{nn}^{011}z_n,\bar z_n\rangle, \end{aligned} \tag{4.14}$$

$$P_+=\int_0^1(1-t)\{\{N,F\},F\}\circ\phi_F^t\,dt+\int_0^1\{R,F\}\circ\phi_F^t\,dt+(P-R)\circ\phi_F^1. \tag{4.15}$$

We shall find a function F of the form

$$
\begin{aligned}
F(\theta, I, w, \bar{w}) &= F_0 + F_1 + F_2 \\
&= \sum_{0<|k|\leqslant K, |l|\leqslant 1} F_{kl00} e^{i\langle k,\theta\rangle} I^l + \sum_{|k|\leqslant K, |n|\leqslant K} (f_n^{k10} w_n + f_n^{k01} \bar{w}_n) e^{i\langle k,\theta\rangle} \\
&\quad + \sum_{|k|\leqslant K, |n+m|\leqslant K} (f_{nm}^{k20} w_n w_m + f_{nm}^{k02} \bar{w}_n \bar{w}_m) e^{i\langle k,\theta\rangle} \\
&\quad + \sum_{|k|\leqslant K, |n-m|\leqslant K} f_{nm}^{k11} w_n \bar{w}_m e^{i\langle k,\theta\rangle} \\
&= \sum_{0<|k|\leqslant K, |l|\leqslant 1} F_{kl00} e^{i\langle k,\theta\rangle} I^l + \sum_{|k|\leqslant K, n} (\langle F_n^{k10}, z_n\rangle + \langle F_n^{k01}, \bar{z}_n\rangle) e^{i\langle k,\theta\rangle} \\
&\quad + \sum_{|k|\leqslant K, n, m} (\langle F_{mn}^{k20} z_n, z_m\rangle + \langle F_{mn}^{k02} \bar{z}_n, \bar{z}_m\rangle) e^{i\langle k,\theta\rangle} \\
&\quad + \sum_{|k|\leqslant K, n, m} \langle F_{mn}^{k11} z_n, \bar{z}_m\rangle e^{i\langle k,\theta\rangle}\,,
\end{aligned}
\tag{4.16}
$$

where

$$
\begin{aligned}
&F_n^{k10} = \begin{pmatrix} f_n^{k10} \\ f_{-n}^{k10} \end{pmatrix}, \quad F_n^{k01} = \begin{pmatrix} f_n^{k01} \\ f_{-n}^{k01} \end{pmatrix} \quad n \leqslant K, \\
&F_{mn}^{k20} = \begin{pmatrix} f_{nm}^{k20} & f_{(-n)m}^{k20} \\ f_{n(-m)}^{k20} & f_{(-n)(-m)}^{k20} \end{pmatrix}, \quad n+m \leqslant K \\
&F_{mn}^{k20} = \begin{pmatrix} 0 & f_{(-n)m}^{k20} \\ f_{n(-m)}^{k20} & 0 \end{pmatrix}, \quad n+m > K, |n-m| \leqslant K, \\
&F_{mn}^{k02} = \begin{pmatrix} f_{nm}^{k02} & f_{(-n)m}^{k02} \\ f_{n(-m)}^{k02} & f_{(-n)(-m)}^{k02} \end{pmatrix}, \quad n+m \leqslant K, \\
&F_{mn}^{k02} = \begin{pmatrix} 0 & f_{(-n)m}^{k02} \\ f_{n(-m)}^{k02} & 0 \end{pmatrix}, \quad n+m > K, |n-m| \leqslant K, \\
&F_{mn}^{k11} = \begin{pmatrix} f_{nm}^{k11} & f_{(-n)m}^{k11} \\ f_{n(-m)}^{k11} & f_{(-n)(-m)}^{k11} \end{pmatrix}, \quad n+m \leqslant K,
\end{aligned}
\tag{4.17}
$$

$$
F_{mn}^{k11} = \begin{pmatrix} f_{nm}^{k11} & 0 \\ 0 & f_{(-n)(-m)}^{k11} \end{pmatrix}, \quad n+m > K, |n-m| \leqslant K,
$$

satisfying the equation

$$\{N, F\} + R - P_{0000} - \langle \omega', I\rangle - \sum_n \langle R^{011}_{nn} z_n, \bar{z}_n\rangle = 0. \tag{4.18}$$

In this section, $|k|$, $|n|$, $|n+m|$ and $|n-m|$ are always smaller than K, which will not be reported again. What's more, since the term $\sum_n \langle R^{011}_{nn} z_n, \bar{z}_n\rangle$ has not been eliminated by symplectic change, so we define $F^{011}_{nn} = 0$, $n \geqslant 0$.

Lemma 4.1. *Eq. (4.18) is equivalent to*

$$\begin{aligned}
F_{kl00} &= i(\langle k, \omega\rangle)^{-1} P_{kl00}, \qquad k \neq 0, |l| \leqslant 1,\\
(\langle k, \omega\rangle I \;-\; A_n) F^{k10}_n &= i R^{k10}_n,\\
(\langle k, \omega\rangle I \;+\; A_n) F^{k01}_n &= i R^{k01}_n,\\
(\langle k, \omega\rangle I \;-\; A_m) F^{k20}_{mn} - F^{k20}_{mn} A_n &= i R^{k20}_{mn},\\
(\langle k, \omega\rangle I \;+\; A_m) F^{k11}_{mn} - F^{k11}_{mn} A_n &= i R^{k11}_{mn}, \;\; |k| + |n-m| \neq 0,\\
(\langle k, \omega\rangle I \;+\; A_m) F^{k02}_{mn} + F^{k02}_{mn} A_n &= i R^{k02}_{mn}.
\end{aligned} \tag{4.19}$$

Proof. Inserting F, defined in (4.16), into (4.18) one sees that (4.18) is equivalent to the following equations [6]

$$\begin{aligned}
\{N, F_0\} + R_0 &= P_{0000} + \langle \omega', I\rangle,\\
\{N, F_1\} + R_1 &= 0,\\
\{N, F_2\} + R_2 &= \sum_n \langle R^{011}_{nn} z_n, \bar{z}_n\rangle.
\end{aligned} \tag{4.20}$$

The first equation in (4.20) is obviously equivalent, by comparing the coefficients, to the first equation in (4.19). To solve $\{N, F_1\} + R_1 = 0$, we note that

$$\begin{aligned}
\{N, F_1\} = i \sum_{k,n} (\langle\, \langle k, \omega\rangle F^{k10}_n, z_n\rangle - \langle A_n z_n, F^{k10}_n\rangle) e^{i\langle k,\theta\rangle}\\
+ i \sum_{k,n} (\langle\, \langle k, \omega\rangle F^{k01}_n, \bar{z}_n\rangle + \langle A_n \bar{z}_n, F^{k01}_n\rangle) e^{i\langle k,\theta\rangle}
\end{aligned}$$

[6] Recall the definition of R_i in (4.8).

$$
\begin{aligned}
&= i\sum_{k,n}\langle(\langle k,\omega\rangle I - A_n)F_n^{k10}, z_n\rangle e^{i\langle k,\theta\rangle}\\
&\quad + i\sum_{k,n}\langle(\langle k,\omega\rangle I + A_n)F_n^{k01}, \bar z_n\rangle e^{i\langle k,\theta\rangle}\ . \qquad (4.21)
\end{aligned}
$$

It follows that F_n^{k10}, F_n^{k01} are determined by the linear algebraic systems

$$
i(\langle k,\omega\rangle I - A_n)F_n^{k10} + R_n^{k10} = 0, \quad n\in\mathbb{N}, k\in\mathbb{Z}^d.
$$

$$
i(\langle k,\omega\rangle I + A_n)F_n^{k01} + R_n^{k01} = 0, \quad n\in\mathbb{N}, k\in\mathbb{Z}^d.
$$

Similarly, from

$$
\begin{aligned}
\{N, F_2\} &= i\sum_{k,n,m}(\langle\ \langle k,\omega\rangle F_{mn}^{k20} z_n, z_m\rangle - \langle F_{mn}^{k20} z_n, A_m z_m\rangle - \langle A_n z_n, (F_{mn}^{k20})^T z_m\rangle)e^{i\langle k,\theta\rangle}\\
&\quad + i\sum_{|k|+|n-m|\neq 0}(\langle\ \langle k,\omega\rangle F_{mn}^{k11} z_n, \bar z_m\rangle + \langle F_{mn}^{k11} z_n, A_m \bar z_m\rangle\\
&\quad - \langle A_n z_n, (F_{mn}^{k11})^T \bar z_m\rangle)e^{i\langle k,\theta\rangle}\\
&\quad + i\sum_{k,n,m}(\langle\ \langle k,\omega\rangle F_{mn}^{k02} \bar z_n, \bar z_m\rangle + \langle F_{mn}^{k02} \bar z_n, A_m \bar z_m\rangle + \langle A_n \bar z_n, (F_{mn}^{k02})^T \bar z_m\rangle)e^{i\langle k,\theta\rangle}\\
&= i\sum_{k,n,m}(\langle\ \langle k,\omega\rangle F_{mn}^{k20} z_n, z_m\rangle - \langle(A_m F_{mn}^{k20} + F_{mn}^{k20} A_n)z_n, z_m\rangle)e^{i\langle k,\theta\rangle}\\
&\quad + i\sum_{|k|+|n-m|\neq 0}(\langle\ \langle k,\omega\rangle F_{mn}^{k11} z_n, \bar z_m\rangle + \langle(A_m F_{mn}^{k11} - F_{mn}^{k11} A_n)z_n, \bar z_m\rangle)e^{i\langle k,\theta\rangle}\\
&\quad + i\sum_{k,n,m}(\langle\ \langle k,\omega\rangle F_{mn}^{k02} \bar z_n, \bar z_m\rangle + \langle(A_m F_{mn}^{k02} + F_{mn}^{k02} A_n)\bar z_n, \bar z_m\rangle)e^{i\langle k,\theta\rangle}\\
&= i\sum_{k,n,m}\langle(\langle k,\omega\rangle F_{mn}^{k20} - A_m F_{mn}^{k20} - F_{mn}^{k20} A_n)z_n, z_m\rangle e^{i\langle k,\theta\rangle}\\
&\quad + i\sum_{|k|+|n-m|\neq 0}\langle(\langle k,\omega\rangle F_{mn}^{k11} + A_m F_{mn}^{k11} - F_{mn}^{k11} A_n)z_n, \bar z_m\rangle e^{i\langle k,\theta\rangle}\\
&\quad + i\sum_{k,n,m}\langle(\langle k,\omega\rangle F_{mn}^{k02} + A_m F_{mn}^{k02} + F_{mn}^{k02} A_n)\bar z_n, \bar z_m\rangle e^{i\langle k,\theta\rangle}\ . \qquad (4.22)
\end{aligned}
$$

It follows that, F_{mn}^{k20}, F_{mn}^{k11} and F_{mn}^{k02} are determined by the following matrix equations

$$(\langle k, \omega\rangle I - A_m)F_{mn}^{k20} - F_{mn}^{k20}A_n = iR_{mn}^{k20}, n, m \in \mathbb{N},\ \ k \in \mathbb{Z}^d,$$

$$(\langle k, \omega\rangle I + A_m)F_{mn}^{k11} - F_{mn}^{k11}A_n = iR_{mn}^{k11}, |k| + |n-m| \neq 0,$$

$$(\langle k, \omega\rangle I + A_m)F_{mn}^{k02} + F_{mn}^{k02}A_n = iR_{mn}^{k02}, n, m \in \mathbb{N},\ \ k \in \mathbb{Z}^d. \qquad \square \tag{4.23}$$

The first three equations in (4.19) are immediately solved in view of (4.1). In order to solve the last three equations in (4.19), we need the following elementary algebraic result from matrix theory.

Lemma 4.2. *Let A, B, C be, respectively, $n \times n, m \times m, n \times m$ matrices, and let X be an $n \times m$ unknown matrix. The matrix equation*

$$AX - XB = C, \tag{4.24}$$

is solvable if and only if $I_m \otimes A - B \otimes I_n$ is nonsingular.

In fact, the matrix equation (4.24) is equivalent to the (bigger) vector equation given by $(I_m \otimes A - B \otimes I_n)X' = C'$ where X', C' are vectors whose elements are just the list (row by row) of the entries of X and C. For a detailed proof we refer the reader to the Appendix in [16].

Remark. Taking the transpose of the fourth equation in (4.19), one sees that $(F_{mn}^{k20})^T$ satisfies the same equation as F_{nm}^{k20}. Then (by the uniqueness of the solution) it follows that $F_{nm}^{k20} = (F_{mn}^{k20})^T$. Similarly, $F_{nm}^{k11} = (F_{mn}^{k11})^T$, $F_{nm}^{k02} = (F_{mn}^{k02})^T$.

4.2. Estimates on the coordinate transformation

We proceed to estimate X_F and ϕ_F^1. We start with the following

Lemma 4.3. *Let $D_i = D_\rho(r_+ + \frac{i}{4}(r - r_+), \frac{i}{4}s)$, $0 < i \leqslant 4$. Then*

$$\|X_F\|_{D_3,\mathcal{O}} \leqslant c\gamma^{-8}(r - r_+)^{-c}K^{8(\tau+1)}\varepsilon. \tag{4.25}$$

Proof. By (4.1), Lemmas 4.1, 4.2 and Lemmas 7.5, 7.6 in the Appendix, we have

$$\begin{aligned} |F_{kl00}|_{\mathcal{O}} &\leqslant |\langle k, \omega\rangle|^{-1}|P_{kl00}|_{\mathcal{O}} < c\gamma^{-8}K^{8\tau+7}\varepsilon e^{-|k|r}s^{2-2|l|}, \quad k \neq 0;\\ \|F_n^{k10}\|_{\mathcal{O}} &\leqslant c\gamma^{-8}K^{8\tau+7}\|R_n^{k10}\|; \end{aligned} \tag{4.26}$$

$$\|F_n^{k01}\|_{\mathcal{O}} \leqslant c\gamma^{-8}K^{8\tau+7}\|R_n^{k01}\|;$$
$$\|F_{mn}^{k20}\|_{\mathcal{O}} \leqslant c\gamma^{-8}K^{8\tau+7}\|R_{mn}^{k20}\|;$$
$$\|F_{mn}^{k11}\|_{\mathcal{O}} \leqslant c\gamma^{-8}K^{8\tau+7}\|R_{mn}^{k11}|, \ |k|+|n-m|\neq 0;$$
$$\|F_{mn}^{k02}\|_{\mathcal{O}} \leqslant c\gamma^{-8}K^{8\tau+7}\|R_{mn}^{k02}\|.$$

It follows that

$$\begin{aligned}\frac{1}{s^2}\|F_\theta\|_{D_3,\mathcal{O}} \leqslant\ & \frac{1}{s^2}\Bigg(\sum_{k,|l|\leqslant 1}|F_{kl00}|s^{2|l|}|k|e^{|k|(r-\frac{1}{4}(r-r_+))}\\ &+\sum_{k,n}\|F_n^{k10}\||z_n||k|e^{|k|(r-\frac{1}{4}(r-r_+))}\\ &+\sum_{k,n}\|F_n^{k01}\||\bar{z}_n||k|e^{|k|(r-\frac{1}{4}(r-r_+))}\\ &+\sum_{k,n,m}\|F_{mn}^{k20}\||z_n||z_m||k|e^{|k|(r-\frac{1}{4}(r-r_+))}\\ &+\sum_{|k|+|n-m|\neq 0}\|F_{mn}^{k11}\||z_n||\bar{z}_m||k|e^{|k|(r-\frac{1}{4}(r-r_+))}\\ &+\sum_{k,n,m}\|F_{mn}^{k02}\||\bar{z}_n||\bar{z}_m||k|e^{|k|(r-\frac{1}{4}(r-r_+))}\Bigg)\\ \leqslant\ & c\gamma^{-8}(r-r_+)^{-c}K^{8(\tau+1)}\|X_R\|\\ \leqslant\ & c\gamma^{-8}(r-r_+)^{-c}K^{8(\tau+1)}\varepsilon. \end{aligned}\tag{4.27}$$

Similarly,

$$\|F_I\|_{D_3,\mathcal{O}} = \sum_{|l|=1}|F_{kl00}|e^{|k|(r-\frac{1}{4}(r-r_+))}\leqslant c\gamma^{-8}(r-r_+)^{-c}K^{8(\tau+1)}\varepsilon.$$

Now we estimate $\|X_{F_1}\|_{D_3,\mathcal{O}}$,

$$\begin{aligned}\|F_{1_{z_n}}\|_{D_3,\mathcal{O}} =\ & \left\|\sum_k F_n^{k10}e^{i\langle k,\theta\rangle}\right\|_{D_3,\mathcal{O}}\\ \leqslant\ & \sum_k \|F_n^{k10}\|e^{|k|(r-\frac{1}{4}(r-r_+))}\\ \leqslant\ & c\gamma^{-8}K^{8(\tau+1)}\sum_k\|R_n^{k10}\|e^{|k|(r-\frac{1}{4}(r-r_+))};\end{aligned}$$

similarly,

$$\|F_{1_{\bar z_n}}\|_{D_3,\mathcal{O}} \leqslant c\gamma^{-8}K^{8(\tau+1)}\sum_k \|R_n^{k01}\| e^{|k|(r-\frac{1}{4}(r-r_+))}.$$

By the definition of the weighted norm,[7] it follows that for $\rho \leqslant r_+$

$$\begin{aligned}\|X_{F_1}\|_{D_3,\mathcal{O}} &\leqslant \frac{c}{s}\left(\sum_n \|F_{1_{z_n}}\|_{D_3,\mathcal{O}} e^{n\rho} + \sum_n \|F_{1_{\bar z_n}}\|_{D_3,\mathcal{O}} e^{n\rho}\right)\\ &\leqslant c\gamma^{-8}(r-r_+)^{-c}K^{8(\tau+1)}\|X_R\|,\\ &\leqslant c\gamma^{-8}(r-r_+)^{-c}K^{8(\tau+1)}\varepsilon. \end{aligned} \tag{4.28}$$

Moreover,

$$\begin{aligned}\|F_{2_{z_n}}\|_{D_3,\mathcal{O}} &= \left\|\sum_{k,m} F_{mn}^{k20} z_m e^{i\langle k,\theta\rangle}\right\|_{D_3,\mathcal{O}} + \left\|\sum_{k,m} F_{mn}^{k11} \bar z_m e^{i\langle k,\theta\rangle}\right\|_{D_3,\mathcal{O}} \\ &\leqslant c\gamma^{-8}K^{8(\tau+1)}\left(\sum_{k,m}\|R_{mn}^{k20}\||z_m| e^{|k|(r-\frac{1}{4}(r-r_+))}\right.\\ &\quad \left. + \sum_{k,m}\|R_{mn}^{k11}\||\bar z_m| e^{|k|(r-\frac{1}{4}(r-r_+))}\right); \end{aligned} \tag{4.29}$$

similarly,

$$\begin{aligned}\|F_{2_{\bar z_n}}\|_{D_3,\mathcal{O}} &\leqslant c\gamma^{-8}K^{8(\tau+1)}\left(\sum_{k,m}\|R_{mn}^{k11}\||z_m| e^{|k|(r-\frac{1}{4}(r-r_+))}\right.\\ &\quad \left. + \sum_{k,m}\|R_{mn}^{k02}\||\bar z_m| e^{|k|(r-\frac{1}{4}(r-r_+))}\right). \end{aligned}$$

[7] Recall (3.15), the definition of the norm.

Hence for $\rho \leqslant r_+$, we have

$$\begin{aligned} \|X_{F_2}\|_{D_3,\mathcal{O}} & \\ &\leqslant \frac{c}{s}\left(\sum_n \|F_{2_{z_n}}\|_{D_3,\mathcal{O}} e^{n\rho} + \sum_n \|F_{2_{\bar{z}_n}}\|_{D_3,\mathcal{O}} e^{n\rho}\right) \\ &\leqslant c\gamma^{-8}(r-r_+)^{-c} K^{8(\tau+1)} \|X_R\| \\ &\leqslant c\gamma^{-8}(r-r_+)^{-c} K^{8(\tau+1)} \varepsilon. \end{aligned} \tag{4.30}$$

The conclusion of the lemma follows from the above estimates. □

In the next lemma, we give some estimates for ϕ_F^t. The following formula (4.31) will be used to prove that our coordinate transformation is well defined. Inequality (4.32) will be used to check the convergence of the iteration.

Lemma 4.4. *Let* $\eta = \varepsilon^{\frac{1}{3}}$, $D_{i\eta} = D_\rho(r_+ + \frac{i}{4}(r - r_+), \frac{i}{4}\eta s), 0 < i \leqslant 4$. *We then have*

$$\phi_F^t : D_{2\eta} \to D_{3\eta}, \quad -1 \leqslant t \leqslant 1, \tag{4.31}$$

if $\varepsilon \ll (\frac{1}{2}\gamma^8 (r - r_+)^c K^{-8(\tau+1)})^{\frac{3}{2}}$. *Moreover,*

$$\|D\phi_F^t - Id\|_{D_{1\eta}} < c\gamma^{-8}(r-r_+)^{-c} K^{8(\tau+1)} \varepsilon. \tag{4.32}$$

Proof. Let

$$\|D^m F\|_{D,\mathcal{O}} = \max\left\{ \left\| \frac{\partial^{|i|+|l|+|\alpha|+|\beta|}}{\partial\theta^i \partial I^l \partial w^\alpha \partial \bar{w}^\beta} F \right\|_{D,\mathcal{O}}, |i| + |l| + |\alpha| + |\beta| = m \geqslant 2 \right\}.$$

Note that F is polynomial in I of order 1, in w, $\bar{w}$ of order 2. From[8] (4.30) and the Cauchy inequality, it follows that

$$\|D^m F\|_{D_2,\mathcal{O}} < c\gamma^{-8}(r-r_+)^{-c} K^{8(\tau+1)} \varepsilon, \tag{4.33}$$

for any $m \geqslant 2$.

To get the estimates for ϕ_F^t, we start from the integral equation,

$$\phi_F^t = id + \int_0^t X_F \circ \phi_F^s \, ds$$

[8] Recall the definition of the weighted norm in (3.15).

so that $\phi_F^t : D_{2\eta} \to D_{3\eta}, \quad -1 \leqslant t \leqslant 1$, which is followed directly from (4.33). Since

$$D\phi_F^t = Id + \int_0^t (DX_F) D\phi_F^s \, ds = Id + \int_0^t J(D^2 F) D\phi_F^s \, ds,$$

where J denotes the standard symplectic matrix $\begin{pmatrix} 0 & -I \\ I & 0 \end{pmatrix}$, it follows that

$$\|D\phi_F^t - Id\| \leqslant 2\|D^2 F\| < c\gamma^{-8}(r - r_+)^{-c} K^{8(\tau+1)} \varepsilon. \tag{4.34}$$

Consequently Lemma 4.4 follows. □

4.3. Estimates for the new normal form

The map ϕ_F^1 defined above transforms H into $H_+ = N_+ + P_+$(see (4.13) and (4.18)) with

$$\begin{aligned} N_+ &= N + P_{0000} + \langle \omega', I \rangle + \sum_n \langle R_{nn}^{011} z_n, \bar{z}_n \rangle \\ &= e_+ + \langle \omega_+, I \rangle + \sum_n \langle A_n^+ z_n, \bar{z}_n \rangle, \end{aligned} \tag{4.35}$$

where [9]

$$e_+ = e + P_{0000}, \;\; \omega_+ = \omega + P_{0l00}(|l| = 1),$$

$$\begin{aligned} A_n^+ &= A_n + R_{nn}^{011} = A_n + \begin{pmatrix} P_{nn}^{011} & P_{(-n)n}^{011} \\ P_{n(-n)}^{011} & P_{(-n)(-n)}^{011} \end{pmatrix} \\ &= \begin{pmatrix} \Omega_n^+ & b_{((-n),n)}^+ \\ b_{(n,(-n))}^+ & \Omega_{-n}^+ \end{pmatrix}, \quad 2n \leqslant K, \end{aligned}$$

$$\begin{aligned} A_n^+ &= A_n + R_{nn}^{011} = A_n + \begin{pmatrix} P_{nn}^{011} & 0 \\ 0 & P_{(-n)(-n)}^{011} \end{pmatrix} \\ &= \begin{pmatrix} \Omega_n^+ & 0 \\ 0 & \Omega_{-n}^+ \end{pmatrix}, \quad 2n > K. \end{aligned}$$

---

[9] Recall the definition of R_{nn}^{011} (4.9).

Remark. In the next KAM step, A_n^+ with $K < 2n \leqslant K_+$ will become non-diagonal matrices.

Now we prove that N_+ shares the same properties as N. By the Assumptions of P and by Cauchy estimates, we have

$$|\omega_+ - \omega|_{\mathcal{O}} < \varepsilon, \quad |P^{011}_{nn}|_{\mathcal{O}}, |P^{011}_{(-n)n}|_{\mathcal{O}}, |P^{011}_{n(-n)}|_{\mathcal{O}} < \varepsilon \tag{4.36}$$

$$|b^+_{((-n),n)}|_{\mathcal{O}}, |b^+_{(n,(-n))}|_{\mathcal{O}} \leqslant \varepsilon_0 e^{-2n\rho}.$$

It follows that

$$|\langle k, \omega + P_{0l00}\rangle^{-1}| \leqslant \frac{|\langle k, \omega\rangle^{-1}|}{1 - |\langle k, \omega\rangle^{-1}|\varepsilon} \leqslant \frac{K_+^\tau}{\gamma}, \tag{4.37}$$

provided that $(K_+K)^\tau \varepsilon < (K_+^\tau - K^\tau)\gamma$.

$$\|(\langle k, \omega + P_{0l00}\rangle I + A_n^+)^{-1}\| \leqslant \frac{\|(\langle k, \omega\rangle I + A_n)^{-1}\|}{1 - \|(\langle k, \omega\rangle I + A_n)^{-1}\|\varepsilon} \leqslant \left(\frac{K_+^\tau}{\gamma}\right)^2, \tag{4.38}$$

provided that $(K_+K)^{2\tau}\varepsilon < (K_+^{2\tau} - K^{2\tau})\gamma^2$. Similarly, we have

$$\|(\langle k, \omega + P_{0l00}\rangle I \pm A_n^+ \otimes I \pm I \otimes A_m^+)^{-1}\| \leqslant \left(\frac{K_+^\tau}{\gamma}\right)^4, n \neq m \tag{4.39}$$

$$\|(\langle k, \omega + P_{0l00}\rangle I + A_n^+ \otimes I + I \otimes A_n^+)^{-1}\| \leqslant \left(\frac{K_+^\tau}{\gamma}\right)^4, \tag{4.40}$$

$$\|(\langle k, \omega + P_{0l00}\rangle I + A_n^+ \otimes I - I \otimes A_n^+)^{-1}\| \leqslant \left(\frac{K_+^\tau}{\gamma}\right)^4, 2n \leqslant K \tag{4.41}$$

provided that $(K_+K)^{4\tau}\varepsilon < (K_+^{4\tau} - K^{4\tau})\gamma^4$. This means that in the next KAM step, small denominator conditions are automatically satisfied for $|k| \leqslant K$. The following bounds will be used for the measure estimates:

$$\sup_{\xi \in \mathcal{O}} \max_{|p| \leqslant 4} \left(\left| \frac{\partial^p (\omega_+ - \omega)}{\partial \xi^p} \right| \right) \leqslant \varepsilon,$$

$$\sup_{\xi \in \mathcal{O}} \max_{|p| \leqslant 4} \left(\left\| \frac{\partial^p (A_n^+ - A_n)}{\partial \xi^p} \right\| \right) \leqslant c\varepsilon. \tag{4.42}$$

And

$$|b^{+}_{((-n),n)}|_{\mathcal{O}}, |b^{+}_{(n,(-n))}|_{\mathcal{O}} \leqslant \varepsilon_0 e^{-2n\rho},$$

will be used for proving the *decay property of F*.

4.4. Estimates for the new perturbation

Since

$$\begin{aligned} P_+ &= \int_0^1 (1-t)\{\{N, F\}, F\} \circ \phi_F^t \, dt + \int_0^1 \{R, F\} \circ \phi_F^t \, dt + (P-R) \circ \phi_F^1 \\ &= \int_0^1 \{R(t), F\} \circ \phi_F^t \, dt + (P-R) \circ \phi_F^1, \end{aligned}$$

where $R(t) = (1-t)(N_+ - N) + tR$. Hence

$$X_{P_+} = \int_0^1 (\phi_F^t)^* X_{\{R(t),F\}} \, dt + (\phi_F^1)^* X_{(P-R)}.$$

According to Lemma 4.4,

$$\|D\phi_F^t - Id\|_{D_{1\eta}} < c\gamma^{-8}(r-r_+)^{-c} K^{8(\tau+1)} \varepsilon, \quad -1 \leqslant t \leqslant 1,$$

thus

$$\|D\phi_F^t\|_{D_{1\eta}} \leqslant 1 + \|D\phi_F^t - Id\|_{D_{1\eta}} \leqslant 2, \quad -1 \leqslant t \leqslant 1.$$

Due to Lemma 7.3,

$$\|X_{\{R(t),F\}}\|_{D_{2\eta}} \leqslant c\gamma^{-8}(r-r_+)^{-c} K^{8(\tau+1)} \eta^{-2} \varepsilon^2,$$

and

$$\|X_{(P-R)}\|_{D_{2\eta}} \leqslant c\eta\varepsilon$$

therefore,

$$\|X_{P_+}\|_{D_\rho(r_+,s_+)} \leqslant c\eta\varepsilon + c\gamma^{-8}(r-r_+)^{-c} K^{8(\tau+1)} \eta^{-2} \varepsilon^2 \leqslant c\varepsilon_+.$$

4.5. Decay property of the new perturbation

Compared with the previous KAM iteration, decay property of new perturbation is added in order to truncate the smaller terms in the next step. Thus the term

$$\sum_{|\imath_1 n+\imath_2 m|>K_+} P_{+(n,m)}^{(\imath_1\imath_2)}(\theta, I) w_n^{\imath_1} w_m^{\imath_2}$$

is small enough to be delayed in the following step. To this end, we have to check the new error term P_+ satisfies (4.5) with ε_+ in place of ε.

Since

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

Due to the assumption, $P-R$ has decay estimates. Moreover, $\{N, F\}$ satisfies the equation (4.18) and R is the truncation of P, hence We only estimate $\{P, F\}, \ldots, \frac{1}{n!}\{\cdots\{P, \underbrace{F\}\ldots, F}_{n}\}$ and then sum up them.

Since

$$\begin{aligned} F &= \sum_{\vec{n}, \vec{\imath}', |\langle \vec{\imath}', \vec{n}\rangle| \leqslant K} F_{\vec{n}}^{\vec{\imath}'}(\theta, I) w_{\vec{n}}^{\vec{\imath}'} \\ &= \sum_{0<|k|\leqslant K, |l|\leqslant 1} F_{kl00} e^{i\langle k,\theta\rangle} I^l + \sum_{|k|\leqslant K, |n|\leqslant K} (f_n^{k10} w_n + f_n^{k01} \bar{w}_n) e^{i\langle k,\theta\rangle} \\ &\quad + \sum_{|k|\leqslant K, |n+m|\leqslant K} (f_{nm}^{k20} w_n w_m + f_{nm}^{k02} \bar{w}_n \bar{w}_m) e^{i\langle k,\theta\rangle} \\ &\quad + \sum_{|k|\leqslant K, |n-m|\leqslant K} f_{nm}^{k11} w_n \bar{w}_m e^{i\langle k,\theta\rangle}, \end{aligned}$$

due to (4.4) and (4.5), one can obtain decay property

$$|P_n^{k10}|_{\mathcal{O}}, |P_n^{k01}|_{\mathcal{O}} \leqslant \varepsilon e^{-|n|\rho} e^{-|k|r}$$

and

$$|P_{nm}^{k20}|_{\mathcal{O}}, |P_{nm}^{k02}|_{\mathcal{O}} \leqslant \varepsilon e^{-|n+m|\rho} e^{-|k|r},$$

$$|P_{nm}^{k11}|_{\mathcal{O}} \leqslant \varepsilon e^{-|n-m|\rho} e^{-|k|r}. \tag{4.43}$$

Then according to Lemmas 7.9 and 7.10, F has decay estimates

$$\begin{aligned} |f_n^{k10}|, |f_n^{k01}| &\leqslant c\gamma^{-8} K^{8(\tau+1)} \varepsilon e^{-n\rho} e^{-|k|r}, n \in \mathbb{Z}, \\ |f_{nm}^{k20}|, |f_{nm}^{k02}| &\leqslant c\gamma^{-8} K^{8(\tau+1)} \varepsilon e^{-|n+m|\rho} e^{-|k|r}, n, m \in \mathbb{Z}, \\ |f_{nm}^{k11}| &\leqslant c\gamma^{-8} K^{8(\tau+1)} \varepsilon e^{-|n-m|\rho} e^{-|k|r}, \end{aligned}$$

using Lemma 7.4 again, we have decay property of F

$$||F_{\vec{n}}^{\vec{\imath}'}(\theta, I)|| \leqslant c\gamma^{-8}(r - r_+)^{-c} K^{8(\tau+1)} \varepsilon e^{-|\langle \vec{\imath}', \vec{n}\rangle|\rho}. \tag{4.44}$$

In addition,

$$P = \sum_{\vec{n},\vec{\imath}} P_{\vec{n}}^{\vec{\imath}}(\theta, I) w_{\vec{n}}^{\vec{\imath}},$$

due to (4.4) and (4.5),

$$||P_{\vec{n}}^{\vec{\imath}}(\theta, I)|| \leqslant \varepsilon e^{-\rho|\langle \vec{\imath}, \vec{n}\rangle|}.$$

According to Lemma 7.11, then on $D_\rho(r_+ + \frac{1}{8}(r - r_+), \frac{1}{8}\eta s + \frac{1}{16}\eta s)$, $\{P, F\} = \sum_{\vec{n},\vec{\imath}''} P_{\vec{n}}^{\vec{\imath}''}(\theta, I) w_{\vec{n}}^{\vec{\imath}''}$ has decay estimates

$$\begin{aligned} ||P_{\vec{n}}^{\vec{\imath}''}(\theta, I)|| &\leqslant c\gamma^{-8}(r - r_+)^{-c} K^{8\tau+9} \eta^{-2} \varepsilon^2 e^{-\rho|\langle \vec{\imath}'', \vec{n}\rangle|} \\ &\leqslant c\gamma^{-8}(r - r_+)^{-c} \varepsilon^{\frac{7}{6}} e^{-\rho|\langle \vec{\imath}'', \vec{n}\rangle|}. \end{aligned}$$

Going on with n steps, then on $D_\rho(r_+ + \frac{1}{2^n}\frac{1}{4}(r - r_+), \frac{1}{8}\eta s + \frac{1}{2^n}\frac{1}{8}\eta s)$, coefficient of $\frac{1}{n!}\{\cdots\{P, \underbrace{F\}\ldots, F}_{n}\}$ has decay estimates $\frac{4^n}{n!}\gamma^{-8n}(r - r_+)^{-cn}\varepsilon^{\frac{n+6}{6}} e^{-\rho|\langle \vec{\imath}'', \vec{n}\rangle|}$. Therefore, $P_+ = \sum_{\vec{n},\vec{\imath}''} P_{+\vec{n}}^{\vec{\imath}''}(\theta, I) w_{\vec{n}}^{\vec{\imath}''}$ has decay property

$$\begin{aligned} ||P_{+\vec{n}}^{\vec{\imath}''}(\theta, I)|| &\leqslant \sum_{n=1}^{\infty} \frac{4^n \gamma^{-8n}(r - r_+)^{-cn}\varepsilon^{\frac{n+6}{6}}}{n!} e^{-\rho|\langle \vec{\imath}'', \vec{n}\rangle|} \\ &\leqslant \varepsilon_+ e^{-\rho|\langle \vec{\imath}'', \vec{n}\rangle|}. \end{aligned} \tag{4.45}$$

Thus decay property of P_+ is obtained, and KAM step is now completed.

5. Iteration lemma and convergence

For any given $s, \varepsilon, r, \gamma$, we define, for all $v \geqslant 1$, the following sequences

$$r_v = r\left(1 - \sum_{i=2}^{v+1} 2^{-i}\right),$$

$$\varepsilon_v = c\gamma^{-8}(r_{v-1} - r_v)^{-c}\varepsilon_{v-1}^{\frac{7}{6}}, \tag{5.1}$$

$$\eta_v = \varepsilon_v^{\frac{1}{3}}, \qquad L_v = L_{v-1} + \varepsilon_{v-1},$$

$$s_v = \frac{1}{8}\eta_{v-1}s_{v-1} = 2^{-3v}\left(\prod_{i=0}^{v-1}\varepsilon_i\right)^{\frac{1}{3}} s_0,$$

$$K_v^d e^{-K_v(r_v - r_{v+1})} = \varepsilon_v^{\frac{1}{6}},$$

where c is the constant, $\gamma = \varepsilon_0^{\frac{1}{50}} \gg \varepsilon_0$, and the parameters $r_0, \varepsilon_0, L_0, s_0$ are defined respectively to be r, ε, L, s. Note that

$$\Psi(r) = \prod_{i=1}^{\infty}[(r_{i-1} - r_i)^{-c}]^{(\frac{6}{7})^i}$$

is a well defined finite function of r.

5.1. Iteration lemma

The preceding analysis may be summarized as follows.

Lemma 5.1. *Suppose that $\varepsilon_0 = \varepsilon(K_0, d, L, \tau, \gamma)$ is small enough. Then the following holds for all $v \geqslant 0$. Let*

$$N_v = e_v + \langle \omega_v(\xi), I\rangle + \sum_n \langle A_n^v(\xi) z_n, \bar{z}_n\rangle,$$

be a normal form, A_n^v with $2n > K_{v-1}$ are diagonal matrices, and parameters ξ satisfy

$$|\langle k, \omega_v\rangle^{-1}| \leqslant \frac{K_v^\tau}{\gamma},$$

$$\|(\langle k,\omega_\nu\rangle I + A_n^\nu)^{-1}\| \leqslant \left(\frac{K_\nu^\tau}{\gamma}\right)^2,$$

$$\|(\langle k,\omega_\nu\rangle I \pm A_n^\nu \otimes I \pm I \otimes A_m^\nu)^{-1}\| \leqslant \left(\frac{K_\nu^\tau}{\gamma}\right)^4, \;\; n \neq m$$

$$\|(\langle k,\omega_\nu\rangle I + A_n^\nu \otimes I + I \otimes A_n^\nu)^{-1}\| \leqslant \left(\frac{K_\nu^\tau}{\gamma}\right)^4,$$

$$\|(\langle k,\omega_\nu\rangle I + A_n^\nu \otimes I - I \otimes A_n^\nu)^{-1}\| \leqslant \left(\frac{K_\nu^\tau}{\gamma}\right)^4, 2n \leqslant K_\nu$$

on a closed set $\mathcal{O}_\nu$ of $\mathbb{R}^d$ for all $0 < |k| \leqslant K_\nu$. Moreover, suppose that $\omega_\nu(\xi)$, $b^\nu_{((-n),n)}(\xi)$, $b^\nu_{(n,(-n))}(\xi)$, $A_n^\nu(\xi)$ are C_W^4 smooth and satisfy

$$\sup_{\xi\in\mathcal{O}_\nu} \max_{|p|\leqslant 4} \left(\left| \frac{\partial^p(\omega_\nu - \omega_{\nu-1})}{\partial \xi^p} \right| \right) \leqslant \varepsilon_{\nu-1},$$

$$|b^\nu_{((-n),n)}(\xi)|_{\mathcal{O}_\nu}, |b^\nu_{(n,(-n))}(\xi)|_{\mathcal{O}_\nu} \leqslant \varepsilon_0 e^{-2n\rho},$$

$$\sup_{\xi\in\mathcal{O}_\nu} \max_{|p|\leqslant 4} \left(\left\| \frac{\partial^p(A_n^\nu - A_n^{\nu-1})}{\partial \xi^p} \right\| \right) \leqslant c\varepsilon_{\nu-1},$$

in the sense of Whitney.

Finally, assume that P_ν satisfies decay estimates (4.5) with $\frac{r_0}{2} < \rho \leqslant r_\nu$ and

$$\|X_{P_\nu}\|_{D_\rho(r_\nu,s_\nu),\mathcal{O}_\nu} \leqslant \varepsilon_\nu.$$

Then there is a subset $\mathcal{O}_{\nu+1} \subset \mathcal{O}_\nu$,

$$\mathcal{O}_{\nu+1} = \mathcal{O}_\nu \Bigg\backslash \left(\left(\bigcup_{K_\nu < |k| \leqslant K_{\nu+1}} \mathcal{R}_k^{\nu+1}(\gamma) \right) \bigcup \mathcal{Q}_{kn}^{\nu+1}(\gamma) \right),$$

where,

$$\mathcal{R}_k^{v+1}(\gamma)=\left\{\xi\in\mathcal{O}_v:\quad \begin{array}{l} |\langle k,\omega_{v+1}\rangle^{-1}|>\frac{K_{v+1}^{\tau}}{\gamma},\ \ \|(\langle k,\omega_{v+1}\rangle I+A_n^{v+1})^{-1}\|>\left(\frac{K_{v+1}^{\tau}}{\gamma}\right)^2, \text{or}\\ \|(\langle k,\omega_{v+1}\rangle I\pm A_n^{v+1}\otimes I\pm I\otimes A_m^{v+1})^{-1}\|>\left(\frac{K_{v+1}^{\tau}}{\gamma}\right)^4, n\neq m, \text{or}\\ \|(\langle k,\omega_{v+1}\rangle I+A_n^{v+1}\otimes I+I\otimes A_n^{v+1})^{-1}\|>\left(\frac{K_{v+1}^{\tau}}{\gamma}\right)^4 \end{array}\right\},$$

$$\mathcal{Q}_{kn}^{v+1}(\gamma)=\left\{\xi\in\mathcal{O}_v:\|(\langle k,\omega_{v+1}\rangle I+A_n^{v+1}\otimes I-I\otimes A_n^{v+1})^{-1}\|>\left(\frac{K_{v+1}^{\tau}}{\gamma}\right)^4,\quad \begin{array}{l} K_v<|k|\leqslant K_{v+1},\\ 2n\leqslant K_{v+1}\ \text{or}\\ 0<|k|\leqslant K_v,\\ K_v<2n\leqslant K_{v+1}\end{array}\right\}$$

with $\omega_{v+1}=\omega_v+P_{0l00}^v$, *and a symplectic transformation of variables*

$$\Phi_v: D_\rho(r_{v+1},s_{v+1})\times\mathcal{O}_v\to D_\rho(r_v,s_v), \tag{5.2}$$

such that $H_{v+1}=H_v\circ\Phi_v$, *defined on* $D_\rho(r_{v+1},s_{v+1})\times\mathcal{O}_{v+1}$, *has the form*

$$H_{v+1}=e_{v+1}+\langle\omega_{v+1},I\rangle+\sum_n\langle A_n^{v+1}z_n,\bar z_n\rangle+P_{v+1}, \tag{5.3}$$

satisfying that A_n^{v+1} *with* $2n>K_v$ *are diagonal matrices and*

$$\sup_{\xi\in\mathcal{O}_{v+1}}\max_{|p|\leqslant4}\left(\left|\frac{\partial^p(\omega_{v+1}-\omega_v)}{\partial\xi^p}\right|\right)\leqslant\varepsilon_v, \tag{5.4}$$

$$|b_{((-n),n)}^{v+1}(\xi)|_{\mathcal{O}_{v+1}},|b_{(n,(-n))}^{v+1}(\xi)|_{\mathcal{O}_{v+1}}\leqslant\varepsilon_0e^{-2n\rho},$$

$$\sup_{\xi\in\mathcal{O}_{v+1}}\max_{|p|\leqslant4}\left(\left\|\frac{\partial^p(A_n^{v+1}-A_n^v)}{\partial\xi^p}\right\|\right)\leqslant c\varepsilon_v,$$

in the sense of Whitney. Moreover, P_{v+1} *satisfies decay estimates* (4.45) *with* $\frac{r_0}{2}<\rho\leqslant r_{v+1}$ *and*

$$\|X_{P_{v+1}}\|_{D_\rho(r_{v+1},s_{v+1}),\mathcal{O}_{v+1}}\leqslant\varepsilon_{v+1}. \tag{5.5}$$

5.2. Convergence

Suppose that the assumptions of Theorem 1 are satisfied to apply the iteration Lemma with $v=0$, recall that

$$\varepsilon_0=\varepsilon, r_0=r, s_0=s, L_0=L, N_0=N, P_0=P, \gamma=\varepsilon^{\frac{1}{50}},$$

$$\mathcal{O}_0=\left\{\xi\in\mathcal{O}: \begin{array}{c} |\langle k,\omega(\xi)\rangle^{-1}|\leqslant\frac{K_0^\tau}{\gamma},\ \ \|(\langle k,\omega(\xi)\rangle I+A_n)^{-1}\|\leqslant\left(\frac{K_0^\tau}{\gamma}\right)^2, \\ \|(\langle k,\omega(\xi)\rangle I\pm A_n\otimes I\pm I\otimes A_m)^{-1}\|\leqslant\left(\frac{K_0^\tau}{\gamma}\right)^4, n\neq m, \\ \|(\langle k,\omega(\xi)\rangle I+A_n\otimes I+I\otimes A_n)^{-1}\|\leqslant\left(\frac{K_0^\tau}{\gamma}\right)^4, \\ \|(\langle k,\omega(\xi)\rangle I+A_n\otimes I-I\otimes A_n)^{-1}\|\leqslant\left(\frac{K_0^\tau}{\gamma}\right)^4,\ \ 2n\leqslant K_0 \end{array}\right\},$$

(with ε small enough). Inductively, we obtain the following sequences:

$$\mathcal{O}_{v+1}\subset\mathcal{O}_v,$$

$$\Psi^v=\Phi_1\circ\cdots\circ\Phi_v: D_\rho(r_{v+1},s_{v+1})\times\mathcal{O}_v\to D_0, v\geqslant 0,$$

$$H\circ\Psi^v=H_{v+1}=N_{v+1}+P_{v+1}.$$

Let $\tilde{\mathcal{O}}=\cap_{v=0}^{\infty}\mathcal{O}_v$. As in [7,12], thanks to Lemma 4.4, we may conclude that N_v, Ψ^v, $D\Psi^v$, ω_v converge uniformly on $D_{\frac{1}{2}r}(\frac{1}{2}r,0)\times\tilde{\mathcal{O}}$ with

$$N_\infty=e_\infty+\langle\omega_\infty,I\rangle+\sum_n\langle A_n^\infty z_n,\bar{z}_n\rangle.$$

Since

$$\varepsilon_{v+1}=c\gamma^{-8}(r_v-r_{v+1})^{-c}\varepsilon_v^{\frac{7}{6}}\leqslant(c\gamma^{-8}\Psi(r)\varepsilon)^{(\frac{7}{6})^v}.$$

it follows that $\varepsilon_{v+1}\to 0$ provided ε is sufficiently small.

Let ϕ_H^t be the flow of X_H, since $H\circ\Psi^v=H_{v+1}$, we have that

$$\phi_H^t\circ\Psi^v=\Psi^v\circ\phi_{H_{v+1}}^t. \tag{5.6}$$

The uniform convergence of Ψ^v, $D\Psi^v$, ω_v, X_{H_v} implies that one can take limit in (5.6) so as to get

$$\phi_H^t\circ\Psi^\infty=\Psi^\infty\circ\phi_{H_\infty}^t. \tag{5.7}$$

on $D_{\frac{1}{2}r}(\frac{1}{2}r,0)\times\tilde{\mathcal{O}}$, with

$$\Psi^{\infty}: D_{\frac{1}{2}r}\left(\frac{1}{2}r,0\right)\times\tilde{\mathcal{O}}\to D_{\rho}(r,s)\times\mathcal{O}.$$

From (5.7) one follows that

$$\phi_H^t(\Psi^{\infty}(\mathbb{T}^d\times\{\xi\}))=\Psi^{\infty}\phi_{N_\infty}^t(\mathbb{T}^d\times\{\xi\})=\Psi^{\infty}(\mathbb{T}^d\times\{\xi\}),$$

for $\xi\in\tilde{\mathcal{O}}$. This means that $\Psi^{\infty}(\mathbb{T}^d\times\{\xi\})$ is an embedded torus invariant for the original perturbed Hamiltonian system at $\xi\in\tilde{\mathcal{O}}$. We remark here that the frequencies $\omega_\infty(\xi)$ associated to $\Psi^{\infty}(\mathbb{T}^d\times\{\xi\})$ is slightly different from $\omega(\xi)$. The normal behavior of the invariant torus is governed by the matrices A_n^{∞}. □

6. Measure estimates

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

$$\mathcal{R}^{\nu+1}=\bigcup_{K_\nu<|k|\leqslant K_{\nu+1},n,m}\left(\mathcal{R}_k^{\nu+1}\bigcup\mathcal{R}_{kn}^{\nu+1}\bigcup\mathcal{R}_{knm}^{\nu+1}\right)\bigcup\mathcal{Q}_{kn}^{\nu+1},$$

the sets $\mathcal{R}_k^{\nu+1}$, $\mathcal{R}_{kn}^{\nu+1}$, $\mathcal{R}_{knm}^{\nu+1}$, $\mathcal{Q}_{kn}^{\nu+1}$ being, respectively,

$$\left\{\xi\in\mathcal{O}_\nu:|\langle k,\omega_{\nu+1}(\xi)\rangle^{-1}|>\frac{K_{\nu+1}^{\tau}}{\gamma}\right\},\tag{6.1}$$

$$\left\{\xi\in\mathcal{O}_\nu:\|(\langle k,\omega_{\nu+1}(\xi)\rangle I+A_n^{\nu+1})^{-1}\|>\left(\frac{K_{\nu+1}^{\tau}}{\gamma}\right)^2\right\},\tag{6.2}$$

$$\left\{\xi\in\mathcal{O}_\nu:\begin{array}{l}\|(\langle k,\omega_{\nu+1}(\xi)\rangle I\pm A_n^{\nu+1}\otimes I\pm I\otimes A_m^{\nu+1})^{-1}\|>\left(\frac{K_{\nu+1}^{\tau}}{\gamma}\right)^4,n\neq m\\ \|(\langle k,\omega_{\nu+1}(\xi)\rangle I+A_n^{\nu+1}\otimes I+I\otimes A_n^{\nu+1})^{-1}\|>\left(\frac{K_{\nu+1}^{\tau}}{\gamma}\right)^4\end{array}\right\},\tag{6.3}$$

$$\left\{\xi\in\mathcal{O}_\nu:\|(\langle k,\omega_{\nu+1}(\xi)\rangle I+A_n^{\nu+1}\otimes I-I\otimes A_n^{\nu+1})^{-1}\|>\left(\frac{K_{\nu+1}^{\tau}}{\gamma}\right)^4,\quad\begin{array}{l}K_\nu<|k|\leqslant K_{\nu+1},\\ 2n\leqslant K_{\nu+1}\ \text{or}\\ 0<|k|\leqslant K_\nu,\\ K_\nu<2n\leqslant K_{\nu+1}\end{array}\right\},$$

recall that $\omega_{\nu+1}(\xi) = \omega(\xi) + \sum_{j=0}^{\nu} P^j_{0l00}(\xi)$ with $|\sum_{j=0}^{\nu} P^j_{0l00}(\xi)|_{\mathcal{O}_\nu} \leqslant \varepsilon$, and

$$\|A_n^{\nu+1}(\xi) - A_n(\xi)\|_{\mathcal{O}_{\nu+1}} \leqslant \sum_{j=0}^{\nu} \|R_{nn}^{011,j}\| \leqslant \varepsilon. \tag{6.4}$$

Remark 2. Different from the former KAM methods, we not only excise the resonant set $\mathcal{Q}_{kn}^{\nu+1}$ with $K_\nu < |k| \leqslant K_{\nu+1}, 2n \leqslant K_{\nu+1}$, but we also excise the resonant set $\mathcal{Q}_{kn}^{\nu+1}$ with $0 < |k| \leqslant K_\nu$, $K_\nu < 2n \leqslant K_{\nu+1}$ (see Lemmas 6.2 and 6.3 below). Note that at the beginning, $A_n = \Omega_n I$, $A_m = \Omega_m I$, $\Omega_n - \Omega_m = n^2 - m^2$ is independent of ξ.

Lemma 6.1. *For any fixed* $K_\nu < |k| \leqslant K_{\nu+1}$, n, m,

$$\mathrm{meas}\left(\mathcal{R}_k^{\nu+1} \bigcup \mathcal{R}_{kn}^{\nu+1} \bigcup \mathcal{R}_{knm}^{\nu+1}\right) < c\frac{\gamma}{K_{\nu+1}^\tau}.$$

Proof. As is well known,

$$\mathrm{meas}(\mathcal{R}_k^{\nu+1}) = \mathrm{meas}\left(\left\{\xi \in \mathcal{O}_\nu : |\langle k, \omega_{\nu+1}(\xi)\rangle| < \frac{\gamma}{K_{\nu+1}^\tau}\right\}\right) < c\frac{\gamma}{K_{\nu+1}^\tau},$$

the set $\mathcal{R}_{kn}^{\nu+1} = \{\xi \in \mathcal{O}_\nu : \|(\langle k, \omega_{\nu+1}(\xi)\rangle I + A_n^{\nu+1})^{-1}\| > (\frac{K_{\nu+1}^\tau}{\gamma})^2\}$ is empty if $n > cK_{\nu+1}$, while if $n \leqslant cK_{\nu+1}$, from Lemmas 7.7, 7.8, we have

$$\mathrm{meas}(\mathcal{R}_{kn}^{\nu+1}) < c\frac{\gamma}{K_{\nu+1}^\tau}.$$

Now we consider the most complicated case $\mathcal{R}_{knm}^{\nu+1}$, $n \neq m$. Here we assume $|\langle k, \omega\rangle \pm \Omega_n \pm \Omega_m| \leqslant 1$, since if $|\langle k, \omega\rangle \pm \Omega_n \pm \Omega_m| \geqslant 1$, small-denominator conditions are automatically satisfied. Set

$$M = \langle k, \omega_{\nu+1}(\xi)\rangle I \pm A_n^{\nu+1} \otimes I \pm I \otimes A_m^{\nu+1}, n \neq m,$$

then 4th order derivative of det M with respect to ξ has lower bound cK_ν^4, therefore according to Lemma 7.7, 7.8,

$$\mathrm{meas}(\mathcal{R}_{knm}^{\nu+1}) \;=\; \mathrm{meas}\left\{\xi \in \mathcal{O}_\nu : \|M^{-1}\| > \left(\frac{K_{\nu+1}^\tau}{\gamma}\right)^4\right\}$$

$$\leqslant \operatorname{meas}\left\{\xi \in \mathcal{O}_\nu : |\det M| < c\left(\frac{\gamma}{K_{\nu+1}^\tau}\right)^4\right\}$$

$$\leqslant c\frac{\gamma}{K_{\nu+1}^\tau}.$$

In conclusion, Lemma 6.1 is obtained.

Lemma 6.2.

$$\operatorname{meas}(\mathcal{Q}_{kn}^{\nu+1}) < c\frac{\gamma}{K_{\nu+1}^{\tau-d-1}}.$$

Proof. We decompose $\mathcal{Q}_{kn}^{\nu+1}$ into two parts $\bar{\mathcal{Q}}_{kn}^{\nu+1}$, $\tilde{\mathcal{Q}}_{kn}^{\nu+1}$, where

$$\bar{\mathcal{Q}}_{kn}^{\nu+1} = \Bigg\{\xi \in \mathcal{O}_\nu : \|(\langle k, \omega_{\nu+1}(\xi)\rangle I + A_n^{\nu+1} \otimes I - I \otimes A_n^{\nu+1})^{-1}\| > \left(\frac{K_{\nu+1}^\tau}{\gamma}\right)^4, K_\nu < |k| \leqslant K_{\nu+1}, 2n \leqslant K_{\nu+1}\Bigg\},$$

$$\tilde{\mathcal{Q}}_{kn}^{\nu+1} = \Bigg\{\xi \in \mathcal{O}_\nu : \|(\langle k, \omega_{\nu+1}(\xi)\rangle I + A_n^{\nu+1} \otimes I - I \otimes A_n^{\nu+1})^{-1}\| > \left(\frac{K_{\nu+1}^\tau}{\gamma}\right)^4, 0 < |k| \leqslant K_\nu, K_\nu < 2n \leqslant K_{\nu+1}\Bigg\}.$$

For $k \neq 0$, $\Omega_n - \Omega_{-n} = 0$ at the beginning, Set

$$\begin{aligned} M &= \langle k, \omega_{\nu+1}(\xi)\rangle I + A_n^{\nu+1} \otimes I - I \otimes A_n^{\nu+1} \\ &= \langle k, \omega_{\nu+1}(\xi)\rangle I + (A_n^{\nu+1} - A_n^0) \otimes I - I \otimes (A_n^{\nu+1} - A_n^0), \end{aligned}$$

then 4th order derivative of det M with respect to ξ has lower bound $c > 0$, therefore according to Lemmas 7.7, 7.8,

$$\begin{aligned} \operatorname{meas}(\mathcal{Q}_{kn}^{\nu+1}) &\leqslant \operatorname{meas}(\bar{\mathcal{Q}}_{kn}^{\nu+1}) + \operatorname{meas}(\tilde{\mathcal{Q}}_{kn}^{\nu+1}) \\ &\leqslant \sum_{|k| \leqslant K_{\nu+1}, |n| \leqslant K_{\nu+1}} \operatorname{meas}\left\{\xi \in \mathcal{O}_\nu : \|M^{-1}\| > \left(\frac{K_{\nu+1}^\tau}{\gamma}\right)^4\right\} \end{aligned}$$

$$\leqslant \sum_{|k|\leqslant K_{\nu+1},|n|\leqslant K_{\nu+1}} \text{meas}\left\{\xi\in\mathcal{O}_\nu : |\det M| < c\left(\frac{\gamma}{K_{\nu+1}^{\tau}}\right)^4\right\}$$
$$\leqslant c \sum_{|k|\leqslant K_{\nu+1},|n|\leqslant K_{\nu+1}} \frac{\gamma}{K_{\nu+1}^{\tau}}$$
$$\leqslant c\frac{\gamma}{K_{\nu+1}^{\tau-d-1}}.$$

As a consequence, Lemma 6.2 follows.

Lemma 6.3.

$$\text{meas}\left(\bigcup_{\nu\geqslant 0}\mathcal{R}^{\nu+1}\right)$$
$$= \text{meas}\left[\bigcup_{\nu}\left(\bigcup_{K_\nu<|k|\leqslant K_{\nu+1},n,m}\left(\mathcal{R}_k^{\nu+1}\bigcup\mathcal{R}_{kn}^{\nu+1}\bigcup\mathcal{R}_{knm}^{\nu+1}\right)\bigcup\mathcal{Q}_{kn}^{\nu+1}\right)\right] < c\gamma.$$

Proof. We estimate

$$\text{meas}\left(\bigcup_{K_\nu<|k|\leqslant K_{\nu+1}}\bigcup_{n\neq m}\mathcal{R}_{knm}^{\nu+1}\right),$$

which is the most complicated case.

By Lemma 6.1, if $K_\nu < |k|\leqslant K_{\nu+1}$ and $n\neq m$, we have

$$\text{meas}\left(\bigcup_{K_\nu<|k|\leqslant K_{\nu+1},n\neq m}\mathcal{R}_{knm}^{\nu+1}\right) = \text{meas}\left(\bigcup_{\substack{K_\nu<|k|\leqslant K_{\nu+1}\\ n\neq m;n,m\leqslant cK_{\nu+1}}}\mathcal{R}_{knm}^{\nu+1}\right) < c\frac{\gamma}{K_{\nu+1}^{\tau-d-2}}. \tag{6.5}$$

Let $\tau > d+3$, by Lemma 6.1 and Lemma 6.2, we have

$$\text{meas}\left(\bigcup_{\nu\geqslant 0}\mathcal{R}^{\nu+1}\right) = \text{meas}\left[\bigcup_{\nu\geqslant 0}\left(\bigcup_{K_\nu<|k|\leqslant K_{\nu+1},n,m}\left(\mathcal{R}_k^{\nu+1}\bigcup\mathcal{R}_{kn}^{\nu+1}\right.\right.\right.$$

$$\times \bigcup \mathcal{R}^{v+1}_{knm}\Big) \bigcup \mathcal{Q}^{v+1}_{kn}\Bigg)\Bigg]$$

$$\leqslant c \sum_{v \geqslant 0} \frac{\gamma}{K_{v+1}} < c\gamma. \tag{6.6}$$

The proof of Lemma 6.3 is finished. □

7. Appendix

Lemma 7.1.

$$\|FG\|_{D_\rho(r,s)} \leqslant \|F\|_{D_\rho(r,s)} \|G\|_{D_\rho(r,s)}.$$

Proof. Since $(FG)_{kl\alpha\beta} = \sum_{k',l',\alpha',\beta'} F_{k-k',l-l',\alpha-\alpha',\beta-\beta'} G_{k'l'\alpha'\beta'}$, we have

$$\begin{aligned}
\|FG\|_{D_\rho(r,s)} &= \sup_{\substack{\|w\|_\rho < s \\ \|\bar{w}\|_\rho < s}} \sum_{k,l,\alpha,\beta} |(FG)_{kl\alpha\beta}| s^{2l} |w^\alpha||\bar{w}^\beta| e^{|k|\rho} \\
&\leqslant \sup_{\substack{\|w\|_\rho < s \\ \|\bar{w}\|_\rho < s}} \sum_{k,l,\alpha,\beta} \sum_{k',l',\alpha',\beta'} |F_{k-k',l-l',\alpha-\alpha',\beta-\beta'} G_{k'l'\alpha'\beta'}| s^{2l} |w^\alpha||\bar{w}^\beta| e^{|k|\rho} \\
&\leqslant \|F\|_{D_\rho(r,s)} \|G\|_{D_\rho(r,s)}
\end{aligned}$$

and the proof is finished. □

Lemma 7.2 (*Cauchy inequalities*).

$$\|F_{\theta_i}\|_{D_\rho(r-\sigma,s)} \leqslant \frac{c}{\sigma} \|F\|_{D_\rho(r,s)},$$

$$\|F_{I_i}\|_{D_\rho(r,\frac{1}{2}s)} \leqslant \frac{c}{s^2} \|F\|_{D_\rho(r,s)},$$

and

$$\|F_{w_n}\|_{D_\rho(r,\frac{1}{2}s)} \leqslant \frac{c}{s} \|F\|_{D_\rho(r,s)} e^{n\rho},$$

$$\|F_{\bar{w}_n}\|_{D_\rho(r,\frac{1}{2}s)} \leqslant \frac{c}{s} \|F\|_{D_\rho(r,s)} e^{|n|\rho}.$$

Let $\{\cdot,\cdot\}$ denote Poisson bracket of smooth functions

$$\{F,G\}=\sum_{i=1}^{d}\left(\frac{\partial F}{\partial I_i}\frac{\partial G}{\partial \theta_i}-\frac{\partial F}{\partial \theta_i}\frac{\partial G}{\partial I_i}\right)+i\sum_{n\neq n_1,\ldots,n_d}\left(\frac{\partial F}{\partial w_n}\frac{\partial G}{\partial \bar{w}_n}-\frac{\partial F}{\partial \bar{w}_n}\frac{\partial G}{\partial w_n}\right)$$

Lemma 7.3. *If*

$$\|X_F\|_{D_\rho(r,s)}<\varepsilon',\ \ \|X_G\|_{D_\rho(r,s)}<\varepsilon'',$$

then

$$\|X_{\{F,G\}}\|_{D_\rho(r-\sigma,\eta s)}<c\sigma^{-1}\eta^{-2}\varepsilon'\varepsilon'',\ \ \eta\ll 1.$$

Proof. According to the definition of the weighted norm for the vector field (see (3.15)),

$$\begin{aligned}\|X_{\{F,G\}}\|_{D_\rho(r-\sigma,\eta s)} &= \left\|\frac{d}{dI}\{F,G\}\right\|_{D_\rho(r-\sigma,\eta s)}+\frac{1}{\eta^2 s^2}\left\|\frac{d}{d\theta}\{F,G\}\right\|_{D_\rho(r-\sigma,\eta s)}\\ &\quad+\frac{1}{\eta s}\sum_{n\neq n_1,\ldots,n_d}\left(\left\|\frac{d}{dw_n}\{F,G\}\right\|_{D_\rho(r-\sigma,\eta s)}e^{|n|\rho}\right.\\ &\quad\left.+\left\|\frac{d}{d\bar{w}_n}\{F,G\}\right\|_{D_\rho(r-\sigma,\eta s)}e^{|n|\rho}\right).\end{aligned}$$

We only explicitly show two terms $\frac{1}{\eta s}\sum_{n\neq n_1,\ldots,n_d}\|\frac{d}{dw_n}\{F,G\}\|_{D_\rho(r-\sigma,\eta s)}e^{|n|\rho}$ and $\frac{1}{\eta^2 s^2}\|\frac{d}{d\theta}\{F,G\}\|_{D_\rho(r-\sigma,\eta s)}$, the remaining terms are achieved analogously.

$$\begin{aligned}\frac{d}{d\theta}\{F,G\} &= \langle F_{I\theta},G_\theta\rangle+\langle F_I,G_{\theta\theta}\rangle-\langle F_{\theta\theta},G_I\rangle-\langle F_\theta,G_{I\theta}\rangle\\ &\quad+\sum_{m\neq n_1,\ldots,n_d}\big(\langle F_{w_m\theta},iG_{\bar{w}_m}\rangle+\langle F_{w_m},iG_{\bar{w}_m\theta}\rangle\\ &\quad-\langle F_{\bar{w}_m\theta},iG_{w_m}\rangle-\langle F_{\bar{w}_m},iG_{w_m\theta}\rangle\big).\end{aligned}$$

By Lemmas 7.1 and 7.2,

$$\begin{aligned}\|\langle F_{I\theta},G_\theta\rangle\|_{D_\rho(r,\frac{1}{2}s)} &< \frac{c}{s^2}\|F_\theta\|\cdot\|G_\theta\|,\\ \|\langle F_I,G_{\theta\theta}\rangle\|_{D_\rho(r-\sigma,s)} &< \frac{c}{\sigma}\|F_I\|\cdot\|G_\theta\|,\end{aligned}$$

$$
\begin{aligned}
\|\langle F_{\theta\theta}, G_I\rangle\|_{D_\rho(r-\sigma,s)} &< \frac{c}{\sigma}\|F_\theta\|\cdot\|G_I\|,\\
\|\langle F_\theta, G_{I\theta}\rangle\|_{D_\rho(r-\sigma,\frac{1}{2}s)} &< \frac{c}{s^2}\|F_\theta\|\cdot\|G_\theta\|,\\
\|\langle F_{w_m\theta}, iG_{\bar{w}_m}\rangle\|_{D_\rho(r,\frac{1}{2}s)} &< \frac{c}{s}\|F_\theta\|\cdot\|G_{\bar{w}_m}\|e^{m\rho},\\
\|\langle F_{w_m}, iG_{\bar{w}_m\theta}\rangle\|_{D_\rho(r,\frac{1}{2}s)} &< \frac{c}{s}\|F_{w_m}\|\cdot\|G_\theta\|e^{m\rho},\\
\|\langle F_{\bar{w}_m\theta}, iG_{w_m}\rangle\|_{D_\rho(r,\frac{1}{2}s)} &< \frac{c}{s}\|F_\theta\|\cdot\|G_{w_m}\|e^{m\rho},\\
\|\langle F_{\bar{w}_m}, iG_{w_m\theta}\rangle\|_{D_\rho(r,\frac{1}{2}s)} &< \frac{c}{s}\|F_{\bar{w}_m}\|\cdot\|G_\theta\|e^{m\rho}.
\end{aligned}
$$

Then

$$
\begin{aligned}
\frac{1}{\eta^2 s^2}\left\|\frac{d}{d\theta}\{F,G\}\right\|_{D_\rho(r-\sigma,\eta s)} &< \frac{c}{\eta^2 s^2}\left\{\frac{1}{s^2}\|F_\theta\|\cdot\|G_\theta\|\right.\\
&\quad +\frac{1}{\sigma}\|F_I\|\cdot\|G_\theta\| + \frac{1}{\sigma}\|F_\theta\|\cdot\|G_I\|\\
&\quad +\frac{1}{s^2}\|F_\theta\|\cdot\|G_\theta\| + \frac{1}{s}\sum_{m\neq n_1,\ldots,n_d}\|F_\theta\|\cdot\|G_{\bar{w}_m}\|e^{m\rho}\\
&\quad +\frac{1}{s}\sum_{m\neq n_1,\ldots,n_d}\|F_{w_m}\|\cdot\|G_\theta\|e^{m\rho}\\
&\quad +\frac{1}{s}\sum_{m\neq n_1,\ldots,n_d}\|F_\theta\|\cdot\|G_{w_m}\|e^{m\rho}\\
&\quad \left.+\frac{1}{s}\sum_{m\neq n_1,\ldots,n_d}\|F_{\bar{w}_m}\|\cdot\|G_\theta\|e^{m\rho}\right\}\\
&< \frac{c}{\sigma\eta^2}\left\{\left(\frac{1}{s^2}\|F_\theta\|\right)\left(\frac{1}{s^2}\|G_\theta\|\right)\right.\\
&\quad +(\|F_I\|)\left(\frac{1}{s^2}\|G_\theta\|\right)\\
&\quad +\left(\frac{1}{s^2}\|F_\theta\|\right)(\|G_I\|)
\end{aligned}
$$

$$
\begin{aligned}
&+\left(\frac{1}{s^2}\|F_\theta\|\right)\left(\frac{1}{s^2}\|G_\theta\|\right)\\
&+\left(\frac{1}{s^2}\|F_\theta\|\right)\left(\frac{1}{s}\sum_{m\neq n_1,\ldots,n_d}\|G_{\bar{w}_m}\|e^{m\rho}\right)\\
&+\left(\frac{1}{s}\sum_{m\neq n_1,\ldots,n_d}\|F_{w_m}\|e^{m\rho}\right)\left(\frac{1}{s^2}\|G_\theta\|\right)\\
&+\left(\frac{1}{s^2}\|F_\theta\|\right)\left(\frac{1}{s}\sum_{m\neq n_1,\ldots,n_d}\|G_{w_m}\|e^{m\rho}\right)\\
&+\left(\frac{1}{s}\sum_{m\neq n_1,\ldots,n_d}\|F_{\bar{w}_m}\|e^{m\rho}\right)\left(\frac{1}{s^2}\|G_\theta\|\right)\Bigg\}\\
< {} & c\sigma^{-1}\eta^{-2}\|X_F\|_{D_\rho(r,s)}\|X_G\|_{D_\rho(r,s)}\\
< {} & c\sigma^{-1}\eta^{-2}\varepsilon'\varepsilon''.
\end{aligned}
$$

Furthermore,

$$
\begin{aligned}
\frac{d}{dw_n}\{F,G\} = {} & \langle F_{Iw_n},G_\theta\rangle+\langle F_I,G_{\theta w_n}\rangle-\langle F_{\theta w_n},G_I\rangle-\langle F_\theta,G_{Iw_n}\rangle\\
&+\sum_{m\neq n_1,\ldots,n_d}\big(\langle F_{w_m w_n},iG_{\bar{w}_m}\rangle+\langle F_{w_m},iG_{\bar{w}_m w_n}\rangle\\
&-\langle F_{\bar{w}_m w_n},iG_{w_m}\rangle-\langle F_{\bar{w}_m},iG_{w_m w_n}\rangle\big).
\end{aligned}
$$

By Lemmas 7.1 and 7.2,

$$
\begin{aligned}
\|\langle F_{Iw_n},G_\theta\rangle\|_{D_\rho(r,\frac{1}{2}s)} &< \frac{c}{s^2}\|F_{w_n}\|\cdot\|G_\theta\|,\\
\|\langle F_I,G_{\theta w_n}\rangle\|_{D_\rho(r-\sigma,s)} &< \frac{c}{\sigma}\|F_I\|\cdot\|G_{w_n}\|,\\
\|\langle F_{\theta w_n},G_I\rangle\|_{D_\rho(r-\sigma,s)} &< \frac{c}{\sigma}\|F_{w_n}\|\cdot\|G_I\|,\\
\|\langle F_\theta,G_{Iw_n}\rangle\|_{D_\rho(r-\sigma,\frac{1}{2}s)} &< \frac{c}{s^2}\|F_\theta\|\cdot\|G_{w_n}\|,\\
\|\langle F_{w_m w_n},iG_{\bar{w}_m}\rangle\|_{D_\rho(r,\frac{1}{2}s)} &< \frac{c}{s}\|F_{w_n}\|\cdot\|G_{\bar{w}_m}\|e^{m\rho},\\
\|\langle F_{w_m},iG_{\bar{w}_m w_n}\rangle\|_{D_\rho(r,\frac{1}{2}s)} &< \frac{c}{s}\|F_{w_m}\|\cdot\|G_{w_n}\|e^{m\rho},
\end{aligned}
$$

$$\|\langle F_{\bar{w}_m w_n}, iG_{w_m}\rangle\|_{D_\rho(r,\frac{1}{2}s)} < \frac{c}{s}\|F_{w_n}\|\cdot\|G_{w_m}\|e^{m\rho},$$

$$\|\langle F_{\bar{w}_m}, iG_{w_m w_n}\rangle\|_{D_\rho(r,\frac{1}{2}s)} < \frac{c}{s}\|F_{\bar{w}_m}\|\cdot\|G_{w_n}\|e^{m\rho}.$$

Then

$$\begin{aligned}
&\frac{1}{\eta s}\sum_{n\neq n_1,\ldots,n_d}\left\|\frac{d}{dw_n}\{F,G\}\right\|_{D_\rho(r-\sigma,\eta s)} e^{|n|\rho}\\
&< \frac{c}{\eta s}\sum_{n\neq n_1,\ldots,n_d}\left\{\frac{1}{s^2}\|F_{w_n}\|\cdot\|G_\theta\|\right.\\
&\quad +\frac{1}{\sigma}\|F_I\|\cdot\|G_{w_n}\| + \frac{1}{\sigma}\|F_{w_n}\|\cdot\|G_I\|\\
&\quad +\frac{1}{s^2}\|F_\theta\|\cdot\|G_{w_n}\| + \frac{1}{s}\sum_{m\neq n_1,\ldots,n_d}\|F_{w_n}\|\cdot\|G_{\bar{w}_m}\|e^{m\rho}\\
&\quad +\frac{1}{s}\sum_{m\neq n_1,\ldots,n_d}\|F_{w_m}\|\cdot\|G_{w_n}\|e^{m\rho}\\
&\quad +\frac{1}{s}\sum_{m\neq n_1,\ldots,n_d}\|F_{w_n}\|\cdot\|G_{w_m}\|e^{m\rho}\\
&\quad \left.+\frac{1}{s}\sum_{m\neq n_1,\ldots,n_d}\|F_{\bar{w}_m}\|\cdot\|G_{w_n}\|e^{m\rho}\right\} e^{|n|\rho}\\
&< \frac{c}{\sigma\eta}\left\{\left(\frac{1}{s}\sum_{n\neq n_1,\ldots,n_d}\|F_{w_n}\|e^{|n|\rho}\right)\left(\frac{1}{s^2}\|G_\theta\|\right)\right.\\
&\quad +(\|F_I\|)\left(\frac{1}{s}\sum_{n\neq n_1,\ldots,n_d}\|G_{w_n}\|e^{|n|\rho}\right)\\
&\quad +\left(\frac{1}{s}\sum_{n\neq n_1,\ldots,n_d}\|F_{w_n}\|e^{|n|\rho}\right)(\|G_I\|)\\
&\quad +\left(\frac{1}{s^2}\|F_\theta\|\right)\left(\frac{1}{s}\sum_{n\neq n_1,\ldots,n_d}\|G_{w_n}\|e^{|n|\rho}\right)
\end{aligned}$$

$$
\begin{aligned}
&+\left(\frac{1}{s}\sum_{n\neq n_1,\ldots,n_d}\|F_{w_n}\|e^{|n|\rho}\right)\left(\frac{1}{s}\sum_{m\neq n_1,\ldots,n_d}\|G_{\bar{w}_m}\|e^{m\rho}\right)\\
&+\left(\frac{1}{s}\sum_{m\neq n_1,\ldots,n_d}\|F_{w_m}\|e^{m\rho}\right)\left(\frac{1}{s}\sum_{n\neq n_1,\ldots,n_d}\|G_{w_n}\|e^{|n|\rho}\right)\\
&+\left(\frac{1}{s}\sum_{n\neq n_1,\ldots,n_d}\|F_{w_n}\|e^{|n|\rho}\right)\left(\frac{1}{s}\sum_{m\neq n_1,\ldots,n_b}\|G_{w_m}\|e^{m\rho}\right)\\
&+\left(\frac{1}{s}\sum_{m\neq n_1,\ldots,n_d}\|F_{\bar{w}_m}\|e^{m\rho}\right)\left(\frac{1}{s}\sum_{n\neq n_1,\ldots,n_d}\|G_{w_n}\|e^{|n|\rho}\right)\Bigg\}\\
&< c\sigma^{-1}\eta^{-1}\|X_F\|_{D_\rho(r,s)}\|X_G\|_{D_\rho(r,s)}\\
&< c\sigma^{-1}\eta^{-1}\varepsilon'\varepsilon''.
\end{aligned}
$$

Thus

$$\|X_{\{F,G\}}\|_{D_\rho(r-\sigma,\eta s)} < c\sigma^{-1}\eta^{-2}\varepsilon'\varepsilon''.$$

In particular, if $\eta \sim \varepsilon^{\frac{1}{3}}$, $\varepsilon', \varepsilon'' \sim \varepsilon$, we have $\|X_{\{F,G\}}\|_{D_\rho(r-\sigma,\eta s)} \sim \varepsilon^{\frac{4}{3}}$. □

Lemma 7.4. *If $f(\theta) = \sum_k f_k e^{i\langle k,\theta\rangle}$ with $|f_k| \leqslant |f|_r e^{-|k|r}$, then*

$$|f(\theta)|_{r-\sigma} \leqslant c\sigma^{-c}|f|_r,\ 0 < \sigma < r.$$

Lemma 7.5. *Let $\mathcal{O}$ be a compact set in $\mathbb{R}^d$ for which (4.1) holds. Suppose that $\omega(\xi)$ is C^4 Whitney-smooth functions in $\xi \in \mathcal{O}$ with derivative bounded by L and $f(\xi)$ is C^4 Whitney-smooth functions in $\xi \in \mathcal{O}$ with C^4_W norm bounded by L. Then*

$$g(\xi) \equiv \frac{f(\xi)}{\langle k, \omega(\xi)\rangle}$$

is C^4 Whitney-smooth in $\mathcal{O}$ with

$$\|g\|_{\mathcal{O}} < c\gamma^{-8}K^{8\tau+7}L.$$

Since in this paper, $A_n(\xi)$ (see (3.18)) are matrices of dimension at most two, then in the following lemmata, the most complicated case that $A_n(\xi)$ are matrices of two dimension is considered, the other cases are same. In addition, when $|\langle k,\omega\rangle + \Omega_n|$ and $|\langle k,\omega\rangle \pm \Omega_n \pm \Omega_m|$ are larger than one, then their inverse can be controlled by a standard Neumann series and corresponding decay estimates are preserved. Without loss of generality, we assume $|\langle k,\omega\rangle + \Omega_n| \leqslant 1$ and $|\langle k,\omega\rangle \pm \Omega_n \pm \Omega_m| \leqslant 1$.

Lemma 7.6. *Let $\mathcal{O}$ be a compact set in $\mathbb{R}^d$ for which (4.1) holds. Suppose that $A_n(\xi)$, $R_n(\xi)$ are, respectively, C^4 Whitney-smooth matrices and vectors, and $\omega(\xi)$ is C^4 Whitney-smooth function with derivatives bounded by L, then*

$$F_n(\xi) = M^{-1}R_n(\xi),$$

is C^4 Whitney-smooth with

$$\|F_n\|_{\mathcal{O}} \leqslant c\gamma^{-8}K^{8\tau+7}L.$$

Where M stands for either $\langle k,\omega\rangle I + A_n$ or $\langle k,\omega\rangle I \pm A_n \otimes I \pm I \otimes A_m$.

Lemma 7.7. *Let M be a non-singular matrix, then*

$$\{\xi : \|M^{-1}\| > h\} \subset \left\{\xi : |\det M| < \frac{c}{h}\right\}.$$

Lemma 7.8. *Suppose that $g(u)$ is a C^p function on the closure $\bar{I}$, where $I \subset \mathbb{R}$ is a finite interval. Let $I_h = \{u : |g(u)| < h\}$, $h > 0$, if for some constant $d > 0$, $|g^{(p)}(u)| \geqslant d$ for all $u \in I$, then $\operatorname{meas}(I_h) \leqslant ch^{\frac{1}{p}}$, where $c = 2(2+3+\cdots+p+d^{-1})$.*

For the proofs of the above lemmata, we refer the readers to the Appendix in [7].

Lemma 7.9. *Let $\mathcal{O}$ be a compact set in $\mathbb{R}^d$ for which (4.1) holds. Suppose that $A_n(\xi)$ and $\omega(\xi)$ are, respectively, C^4 Whitney-smooth matrices and function with derivatives bounded by L, and*

$$R_n(\xi) = (P_n(\xi), P_{-n}(\xi))^T, \quad |P_n(\xi)|_{\mathcal{O}}, |P_{-n}(\xi)|_{\mathcal{O}} \leqslant \varepsilon e^{-|n|\rho}.$$

Then

$$F_n(\xi) = (f_n(\xi), f_{-n}(\xi))^T = (\langle k,\omega\rangle I + A_n)^{-1}R_n$$

is C^4 Whitney-smooth with

$$|f_n|_{\mathcal{O}}, |f_{-n}|_{\mathcal{O}} \leqslant c\gamma^{-8}K^{8\tau+7}\varepsilon e^{-|n|\rho}.$$

Proof. Setting $M = \langle k, \omega\rangle I + A_n$. Let M^* denote the adjoint matrix of M, then

$$M^* = \begin{pmatrix} \langle k, \omega\rangle + \Omega_{-n} & -b_{((-n),n)} \\ -b_{(n,(-n))} & \langle k, \omega\rangle + \Omega_n \end{pmatrix},$$

consequently $M^{-1} = \frac{M^*}{\det M}$ and

$$f_n = \frac{1}{\det M}((\langle k, \omega\rangle + \Omega_{-n})P_n + (-b_{((-n),n)})P_{-n})$$

$$f_{-n} = \frac{1}{\det M}((-b_{(n,(-n))})P_n + (\langle k, \omega\rangle + \Omega_n)P_{-n}). \tag{7.1}$$

According to (4.3), (4.1), (7.1) and the assumptions, one can obtain

$$|f_n|_{\mathcal{O}}, |f_{-n}|_{\mathcal{O}} \leqslant c\gamma^{-8}K^{8\tau+7}\varepsilon e^{-|n|\rho},$$

lemma follows.

Lemma 7.10. *Let $\mathcal{O}$ be a compact set in $\mathbb{R}^d$ for which (4.1) holds. Suppose that $A_n(\xi)$ and $\omega(\xi)$ are, respectively, C^4 Whitney-smooth matrices and function with derivatives bounded by L, and*

$$R_{mn}(\xi) = (P_{nm}(\xi), P_{n(-m)}(\xi), P_{(-n)m}(\xi), P_{(-n)(-m)}(\xi))^T,$$

$$|P_{nm}(\xi)|_{\mathcal{O}}, |P_{(-n)(-m)}(\xi)|_{\mathcal{O}} \leqslant \varepsilon e^{-|n-m|\rho},$$

$$|P_{n(-m)}(\xi)|_{\mathcal{O}}, |P_{(-n)m}(\xi)|_{\mathcal{O}} \leqslant \varepsilon e^{-|n+m|\rho},$$

Then

$$F_{mn}(\xi) = (f_{nm}(\xi), f_{n(-m)}(\xi), f_{(-n)m}(\xi), f_{(-n)(-m)}(\xi))^T$$

$$= (\langle k, \omega\rangle I + A_n \otimes I - I \otimes A_m)^{-1} R_{mn}$$

is C^4 Whitney-smooth with

$$|f_{nm}|_{\mathcal{O}}, |f_{(-n)(-m)}|_{\mathcal{O}} \leqslant c\gamma^{-8}K^{8\tau+7}\varepsilon e^{-|n-m|\rho},$$

$$|f_{n(-m)}|_{\mathcal{O}}, |f_{(-n)m}|_{\mathcal{O}} \leqslant c\gamma^{-8}K^{8\tau+7}\varepsilon e^{-|n+m|\rho}.$$

Proof. Setting $M = \langle k, \omega\rangle I + A_n \otimes I - I \otimes A_m$. Let M^* denote the adjoint matrix of M,

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

By elementary calculation together with (4.3), (4.1) and the assumptions, one has

$$\begin{aligned} |m_{11}|_{\mathcal{O}}, |m_{22}|_{\mathcal{O}}, |m_{33}|_{\mathcal{O}}, |m_{44}|_{\mathcal{O}} &\leqslant cK^4, \\ |m_{12}|_{\mathcal{O}}, |m_{21}|_{\mathcal{O}}, |m_{34}|_{\mathcal{O}}, |m_{43}|_{\mathcal{O}} &\leqslant cK^4 e^{-2|m|\rho}, \\ |m_{13}|_{\mathcal{O}}, |m_{24}|_{\mathcal{O}}, |m_{31}|_{\mathcal{O}}, |m_{42}|_{\mathcal{O}} &\leqslant cK^4 e^{-2|n|\rho}, \\ |m_{14}|_{\mathcal{O}}, |m_{23}|_{\mathcal{O}}, |m_{32}|_{\mathcal{O}}, |m_{41}|_{\mathcal{O}} &\leqslant cK^4 e^{-2|n|\rho - 2|m|\rho}. \end{aligned} \tag{7.2}$$

Therefore

$$\begin{aligned} f_{nm} &= \frac{1}{\det\ M}(m_{11}P_{nm} + m_{12}P_{n(-m)} + m_{13}P_{(-n)m} + m_{14}P_{(-n)(-m)}) \\ f_{n(-m)} &= \frac{1}{\det\ M}(m_{21}P_{nm} + m_{22}P_{n(-m)} + m_{23}P_{(-n)m} + m_{24}P_{(-n)(-m)}) \\ f_{(-n)m} &= \frac{1}{\det\ M}(m_{31}P_{nm} + m_{32}P_{n(-m)} + m_{33}P_{(-n)m} + m_{34}P_{(-n)(-m)}) \\ f_{(-n)(-m)} &= \frac{1}{\det\ M}(m_{41}P_{nm} + m_{42}P_{n(-m)} + m_{43}P_{(-n)m} + m_{44}P_{(-n)(-m)}) \end{aligned}$$

According to (3.18), (4.3), (4.1), (7.2) and the assumptions, we have

$$\begin{aligned} |f_{nm}|_{\mathcal{O}}, |f_{(-n)(-m)}|_{\mathcal{O}} &\leqslant c\gamma^{-8}K^{8\tau+7}\varepsilon e^{-|n-m|\rho}, \\ |f_{n(-m)}|_{\mathcal{O}}, |f_{(-n)m}|_{\mathcal{O}} &\leqslant c\gamma^{-8}K^{8\tau+7}\varepsilon e^{-|n+m|\rho}. \end{aligned}$$

Thus the proof of Lemma is finished. □

Lemma 7.11. *If on* $D_\rho(r,s)$, $P = \sum_{\vec{n},\vec{\imath}} P_{\vec{n}}^{\vec{\imath}}(\theta, I) w_{\vec{n}}^{\vec{\imath}}$, $F = \sum_{\substack{\vec{n},\vec{\imath}' \\ |\langle\vec{\imath}',\vec{n}\rangle| \leqslant K}} F_{\vec{n}}^{\vec{\imath}'}(\theta, I) w_{\vec{n}}^{\vec{\imath}'}$ *with* $||P_{\vec{n}}^{\vec{\imath}}(\theta, I)|| \leqslant ce^{-|\langle\vec{\imath},\vec{n}\rangle|\rho}$, $||F_{\vec{n}}^{\vec{\imath}'}(\theta, I)|| \leqslant ce^{-|\langle\vec{\imath}',\vec{n}\rangle|\rho}$, *let* $\dot{\vec{n}}$ *denote vector* $\vec{n}$ *without* nth *entry and* $\dot{\vec{\imath}}$ *denote vector* $\vec{\imath}$ *without* nth *entry, then on* $D_\rho(r-\sigma, \frac{1}{2}s)$,

$$\{P, F\} = \sum_{\substack{\vec{n},\vec{\imath},\vec{\imath}'\\ |\langle \vec{\imath}',\vec{n}\rangle| \leqslant K}} \left(\left\langle \frac{\partial P_{\vec{n}}^{\vec{\imath}}(\theta, I)}{\partial I}, \frac{\partial F_{\vec{n}}^{\vec{\imath}'}(\theta, I)}{\partial \theta}\right\rangle - \left\langle \frac{\partial P_{\vec{n}}^{\vec{\imath}}(\theta, I)}{\partial \theta}, \frac{\partial F_{\vec{n}}^{\vec{\imath}'}(\theta, I)}{\partial I}\right\rangle\right) w_{\vec{n}}^{\vec{\imath}} w_{\vec{n}}^{\vec{\imath}'}$$
$$+i \sum_{\dot{\vec{n}},\dot{\vec{\imath}},\dot{\vec{\imath}}'} \sum_{\substack{n\\ |\langle \vec{\imath}',\vec{n}\rangle| \leqslant K}} P_{\vec{n}}^{\vec{\imath}}(\theta, I) F_{\vec{n}}^{\vec{\imath}'}(\theta, I) \left(\frac{\partial w_{\vec{n}}^{\vec{\imath}}}{\partial w_n}\frac{\partial w_{\vec{n}}^{\vec{\imath}'}}{\partial \bar{w}_n} - \frac{\partial w_{\vec{n}}^{\vec{\imath}}}{\partial \bar{w}_n}\frac{\partial w_{\vec{n}}^{\vec{\imath}'}}{\partial w_n}\right)$$

have decay property, i.e.,

$$\left\|\left\langle \frac{\partial P_{\vec{n}}^{\vec{\imath}}(\theta, I)}{\partial I}, \frac{\partial F_{\vec{n}}^{\vec{\imath}'}(\theta, I)}{\partial \theta}\right\rangle - \left\langle \frac{\partial P_{\vec{n}}^{\vec{\imath}}(\theta, I)}{\partial \theta}, \frac{\partial F_{\vec{n}}^{\vec{\imath}'}(\theta, I)}{\partial I}\right\rangle\right\| \leqslant c\sigma^{-1}s^{-2}e^{-|\langle \vec{\imath}+\vec{\imath}',\vec{n}\rangle|\rho}$$
$$\left\|\sum_{\substack{n\\ |\langle \vec{\imath}',\vec{n}\rangle| \leqslant K}} P_{\vec{n}}^{\vec{\imath}}(\theta, I) F_{\vec{n}}^{\vec{\imath}'}(\theta, I)\right\| \leqslant cKe^{-|\langle \dot{\vec{\imath}}+\dot{\vec{\imath}}',\vec{n}\rangle|\rho}$$

Proof. According to Lemma 7.2,

$$\left\|\frac{\partial P_{\vec{n}}^{\vec{\imath}}(\theta, I)}{\partial I}\right\|_{D_\rho(r,\frac{1}{2}s)} \leqslant \frac{c}{s^2}||P_{\vec{n}}^{\vec{\imath}}(\theta, I)||_{D_\rho(r,s)},$$
$$\left\|\frac{\partial F_{\vec{n}}^{\vec{\imath}'}(\theta, I)}{\partial \theta}\right\|_{D_\rho(r-\sigma,s)} \leqslant \frac{c}{\sigma}||F_{\vec{n}}^{\vec{\imath}'}(\theta, I)||_{D_\rho(r,s)},$$

and

$$\left\|\frac{\partial P_{\vec{n}}^{\vec{\imath}}(\theta, I)}{\partial \theta}\right\|_{D_\rho(r-\sigma,s)} \leqslant \frac{c}{\sigma}||P_{\vec{n}}^{\vec{\imath}}(\theta, I)||_{D_\rho(r,s)},$$
$$\left\|\frac{\partial F_{\vec{n}}^{\vec{\imath}'}(\theta, I)}{\partial I}\right\|_{D_\rho(r,\frac{1}{2}s)} \leqslant \frac{c}{s^2}||F_{\vec{n}}^{\vec{\imath}'}(\theta, I)||_{D_\rho(r,s)}.$$

Hence

$$\left\|\left\langle \frac{\partial P_{\vec{n}}^{\vec{\imath}}(\theta,I)}{\partial I}, \frac{\partial F_{\vec{n}}^{\vec{\imath}'}(\theta,I)}{\partial \theta}\right\rangle - \left\langle \frac{\partial P_{\vec{n}}^{\vec{\imath}}(\theta,I)}{\partial \theta}, \frac{\partial F_{\vec{n}}^{\vec{\imath}'}(\theta,I)}{\partial I}\right\rangle\right\|_{D_\rho(r-\sigma,\frac{1}{2}s)}$$
$$\leqslant c\sigma^{-1}s^{-2}e^{-|\langle \vec{\imath}+\vec{\imath}',\vec{n}\rangle|\rho}.$$

In addition, due to Poisson bracket, if $\imath_n$ is nth entry of vector $\vec{\imath}$, then $-\imath_n$ is nth entry of vector $\vec{\imath}'$, consequently,

$$\begin{aligned}\left\|\sum_{\substack{n\\ |\langle \vec{\imath}',\vec{n}\rangle|\leqslant K}} P_{\vec{n}}^{\vec{\imath}}(\theta,I)F_{\vec{n}}^{\vec{\imath}'}(\theta,I)\right\| &\leqslant c \sum_{\substack{n\\ |\langle \dot{\vec{\imath}}',\dot{\vec{n}}\rangle|-K\leqslant |n|\leqslant |\langle \dot{\vec{\imath}}',\dot{\vec{n}}\rangle|+K}} e^{-|\langle \vec{\imath},\vec{n}\rangle|\rho}e^{-|\langle \vec{\imath}',\vec{n}\rangle|\rho}\\
&\leqslant c \sum_{\substack{n\\ |\langle \dot{\vec{\imath}}',\dot{\vec{n}}\rangle|-K\leqslant |n|\leqslant |\langle \dot{\vec{\imath}}',\dot{\vec{n}}\rangle|+K}} e^{-|\langle \vec{\imath}+\vec{\imath}',\vec{n}\rangle|\rho}\\
&\leqslant c \sum_{\substack{n\\ |\langle \dot{\vec{\imath}}',\dot{\vec{n}}\rangle|-K\leqslant |n|\leqslant |\langle \dot{\vec{\imath}}',\dot{\vec{n}}\rangle|+K}} e^{-|\langle \dot{\vec{\imath}}+\dot{\vec{\imath}}',\vec{n}\rangle|\rho}\\
&\leqslant cKe^{-|\langle \dot{\vec{\imath}}+\dot{\vec{\imath}}',\vec{n}\rangle|\rho}.\end{aligned}$$

Thus Lemma 7.11 follows. □

Acknowledgments

We would like to thank Shui-Nee Chow for warm-hearted helping and worthwhile suggestion. We would also like to thank the referee for her or his valuable remarks and suggestions which have helped to improve the quality of the paper.

References

[1] D. Bambusi, Birkhoff normal form for some nonlinear PDEs, Commun. Math. Phys. 234 (2003) 253–285.
[2] D. Bambusi, An averaging theorem for quasilinear Hamiltonian PDEs, Ann. Henri Poincare 4 (2003) 685–712.
[3] J. Bourgain, Construction of quasi-periodic solutions for Hamiltonian perturbations of linear equations and applications to nonlinear PDE, International Mathematics Research Notices, 1994, pp. 475–497.
[4] J. Bourgain, Construction of periodic solutions of nonlinear wave equations in higher dimension, Geom. Funct. Anal. 5 (1995) 629–639.
[5] J. Bourgain, Quasiperiodic solutions of Hamiltonian perturbations of 2D linear Schrödinger equations, Ann. Math. 148 (1998) 363–439.

[6] J. Bourgain, Green's function estimates for lattice Schrödinger operators and applications, 2000.
[7] L. Chierchia, J. You, KAM tori for 1D nonlinear wave equations with periodic boundary conditions, Commun. Math. Phys. 211 (2000) 498–525.
[8] W. Craig, C.E. Wayne, Newton's method and periodic solutions of nonlinear wave equations, Comm. Pure. Appl. Math. 46 (1993) 1409–1498.
[9] J. Fröhlich, T. Spencer, Absence of diffusion in the Anderson tight binding model for large disorder or lower energy, Commun. Math. Phys. 88 (1983) 151–184.
[10] S.B. Kuksin, Nearly integrable infinite dimensional Hamiltonian systems, Lecture Notes in Mathematics, vol. 1556, Springer, Berlin, 1993.
[11] S.B. Kuksin, J. Pöschel, Invariant Cantor manifolds of quasiperiodic oscillations for a nonlinear Schrödinger equation, Ann. Math. 143 (1996) 149–179.
[12] J. Pöschel, Quasi-periodic solutions for a nonlinear wave equation, Comm. Math. Helv. 71 (1996) 269–296.
[13] J. Pöschel, A KAM theorem for some nonlinear partial differential equations, Ann. Sc. Norm. Sup. Pisa CI. Sci. 23 (1996) 119–148.
[14] J. Pöschel, E. Trubowitz, Inverse Spectral Theory, Academic Press, Inc., New York, 1987.
[15] C.E. Wayne, Periodic and quasi-periodic solutions for nonlinear wave equations via KAM theory, Commun. Math. Phys. 127 (1990) 479–528.
[16] J. You, Perturbation of lower dimensional tori for Hamiltonian systems, J. Differential Equations 152 (1999) 1–29.