# Quasi-periodic Solutions for One-dimensional Nonlinear Lattice Schrödinger Equation with Tangent Potential*

Jiansheng Geng, Zhiyan Zhao
Department of Mathematics
Nanjing University, Nanjing 210093, P.R.China
Email: jgeng@nju.edu.cn; zyqiao1985@gmail.com

**Abstract**

In this paper, we construct time quasi-periodic solutions for the nonlinear lattice Schrödinger equation

$$\mathrm{i}\dot{q}_n + \epsilon(q_{n+1} + q_{n-1}) + \tan \pi(n\tilde{\alpha} + x)q_n + \epsilon|q_n|^2 q_n = 0, \quad n \in \mathbb{Z},$$

where $\tilde{\alpha}$ satisfies a certain Diophantine condition and $x \in \mathbb{R}/\mathbb{Z}$. We prove that for $\epsilon$ sufficiently small, the equation admits a family of small-amplitude time quasi-periodic solutions for "most" of $x$ belonging to $\mathbb{R}/\mathbb{Z}$.

## 1 Introduction and main result

During the past two decades or so, the celebrated KAM (Kolmogorov-Arnold-Moser) theory and the CWB(Craig-Wayne-Bourgain) method were successfully generalized to infinite dimensional Hamiltonian systems, motivated by the construction of quasi-periodic solutions for Hamiltonian partial differential equations(see [1, 11, 16, 17, 19, 20, 21, 22, 23, 24, 25, 26, 28, 29, 30, 31, 32, 35, 36, 37, 41, 43, 44, 45, 46] for the KAM method, and [4, 5, 6, 7, 8, 9, 13] for the CWB method). In this paper, we focus on the nonlinear lattice Schrödinger equation

$$\mathrm{i}\dot{q}_n + \epsilon(q_{n+1} + q_{n-1}) + \tan \pi(n\tilde{\alpha} + x)q_n + \epsilon|q_n|^2 q_n = 0, \quad n \in \mathbb{Z}, \tag{1.1}$$

where $\tilde{\alpha} \in \mathbb{R}$ satisfies the Diophantine condition, i.e., there exist constants $\tilde{\tau} > 1$, $\tilde{\gamma} > 0$ such that

$$|n\tilde{\alpha}|_1 \geq \frac{\tilde{\gamma}}{|n|^{\tilde{\tau}}}, \quad n \neq 0, \tag{1.2}$$

with $|x|_1$ the absolute value of $x$ modulo 1 defined so that $0 \leq |x|_1 \leq \frac{1}{2}$.

We start with some physical motivation for studying Equation (1.1). The time-dependent Maryland model, i.e., the linear Schrödinger equation

$$\mathrm{i}\dot{q}_n + \epsilon(q_{n+1} + q_{n-1}) + \tan \pi(n\tilde{\alpha} + x)q_n = 0,$$

*This work is partially supported by NSFC grant 11271180. It is also partially supported by a Project Funded by the Priority Academic Program Development of Jiangsu Higher Education Institutions.

describes the motion of particles or waves in some quasi-crystal material, where $n$ is the primary lattice site index, the Diophantine number $\tilde{\alpha} \in \mathbb{R}$ is some ratio between the wavenumbers of two lattices, $x \in \mathbb{R}/\mathbb{Z}$ is an arbitrary phase, and $q_n$ is a complex variable whose modulus square gives the probability of finding a particle at the lattice site $n$. It is important in the study of Bose-Einstein condensation and nonlinear optics. When we consider the interactions(nonlinearities) additionally, we can start from the Gross-Pitaevskii(GP) equation[27, 34] in Hartree-Fock theory, and get a generalized Maryland model which includes an additional nonlinear term that represents the mean-field interaction. The Hamiltonian is

$$H = \sum_{n\in\mathbb{Z}} \left[ \epsilon(q_{n+1}\bar{q}_n + \bar{q}_{n+1}q_n) + \tan\pi(n\tilde{\alpha} + x)|q_n|^2 + \frac{1}{2}\epsilon|q_n|^4 \right],$$

and the equation of motion is generated by $\mathrm{i}\dot{q}_n = -\frac{\partial H}{\partial \bar{q}_n}$, yielding the nonlinear Schrödinger equation

$$\mathrm{i}\dot{q}_n + \epsilon(q_{n+1} + q_{n-1}) + \tan\pi(n\tilde{\alpha} + x)q_n + \epsilon|q_n|^2 q_n = 0, \quad n \in \mathbb{Z},$$

which can be considered as the GP equation on a discretized lattice. Similar versions of a discretized GP equation have been already used to investigate the dynamics of condensates in different situations(see, for instance, [40]). Other physical motivations can be found in Section 4.2 of Reference [18].

As a mathematical model, the spectral property of the linear problem has been thoroughly studied(see [2, 12, 39] and Section 10.3 of [14]). Bellissard-Lima-Scoppola[2] investigated the linear operator on $\ell^2(\mathbb{Z}^d)$,

$$(L_x q)_n = -\epsilon \sum_{m\in\mathbb{Z}^d} a(n-m)q_m + \tan\pi(\langle n, \tilde{\alpha}\rangle + x)q_n,$$

where $\tilde{\alpha} \in \mathbb{R}^d$ is a given Diophantine vector, and $a(n)$ decays exponentially with $|n|$. Clearly, there exist $\tilde{\gamma} > 0$ and $\tilde{\tau} > d$ such that

$$|\tan\pi(\langle m, \tilde{\alpha}\rangle + x) - \tan\pi(\langle n, \tilde{\alpha}\rangle + x)| \geq \frac{\tilde{\gamma}}{|m-n|^{\tilde{\tau}}}, \quad \forall m \neq n, \tag{1.3}$$

They have shown that, if $\epsilon$ is small enough, then for a.e. $x \in \mathbb{R}/\mathbb{Z}$, $L_x$ has only pure point spectrum with exponentially localized states, and a dense set of eigenvalues in the real line. This estimate (1.3) is exactly the condition needed in a perturbation theory to avoid a tunneling effect at large distance. Thanks to this work, we can diagonalize the linear Schrödinger operator to avoid the difficulty brought by the coupling term $\epsilon(q_{n+1} + q_{n-1})$ in Equation (1.1). We shall give a precise statement of this result of [2] before proof of the main theorem.

From the perspective of Hamiltonian PDE's, there are also some related works. Craig-Wayne[13] retrieved the origination of the KAM method - Newtonian iteration method together with the Lyapunov-Schmidt decomposition which involves the Green's function analysis and the control of the inverse of infinite matrices with small eigenvalues. They succeeded in constructing periodic solutions of the one-dimensional semi-linear wave equations with periodic boundary conditions. Bourgain[4, 5, 6, 7, 8] further developed the

Craig-Wayne's method and proved the existence of quasi-periodic solutions for Hamiltonian PDE's in higher dimensional spaces with Dirichlet boundary conditions or periodic boundary conditions. In a similar way, Bourgain-Wang[9] constructed time quasi-periodic solutions to the nonlinear random Schrödinger equation

$$\mathrm{i}\dot{q}_n = \epsilon(\Delta q)_n + V_n q_n + \delta|q_n|^{2p} q_n \ \ (p > 0), \quad n \in \mathbb{Z}^d, \ \ t \in \mathbb{R},$$

with $\epsilon, \delta$ sufficiently small, and $\{V_j\}_{j\in\mathbb{Z}^d}$, the potential, is a family of time-independent independent identically distributed(i.i.d.) random variables. We point out that the Craig-Wayne-Bourgain's method allows one to avoid explicitly using the Hamiltonian structure of the systems. We will not introduce their approach in detail. The reader is referred to [4, 5, 6, 7, 8, 9, 13].

Comparing with Craig-Wayne-Bourgain's approach, the KAM approach has its own advantages. Besides obtaining the existence results, it allows one to construct a local normal form in a neighborhood of the obtained solutions, and this is useful for better understanding of the dynamics. For example, one can obtain the linear stability and zero Lyapunov exponents. The KAM method was successfully applied by Kuksin[29] and Wayne[41] (see also [30, 32, 36, 37]) to, as typical examples, one-dimensional semi-linear Schrödinger equations

$$\mathrm{i}u_t - u_{xx} + mu = f(u),$$

and wave equations

$$u_{tt} - u_{xx} + mu = f(u),$$

with Dirichlet boundary conditions. Geng-You[21, 22] proved that the higher-dimensional nonlinear beam equations and nonlocal Schrödinger equations admit small-amplitude linearly stable quasi-periodic solutions. The breakthrough of constructing quasi-periodic solutions for more interesting higher dimensional Schrödinger equation by modified KAM method was made recently by Eliasson-Kuksin [17]. They proved that the higher dimensional nonlinear Schrödinger equations admit small-amplitude linearly stable quasi–periodic solutions. Recently, quasi-periodic solutions of two-dimensional cubic Schrödinger equation

$$\mathrm{i}u_t - \triangle u + |u|^2 u = 0, \qquad x \in \mathbb{T}^2, \ t \in \mathbb{R},$$

with periodic boundary conditions are obtained by Geng-Xu-You [19]. By appropriately choosing tangential sites $\{i_1, \cdots, i_b\} \in \mathbb{Z}^2$, the authors proved that the above nonlinear Schrödinger equation admits a family of small-amplitude quasi-periodic solutions.

However, all the KAM results mentioned above fail in dealing with the dense point spectrum. In this paper, we try to attack this case. Concretely, we consider Equation (1.1) as a model, note that $\{\tan \pi(n\tilde{\alpha} + x)\}_{n\in\mathbb{Z}}$ is dense on the real line when $\tilde{\alpha}$ is an irrational number. We shall give an abstract KAM theorem which can be applied to an equation deriving from Equation (1.1), via some suitable change of variables, and use the theorem to construct the quasi-periodic solutions for Equation (1.1). To establish the KAM theorem, we have to impose further restrictions both on the unperturbed part and on the perturbation. In the existent infinite dimensional KAM theorems, e.g., Kuksin[29], Pöschel[37], Wayne[41], Eliasson-Kuksin[17], Geng-Viveros-Yi[26], Geng-Xu-You[19], some assumptions on the regularity of the frequencies and the perturbation are required (See $(\mathbf{A1}) - (\mathbf{A4})$ in Section 2). In addition, we also assume that the perturbation has a special form defined in $(\mathbf{A5})$ in Section 2, which is called gauge invariance. In fact, the condition (**A5**) means the $l^2$ norm $(\sum |q_n|^2)^{\frac{1}{2}}$ is a conserved quantity. With such a special form, our proof is simplified, compared with previous KAM theorems, because some terms, which can not be eliminated easily, are zero(see (4.2) in Subsection 4.1).

Now we are going to state our main result. Consider the lattice Schrödinger equation

$$\mathrm{i}\dot{q}_n + \epsilon(q_{n+1} + q_{n-1}) + \tan\pi(n\tilde{\alpha} + x)q_n + \epsilon|q_n|^2 q_n = 0, \quad n \in \mathbb{Z}, \tag{1.4}$$

where $\tilde{\alpha}$ satisfies the Diophantine condition (1.2), and $x$ belongs to the full-measure subset

$$\mathcal{X} := \left\{ x \in \mathbb{R}/\mathbb{Z} : n\tilde{\alpha} + x \neq \frac{1}{2}, \ \ \forall n \in \mathbb{Z} \right\}.$$

**Theorem 1** *For $\mathcal{J} = \{n_1, \cdots, n_b\} \subset \mathbb{Z}$, $b > 1$, and $\kappa > 0$, given an initial datum $q_{\mathbb{Z}}(0) = (q_n(0))_{n\in\mathbb{Z}}$ supported in $\mathcal{J}$ with $q_{\mathbb{Z}}(0) \in \epsilon^{\frac{\kappa}{2}} \cdot [0,1]^b$. There is a sufficiently small positive number $\epsilon_* = \epsilon_*(\tilde{\alpha}, \kappa, \mathcal{J})$, such that if $0 < \epsilon < \epsilon_*$, one can find a subset $\mathcal{X}_\epsilon$ of $\mathcal{X}$ with*

$$\mathrm{mes}(\mathcal{X} \setminus \mathcal{X}_\epsilon) < \epsilon^{\vartheta} \ \ for \ \ some \ \ 0 < \vartheta < 1$$

*such that the following holds for fixed $x \in \mathcal{X}_\epsilon$.*

*There exists a Cantor set $\mathcal{O}_\epsilon = \mathcal{O}_\epsilon(x) \subset [0,1]^b$ with*

$$|[0,1]^b \setminus \mathcal{O}_\epsilon| \to 0 \ \ as \ \ \epsilon \to 0,^1$$

*such that if $q_{\mathbb{Z}}(0) \in \epsilon^{\frac{\kappa}{2}} \cdot \mathcal{O}_\epsilon$, $q_{\mathbb{Z}}(t) = (q_n(t))_{n\in\mathbb{Z}} \in \ell^1$ is a small-amplitude b-frequency quasi-periodic solution of Equation (1.4), with the frequencies slightly deformed from*

$$(\tan\pi(n_1\tilde{\alpha} + x), \ \cdots, \ \tan\pi(n_b\tilde{\alpha} + x)).$$

**Remark 1.1** *The nonlinear term $\epsilon|q_n|^2 q_n$ in Equation (1.4) has its physical meaning, but its special form in the Hamiltonian, i.e., $\epsilon|q_n|^4$, is not essential, as long as it is finite-range or sufficiently short-range and of bounded degree, for example, $\epsilon|q_n|^4$ can be replaced by*

$$\epsilon|q_n|^4 + \epsilon|q_n|^2 \bar{q}_n q_{n+1} + \epsilon|q_n|^2 q_n \bar{q}_{n+1}$$

*in the finite-range case and*

$$\epsilon|q_n|^2 \sum_k e^{-\varrho|n-k|}|q_k|^4$$

*in the short-range case.*

**Remark 1.2** *In the above theorem, we construct time quasi–periodic solutions for a corresponding appropriate set of small initial data with compact support, which means that for such initial data, the corresponding solutions are bounded in $\ell^1$. Clearly such initial data are a subset of all small initial data. It should be very interesting whether one can prove the similar result like that in [10, 42].*

The rest of this paper is organized as follows. We present the abstract KAM theorem, which can be applied to an equation which conjugates with Equation (1.1) in Section 2, and prove Theorem 1 via this KAM theorem in Section 3. In Section 4, we give the details for one step of the KAM iteration. The proof of the abstract KAM theorem is completed in Section 5 by an iteration lemma, giving a convergence result, and finally conducting the measure estimates of the remaining parameters.

[1] Hereafter, we use the symbol $|\mathcal{O}|$ to denote the Lebesgue measure of $\mathcal{O} \subset \mathbb{R}^b$.

## 2 An abstract KAM theorem

### 2.1 Function space norms and gauge invariance

Given $\mathbb{Z}_1 \subset \mathbb{Z}$, and $d, \rho > 0$, let $\ell^1_{d,\rho}(\mathbb{Z}_1)$ be the space of summable complex-valued sequences $q = (q_n)_{n\in\mathbb{Z}_1}$, with the norm

$$\|q\|_{d,\rho} := \sum_{n\in\mathbb{Z}_1} |q_n|\langle n\rangle^d e^{\rho|n|} < \infty,$$

where $\langle n\rangle := \sqrt{1+n^2}$. For $r, s > 0$, let $\mathcal{D}_{d,\rho}(r,s)$ be the complex $b$-dimensional neighborhood of $\mathbb{T}^b \times \{0\} \times \{0\} \times \{0\}$ in $\mathbb{T}^b \times \mathbb{R}^b \times \ell^1_{d,\rho}(\mathbb{Z}_1) \times \ell^1_{d,\rho}(\mathbb{Z}_1)$, i.e.,

$$\mathcal{D}_{d,\rho}(r,s) := \{(\theta, I, q, \bar{q}) : |\mathrm{Im}\theta| = |\mathrm{Im}(\theta_1, \cdots, \theta_b)| < r,\ |I| < s^2,\ \|q\|_{d,\rho} = \|\bar{q}\|_{d,\rho} < s\},$$

where $|\cdot|$ denotes the $\ell^1$-norm of complex vectors.

Given a real-analytic function $F(\theta, I, q, \bar{q}; \xi)$ on $\mathcal{D} = \mathcal{D}_{d,\rho}(r,s)$, $C^1_W$(i.e., $C^1$ in the sense of Whitney) dependent on a parameter $\xi \in \mathcal{O}$,[2] a closed region in $\mathbb{R}^b$. We expand $F$ into the Taylor-Fourier series with respect to $\theta, I, q, \bar{q}$:

$$F(\theta, I, q, \bar{q}; \xi) = \sum_{\alpha,\beta} F_{\alpha\beta}(\theta, I; \xi) q^\alpha \bar{q}^\beta,$$

where, for multi-indices $\alpha := \sum_{n\in\mathbb{Z}_1} \alpha_n e_n$, $\beta := \sum_{n\in\mathbb{Z}_1} \beta_n e_n$, $\alpha_n, \beta_n \in \mathbb{N}$, with finitely many non-vanishing components,

$$F_{\alpha\beta}(\theta, I; \xi) = \sum_{k\in\mathbb{Z}^b,\, l\in\mathbb{N}^b} F_{kl\alpha\beta}(\xi) I^l e^{\mathrm{i}\langle k,\theta\rangle}, \quad q^\alpha \bar{q}^\beta = \prod_{(\alpha_n,\beta_n)\neq(0,0)} q_n^{\alpha_n} \bar{q}_n^{\beta_n}.$$

(Here $e_n$ denotes the vector with the $n^{\mathrm{th}}$ component being 1 and the other components being zero.)

**Definition 2.1** *For each non-zero multi–index $(\alpha, \beta) = (\alpha_n, \beta_n)_{n\in\mathbb{Z}_1}$, $\alpha_n, \beta_n \in \mathbb{N}$, with finitely many non-vanishing components, we define*

$$\mathrm{supp}(\alpha, \beta) := \{n \in \mathbb{Z}_1 : (\alpha_n, \beta_n) \neq (0,0)\},$$

$$n^+_{\alpha\beta} := \max\{n \in \mathrm{supp}(\alpha,\beta)\}, \quad n^-_{\alpha\beta} := \min\{n \in \mathrm{supp}(\alpha,\beta)\}, \quad n^*_{\alpha\beta} := \max\{|n^+_{\alpha\beta}|, |n^-_{\alpha\beta}|\},$$

*and $|\alpha| := \sum_{n\in\mathbb{Z}_1} \alpha_n$, $|\beta| := \sum_{n\in\mathbb{Z}_1} \beta_n$.*

*In particular, for $|\alpha| = |\beta| = 0$, define $n^+_{\alpha\beta} = n^-_{\alpha\beta} = n^*_{\alpha\beta} = 0$.*

**Remark 2.1** *The notations above are closely related to the notations of support and diameter for the monomials in [10] and [42]. The decay properties of functions on phase space in terms of the index $n$ are important to this study.*

[2] In the rest of the paper, all dependencies on $\xi$ are assumed of class $C^1_W$, thus all derivatives with respective to the parameter $\xi \in \mathcal{O}$ will be interpreted in this sense.

With $|\partial_\xi F_{kl\alpha\beta}| := \sum_{i=1}^b |\partial_{\xi_i} F_{kl\alpha\beta}|$ and $|F_{kl\alpha\beta}|_{\mathcal{O}} := \sup_{\xi\in\mathcal{O}} (|F_{kl\alpha\beta}| + |\partial_\xi F_{kl\alpha\beta}|)$, let

$$\|F_{\alpha\beta}\|_{\mathcal{O}} := \sum_{k,l} |F_{kl\alpha\beta}|_{\mathcal{O}}\ |I^l| e^{|k||\mathrm{Im}\theta|}, \quad \|F\|_{\mathcal{O}} := \sum_{k,l,\alpha,\beta} |F_{kl\alpha\beta}|_{\mathcal{O}}\ |I^l| e^{|k||\mathrm{Im}\theta|}\ |q^\alpha||\bar{q}^\beta|.$$

Define the weighted norm of $F$ as

$$\|F\|_{\mathcal{D},\mathcal{O}} := \sup_{\mathcal{D}} \|F\|_{\mathcal{O}}.^3 \tag{2.1}$$

For the Hamiltonian vector field $X_F = (\partial_I F, -\partial_\theta F, (-\mathrm{i}\partial_{q_n} F)_{n\in\mathbb{Z}_1}, (\mathrm{i}\partial_{\bar{q}_n} F)_{n\in\mathbb{Z}_1})$ associated with $F$ on $\mathcal{D}\times\mathcal{O}$, define its norm by

$$\|X_F\|_{\mathcal{D},\mathcal{O}} := \|\partial_I F\|_{\mathcal{D},\mathcal{O}} + \frac{1}{s^2}\|\partial_\theta F\|_{\mathcal{D},\mathcal{O}} + \sup_{\mathcal{D}} \frac{1}{s} \sum_{n\in\mathbb{Z}_1} (\|\partial_{q_n} F\|_{\mathcal{O}} + \|\partial_{\bar{q}_n} F\|_{\mathcal{O}})\, \langle n\rangle^d e^{|n|\rho}.$$

Sometimes, for the sake of notational simplification, we shall not write the subscript $\mathcal{O}$ in the norms defined above if it is obvious enough.

In what follows in the formulations and proofs of various assertions, we shall encounter absolute constants depending on the Hamiltonian, the dimension and so on. All such constants will be denoted by $c, c_1, c_2, \cdots$, and sometimes even different constants will be denoted by the same symbol.

For $d, \rho, r, s > 0$, let $F, G$ be two real-analytic functions on $\mathcal{D} = \mathcal{D}_{d,\rho}(r,s)$, both of which $C^1_W$ depend on the parameter $\xi\in\mathcal{O}$.

**Lemma 2.1** *The norm $\|\cdot\|_{\mathcal{D},\mathcal{O}}$ satisfies the Banach algebraic property, i.e.,*

$$\|FG\|_{\mathcal{D},\mathcal{O}} \le \|F\|_{\mathcal{D},\mathcal{O}} \|G\|_{\mathcal{D},\mathcal{O}}.$$

*Proof:* Since $(FG)_{kl\alpha\beta} = \sum_{\substack{\check{k}+\hat{k}=k,\ \check{l}+\hat{l}=l \\ \check{\alpha}+\hat{\alpha}=\alpha,\ \check{\beta}+\hat{\beta}=\beta}} F_{\check{k}\check{l}\check{\alpha}\check{\beta}} G_{\hat{k}\hat{l}\hat{\alpha}\hat{\beta}}$, we have that

$$\begin{aligned}
\|FG\|_{\mathcal{D},\mathcal{O}} &= \sup_{\mathcal{D}} \sum_{k,l,\alpha,\beta} |(FG)_{kl\alpha\beta}|_{\mathcal{O}} |q^\alpha||\bar{q}^\beta||I^l| e^{|k||\mathrm{Im}\theta|} \\
&\le \sup_{\mathcal{D}} \sum_{k,l,\alpha,\beta} \sum_{\substack{\check{k}+\hat{k}=k,\ \check{l}+\hat{l}=l \\ \check{\alpha}+\hat{\alpha}=\alpha,\ \check{\beta}+\hat{\beta}=\beta}} |F_{\check{k}\check{l}\check{\alpha}\check{\beta}} G_{\hat{k}\hat{l}\hat{\alpha}\hat{\beta}}|_{\mathcal{O}} |q^\alpha||\bar{q}^\beta||I^l| e^{(|\check{k}|+|\hat{k}|)|\mathrm{Im}\theta|} \\
&\le \|F\|_{\mathcal{D},\mathcal{O}} \|G\|_{\mathcal{D},\mathcal{O}}.
\end{aligned}$$

■

**Lemma 2.2 (Generalized Cauchy Inequalities)** *The various components of the Hamiltonian vector field $X_F$ satisfy: for any $0 < r' < r$, $0 < \rho' < \rho$,*

$$\|\partial_\theta F\|_{\mathcal{D}_{d,\rho}(r',s)} \le \frac{c}{r-r'} \|F\|_{\mathcal{D}},$$

$$\|\partial_I F\|_{\mathcal{D}_{d,\rho}(r,\frac{s}{2})} \le \frac{c}{s^2} \|F\|_{\mathcal{D}},$$

$$\sup_{\mathcal{D}_{d,\rho}(r,\frac{s}{2})} \sum_{n\in\mathbb{Z}_1} (\|\partial_{q_n} F\|_{\mathcal{O}} + \|\partial_{\bar{q}_n} F\|_{\mathcal{O}})\, \langle n\rangle^d e^{|n|\rho'} \le \frac{c}{s(\rho-\rho')} \|F\|_{\mathcal{D}}.$$

[3] In the case of a vector-valued function $F : \mathcal{D}\times\mathcal{O} \to \mathbb{C}^b (b < +\infty)$, the norm can be defined as $\|F\|_{\mathcal{D},\mathcal{O}} := \sum_{i=1}^b \|F_i\|_{\mathcal{D},\mathcal{O}}$.

*Proof:* We only prove the third inequality, with others shown analogously. Given $\omega \in \ell^1_{d,\rho}(\mathbb{Z}_1) \setminus \{0\}$, $f(t) = F(\cdot, \cdot, q + t\omega, \cdot)$ is an analytic function on the the complex disc $\{z \in \mathbb{C} : |z| < \frac{s}{\|\omega\|_{d,\rho}}\}$. Hence

$$|f'(0)| = \left| \sum_{n \in \mathbb{Z}_1} \omega_n \cdot \partial_{q_n} F \right| \leq \frac{c}{s} \|F\|_{\mathcal{D}} \cdot \|\omega\|_{d,\rho},$$

by the usual Cauchy inequality. As a linear operator on $\ell^1_{d,\rho}(\mathbb{Z}_1)$, $\partial_q F$ satisfies

$$\|\partial_q F\|_{\text{op}} := \sup_{\omega \neq 0} \frac{\left|\sum_{n \in \mathbb{Z}_1} \omega_n \cdot \partial_{q_n} F\right|}{\|\omega\|_{d,\rho}} \leq \frac{c}{s} \|F\|_{\mathcal{D}}.$$

Let $\|\omega\|_{d,\rho} = \frac{s}{2}$, then

$$|\partial_{q_n} F| \leq \sup_{\|\omega\|_{d,\rho} = \frac{s}{2}} \frac{|\partial_{q_n} F| \cdot |\omega_n|}{\|\omega\|_{d,\rho}} \leq \frac{\|\partial_q F\|_{\text{op}} |\omega_n|}{\frac{s}{2}} \leq \frac{c}{s} \|F\|_{\mathcal{D}} \langle n \rangle^{-d} e^{-|n|\rho}.$$

Hence, for any $0 < \rho' < \rho$,

$$\sum_{n \in \mathbb{Z}_1} |\partial_{q_n} F| \langle n \rangle^d e^{|n|\rho'} \leq \sum_{n \in \mathbb{Z}_1} \frac{c}{s} \|F\|_{\mathcal{D}} e^{-|n|(\rho - \rho')} \leq \frac{c}{s(\rho - \rho')} \|F\|_{\mathcal{D}}.$$

With $\tilde{F} = \sum_{k,l,\alpha,\beta} (\partial_\xi F_{kl\alpha\beta}) I^l e^{\mathrm{i}\langle k, \theta \rangle} q^\alpha \bar{q}^\beta$, it can be proved similarly that

$$\sum_{n \in \mathbb{Z}_1} |\partial_{q_n} \tilde{F}| e^{|n|\rho'} \leq \frac{c}{s(\rho - \rho')} \|F\|_{\mathcal{D}}.$$

Since in the process above, $\xi \in \mathcal{O}$ and $(\theta, I, q, \bar{q}) \in \mathcal{D}_{d,\rho}(r, \frac{s}{2})$ are arbitrarily chosen, this inequality is proved. ■

Let $\{\cdot, \cdot\}$ denote the Poisson bracket of smooth functions, i.e.,

$$\{F, G\} = \langle \partial_I F, \partial_\theta G \rangle - \langle \partial_\theta F, \partial_I G \rangle + \mathrm{i} \sum_{n \in \mathbb{Z}_1} \left( \partial_{q_n} F \cdot \partial_{\bar{q}_n} G - \partial_{\bar{q}_n} F \cdot \partial_{q_n} G \right).$$

**Lemma 2.3** *If $\|X_F\|_{\mathcal{D}} < \varepsilon'$, $\|X_G\|_{\mathcal{D}} < \varepsilon''$, then*

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

*for any $0 < \sigma < r$ and $0 < \eta \ll 1$.*

The proof is similar to that of Lemma 7.3 in [20].

**Definition 2.2** *The function $F(\theta, I, q, \bar{q}; \xi)$ is said to have* **gauge invariance**, *if*

$$F_{kl\alpha\beta}(\xi) \equiv 0, \quad \textit{when } k_1 + k_2 + \cdots + k_b + |\alpha| - |\beta| \neq 0.$$

**Remark 2.2** *This property means the $l^2$ norm $(\sum |q_n|^2)^{\frac{1}{2}}$ is a conserved quantity. It is also related to the fact that solutions of the original equation are invariant with respect to rotations in the complex plane.*

**Lemma 2.4** *If both of $F$ and $G$ have gauge invariance, then $\{F,G\}$ has gauge invariance.*

*Proof:* $F$ and $G$ can be written as

$$F=\sum_{k,\alpha,\beta}F_{k\alpha\beta}(I;\xi)e^{\mathrm{i}\langle k,\theta\rangle}q^{\alpha}\bar{q}^{\beta},\quad G=\sum_{k,\alpha,\beta}G_{k\alpha\beta}(I;\xi)e^{\mathrm{i}\langle k,\theta\rangle}q^{\alpha}\bar{q}^{\beta},$$

with $F_{k\alpha\beta}=G_{k\alpha\beta}\equiv 0$ if $\sum_{j=1}^{b}k_j+|\alpha|-|\beta|\neq 0$. By a simple calculation, we have

$$\{F,G\}_{k\alpha\beta} = \mathrm{i}\sum_{\substack{\check{k}+\hat{k}=k\\ \check{\alpha}+\hat{\alpha}=\alpha\\ \check{\beta}+\hat{\beta}=\beta}}\Big(\langle\partial_I F_{\check{k}\check{\alpha}\check{\beta}},\hat{k}\rangle G_{\hat{k}\hat{\alpha}\hat{\beta}}-\langle\check{k},\partial_I G_{\hat{k}\hat{\alpha}\hat{\beta}}\rangle F_{\check{k}\check{\alpha}\check{\beta}}\Big) \tag{2.2}$$

$$+\mathrm{i}\sum_{\substack{\check{k}+\hat{k}=k\\ \check{\alpha}+\hat{\alpha}=\alpha\\ \check{\beta}+\hat{\beta}=\beta}}\sum_{m\in\mathbb{Z}}\Big(F_{\check{k}(\check{\alpha}+e_m)\check{\beta}}G_{\hat{k}\hat{\alpha}(\hat{\beta}+e_m)}-F_{\check{k}\check{\alpha}(\check{\beta}+e_m)}G_{\hat{k}(\hat{\alpha}+e_m)\hat{\beta}}\Big). \tag{2.3}$$

Assume $\sum_{j=1}^{b}k_j+|\alpha|-|\beta|\neq 0$. Then, in the summation above, it is impossible that

$$\sum_{j=1}^{b}\check{k}_j+|\check{\alpha}|-|\check{\beta}|=\sum_{j=1}^{b}\hat{k}_j+|\hat{\alpha}|-|\hat{\beta}|=0,$$

or

$$\sum_{j=1}^{b}\check{k}_j+|\check{\alpha}+e_m|-|\check{\beta}|=\sum_{j=1}^{b}\hat{k}_j+|\hat{\alpha}|-|\hat{\beta}+e_m|=0,$$

$$\sum_{j=1}^{b}\check{k}_j+|\check{\alpha}|-|\check{\beta}+e_m|=\sum_{j=1}^{b}\hat{k}_j+|\hat{\alpha}+e_m|-|\hat{\beta}|=0.$$

This means, in (2.2) and (2.3), each term $\equiv 0$. Thus Lemma 2.4 is obtained. ■

## 2.2 Statement of the abstract KAM theorem

Associated with the symplectic structure $dI\wedge d\theta+\mathrm{i}\sum_{n\in\mathbb{Z}_1}dq_n\wedge d\bar{q}_n$, $\mathbb{Z}_1\subset\mathbb{Z}$, we consider the following family of real-analytic Hamiltonians

$$H=N+P=e(\xi)+\langle\omega(\xi),I\rangle+\sum_{n\in\mathbb{Z}_1}\Omega_n(\xi)q_n\bar{q}_n+P(\theta,I,q,\bar{q};\xi), \tag{2.4}$$

on some $\mathcal{D}=\mathcal{D}_{d,\rho}(r,s)$, parametrized by $\xi\in\mathcal{O}\subset[0,1]^b$.

Clearly, when $P\equiv 0$, the Hamiltonian reduces to $N$ which is completely integrable and admits a family of special quasi-periodic solutions $(\theta,0,0,0)\to(\theta+\omega t,0,0,0)$, corresponding to invariant $b-$tori in the phase space. To show the persistence of most of these $b-$tori(in Lebesgue measure sense), we need to impose the following conditions on the frequencies $\omega$, $\Omega_n$ and the perturbation $P$.

**(A1)** *Nondegeneracy of tangential frequencies:* The map $\xi\to\omega(\xi)$ is a $C^1_W$ diffeomorphism between $\mathcal{O}$ and its image.

**(A2)** *Regularity of normal frequencies:* For each $n \in \mathbb{Z}_1$, $\Omega_n$ is a $C^1_W$ function of $\xi$ with $\sup_{\xi\in\mathcal{O}} |\partial_\xi \Omega_n| \ll 1$.

**(A3)** *Regularity of the perturbation:* The perturbation $P$ is real-analytic in $\theta, I, q, \bar{q}$ and $C^1_W$ smoothly parametrized by $\xi \in \mathcal{O}$.

**(A4)** *Decay property of the perturbation:* $P$ can be decomposed as $\breve{P} + \acute{P}$, where

$$\breve{P} = \breve{P}(\theta, I, q, \bar{q}; \xi) = \sum_{\alpha,\beta} \breve{P}_{\alpha\beta} q^\alpha \bar{q}^\beta = \sum_{\substack{(k,l)\neq 0\\ \alpha,\beta}} P_{kl\alpha\beta} I^l e^{\mathrm{i}\langle k,\theta\rangle} q^\alpha \bar{q}^\beta,$$

$$\acute{P} = \acute{P}(q, \bar{q}; \xi) = \sum_{\alpha,\beta} \acute{P}_{\alpha\beta} q^\alpha \bar{q}^\beta = \sum_{\alpha,\beta} P_{00\alpha\beta} q^\alpha \bar{q}^\beta,$$

with

$$\|\breve{P}_{\alpha\beta}\|_{\mathcal{D},\mathcal{O}} \leq \begin{cases} \varepsilon e^{-\rho n^*_{\alpha\beta}}, & |\alpha| + |\beta| \leq 2 \\ e^{-\rho n^*_{\alpha\beta}}, & |\alpha| + |\beta| \geq 3 \end{cases}, \tag{2.5}$$

$$\|\acute{P}_{\alpha\beta}\|_{\mathcal{D},\mathcal{O}} \leq \begin{cases} \varepsilon e^{-\rho n^*_{\alpha\beta}}, & |\alpha| + |\beta| \leq 2 \\ e^{-\rho (n^+_{\alpha\beta} - n^-_{\alpha\beta})}, & |\alpha| + |\beta| \geq 3 \end{cases}. \tag{2.6}$$

**(A5)** *Gauge invariance of the perturbation:* For $P = \sum_{\substack{k\in\mathbb{Z}^b,\, l\in\mathbb{N}^b\\ \alpha,\beta}} P_{kl\alpha\beta} I^l e^{\mathrm{i}\langle k,\theta\rangle} q^\alpha \bar{q}^\beta$, we have

$$P_{kl\alpha\beta} \equiv 0 \quad \text{if} \quad \sum_{j=1}^{b} k_j + |\alpha| - |\beta| \neq 0.$$

**Theorem 2** *Assume that the Hamiltonian $H$ in (2.4) satisfies* (**A1**) $-$ (**A5**). *There is a positive constant $\varepsilon_* = \varepsilon_*(\omega, \Omega_n, \varepsilon, r, s, d, \rho)$ such that if $\|X_P\|_{\mathcal{D},\mathcal{O}} < \varepsilon \leq \varepsilon_*$, then there exists a Cantor set $\mathcal{O}_\varepsilon \subset \mathcal{O}$ with $|\mathcal{O} \setminus \mathcal{O}_\varepsilon| \to 0$ as $\varepsilon \to 0$ such that*

(a) *there exists a $C^1_W$ map $\tilde{\omega} : \mathcal{O}_\varepsilon \to \mathbb{R}^b$, such that $|\tilde{\omega} - \omega|_{\mathcal{O}_\varepsilon} \to 0$ as $\varepsilon \to 0$;*

(b) *there exists a map $\Psi : \mathbb{T}^b \times \mathcal{O}_\varepsilon \to \mathcal{D}_{d,0}(r/4, 0)$, real-analytic in $\theta \in \mathbb{T}^b$ and $C^1_W$ parametrized by $\xi \in \mathcal{O}$, such that $\|\Psi - \Psi_0\|_{\mathcal{D}_{d,0}(r/4,0),\mathcal{O}_\varepsilon} \to 0$ as $\varepsilon \to 0$, where $\Psi_0$ is the trivial embedding: $\mathbb{T}^b \times \mathcal{O} \to \mathbb{T}^b \times \{0\} \times \{0\} \times \{0\}$;*

(c) *for any $\theta \in \mathbb{T}^b$ and $\xi \in \mathcal{O}_\varepsilon$, $\Psi(\theta + \tilde{\omega}(\xi)t, \xi) = (\theta + \tilde{\omega}(\xi)t, I(t), q(t), \bar{q}(t))$ is a $b$-frequency quasi-periodic solution of equations of motion associated with the Hamiltonian (2.4);*

(d) *for each $t$, $q(t) = (q_n(t))_{n\in\mathbb{Z}_1} \in \ell^1_{d,0}(\mathbb{Z}_1)$.*

# 3 Proof of Theorem 1

## 3.1 Diagonalization of the linear operator

First, we consider the Schrödinger operators on $\ell^2(\mathbb{Z})$

$$(L_x q)_n = \epsilon(q_{n-1} + q_{n+1}) + \tan \pi(n\tilde{\alpha} + x) q_n, \quad x \in \mathcal{X}, \tag{3.1}$$

which can be interpreted as an infinite dimensional matrix, with the matrix entry

$$(L_x)_{nm} = \begin{cases} \tan \pi(n\tilde{\alpha} + x), & n = m \\ \epsilon, & n - m = \pm 1 \\ 0, & \text{otherwise} \end{cases},$$

where $\tilde{\alpha} \in \mathbb{R}$ satisfies the Diophantine condition (1.2).

**Theorem 3** *(Bellissard-Lima-Scoppola [2]) Consider the Schrödinger operators $L_x$ defined in (3.1) on $\ell^2(\mathbb{Z})$. There exists a positive constant $\epsilon_0 = \epsilon_0(\tilde{\alpha})$, such that if $0 < \epsilon < \epsilon_0$, then the following holds for every $x \in \mathcal{X}$.*

*There is a periodic-one meromophic function $\hat{V}$ on $\{z \in \mathbb{C} : |\mathrm{Im} z| < R\}$ for some $R > 0$ satisfying*

- *The poles of $\hat{V}$ are $\left\{n + \frac{1}{2} : n \in \mathbb{Z}\right\}$,*
- *$\hat{V}(x) - \tan \pi x$ is real-analytic on $\mathbb{R}/\mathbb{Z}$, with $\sup_{x \in \mathbb{R}/\mathbb{Z}} |\hat{V}(x) - \tan \pi x| \leq \epsilon$,*

*and an orthogonal transform $U_x : \ell^2(\mathbb{Z}) \to \ell^2(\mathbb{Z})$ with*

$$|(U_x - I_{\mathbb{Z}})_{mn}| \leq \epsilon e^{-2|m-n|}, \tag{3.2}$$

*such that $U_x^* L_x U_x = diag\{\hat{V}(n\tilde{\alpha} + x)\}$.*

This theorem (in its original form) is due to Bellissard-Lima-Scoppola[2]. The detailed statement will be given in Appendix A.1.

**Remark 3.1** *Typically, there is a polynomial or exponential factor in front of the exponential decay in (3.2), which is called semi-uniform localized eigenstate(SULE). For example, the random Schrödinger operator and the almost Mathieu operator exhibit such a phenomenon. It is necessary to point out that, the method needed to investigate such models is totally different from that of the present paper, because there are infinitely many resonances.*

*Compared with SULE, the uniform localized eigenstate in (3.2) is not generic[15]. Correspondingly, the Maryland model is a special quasi-crystal model. However, in the presence of nonlinearity, many problems related to the model are still unsolved and attract plenty of attention.*

### 3.2 The Hamiltonian

Consider Equation(1.4). For every $x \in \mathcal{X}$, after the coordinate transformation

$$q_{\mathbb{Z}} = U_x \tilde{q}_{\mathbb{Z}},$$

with $U_x$ given in Theorem 3, there is no difference in the linear part, and the new Hamiltonian has the form

$$H(\tilde{q}_{\mathbb{Z}}, \bar{\tilde{q}}_{\mathbb{Z}}) = \Lambda + G := \sum_{n \in \mathbb{Z}} \hat{V}_n |\tilde{q}_n|^2 + \frac{1}{2}\epsilon \sum_{i,j,n,m \in \mathbb{Z}} u_{ijnm} \tilde{q}_i \bar{\tilde{q}}_j \tilde{q}_n \bar{\tilde{q}}_m, \tag{3.3}$$

where $\hat{V}_n = \hat{V}_n(x) := \hat{V}(n\tilde{\alpha} + x)$. The off-diagonal decay of $U_x$ in (3.2) implies the short-range estimates of coefficients $u_{ijnm}$, i.e.,

$$|u_{ijnm}| < ce^{-2(\max\{i,j,n,m\}-\min\{i,j,n,m\})}. \tag{3.4}$$

Indeed, for fixed $x \in \mathcal{X}$, we can calculate that

$$u_{ijnm} = \sum_{l\in\mathbb{Z}} (U_x)_{li}\overline{(U_x)_{lj}}(U_x)_{ln}\overline{(U_x)_{lm}}. \tag{3.5}$$

Without loss of generality, assume that $i \le j \le n \le m$, then

$$\begin{aligned} |u_{ijnm}| &\le c\sum_{l\in\mathbb{Z}} e^{-2(|i-l|+|j-l|+|n-l|+|m-l|)} \\ &\le ce^{-2(m-i)}\sum_{l\in\mathbb{Z}} e^{-2(|j-l|+|n-l|)} \\ &\le ce^{-2(m-i)}. \end{aligned}$$

Now we fix $\mathcal{J} = \{n_1, \cdots, n_b\} \subset \mathbb{Z}$, and $\mathbb{Z}_1 = \mathbb{Z} \setminus \mathcal{J}$. When $\epsilon$ is sufficiently small, we have $|n_i| \le \frac{\kappa}{6}|\ln\epsilon|$ for $i = 1, \cdots, b$.

Fix $r, d > 0$ and $\rho = \frac{1}{4}$, $s \le \epsilon^{\frac{2}{3}\kappa}$. Define $\mathcal{D} = \mathcal{D}_{d,\rho}(r,s)$ as in Subsection 2.1. Before introducing action-angle variables and parameters, we need to transform $H$ into a Hamiltonian with a nice normal form. Hereafter, we will write the variable $q_{\mathbb{Z}}$ instead of $\tilde{q}_{\mathbb{Z}}$ in the Hamiltonian for convenience.

**Proposition 1** *For $\epsilon$ sufficiently small, there exists a subset $\mathcal{X}_\epsilon$ of $\mathcal{X}$ with*

$$\mathrm{mes}(\mathcal{X} \setminus \mathcal{X}_\epsilon) < \epsilon^{\vartheta} \quad \textit{for some} \quad 0 < \vartheta < 1,$$

*such that for every $x \in \mathcal{X}_\epsilon$, there is a symplectic transformation $\Psi = \Psi(x)$, which transforms $H$ in (3.3) into*

$$\begin{aligned} H \circ \Psi &= N + P \\ &:= e(\xi) + \langle \omega(\xi), I\rangle + \sum_{n\in\mathbb{Z}_1} \Omega_n(\xi) q_n \bar{q}_n + P(\theta, I, q, \bar{q}; \xi), \end{aligned} \tag{3.6}$$

*a real-analytic Hamiltonian on $\mathcal{D}$, $C^1_W$ parametrized by $\xi \in \mathcal{O} := [\epsilon^{\frac{\kappa}{12}}, 1]^b$. Here,*

- *$\omega$ is a $C^1_W$ diffeomorphism between $\mathcal{O}$ and its image,*
- *for each $n \in \mathbb{Z}_1$, $\Omega_n$ is a $C^1_W$ function of $\xi$ with $\sup_{\xi\in\mathcal{O}} |\partial_\xi \Omega_n| \le \epsilon$.*

*Moreover, $P$ has gauge invariance, and can be decomposed as $\breve{P} + \acute{P}$ with*

$$\breve{P} = \breve{P}(\theta, I, q, \bar{q}; \xi) = \sum_{\alpha,\beta} \breve{P}_{\alpha\beta} q^\alpha \bar{q}^\beta = \sum_{\substack{(k,l)\neq 0\\ \alpha,\beta}} P_{kl\alpha\beta} I^l e^{\mathrm{i}\langle k,\theta\rangle} q^\alpha \bar{q}^\beta,$$

$$\acute{P} = \acute{P}(q, \bar{q}; \xi) = \sum_{\alpha,\beta} \acute{P}_{\alpha\beta} q^\alpha \bar{q}^\beta = \sum_{\alpha,\beta} P_{00\alpha\beta} q^\alpha \bar{q}^\beta,$$

*satisfying*

$$\|\breve{P}_{\alpha\beta}\|_{\mathcal{D},\mathcal{O}} \le \begin{cases} \epsilon^{\frac{\kappa}{4}} e^{-\frac{1}{2}n^*_{\alpha\beta}}, & |\alpha|+|\beta|\le 2 \\ e^{-\frac{1}{2}n^*_{\alpha\beta}}, & |\alpha|+|\beta|\ge 3 \end{cases}, \tag{3.7}$$

$$\|\acute{P}_{\alpha\beta}\|_{\mathcal{D},\mathcal{O}} \le \begin{cases} \epsilon^{\frac{\kappa}{4}} e^{-\frac{1}{2}n^*_{\alpha\beta}}, & |\alpha|+|\beta|\le 2 \\ e^{-\frac{1}{2}(n^+_{\alpha\beta}-n^-_{\alpha\beta})}, & |\alpha|+|\beta|\ge 3 \end{cases}. \tag{3.8}$$

*Proof:* We decompose the proof into the following parts.

- **Symplectic changes of variables**

According to the form of $H = \Lambda + G$, let

$$T(q_{\mathbb{Z}},\, \bar{q}_{\mathbb{Z}}) = \frac{1}{2}\epsilon \sum_{|i|,|j|,|n|,|m|\le\kappa|\ln\epsilon|} u_{ijnm} q_i \bar{q}_j q_n \bar{q}_m,$$

$$F(q_{\mathbb{Z}},\, \bar{q}_{\mathbb{Z}}) = \frac{\mathrm{i}}{2}\epsilon \sum_{\substack{\hat{V}_i-\hat{V}_j+\hat{V}_n-\hat{V}_m\neq 0\\ |i|,|j|,|n|,|m|\le\kappa|\ln\epsilon|}} \frac{u_{ijnm}}{\hat{V}_i-\hat{V}_j+\hat{V}_n-\hat{V}_m} q_i \bar{q}_j q_n \bar{q}_m,$$

and $\Psi^1_F$ be the time-one map of the flow of associated Hamiltonian systems. For fixed $i, j, n, m \in \mathbb{Z}$ with $|i|, |j|, |n|, |m| \le \kappa|\ln\epsilon|$, consider the function

$$V_{i,j,n,m}(x) := \hat{V}_i(x) - \hat{V}_j(x) + \hat{V}_n(x) - \hat{V}_m(x).$$

Since $\epsilon$ is small enough, by Lemma 3.1 below, there exists a subset $\mathcal{X}_\epsilon$ of $\mathcal{X}$ with

$$\mathrm{mes}(\mathcal{X}\setminus\mathcal{X}_\epsilon) \le \epsilon^{\vartheta} \ \text{ for some } \ 0<\vartheta<1,$$

such that if $x \in \mathcal{X}_\epsilon$ and $\{i,n\} \neq \{j,m\}$, then $|V_{i,j,n,m}(x)| \ge \epsilon^{\frac{1}{4}}$. This guarantees that there is a uniform lower bound for the denominators in coefficients of $F$.

In view of the homological equation

$$\{\Lambda, F\} + T = \frac{1}{2}\epsilon \sum_{|i|,|j|\le\kappa|\ln\epsilon|} u_{iijj}|q_i|^2|q_j|^2,$$

we know that the change of variables $\Psi^1_F$ sends $H$ to

$$H\circ\Psi^1_F = \sum_{i\in\mathbb{Z}} \hat{V}_i|q_i|^2 + \frac{1}{2}\epsilon \sum_{|i|,|j|\le\kappa|\ln\epsilon|} u_{iijj}|q_i|^2|q_j|^2 + \tilde{R}, \tag{3.9}$$

where

$$\begin{aligned}\tilde{R} \;=\;& G - T + \{G,F\} + \frac{1}{2!}\{\{\Lambda,F\},F\} + \frac{1}{2!}\{\{G,F\},F\} + \cdots \\ &+\frac{1}{n!}\{\cdots\{\Lambda,\underbrace{F\}\cdots,F}_{n}\} + \frac{1}{n!}\{\cdots\{G,\underbrace{F\}\cdots,F}_{n}\} + \cdots.\end{aligned}$$

Expand $\tilde{R}$ as $\tilde{R}=\sum_{\alpha',\beta'}\tilde{R}_{\alpha'\beta'}q_{\mathbb{Z}}^{\alpha'}\bar{q}_{\mathbb{Z}}^{\beta'}$. Here $(\alpha',\beta')=(\alpha_n,\beta_n)_{n\in\mathbb{Z}}$, with finitely many non-vanishing components, for which notations $\operatorname{supp}(\alpha',\beta')$, $n^{+}_{\alpha'\beta'}$, $n^{-}_{\alpha'\beta'}$, $n^{*}_{\alpha'\beta'}$ and $|\alpha'|$, $|\beta'|$ can be defined as in Definition 2.1. By the construction of $\tilde{R}$, we have

$$\tilde{R}_{\alpha'\beta'}=0 \quad \text{if} \quad |\alpha'|+|\beta'|<4 \quad \text{or} \quad |\alpha'|\neq|\beta'|, \tag{3.10}$$

and

$$\tilde{R}_{\alpha'\beta'}=0 \quad \text{if} \quad |\alpha'|+|\beta'|=4 \quad \text{and} \quad n^{*}_{\alpha'\beta'}\leq\kappa|\ln\epsilon|. \tag{3.11}$$

Moreover, by applying Lemma 3.2 below iteratively,

$$|\tilde{R}_{\alpha'\beta'}|\leq\epsilon e^{-2(n^{+}_{\alpha'\beta'}-n^{-}_{\alpha'\beta'})}.$$

- **Introduction of action-angle variables**

Introducing the action-angle variables in the tangential space

$$q_i=\sqrt{I_i+\xi_i}e^{\mathrm{i}\theta_i}, \quad \bar{q}_i=\sqrt{I_i+\xi_i}e^{-\mathrm{i}\theta_i}, \quad i\in\mathcal{J},$$

where $(\theta,I)=(\theta_{n_1},\cdots,\theta_{n_b},I_{n_1},\cdots,I_{n_b})$ are the standard action-angle variables in the $(q_n,\bar{q}_n)_{n\in\mathcal{J}}$-space around $\xi$, with $\xi=(\xi_{n_1},\cdots,\xi_{n_b})\in\epsilon^{\kappa}[\epsilon^{\frac{\kappa}{12}},1]^b$ a parameter, and

$$(q,\bar{q})=(q_n,\bar{q}_n)_{n\in\mathbb{Z}_1}$$

the remaining normal variables. Then the Hamiltonian in (3.9) becomes

$$\begin{aligned}
H\circ\Psi_F^1 &= \sum_{i\in\mathcal{J}}\hat{V}_i(I_i+\xi_i)+\sum_{i\in\mathbb{Z}_1}\hat{V}_i|q_i|^2+\frac{1}{2}\epsilon\sum_{i\in\mathcal{J}}u_{iiii}(I_i+\xi_i)^2\\
&\quad+\frac{1}{2}\epsilon\sum_{\substack{i\in\mathcal{J},j\in\mathbb{Z}_1\\|j|\leq\kappa|\ln\epsilon|}}u_{iijj}(I_i+\xi_i)|q_j|^2+\frac{1}{2}\epsilon\sum_{\substack{i,j\in\mathcal{J}\\i\neq j}}u_{iijj}(I_i+\xi_i)(I_j+\xi_j)\\
&\quad+\frac{1}{2}\epsilon\sum_{\substack{i,j\in\mathbb{Z}_1\\|i|,|j|\leq\kappa|\ln\epsilon|}}u_{iijj}|q_i|^2|q_j|^2+\tilde{R}\\
&= \sum_{i\in\mathcal{J}}\hat{V}_iI_i+\sum_{i\in\mathbb{Z}_1}\hat{V}_i|q_i|^2+\epsilon\sum_{i\in\mathcal{J}}u_{iiii}\xi_iI_i+\frac{1}{2}\epsilon\sum_{\substack{i,j\in\mathcal{J}\\i\neq j}}u_{iijj}(\xi_iI_j+\xi_jI_i)\\
&\quad+\frac{1}{2}\epsilon\sum_{\substack{i\in\mathcal{J},j\in\mathbb{Z}_1\\|j|\leq\kappa|\ln\epsilon|}}u_{iijj}\xi_i|q_j|^2+\left(\sum_{i\in\mathcal{J}}\hat{V}_i\xi_i+\frac{1}{2}\epsilon\sum_{i\in\mathcal{J}}u_{iiii}\xi_i^2+\frac{1}{2}\epsilon\sum_{\substack{i,j\in\mathcal{J}\\j\neq i}}u_{iijj}\xi_i\xi_j\right)\\
&\quad+R,
\end{aligned}$$

where

$$R=\tilde{R}+\frac{1}{2}\epsilon\sum_{i\in\mathcal{J}}u_{iiii}I_i^2+\frac{1}{2}\epsilon\sum_{\substack{i,j\in\mathcal{J}\\i\neq j}}u_{iijj}I_iI_j+\frac{1}{2}\epsilon\sum_{\substack{i\in\mathcal{J},j\in\mathbb{Z}_1\\|j|\leq\kappa|\ln\epsilon|}}u_{iijj}I_i|q_j|^2.$$

By the scaling in time

$$\theta\to\theta, \quad I\to\epsilon^{\frac{4}{3}\kappa}I, \quad q\to\epsilon^{\frac{2}{3}\kappa}q, \quad \bar{q}\to\epsilon^{\frac{2}{3}\kappa}\bar{q}, \quad \xi\to\epsilon^{\kappa}\xi, \tag{3.12}$$

we finally arrive at the rescaled Hamiltonian

$$H\circ\Psi_F^1=\epsilon^{-(1+\frac{7}{3}\kappa)}(H\circ\Psi_F^1)(\theta,\epsilon^{\frac{4}{3}\kappa}I,\epsilon^{\frac{2}{3}\kappa}q,\epsilon^{\frac{2}{3}\kappa}\bar{q};\epsilon^{\kappa}\xi)=N+P,$$

where $N=e+\langle\omega,I\rangle+\sum_{n\in\mathbb{Z}_1}\Omega_n|q_n|^2$, with

$$\omega_i(\xi) = \epsilon^{-(1+\kappa)}\hat{V}_i+u_{iiii}\xi_i+\frac{1}{2}\sum_{\substack{j\in\mathcal{J}\\ j\neq i}}u_{iijj}\xi_j,\quad i\in\mathcal{J}, \tag{3.13}$$

$$\Omega_n(\xi) = \begin{cases}\epsilon^{-(1+\kappa)}\hat{V}_n+\frac{1}{2}\sum_{i\in\mathcal{J}}u_{iinn}\xi_i, & |n|\le\kappa|\ln\epsilon|\\ \epsilon^{-(1+\kappa)}\hat{V}_n, & |n|>\kappa|\ln\epsilon|\end{cases},\quad n\in\mathbb{Z}_1 \tag{3.14}$$

and $P=\epsilon^{-(1+\frac{7}{3}\kappa)}R(\theta,\epsilon^{\frac{4}{3}\kappa}I,\epsilon^{\frac{2}{3}\kappa}q,\epsilon^{\frac{2}{3}\kappa}\bar{q};\epsilon^{\kappa}\xi)$.

- **Properties of the Hamiltonian $N$**

In view of (3.13), the $b\times b$ matrix $\frac{\partial\omega}{\partial\xi}$ satisfies that

$$\left(\frac{\partial\omega}{\partial\xi}\right)_{ij}=\begin{cases}\sum_{l\in\mathbb{Z}}|(U_x)_{il}|^4, & j=i\\ \frac{1}{2}\sum_{l\in\mathbb{Z}}|(U_x)_{il}|^2|(U_x)_{jl}|^2, & j\neq i\end{cases},\quad i,j\in\mathcal{J},$$

since $u_{iijj}=\sum_{l\in\mathbb{Z}}|(U_x)_{il}|^2|(U_x)_{jl}|^2$ as is shown in (3.5). By (3.2), we have

$$|(U_x)_{ii}-1|<\epsilon\ \text{ and }\ |(U_x)_{il}|\le\epsilon e^{-2|i-l|},\ \ l\neq i.$$

Hence, $\sum_{l\in\mathbb{Z}}|(U_x)_{il}|^4>c(1-\epsilon)^4$, while $\sup_{i\neq j}\sum_{l\in\mathbb{Z}}|(U_x)_{il}|^2|(U_x)_{jl}|^2\le c\epsilon^2$. The diagonal dominance of $\frac{\partial\omega}{\partial\xi}$, which is deduced from the smallness of $\epsilon$, implies that $\omega$ is a $C_W^1$ diffeomorphism between $\mathcal{O}$ and its image.

The formulation of $\Omega_n$ given in (3.14) implies that $\partial_\xi\Omega_n=0$ for $|n|>\kappa|\ln\epsilon|$. As for the case $|n|\le\kappa|\ln\epsilon|$, we have

$$|\partial_{\xi_i}\Omega_n|=\frac{1}{2}\sum_{l\in\mathbb{Z}}|(U_x)_{il}|^2|(U_x)_{nl}|^2\le c\epsilon^2,\quad i\in\mathcal{J}.$$

- **Properties of the Hamiltonian $P$**

By (3.10), each non-zero term of $\tilde{R}$ can be rewritten as

$$\tilde{R}_{\alpha'\beta'}q_{\mathbb{Z}}^{\alpha'}\bar{q}_{\mathbb{Z}}^{\beta'}=\tilde{R}_{\alpha'\beta'}q_{\mathcal{J}}^{\alpha_{\mathcal{J}}}\bar{q}_{\mathcal{J}}^{\beta_{\mathcal{J}}}q^{\alpha}\bar{q}^{\beta},\quad |\alpha'|+|\beta'|\ge 4,\ \ |\alpha'|=|\beta'|,$$

where $\alpha_{\mathcal{J}}=(\alpha_n)_{n\in\mathcal{J}}$, $\beta_{\mathcal{J}}=(\beta_n)_{n\in\mathcal{J}}$, and $q_{\mathcal{J}}=(q_n)_{n\in\mathcal{J}}$, $\bar{q}_{\mathcal{J}}=(\bar{q}_n)_{n\in\mathcal{J}}$, then the introduction of action-angle variables brings us

$$\tilde{R}_{\alpha'\beta'}\left(\prod_{n\in\mathcal{J}}\left(\sqrt{I_n+\xi_n}\right)^{\alpha_n+\beta_n}e^{\mathrm{i}(\alpha_n-\beta_n)\theta_n}\right)q^{\alpha}\bar{q}^{\beta},$$

which, after the scaling (3.12), becomes

$$\mathcal{E}\tilde{R}_{\alpha'\beta'}\left(\prod_{n\in\mathcal{J}}\left(\sqrt{\epsilon^{\frac{\kappa}{3}}I_n+\xi_n}\right)^{\alpha_n+\beta_n}e^{\mathrm{i}(\alpha_n-\beta_n)\theta_n}\right)q^{\alpha}\bar{q}^{\beta}, \tag{3.15}$$

where $\mathcal{E} = \epsilon^{-(1+\frac{7}{3}\kappa)}\epsilon^{\frac{\kappa}{2}(|\alpha_{\mathcal{J}}|+|\beta_{\mathcal{J}}|)+\frac{2}{3}\kappa(|\alpha|+|\beta|)}$. As a term of $P = \sum_{k,\alpha,\beta} P_{k\alpha\beta}(I)e^{\mathrm{i}\langle k,\theta\rangle}q^{\alpha}\bar{q}^{\beta}$, this means,

$$\sum_{j=1}^{b} k_j = \sum_{n\in\mathcal{J}} (\alpha_n - \beta_n).$$

Then $\sum_{j=1}^{b} k_j + |\alpha| - |\beta|$ equals to its initial value $\sum_{n\in\mathbb{Z}} \alpha_n - \sum_{n\in\mathbb{Z}} \beta_n = |\alpha'| - |\beta'|$. Thus, by (3.10),

$$P_{k\alpha\beta} \equiv 0 \ \text{ if } \ \sum_{j=1}^{b} k_j + |\alpha| - |\beta| = |\alpha'| - |\beta'| \neq 0.$$

The gauge invariance of $P$ is deduced by expanding $P_{k\alpha\beta}$ with respect to $I$.

We need to verify the decay property of $P$. Decompose $P$ as $P = \breve{P} + \acute{P}$, which has been given in the proposition.

1) $|\alpha_{\mathcal{J}}| + |\beta_{\mathcal{J}}| = 0$
In this case, $|\alpha'| + |\beta'| = |\alpha| + |\beta| \geq 4$ in view of (3.10), and the term in (3.15) is $\epsilon^{-(1+\frac{7}{3}\kappa)}\epsilon^{\frac{2}{3}\kappa(|\alpha|+|\beta|)}\tilde{R}_{\alpha'\beta'}q^{\alpha}\bar{q}^{\beta}$. This is a higher-order term of $\acute{P}$, with its coefficient smaller than

$$\epsilon^{\frac{\kappa}{3}-1}|\tilde{R}_{\alpha'\beta'}| \leq \epsilon^{\frac{\kappa}{3}-1} \cdot \epsilon e^{-2(n^{+}_{\alpha'\beta'}-n^{-}_{\alpha'\beta'})} \leq \epsilon^{\frac{\kappa}{3}} e^{-2(n^{+}_{\alpha'\beta'}-n^{-}_{\alpha'\beta'})}. \tag{3.16}$$

2) $|\alpha_{\mathcal{J}}| + |\beta_{\mathcal{J}}| \geq 1$
This means $\mathrm{supp}(\alpha', \beta') \cap [-\frac{\kappa}{6}|\ln\epsilon|, \frac{\kappa}{6}|\ln\epsilon|] \neq \emptyset$, i.e., there exists $|n| \leq \frac{\kappa}{6}|\ln\epsilon|$ such that $(\alpha'_n, \beta'_n) \neq (0,0)$, then we have that

$$n^{*}_{\alpha'\beta'} - \frac{\kappa}{6}|\ln\epsilon| \leq n^{*}_{\alpha'\beta'} - |n| \leq n^{+}_{\alpha'\beta'} - n^{-}_{\alpha'\beta'}.$$

Hence,

$$|\tilde{R}_{\alpha'\beta'}| \leq \epsilon e^{-2(n^{+}_{\alpha'\beta'}-n^{-}_{\alpha'\beta'})} \leq \epsilon e^{\frac{\kappa}{3}|\ln\epsilon|}e^{-2n^{*}_{\alpha'\beta'}} = \epsilon^{1-\frac{\kappa}{3}}e^{-2n^{*}_{\alpha'\beta'}}.$$

By (3.10), we can consider Case 2) in the following two situations.

– If $|\alpha'| + |\beta'| \geq 6$, then $\frac{\kappa}{2}(|\alpha_{\mathcal{J}}| + |\beta_{\mathcal{J}}|) + \frac{2}{3}\kappa(|\alpha| + |\beta|) \geq 3\kappa$ and $\mathcal{E} \leq \epsilon^{\frac{2}{3}\kappa-1}$. This means the coefficient is not more than

$$\mathcal{E}|\tilde{R}_{\alpha'\beta'}| \leq \epsilon^{\frac{2}{3}\kappa-1} \cdot \epsilon^{1-\frac{\kappa}{3}}e^{-2n^{*}_{\alpha'\beta'}} \leq \epsilon^{\frac{\kappa}{3}}e^{-2n^{*}_{\alpha'\beta'}}. \tag{3.17}$$

– If $|\alpha'| + |\beta'| = 4$, then by (3.11), $n^{*}_{\alpha'\beta'} > \kappa|\ln\epsilon|$, and hence

$$|\tilde{R}_{\alpha'\beta'}| \leq \epsilon^{1-\frac{\kappa}{3}}e^{-\kappa|\ln\epsilon|}e^{-n^{*}_{\alpha'\beta'}} = \epsilon^{1+\frac{2}{3}\kappa}e^{-n^{*}_{\alpha'\beta'}}.$$

This means the coefficient in (3.15) is not more than

$$\mathcal{E}|\tilde{R}_{\alpha'\beta'}| \leq \epsilon^{-(1+\frac{7}{3}\kappa)}\epsilon^{2\kappa}\epsilon^{1+\frac{2}{3}\kappa}e^{-n^{*}_{\alpha'\beta'}} \leq \epsilon^{\frac{\kappa}{3}}e^{-n^{*}_{\alpha'\beta'}}. \tag{3.18}$$

Thus, Case 2), the coefficient of $q^\alpha \bar{q}^\beta$ in (3.15) can be controlled as

$$\left\| \mathcal{E}\tilde{R}_{\alpha'\beta'} \left( \prod_{n\in\mathcal{J}} \left( \sqrt{\epsilon^{\frac{\kappa}{3}} I_n + \xi_n} \right)^{\alpha_n+\beta_n} e^{\mathrm{i}(\alpha_n-\beta_n)\theta_n} \right) \right\|_{\mathcal{D},\mathcal{O}} \leq \epsilon^{\frac{\kappa}{4}} e^{-n^*_{\alpha'\beta'}}.$$

In expanding $\sqrt{I_n+\xi_n}$ around $\xi_n$, we need to keep $\xi_n$ apart from 0 to avoid singularity. This is why we choose $\xi \in [\epsilon^{\frac{\kappa}{12}}, 1]^b$(after scaling).

There is no doubt that terms of $\breve{P}$ are all generated in Case 2), so, applying the basic fact $\mathrm{supp}(\alpha,\beta) \subset \mathrm{supp}(\alpha',\beta')$,

$$\|\breve{P}_{\alpha\beta}\|_{\mathcal{D},\mathcal{O}} \leq \epsilon^{\frac{\kappa}{4}} e^{-n^*_{\alpha'\beta'}} \leq \epsilon^{\frac{\kappa}{4}} e^{-n^*_{\alpha\beta}},$$

which implies (3.7).

Terms of $\acute{P}$ come from both cases. When the term in (3.15) satisfies that $\alpha_{\mathcal{J}} = \beta_{\mathcal{J}}$, by expanding $\sqrt{I_n+\xi_n}$ around $\xi_n$ we can obtain

$$\mathcal{E}\tilde{R}_{\alpha'\beta'} \left( \prod_{n\in\mathcal{J}} \left( \sqrt{\xi_n} \right)^{\alpha_n+\beta_n} \right) q^\alpha \bar{q}^\beta,$$

which contributes one term to $\acute{P}$ due to cancelation of angle variables. As in Case 2), the corresponding coefficient is not more than $\epsilon^{\frac{\kappa}{4}} e^{-n^*_{\alpha\beta}}$, which can be replaced by $\epsilon^{\frac{\kappa}{4}} e^{-\frac{1}{2}(n^+_{\alpha\beta} - n^-_{\alpha\beta})}$ as we need, since $\frac{1}{2}(n^+_{\alpha\beta} - n^-_{\alpha\beta}) \leq n^*_{\alpha\beta}$. Together with (3.16), (3.8) is proved. ■

Combing (3.16) – (3.18) together, we have

$$\|X_P\|_{\mathcal{D}_{d,\rho}(r,s),\mathcal{O}} \leq \varepsilon := \epsilon^{\frac{\kappa}{8}}.$$

To this stage, we have that all the assumptions of Theorem 2 hold for (3.6), which conjugates with (1.4). Thus, Theorem 1 follows from Theorem 2.

We have applied several conclusions directly in proving Proposition 1. Now we give their precise statements. The first lemma shows that the function

$$V_{i,j,n,m}(x) = \hat{V}(x+i\tilde{\alpha}) - \hat{V}(x+j\tilde{\alpha}) + \hat{V}(x+n\tilde{\alpha}) - \hat{V}(x+m\tilde{\alpha})$$

on $\mathcal{X}$ is not identically zero, if $|i|,|j|,|n|,|m| \leq \kappa|\ln\epsilon|$ and $\{i,n\} \neq \{j,m\}$.

**Lemma 3.1** *For $\epsilon$ sufficiently small, there exists a subset $\mathcal{X}_\epsilon$ of $\mathcal{X}$ with*

$$\mathrm{mes}(\mathcal{X} \setminus \mathcal{X}_\epsilon) < \epsilon^\vartheta \quad for \ \ some \ \ 0 < \vartheta < 1,$$

*such that for any $|i|,|j|,|n|,|m| \leq \kappa|\ln\epsilon|$ and $\{i,n\} \neq \{j,m\}$, we have*

$$|V_{i,j,n,m}(x)| \geq \epsilon^{\frac{1}{4}}, \quad \forall x \in \mathcal{X}_\epsilon. \tag{3.19}$$

The proof of Lemma 3.1 is very similar to Appendix A in [22], and the measure estimate is an analogue with Lemma 5.3 in [33]. For the sake of completeness, we give its proof in Appendix A.2.

The next lemma implies that the property (3.4) about the coefficients of the Hamiltonian is preserved under the poisson bracket.

**Lemma 3.2** *Consider two real-analytic functions*[4]

$$G(q_{\mathbb{Z}},\bar{q}_{\mathbb{Z}})=\sum_{\alpha,\beta}G_{\alpha\beta}q_{\mathbb{Z}}^{\alpha}\bar{q}_{\mathbb{Z}}^{\beta},\quad F(q_{\mathbb{Z}},\bar{q}_{\mathbb{Z}})=\sum_{\substack{\alpha,\beta\\ n^+_{\alpha\beta}-n^-_{\alpha\beta}\le M}}F_{\alpha\beta}q_{\mathbb{Z}}^{\alpha}\bar{q}_{\mathbb{Z}}^{\beta},$$

*with*

$$|G_{\alpha\beta}|\le c_G e^{-\sigma(n^+_{\alpha\beta}-n^-_{\alpha\beta})},\quad |F_{\alpha\beta}|\le c_F e^{-\sigma(n^+_{\alpha\beta}-n^-_{\alpha\beta})},$$

*for some positive $c_G$, $c_F$ and $\sigma$. We have that*

$$K(q_{\mathbb{Z}},\bar{q}_{\mathbb{Z}})=\mathrm{i}\sum_{n\in\mathbb{Z}}(\partial_{q_n}F\cdot\partial_{\bar{q}_n}G-\partial_{\bar{q}_n}F\cdot\partial_{q_n}G)=\sum_{\alpha,\beta}K_{\alpha\beta}q_{\mathbb{Z}}^{\alpha}\bar{q}_{\mathbb{Z}}^{\beta}$$

*satisfies*

$$|K_{\alpha\beta}|\le c\cdot M^2c_Gc_Fe^{-\sigma(n^+_{\alpha\beta}-n^-_{\alpha\beta})}.$$

*Proof:* A straightforward calculation yields that

$$K_{\alpha\beta}=\mathrm{i}\sum_{\mathcal{S}}\left(G_{\check{\alpha}+e_n,\check{\beta}}F_{\hat{\alpha},\hat{\beta}+e_n}-G_{\check{\alpha},\check{\beta}+e_n}F_{\hat{\alpha}+e_n,\hat{\beta}}\right), \tag{3.20}$$

with the summation notation

$$\mathcal{S}=\left\{\begin{array}{l} n\in\mathbb{Z},\quad (\check{\alpha},\check{\beta})+(\hat{\alpha},\hat{\beta})=(\alpha,\beta),\\ n^+_{\hat{\alpha},\hat{\beta}+e_n}-n^-_{\hat{\alpha},\hat{\beta}+e_n}\le M\ \text{ or }\ n^+_{\hat{\alpha}+e_n,\hat{\beta}}-n^-_{\hat{\alpha}+e_n,\hat{\beta}}\le M\end{array}\right\}.$$

For $G_{\check{\alpha}+e_n,\check{\beta}}F_{\hat{\alpha},\hat{\beta}+e_n}$ in (3.20), note that

$$n^+_{\alpha\beta}\le\max\{n^+_{\check{\alpha}+e_n,\check{\beta}},n^+_{\hat{\alpha},\hat{\beta}+e_n}\},\quad n^-_{\alpha\beta}\ge\max\{n^-_{\check{\alpha}+e_n,\check{\beta}},n^-_{\hat{\alpha},\hat{\beta}+e_n}\},$$

then

$$n^+_{\check{\alpha}+e_n,\check{\beta}}-n^-_{\check{\alpha}+e_n,\check{\beta}}+n^+_{\hat{\alpha},\hat{\beta}+e_n}-n^-_{\hat{\alpha},\hat{\beta}+e_n}\ge n^+_{\alpha\beta}-n^-_{\alpha\beta}.$$

Hence

$$|G_{\check{\alpha}+e_n,\check{\beta}}F_{\hat{\alpha},\hat{\beta}+e_n}|\le c_Gc_Fe^{-\sigma(n^+_{\check{\alpha}+e_n,\check{\beta}}-n^-_{\check{\alpha}+e_n,\check{\beta}})}e^{-\sigma(n^+_{\hat{\alpha},\hat{\beta}+e_n}-n^-_{\hat{\alpha},\hat{\beta}+e_n})}\le c_Gc_Fe^{-\sigma(n^+_{\alpha\beta}-n^-_{\alpha\beta})}.$$

Doing the same for $G_{\check{\alpha},\check{\beta}+e_n}F_{\hat{\alpha}+e_n,\hat{\beta}}$ in (3.20), and noting that $K_{\alpha\beta}$ is a finite sum in view of the definition of $\mathcal{S}$, we have completed the proof of this lemma. ∎

[4]Here we use $(\alpha,\beta)$ instead of $(\alpha',\beta')$ to denote $(\alpha_n,\beta_n)_{n\in\mathbb{Z}}$ for convenience.

## 4 KAM step

The remaining sections are devoted to the proof of Theorem 2. In this section we present the KAM iteration scheme applied to (2.4). This is a succession of infinitely many steps, to eliminate lower-order $\theta$-dependent terms in $P$. At each KAM step, the perturbation is made smaller at the cost of excluding a small-measure set of parameters. It will be shown that the KAM iterations converge and that, in the end, the total measure of the set of parameters that has been excluded is small.

### 4.1 Normal form

In order to perform the KAM iteration scheme, we shall first write the Hamiltonian (2.4) into a normal form that is more convenient for this purpose. For simplicity, we only outline the derivation of the normal form. Detailed construction and estimation is similar to those for the general KAM step which we will show later.

To begin the KAM iteration, we set $r_0 = \frac{r}{2}$, $\varepsilon_0 = \varepsilon^{\frac{5}{4}}$, and $K_0 = 2|\ln\varepsilon|\rho^{-1}$, $\rho_0 = K_0^{-1}$. Let $s_0$ be such that $0 < s_0 < \min\{\varepsilon_0, s\}$, and define $\mathcal{D}_0 = \mathcal{D}_{d,\rho_0}(r_0, s_0)$.

Consider terms of $\check{P}$ and $\acute{P}$. According to (2.5) and (2.6) in the assumption (**A4**) and the definition of norm (2.1), we have that coefficients of

$$\check{P} = \sum_{\substack{(k,l)\neq 0\\ \alpha,\beta}} P_{kl\alpha\beta} I^l e^{\mathrm{i}\langle k,\theta\rangle} q^\alpha \bar{q}^\beta, \quad \acute{P} = \sum_{\alpha,\beta} P_{00\alpha\beta} q^\alpha \bar{q}^\beta$$

satisfy that

$$|P_{kl\alpha\beta}|_{\mathcal{O}} \leq \varepsilon e^{-\rho n^*_{\alpha\beta}} e^{-|k|r}, \quad \forall k \in \mathbb{Z}^b, \ \ 2|l| + |\alpha| + |\beta| \leq 2. \tag{4.1}$$

Decompose $P$ as $P = R + (P - R)$, with

$$R := \sum_{\substack{n^*_{\alpha\beta} \leq K_0\\ 2|l|+|\alpha|+|\beta|\leq 2}} P_{kl\alpha\beta} e^{\mathrm{i}\langle k,\theta\rangle} q^\alpha \bar{q}^\beta,$$

and then

$$P - R = \sum_{\substack{k,l,n^*_{\alpha\beta} > K_0\\ 1\leq 2|l|+|\alpha|+|\beta|\leq 2}} P_{kl\alpha\beta} e^{\mathrm{i}\langle k,\theta\rangle} I^l q^\alpha \bar{q}^\beta + \sum_{\substack{k,l\\ 2|l|+|\alpha|+|\beta|\geq 3}} P_{kl\alpha\beta} e^{\mathrm{i}\langle k,\theta\rangle} I^l q^\alpha \bar{q}^\beta.$$

It follows, from (4.1) and the definition of the vector field norm, that one can make $s_0$ small enough so that

$$\|X_{P-R}\|_{\mathcal{D}_0,\mathcal{O}} \leq \frac{1}{2}\varepsilon_0 = \frac{1}{2}\varepsilon^{\frac{5}{4}}.$$

We can rewrite $R$ as

$$\begin{aligned} R \ &= \ \sum_{\substack{k\\ |l|\leq 1}} P_{kl00} e^{\mathrm{i}\langle k,\theta\rangle} I^l + \sum_{\substack{k\\ |n|\leq K_0}} (P^{k10}_n q_n + P^{k01}_n \bar{q}_n) e^{\mathrm{i}\langle k,\theta\rangle} \\ &\quad + \sum_{\substack{k\\ |n|,|m|\leq K_0}} (P^{k20}_{nm} q_n q_m + P^{k11}_{nm} q_n \bar{q}_m + P^{k02}_{nm} \bar{q}_n \bar{q}_m) e^{\mathrm{i}\langle k,\theta\rangle}, \end{aligned}$$

where

$$P_n^{k10} := P_{k0e_n0}, \qquad P_n^{k01} := P_{k00e_n},$$
$$P_{nm}^{k20} := P_{kl(e_n+e_m)0}, \quad P_{nm}^{k11} := P_{kle_ne_m}, \quad P_{nm}^{k02} := P_{kl0(e_n+e_m)}.$$

The gauge invariance of $P$ implies that for all $n, m \in \mathbb{Z}_1$,

$$P_n^{010},\, P_n^{001},\, P_{nm}^{020},\, P_{nm}^{002} \equiv 0. \tag{4.2}$$

To handle terms of $R$, we need to construct a symplectic transformation $\Phi_* = \Phi_{F_*}^1$ defined as the time-1 map of the Hamiltonian flow associated with a real-analytic Hamiltonian $F_*$ of the form

$$\begin{aligned} F_* \;=\; & \sum_{\substack{k\neq 0\\ |l|\le 1}} F_{kl00} e^{\mathrm{i}\langle k,\theta\rangle} I^l + \sum_{\substack{k\neq 0\\ |n|\le K_0}} (F_n^{k10} q_n + F_n^{k01}\bar{q}_n) e^{\mathrm{i}\langle k,\theta\rangle} \\ & + \sum_{\substack{k\neq 0\\ |n|,|m|\le K_0}} (F_{nm}^{k20} q_n q_m + F_{nm}^{k11} q_n \bar{q}_m + F_{nm}^{k02} \bar{q}_n \bar{q}_m) e^{\mathrm{i}\langle k,\theta\rangle}, \end{aligned}$$

such that all non–resonant terms

$$\begin{aligned} & P_{kl00} I^l e^{\mathrm{i}\langle k,\theta\rangle}, \quad k\neq 0, \quad |l|\le 1, \\ & P_{k0\alpha\beta} e^{\mathrm{i}\langle k,\theta\rangle} q^\alpha \bar{q}^\beta, \quad k\neq 0, \quad n_{\alpha\beta}^* \le K_0, \quad 1\le |\alpha|+|\beta|\le 2, \end{aligned}$$

will be eliminated, and terms

$$P_{0l00} I^l, \quad |l|\le 1; \quad P_{nm}^{011} q_n \bar{q}_m, \quad |n|,|m|\le K_0,$$

will be added to the normal form part of the new Hamiltonian. More precisely, we shall construct $\Phi_{F_*}^1$ such that $F_*$ satisfies the homological equation

$$\{N, F_*\} + R = \sum_{|l|\le 1} P_{0l00} I^l + \sum_{|n|,|m|\le K_0} P_{nm}^{011} q_n \bar{q}_m.$$

One can show that it is solvable on the parameter set

$$\mathcal{O}_0 = \left\{ \xi\in\mathcal{O} : \begin{array}{l} |\langle k,\omega\rangle| \ge \frac{\gamma_0}{|k|^\tau}, \\ |\langle k,\omega\rangle + \Omega_n| \ge \frac{\gamma_0}{|k|^\tau K_0^2}, \\ |\langle k,\omega\rangle + \Omega_n + \Omega_m| \ge \frac{\gamma_0}{|k|^\tau K_0^4}, \\ |\langle k,\omega\rangle + \Omega_n - \Omega_m| \ge \frac{\gamma_0}{|k|^\tau K_0^4}, \end{array} \quad k\neq 0, \quad |n|,|m|\le K_0 \right\}.$$

By virtue of (4.2), which is guaranteed by gauge invariance of $P$, we need not consider the lower bound of $|\Omega_n|$ or $|\Omega_n \pm \Omega_m|$.

The parameter set satisfies that $|\mathcal{O}\setminus\mathcal{O}_0| = O(\gamma_0)$. Indeed, by the assumptions on $\omega$ and $\Omega_n$, we have

$$|\partial_\xi(\langle k,\omega\rangle + \Omega_m \pm \Omega_n)| \ge c|k|.$$

Therefore, by excluding some parameter set with measure $O(\gamma_0)$, we have that

$$|\langle k,\omega\rangle + \Omega_m \pm \Omega_n| \ge \frac{\gamma_0}{|k|^\tau K_0^4}.$$

The other conditions can be handled similarly.

With $\Phi_* = \Phi^1_{F_*}$, the Hamiltonian (2.4) can be transformed into $H_0 = H \circ \Phi_* = N_0 + P_0$ with

$$N_0 = e_0(\xi) + \langle \omega_0(\xi), I\rangle + \langle A_0(\xi) z_0, \bar{z}_0\rangle + \sum_{|n|>K_0} \Omega_n(\xi) q_n \bar{q}_n,$$
$$P_0 = \breve{P}_0 + \acute{P}_0 = \sum_{\alpha,\beta} \breve{P}^0_{\alpha\beta}(\theta, I; \xi) q^\alpha \bar{q}^\beta + \sum_{\alpha,\beta} \acute{P}^0_{\alpha\beta}(\xi) q^\alpha \bar{q}^\beta,$$

where $z_0 = (q_n)_{|n|\le K_0}$, $\bar{z}_0 = (\bar{q}_n)_{|n|\le K_0}$ and

$$e_0(\xi) = e(\xi) + P_{0000}(\xi),$$
$$\omega_0(\xi) = \omega(\xi) + P_{0l00(|l|=1)}(\xi),$$
$$\langle A_0(\xi) z_0, \bar{z}_0\rangle = \sum_{|n|\le K_0} \Omega_n(\xi) q_n \bar{q}_n + \sum_{|n|,|m|\le K_0} P^{011}_{nm}(\xi) q_n \bar{q}_m.$$

Moreover, $P_0$ satisfies $\|X_{P_0}\|_{\mathcal{D}_0,\mathcal{O}_0} \le \varepsilon^{\frac{5}{4}} = \varepsilon_0$ and

$$\|\breve{P}^0_{\alpha\beta}\|_{\mathcal{D}_0,\mathcal{O}_0} \le \begin{cases} \varepsilon_0 e^{-\rho_0 n^*_{\alpha\beta}}, & |\alpha|+|\beta| \le 2 \\ e^{-\rho_0 n^*_{\alpha\beta}}, & |\alpha|+|\beta| \ge 3 \end{cases},$$
$$\|\acute{P}^0_{\alpha\beta}\|_{\mathcal{D}_0,\mathcal{O}_0} \le \begin{cases} \varepsilon_0 e^{-\rho_0 n^*_{\alpha\beta}}, & |\alpha|+|\beta| \le 2 \\ e^{-\rho_0 (n^+_{\alpha\beta} - n^-_{\alpha\beta})}, & |\alpha|+|\beta| \ge 3 \end{cases}.$$

We shall prove that the decay property is preserved during the KAM iteration in Subsection 4.4.

Suppose that, we have arrived at the $\nu^{\text{th}}$ KAM step, and we consider the Hamiltonian $H_\nu = N_\nu + P_\nu$, which is real-analytic on $\mathcal{D}_\nu = \mathcal{D}_{d,\rho_\nu}(r_\nu, s_\nu)$, and $C^1_W$ parametrized by $\xi \in \mathcal{O}_\nu$, with

$$N_\nu = e_\nu(\xi) + \langle \omega_\nu(\xi), I\rangle + \langle A_\nu(\xi) z_\nu, \bar{z}_\nu\rangle + \sum_{|n|>K_\nu} \Omega_n(\xi) q_n \bar{q}_n,$$
$$P_\nu = \breve{P}_\nu + \acute{P}_\nu = \sum_{\alpha,\beta} \breve{P}^\nu_{\alpha\beta}(\theta, I; \xi) q^\alpha \bar{q}^\beta + \sum_{\alpha,\beta} \acute{P}^\nu_{\alpha\beta}(\xi) q^\alpha \bar{q}^\beta,$$

where $z_\nu = (q_n)_{|n|\le K_\nu}$, $\bar{z}_\nu = (\bar{q}_n)_{|n|\le K_\nu}$. Moreover, $P_\nu$ satisfies that $\|X_{P_\nu}\|_{\mathcal{D}_\nu,\mathcal{O}_\nu} < \varepsilon_\nu$ and

$$\|\breve{P}^\nu_{\alpha\beta}\|_{\mathcal{D}_\nu,\mathcal{O}_\nu} \le \begin{cases} \varepsilon_\nu e^{-\rho_\nu n^*_{\alpha\beta}}, & |\alpha|+|\beta| \le 2 \\ e^{-\rho_\nu n^*_{\alpha\beta}}, & |\alpha|+|\beta| \ge 3 \end{cases}, \tag{4.3}$$
$$\|\acute{P}^\nu_{\alpha\beta}\|_{\mathcal{D}_\nu,\mathcal{O}_\nu} \le \begin{cases} \varepsilon_\nu e^{-\rho_\nu n^*_{\alpha\beta}}, & |\alpha|+|\beta| \le 2 \\ e^{-\rho_\nu (n^+_{\alpha\beta} - n^-_{\alpha\beta})}, & |\alpha|+|\beta| \ge 3 \end{cases}. \tag{4.4}$$

In what follows, we shall construct a subset $\mathcal{O}_{\nu+1} \subset \mathcal{O}_\nu$, and a symplectic transformation $\Phi_\nu : \mathcal{D}_{\nu+1} \to \mathcal{D}_\nu$, so that the Hamiltonian $H_{\nu+1} = H_\nu \circ \Phi_\nu = N_{\nu+1} + P_{\nu+1}$, $C^1_W$ parametrized by $\xi \in \mathcal{O}_{\nu+1}$, has similar properties with $H_\nu$, and

$$\|X_{P_{\nu+1}}\|_{\mathcal{D}_{\nu+1},\mathcal{O}_{\nu+1}} \le \varepsilon_\nu^{\frac{5}{4}} = \varepsilon_{\nu+1}.$$

From now on, to simplify notations, the subscripts (or superscripts) "$\nu$" of quantities at the $\nu^{\rm th}$ step are neglected, and the corresponding quantities at the $(\nu+1)^{\rm th}$ step are labeled with "+". In addition, all constants labeled with $c$, $c_0$, $c_1$, $\cdots$ are positive and independent of the iteration step.

Let $K_+ = 2|\ln\varepsilon|K$. In the KAM step detailed below, terms with $(q_n, \bar{q}_n)_{K<|n|\le K_+}$ will be added to the new normal components $z_+$, $\bar{z}_+$. To facilitate the calculations when solving a homological equation later on, we will also adopt the following expression of $N$,

$$\begin{aligned} N &= e(\xi) + \langle \omega(\xi), I\rangle + \langle A(\xi)z, \bar{z}\rangle + \sum_{K<|n|\le K_+} \Omega_n(\xi) q_n\bar{q}_n + \sum_{|n|>K_+} \Omega_n(\xi) q_n\bar{q}_n \\ &= e(\xi) + \langle \omega(\xi), I\rangle + \langle \tilde{A}(\xi)z_+, \bar{z}_+\rangle + \sum_{|n|>K_+} \Omega_n(\xi) q_n\bar{q}_n, \end{aligned}$$

where $\tilde{A}$ is a Hermitian matrix with $\dim(\tilde{A}) \le 2K_+ + 1$ given by

$$\tilde{A} = \begin{pmatrix} A & 0 \\ 0 & \Omega_n \end{pmatrix}_{K<|n|\le K_+} \tag{4.5}$$

and $z_+ = (q_n)_{|n|\le K_+}$, $\bar{z}_+ = (\bar{q}_n)_{|n|\le K_+}$.

### 4.2 Truncation and homological equation

Expand $\breve{P}$ and $\acute{P}$ into their Taylor-Fourier series,

$$\breve{P} = \sum_{\substack{(k,l)\neq 0\\ \alpha,\beta}} P_{kl\alpha\beta} e^{{\rm i}\langle k,\theta\rangle} I^l q^\alpha \bar{q}^\beta, \quad \acute{P} = \sum_{\alpha,\beta} P_{00\alpha\beta} q^\alpha \bar{q}^\beta.$$

By (4.3) and (4.4), and the definition of norm $\|\cdot\|_{\mathcal{D},\mathcal{O}}$,

$$|P_{kl\alpha\beta}|_{\mathcal{O}} \le \varepsilon e^{-\rho n^*_{\alpha\beta}} e^{-|k|r}, \quad \forall k\in\mathbb{Z}^b, \;\; 2|l| + |\alpha| + |\beta| \le 2. \tag{4.6}$$

Associated with terms in the normal form $N$, let $R$ be the following truncation of $P$:

$$R(\theta, I, z_+, \bar{z}_+) = \sum_{\substack{2|l|+|\alpha|+|\beta|\le 2\\ n^*_{\alpha\beta}\le K_+}} P_{kl\alpha\beta} e^{{\rm i}\langle k,\theta\rangle} I^l q^\alpha \bar{q}^\beta = R_0 + R_1 + R_2,$$

with

$$\begin{aligned} R_0 &= \sum_{\substack{k\\ |l|\le 1}} P_{kl00} e^{{\rm i}\langle k,\theta\rangle} I^l, \\ R_1 &= \sum_{\substack{k\\ |n|\le K_+}} (P^{k10}_n q_n + P^{k01}_n \bar{q}_n) e^{{\rm i}\langle k,\theta\rangle} =: \sum_k (\langle R^{k10}, z_+\rangle + \langle R^{k01}, \bar{z}_+\rangle) e^{{\rm i}\langle k,\theta\rangle} \\ R_2 &= \sum_{\substack{k\\ |n|,|m|\le K_+}} (P^{k20}_{nm} q_n q_m + P^{k11}_{nm} q_n\bar{q}_m + P^{k02}_{nm} \bar{q}_n\bar{q}_m) e^{{\rm i}\langle k,\theta\rangle} \\ &=: \sum_k (\langle R^{k20} z_+, z_+\rangle + \langle R^{k11} z_+, \bar{z}_+\rangle + \langle R^{k02} \bar{z}_+, \bar{z}_+\rangle) e^{{\rm i}\langle k,\theta\rangle}, \end{aligned}$$

where $R^{k10}$, $R^{k01}$, $R^{k20}$, $R^{k11}$, $R^{k02}$ are defined as

$$R^{k10} := \left(P_n^{k10}\right)_{|n|\le K_+}, \qquad R^{k01} := \left(P_n^{k01}\right)_{|n|\le K_+},$$
$$R^{k20} := \left(P_{nm}^{k20}\right)_{|n|,|m|\le K_+}, \quad R^{k11} := \left(P_{nm}^{k11}\right)_{|n|,|m|\le K_+}, \quad R^{k01} := \left(P_{nm}^{k01}\right)_{|n|,|m|\le K_+}.$$

Since $\bar{P} = P$, it is clear that

$$\begin{array}{lll} \overline{P_{(-k)l00}} = P_{kl00}, & \overline{R^{(-k)10}} = R^{k01}, & \overline{R^{(-k)01}} = R^{k10}, \\ \overline{R^{(-k)20}} = R^{k02}, & \overline{R^{(-k)11}}^{\top} = R^{k11}, & \overline{R^{(-k)02}} = R^{k20}. \end{array} \tag{4.7}$$

From our definition of norms, it follows that

$$\|X_R\|_{\mathcal{D},\mathcal{O}} \le \|X_P\|_{\mathcal{D},\mathcal{O}} \le \varepsilon.$$

Let $\rho_+ = K_+^{-1}$, $r_+ = \frac{r}{2} + \frac{r_0}{4}$ and $\eta = \varepsilon^{\frac{1}{4}}$. Since

$$P - R = \sum_{\substack{k,l \\ 2|l|+|\alpha|+|\beta|\ge 3}} P_{kl\alpha\beta} e^{\mathrm{i}\langle k,\theta\rangle} I^l q^\alpha \bar{q}^\beta + \sum_{\substack{k,l,\, n^*_{\alpha\beta} > K_+ \\ 2|l|+|\alpha|+|\beta|\le 2}} P_{kl\alpha\beta} e^{\mathrm{i}\langle k,\theta\rangle} I^l q^\alpha \bar{q}^\beta, \tag{4.8}$$

combining with (4.6), there exists $c_1 > 0$ such that

$$\|X_{P-R}\|_{\mathcal{D}_{d,\rho_+}(r_+ + \frac{r-r_+}{2}, \eta s), \mathcal{O}} \le \varepsilon \sum_{|n|>K_+} e^{-(\rho-\rho_+)|n|} + c_1 \eta s \le \frac{1}{4}\varepsilon^{\frac{5}{4}}, \tag{4.9}$$

provided that

**(C1)** $e^{-(\rho-\rho_+)K_+} \le \frac{1}{8}\varepsilon^{\frac{1}{4}}, \quad c_1 s \le \frac{1}{8}\varepsilon.$

We are going to construct a Hamiltonian $F$, defined on a new domain $\mathcal{D}_+ = \mathcal{D}_{d,\rho_+}(r_+, s_+)$ such that, the time-1 map $\Phi = \Phi_F^1$ associated with the Hamiltonian vector field $X_F$, is a (symplectic) map from $\mathcal{D}_+$ to $\mathcal{D}$ which transforms $H$ into $H_+$, the Hamiltonian in the next KAM cycle. Let $F$ be of the form

$$F(\theta, I, z_+, \bar{z}_+) \ = \ F_0 + F_1 + F_2,$$

with

$$\begin{aligned} F_0 \ &= \ \sum_{\substack{k\ne 0 \\ |l|\le 1}} F_{kl00} e^{\mathrm{i}\langle k,\theta\rangle} I^l, \\ F_1 \ &= \ \sum_{\substack{k\ne 0 \\ |n|\le K_+}} (F_n^{k10} q_n + F_n^{k01} \bar{q}_n) e^{\mathrm{i}\langle k,\theta\rangle} =: \sum_{k\ne 0} (\langle F^{k10}, z_+\rangle + \langle F^{k01}, \bar{z}_+\rangle) e^{\mathrm{i}\langle k,\theta\rangle}, \\ F_2 \ &= \ \sum_{\substack{k\ne 0 \\ |n|,|m|\le K_+}} (F_{nm}^{k20} q_n q_m + F_{nm}^{k11} q_n \bar{q}_m + F_{nm}^{k02} \bar{q}_n \bar{q}_m) e^{\mathrm{i}\langle k,\theta\rangle} \\ &=: \ \sum_{k\ne 0} (\langle F^{k20} z_+, z_+\rangle + \langle F^{k11} z_+, \bar{z}_+\rangle + \langle F^{k02} \bar{z}_+, \bar{z}_+\rangle) e^{\mathrm{i}\langle k,\theta\rangle}, \end{aligned}$$

and satisfy the homological equation

$$\{N, F\} + R = e' + \langle \omega', I \rangle + \langle R^{011} z_+, \bar{z}_+ \rangle, \tag{4.10}$$

where $e' = P_{0000}$ and $\omega' = P_{0l00}(|l| = 1)$. By simple comparison of coefficients, we can see Equation (4.10) is equivalent to the following system

$$\langle k, \omega \rangle F_{kl00} = \mathrm{i} P_{kl00}, \tag{4.11}$$

$$(\langle k, \omega \rangle I - \tilde{A}) F^{k10} = \mathrm{i} R^{k10}, \tag{4.12}$$

$$(\langle k, \omega \rangle I + \tilde{A}) F^{k01} = \mathrm{i} R^{k01}, \tag{4.13}$$

$$(\langle k, \omega \rangle I - \tilde{A}) F^{k20} - F^{k20} \tilde{A} = \mathrm{i} R^{k20}, \tag{4.14}$$

$$(\langle k, \omega \rangle I - \tilde{A}) F^{k11} + F^{k11} \tilde{A} = \mathrm{i} R^{k11}, \tag{4.15}$$

$$(\langle k, \omega \rangle I + \tilde{A}) F^{k02} + F^{k02} \tilde{A} = \mathrm{i} R^{k02} \tag{4.16}$$

for every $k \neq 0$ and $|l| \leq 1$.

Since $\tilde{A}$ is Hermitian, there is a unitary matrix $Q$ such that

$$Q^* \tilde{A} Q = \Lambda := \mathrm{diag}\{\mu_j\}_{|j| \leq K_+},$$

where $\{\mu_j\}_{|j| \leq K_+}$ denote the eigenvalues of $\tilde{A}$. In addition, by (4.5), the eigenvalues of $A$ are all labeled with $|j| \leq K$, and $\mu_j = \Omega_j$ for $K < |j| \leq K_+$. Due to the block-diagonal structure of $\tilde{A}$ in (4.5), we have that

$$Q_{mn} \equiv 0 \;\text{ if }\; |m - n| > 2K + 1. \tag{4.17}$$

Indeed, the diagonalization of $\tilde{A}$ is just the diagonalization of $A$.

Define the new parameter set $\mathcal{O}_+ \subset \mathcal{O}$ as

$$\mathcal{O}_+ := \left\{ \xi \in \mathcal{O} : \begin{array}{l} |\langle k, \omega \rangle| > \frac{\gamma}{|k|^\tau}, \\ |\langle k, \omega \rangle I + \mu_n| > \frac{\gamma}{|k|^\tau K_+^2}, \\ |\langle k, \omega \rangle I + \mu_n + \mu_m| > \frac{\gamma}{|k|^\tau K_+^4}, \\ |\langle k, \omega \rangle I + \mu_n - \mu_m| > \frac{\gamma}{|k|^\tau K_+^4}, \end{array} \quad k \neq 0, \;\; |n|, |m| \leq K_+ \right\}.$$

The same as the construction of $\mathcal{O}_0$ in Subsection 4.1, we need not consider the lower bound of $|\mu_n|$ or $|\mu_n \pm \mu_m|$, in view of gauge invariance of $P$.

Obviously, (4.11) can be solved on $\mathcal{O}_+$. As for solvability of (4.12) − (4.16), let us define the vectors $\tilde{R}^{k10}$, $\tilde{R}^{k01}$ and the matrices $\tilde{R}^{k20}$, $\tilde{R}^{k11}$, $\tilde{R}^{k02}$ as

$$\begin{aligned} &\tilde{R}^{k10} := Q^* R^{k10}, \quad\;\; \tilde{R}^{k01} := Q^* R^{k01}, \\ &\tilde{R}^{k20} := Q^* R^{k20} Q, \quad \tilde{R}^{k11} := Q^* R^{k11} Q, \quad \tilde{R}^{k02} := Q^* R^{k02} Q. \end{aligned}$$

for $k \neq 0$. We consider the equations

$$\begin{aligned} &(\langle k, \omega \rangle I - \Lambda) \tilde{F}^{k10} = \mathrm{i} \tilde{R}^{k10}, \\ &(\langle k, \omega \rangle I + \Lambda) \tilde{F}^{k01} = \mathrm{i} \tilde{R}^{k01}, \\ &(\langle k, \omega \rangle I - \Lambda) \tilde{F}^{k20} - \tilde{F}^{k20} \Lambda = \mathrm{i} \tilde{R}^{k20}, \\ &(\langle k, \omega \rangle I - \Lambda) \tilde{F}^{k11} + \tilde{F}^{k11} \Lambda = \mathrm{i} \tilde{R}^{k11}, \\ &(\langle k, \omega \rangle I + \Lambda) \tilde{F}^{k02} + \tilde{F}^{k02} \Lambda = \mathrm{i} \tilde{R}^{k02}. \end{aligned}$$

These equations is equivalent to

$$
\begin{aligned}
(\langle k,\omega\rangle I-\mu_n)\tilde F^{k10}_n &= \mathrm{i}\tilde R^{k10}_n,\\
(\langle k,\omega\rangle I+\mu_n)\tilde F^{k01}_n &= \mathrm{i}\tilde R^{k01}_n,\\
(\langle k,\omega\rangle I-\mu_n-\mu_m)\tilde F^{k20}_{nm} &= \mathrm{i}\tilde R^{k20}_{nm},\\
(\langle k,\omega\rangle I-\mu_n+\mu_m)\tilde F^{k11}_{nm} &= \mathrm{i}\tilde R^{k11}_{nm},\\
(\langle k,\omega\rangle I+\mu_n+\mu_m)\tilde F^{k02}_{nm} &= \mathrm{i}\tilde R^{k02}_{nm},
\end{aligned}
$$

for $k\neq 0$, $|n|,|m|\le K_+$, which can be solved on $\mathcal{O}_+$. Then (4.12) – (4.16) are also solved with

$$
\begin{aligned}
&F^{k10}:=Q\tilde F^{k10}, && F^{k01}:=Q\tilde F^{k01},\\
&F^{k20}:=Q\tilde F^{k20}Q^*, && F^{k11}:=Q\tilde F^{k11}Q^*, \quad F^{k02}:=Q\tilde F^{k02}Q^*.
\end{aligned}
$$

By (4.7), it is easy to show that

$$
\begin{aligned}
&\overline{F_{(-k)l00}}=F_{kl00}, && \overline{F^{(-k)10}}=F^{k01}, && \overline{F^{(-k)01}}=F^{k10},\\
&\overline{F^{(-k)20}}=F^{k02}, && (F^{(-k)11})^*=F^{k11}, && \overline{F^{(-k)02}}=F^{k20}.
\end{aligned}
$$

Thus $\bar F = F$.

## 4.3 Property of the coordinate transformation

**Lemma 4.1** *$F$ has gauge invariance, and for $\varepsilon$ sufficiently small, the coefficients of $F$ satisfy that*

$$|F_{kl00}|_{\mathcal{O}_+}\le \varepsilon^{\frac56}|k|^{2\tau+1}e^{-|k|r}, \tag{4.18}$$

$$|F^{k10}_n|_{\mathcal{O}_+},|F^{k01}_n|_{\mathcal{O}_+}\le \varepsilon^{\frac56}|k|^{2\tau+1}e^{-|k|r}e^{-\rho|n|}, \tag{4.19}$$

$$|F^{k20}_{nm}|_{\mathcal{O}_+},|F^{k11}_{nm}|_{\mathcal{O}_+},|F^{k02}_{nm}|_{\mathcal{O}_+}\le \varepsilon^{\frac56}|k|^{2\tau+1}e^{-|k|r}e^{-\rho\max\{|n|,|m|\}}. \tag{4.20}$$

*Proof:* Let us first consider $F^{k20}_{mn}$ for instance, with other terms in (4.19) and (4.20) analogous. By the construction above, we can present $F^{k20}_{mn}$ as

$$F^{k20}_{nm}=\mathrm{i}\sum_{\mathcal{F}}\frac{Q_{nn_1}Q^*_{n_1n_2}R^{k20}_{n_2n_3}Q_{n_3n_4}Q^*_{n_4m}}{\langle k,\omega\rangle-\mu_{n_1}-\mu_{n_4}}, \tag{4.21}$$

where the summation notation $\mathcal{F}$ denotes

$$\left\{\begin{array}{l}|n_1|,|n_2|,|n_3|,|n_4|\le K_+,\\ |n_1-n|,\ |n_2-n_1|\le 2K+1,\quad |n_4-m|,\ |n_3-n_4|\le 2K+1\end{array}\right\},$$

by virtue of the structure of $Q$ in (4.17). Then by (4.6),

$$\sup_{\xi\in\mathcal{O}_+}|F^{k20}_{nm}(\xi)|\le c(\gamma^{-1}|k|^\tau K_+^4)K^4e^{(2K+1)\rho}\varepsilon e^{-\rho\max\{|n|,|m|\}}e^{-|k|r}.$$

Here we have applied the property of the orthogonal matrix $Q$, and used the factor $e^{(2K+1)\rho}$ to recover the exponential decay.

To estimate $|\partial_{\xi_j} F^{k20}_{nm}|$, we need to differentiate both sides of (4.14) with respect to $\xi_j$, $j = 1, 2, \cdots, b$. Then we obtain the equation about $\partial_{\xi_j} F^{k20}$

$$(\langle k, \omega\rangle I - \tilde{A})(\partial_{\xi_j} F^{k20}) - (\partial_{\xi_j} F^{k20})\tilde{A} = G^{k20}_{\xi_j},$$

which can be solved by diagonalizing $\tilde{A}$ via $Q$ as above, where

$$G^{k20}_{\xi_j} := \mathrm{i}\partial_{\xi_j} R^{k20} + F^{k20}(\partial_{\xi_j}\tilde{A}) - [\partial_{\xi_j}(\langle k, \omega\rangle I - \tilde{A})]F^{k20}.$$

Just like (4.21), we get the formulation

$$\partial_{\xi_j} F^{k20}_{nm} = \sum_{\mathcal{F}} \frac{Q_{nn_1} Q^*_{n_1 n_2} (G^{k20}_{\xi_j})_{n_2 n_3} Q_{n_3 n_4} Q^*_{n_4 m}}{\langle k, \omega\rangle - \mu_{n_1} - \mu_{n_4}}.$$

By the decay property of $R^{k20}$ and the construction of $\tilde{A}$, we have that

$$\sup_{\xi\in\mathcal{O}_+} |(G^{k20}_{\xi_j})_{nm}| \leq c(\gamma^{-1}|k|^{\tau+1}K_+^4)K^5 e^{(4K+2)\rho}\varepsilon e^{-\rho\max\{|n|,|m|\}} e^{-|k|r}.$$

Thus there exists $c_2 > 0$ such that

$$\begin{aligned} & \sup_{\xi\in\mathcal{O}_+} (|F^{k20}_{nm}| + |\partial_\xi F^{k20}_{nm}|) \\ \leq\ & c_2(\gamma^{-2}|k|^{2\tau+1}K_+^8)K^9 e^{(6K+3)\rho}\varepsilon e^{-\rho\max\{|n|,|m|\}} e^{-|k|r} \\ \leq\ & \varepsilon^{\frac{5}{6}}|k|^{2\tau+1} e^{-\rho\max\{|n|,|m|\}} e^{-|k|r}. \end{aligned}$$

It is easy to see that

$$|F_{kl00}|_{\mathcal{O}_+} \leq |\langle k, \omega\rangle|^{-2}|k||P_{kl00}|_{\mathcal{O}_+} \leq \gamma^{-2}|k|^{2\tau+1}e^{-|k|r}\varepsilon, \quad k \neq 0, \quad |l| \leq 1,$$

by the definition of $\mathcal{O}_+$. Thus, (4.18) – (4.20) hold under the assumption

**(C2)** $c_2\gamma^{-2}K_+^8 K^9 e^{(6K+3)\rho}\varepsilon^{\frac{1}{6}} \leq 1$.

Suppose that $\sum_{j=1}^b k_j + 2 \neq 0$, which means $R^{k20} \equiv 0$. By the formulation of $F^{k20}_{mn}$ in (4.21), $F^{k20} \equiv 0$. Doing the same thing for $F^{k11}$, $F^{k02}$, $F^{k10}$, $F^{k01}$ as above, we obtain the gauge invariance of $F$. ■

We proceed to estimate the norm of $X_F$ and to study properties of $\Phi^1_F$, on domains $\mathcal{D}_i := \mathcal{D}_{d,\rho_+}(r_+ + \frac{i}{4}(r - r_+), \frac{i}{4}s)$, $i = 1, 2, 3, 4$.

**Lemma 4.2** *For $\varepsilon$ sufficiently small, we have $\|X_F\|_{\mathcal{D}_3,\mathcal{O}_+} \leq \varepsilon^{\frac{4}{5}}$.*

*Proof:* In view of (4.18) – (4.20), it follows that

$$\frac{1}{s^2}\|\partial_\theta F\|_{\mathcal{D}_3,\mathcal{O}_+}, \|\partial_I F\|_{\mathcal{D}_3,\mathcal{O}_+} \leq c(r - r_+)^{-(2\tau+b+1)}\varepsilon^{\frac{5}{6}},$$

and

$$
\begin{aligned}
& \sup_{\mathcal{D}_3} \frac{1}{s} \sum_{n\in\mathbb{Z}_1} \left(\|\partial_{q_n} F\|_{\mathcal{O}_+} + \|\partial_{\bar{q}_n} F\|_{\mathcal{O}_+}\right) \langle n\rangle^d e^{\rho_+|n|} \\
\leq\ & \sup_{\mathcal{D}_3} \frac{c}{s} \sum_{\substack{k\neq 0\\ |n|\leq K_+}} \left(|F_n^{k10}|_{\mathcal{O}_+} + |F_n^{k01}|_{\mathcal{O}_+}\right) e^{|k|(r-\frac{1}{4}(r-r_+))} \langle n\rangle^d e^{\rho_+|n|} \\
& + \sup_{\mathcal{D}_3} \frac{c}{s} \sum_{\substack{k\neq 0\\ |n|,|m|\leq K_+}} \left(|F_{mn}^{k20}|_{\mathcal{O}_+} + |F_{mn}^{k11}|_{\mathcal{O}_+} + |F_{mn}^{k02}|_{\mathcal{O}_+}\right)|q_m| e^{|k|(r-\frac{1}{4}(r-r_+))} \langle n\rangle^d e^{\rho_+|n|} \\
\leq\ & c(r-r_+)^{-(2\tau+b+1)} K_+^d e^{\rho_+ K_+} \varepsilon^{\frac{5}{6}}.
\end{aligned}
$$

Putting together the estimates above, there exists a constant $c_3$ such that

$$
\|X_F\|_{\mathcal{D}_3,\mathcal{O}_+} \leq c_3 (r-r_+)^{-(2\tau+b+1)} K_+^d e^{\rho_+ K_+} \varepsilon^{\frac{5}{6}}.
$$

Moreover, if

**(C3)** $c_3(r-r_+)^{-(2\tau+b+1)} K_+^d e^{\rho_+ K_+} \varepsilon^{\frac{1}{30}} \leq 1,$

then Lemma 4.2 follows. ∎

Now let $\mathcal{D}_{i\eta} := \mathcal{D}_{d,\rho_+}(r_+ + \frac{i}{4}(r-r_+), \frac{i}{4}\eta s)$, $i = 1,2,3,4$.

**Lemma 4.3** *For $\varepsilon$ sufficiently small, we have $\Phi_F^t : \mathcal{D}_{2\eta} \to \mathcal{D}_{3\eta}$, $-1 \leq t \leq 1$, and moreover,*

$$
\|D\Phi_F^t - Id\|_{\mathcal{D}_{1\eta}} < 2\varepsilon^{\frac{4}{5}}.
$$

*Proof:* Let

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

Notice that $F$ is a polynomial of order 1 in $I$ and of order 2 in $z_+$, $\bar{z}_+$. It thus follows from Lemma 4.2 and Cauchy inequality (Lemma 2.2 in Section 2) that

$$
\|D^m F\|_{\mathcal{D}_2,\mathcal{O}_+} < \varepsilon^{\frac{4}{5}}, \quad \forall m \geq 2.
$$

Using the integral equation

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

and Lemma 4.2, one sees easily that $\Phi_F^t : \mathcal{D}_{2\eta} \to \mathcal{D}_{3\eta}$, $-1 \leq t \leq 1$. Moreover, 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, it follows that

$$
\|D\Phi_F^t - Id\|_{\mathcal{D}_{1\eta}} \leq 2\|D^2 F\|_{\mathcal{D}_{2\eta}} \leq 2\varepsilon^{\frac{4}{5}}.
$$

∎

### 4.4 Estimation for the new Hamiltonian

Let $\Phi = \Phi_F^1$ , $s_+ = \frac{1}{8}\eta s$ , $\mathcal{D}_+ = \mathcal{D}_{d,\rho_+}(r_+, s_+)$ and

$$N_+ = e_+ + \langle \omega_+, I\rangle + \langle A_+ z_+, \bar{z}_+\rangle + \sum_{|n|>K_+} \Omega_n q_n \bar{q}_n,$$

where $e_+ = e + e'$, $\omega_+ = \omega + \omega'$, $A_+ = \tilde{A} + R^{011}$. Then $\Phi : \mathcal{D}_+ \to \mathcal{D}$ and, by Taylor's second-order formula,

$$\begin{aligned} H_+ &:= H \circ \Phi = (N + R) \circ \Phi + (P - R) \circ \Phi \\ &= N + \{N, F\} + R + \int_0^1 (1 - t)\{\{N, F\}, F\} \circ \Phi_F^t \, dt \\ &\quad + \int_0^1 \{R, F\} \circ \Phi_F^t \, dt + (P - R) \circ \Phi_F^1 \\ &= N + \{N, F\} + R + P_+ \\ &= N_+ + P_+ + \{N, F\} + R - e' - \langle \omega', I\rangle - \langle R^{011} z_+, \bar{z}_+\rangle \\ &= N_+ + P_+, \end{aligned}$$

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

The new normal form $N_+$ has properties similar to those of $N$. Observe that, since $\tilde{A}^* = \tilde{A}$ and $(R^{011})^* = R^{011}$, we have $A_+^* = A_+$, i.e., $A_+$ is a Hermitian matrix. Then, from the assumptions on $\breve{P}$ and $\acute{P}$, we further have that

$$|\omega_+ - \omega|_{\mathcal{O}_+} \le \varepsilon, \quad |(A_+ - \tilde{A})_{nm}|_{\mathcal{O}_+} \le \varepsilon e^{-\rho \max\{|n|,|m|\}}, \tag{4.22}$$

which will be used for the measure estimates. The eigenvalues of $A_+$, $\{\mu_j^+\}_{|j|\le K_+}$, can be labeled with $|\mu_j^+ - \mu_j|_{\mathcal{O}_+} \le c\varepsilon$ in view of the min-max principle[38].

Let $R(t) = (1 - t)(N_+ - N) + tR$. Then $P_+$ can be rewritten as

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

Hence, $X_{P_+} = \int_0^1 (\Phi_F^t)^* X_{\{R(t),F\}} \, dt + (\Phi_F^1)^* X_{(P-R)}$. By Lemma 4.3,

$$\|D\Phi_F^t\|_{\mathcal{D}_{1\eta}} \le 1 + \|D\Phi_F^t - I\|_{\mathcal{D}_{1\eta}} \le 2, \quad -1 \le t \le 1.$$

Furthermore, by Lemma 2.3, we also have

$$\|X_{\{R(t),F\}}\|_{\mathcal{D}_{2\eta}} \le c\eta^{-2}\varepsilon^{\frac{9}{5}} = \frac{1}{4}\varepsilon^{\frac{5}{4}}.$$

Then, combining with (4.9), $\|X_{P_+}\|_{\mathcal{D}_+,\mathcal{O}_+} \le \varepsilon^{\frac{5}{4}} = \varepsilon_+$.

Note that

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

The reality of $P_+$ is verified easily because, for any two function $F$ and $G$ satisfying $\bar{F} = F$ and $\bar{G} = G$ respectively, their Poisson bracket $\{F,G\}$ satisfies $\overline{\{F,G\}} = \{\overline{F},\overline{G}\} = \{F,G\}$.

It has been proved that the gauge invariance is preserved during the KAM iteration by Lemma 2.4, so we only need to examine the decay property of $P_+$. More precisely, if we decompose $P_+$ as $P_+ = \breve{P}_+ + \acute{P}_+$ with

$$\breve{P}_+ = \sum_{\alpha,\beta} \breve{P}^+_{\alpha\beta}(\theta, I;\xi) q^\alpha \bar{q}^\beta, \quad \acute{P}_+ = \sum_{\alpha,\beta} \acute{P}^+_{\alpha\beta}(\xi) q^\alpha \bar{q}^\beta,$$

we will show that

$$\|\breve{P}^+_{\alpha\beta}\|_{\mathcal{D}_+,\mathcal{O}_+} \le \begin{cases} \varepsilon_+ e^{-\rho_+ n^*_{\alpha\beta}}, & |\alpha|+|\beta| \le 2 \\ e^{-\rho_+ n^*_{\alpha\beta}}, & |\alpha|+|\beta| \ge 3 \end{cases},$$

$$\|\acute{P}^+_{\alpha\beta}\|_{\mathcal{D}_+,\mathcal{O}_+} \le \begin{cases} \varepsilon_+ e^{-\rho_+ n^*_{\alpha\beta}}, & |\alpha|+|\beta| \le 2 \\ e^{-\rho_+ (n^+_{\alpha\beta} - n^-_{\alpha\beta})}, & |\alpha|+|\beta| \ge 3 \end{cases}.$$

For terms of $P - R$ in (4.8), we have

$$\|\breve{P}_{\alpha\beta}\|_{\mathcal{D}_+,\mathcal{O}_+} \le e^{-\rho n^*_{\alpha\beta}}, \ \ \|\acute{P}_{\alpha\beta}\|_{\mathcal{D}_+,\mathcal{O}_+} \le e^{-\rho(n^+_{\alpha\beta} - n^-_{\alpha\beta})}, \quad |\alpha|+|\beta| \ge 3.$$

If $|\alpha|+|\beta| \le 2$, then by **(C1)** and $n^*_{\alpha\beta} > K_+$,

$$\|\breve{P}_{\alpha\beta}\|_{\mathcal{D}_+,\mathcal{O}_+}, \ \|\acute{P}_{\alpha\beta}\|_{\mathcal{D}_+,\mathcal{O}_+} \le \varepsilon e^{-\rho n^*_{\alpha\beta}} \le \varepsilon e^{-(\rho-\rho_+)K_+} \cdot e^{-\rho_+ n^*_{\alpha\beta}} \le \frac{1}{2}\varepsilon_+ e^{-\rho_+ n^*_{\alpha\beta}}.$$

Here we applied the estimate $|I| \le s_+ \le \frac{1}{8}\varepsilon_+$ to handle the case that $|\alpha|+|\beta| \le 2$ and $2|l| + |\alpha| + |\beta| \ge 3$.

The decay property of remaining terms, which are made up of several Poisson brackets, is covered by the following lemma.

**Lemma 4.4** *For $\varepsilon$ sufficiently small, we have*

$$\|\{P,F\}_{\alpha\beta}\|_{\mathcal{D}_{3\eta},\mathcal{O}_+} \le \frac{1}{4}\varepsilon^{\frac{1}{4}} \begin{cases} \varepsilon e^{-\rho n^*_{\alpha\beta}}, & |\alpha|+|\beta| \le 2 \\ e^{-\rho n^*_{\alpha\beta}}, & |\alpha|+|\beta| \ge 3 \end{cases}.$$

*Proof:* A straightforward calculation yields that

$$\{P,F\}_{\alpha\beta} = \mathrm{i} \sum_{\substack{|n|\le K_+ \\ (\check{\alpha},\check{\beta})+(\hat{\alpha},\hat{\beta})=(\alpha,\beta)}} \left(P_{\check{\alpha}+e_n,\check{\beta}} F_{\hat{\alpha},\hat{\beta}+e_n} - P_{\check{\alpha},\check{\beta}+e_n} F_{\hat{\alpha}+e_n,\hat{\beta}}\right) \tag{4.23}$$

$$+ \sum_{(\check{\alpha},\check{\beta})+(\hat{\alpha},\hat{\beta})=(\alpha,\beta)} \left\{P_{\check{\alpha}\check{\beta}}, F_{\hat{\alpha}\hat{\beta}}\right\}. \tag{4.24}$$

In view of Lemma 4.1, we know that $\|F_{\alpha\beta}\|_{\mathcal{D}_3,\mathcal{O}_+} \le \varepsilon^{\frac{4}{5}} e^{-\rho n^*_{\alpha\beta}}$.

(1) Terms in (4.23)

Let us first consider the term $P_{\check{\alpha}+e_n,\check{\beta}} F_{\hat{\alpha},\hat{\beta}+e_n}$, which contains $\breve{P}_{\check{\alpha}+e_n,\check{\beta}} F_{\hat{\alpha},\hat{\beta}+e_n}$ and $\acute{P}_{\check{\alpha}+e_n,\check{\beta}} F_{\hat{\alpha},\hat{\beta}+e_n}$. In view of the construction of $F$, we have that

$$|\hat{\alpha}| + |\hat{\beta} + e_n| = 1 \ \text{ or } \ 2. \tag{4.25}$$

i) $|\alpha| + |\beta| \leq 2$

In this case, $|\check{\alpha} + e_n| + |\check{\beta}| = |\alpha| + |\beta| + 1 - (|\hat{\alpha}| + |\hat{\beta}|) \leq 3$.

- If $|\check{\alpha} + e_n| + |\check{\beta}| \leq 2$, then, noting that $n^*_{\alpha\beta} \leq \max\{n^*_{\check{\alpha}+e_n,\check{\beta}},\, n^*_{\hat{\alpha},\hat{\beta}+e_n}\}$, we have

$$\begin{aligned} \|\breve{P}_{\check{\alpha}+e_n,\check{\beta}} F_{\hat{\alpha},\hat{\beta}+e_n}\|_{\mathcal{D}_3,\mathcal{O}_+},\ \|\acute{P}_{\check{\alpha}+e_n,\check{\beta}} F_{\hat{\alpha},\hat{\beta}+e_n}\|_{\mathcal{D}_3,\mathcal{O}_+} &\leq \varepsilon e^{-\rho n^*_{\check{\alpha}+e_n,\check{\beta}}} \cdot \varepsilon^{\frac{4}{5}} e^{-\rho n^*_{\hat{\alpha},\hat{\beta}+e_n}} \\ &\leq \varepsilon^{\frac{9}{5}} e^{-\rho n^*_{\alpha\beta}}. \end{aligned} \tag{4.26}$$

- If $|\check{\alpha} + e_n| + |\check{\beta}| = 3$, then gauge invariance of $P$ implies $\acute{P}_{\check{\alpha}+e_n,\check{\beta}} = 0$. By (4.25), we can see that the only case, in which a higher-order term of $P$ is transformed into a lower-order term of $\{P, F\}$(indeed only $\{\breve{P}, F\}$), is $(\hat{\alpha}, \hat{\beta}) = (0, 0)$, $(\check{\alpha}, \check{\beta}) = (\alpha, \beta)$. By the definition of norm $\|X_F\|_{\mathcal{D}_3,\mathcal{O}}$ and the decay property of $P$,

$$\|\breve{P}_{\alpha+e_n,\beta}\|_{\mathcal{D}_3,\mathcal{O}} \leq e^{-\rho n^*_{\alpha+e_n,\beta}}, \quad \|F_{0,e_n}\|_{\mathcal{D}_3,\mathcal{O}_+} \leq cs\varepsilon^{\frac{4}{5}} e^{-\rho|n|}.$$

Thus, noting that $n^*_{\alpha\beta} \leq \max\{n^*_{\alpha+e_n,\beta},\, |n|\}$, we have

$$\|\breve{P}_{\alpha+e_n,\beta} F_{0,e_n}\|_{\mathcal{D}_3,\mathcal{O}_+} \leq cs\varepsilon^{\frac{4}{5}} e^{-\rho n^*_{\alpha\beta}} \leq c\varepsilon^{\frac{9}{5}} e^{-\rho n^*_{\alpha\beta}}. \tag{4.27}$$

ii) $|\alpha| + |\beta| \geq 3$

In this case, $|\check{\alpha} + e_n| + |\check{\beta}| \geq 3$. By the same argument as above, noting that $n^*_{\alpha\beta} \leq \max\{n^*_{\check{\alpha}+e_n,\check{\beta}},\, n^*_{\hat{\alpha},\hat{\beta}+e_n}\}$, or $n^*_{\alpha\beta} \leq n^+_{\check{\alpha}+e_n,\check{\beta}} - n^-_{\check{\alpha}+e_n,\check{\beta}} + n^*_{\hat{\alpha},\hat{\beta}+e_n}$,

$$\|\breve{P}_{\check{\alpha}+e_n,\check{\beta}} F_{\hat{\alpha},\hat{\beta}+e_n}\|_{\mathcal{D}_3,\mathcal{O}_+} \leq e^{-\rho n^*_{\check{\alpha}+e_n,\check{\beta}}} \cdot \varepsilon^{\frac{4}{5}} e^{-\rho n^*_{\hat{\alpha},\hat{\beta}+e_n}} \leq \varepsilon^{\frac{4}{5}} e^{-\rho n^*_{\alpha\beta}}, \tag{4.28}$$

$$\|\acute{P}_{\check{\alpha}+e_n,\check{\beta}} F_{\hat{\alpha},\hat{\beta}+e_n}\|_{\mathcal{D}_3,\mathcal{O}_+} \leq e^{-\rho(n^+_{\check{\alpha}+e_n,\check{\beta}} - n^-_{\check{\alpha}+e_n,\check{\beta}})} \cdot \varepsilon^{\frac{4}{5}} e^{-\rho n^*_{\hat{\alpha},\hat{\beta}+e_n}} \leq \varepsilon^{\frac{4}{5}} e^{-\rho n^*_{\alpha\beta}}. \tag{4.29}$$

Doing the same for $P_{\check{\alpha},\check{\beta}+e_n} F_{\hat{\alpha}+e_n,\hat{\beta}}$, we finish estimates for terms in (4.23).

(2) Terms in (4.24)

By Lemma 2.2 and the inequality $n^*_{\alpha\beta} \leq \max\{n^*_{\check{\alpha}\check{\beta}},\, n^*_{\hat{\alpha}\hat{\beta}}\}$, we have

$$\|\{P_{\check{\alpha}\check{\beta}}, F_{\hat{\alpha}\hat{\beta}}\}\|_{\mathcal{D}_{3\eta},\mathcal{O}_+} \leq c(r - r_+)^{-1}\eta^{-2} \begin{cases} \varepsilon^{\frac{9}{5}} e^{-\rho n^*_{\alpha\beta}}, & |\alpha| + |\beta| \leq 2 \\ \varepsilon^{\frac{4}{5}} e^{-\rho n^*_{\alpha\beta}}, & |\alpha| + |\beta| \geq 3 \end{cases}. \tag{4.30}$$

Combining (4.26) – (4.30), there exists $c_4 > 0$ such that

$$\|\{P, F\}_{\alpha\beta}\|_{\mathcal{D}_{3\eta},\mathcal{O}_+} \leq c_4(r - r_+)^{-1}\eta^{-2}K_+^2 \begin{cases} \varepsilon^{\frac{9}{5}} e^{-\rho n^*_{\alpha\beta}}, & |\alpha| + |\beta| \leq 2 \\ \varepsilon^{\frac{4}{5}} e^{-\rho n^*_{\alpha\beta}}, & |\alpha| + |\beta| \geq 3 \end{cases},$$

applying the fact that $|\hat{\alpha}| + |\hat{\beta}| \leq 2$. Moreover, if

**(C4)** $c_4(r-r_+)^{-1}K_+^2\varepsilon^{\frac{1}{20}} \le \frac{1}{4}$,

then Lemma 4.4 follows. ■

For $Y = P_+ - (P-R) = \sum_{\alpha,\beta} Y_{\alpha\beta}q^\alpha\bar{q}^\beta$, which is made up with iterated Poisson brackets, we can estimate them as above, and obtain

$$\|Y_{\alpha\beta}\|_{\mathcal{D}_+,\mathcal{O}_+} \le \begin{cases} \frac{1}{2}\varepsilon_+ e^{-\rho n^*_{\alpha\beta}}, & |\alpha|+|\beta|\le 2 \\ \varepsilon^{\frac{1}{5}} e^{-\rho n^*_{\alpha\beta}}, & |\alpha|+|\beta|\ge 3 \end{cases}.$$

for $\varepsilon$ sufficiently small. If we decompose $Y$ into $\breve{Y}$ and $\acute{Y}$, with

$$\breve{Y} = \sum_{\alpha,\beta} \breve{Y}_{\alpha\beta}(\theta, I;\xi)q^\alpha\bar{q}^\beta, \quad \acute{Y} = \sum_{\alpha,\beta} \acute{Y}_{\alpha\beta}(\xi)q^\alpha\bar{q}^\beta,$$

then

$$\|\breve{Y}_{\alpha\beta}\|_{\mathcal{D}_+,\mathcal{O}_+} \le \begin{cases} \frac{1}{2}\varepsilon_+ e^{-\rho_+ n^*_{\alpha\beta}}, & |\alpha|+|\beta|\le 2 \\ \varepsilon^{\frac{1}{5}} e^{-\rho_+ n^*_{\alpha\beta}}, & |\alpha|+|\beta|\ge 3 \end{cases},$$

$$\|\acute{Y}_{\alpha\beta}\|_{\mathcal{D}_+,\mathcal{O}_+} \le \begin{cases} \frac{1}{2}\varepsilon_+ e^{-\rho_+ n^*_{\alpha\beta}}, & |\alpha|+|\beta|\le 2 \\ \varepsilon^{\frac{1}{5}} e^{-\rho_+ (n^+_{\alpha\beta}-n^-_{\alpha\beta})}, & |\alpha|+|\beta|\ge 3 \end{cases}.$$

applying the basic facts $\frac{1}{2}(n^+_{\alpha\beta}-n^-_{\alpha\beta}) \le n^*_{\alpha\beta}$ and $\rho_+ < \frac{\rho}{2}$.

This completes one step of KAM iterations.

## 5 Proof of the KAM theorem

Let $r_0$, $s_0$, $\rho_0$, $\varepsilon_0$, $\gamma_0$, $K_0$, $\mathcal{O}_0$, $H_0$, $N_0$, $P_0$ be as given in Subsection 4.1. For $\nu = 1, 2, \cdots$, define the following sequences:

$$\varepsilon_\nu = \varepsilon_{\nu-1}^{\frac{5}{4}} = \varepsilon_0^{\left(\frac{5}{4}\right)^\nu}, \quad \eta_\nu = \varepsilon_\nu^{\frac{1}{4}}, \quad \gamma_\nu = \varepsilon_\nu^{\frac{1}{16}}, \quad K_\nu = 2|\ln\varepsilon_{\nu-1}|K_{\nu-1}, \quad \rho_\nu = K_\nu^{-1},$$

$$r_\nu = r_0\left(1-\sum_{i=2}^{\nu+1}2^{-i}\right), \quad s_\nu = \frac{1}{8}\eta_{\nu-1}s_{\nu-1} = 2^{-3\nu}\left(\prod_{i=0}^{\nu-1}\varepsilon_i\right)^{\frac{1}{4}} s_0.$$

Consider $H_\nu = N_\nu + P_\nu$ on $\mathcal{D}_\nu = \mathcal{D}_{d,\rho_\nu}(r_\nu, s_\nu)$, with

$$\begin{aligned} N_\nu &= e_\nu(\xi) + \langle\omega_\nu(\xi), I\rangle + \langle A_\nu(\xi)z_\nu, \bar{z}_\nu\rangle + \sum_{|n|>K_\nu}\Omega_n(\xi)q_n\bar{q}_n \\ &= e_\nu(\xi) + \langle\omega_\nu(\xi), I\rangle + \langle\tilde{A}_\nu(\xi)z_{\nu+1}, \bar{z}_{\nu+1}\rangle + \sum_{|n|>K_{\nu+1}}\Omega_n(\xi)q_n\bar{q}_n, \\ P_\nu &= \breve{P}_\nu + \acute{P}_\nu = \sum_{\alpha,\beta}\breve{P}^\nu_{\alpha\beta}(\theta, I;\xi)q^\alpha\bar{q}^\beta + \sum_{\alpha,\beta}\acute{P}^\nu_{\alpha\beta}(\xi)q^\alpha\bar{q}^\beta \end{aligned}$$

where $z_\nu = (q_n)_{|n|\le K_\nu}$, $\bar{z}_\nu = (\bar{q}_n)_{|n|\le K_\nu}$, and

$$\tilde{A}_\nu = \begin{pmatrix} A_\nu & 0 \\ 0 & \Omega_n \end{pmatrix}_{K_\nu<|n|\le K_{\nu+1}}$$

whose eigenvalues are $\{\mu_j^\nu\}_{|j|\le K_{\nu+1}}$, with $\{\mu_j^\nu\}_{|j|\le K_\nu}$ being eigenvalues of $A_\nu$ and $\mu_j^\nu = \Omega_j$ for $K_\nu < |j| \le K_{\nu+1}$. Let

$$\mathcal{O}_{\nu+1} = \left\{ \xi \in \mathcal{O}_\nu : \begin{array}{l} |\langle k, \omega_\nu\rangle| > \frac{\gamma_\nu}{|k|^\tau} \\ |\langle k, \omega_\nu\rangle + \mu_n^\nu| > \frac{\gamma_\nu}{|k|^\tau K_{\nu+1}^2}, \\ |\langle k, \omega_\nu\rangle + \mu_n^\nu + \mu_m^\nu| \le \frac{\gamma_\nu}{|k|^\tau K_{\nu+1}^4}, \\ |\langle k, \omega_\nu\rangle + \mu_n^\nu - \mu_m^\nu| \le \frac{\gamma_\nu}{|k|^\tau K_{\nu+1}^4}, \end{array} \quad k \neq 0, \quad |n|, |m| \le K_{\nu+1} \right\}.$$

## 5.1 Iteration Lemma

The preceding analysis may be summarized in the following

**Lemma 5.1** *There exists $\varepsilon_0$ sufficiently small such that the following holds for all $\nu = 0, 1, \cdots$.*

(a) *$H_\nu = N_\nu + P_\nu$ is real-analytic on $\mathcal{D}_\nu$, $C_W^1$ parametrized by $\xi \in \mathcal{O}_\nu$, and*

$$|\omega_{\nu+1} - \omega_\nu|_{\mathcal{O}_{\nu+1}}, \quad |(A_{\nu+1} - \tilde{A}_\nu)_{nm}|_{\mathcal{O}_{\nu+1}} \le \varepsilon_\nu e^{-\rho_\nu \max\{|n|,|m|\}}.$$

*Moreover, $P_\nu$ has gauge invariance, and $\|X_{P_\nu}\|_{\mathcal{D}_\nu,\mathcal{O}_\nu} \le \varepsilon_\nu$,*

$$\|\breve{P}_{\alpha\beta}^\nu\|_{\mathcal{D}_\nu,\mathcal{O}_\nu} \le \left\{ \begin{array}{ll} \varepsilon_\nu e^{-\rho_\nu n_{\alpha\beta}^*}, & |\alpha| + |\beta| \le 2 \\ e^{-\rho_\nu n_{\alpha\beta}^*}, & |\alpha| + |\beta| \ge 3 \end{array} \right.,$$

$$\|\acute{P}_{\alpha\beta}^\nu\|_{\mathcal{D}_\nu,\mathcal{O}_\nu} \le \left\{ \begin{array}{ll} \varepsilon_\nu e^{-\rho_\nu n_{\alpha\beta}^*}, & |\alpha| + |\beta| \le 2 \\ e^{-\rho_\nu (n_{\alpha\beta}^+ - n_{\alpha\beta}^-)}, & |\alpha| + |\beta| \ge 3 \end{array} \right. .$$

(b) *There is a symplectic transformation $\Phi_\nu : \mathcal{D}_{\nu+1} \to \mathcal{D}_\nu$ with*

$$\|D\Phi_\nu - Id\|_{\mathcal{D}_{\nu+1},\mathcal{O}_{\nu+1}} \le \varepsilon_\nu^{\frac{4}{5}}$$

*such that $H_{\nu+1} = H_\nu \circ \Phi_\nu$.*

*Proof:* Let $c_0 = e^{10} \max\{c_1, c_2, c_3, c_4\}$. We need to verify the assumptions $(\mathbf{C1}) - (\mathbf{C4})$ for all $\nu = 0, 1, \cdots$. Noting that $r_\nu - r_{\nu+1} = \frac{r_0}{2^{\nu+2}}$ and $\rho_\nu K_\nu = 1$, it is sufficient for us to check:

**(D1)** $c_0 s_\nu \le \varepsilon_\nu$,

**(D2)** $c_0 r_0^{-(2\tau+b+1)} 2^{(\nu+2)(2\tau+b+1)} K_{\nu+1}^{d+20} \le \varepsilon_\nu^{-\frac{1}{30}}$,

for all $\nu = 0, 1, \cdots$.

By the choice of $s_0$, the condition $(\mathbf{D1})$ clearly holds for $\nu = 0$. Suppose that it holds for some $\nu$. Then it is easy to see that

$$c_0 s_{\nu+1} = 2^{-3} \varepsilon_\nu^{\frac{1}{4}} \cdot c_0 s_\nu < 2^{-3} \varepsilon_\nu^{\frac{1}{4}} \cdot \varepsilon_\nu < \varepsilon_{\nu+1}.$$

Hence $(\mathbf{D1})$ holds for all $\nu$.

As for **(D2)**, let us take $\varepsilon_0$ sufficiently small such that

$$c_0 r_0^{-(2\tau+b+1)} 2^{(2\tau+b+1)} (2K_0|\ln\varepsilon_0|)^{d+20} \leq \varepsilon_0^{-\frac{1}{30}},$$

then **(D2)** holds for $\nu = 0$. Since for $\nu = 0, 1, \cdots$,

$$K_{\nu+1} = 2K_\nu|\ln\varepsilon_\nu| = 2^{\nu+1}K_0 \prod_{i=0}^{\nu} |\ln\varepsilon_i| = K_0(2|\ln\varepsilon_0|)^{\nu+1} \left(\frac{5}{4}\right)^{\frac{(\nu+1)\nu}{2}},$$

while $\varepsilon_\nu^{-\frac{1}{30}} = \left(\varepsilon_0^{-\frac{1}{30}}\right)^{\left(\frac{5}{4}\right)^\nu}$. This means that the right side of **(D2)** grows with $\nu$ much faster than the left side. Thus, (**D2**) holds true. ■

## 5.2 Convergence

Define $\Psi^\nu = \Phi_* \circ \Phi_0 \circ \Phi_1 \circ \cdots \circ \Phi_{\nu-1}$, $\nu = 1, 2, \cdots$. An induction argument shows that $\Psi^\nu : \mathcal{D}_{\nu+1} \to \mathcal{D}_0$ and

$$H_0 \circ \Psi^\nu = H_\nu = N_\nu + P_\nu, \quad \nu = 1, 2, \cdots.$$

Let $\mathcal{O}_\varepsilon = \cap_{\nu=0}^\infty \mathcal{O}_\nu$. Using Lemma 4.3 and standard arguments (e.g., [30, 36]), it concludes that $H_\nu$, $N_\nu$, $P_\nu$ and $\Psi^\nu$ converge uniformly on $\mathcal{D}_{d,0}(\frac{1}{2}r_0, 0) \times \mathcal{O}_\varepsilon$ to, say, $H_\infty$, $N_\infty$, $P_\infty$ and $\Psi^\infty$, respectively, in which case it is clear that

$$N_\infty = e_\infty + \langle \omega_\infty, I\rangle + \langle A_\infty z_\infty, \bar{z}_\infty\rangle.$$

Since $\varepsilon_\nu = \varepsilon_0^{(\frac{5}{4})^\nu}$, we have, by Lemma 5.1, that $X_{P_\infty}|_{\mathcal{D}_{d,0}(\frac{1}{2}r_0,0)\times\mathcal{O}_\varepsilon} = 0$.

Since $H_0 \circ \Psi^\nu = H_\nu$, we have $\Phi_{H_0}^t \circ \Psi^\nu = \Psi^\nu \circ \Phi_{H_\nu}^t$, with $\Phi_{H_0}^t$ denoting the flow of the Hamiltonian vector field $X_{H_0}$. The uniform convergence of $\Psi^\nu$ and $X_{H_\nu}$ implies that one can pass the limit in the above and conclude that

$$\Phi_{H_0}^t \circ \Psi^\infty = \Psi^\infty \circ \Phi_{H_\infty}^t, \quad \Psi^\infty : \mathcal{D}_{d,0}(\frac{1}{2}r_0, 0) \to \mathcal{D}.$$

Hence,

$$\Phi_{H_0}^t(\Psi^\infty(\mathbb{T}^b \times \{\xi\})) = \Psi^\infty \Phi_{N_\infty}^t(\mathbb{T}^b \times \{\xi\}) = \Psi^\infty(\mathbb{T}^b \times \{\xi\}), \quad \forall \xi \in \mathcal{O}_\varepsilon.$$

This means that $\Psi^\infty(\mathbb{T}^b \times \{\xi\})$ is an embedded invariant torus of the original perturbed Hamiltonian system at $\xi \in \mathcal{O}_\varepsilon$. Moreover, the frequencies $\omega_\infty(\xi)$ associated with $\Psi^\infty(\mathbb{T}^b \times \{\xi\})$ are slightly deformed from the unperturbed ones, $\omega(\xi)$.

## 5.3 Measure estimates

At the $\nu^{\text{th}}$ step of KAM iteration, we need to exclude the following resonant parameter set

$$\mathcal{R}_k^\nu := \mathcal{R}_k^{\nu 1} \bigcup \left( \bigcup_{|n|\leq K_{\nu+1}} \mathcal{R}_{kn}^{\nu 2} \right) \bigcup \left( \bigcup_{|n|,|m|\leq K_{\nu+1}} \mathcal{R}_{knm}^{\nu 3} \right) \bigcup \left( \bigcup_{|n|,|m|\leq K_{\nu+1}} \mathcal{R}_{knm}^{\nu 4} \right),$$

for all $k \neq 0$, where

$$\begin{aligned}
\mathcal{R}_k^{\nu 1} &:= \left\{\xi \in \mathcal{O}_\nu : |\langle k, \omega_\nu\rangle| < \frac{\gamma_\nu}{|k|^\tau}\right\}, \\
\mathcal{R}_{kn}^{\nu 2} &:= \left\{\xi \in \mathcal{O}_\nu : |\langle k, \omega_\nu\rangle + \mu_n^\nu| < \frac{\gamma_\nu}{|k|^\tau K_{\nu+1}^2}\right\}, \\
\mathcal{R}_{knm}^{\nu 3} &:= \left\{\xi \in \mathcal{O}_\nu : |\langle k, \omega_\nu\rangle + \mu_n^\nu + \mu_m^\nu| < \frac{\gamma_\nu}{|k|^\tau K_{\nu+1}^4}\right\}, \\
\mathcal{R}_{knm}^{\nu 4} &:= \left\{\xi \in \mathcal{O}_\nu : |\langle k, \omega_\nu\rangle + \mu_n^\nu - \mu_m^\nu| < \frac{\gamma_\nu}{|k|^\tau K_{\nu+1}^4}\right\}.
\end{aligned}$$

It is clear that $\mathcal{O}_0 \setminus \mathcal{O}_\varepsilon \subseteq \bigcup_{\nu\geq 0}\bigcup_{k\neq 0}\mathcal{R}_k^\nu$.

As eigenvalues of the Hermitian matrix $\tilde{A}_\nu$, it is well-known that $\{\mu_n^\nu\}_{|n|\leq K_{\nu+1}}$ $C_W^1$ depend on $\xi$ and there exist orthonormal eigenvectors $\psi_n^\nu$ corresponding to $\mu_n^\nu$, $C_W^1$ depending on $\xi$ (see e.g. [13]). It follows that $\mu_n^\nu = \langle \tilde{A}_\nu \psi_n^\nu, \bar{\psi}_n^\nu\rangle$ and

$$\partial_{\xi_j}\mu_n^\nu = \langle (\partial_{\xi_j}\tilde{A}_\nu)\psi_n^\nu, \bar{\psi}_n^\nu\rangle, \quad j = 1, \cdots, b.$$

Recalling that $\omega_0$ is a diffeomorphism of $\xi$, and $\sup_{\xi\in\mathcal{O}}|\partial_\xi \Omega_n| \ll 1$, together with the estimates in (4.22), we have

$$|\partial_\xi(\langle k, \omega_\nu\rangle + \mu_m^\nu - \mu_n^\nu)| \geq |\partial_\xi(\langle k, \omega_0\rangle + \Omega_n - \Omega_m)| - \varepsilon_0^{\frac{1}{2}}|k| - \varepsilon_0^{\frac{1}{2}} = O(|k|)$$

for the set $\mathcal{R}_{knm}^{\nu 4}$. The cases for $\mathcal{R}_k^{\nu 1}$, $\mathcal{R}_{kn}^{\nu 2}$, $\mathcal{R}_{knm}^{\nu 3}$ can be handled in an entirely analogous way. Thus for fixed $k \neq 0$,

$$\left|\mathcal{R}_k^{\nu 1} \bigcup \left(\bigcup_{|n|\leq K_{\nu+1}} \mathcal{R}_{kn}^{\nu 2}\right) \bigcup \left(\bigcup_{|n|,|m|\leq K_{\nu+1}} \mathcal{R}_{knm}^{\nu 3}\right) \bigcup \left(\bigcup_{|n|,|m|\leq K_{\nu+1}} \mathcal{R}_{knm}^{\nu 4}\right)\right| \leq \frac{c\gamma_\nu}{|k|^{\tau+1}}.$$

Since $\tau \geq b$, we have that

$$|\mathcal{O}_0 \setminus \mathcal{O}_\varepsilon| \leq \left|\bigcup_{\nu\geq 0}\bigcup_{k\neq 0}\mathcal{R}_k^\nu\right| \leq c\sum_{\nu\geq 0}\sum_{k\neq 0}\frac{\gamma_\nu}{|k|^{\tau+1}} = c\sum_{\nu\geq 0}\gamma_\nu \sim \gamma_0 = \varepsilon_0^{\frac{1}{16}}.$$

# A Appendix

## A.1 The original form of Theorem 3

Given $R > 0$, $\mathcal{H}_R$ denotes the set of period-one holomorphic bounded functions $f$ on

$$\mathcal{S}_R = \{z \in \mathbb{C} : |\mathrm{Im} z| < R\},$$

equipped with the sup-norm

$$\|f\|_R = \sup_{z\in\mathcal{S}_R}|f(z)|.$$

$\mathcal{P}_R$ denotes the set of period-one meromorphic functions $f$ on $\mathcal{S}_R$ such that there is a constant $c > 0$ with

$$|f(z) - f(z-a)| \geq c|a|_1, \quad \forall a \in \mathbb{R}, \quad \forall z \in \mathcal{S}_R, \tag{A.1}$$

where $|\cdot|_1$ is defined as in (1.2). Then $|f|_R$ is defined as the biggest possible value of $c$ in (A.1). It is obvious the function $V(x) = \tan \pi x$ belongs to $\mathcal{P}_R$ for any $R > 0$, with $|V|_R \geq 1$.

For $\sigma > 0$, $R > 0$ and $\tilde{\alpha} \in \mathbb{R}^d$ satisfying the Diophantine condition, i.e., there exist $\tilde{\gamma} > 0$, $\tilde{\tau} > d$ such that

$$|\langle n, \tilde{\alpha}\rangle|_1 > \frac{\tilde{\gamma}}{|n|^{\tilde{\tau}}}, \quad \forall n \in \mathbb{Z}^d \setminus \{0\},$$

let $\mathcal{U}^{\tilde{\alpha}}_{R,\sigma}$ denote the Banach $*$-algebra of kernels $m = \{m(z,n)\}_{n\in\mathbb{Z}^d, z\in\mathcal{S}_R}$, where for each $n \in \mathbb{Z}^d$, the map $z \mapsto m(z,n)$ belongs to $\mathcal{H}_R$(or $\mathcal{P}_R$), and

$$\|m\|'_{R,\sigma} := \sup_{z\in\mathcal{S}_R} \sum_{n\in\mathbb{Z}^d} |m(z,n)| e^{\sigma|n|}$$

is finite. (We need to exclude a subset of $\mathcal{S}_R$ with measure zero in the case that $m(\cdot, n) \in \mathcal{P}_R$ and there is some poles in $\mathcal{S}_R$.) The $*$-algebraic structure is defined by

$$\begin{aligned}(m_1 \cdot m_2)(z,n) &:= \sum_{l\in\mathbb{Z}^d} m_1(z,l) m_2(z - \langle l, \tilde{\alpha}\rangle, n-l), \\ m^*(z,n) &:= \overline{m(\bar{z} - \langle n, \tilde{\alpha}\rangle, -n)}.\end{aligned}$$

Then the norm is defined by

$$\|m\|_{R,\sigma} = \max\{\|m\|'_{R,\sigma}, \|m^*\|'_{R,\sigma}\}.$$

For example, if $g \in \mathcal{H}_R$(or $g \in \mathcal{P}_R$) then $g$ can be considered as an element of $\mathcal{U}^{\tilde{\alpha}}_{R,\sigma}$, by putting:

$$g(z,n) := g(z)\delta_{n,0}.$$

Such a kernel is called *diagonal.* If $e \in \mathbb{Z}^d$, $u_e$ is the kernel

$$u_e(z,n) := \delta_{n,e}.$$

One can easily see that $u_0$ is an identity and

$$u_e^* u_e = u_e u_e^* = u_0, \quad \forall e \in \mathbb{Z}^d.$$

The Laplace kernel is then given by

$$\Delta = \sum_{e=\pm 1} u_e.$$

A canonical set of representations of $\mathcal{U}^{\tilde{\alpha}}_{R,\sigma}$ in $\ell^2(\mathbb{Z}^d)$ is given by

$$[\Pi_z(m)\psi](n) = \sum_{l\in\mathbb{Z}^d} m(z - \langle n, \tilde{\alpha}\rangle, l-n)\psi(l),$$

where $\psi \in \ell^2(\mathbb{Z}^d)$, $z \in \mathcal{S}_R$ and $m \in \mathcal{U}^{\tilde{\alpha}}_{R,\sigma}$. Actually, $\Pi_z(m)$ can be seen as an infinite matrix, with its matrix elements $[\Pi_z(m)]_{ln} = m(z - \langle n, \tilde{\alpha}\rangle, l-n)$.

**Theorem 4 (Theorem 1 of [2])** *Given $R > 0$, $r > 0$, and $\tilde{\alpha} \in \mathbb{R}^d$ satisfying the Diophantine condition, i.e., for all $n \in \mathbb{Z}^d \setminus \{0\}$*

$$|\langle n, \tilde{\alpha}\rangle|_1 \geq \frac{\tilde{\gamma}}{|n|^{\tilde{\tau}}},$$

*for some $\tilde{\gamma} > 0$ and $\tilde{\tau} > d$. If $V \in \mathcal{P}_R$, there is a positive constant $\varepsilon_c$, depending on $R$, $\sigma$, $\tilde{\gamma}$, $\tilde{\tau}$ and $|V|_R$ only such that if $m \in \mathcal{U}^{\tilde{\alpha}}_{R,\sigma}$, $\|m\|_{R,\sigma} < \varepsilon_c$, there exists an invertible element $u \in \mathcal{U}^{\tilde{\alpha}}_{R,\sigma}$ and $\hat{V} \in \mathcal{P}_{R/2}$ with*

$$u(V+m)u^{-1} = \hat{V}, \tag{A.2}$$

$$\max\{\|u - Id\|_{R/2,\,\sigma/2}, \|u^{-1} - Id\|_{R/2,\,\sigma/2}\} \leq c\|m\|_{R,\sigma}, \tag{A.3}$$

$$V - \hat{V} \in \mathcal{H}_{R/2}, \quad \|V - \hat{V}\|_{R/2} \leq \|m\|_{R,\sigma}, \tag{A.4}$$

$$|\hat{V}|_{R/2} \geq \frac{1}{2}|V|_R. \tag{A.5}$$

*If in addition $m + V$ is self-adjoint, then $u$ is unitary and $\hat{V} = \hat{V}^*$.*

**Corollary 1 (Corollary 1 of [2])** *Let $m$ and $V$ be as in the previous theorem. Then the operator $H_z = \prod_z(m+V)$ has a complete set of eigenvectors which are exponentially localized. The corresponding eigenvalues are the set*

$$\{\hat{V}(z - \langle n, \tilde{\alpha}\rangle) : z - \langle n, \tilde{\alpha}\rangle \text{ is not a pole of } V, \ \ n \in \mathbb{Z}^d\}.$$

Now, for $d = 1$, $\sigma = 4$ and arbitrary $R > 0$, consider the Schrödinger operator on $\ell^2(\mathbb{Z})$

$$(L_x q)_n = (\epsilon\Delta q)_n + \tan\pi(n\tilde{\alpha} + x)q_n = \epsilon(q_{n-1} + q_{n+1}) + \tan\pi(n\tilde{\alpha} + x)q_n, \quad x \in \mathcal{X}.$$

In the set-up above, it can be expressed as $\Pi_x(\epsilon\Delta + V)$. Obviously, $\|\epsilon\Delta\|_{R,\sigma} < c\epsilon$. Theorem 4 implies that if $\epsilon$ is sufficiently small, then for every $x \in \mathcal{X} \subset \mathcal{S}_R$, there is an orthogonal transformation $U_x = \Pi_x(u)$ on $\ell^2(\mathbb{Z})$ such that

$$U_x^* L_x U_x = \mathrm{diag}\{\hat{V}(x + n\tilde{\alpha})\}_{n\in\mathbb{Z}},$$

where $u \in \mathcal{U}^{\tilde{\alpha}}_{R,\sigma}$, $\hat{V} \in \mathcal{P}_{R/2}$ with $g(z) := \hat{V}(z) - \tan\pi z$ contained in $\mathcal{H}_{R/2}$ and $\|g\|_{R/2} < c\epsilon$. By Corollary 1, $\{\hat{V}(x + n\tilde{\alpha})\}_{n\in\mathbb{Z}} \subset \mathbb{R}$ is exactly the set of the eigenvalues of the operator $L_x$. By (A.3) in Theorem 4, the infinite matrix $U_x$ has off-diagonal decay, i.e., the matrix elements $(U_x - I_{\mathbb{Z}})_{mn}$ satisfy

$$|(U_x - I_{\mathbb{Z}})_{mn}| = |u(x - n\tilde{\alpha}, m - n) - \delta_{mn}| \leq c\epsilon e^{-2|m-n|}.$$

Setting several constants $c = 1$ for convenience, we obtain the content of Theorem 3.

## A.2 Proof of Lemma 3.1

For $|i|, |j|, |n|, |m| \leq \kappa |\ln \epsilon|$, we consider the function

$$V^0_{i,j,n,m}(x) := \tan \pi(x + i\tilde{\alpha}) - \tan \pi(x + j\tilde{\alpha}) + \tan \pi(x + n\tilde{\alpha}) - \tan \pi(x + m\tilde{\alpha})$$

on $\mathbb{R}/\mathbb{Z}$. To get the lower bound in (3.19), it is sufficient to show that

$$|V^0_{i,j,n,m}(x)| \geq 2\epsilon^{\frac{1}{4}}$$

on some subset of $\mathbb{R}/\mathbb{Z}$, since $\sup_{x\in\mathbb{R}/\mathbb{Z}} |\hat{V}(x) - \tan \pi x| \leq \epsilon$.

It is necessary to restrict the functions on the subset $\mathcal{X}_0 = \mathcal{X}_0' \cap \mathcal{X}_0'' \subset \mathbb{R}/\mathbb{Z}$, with the necessity clear somewhat later, where

$$\mathcal{X}_0' := \{x \in \mathbb{R}/\mathbb{Z} : \left| x + n\tilde{\alpha} - \frac{1}{2} \right| \geq \epsilon^{\frac{1}{1200}}, \quad \forall |n| \leq \kappa |\ln \epsilon|\},$$
$$\mathcal{X}_0'' := \{x \in \mathbb{R}/\mathbb{Z} : |\tan \pi(x + n\tilde{\alpha})| \geq \epsilon^{\frac{1}{1200}}, \quad \forall |n| \leq \kappa |\ln \epsilon|\}.$$

Hence on $\mathcal{X}_0$, for $|n| \leq \kappa |\ln \epsilon|$,

$$\epsilon^{\frac{1}{1200}} \leq |\tan \pi(x + n\tilde{\alpha})| \leq \left| \tan \pi \left( \frac{1}{2} - \epsilon^{\frac{1}{1200}} \right) \right| = \left| \tan \epsilon^{\frac{1}{1200}} \pi \right|^{-1} \leq c\epsilon^{-\frac{1}{1200}}, \qquad \text{(A.6)}$$

if $\epsilon$ is sufficiently small. Then $V^0_{i,j,n,m}(x)$ are all bounded piecewise smooth functions on $\mathcal{X}_0$. It is easy to see that there is at most $c\kappa|\ln \epsilon|$ many connected components contained in $\mathcal{X}_0$ and

$$\text{mes}(\mathbb{R}/\mathbb{Z} \setminus (\mathcal{X}_0' \cap \mathcal{X}_0'')) \leq c\kappa |\ln \epsilon| \cdot \epsilon^{\frac{1}{1200}} < \epsilon^{\frac{1}{1400}}$$

for $\epsilon$ sufficiently small.

It is clear $\{i, n\} = \{j, m\}$ implies that $V^0_{i,j,n,m} \equiv 0$, so we assume that $\{i, n\} \neq \{j, m\}$. If, in addition, $\{i, n\} \cap \{j, m\} \neq \emptyset$, then the intersection has a single element. Assume that $i = j$ without loss of generality, then $n \neq m$ and

$$V^0_{i,j,n,m}(x) = \tan \pi(x + n\tilde{\alpha}) - \tan \pi(x + m\tilde{\alpha}). \qquad \text{(A.7)}$$

Thus, we have

$$|V^0_{i,j,n,m}(x)| \geq \pi |(n - m)\tilde{\alpha}|_1 \geq \frac{\pi\tilde{\gamma}}{(2\kappa)^{\tilde{\tau}} |\ln \epsilon|^{\tilde{\tau}}} \geq \epsilon^{\frac{1}{1200}}. \qquad \text{(A.8)}$$

The case $\{i, n\} \cap \{j, m\} = \emptyset$ is much more complex, which can be decomposed into the following four subcases:

**(S1)** $\{i, n\} \cap \{j, m\} = \emptyset$ with $i \neq n$ and $j \neq m$;

**(S2)** $\{i, n\} \cap \{j, m\} = \emptyset$ with $i = n$ and $j \neq m$;

**(S3)** $\{i, n\} \cap \{j, m\} = \emptyset$ with $i \neq n$ and $j = m$;

**(S4)** $\{i, n\} \cap \{j, m\} = \emptyset$ with $i = n$ and $j = m$.

We only need to consider the subcases (**S1**) $-$ (**S3**), since in the subcase **(S4)**,

$$V^0_{i,j,n,m}(x) = 2(\tan\pi(x+n\tilde{\alpha}) - \tan\pi(x+m\tilde{\alpha})),$$

which is the same as in (A.7). Corresponding to (**S1**) $-$ (**S3**), let

$$B_1(x) := \begin{pmatrix} \tan\pi(x+i\tilde{\alpha}) & \tan\pi(x+j\tilde{\alpha}) & \tan\pi(x+n\tilde{\alpha}) & \tan\pi(x+m\tilde{\alpha}) \\ \tan^2\pi(x+i\tilde{\alpha}) & \tan^2\pi(x+j\tilde{\alpha}) & \tan^2\pi(x+n\tilde{\alpha}) & \tan^2\pi(x+m\tilde{\alpha}) \\ \tan^3\pi(x+i\tilde{\alpha}) & \tan^3\pi(x+j\tilde{\alpha}) & \tan^3\pi(x+n\tilde{\alpha}) & \tan^3\pi(x+m\tilde{\alpha}) \\ \tan^4\pi(x+i\tilde{\alpha}) & \tan^4\pi(x+j\tilde{\alpha}) & \tan^4\pi(x+n\tilde{\alpha}) & \tan^4\pi(x+m\tilde{\alpha}) \end{pmatrix},$$

and

$$B_2(x) := \begin{pmatrix} \tan\pi(x+i\tilde{\alpha}) & \tan\pi(x+j\tilde{\alpha}) & \tan\pi(x+m\tilde{\alpha}) \\ \tan^2\pi(x+i\tilde{\alpha}) & \tan^2\pi(x+j\tilde{\alpha}) & \tan^2\pi(x+m\tilde{\alpha}) \\ \tan^3\pi(x+i\tilde{\alpha}) & \tan^3\pi(x+j\tilde{\alpha}) & \tan^3\pi(x+m\tilde{\alpha}) \end{pmatrix},$$

$$B_3(x) := \begin{pmatrix} \tan\pi(x+i\tilde{\alpha}) & \tan\pi(x+n\tilde{\alpha}) & \tan\pi(x+m\tilde{\alpha}) \\ \tan^2\pi(x+i\tilde{\alpha}) & \tan^2\pi(x+n\tilde{\alpha}) & \tan^2\pi(x+m\tilde{\alpha}) \\ \tan^3\pi(x+i\tilde{\alpha}) & \tan^3\pi(x+n\tilde{\alpha}) & \tan^3\pi(x+m\tilde{\alpha}) \end{pmatrix}.$$

**Lemma A.1** *Given $|i|, |j|, |n|, |m| \le \kappa|\ln\epsilon|$. If $\epsilon$ is sufficiently small, then for any $x \in \mathcal{X}_0$, we have*

- *when* **(S1)** *holds,* $|\det(B_1(x))| \ge \epsilon^{\frac{1}{120}}$*;*
- *when* **(S2)** *holds,* $|\det(B_2(x))| \ge \epsilon^{\frac{1}{200}}$*;*
- *when* **(S3)** *holds,* $|\det(B_3(x))| \ge \epsilon^{\frac{1}{200}}$*.*

*Proof:* The determinant of $B_1(x)$ can be written as

$$\tan\pi(x+i\tilde{\alpha})\cdot\tan\pi(x+j\tilde{\alpha})\cdot\tan\pi(x+n\tilde{\alpha})\cdot\tan\pi(x+m\tilde{\alpha})\cdot\det(\tilde{B}_1(x)),$$

with $\tilde{B}_1(x)$ the Vandermonde matrix

$$\begin{pmatrix} 1 & 1 & 1 & 1 \\ \tan\pi(x+i\tilde{\alpha}) & \tan\pi(x+j\tilde{\alpha}) & \tan\pi(x+n\tilde{\alpha}) & \tan\pi(x+m\tilde{\alpha}) \\ \tan^2\pi(x+i\tilde{\alpha}) & \tan^2\pi(x+j\tilde{\alpha}) & \tan^2\pi(x+n\tilde{\alpha}) & \tan^2\pi(x+m\tilde{\alpha}) \\ \tan^3\pi(x+i\tilde{\alpha}) & \tan^3\pi(x+j\tilde{\alpha}) & \tan^3\pi(x+n\tilde{\alpha}) & \tan^3\pi(x+m\tilde{\alpha}) \end{pmatrix}.$$

Then, when **(S1)** holds, we can obtain that $|\det(B_1(x))| \ge \epsilon^{\frac{1}{120}}$, by (A.6) and (A.8), combining with

$$\det\tilde{B}_1(x) = \prod_{\substack{n_1,n_2\in\{i,j,n,m\}\\ n_1<n_2}} (\tan\pi(x+n_1\tilde{\alpha}) - \tan\pi(x+n_2\tilde{\alpha})).$$

As for the subcases **(S2)** and **(S3)**, there is no doubt that $|\det(B_2(x))|, |\det(B_3(x))| \ge \epsilon^{\frac{1}{200}}$, which can be proved in the same way as above. ■

For $s \in \{0,1,2,3\}$, let

$$\tilde{u}^{(s)}(x) = \Big(V^{(s)}(x+i\tilde{\alpha}),\ \ V^{(s)}(x+j\tilde{\alpha}),\ \ V^{(s)}(x+n\tilde{\alpha}),\ \ V^{(s)}(x+m\tilde{\alpha})\Big)^{\top} \in \mathbb{R}^4,$$

where $V(x) := \tan \pi x$, $V^{(s)}$ is its $s^{\text{th}}-$order derivative and $V^{(0)}$ means the function $V$ itself in particular. We can calculate that

$$\begin{array}{rcl} V^{(1)}(x) & = & \pi + \pi \tan^2 \pi x, \\ V^{(2)}(x) & = & 2\pi^2 \tan \pi x + 2\pi^2 \tan^3 \pi x, \\ V^{(3)}(x) & = & 2\pi^3 + 8\pi^3 \tan^2 \pi x + 6\pi^3 \tan^4 \pi x. \end{array}$$

Moreover, if $\epsilon$ is sufficiently small, then for $x \in \mathcal{X}_0$, we have that

$$|V^{(0)}(x)| \le c\epsilon^{-\frac{1}{1200}}, \quad \left|V^{(1)}(x)\right| \le c\epsilon^{-\frac{1}{600}}, \quad \left|V^{(2)}(x)\right| \le c\epsilon^{-\frac{1}{400}}, \quad \left|V^{(3)}(x)\right| \le c\epsilon^{-\frac{1}{300}}.$$

Indeed, it can be checked that for $s = 0,1,2,\cdots$,

$$\left|V^{(s)}(x)\right| \le c\epsilon^{-\frac{s+1}{1200}}, \tag{A.9}$$

where $c = c(s)$ grows exponentially in $s$. Let

$$\begin{array}{ll} u^{(0)}(x) = \tilde{u}^{(0)}(x), & u^{(1)}(x) = \tilde{u}^{(1)}(x) - \pi(1,1,1,1)^{\top}, \\ u^{(2)}(x) = \tilde{u}^{(2)}(x), & u^{(3)}(x) = \tilde{u}^{(3)}(x) - 2\pi^3(1,1,1,1)^{\top}. \end{array}$$

Thus the determinant of the $4 \times 4$ matrix $(u^{(0)}(x),\, u^{(1)}(x),\, u^{(2)}(x),\, u^{(3)}(x))$ equals to $c \cdot \det(B_1(x))$, where $B_1(x)$ is defined as in Lemma A.1.

We need to arrive at some transversality conditions, which are elaborated in Corollary 2, by virtue of the following lemma .

**Lemma A.2 (Proposition of appendix B in [3])** *Let $u^{(0)},\cdots,u^{(L-1)}$ be $L$ independent vectors in $\mathbb{R}^L$ with $\|u^{(s)}\|_{\ell^1} \le 1$. Let $v \in \mathbb{R}^L$ be an arbitrary vector, then there exists $s \in \{0,\cdots,L-1\}$, such that*

$$|\langle v, u^{(s)}\rangle| \ge L^{-\frac{3}{2}} \|v\|_{\ell^1} \det U,$$

*where $\det U$ is the determinant of the matrix formed by the components of the vectors $u^{(s)}$, and $\langle \cdot,\cdot \rangle$ is the usual scalar product.*

For the proof see [3].

**Corollary 2** *Given $|i|, |j|, |n|, |m| \le \kappa |\ln \epsilon|$, and $\{i,n\} \cap \{j,m\} = \emptyset$. If $\epsilon$ is sufficiently small, then for any $x \in \mathcal{X}_0$, we have*

- *when* **(S1)** *holds, there exists $s \in \{0,1,2,3\}$ such that*

$$\left|V^{0(s)}_{i,j,n,m}(x)\right| \ge c\epsilon^{\frac{1}{60}}; \tag{A.10}$$

- *when* **(S2)** *or* **(S3)** *holds, there exists* $s \in \{0,1,2\}$ *such that*

$$\left|V^{0(s)}_{i,j,n,m}(x)\right| \geq c\epsilon^{\frac{1}{100}}. \tag{A.11}$$

*Proof:* Consider the vectors

$$\bar{u}^{(s)}(x) = \begin{cases} \frac{u^{(s)}(x)}{\|u^{(s)}(x)\|_{\ell^1}}, & \|u^{(s)}(x)\|_{\ell^1} > 1 \\ u^{(s)}(x), & \|u^{(s)}(x)\|_{\ell^1} \leq 1 \end{cases}, \quad s = 0,1,2,3.$$

In view of (A.9),

$$|\det(U(x))| > c\left(\prod_{s=0}^{3} \frac{1}{\max\{\|u^{(s)}(x)\|_{\ell^1}, 1\}}\right)|\det(B_1(x))| > c(\epsilon^{\frac{1}{1200}})^{10} \cdot \epsilon^{\frac{1}{120}} > c\epsilon^{\frac{1}{60}},$$

for $x \in \mathcal{X}_0$. Apply Lemma A.2 with $v = (1,-1,1,-1)$, thus we get that there exists $s \in \{0,1,2,3\}$ such that

$$\left|V^{0(s)}_{i,j,n,m}(x)\right| = |\langle v, \tilde{u}^{(s)}(x)\rangle| = |\langle v, u^{(s)}(x)\rangle| \geq |\langle v, \bar{u}^{(s)}(x)\rangle| \geq c \cdot 4^{-\frac{3}{2}}\epsilon^{\frac{1}{60}}\|v\|_{\ell^1} = c\epsilon^{\frac{1}{60}}.$$

As for the subcases **(S2)** and **(S3)**, we can tackle with them similarly, applying Lemma A.2 with $v = (2,-1,-1)$ and $v = (1,1,-2)$ respectively, together with the corresponding conclusion Lemma A.1. ■

From now on, we set the constant $c = 1$ in (A.10) and (A.11) for convenience. The proof of Lemma 3.1 ends with the following lemma.

**Lemma A.3** *For $\epsilon$ sufficiently small, there is a subset $\mathcal{X}_\epsilon$ of $\mathcal{X}_0$ with*

$$\mathrm{mes}(\mathcal{X}_0 \setminus \mathcal{X}_\epsilon) < \epsilon^{\frac{1}{50}}$$

*such that for any $|i|, |j|, |n|, |m| \leq \kappa|\ln\epsilon|$ and $\{i,n\} \neq \{j,m\}$,*

$$|V^0_{i,j,n,m}(x)| \geq 2\epsilon^{\frac{1}{4}}, \quad x \in \mathcal{X}_\epsilon. \tag{A.12}$$

*Proof:* Fix $|i|, |j|, |n|, |m| \leq \kappa|\ln\epsilon|$ and $\{i,n\} \neq \{j,m\}$. Let us demonstrate that

$$\mathrm{mes}(\{x \in \mathcal{X}_0 : |V^0_{i,j,n,m}(x)| < 2\epsilon^{\frac{1}{4}}\}) < \epsilon^{\frac{1}{45}}.$$

We only deal with the subcase (**S1**), with the others done similarly. By Corollary 2, for each $x \in \mathcal{X}_0$, we have

$$\max_{0\leq s\leq 3}\left|V^{0(s)}_{i,j,n,m}(x)\right| \geq \epsilon^{\frac{1}{60}}.$$

Let $A := \max\limits_{0\leq s\leq 4}\sup\limits_{x\in\mathcal{X}_0}\left|V^{0(s)}_{i,j,n,m}(x)\right|$. In view of (A.9), $A \leq c\epsilon^{-\frac{1}{240}}$.

We first consider the function $V^0_{i,j,n,m}$ on $(a,b)$, one of the connected components of $\mathcal{X}_0$. Partition $(a,b)$ in about $2\epsilon^{-\frac{1}{24}}$ many intervals of length no more than $\frac{1}{2}\epsilon^{\frac{1}{24}}$. Choose one of such intervals, say $I$. Then either $|V^0_{i,j,n,m}(x)| \geq 2\epsilon^{\frac{1}{4}}$ for all $x \in I$, so we are done with the interval $I$, or there is some $x_0 \in I$ such that $|V^0_{i,j,n,m}(x_0)| < 2\epsilon^{\frac{1}{4}}$. In this case, for some $1 \le s \le 3$, $\left|V^{0(s)}_{i,j,n,m}(x_0)\right| \ge \epsilon^{\frac{1}{60}}$ by Corollary 2. Let us say $s = 3$, which is considered as the most complex case, so $\left|V^{0(3)}_{i,j,n,m}(x_0)\right| \ge \epsilon^{\frac{1}{60}}$. Since for $x \in I$,

$$\left|V^{0(3)}_{i,j,n,m}(x) - V^{0(3)}_{i,j,n,m}(x_0)\right| \le \sup_{y\in I}\left|V^{0(4)}_{i,j,n,m}(y)\right| \cdot |x - x_0| \le A|I| < \frac{1}{2}\epsilon^{\frac{1}{60}},$$

we obtain that $\left|V^{0(3)}_{i,j,n,m}(x)\right| \ge \frac{1}{2}\epsilon^{\frac{1}{60}}$.

Now we analyze $V^{0(2)}_{i,j,n,m}$ on $I$. If there is some $x_1 \in I$ such that $\left|V^{0(2)}_{i,j,n,m}(x_1)\right| < \epsilon^{\frac{1}{12}}$, then for every $x \in I$ with $|x - x_1| > 4\epsilon^{\frac{1}{15}}$, there is some $y \in I$ such that

$$\left|V^{0(2)}_{i,j,n,m}(x) - V^{0(2)}_{i,j,n,m}(x_1)\right| = \left|V^{0(3)}_{i,j,n,m}(y)\right| \cdot |x - x_1| \ge \frac{1}{2}\epsilon^{\frac{1}{60}} \cdot 4\epsilon^{\frac{1}{15}} = 2\epsilon^{\frac{1}{12}}.$$

Hence there exists an interval $I_1 \subset I$, which contains $x_1$, with $|I_1| \le 4\epsilon^{\frac{1}{15}}$, so that if $x \in I \setminus I_1$, then $\left|V^{0(2)}_{i,j,n,m}(x)\right| \ge \epsilon^{\frac{1}{12}}$.

We then consider $V^{0(1)}_{i,j,n,m}$ on $I \setminus I_1$, which has at most two connected components, denoted by $J_1$ and $J_2$. If there is some $x_2 \in J_1$ such that $\left|V^{0(1)}_{i,j,n,m}(x_2)\right| < \epsilon^{\frac{1}{6}}$, then for each $x \in J_1$ with $|x - x_2| > 2\epsilon^{\frac{1}{12}}$, there is some $y \in J_1$ such that

$$\left|V^{0(1)}_{i,j,n,m}(x) - V^{0(1)}_{i,j,n,m}(x_2)\right| = \left|V^{0(2)}_{i,j,n,m}(y)\right| \cdot |x - x_2| \ge \epsilon^{\frac{1}{12}} \cdot 2\epsilon^{\frac{1}{12}} = 2\epsilon^{\frac{1}{6}}.$$

Therefore, we obtain an interval $I_2 \subset J_1 \subset I \setminus I_1$ with $|I_2| \le 2\epsilon^{\frac{1}{12}}$, so that if $x \in J_1 \setminus I_2$, then $\left|V^{0(1)}_{i,j,n,m}(x)\right| \ge \epsilon^{\frac{1}{6}}$. Doing the same for $J_2$, we get an interval $I_3 \subset J_2 \subset I \setminus I_1$, with $|I_3| \le 2\epsilon^{\frac{1}{12}}$, such that if $x \in I \setminus (I_1 \cup I_2 \cup I_3)$, then $\left|V^{0(1)}_{i,j,n,m}(x)\right| \ge \epsilon^{\frac{1}{6}}$.

It is clear that there is at most four connected components contained in $I\setminus(I_1\cup I_2\cup I_3)$, say $J'_1$, $J'_2$, $J'_3$ and $J'_4$. If there is some $x'_1 \in J'_1$ such that $\left|V^{0}_{i,j,n,m}(x'_1)\right| < 2\epsilon^{\frac{1}{4}}$, then for each $x \in J'_1$ with $|x - x'_1| > 4\epsilon^{\frac{1}{12}}$, there is some $y \in J'_1$ such that

$$|V^{0}_{i,j,n,m}(x) - V^{0}_{i,j,n,m}(x'_1)| = \left|V^{0(1)}_{i,j,n,m}(y)\right| \cdot |x - x'_1| \ge \epsilon^{\frac{1}{6}} \cdot 4\epsilon^{\frac{1}{12}} = 4\epsilon^{\frac{1}{4}}.$$

Therefore, we obtain an interval $I'_1 \subset J'_1 \subset I \setminus (I_1 \cup I_2 \cup I_3)$, which contains $x'_1$, with $|I'_1| \le 4\epsilon^{\frac{1}{12}}$, so that if $x \in J'_1 \setminus I'_1$, then $|V^{0}_{i,j,n,m}(x)| \ge 2\epsilon^{\frac{1}{4}}$. Doing the same for $J'_2$, $J'_3$ and $J'_4$, we get intervals $I'_2$, $I'_3$ and $I'_4$, with $I'_k \subset J'_k \subset I \setminus (I_1 \cup I_2 \cup I_3)$ and $|I'_k| \le 4\epsilon^{\frac{1}{12}}$, $k = 2, 3, 4$, such that if $x \in \bigcup_{k=1}^{4} (J'_k \setminus I'_k)$, then

$$|V^{0}_{i,j,n,m}(x)| \ge 2\epsilon^{\frac{1}{4}}.$$

Hence, (A.12) holds on $I$ after excluding a subset with measure less than $5\epsilon^{\frac{1}{15}}$ since $\epsilon$ is sufficiently small. On the whole set $\mathcal{X}_0$, which is a finite union of no more than $c\kappa|\ln\epsilon|\cdot\epsilon^{-\frac{1}{24}}$ many intervals such as $I$, we need to exclude a subset with measure less than

$$c\kappa|\ln\epsilon| \cdot \epsilon^{-\frac{1}{24}} \cdot \epsilon^{\frac{1}{15}} < \epsilon^{\frac{1}{45}}.$$

Since the subscripts satisfy that $|i|, |j|, |n|, |m| \le \kappa|\ln\epsilon|$, the measure of the subset of parameters we exclude is less than $c\kappa^4|\ln\epsilon|^4 \cdot \epsilon^{\frac{1}{45}} < \epsilon^{\frac{1}{50}}$. ■

**Acknowledgements.** The authors are grateful to Jiangong You for his valuable discussions. We also appreciate the anonymous referees for helpful suggestions in modifying this manuscript.

## References

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

[2] Bellissard, J., Lima, R., Scoppola, E.: Localization in $\nu$-dimensional incommensurate structures. Commun. Math. Phys. **88**, 465–477(1983).

[3] Benettin, G., Galgani, L., Giorgilli, A.: A proof of Nekhoroshev's theorem for the stability times in nearly integrable Hamiltonian systems. Celest. Mech. **37**, 1–25(1985).

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

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

[6] Bourgain, J.: Quasiperiodic solutions of Hamiltonian perturbations of 2D linear Schrödinger equations. Ann. Math. **148**, 363–439(1998).

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

[8] Bourgain, J.: *Nonlinear Schrödinger equations. Park City Series*, **5**. Providence, RI: American Mathematical Society(1999).

[9] Bourgain, J., Wang, W. M.: Quasi-periodic solutions of nonlinear random Schrödinger equations. J. Eur. Math. Soc. **10**, 1–45(2008).

[10] Bourgain, J., Wang, W. M.: Diffusion bound for a nonlinear Schrödinger equation. *Mathematical Aspect of Nonlinear Dispersive Equations.*, 21–42. Princeton NJ: Princeton University Press(2007).

[11] Chierchia, L., You, J.: KAM tori for 1D nonlinear wave equations with periodic boundary conditions. Commun. Math. Phys. **211**, 498–525(2000).

[12] Craig, W.: Pure point spectrum for discrete almost periodic Schrödinger operators. Commun. Math. Phys. **88**, 113–131(1983).

[13] Craig, W., Wayne, C. E.: Newton's method and periodic solutions of nonlinear wave equations. Commun. Pure. Appl. Math. **46**, 1409–1498(1993).

[14] Cycon, H. L., Froese, R. G., Kirsch, W., Simon, B.: Schrödinger operators, with application to quantum mechanics and global geometry. Springer-Verlag(1987).

[15] del Rio, R., Jitomirskaya, S., Last, Y., Simon, B.: Operators with singular continuous spectrum, iv. Hausdorff dimensions, rank one perturbations, and localization. J. Anal. Math. **69**, 153–200(1996).

[16] Eliasson, L. H.: Perturbations of stable invariant tori for Hamiltonian systems. Ann. Sc. Norm. Sup. Pisa **15**, 115–147(1988).

[17] Eliasson, L. H., Kuksin, S. B.: KAM for the non-linear Schrödinger equation. Ann. Math., **172**, 371–435(2010).

[18] Fröhlich, J., Specer, T., Wayne, C. E.: Localization in disordered, nonlinear dynamical systems. J. Stat. Phys. **42**, 247–274(1986).

[19] Geng, J., Xu, X., You, J.: An infinite dimensional KAM theorem and its application to the two dimensional cubic Schrödinger equation. Advances in Mathematics, **226**, 5361–5402(2011).

[20] Geng, J., You, J.: A KAM theorem for one dimensional Schrödinger equation with periodic boundary conditions. J. Diff. Eqs. **209**, 1–56(2005).

[21] Geng, J., You, J.: A KAM theorem for Hamiltonian partial differential equationsin higher dimensional spaces. Commun. Math. Phys. **262**, 343–372(2006).

[22] Geng, J., You, J.: KAM tori for higher dimensional beam equations with constant potentials. Nonlinearity **19**, 2405–2423(2006).

[23] Geng, J., You, J.: A KAM theorem for higher dimensional nonlinear Schrödinger equations, J. Dynam. Differential Equations **25**, 451–476(2013).

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

[25] Geng, J., Yi, Y.: A KAM theorem for Hamiltonian networks with long ranged couplings. Nonlinearity **20**, 1313–1342(2007).

[26] Geng, J., Viveros, J., Yi, Y.: Quasi-periodic breathers in Hamiltonian networks of long-range coupling. Physical D. **237**, 2866–2892(2008).

[27] Gross, E. P.: Structure of a quantized vortex in boson systems. Nuovo Cimento **20**, 454–477(1961).

[28] Kappeler, T., Pöschel, J.: *KdV & KAM*, Springer(2003).

[29] Kuksin, S. B.: Hamiltonian perturbations of infinite–dimensional linear systems with an imaginary spectrum. Funct. Anal. Appl. **21**, 192–205(1987).

[30] Kuksin, S. B.: *Nearly integrable infinite dimensional Hamiltonian systems. Lecture Notes in Mathematics* **1556**, Springer(1993).

[31] Kuksin, S. B.: A KAM-theorem for equations of the Korteweg–de Vries type. Rev. Math. Phys. **10**, 1–64(1998).

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

[33] Klein, S.: Anderson localization for the discrete one-dimensional quasi-periodic Schrödinger operator with potential defined by a Gevrey-class function. J. Funct. Anal. **218**, 255–292(2005).

[34] Pitaevskii, L. P.: Vortex lines in an imperfect Bose gas. Sov. Phys. JETP **13**, 451(1961).

[35] Pöschel, J.: On elliptic lower dimensional tori in Hamiltonian systems. Math. Z. **202**, 559–608(1989).

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

[37] Pöschel, J.: A KAM Theorem for some nonlinear partial differential equations. Ann. Sc. Norm. Sup. Pisa Cl. Sci. **23**, 119–148(1996).

[38] Reed, M., Simon, B.: *Methods of modern mathematical physics. IV:* Analysis of operators. New York-San Francisco-London: Academic Press (1978).

[39] Simon, B.: Almost periodic Schrödinger operators: IV. the Maryland model. Annal. Phys. **159**, 157–183(1985).

[40] Trombettoni, A., Smerzi, A.: Discrete solitons and breathers with dilute Bose-Einstein condensates. Phys. Rev. Lett. **86**, 2353–2356(2001).

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

[42] Wang, W. M., Zhang, Z.: Long time Anderson localization for the nonlinear random Schrödinger equation. J. Stat. Phys. **134**, 953–968(2009).

[43] Xu, J., Qiu, Q., You, J.: A KAM theorem of degenerate infinite dimensional Hamiltonian systems (I). Sci. China Ser. A **39**, 372–383(1996).

[44] Xu, J., Qiu, Q., You, J.: A KAM theorem of degenerate infinite dimensional Hamiltonian systems (II). Sci. China Ser. A **39**, 384–394(1996).

[45] Yuan, X.: Construction of quasi-periodic breathers via KAM technique. Commun. Math. Phys. **226**, 61–100(2002).

[46] Yuan, X.: Quasi-periodic solutions of completely resonant nonlinear wave equations. J. Diff. Eqs. **230**, 213–274(2006).