ANDERSON LOCALIZATION

IN DISORDERED DYNAMICAL SYSTEMS

Dissertation Submitted
Nanjing University
in partial fufillment of the requirement
for the degree of
Doctor of Science

POSTGRADUATE:	Zhiyan Zhao
MENTOR:	Professor Jiangong You
SPECIALIZATION:	Mathematics
UNIVERSITY:	Nanjing University

April, 2013

Contents

Acknowledgements

THESIS: ANDERSON LOCALIZATION IN DISORDERED DYNAMICAL SYSTEMS

SPECIALIZATION: Mathematics

POSTGRADUATE: Zhiyan Zhao

MENTOR: Professor Jiangong You

Abstract

In this thesis, we try to explain and investigate Anderson localization, an intriguing physical phenomenon, from the perspective of mathematics. The disordered systems we consider are two quasi-crystal models, i.e.,

- one-dimensional nonlinear Maryland model:

$$\mathrm{i}\dot{q}_n = \epsilon(q_{n+1} + q_{n-1}) + \tan \pi(x + n\tilde{\alpha})q_n + \epsilon|q_n|^2 q_n, \quad n \in \mathbb{Z}, \tag{0.1}$$

  where $x \in \mathbb{R}/\mathbb{Z}$, and $\tilde{\alpha} \in \mathbb{R}$ is some fixed Diophantine vector;

- one-dimensional nonlinear quasi-periodic Schrödinger equation:

$$\mathrm{i}\dot{q}_n = \epsilon(q_{n+1} + q_{n-1}) + V(x + n\tilde{\alpha})q_n + |q_n|^2 q_n, \quad n \in \mathbb{Z}, \tag{0.2}$$

  where V is a nonconstant real-analytic function on $\mathbb{R}/\mathbb{Z}$, and $\tilde{\alpha}$ is some fixed Diophantine number.

In the first chapter, we take the ergodic Schrödinger operator as the main object of study, to explain localization in *linear* disorder systems. Some concepts in the spectral theory of operators, e.g., exponential localization, dynamical localization, will be given in this chapter. For three significant models, i.e., linear Anderson model, linear Maryland model and one-dimensional linear quasi-periodic Schrödinger operator, we shall state the corresponding conclusions respectively.

In the second chapter, we consider the one-dimensional nonlinear Maryland model. We shall prove that, for "most" compactly-supported small-amplitude initial data $(q_n(0))_{n\in\mathbb{Z}}$, if ϵ is sufficiently small, then for "most" $x \in \mathbb{R}/\mathbb{Z}$, the solution $(q_n(t))_{n\in\mathbb{Z}}$ of Equation (0.1) satisfies: for any fixed $s > 0$, the diffusion norm

$$\sum_{n\in\mathbb{Z}} n^{2s}|q_n(t)|^2$$

is uniformly bounded with respect to t.

In the third chapter, we consider the one-dimensional nonlinear quasi-periodic Schrödinger equation. For "most" compactly-supported initial data $(q_n(0))_{n\in\mathbb{Z}}$, if ϵ is sufficiently small, then for a.e. $x \in \mathbb{R}/\mathbb{Z}$, the solution $(q_n(t))_{n\in\mathbb{Z}}$ of Equation (0.2) satisfies: for any fixed $s > 0$,

$$\sup_t \sum_{n\in\mathbb{Z}} n^{2s}|q_n(t)|^2 < \infty.$$

Key Words: disordered medium; Anderson localization; nonliear Schrödinger equation; perturbation; KAM

THÈSE: LOCALISATION D'ANDERSON DANS LES SYSTÈMES DYNAMIQUES DÉSORDEONNÉS

DISCIPLINE: Mathématiques

PRÉSENTÉE PAR: Zhiyan ZHAO

DIRIGÉE PAR: Professeur Jiangong YOU

Résumé

Dans cette thèse, on essaye d'expliquer et d'étudier, du point de vue mathématique, la localisation d'Anderson, qui est un phénomène physique intéressant. Les systèmes désordonnés qu'on considère sont deux modèle de quasi-cristal, c'est-à-dire

- le modèle de Maryland non linéaire unidimensionnel:

$$i\dot{q}_n = \epsilon(q_{n+1} + q_{n-1}) + \tan\pi(x + n\tilde{\alpha})q_n + \epsilon|q_n|^2 q_n, \quad n \in \mathbb{Z}, \tag{0.3}$$

  avec $x \in \mathbb{R}/\mathbb{Z}$ et $\tilde{\alpha} \in \mathbb{R}$ un nombre diophantien fixé;

- l'équation de Schrödinger quasi-périodique non linéaire unidimensionnelle:

$$i\dot{q}_n = \epsilon(q_{n+1} + q_{n-1}) + V(x + n\tilde{\alpha})q_n + |q_n|^2 q_n, \quad n \in \mathbb{Z}, \tag{0.4}$$

  avec V une fonction analytique réelle non constante sur $\mathbb{R}/\mathbb{Z}$ et $\tilde{\alpha} \in \mathbb{R}$ un nombre diophantien fixé.

Dans le premier chapitre, on prend l'opérateur de Schrödinger ergodique comme l'objet de recherche principal, et explique la localisation dans les systèmes désordonnés *linéaires*. Quelques concepts dans la théorie spectrale de l'opérateur, par exemple, localisation exponentielle, localisation dynamique, seront donnés dans ce chapitre. Pour trois modèles importants, c'est-à-dire le modèle d'Anderson linéaire, le modèle de Maryland linéaire et l'opérateur de Schrödinger quasi-périodique linéaire unidimensionnel, on va énoncer les conclusions correspondantes respectivement.

Dans le deuxième chapitre, on considère le modèle de Maryland non linéaire unidimensionnel. On va prouver que pour la plupart des données initiales $(q_n(0))_{n\in\mathbb{Z}}$ avec le support compact et la petite amplitude, si ϵ est suffisamment petit, alors pour la plupart $x \in \mathbb{R}/\mathbb{Z}$, la solution $(q_n(t))_{n\in\mathbb{Z}}$ d'Équation (0.3) satisfait à: pour tout $s > 0$ fixé, la norme de diffusion

$$\sum_{n\in\mathbb{Z}} n^{2s}|q_n(t)|^2$$

est uniformément bornée par rapport à t.

Dans le troisième chapitre, on considère l'équation de Schrödinger quasi-périodique non linéaire unidimensionnelle. Pout la plupart des données initiales $(q_n(0))_{n\in\mathbb{Z}}$ avec le support compact, si ϵ est suffisamment petit, alors pour presque tout $x \in \mathbb{R}/\mathbb{Z}$, la solution $(q_n(t))_{n\in\mathbb{Z}}$ d'Equation (0.4) satisfait à: pour tout $s > 0$ fixé,

$$\sup_t \sum_{n\in\mathbb{Z}} n^{2s}|q_n(t)|^2 < \infty.$$

Mots-clés: milieu désordonné; localisation d'Anderson; équation de Schrödinger non linéaire; perturbation; KAM

Introduction

§0.1 Physical background

Localization of particles and waves in disordered media is one of the most intriguing phenomena in solid-state physics. This phenomenon was first analyzed by the American physicist P.W.Anderson[3], a Nobel prize winner in Physics. In Anderson's model, the disorderedness of the medium is generated by the random potential. He studied the transport of non-interacting electrons in such crystal lattice. If the amplitude of disorder becomes higher than a critical value, the diffusion in the lattice of an initially localized wavepacket is suppressed. After the work of Anderson, there are still several physicists who got the Nobel prize in Physics because of their outstanding contribution on the research of localization. In 2012, American physicist S.Haroche and American physicist D.Wineland have won this prize for ground-breaking experimental methods that enable measuring and manipulation of individual quantum systems.

In recent years, some media with relatively weak disorderedness, e.g., the quasi-crystal, have been concerned by physicists. Normally, such media can be introduced by a quasi-periodic potential. As an important model in Bose-Einstein condensate and optics, Maryland model[5] and Aubry-André model[4](also called Harper model) are typical examples. Anderson localization in such linear systems, especially in the one-dimensional case, has been thoroughly studied[42], and rigorous mathematical results have been established[32].

As a well-known model in mathematical physics, the almost Mathieu operator $H_{x,\lambda,\tilde{\alpha}}$ acting on $\ell^2(\mathbb{Z})$ is defined by

$$(H_{x,\lambda,\tilde{\alpha}}\psi)_n = (\psi_{n+1} + \psi_{n-1}) + \lambda \cos 2\pi(x + n\tilde{\alpha})\psi_n, \quad n \in \mathbb{Z},$$

where n is the primary lattice site index, $\tilde{\alpha}$ is some ratio between the wavenumbers of two lattices, $x \in \mathbb{R}/\mathbb{Z}$ is an arbitrary phase, and ψ_n is a complex variable whose modulus square gives the probability of finding a particle at the lattice site n. With $\tilde{\alpha}$ a fixed Diophantine number, for a.e. x and λ large enough, $H_{x,\lambda,\tilde{\alpha}}$ exhibits *dynamical* *localization*[23, 25], i.e., for any $\psi \in \ell^2(\mathbb{Z})$ with compact support and arbitrary $d > 0$,

$$\sup_t r^{(d)}(t) := \sup_t \sum_{n\in\mathbb{Z}} n^{2d} |(e^{\mathrm{i}H_{x,\lambda,\tilde{\alpha}}t}\psi)_n|^2 < \infty.$$

This means, at any moment, there is not too much energy transfer.

In particular, there exists a transparent transition between diffusion and localization for the almost Mathieu operator. From the perspective of spectral theory, it is shown by Jitomirskaya[32] that, for a.e. x, $H_{x,\lambda,\tilde{\alpha}}$ has

1. $\lambda > 2$: only pure point spectrum with exponentially decaying eigenfunctions;
2. $\lambda = 2$: purely singular continuous spectrum;
3. $\lambda < 2$: purely absolutely continuous spectrum.

There is a perfect agreement with this conclusion in some experiments(e.g., [30]). For $\tilde{\alpha} = \frac{\sqrt{5}-1}{2}$, with an initial $\delta-$function wavepacket, the asymptotic spreading of the wavepacket width $r^{(1)}(t)$ can be parametrized as $r^{(1)}(t) \sim t^{\gamma}$, and one finds three different regimes

1. $\lambda > 2$: localized regime, $\gamma = 0$;
2. $\lambda = 2$: sub-diffusive, $\gamma \sim 0.5$;
3. $\lambda < 2$: ballistic regime, $\gamma = 1$.

However, the situation is much less clear in the presence of interactions(nonlinearities). It strongly influences the possibility to observe the localization induced by disorderedness. One can start from the Gross-Pitaevskii(GP) equation[28, 38] in Hartree-Fock theory, and get a generalized Aubry-André model which includes an additional nonlinear term that represents the mean-field interaction. The Hamiltonian is

$$H = \sum_{n\in\mathbb{Z}} \left[(\psi_{n+1}\bar{\psi}_n + \bar{\psi}_{n+1}\psi_n) + \lambda \cos 2\pi(n\tilde{\alpha} + x)|\psi_n|^2 + \frac{1}{2}\beta|\psi_n|^4 \right],$$

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

$$\mathrm{i}\dot{\psi}_n + (\psi_{n+1} + \psi_{n-1}) + \lambda \cos 2\pi(n\tilde{\alpha} + x)\psi_n + \beta|\psi_n|^2\psi_n = 0, \quad n \in \mathbb{Z}, \tag{0.5}$$

that 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, [46]).

It is shown experimentally by Larcher-Dalfovo-Modugno[35] that, if the condensate initially occupies a single lattice site, i.e., a $\delta-$function $\psi_n(0) = \delta_{n,0}$, the dynamics of the interacting gas is dominated by self-trapping in a wide range of parameters, even for weak interaction. Conversely, if the diffusion starts from a Gaussian wavepacket of width σ, $\psi_n(0) = ce^{-\frac{n^2}{2\sigma^2}}$, then self-trapping is significantly suppressed and the destruction of localization by interaction is more easily observable. So, in the nonlinear systems, the different forms of the initial state influences the formation of localization, which is totally different from the linear case.

§0.2 Related works on mathematics

In the theory of mathematical physics, localization in disordered, nonlinear dynamical systems was initiated by Fröhlich-Spencer-Wayne[19](Similar work was also done by Pöschel[39] and Vittot-Bellissard[47]), who constructed infinite-dimensional, compact invariant tori for a large class of non-coupling systems

$$\mathrm{i}\dot{q}_n + V_n q_n + \sum_{m\in\mathbb{Z}} \epsilon_{mn}(q_m + \bar{q}_m)^2 q_n = 0, \quad n \in \mathbb{Z},$$

via the KAM techniques, where $\{V_n\}_{n\in\mathbb{Z}}$ are i.i.d. random variables, ϵ_{mn} are sufficiently small and vanish for $|m-n|$ large enough. Solutions which lie on such tori are localized for all times.

Besides the conclusion, they raised the following conjecture in that paper.

Conjecture.[19] Consider the equation

$$\mathrm{i}\dot{q}_n + \epsilon(q_{n+1} + q_{n-1}) + V_n q_n + \delta|q_n|^2 q_n = 0, \quad n \in \mathbb{Z}, \tag{0.6}$$

with $\{V_n\}_{n\in\mathbb{Z}}$ i.i.d. random variables. If ϵ and δ are small enough, with the equation in a large class, then for "most" initial conditions ("Most", e.g., with respect to the uniform measure on finite-dimensional unit spheres.), $q(0) = (q_n(0))_{n\in\mathbb{Z}}$, of finite support, the solutions $q(t) = (q_n(t))_{n\in\mathbb{Z}}$ of (0.6) satisfy

$$\overline{\lim_{t\to\infty}} t^{-1} \sum_{n\in\mathbb{Z}} n^2 |q_n(t)|^2 = 0.$$

Recently, there are several breakthroughs on such problem. For a large class of equation in (0.6), Bourgain-Wang[10] constructed a quasi-periodic solution when ϵ, δ are sufficiently small. The precise statement is

Theorem 0.1 *[10] Consider the nonlinear random Schrödinger equation(0.6). Fix* $\mathcal{J} = \{n_1, \cdots, n_b\} \in \mathbb{Z}$, $b > 1$, *and let* $\omega = (V_{n_1}, \cdots, V_{n_b}) \in \mathbb{R}^{\mathcal{J}}$. *When* $\epsilon = \delta = 0$, *the equation above has solutions*

$$u_0(y,t) = \sum_{j=1}^{b} a_k e^{-iV_{n_j}t} \delta_{n_j}(y), \quad y \in \mathbb{Z},$$

with $a = (a_1, \cdots, a_b)$ *satisfying that* $\sum_{j=1}^{b} |a_j|$ *is sufficiently small.*

For $0 < \epsilon \ll 1$, *there exists* $X_\epsilon \subset \mathbb{R}^{\mathbb{Z}} \setminus \mathbb{R}^{\mathcal{J}}$ *of positive probability such that for* $0 < \delta \ll 1$, *if we fix* $x \in X_\epsilon$, *there exists a Cantor set* $\mathcal{G}_{\epsilon,\delta}(x,a) \in \mathbb{R}^{\mathcal{J}}$ *of positive measure and a smooth function* $\omega_{\epsilon,\delta}(x,a)$ *defined on* $\mathcal{G}_{\epsilon,\delta}(x,a)$ *such that if* $\omega \in \mathcal{G}_{\epsilon,\delta}(x,a)$ *then*

$$u_{\epsilon,\delta,x}(y,t) = \sum_{(n,k)\in\mathbb{Z}^{d+b}} \hat{u}(n,k) e^{-i\langle k,\,\omega_{\epsilon,\delta}\rangle t} \delta_j(y)$$

is a solution to Eq. (0.6), with

$$\begin{aligned}
&\hat{u}(n_j, -e_j) = a_j, \quad k = 1, \cdots, b, \\
&\sum_{(n,k)\notin\mathcal{S}} e^{c(|n|+|k|)} |\hat{u}(n,k)| < \sqrt{\epsilon + \delta}, \\
&|\omega - \omega_{\epsilon,\delta}| < c(\epsilon + \delta),
\end{aligned}$$

for some $c > 0$, *where* $\{e_j\}_{j=1}^{b}$ *are the basis vectors for* $\mathbb{Z}^b$ *and* $\mathcal{S} = \{n_j, -e_j\}_{j=1}^{b} \subset \mathbb{Z}^{1+b}$. *The sets* X_ϵ *and* $\mathcal{G}_{\epsilon,\delta}(x,a)$ *satisfy*

$$\text{Prob} X_\epsilon \to 1, \ \text{mes}\mathbb{R}^{\mathcal{J}} \setminus \mathcal{G}_{\epsilon,\delta}(x,a) \to 0, \quad \text{as } \epsilon + \delta \to 0.$$

Corollary 0.1 *For* $0 < \epsilon, \delta \ll 1$, *there exists* $X_{\epsilon,\delta} \subset \mathbb{R}^{\mathbb{Z}^d}$ *of positive probability, satisfying*

$$\text{Prob} X_\epsilon \to 1, \quad \text{as } \epsilon + \delta \to 0,$$

such that for initial amplitudes a sufficiently small, there are quasi-periodic solutions to (0.6).

Remark 0.1 *The conclusion above has explained the conjecture of Fröhlich-Spencer-Wayne from another perspective. But for the conjecture "one equation has multiple well-localized solutions", it has not given a direct description.*

§0.3 Results of this thesis

This thesis aims to analyze the nonlinear equations which have the forms similar to (0.5), in order to investigate localization in nonlinear disordered dynamical systems. Inspired by the experimental conclusion of [35], and the conjecture of Fröhlich-Spencer-Wayne, we give rigorous mathematical arguments for localization in two important nonlinear quasi-crystal models.

(1) Consider the one-dimensional nonlinear Maryland model

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

where $\tilde{\alpha} \in \mathbb{R}^d$ is some fixed Diophantine number and $x \in \mathbb{R}/\mathbb{Z}$. For "most" compactly-supported small-amplitude initial data $(q_n(0))_{n\in\mathbb{Z}}$, if ϵ is sufficiently s-mall, then for "most" $x \in \mathbb{R}/\mathbb{Z}$, the solution $(q_n(t))_{n\in\mathbb{Z}}$ of the equation above satisfies: for any fixed $s > 0$, the diffusion norm

$$\sum_{n\in\mathbb{Z}} n^{2s}|q_n(t)|^2$$

is uniformly bounded with respect to t. See Theorem 2.1 for the precise statement.

(2) Consider one-dimensional nonlinear quasi-periodic Schrödinger equation:

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

where $\tilde{\alpha}$ is also a fixed Diophantine number, and V is a nonconstant real-analytic function on $\mathbb{R}/\mathbb{Z}$. For "most" compactly-supported initial data $(q_n(0))_{n\in\mathbb{Z}}$, if ϵ is sufficiently small, then for a.e. $x \in \mathbb{R}/\mathbb{Z}$, the solution $(q_n(t))_{n\in\mathbb{Z}}$ of the equation above satisfies: for any fixed $s > 0$,

$$\sup_t \sum_{n\in\mathbb{Z}} n^{2s}|q_n(t)|^2 < \infty.$$

See Theorem 3.1 for the precise statement.

In the formulations and proofs of various assertions of this thesis, 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.

第一章　Localization of linear Schrödinger operators

Consider the linear Schrödinger equation

$$\mathrm{i}\dot{q}_n = \epsilon(\Delta q)_n + V_n q_n, \quad n \in \mathbb{Z}^d, \tag{1.1}$$

where $d \geq 1$, Δ denotes the discrete Laplacian, i.e.,

$$\Delta_{ij} = \begin{cases} 1, & |i-j|_{\ell^1} = 1 \\ 0, & |i-j|_{\ell^1} \neq 1 \end{cases},$$ [1]

$\{V_n\}_{n\in\mathbb{Z}^d}$ satisfies some disordered condition and independent of time t. The property of its solution is completely determined by the linear operator H on $\ell^2(\mathbb{Z}^d)$:

$$(Hq)_n = \epsilon(\Delta q)_n + V_n q_n, \quad n \in \mathbb{Z}^d,$$

so localization for Eq. (1.1) can be also interpreted as localization for the operator H.

§1.1　Localization for ergodic operators

Definition 1.1 *Given a probability space* $(\Omega, \mathcal{F}, \mathbb{P})$*. We call a family of linear operators*

$$H_\theta : \ell^2(\mathbb{Z}^d) \to \ell^2(\mathbb{Z}^d), \quad \theta \in \Omega$$

is $\mathbb{Z}^d-$**ergodic**, *if there exists a family of ergodic measure-preserving transformations* $\{T_i\}_{i\in\mathbb{Z}^d}$ *on* $(\Omega, \mathcal{F}, \mathbb{P})$ *satisfying*

(1) Any T_i-*invariant subset* A *of* Ω *satisfies* $\mathbb{P}(A) = 0$ *or* 1*;*

(2) $H_{T_i\theta} = U_i H_\theta U_i^*$*, where the unitary operator* $U_i : \ell^2(\mathbb{Z}^d) \to \ell^2(\mathbb{Z}^d)$ *is defined by* $(U_i q)_n = q_{n-i}$*.*

We use $\sigma(H)(\sigma_{ac}(H)$, $\sigma_{sc}(H)$, $\sigma_{pp}(H))$ to denote the spectrum(absolutely continuous, singular continuous spectrum, pure point spectrum) of H. In the spectral theory of ergodic operators, we have the following elementary conclusion.

[1] From now on, we use $|\cdot|$ to denote the ℓ^1-distance on $\mathbb{Z}^d$.

Theorem 1.1 *(Pastur[36]) If H_θ is a family of $\mathbb{Z}^d-$ergodic self-adjoint operator, then there exists a closed set $\Sigma \subset \mathbb{R}$ such that, with $\mathbb{P}-$probability* 1,

$$\sigma(H_\theta) = \Sigma.$$

Moreover, there are closed subsets Σ_{ac}, Σ_{sc}, $\Sigma_{pp} \in \mathbb{R}$ such that, with $\mathbb{P}-$probability 1,

$$\sigma_{ac}(H_\theta) = \Sigma_{ac}, \quad \sigma_{sc}(H_\theta) = \Sigma_{sc}, \quad \sigma_{pp}(H_\theta) = \Sigma_{pp}.$$

We can study localization from the perspective of spectral theory or dynamical systems. Due to the different perspectives, we have different ways to define it. In this paper, we focus on the following three localizations.

Definition 1.2 *For a family of $\mathbb{Z}^d-$ergodic operator $H_\theta : \ell^2(\mathbb{Z}^d) \to \ell^2(\mathbb{Z}^d)$, we call that H_θ exhibits*

(1) **Spectral localization(SL)**
if, with $\mathbb{P}-$probability 1, H_θ has only pure point spectrum, i.e., $\Sigma = \Sigma_{pp}$ and $\Sigma_{ac} = \Sigma_{sc} = \emptyset$.

(2) **Exponential localization(EL)**
if H_θ exhibits spectral localization and, with $\mathbb{P}-$probability 1, its eigenfunctions are exponentially decaying.

(3) **Dynamical localization(DL)**
if, with $\mathbb{P}-$probability 1, for any compactly-supported $\psi \in \ell^2(\mathbb{Z}^d)$,

$$\sup_t \sum_{n\in\mathbb{Z}^d} |n|^{2s} |(e^{-\mathrm{i}tH_\theta}\psi)_n|^2 < \infty, \quad \forall s > 0.$$

The three localizations have the following implications:

$$(DL) \Rightarrow (EL) \Rightarrow (SL).$$

It is worth noting that, $(EL) \not\Rightarrow (DL)$(Refer to [15] for the construction of counter examples).

There is a well-known and important sufficient condition for dynamical localization

Definition 1.3 *If the family of $\mathbb{Z}^d-$ergodic operator H_θ exhibits (SL), and, with $\mathbb{P}-$probability 1, for its eigenvalue μ_n, $n \in \mathbb{Z}^d$, the corresponding eigenvector $\psi^n = (\psi^n_j)_{j\in\mathbb{Z}^d}$ satisfies*

$$|\psi^n_j| \le c_\sigma e^{\sigma|x_n|} e^{-r|j-x_n|}, \quad \forall \sigma > 0,$$

for some $r > 0$ and $|x_n| \sim |n|^{1/d}$, then we call that H_θ exhibits **semi-uniform localized eigenstates(SULE)**. *Furthermore, if*

$$|\psi^n_j| \le c_\sigma e^{-r|j-x_n|}, \quad \forall \sigma > 0,$$

then we call that H_θ exhibits **uniform localized eigenstates(ULE)**.

Obviously, (ULE)⇒(SULE).

Theorem 1.2 *(Rio-Jitormirskaya-Last-Simon[15]) If the family of $\mathbb{Z}^d-$ergodic operator H_θ exhibits semi-uniform localized eigenstates, then H_θ exhibits dynamical localization.*

Moreover, there are still some related conditions for dynamical localization, sufficient or necessary. See [45] for details.

§1.2 Linear Schrödinger operators

§1.2.1 Anderson model

Consider Anderson model $H : \ell^2(\mathbb{Z}^d) \to \ell^2(\mathbb{Z}^d)$,

$$(Hq)_n = \epsilon(\Delta q)_n + V_n q_n, \quad n \in \mathbb{Z}^d, \tag{1.2}$$

where Δ is the discrete Laplacian, and $\{V_n\}_{n\in\mathbb{Z}^d}$ is a family of independently identically distributed random variables, with the common distribution:

$$g = \tilde{g}(V_n)dV_n, \quad \tilde{g} \in L^\infty.$$

We also assume supp$g$ is a bounded set. The probability space is taken to be $\mathbb{R}^{\mathbb{Z}^d}$ with measure

$$\prod_{n\in\mathbb{Z}^d} g(V_n) = \prod_{n\in\mathbb{Z}^d} \tilde{g}(V_n)dV_n, \quad \tilde{g} \in L^\infty.$$

It is easy to verify that H is a $\mathbb{Z}^d-$ergodic self-adjoint operator, where $\{V_n\}_{n\in\mathbb{Z}^d}$ is the random variable in the probability space. For the spectrum of H, with probability 1, we have(by [14, 37])

$$\sigma(H) = [-2\epsilon d, 2\epsilon d] + \mathrm{supp} g.$$

Anderson model has been interested by mathematicians and physicists. About localization of Anderson model, there are plenty of well-known works[1, 2, 17, 18, 24, 26, 27, 48].

Theorem 1.3 *(Germinet–De Bièvre[24]) Consider Anderson model (1.2).*

- *When $d = 1$, H has dynamical localization.*
- *When $d > 1$, if ϵ is sufficiently small, then H has dynamical localization.*

§1.2.2 Maryland model

In this subsection, we will describe localization in the linear Maryland model in detail, i.e., consider the linear Schödinger operator $L = L(x)$ on $\ell^2(\mathbb{Z}^d)$:

$$(Lq)_n = \epsilon(\Delta q)_n + \tan(x + \langle n, \tilde{\alpha}\rangle)q_n, \quad n \in \mathbb{Z}^d,$$

where $\tilde{\alpha} \in \mathbb{R}^d$ satisfies the Diophantine condition: there exist constants $\tilde{\tau} > d$, $\tilde{\gamma} > 0$ such that

$$|\langle n, \tilde{\alpha}\rangle|_1 \geq \frac{\tilde{\gamma}}{|n|^{\tilde{\tau}}}, \quad \forall n \in \mathbb{Z}^d \setminus \{0\}, \tag{1.3}$$

and x comes from the full measure subset of $\mathbb{R}/\mathbb{Z}$:

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

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

$$L_{mn} = \begin{cases} \tan\pi(x + \langle n, \tilde{\alpha}\rangle), & m = n \\ \epsilon, & |m-n| = 1 \\ 0, & \text{otherwise} \end{cases}.$$

According to the property of tangent function and Diophantine vector, we know

$$|\tan\pi(x + \langle m, \tilde{\alpha}\rangle) - \tan\pi(x + \langle n, \tilde{\alpha}\rangle)| \geq \frac{\tilde{\gamma}}{|m-n|^{\tilde{\tau}}}, \quad m - n \in \mathbb{Z}^d \setminus \{0\}.$$

Hence, the operator L simulates the medium without resonance. This is widely and deeply applied in the KAM iteration.

Theorem 1.4 *Consider the linear Schrödinger operators on* $\ell^2(\mathbb{Z}^d)$*:*

$$(Lq)_n = \epsilon(\Delta q)_n + \tan\pi(x + \langle n, \tilde{\alpha}\rangle)q_n, \quad n \in \mathbb{Z}^d, \quad x \in \mathcal{X}, \tag{1.4}$$

where $\tilde{\alpha} \in \mathbb{R}^d$ *satisfies Diophantine condition(1.3). There exists a positive constant* $\epsilon_0 = \epsilon_0(\tilde{\alpha})$*, such that if* $0 < \epsilon < \epsilon_0$*, then the following holds.*

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

- *The poles of* $\hat{V}$ *are* $x = k + \frac{1}{2}$*,* $k \in \mathbb{Z}$*,*
- $\hat{V}(x) - \tan\pi x$ *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 for every $x \in \mathcal{X}$*, there is an orthogonal transform* $U : \ell^2(\mathbb{Z}) \to \ell^2(\mathbb{Z})$ *with*

$$|(U - I_{\mathbb{Z}^d})_{mn}| \leq c_L \epsilon e^{-2|m-n|}, \tag{1.5}$$

such that $U^*LU = \mathrm{diag}\{\hat{V}(x + \langle n, \tilde{\alpha}\rangle)\}_{n\in\mathbb{Z}^d}$*.*

This theorem (in its original form) is due to Bellissard-Lima-Scoppola[5]. The proof will be given in 附录三.

Corollary 1.1 *[5] Let* $0 < \epsilon < \epsilon_0$ *as in Theorem 1.4. For every* $x \in \mathcal{X}$*, the operator* $L = L(x)$ *has a complete family of exponentially decaying eigenvectors, and the set of eigenvalues is* $\{\hat{V}(x + \langle n, \tilde{\alpha}\rangle)\}_{n\in\mathbb{Z}^d}$*.*

The operator L has some other important property, we can find the statements ans proofs in references[5, 12, 15, 43].

According to the decay property of eigenvectors of L given as in (1.5), we can see that L exhibits uniformly localized eigenstates. Combining with Theorem 1.2, we get

Corollary 1.2 *Let* $0 < \epsilon < \epsilon_0$ *as in Theorem 1.4. Then the linear Schrödinger operator* L *exhibits dynamical localization.*

§1.2.3 Quasi-periodic Schrödinger operator

We consider the one-dimensional quasi-periodic Schrödinger operator $T = T(x)$: $\ell^2(\mathbb{Z}) \to \ell^2(\mathbb{Z})$,

$$(Tq)_n = \epsilon(q_{n+1} + q_{n-1}) + V(x + n\tilde{\alpha})q_n, \quad n \in \mathbb{Z}, \tag{1.6}$$

where $\tilde{\alpha} \in \mathbb{R}^1$ satisfies the Diophantine condition(1.3), and V is a nonconstant real-analytic function on $\mathbb{R}/\mathbb{Z}$. As in [16], the potential function V is a Gevrey function, i.e., there are C, $L > 0$ such that

$$\sup_{x\in\mathbb{R}/\mathbb{Z}} |\partial^m V(x)| \leq CL^m m!, \quad m \geq 0, \tag{1.7}$$

and there are also $\tilde{\xi}$, $\tilde{s} > 0$ satisfying the transversality condition

$$\max_{0\leq m\leq\tilde{s}} |\partial_\varphi^m (V(x+\varphi) - V(x))| \geq \tilde{\xi} > 0, \quad \forall x, \forall\varphi, \tag{1.8}$$

$$\max_{0\leq m\leq\tilde{s}} |\partial_x^m (V(x+\varphi) - V(x))| \geq \tilde{\xi}|\varphi|_1, \quad \forall x, \forall\varphi, \tag{1.9}$$

Clearly, the case $V(x) = \cos 2\pi x$ is included.

By [16] we know that if ϵ is sufficiently small, T has only pure point spectrum.

Theorem 1.5 *(Eliasson[16]) There exists $\epsilon_0 = \epsilon_0(V, \tilde{\alpha})$ such that if $0 < \epsilon < \epsilon_0$, then for a.e. $x \in \mathbb{R}/\mathbb{Z}$, the spectrum of T is purely pure point, with a complete set of eigenvectors in $\ell^2(\mathbb{Z})$. Moreover, the measure of the set $[\inf V, \sup V] \setminus \sigma(T_x)$ goes to 0 as $\epsilon \to 0$.*

About pure point spectrum and localization of quasi-periodic Schrödinger operators, there are still a lot of other works, e.g., [8, 11, 20, 31, 33, 44]. Since the idea of proof of [16] plays an important role in considering the nonlinear problem, we will describe it in detail.

Let us start with some notations for infinite-dimensional matrices. Given an infinite-dimensional matrix D, with $D_{mn} \in \mathbb{R}$ the $(m, n)^{\text{th}}$ entry, for a subset $\Lambda \subset \mathbb{Z}$, we define $\Lambda^\perp := \mathbb{Z} \setminus \Lambda$,

$$\mathbb{R}^\Lambda := \{n \in \mathbb{R}^\mathbb{Z} : n_i = 0 \ \text{ if } \ i \notin \Lambda\}, \quad D_\Lambda := \begin{cases} D_{mn}, & m, n \in \Lambda \\ \delta_{mn}, & \text{otherwise} \end{cases}.$$

Then $D_\Lambda : \mathbb{R}^\Lambda + \mathbb{R}^{\Lambda^\perp} \to \mathbb{R}^\Lambda + \mathbb{R}^{\Lambda^\perp}$, acts as $\mathbb{R}^\Lambda \hookrightarrow \mathbb{R}^\mathbb{Z} \xrightarrow{D} \mathbb{R}^\mathbb{Z} \xrightarrow{\perp \text{proj}} \mathbb{R}^\Lambda$ on the first component and as the identity on the second component. (When there is no risk for confusion, we will use D_Λ also to denote its first component.)

Let $D_0 = \text{diag}\{V(x+n\tilde{\alpha})\}_{n\in\mathbb{Z}}$ and $Z_0 = \epsilon\Delta$ with Δ the discrete Laplacian. With $\varepsilon_0 = \epsilon^{\frac{1}{4}}$, $\sigma_0 = 1$ and any

$$M_0 \geq \max\left\{2^{\tilde{s}+4} C \frac{L^{\tilde{s}+1}((\tilde{s}+1)!)^2}{\tilde{\xi}},\, 2\tilde{\tau},\, 8\right\}, \quad N_0 \geq 1, \quad \rho_0 = N_0^{-1},$$

one can define the following sequences as in [16],

$$\begin{aligned} M_{\nu+1} &= M_\nu^{\tilde{s}M_\nu^3}, \quad a_\nu = \frac{1}{\tilde{\tau}} M_\nu^{-3\tilde{s}M_\nu^3}, \quad \varepsilon_{\nu+1} = \varepsilon_\nu^{\frac{1}{2}\varepsilon_\nu^{-a_\nu/2}}, \\ N_{\nu+1} &= \varepsilon_\nu^{-a_\nu}, \qquad \rho_{\nu+1} = \varepsilon_\nu^{a_\nu}, \qquad\quad \sigma_{\nu+1} = \frac{1}{3}\rho_\nu. \end{aligned} \tag{1.10}$$

Afterwards, when we consider the nonlinear model, these sequences of parameters will appear in the KAM iteration in 第三章.

Theorem 1.6 *Let $0 < \epsilon < \epsilon_0$ as in Theorem 1.4. The following holds for one-dimensional quasi-periodic Schrödinger operator T.*

Fix any $x \in \mathbb{R}/\mathbb{Z}$. There exists a sequence of orthogonal matrices U_ν, $\nu = 1, 2, \cdots$, with

$$|(U_\nu - I_\mathbb{Z})_{mn}| \leq \varepsilon_0^{\frac{1}{2}} e^{-\frac{3}{2}\sigma_\nu |m-n|},$$

such that $U_\nu^*(D_0 + Z_0)U_\nu = D_\nu + Z_\nu$, where Z_ν is a symmetric matrix satisfying

$$|(Z_\nu)_{mn}| \leq \varepsilon_\nu e^{-\rho_\nu |m-n|},$$

and D_ν is a symmetric matrix which can be block-diagonalized via an orthogonal matrix Q_ν with

$$(Q_\nu)_{mn} = 0 \;\; if \;\; |m-n| > N_\nu.$$

More precisely, there is a disjoint decomposition $\bigcup_j \Lambda_j^\nu = \mathbb{Z}$ such that

$$\tilde{D}^\nu = Q_\nu^* D_\nu Q_\nu = \prod_j \tilde{D}^\nu_{\Lambda_j^\nu} \;\; with \;\; \sharp\Lambda_j^\nu \leq M_\nu, \;\; \text{diam}\Lambda_j^\nu \leq M_\nu N_\nu, \;\; \forall j.[2]$$

[2] The disjoint decomposition defines an equivalence relation $m \sim n$ on the integers and, for each $n \in \mathbb{Z}$, we denote its equivalence class by $\Lambda^\nu(n)$.

Moreover, there exists a full-measure subset $\tilde{\mathcal{X}} \subset \mathbb{R}/\mathbb{Z}$ *such that if we fix* $x \in \tilde{\mathcal{X}}$*, then for each* $k \in \mathbb{Z}$*, there is a* $\nu_0(k)$ *such that*

$$\Lambda^{\nu+1}(k) = \Lambda^{\nu}(k), \quad \forall \nu \geq \nu_0(k).$$

In 附录四, we shall give an outline of the proof.

We can also consider higher-dimensional quasi-periodic Schrödinger operator $\mathcal{H} : \ell^2(\mathbb{Z}^d) \to \ell^2(\mathbb{Z}^d)$, $d > 1$,

$$\mathcal{H} = \epsilon(\Delta q)_n + V(x_1 + n_1\alpha_1, \cdots, x_d + n_d\alpha_d)q_n, \quad x = (x_1, \cdots, x_d) \in \mathbb{R}^d/\mathbb{Z}^d,$$

where V is nonconstant real-analytic on $\mathbb{R}^d/\mathbb{Z}^d$. Bourgain-Goldstein-Schlag[9] has proven localization for the case $d = 2$, and Bourgain[7] generalized the conclusion into the case of arbitrary dimension.

Theorem 1.7 *[7] Fix any* $x \in \mathbb{R}^d/\mathbb{Z}^d$*. For any* $\delta > 0$*, there exists* $\epsilon_0 = \epsilon_0(V, \delta)$ *such that if* $0 < \epsilon < \epsilon_0$*, then there is a subset* $\Omega = \Omega(\epsilon, V) \subset \mathbb{R}^d/\mathbb{Z}^d$ *satisfying*

$$\text{mes}(\mathbb{R}^d/\mathbb{Z}^d \setminus \Omega) < \delta,$$

such that for $\alpha = (\alpha_1, \cdots, \alpha_d) \in \Omega$*,* $\mathcal{H}$ *exhibits exponential localization and dynamical localization.* [3]

[3]The definitions here of exponential localization and dynamical localization have been modified into for one operator but not for an ergodic family.

第二章　Localization in one-dimensional nonlinear Maryland model

Based on the conclusion of Bellissard-Lima-Scoppola[5], we consider the one-dimensional nonlinear Maryland model

$$\mathrm{i}\dot{q}_n = \epsilon(q_{n+1} + q_{n-1}) + \tan(x + n\tilde{\alpha})q_n + \epsilon|q_n|^2 q_n, \quad n \in \mathbb{Z}, \tag{2.1}$$

in this chapter, where $\tilde{\alpha} \in \mathbb{R}$ satisfies Diophantine condition(1.3), and x belongs to the full-measure subset of $\mathbb{R}/\mathbb{Z}$

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

§2.1　Statement of the result

First of all, we make a suitable coordinate transformation for Eq. (2.1), and then establish an abstract KAM theorem, which can be applied to the transformed system to study its localization. Localization for (2.1) can be derived by property of conjugate. It is worth mentioning that, after the transformation, the parameter for establishing the KAM theorem only comes from the nonlinearity. Indeed, the feasibility of the initial coordinate transformation is guaranteed by the special property of the tangent function and Diophantine number.

Theorem 2.1 *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

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

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

such that if $q_{\mathbb{Z}}(0) \in \epsilon^{\frac{\kappa}{2}} \cdot \mathcal{O}_\epsilon$*, the solution* $q_{\mathbb{Z}}(t) = (q_n(t))_{n\in\mathbb{Z}}$ *of Eq. (2.1) satisfies*

$$\sup_t \sum_{n\in\mathbb{Z}} n^{2d}|q_n(t)|^2 < \infty, \quad \forall d > 0.$$

Moreover, for every $n \in \mathbb{Z}$*,* $q_n(t)$ *is quasi-periodic with respect to* t*.*

Remark 2.1 *As stated in §0.1, the nonlinear term* $\epsilon|q_n|^2 q_n$ *in Eq. (2.1) 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.

§2.2 An abstract infinite-dimensional KAM theorem

§2.2.1 Function spaces and norms

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 $\mathbb{T}^b \times \{I = 0\} \times \{q = 0\} \times \{\bar{q} = 0\}$ 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 ℓ^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

[2]In the rest of the paper, all dependencies on ξ 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.

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^{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*

$$\begin{aligned} \operatorname{supp}(\alpha, \beta) &:= \{n \in \mathbb{Z}_1 : (\alpha_n, \beta_n) \neq (0, 0)\}, \\ n^+_{\alpha\beta} &:= \max\{n \in \operatorname{supp}(\alpha, \beta)\}, \\ n^-_{\alpha\beta} &:= \min\{n \in \operatorname{supp}(\alpha, \beta)\}, \\ n^*_{\alpha\beta} &:= \max\{|n^+_{\alpha\beta}|, |n^-_{\alpha\beta}|\}, \end{aligned}$$

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 $n^+_{\alpha\beta} = n^-_{\alpha\beta} = n^*_{\alpha\beta} = 0$.

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

$$\|F\|_{\mathcal{D},\mathcal{O}} := \sup_{\mathcal{D}} \|F\|_{\mathcal{O}}.^3 \tag{2.2}$$

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

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

[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 $\|F\|_{\mathcal{D},\mathcal{O}} := \sum_{i=1}^b \|F_i\|_{\mathcal{D},\mathcal{O}}$.

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.

§2.2.2 Statement of the KAM Theorem

First, for the integrable Hamiltonian, C_W^1 parametrized by $\xi \in \mathcal{O}$, with the following form

$$\mathcal{N} = e(\xi) + \langle \omega(\xi), I \rangle + \sum_{n\in\mathbb{Z}_1} \Omega_n(\xi) q_n \bar{q}_n, \quad \mathbb{Z}_1 \subset \mathbb{Z},$$

the phase space is equipped with the symplectic structure $dI \wedge d\theta + \mathrm{i}\sum_{n\in\mathbb{Z}_1} dq_n \wedge d\bar{q}_n$. For every $\xi \in \mathcal{O}$, the corresponding Hamilton equation of motion to $\mathcal{N}$ is

$$\frac{d\theta}{dt} = \omega, \quad \frac{dI}{dt} = 0, \quad \frac{dq_n}{dt} = -\mathrm{i}\Omega_n q_n, \quad \frac{d\bar{q}_n}{dt} = \mathrm{i}\Omega_n \bar{q}_n, \quad n \in \mathbb{Z}_1.$$

It 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.

Now we consider the following family of perturbed Hamiltonians

$$H = \mathcal{N} + P = e(\xi) + \langle \omega(\xi), I \rangle + \sum_{n\in\mathbb{Z}_1} \Omega_n(\xi) q_n \bar{q}_n + P(\theta, I, q, \bar{q}; \xi). \tag{2.3}$$

on some $\mathcal{D} = \mathcal{D}_{d,\rho}(r, s)$. We try to show the persistence of "most" of these $b-$tori for $H = \mathcal{N} + P$, provided that $\|X_P\|_{\mathcal{D},\mathcal{O}}$ is sufficiently small, and the solutions of the equation of motion on these tori are always well-localized.

Before stating the KAM theorem, we need to impose the following conditions on the frequencies ω, Ω_n and the perturbation P.

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

(A2) *Regularity of normal frequencies:* For each $n \in \mathbb{Z}_1$, Ω_n is a C_W^1 function of ξ 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_W^1 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

$$\|\breve{P}_{\alpha\beta}\|_{\mathcal{D},\mathcal{O}} \leq \begin{cases} \varepsilon e^{-\rho n^*_{\alpha\beta}}, & |\alpha| + |\beta| \leq 2 \\ e^{-\rho n^*_{\alpha\beta}}, & |\alpha| + |\beta| \geq 3 \end{cases}, \tag{2.4}$$

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

(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 2.2 *Assume that the Hamiltonian H in (2.3) satisfies* $(\mathbf{A1}) - (\mathbf{A5})$. *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 $\Psi : \mathbb{T}^b \times \mathcal{O}_\varepsilon \to \mathcal{D}_{d,0}(r/2, 0)$, 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 Ψ_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 H;*

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

Remark 2.2 *The statement* (d) *of the theorem above implies that*

$$\sup_t \sum_{n\in\mathbb{Z}} n^{2d} |q_n(t)|^2 < c \left(\sup_t \sum_{n\in\mathbb{Z}} \langle n\rangle^d |q_n(t)| \right)^2 < \infty,$$

which shows the conclusion of Theorem 2.1.

§2.3 Hamiltonian and normal form

Back to Eq. (2.1), we fix $x \in \mathcal{X}$. After the coordinate transformation $q_{\mathbb{Z}} = U\tilde{q}_{\mathbb{Z}}$, with U given in Theorem 1.4, there is no difference in the linear part, and the new equation transformed from the nonlinear equation (2.1) corresponds to the following Hamiltonian

$$H(\tilde{q}_{\mathbb{Z}}, \bar{\tilde{q}}_{\mathbb{Z}}) = \Lambda + G := \sum_{n\in\mathbb{Z}} \hat{V}_n |\tilde{q}_n|^2 + \frac{1}{2}\epsilon \sum_{i,j,m,n\in\mathbb{Z}} u_{ijnm}\tilde{q}_i\bar{\tilde{q}}_j\tilde{q}_m\bar{\tilde{q}}_n, \tag{2.6}$$

where $\hat{V}_n = \hat{V}_n(x) := \hat{V}(x + n\tilde{\alpha})$. The off-diagonal decay of U in (1.5) implies the short-range estimates of coefficients u_{ijnm}, i.e.,

$$|u_{ijmn}| < ce^{-2(\max\{i,j,m,n\}-\min\{i,j,m,n\})}. \tag{2.7}$$

Indeed, we can calculate that

$$u_{ijmn} = \sum_{l\in\mathbb{Z}} U_{li}\bar{U}_{lj}U_{lm}\bar{U}_{ln}. \tag{2.8}$$

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

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

We fix the tangential directions $\mathcal{J} = \{n_1, \cdots, n_b\} \subset \mathbb{Z}$, and $\mathbb{Z}_1 = \mathbb{Z} \setminus \mathcal{J}$. When ϵ 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 §2.2.1. 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 2.1 *For ϵ sufficiently small, there exists a subset $\mathcal{X}_\epsilon$ of $\mathcal{X}$ with*

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

such that for every $x \in \mathcal{X}_\epsilon$, there is a symplectic transformation $\Psi = \Psi(x)$, which transforms H in (2.6) into

$$\begin{aligned} H \circ \Psi &= \mathcal{N} + P \\ &:= e(\xi) + \langle \omega(\xi), I \rangle + \sum_{n \in \mathbb{Z}_1} \Omega_n(\xi) q_n \bar{q}_n + P(\theta, I, q, \bar{q}; \xi), \end{aligned} \tag{2.9}$$

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

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

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

$$\begin{aligned} \breve{P} &= \breve{P}(\theta, I, q, \bar{q}; \xi) = \sum_{\alpha,\beta} \breve{P}_{\alpha\beta} q^\alpha \bar{q}^\beta = \sum_{\substack{(k,l)\neq 0 \\ \alpha,\beta}} P_{kl\alpha\beta} I^l e^{\mathrm{i}\langle k,\theta\rangle} q^\alpha \bar{q}^\beta, \\ \acute{P} &= \acute{P}(q, \bar{q}; \xi) = \sum_{\alpha,\beta} \acute{P}_{\alpha\beta} q^\alpha \bar{q}^\beta = \sum_{\alpha,\beta} P_{00\alpha\beta} q^\alpha \bar{q}^\beta, \end{aligned}$$

satisfying

$$\|\breve{P}_{\alpha\beta}\|_{\mathcal{D},\mathcal{O}} \le \begin{cases} \epsilon^{\frac{\kappa}{4}} e^{-\frac{1}{2} n^*_{\alpha\beta}}, & |\alpha| + |\beta| \le 2 \\ e^{-\frac{1}{2} n^*_{\alpha\beta}}, & |\alpha| + |\beta| \ge 3 \end{cases}, \tag{2.10}$$

$$\|\acute{P}_{\alpha\beta}\|_{\mathcal{D},\mathcal{O}} \le \begin{cases} \epsilon^{\frac{\kappa}{4}} e^{-\frac{1}{2} n^*_{\alpha\beta}}, & |\alpha| + |\beta| \le 2 \\ e^{-\frac{1}{2}(n^+_{\alpha\beta} - n^-_{\alpha\beta})}, & |\alpha| + |\beta| \ge 3 \end{cases}. \tag{2.11}$$

Proof. We decompose the proof into the following parts.

- **Construction of symplectic changes of variables**

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

$$\begin{aligned} T(q_{\mathbb{Z}}, \bar{q}_{\mathbb{Z}}) &= \frac{1}{2}\epsilon \sum_{|i|,|j|,|m|,|n| \le \kappa |\ln \epsilon|} u_{ijmn} q_i \bar{q}_j q_m \bar{q}_n, \\ F(q_{\mathbb{Z}}, \bar{q}_{\mathbb{Z}}) &= \frac{\mathrm{i}}{2}\epsilon \sum_{\substack{\hat{V}_i - \hat{V}_j + \hat{V}_m - \hat{V}_n \neq 0 \\ |i|,|j|,|m|,|n| \le \kappa |\ln \epsilon|}} \frac{u_{ijmn}}{\hat{V}_i - \hat{V}_j + \hat{V}_m - \hat{V}_n} q_i \bar{q}_j q_m \bar{q}_n, \end{aligned}$$

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

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

Since ϵ is small enough, by Lemma 2.1 below, there exists a subset $\mathcal{X}_\epsilon$ of $\mathcal{X}$ with

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

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

In view of the homological equation

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

we know that the change of variables Ψ_F^1 sends H to

$$H \circ \Psi_F^1 = \sum_{i\in\mathbb{Z}} \hat{V}_i|q_i|^2 + \frac{1}{2}\epsilon \sum_{|i|,|j| \leq \kappa|\ln\epsilon|} u_{iijj}|q_i|^2|q_j|^2 + \tilde{R}, \tag{2.12}$$

where

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

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 $\text{supp}(\alpha', \beta')$, $n^+_{\alpha'\beta'}$, $n^-_{\alpha'\beta'}$, $n^*_{\alpha'\beta'}$ and $|\alpha'|$, $|\beta'|$ can be defined as in Definition 2.1. By the construction of $\tilde{R}$, we have

$$\tilde{R}_{\alpha'\beta'} = 0, \quad |\alpha'| \neq |\beta'|, \tag{2.13}$$

$$\tilde{R}_{\alpha'\beta'} = 0, \quad |\alpha'| + |\beta'| < 4, \tag{2.14}$$

$$\tilde{R}_{\alpha'\beta'} = 0, \quad |\alpha'| + |\beta'| = 4, \;\; n^*_{\alpha'\beta'} \leq \kappa|\ln\epsilon|. \tag{2.15}$$

Moreover, by applying Lemma 2.2 below iteratively, [4]

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

[4] For convenience of expression, we assume that the constant in Theorem 1.4 $c_L = 1$.

- **Introduction of action-angle variables**

Introducing the action-angle variables and the amplitude parameters to the Hamiltonian (2.12),

$$q_n = \sqrt{I_n + \xi_n} e^{\mathrm{i}\theta_n}, \quad \bar{q}_n = \sqrt{I_n + \xi_n} e^{-\mathrm{i}\theta_n}, \quad n \in \mathcal{J},$$

where $(I, \theta) = (I_{n_1}, \cdots, I_{n_b}, \theta_{n_1}, \cdots, \theta_{n_b})$ are the standard action-angle variables in the $(q_n, \bar{q}_n)_{n\in\mathcal{J}}$-space around ξ, with $\xi = (\xi_{n_1}, \cdots, \xi_{n_b}) \in \epsilon^{\kappa}[\epsilon^{\frac{\kappa}{12}}, 1]^b$ the amplitude parameter, and $(q, \bar{q}) = (q_n, \bar{q}_n)_{n\in\mathbb{Z}_1}$. Then the Hamiltonian in (2.12) becomes

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

where

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

By the scaling in time

$$\theta \to \theta, \quad I \to \epsilon^{\frac{4}{3}\kappa} I, \quad q \to \epsilon^{\frac{2}{3}\kappa} q, \quad \bar{q} \to \epsilon^{\frac{2}{3}\kappa} \bar{q}, \quad \xi \to \epsilon^{\kappa}\xi, \tag{2.16}$$

we finally arrive at the rescaled Hamiltonian

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

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

$$e = \epsilon^{-(1+\frac{4}{3}\kappa)} \sum_{i\in\mathcal{J}} \hat{V}_i \xi_i + \frac{1}{2}\epsilon^{-\frac{\kappa}{3}} \sum_{i\in\mathcal{J}} u_{iiii}\xi_i^2 + \frac{1}{2}\epsilon^{-\frac{\kappa}{3}} \sum_{\substack{i,j\in\mathcal{J}\\ i\neq j}} u_{iijj}\xi_i\xi_j, \tag{2.17}$$

$$\omega_i(\xi) = \epsilon^{-(1+\kappa)}\hat{V}_i + u_{iiii}\xi_i + \frac{1}{2}\sum_{\substack{j\in\mathcal{J}\\ j\neq i}} u_{iijj}\xi_j, \quad i \in \mathcal{J}, \tag{2.18}$$

$$\Omega_n(\xi) = \left\{ \begin{array}{ll} \epsilon^{-(1+\kappa)}\hat{V}_n + \frac{1}{2}\sum_{i\in\mathcal{J}} u_{iinn}\xi_i, & |n| \leq \kappa|\ln\epsilon| \\ \epsilon^{-(1+\kappa)}\hat{V}_n, & |n| > \kappa|\ln\epsilon| \end{array} \right., \quad n \in \mathbb{Z}_1 \tag{2.19}$$

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 new Hamiltonian $\mathcal{N} + P$**

As shown in (2.8), $u_{iijj} = \sum_{l\in\mathbb{Z}} |U_{il}|^2|U_{jl}|^2$, so, in view of (2.18), the $b \times b$ matrix $\frac{\partial\omega}{\partial\xi}$ satisfies that

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

By (1.5), we have

$$|U_{ii} - 1| < \epsilon, \quad |U_{il}| \leq \epsilon e^{-2|i-l|}, \ \ l \neq i.$$

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

The formulation of Ω_n given in (2.19) implies that $\partial_\xi\Omega_n = 0$ for $|n| > \kappa|\ln\epsilon|$. As for the case $|n| \leq \kappa|\ln\epsilon|$, we have

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

For $n \in \mathbb{Z}_1$, the formulation of Ω_n given in (2.19) implies that $\partial_\xi\Omega_n = 0$ for $|n| > \kappa|\ln\epsilon|$. As for the case $|n| \leq \kappa|\ln\epsilon|$, we have

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

By (2.13) and (2.14), each non-zero term of $\tilde{R}$ can be rewritten as

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

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 terms like

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

which, after the scaling (2.16), becomes

$$\mathcal{E}\tilde{R}_{\alpha'\beta'}\left(\prod_{n\in\mathcal{J}}\left(\sqrt{\epsilon^{\frac{\kappa}{3}}I_n+\xi_n}\right)^{\alpha_n+\beta_n}e^{\mathrm{i}(\alpha_n-\beta_n)\theta_n}\right)q^{\alpha}\bar{q}^{\beta}, \tag{2.20}$$

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 (2.13),

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

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 (2.14), and the term in (2.20) 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

$$\epsilon^{\frac{\kappa}{3}-1}|\tilde{R}_{\alpha'\beta'}|\leq\epsilon^{\frac{\kappa}{3}-1}\cdot\epsilon e^{-2(n^{+}_{\alpha'\beta'}-n^{-}_{\alpha'\beta'})}\leq\epsilon^{\frac{\kappa}{3}}e^{-2(n^{+}_{\alpha'\beta'}-n^{-}_{\alpha'\beta'})}. \tag{2.21}$$

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 (2.14), 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

$$\mathcal{E}|\tilde{R}_{\alpha'\beta'}| \leq \epsilon^{\frac{2}{3}\kappa-1} \cdot \epsilon^{1-\frac{\kappa}{3}} e^{-2n^*_{\alpha'\beta'}} \leq \epsilon^{\frac{\kappa}{3}} e^{-2n^*_{\alpha'\beta'}}. \tag{2.22}$$

- If $|\alpha'|+|\beta'| = 4$, then by (2.15), $n^*_{\alpha'\beta'} > \kappa|\ln\epsilon|$, 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 (2.20) is not more than

$$\mathcal{E}|\tilde{R}_{\alpha'\beta'}| \leq \epsilon^{-(1+\frac{7}{3}\kappa)} \epsilon^{2\kappa} \epsilon^{1+\frac{2}{3}\kappa} e^{-n^*_{\alpha'\beta'}} \leq \epsilon^{\frac{\kappa}{3}} e^{-n^*_{\alpha'\beta'}}. \tag{2.23}$$

Thus, Case 2), the coefficient of $q^\alpha \bar{q}^\beta$ in (2.20) 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 ξ_n, we need to keep ξ_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 (2.10).

Terms of $\acute{P}$ come from both cases. When the term in (2.20) satisfies that $\alpha_{\mathcal{J}} = \beta_{\mathcal{J}}$, by expanding $\sqrt{I_n+\xi_n}$ around ξ_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 (2.21), (2.11) is proved. □

Combing (2.21) − (2.23) together, we have

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

To this stage, we have that all the assumptions of Theorem 2.2 hold for (2.9), which conjugates with (2.1). Thus, Theorem 2.1 follows from Theorem 2.2.

We have applied several conclusions directly in proving Proposition 2.1. Now we give their precise statements. The first lemma shows that the function

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

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

Lemma 2.1 *For ϵ sufficiently small, there exists a subset $\mathcal{X}_\epsilon$ of $\mathcal{X}$ with*

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

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

$$|V_{i,j,m,n}(x)| \ge \epsilon^{\frac{1}{4}}, \quad \forall x\in\mathcal{X}_\epsilon. \tag{2.24}$$

The proof of Lemma 2.1 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 附录五.

The next lemma implies that the property (2.7) about the coefficients of the Hamiltonian is preserved under the poisson bracket.

Lemma 2.2 *Consider two real-analytic functions*[5]

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

with

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

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

for some positive c_G, c_F and σ. 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

$$K_{\alpha\beta}=\mathrm{i}\sum_{\mathcal{S}}\left(G_{\check{\alpha}+e_n,\check{\beta}}F_{\hat{\alpha},\hat{\beta}+e_n}-G_{\check{\alpha},\check{\beta}+e_n}F_{\hat{\alpha}+e_n,\hat{\beta}}\right), \tag{2.25}$$

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 (2.25), 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 (2.25), 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. □

§2.4 KAM iteration

The remaining sections of this chapter are devoted to the proof of Theorem 2.2. In this section we present the KAM iteration scheme applied to (2.3). This is a succession of infinitely many symplectic transformations. We will show that, under these symplectic transformations, the perturbation is made smaller at the cost of excluding a small-measure set of parameters and some weight of exponent. It will be shown in the next section that the sequence of the symplectic transformations converges and, to finish the proof of Theorem 2.2, the total measure of the set of parameters that has been excluded is small.

§2.4.1 Normal form

In order to perform the KAM iteration scheme, we shall first write the Hamiltonian (2.3) 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 $\mathcal{D}_0 = \mathcal{D}_{d,\rho_0}(r_0, s_0)$.

We first consider the lower-order terms of $\breve{P}$ and $\acute{P}$. According to (2.4) and (2.5) in the assumption (**A4**) and the definition of norm (2.2), we have that coefficients of

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

satisfy that

$$|P_{kl\alpha\beta}|_{\mathcal{O}} \le \varepsilon e^{-\rho n^*_{\alpha\beta}} e^{-|k|r}, \quad \forall k \in \mathbb{Z}^b, \ \ 2|l| + |\alpha| + |\beta| \le 2. \tag{2.26}$$

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

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

and then

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

It follows, from (2.26) and the definition of the vector field norm, that one can make s_0 small enough so that

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

We can rewrite R as

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

where

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

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

$$P_n^{010},\, P_n^{001},\, P_{mn}^{020},\, P_{mn}^{002} \equiv 0. \tag{2.27}$$

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

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

such that all non–resonant terms

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

will be eliminated, and terms

$$P_{0l00} I^l, \quad |l|\leq 1; \quad P_{mn}^{011} q_m \bar{q}_n, \quad |m|,|n|\leq K_0,$$

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

$$\{\mathcal{N}, F_*\} + R = \sum_{|l|\leq 1} P_{0l00} I^l + \sum_{|m|,|n|\leq K_0} P_{mn}^{011} q_m \bar{q}_n.$$

One can show that it is solvable on the parameter set

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

By virtue of (2.27), which is guaranteed by gauge invariance of P, we need not consider the lower bound of $|\Omega_n|$ or $|\Omega_m \pm \Omega_n|$.

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

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

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

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

The other conditions can be handled similarly.

With $\Phi_* = \Phi^1_{F_*}$, the Hamiltonian (2.3) can be transformed into the following system on $\mathcal{D}_0 := \mathcal{D}_{d,\rho_0}(r_0, s_0)$:

$$H_0 = H \circ \Phi_* = \mathcal{N}_0 + P_0,$$

with $\mathcal{N}_0$ and P_0 given as

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

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

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

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

$$\|\breve{P}^0_{\alpha\beta}\|_{\mathcal{D}_0,\mathcal{O}_0} \leq \begin{cases} \varepsilon_0 e^{-\rho_0 n^*_{\alpha\beta}}, & |\alpha| + |\beta| \leq 2 \\ e^{-\rho_0 n^*_{\alpha\beta}}, & |\alpha| + |\beta| \geq 3 \end{cases},$$

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

We shall prove that the decay property is preserved during the KAM iteration in the following subsection.

Suppose that, we have arrived at the ν^{th} KAM step, and we consider the Hamiltonian $H_\nu = \mathcal{N}_\nu + P_\nu$, which is real-analytic on $\mathcal{D}_\nu = \mathcal{D}_{d,\rho_\nu}(r_\nu, s_\nu)$, and C^1_W parametrized by $\xi \in \mathcal{O}_\nu(\mathcal{O}_\nu \subset \mathcal{O}_0$ is a parameter set), 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_ν satisfies that $\|X_{P_\nu}\|_{\mathcal{D}_\nu,\mathcal{O}_\nu} < \varepsilon_\nu$ and

$$\|\breve{P}^\nu_{\alpha\beta}\|_{\mathcal{D}_\nu,\mathcal{O}_\nu} \le \begin{cases} \varepsilon_\nu e^{-\rho_\nu n^*_{\alpha\beta}}, & |\alpha| + |\beta| \le 2 \\ e^{-\rho_\nu n^*_{\alpha\beta}}, & |\alpha| + |\beta| \ge 3 \end{cases}, \tag{2.28}$$

$$\|\acute{P}^\nu_{\alpha\beta}\|_{\mathcal{D}_\nu,\mathcal{O}_\nu} \le \begin{cases} \varepsilon_\nu e^{-\rho_\nu n^*_{\alpha\beta}}, & |\alpha| + |\beta| \le 2 \\ e^{-\rho_\nu (n^+_{\alpha\beta} - n^-_{\alpha\beta})}, & |\alpha| + |\beta| \ge 3 \end{cases}. \tag{2.29}$$

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 $H_{\nu+1} = H_\nu \circ \Phi_\nu = \mathcal{N}_{\nu+1} + P_{\nu+1}$, C^1_W parametrized by $\xi \in \mathcal{O}_{\nu+1}$, has similar properties with H_ν, 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) "ν" of quantities at the ν^{th} step are neglected, and the corresponding quantities at the $(\nu+1)^{\text{th}}$ 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 the normal form $\mathcal{N}$,

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

where $z_+ = (q_n)_{|n|\le K_+}$, $\bar z_+ = (\bar q_n)_{|n|\le K_+}$, and $\tilde A$ is a Hermitian matrix with $\dim(\tilde A) \le 2K_+ + 1$ given by

$$\tilde A = \begin{pmatrix} A & 0 \\ 0 & \Omega_n \end{pmatrix}_{K<|n|\le K_+}. \tag{2.30}$$

§2.4.2 Truncation and homological equation

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

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

By (2.28) and (2.29), and the definition of norm $\|\cdot\|_{\mathcal{D},\mathcal{O}}$,

$$|P_{kl\alpha\beta}|_{\mathcal{O}} \le \varepsilon e^{-\rho n^*_{\alpha\beta}} e^{-|k|r}, \quad \forall k \in \mathbb{Z}^b, \ \ 2|l| + |\alpha| + |\beta| \le 2. \tag{2.31}$$

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

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

with

$$\begin{aligned}
R_0 &= \sum_{\substack{k\\ |l|\le 1}} P_{kl00} e^{\mathrm{i}\langle k,\theta\rangle} I^l, \\
R_1 &= \sum_{\substack{k\\ |n|\le K_+}} (P_n^{k10} q_n + P_n^{k01} \bar q_n) e^{\mathrm{i}\langle k,\theta\rangle} = \sum_k (\langle R^{k10}, z_+\rangle + \langle R^{k01}, \bar z_+\rangle) e^{\mathrm{i}\langle k,\theta\rangle} \\
R_2 &= \sum_{\substack{k\\ |m|,|n|\le K_+}} (P_{mn}^{k20} q_m q_n + P_{mn}^{k11} q_m \bar q_n + P_{mn}^{k02} \bar q_m \bar q_n) e^{\mathrm{i}\langle k,\theta\rangle} \\
&= \sum_k (\langle R^{k20} z_+, z_+\rangle + \langle R^{k11} z_+, \bar z_+\rangle + \langle R^{k02} \bar z_+, \bar z_+\rangle) e^{\mathrm{i}\langle k,\theta\rangle}.
\end{aligned}$$

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

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

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

$$\begin{aligned}
&\overline{P_{(-k)l00}} = P_{kl00}, \quad \overline{R^{(-k)10}} = R^{k01}, \quad \overline{R^{(-k)01}} = R^{k10}, \\
&\overline{R^{(-k)20}} = R^{k02}, \quad \overline{R^{(-k)11}}^{\top} = R^{k11}, \quad \overline{R^{(-k)02}} = R^{k20}.
\end{aligned} \tag{2.32}$$

From our definition of norms, it follows that

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

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

$$P - R = \sum_{\substack{k,l \\ 2|l|+|\alpha|+|\beta|\geq 3}} P_{kl\alpha\beta} e^{\mathrm{i}\langle k,\theta\rangle} I^l q^\alpha \bar{q}^\beta + \sum_{\substack{k,l,\, n^*_{\alpha\beta}>K_+ \\ 2|l|+|\alpha|+|\beta|\leq 2}} P_{kl\alpha\beta} e^{\mathrm{i}\langle k,\theta\rangle} I^l q^\alpha \bar{q}^\beta, \tag{2.33}$$

combining with (2.31), there exists $c_1 > 0$ such that

$$\|X_{P-R}\|_{\mathcal{D}_{d,\rho_+}(r_+ + \frac{r-r_+}{2}, \eta s),\mathcal{O}} \leq \varepsilon \sum_{|n|>K_+} e^{-\rho|n|} + c_1 \eta s \leq \frac{1}{4}\varepsilon^{\frac{5}{4}}, \tag{2.34}$$

provided that
(C1): $e^{-(\rho-\rho_+)K_+} \leq \frac{1}{8}\varepsilon^{\frac{1}{4}}$ and $c_1 s \leq \frac{1}{8}\varepsilon$.

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

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

and satisfy the homological equation

$$\{\mathcal{N}, F\} + R = e' + \langle \omega', I\rangle + \langle R^{011} z_+, \bar{z}_+\rangle. \tag{2.35}$$

where $e' = P_{0000}$ and $\omega' = P_{0l00}(|l| = 1)$. By simple comparison of coefficients, we can see Equation (2.35) is equivalent to the following system

$$\langle k, \omega\rangle F_{kl00} = \mathrm{i}P_{kl00}, \tag{2.36}$$

$$((\langle k, \omega\rangle I - \tilde{A})F^{k10} = \mathrm{i}R^{k10}, \tag{2.37}$$

$$((\langle k, \omega\rangle I + \tilde{A})F^{k01} = \mathrm{i}R^{k01}, \tag{2.38}$$

$$((\langle k, \omega\rangle I - \tilde{A})F^{k20} - F^{k20}\tilde{A} = \mathrm{i}R^{k20}, \tag{2.39}$$

$$((\langle k, \omega\rangle I - \tilde{A})F^{k11} + F^{k11}\tilde{A} = \mathrm{i}R^{k11}, \tag{2.40}$$

$$((\langle k, \omega\rangle I + \tilde{A})F^{k02} + F^{k02}\tilde{A} = \mathrm{i}R^{k02}. \tag{2.41}$$

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 (2.30), 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 $\tilde{A}$, we have that

$$Q_{mn} \equiv 0 \ \text{ if } \ |m - n| > 2K + 1. \tag{2.42}$$

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

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

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

The same as the construction of $\mathcal{O}_0$ in Subsection §2.4.1, we need not consider the lower bound of $|\mu_n|$ or $|\mu_n \pm \mu_m|$, in view of gauge invariance of P.

Obviously, (2.36) can be solved on $\mathcal{O}_+$. As for solvability of (2.37) – (2.41), let us define the vectors $\tilde{R}^{k10}$, $\tilde{R}^{k01}$ and the matrices $\tilde{R}^{k20}$, $\tilde{R}^{k11}$, $\tilde{R}^{k02}$ as

$$\tilde{R}^{k10} := Q^*R^{k10}, \quad \tilde{R}^{k01} := Q^*R^{k01},$$
$$\tilde{R}^{k20} := Q^*R^{k20}Q, \quad \tilde{R}^{k11} := Q^*R^{k11}Q, \quad \tilde{R}^{k02} := Q^*R^{k02}Q$$

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 are equivalent to

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

for $k \neq 0$, $|m|, |n| \leq K_+$, which can be solved on $\mathcal{O}_+$. Then (2.37) – (2.41) are also solved with

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

By (2.32), it is easy to show that

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

Thus $\bar{F} = F$.

§2.4.3 Property of the coordinate transformation

Lemma 2.3 *F has gauge invariance, and for ε sufficiently small, the coefficients of F satisfy that*

$$|F_{kl00}|_{\mathcal{O}_+} \leq \varepsilon^{\frac{5}{6}}|k|^{2\tau+1}e^{-|k|r}, \tag{2.43}$$

$$|F^{k10}_n|_{\mathcal{O}_+}, |F^{k01}_n|_{\mathcal{O}_+} \leq \varepsilon^{\frac{5}{6}}|k|^{2\tau+1}e^{-|k|r}e^{-\rho|n|}, \tag{2.44}$$

$$|F^{k20}_{mn}|_{\mathcal{O}_+}, |F^{k11}_{mn}|_{\mathcal{O}_+}, |F^{k02}_{mn}|_{\mathcal{O}_+} \leq \varepsilon^{\frac{5}{6}}|k|^{2\tau+1}e^{-|k|r}e^{-\rho\max\{|m|,|n|\}}, \tag{2.45}$$

$\forall\ k \neq 0$, $|l| \leq 1$ *and* $|m|, |n| \leq K_+$.

Proof. Let us first consider F_{mn}^{k20} for instance, with other terms in (2.44) and (2.45) analogous. By the construction above, we can present F_{mn}^{k20} as

$$F_{nm}^{k20} = \mathrm{i}\sum_{\mathcal{F}} \frac{Q_{nn_1}Q_{n_1n_2}^{*}R_{n_2n_3}^{k20}Q_{n_3n_4}Q_{n_4m}^{*}}{\langle k,\omega\rangle - \mu_{n_1} - \mu_{n_4}}, \tag{2.46}$$

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

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

by virtue of the structure of Q in (3.21). Then by (2.31),

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

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 $|\partial_{\xi_j}F_{mn}^{k20}|$, we need to differentiate both sides of (2.39) with respect to ξ_j, $j = 1, 2, \cdots, b$. Then we obtain the equation about $\partial_{\xi_j}F^{k20}$

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

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

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

This equation also can be solved by diagonalizing $\tilde{A}$ via Q Just like (2.46), we get the formulation

$$\partial_{\xi_j}F_{mn}^{k20} = \sum_{\mathcal{F}} \frac{Q_{nn_1}Q_{n_1n_2}^{*}(G_{\xi_j}^{k20})_{n_2n_3}Q_{n_3n_4}Q_{n_4m}^{*}}{\langle k,\omega\rangle - \mu_{n_1} - \mu_{n_4}}.$$

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

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

Thus there exists $c_2 > 0$ such that

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

From the definition of $\mathcal{O}_+$, it is easy to see that

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

Thus, (2.43) – (2.45) hold under the assumption
(C2): $c_2\gamma^{-2}K_+^8K^9e^{(6K+3)\rho}\varepsilon^{\frac{1}{6}}\leq 1.$

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 (2.46), $F^{k20}\equiv 0$. 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 Φ^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 2.4 *For ε sufficiently small, we have $\|X_F\|_{\mathcal{D}_3,\mathcal{O}_+}\leq\varepsilon^{\frac{4}{5}}$.*

Proof. In view of (2.43) – (2.45), 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

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

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 2.4 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 2.5 *For ε sufficiently small, we have $\Phi_F^t : \mathcal{D}_{2\eta} \to \mathcal{D}_{3\eta}$, $-1 \le t \le 1$, and moreover, $\|D\Phi_F^t - Id\|_{\mathcal{D}_{1\eta}} < 2\varepsilon^{\frac{4}{5}}$.*

Proof. Let

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

Notice that F is a polynomial of order 1 in I and of order 2 in z_+, $\bar{z}_+$. It thus follows from Lemma 2.4 and Cauchy inequality that

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

Using the integral equation

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

and Lemma 2.4, one sees easily that $\Phi_F^t : \mathcal{D}_{2\eta} \to \mathcal{D}_{3\eta}$, $-1 \le t \le 1$. 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

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

□

§2.4.4 The new Hamiltonian

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

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

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,

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

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

The new normal form $\mathcal{N}_+$ has properties similar to those of $\mathcal{N}$. 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

$$|\omega_+ - \omega|_{\mathcal{O}_+} \le \varepsilon, \quad |(A_+ - \tilde{A})_{mn}|_{\mathcal{O}_+} \le \varepsilon e^{-\rho \max\{|m|,|n|\}}. \tag{2.47}$$

Let $R(t) = (1-t)(\mathcal{N}_+ - \mathcal{N}) + tR$. Then P_+ can be rewritten as

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

Hence, $X_{P_+} = \int_0^1 (\Phi_F^t)^* X_{\{R(t),F\}} \, dt + (\Phi_F^1)^* X_{(P-R)}$. 由引理2.5，

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

Furthermore, by Lemma A.3，

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

Then, combining with (3.29), $\|X_{P_+}\|_{\mathcal{D}_+, \mathcal{O}_+} \le \varepsilon^{\frac{5}{4}} = \varepsilon_+$.

Note that

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

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 A.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 (2.33), we have

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

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 2.6 *For ε 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

$$\{P, F\}_{\alpha\beta} = \mathrm{i} \sum_{\substack{|n| \le K_+ \\ (\breve{\alpha},\breve{\beta}) + (\hat{\alpha},\hat{\beta}) = (\alpha,\beta)}} \left(P_{\breve{\alpha}+e_n,\breve{\beta}} F_{\hat{\alpha},\hat{\beta}+e_n} - P_{\breve{\alpha},\breve{\beta}+e_n} F_{\hat{\alpha}+e_n,\hat{\beta}} \right) \tag{2.48}$$

$$+ \sum_{(\breve{\alpha},\breve{\beta}) + (\hat{\alpha},\hat{\beta}) = (\alpha,\beta)} \left\{ P_{\breve{\alpha}\breve{\beta}}, F_{\hat{\alpha}\hat{\beta}} \right\}. \tag{2.49}$$

In view of Lemma 2.3, we know that $\|F_{\alpha\beta}\|_{\mathcal{D}_3,\mathcal{O}_+} \leq \varepsilon^{\frac{4}{5}} e^{-\rho n^*_{\alpha\beta}}$.

(1) Terms in (2.48)

Let us first consider the term $P_{\breve{\alpha}+e_n,\breve{\beta}} F_{\hat{\alpha},\hat{\beta}+e_n}$, which contains $\breve{P}_{\breve{\alpha}+e_n,\breve{\beta}} F_{\hat{\alpha},\hat{\beta}+e_n}$ and $\acute{P}_{\breve{\alpha}+e_n,\breve{\beta}} F_{\hat{\alpha},\hat{\beta}+e_n}$. In view of the construction of F, we have that $|\hat{\alpha}| + |\hat{\beta} + e_n| = 1$ or 2.

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

In this case,

$$|\breve{\alpha} + e_n| + |\breve{\beta}| = |\alpha| + |\beta| + 1 - (|\hat{\alpha}| + |\hat{\beta}|) \leq 3.$$

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

$$\begin{aligned} \|\breve{P}_{\breve{\alpha}+e_n,\breve{\beta}} F_{\hat{\alpha},\hat{\beta}+e_n}\|_{\mathcal{D}_3,\mathcal{O}_+},\ \|\acute{P}_{\breve{\alpha}+e_n,\breve{\beta}} F_{\hat{\alpha},\hat{\beta}+e_n}\|_{\mathcal{D}_3,\mathcal{O}_+} &\leq \varepsilon e^{-\rho n^*_{\breve{\alpha}+e_n,\breve{\beta}}} \cdot \varepsilon^{\frac{4}{5}} e^{-\rho n^*_{\hat{\alpha},\hat{\beta}+e_n}} \\ &\leq \varepsilon^{\frac{9}{5}} e^{-\rho n^*_{\alpha\beta}}. \end{aligned} \tag{2.50}$$

- If $|\breve{\alpha} + e_n| + |\breve{\beta}| = 3$, then gauge invariance of P implies $\acute{P}_{\breve{\alpha}+e_n,\breve{\beta}} = 0$. By the construction of F, 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 $(\hat{\alpha}, \hat{\beta}) = (0, 0)$, $(\breve{\alpha}, \breve{\beta}) = (\alpha, \beta)$. 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

$$\|\breve{P}_{\alpha+e_n,\beta} F_{0,e_n}\|_{\mathcal{D}_3,\mathcal{O}_+} \leq cs\varepsilon^{\frac{4}{5}} e^{-\rho n^*_{\alpha\beta}} \leq c\varepsilon^{\frac{9}{5}} e^{-\rho n^*_{\alpha\beta}}. \tag{2.51}$$

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

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

$$\|\breve{P}_{\breve{\alpha}+e_n,\breve{\beta}} F_{\hat{\alpha},\hat{\beta}+e_n}\|_{\mathcal{D}_3,\mathcal{O}_+} \leq e^{-\rho n^*_{\breve{\alpha}+e_n,\breve{\beta}}} \cdot \varepsilon^{\frac{4}{5}} e^{-\rho n^*_{\hat{\alpha},\hat{\beta}+e_n}} \leq \varepsilon^{\frac{4}{5}} e^{-\rho n^*_{\alpha\beta}}, \tag{2.52}$$

$$\|\acute{P}_{\breve{\alpha}+e_n,\breve{\beta}} F_{\hat{\alpha},\hat{\beta}+e_n}\|_{\mathcal{D}_3,\mathcal{O}_+} \leq e^{-\rho(n^+_{\breve{\alpha}+e_n,\breve{\beta}} - n^-_{\breve{\alpha}+e_n,\breve{\beta}})} \cdot \varepsilon^{\frac{4}{5}} e^{-\rho n^*_{\hat{\alpha},\hat{\beta}+e_n}} \leq \varepsilon^{\frac{4}{5}} e^{-\rho n^*_{\alpha\beta}}. \tag{2.53}$$

Doing the same for $P_{\breve{\alpha},\breve{\beta}+e_n} F_{\hat{\alpha}+e_n,\hat{\beta}}$, we finish estimates for terms in (2.48).

(2) Terms in (2.49)

By Lemma A.2 and the inequality $n^*_{\alpha\beta} \le \max\{n^*_{\breve{\alpha}\breve{\beta}}, n^*_{\hat{\alpha}\hat{\beta}}\}$, we have

$$\|\{P_{\breve{\alpha}\breve{\beta}}, F_{\hat{\alpha}\hat{\beta}}\}\|_{\mathcal{D}_{3\eta},\mathcal{O}_+} \le c(r-r_+)^{-1}\eta^{-2}\begin{cases} \varepsilon^{\frac{9}{5}}e^{-\rho n^*_{\alpha\beta}}, & |\alpha|+|\beta|\le 2 \\ \varepsilon^{\frac{4}{5}}e^{-\rho n^*_{\alpha\beta}}, & |\alpha|+|\beta|\ge 3 \end{cases}. \tag{2.54}$$

Combining (2.50) – (2.54), there exists $c_4 > 0$ such that

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

provided that
(C4): $c_4(r-r_+)^{-1}K_+^2\varepsilon^{\frac{1}{20}} \le \frac{1}{4}$,
then Lemma 3.4 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 ε 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, applying the basic facts $\frac{1}{2}(n^+_{\alpha\beta} - n^-_{\alpha\beta}) \le n^*_{\alpha\beta}$ and $\rho_+ < \frac{\rho}{2}$,

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

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

This completes one step of KAM iterations.

§2.5 Proof of Theorem 2.2

Let $r_0,\ s_0,\ \rho_0,\ \varepsilon_0,\ \gamma_0,\ K_0,\ \mathcal{O}_0,\ H_0,\ \mathcal{N}_0,\ P_0$ be as given in Subsection §2.4. 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 $H_\nu = \mathcal{N}_\nu + P_\nu$ on $\mathcal{D}_\nu = \mathcal{D}_{d,\rho_\nu}(r_\nu, s_\nu)$, with

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

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_ν and $\mu_j^\nu = \Omega_j$ for $K_\nu < |j| \le K_{\nu+1}$. Let

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

§2.5.1 Iteration Lemma

The preceding analysis may be summarized in the following

Lemma 2.7 *There exists ε_0 sufficiently small such that the following holds for all $\nu = 0, 1, \cdots$.*

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

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

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

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

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

(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 (**C1**)$-$(**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 (**D1**) clearly holds for $\nu = 0$. Suppose that it holds for some ν. 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 (**D1**) holds for all ν.

As for **(D2)**, let us take ε_0 sufficiently small such that

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

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 ν much faster than the left side. Thus, (**D2**) holds true. □

§2.5.2 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

$$H_0 \circ \Psi^\nu = H_\nu = \mathcal{N}_\nu + P_\nu, \quad \nu = 1, 2, \cdots.$$

Let $\mathcal{O}_\varepsilon = \cap_{\nu=0}^{\infty} \mathcal{O}_\nu$. Using Lemma 2.5 and standard arguments (e.g., [34, 40]), it concludes that H_ν, $\mathcal{N}_\nu$, P_ν and Ψ^ν converge uniformly on $\mathcal{D}_{d,0}(\frac{1}{2}r_0, 0) \times \mathcal{O}_\varepsilon$ to, say, H_∞, $\mathcal{N}_\infty$, P_∞ and Ψ^∞, respectively, in which case it is clear that

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

Since $\varepsilon_\nu = \varepsilon_0^{(\frac{5}{4})^\nu}$, we have, by Lemma A.1, that

$$X_{P^\infty}|_{\mathcal{D}_{d,0}(\frac{1}{2}r_0,0)\times\mathcal{O}_\varepsilon} \equiv 0.$$

Since $H_0 \circ \Psi^\nu = H_\nu$, we have

$$\Phi^t_{H_0} \circ \Psi^\nu = \Psi^\nu \circ \Phi^t_{H_\nu},$$

with $\Phi^t_{H_0}$ denoting the flow of the Hamiltonian vector field X_{H_0}. The uniform convergence of Ψ^ν and X_{H_ν} implies that one can pass the limit in the above and conclude that, on $\mathcal{D}_{d,0}(\frac{1}{2}r_0, 0) \times \mathcal{O}_\varepsilon$,

$$\Phi^t_{H_0} \circ \Psi^\infty = \Psi^\infty \circ \Phi^t_{H_\infty}.$$

Hence, for all $\xi \in \mathcal{O}_\varepsilon$,

$$\Phi^t_{H_0}(\Psi^\infty(\mathbb{T}^b \times \{\xi\})) \;=\; \Psi^\infty \Phi^t_{\mathcal{N}_\infty}(\mathbb{T}^b \times \{\xi\}) \;=\; \Psi^\infty(\mathbb{T}^b \times \{\xi\}).$$

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, $\omega(\xi)$.

§2.5.3 Measure estimates

At the ν^{th} step of KAM iteration, we need to exclude the following resonant parameter set for $k \neq 0$,

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

where

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

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

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 ξ and there exist orthonormal eigenvectors ψ_n^ν corresponding to μ_n^ν, C_W^1 depending on ξ (see e.g. [13]). It follows that $\mu_n^\nu = \langle \tilde{A}_\nu \psi_n^\nu, \bar{\psi}_n^\nu \rangle$ and

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

In view of the construction of $\tilde{A}_\nu$, together with the estimates in (2.47), 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

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

第三章 Localization in one-dimensional nonlinear Schrödinger equation

In this chapter, we consider the lattice Schrödinger equation

$$\mathrm{i}\dot{q}_n = \epsilon(q_{n+1} + q_{n-1}) + V(n\tilde{\alpha} + x)q_n + |q_n|^2 q_n, \quad n \in \mathbb{Z}, \tag{3.1}$$

where $\tilde{\alpha} \in \mathbb{R}$ satisfies the Diophantine condition(1.3), and V is a nonconstant real-analytic function on $\mathbb{R}/\mathbb{Z}$.

§3.1 Statement of the result

Based on the KAM mechanism of Eliasson[16] in Theorem 1.6, we construct an abstract KAM theorem, and apply this theorem to prove well-localization of Equation (3.1) for typical initial data. From the KAM perspective, the main technical challenges in this work are the following:

i) Unlike the model in [19], we need to tackle with the second order perturbation in the Hamiltonian;

ii) Different from the method in [10], our proof is developed from the traditional KAM method;

iii) Compared with the work in the previous chapter, the main difficulty is that the corresponding linear operator has dense point spectrum with *infinitely many resonances.*

Theorem 3.1 *Given an integer $b > 1$, and any $\mathcal{J} = \{n_1, \cdots, n_b\} \subset \mathbb{Z}$. Assume that the support of the initial datum $q_{\mathbb{Z}}(0) = (q_n(0))_{n\in\mathbb{Z}}$ is $\mathcal{J}$ and $q_{\mathbb{Z}}(0) \in [0,1]^b$. There exists a sufficiently small $\epsilon_* = \epsilon_*(V, \tilde{\alpha}, \mathcal{J})$, such that if $0 < \epsilon < \epsilon_*$, then the following holds for a.e. $x \in \mathbb{R}/\mathbb{Z}$.*

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 $\epsilon \to 0$ such that the solution $q(t) = (q_n(t))_{n\in\mathbb{Z}}$ of Equation (3.1), with initial datum $q(0) \in \mathcal{O}_\epsilon$, *satisfies, for any fixed* $d > 0$,

$$\sup_t \sum_{n\in\mathbb{Z}} n^{2d}|q_n(t)|^2 < \infty.$$

Moreover, for each $n \in \mathbb{Z}$, $q_n(t)$ *is quasi-periodic in time.*

Remark 3.1 *The quasi-periodic solutions we obtained are not necessarily small-amplitude, since the nonlinearity* $|q_n|^2 q_n$ *is integrable.*

Remark 3.2 *Smallness assumption on* ϵ *is necessary, otherwise the result is not true even for the linear problem. This is different from the random potential case.*

§3.2 An abstract infinite-dimensional KAM theorem and its application

§3.2.1 Statement of the KAM theorem

We still use the notations and norms in subsection §2.2.1. we consider the perturbed Hamiltonian

$$\begin{aligned} H &= \mathcal{N} + \breve{P} + P \\ &= e(x,\xi) + \langle \omega(x,\xi), I\rangle + \langle \Omega(x,\xi)q, \bar{q}\rangle + \breve{P}(q,\bar{q};x) + P(\theta, I, q, \bar{q}; x, \xi). \quad (3.2) \end{aligned}$$

defined on the domain $\mathcal{D} = \mathcal{D}_{d,\rho}(r,s)$. Our goal is to prove that, for a.e. $x \in \mathbb{R}/\mathbb{Z}$, the Hamiltonian H admits invariant tori for "most" of the parameter $\xi \in \mathcal{O} = \mathcal{O}(x)$, provided that $\|X_{\breve{P}+P}\|_{\mathcal{D},\mathcal{O}}$ is sufficiently small. From now on, we shall not report x for convenience if it is irrelevant.

We need to impose some conditions on ω, Ω, and the perturbations $\breve{P} + P$.

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

(A2) *Regularity of* Ω: $\Omega = T + A + W$.

– T is the symmetric matrix defined in (1.6), independent of ξ. More precisely,

$$T = \operatorname{diag}\{V(x + n\tilde{\alpha})\}_{n\in\mathbb{Z}} + \epsilon\Delta,$$

with V and $\tilde{\alpha}$ as in Equation (1.6).

– A is Hermitian, independent of ξ, satisfying

$$|A_{mn}| \le \begin{cases} c, & |m|, |n| \le \hat{N} \\ 0, & \text{otherwise} \end{cases} \tag{3.3}$$

for some positive $\hat{N}$.

– W is C^1_W parametrized by $\xi \in \mathcal{O}$, with

$$|W_{mn}|_{\mathcal{O}} \le \begin{cases} pe^{-\sigma \max\{|m|, |n|\}}, & |m|, |n| \le N \\ 0, & \text{otherwise} \end{cases} \tag{3.4}$$

for some positive $p \ll 1$, $\sigma \gg \rho$ and sufficiently large N.

Moreover, there exists a subset $\mathcal{J} \subset \mathbb{Z}$ such that

$$\Omega_{mn} \equiv 0 \quad \text{if} \quad m \quad \text{or} \quad n \in \mathcal{J}. \tag{3.5}$$

(A3) *Short range of $\breve{P}$*: $\breve{P}(q, \bar{q}) = \sum_{|\alpha|=|\beta|\ge 2} \breve{P}_{\alpha\beta} q^\alpha \bar{q}^\beta$ is real-analytic in $q, \bar{q}$, and independent of ξ, with

$$|\breve{P}_{\alpha\beta}| \le e^{-\rho(n^+_{\alpha\beta} - n^-_{\alpha\beta})}, \quad |\alpha| = |\beta| \ge 2, \tag{3.6}$$

$$\partial_{q_n} \breve{P} = \partial_{\bar{q}_n} \breve{P} \equiv 0, \quad \forall n \in \mathcal{J}. \tag{3.7}$$

(A4) *Decay property of P*: $P = \sum_{\alpha,\beta} P_{\alpha\beta}(\theta, I; \xi) q^\alpha \bar{q}^\beta$ is real-analytic in $\theta, I, q, \bar{q}$, C^1_W parametrized by $\xi \in \mathcal{O}$, and

$$\|P_{\alpha\beta}\|_{\mathcal{D},\mathcal{O}} \le \begin{cases} \varepsilon e^{-\rho n^*_{\alpha\beta}}, & |\alpha| + |\beta| \le 2 \\ e^{-\rho n^*_{\alpha\beta}}, & |\alpha| + |\beta| \ge 3 \end{cases}, \tag{3.8}$$

$$\partial_{q_n} P = \partial_{\bar{q}_n} P \equiv 0, \quad \forall n \in \mathcal{J}. \tag{3.9}$$

(A5) *Gauge invariance of P*: 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

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

Theorem 3.2 *Consider the Hamiltonian H in (3.2), with* (**A1**)−(**A5**) *satisfied. There is a positive constant*

$$\varepsilon_* = \varepsilon_*(\omega, V, \tilde{\alpha}, \hat{N}, p, \sigma, N, r, s, d, \rho)$$

such that if $\|X_{\breve{P}+P}\|_{\mathcal{D},\mathcal{O}} \le \varepsilon \le \varepsilon_*$*, then for a.e.* $x \in \mathbb{R}/\mathbb{Z}$*, there exists a Cantor set* $\mathcal{O}_\varepsilon = \mathcal{O}_\varepsilon(x) \subset \mathcal{O}(x)$ *with* $|\mathcal{O} \setminus \mathcal{O}_\varepsilon| \to 0$ *as* $\varepsilon \to 0$*, such that the following holds.*

(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* $\Psi : \mathbb{T}^b \times \mathcal{O}_\varepsilon \to \mathcal{D}_{d,0}(r/2, 0)$*, real-analytic in* $\theta \in \mathbb{T}^b$ *and* C^1_W *parametrized by* $\xi \in \mathcal{O}$*, such that* $\|\Psi - \Psi_0\|_{\mathcal{D}_{d,0}(r/2,0),\mathcal{O}_\varepsilon} \to 0$ *as* $\varepsilon \to 0$*, where* Ψ_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 equation of motion associated with the Hamiltonian H.*

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

Remark 3.3 *In case that H satisfies* (**A1**) − (**A5**) *at the first step, all assumptions hold for the Hamiltonian at each KAM step (with suitable parameters).*

§3.2.2 Application to Eq. (3.1)

The Hamiltonian associated with Eq. (3.1) is

$$H = \sum_{n\in\mathbb{Z}} V(x + n\tilde{\alpha}) q_n \bar{q}_n + \epsilon \sum_{n\in\mathbb{Z}} \bar{q}_n (q_{n+1} + q_{n-1}) + \frac{1}{2} \sum_{n\in\mathbb{Z}} |q_n|^4. \tag{3.10}$$

Fix $\mathcal{J} = \{n_1, \cdots, n_b\} \subset \mathbb{Z}$, and $\mathbb{Z}_1 = \mathbb{Z} \setminus \mathcal{J}$. Let $\varepsilon = \epsilon^{\frac{1}{4}}$, with ϵ sufficiently small such that

$$|n_i| \le |\ln \varepsilon| = \frac{1}{4} |\ln \epsilon|, \quad i = 1, \cdots, b.$$

We introduce action-angle variables and amplitude parameters to the Hamiltonian (3.10),

$$q_n = \sqrt{I_n + \xi_n} e^{\mathrm{i}\theta_n}, \quad \bar{q}_n = \sqrt{I_n + \xi_n} e^{-\mathrm{i}\theta_n}, \quad n \in \mathcal{J},$$

where $(I, \theta) = (I_{n_1}, \cdots, I_{n_b}, \theta_{n_1}, \cdots, \theta_{n_b})$ are the standard action-angle variables in the $(q_n, \bar{q}_n)_{n\in\mathcal{J}}$-space around ξ, with $\xi = (\xi_{n_1}, \cdots, \xi_{n_b}) \in \mathcal{O} = [\epsilon^{\frac{1}{12}}, 1] \subset [0, 1]^b$ the amplitude parameter, and $(q,\bar q)=(q_n,\bar q_n)_{n\in\mathbb{Z}_1}$. Then the Hamiltonian (3.10) becomes

$$H=\mathcal{N}(\theta,I,q,\bar q;x,\xi)+\breve P(q,\bar q)+P(\theta,I,q,\bar q;\xi),$$

with

$$\begin{aligned}
\mathcal{N}(\theta,I,q,\bar q;x,\xi) &:= \sum_{n\in\mathcal{J}}(V(x+n\tilde\alpha)\xi_n+\frac{1}{2}\xi_n^2)+\sum_{n\in\mathcal{J}}(V(x+n\tilde\alpha)+\xi_n)I_n\\
&\quad+\sum_{n\in\mathbb{Z}_1}V(x+n\tilde\alpha)|q_n|^2+\epsilon\sum_{\substack{n\in\mathbb{Z}_1\\ n+1\in\mathbb{Z}_1,}}(\bar q_nq_{n+1}+q_n\bar q_{n+1}),\\
\breve P(q,\bar q) &:= \frac{1}{2}\sum_{n\in\mathbb{Z}_1}|q_n|^4,\\
P(\theta,I,q,\bar q;\xi) &:= \frac{1}{2}\sum_{n\in\mathcal{J}}I_n^2+\epsilon\sum_{\substack{m\in\mathcal{J},\,n\in\mathbb{Z}_1\\ |m-n|=1}}\sqrt{I_m+\xi_m}(e^{-\mathrm{i}\theta_m}q_n+e^{\mathrm{i}\theta_m}\bar q_n)\\
&\quad+\epsilon\sum_{\substack{m,\,n\in\mathcal{J}\\ |m-n|=1}}\sqrt{I_m+\xi_m}\sqrt{I_n+\xi_n}(e^{-\mathrm{i}(\theta_m-\theta_n)}+e^{\mathrm{i}(\theta_m-\theta_n)}).
\end{aligned}$$

After introducing the action-angle variables, we find that the structure of the linear operator T in (3.10) has been destroyed. To overcome this disadvantage, we need to add b variables $q'_{n_1},\cdots,q'_{n_b}$ and the corresponding conjugates $\bar q'_{n_1},\cdots,\bar q'_{n_b}$ into this system. For convenience, omit the prime of the newly-added variables and still use q to denote $(q_n)_{n\in\mathbb{Z}}$, since there is no confusion. We then rewrite $\mathcal{N}$ as

$$\begin{aligned}
\mathcal{N} &= \sum_{n\in\mathcal{J}}(V(x+n\tilde\alpha)\xi_n+\frac{1}{2}\xi_n^2)+\sum_{n\in\mathcal{J}}(V(x+n\tilde\alpha)+\xi_n)I_n\\
&\quad+\left[\sum_{n\in\mathbb{Z}_1}V(x+n\tilde\alpha)|q_n|^2+\sum_{n\in\mathcal{J}}V(x+n\tilde\alpha)|q_n|^2\right]+\epsilon\sum_{n\in\mathbb{Z}}(\bar q_nq_{n+1}+q_n\bar q_{n+1})\\
&\quad-\sum_{n\in\mathcal{J}}V(x+n\tilde\alpha)|q_n|^2-\epsilon\sum_{\{n,n+1\}\cap\mathcal{J}\neq\emptyset}(\bar q_nq_{n+1}+q_n\bar q_{n+1})\\
&= e(x,\xi)+\langle\omega(x,\xi),I\rangle+\langle T(x)q,\bar q\rangle+\langle A(x)q,\bar q\rangle,
\end{aligned}$$

with $e(x,\xi) := \sum_{n\in\mathcal{J}} V(x+n\tilde{\alpha})\xi_n + \frac{1}{2}\sum_{n\in\mathcal{J}}\xi_n^2$, and

$$\omega(x,\xi) := (V(x+n_1\tilde{\alpha})+\xi_{n_1},\cdots,V(x+n_b\tilde{\alpha})+\xi_{n_b}), \tag{3.11}$$

$$T_{mn}(x) := \begin{cases} V(x+m\tilde{\alpha}), & m=n \\ \epsilon, & m-n=\pm 1 \\ 0, & \text{otherwise} \end{cases}, \tag{3.12}$$

$$A_{mn}(x) := \begin{cases} -V(x+m\tilde{\alpha}), & m=n, \quad m\in\mathcal{J} \\ -\epsilon, & m-n=\pm 1, \quad m\in\mathcal{J} \\ 0, & \text{otherwise} \end{cases}. \tag{3.13}$$

Now, on some domain $\mathcal{D}_{d,\rho}(r,s)$, the regularity of $\breve{P}+P$ holds true:

Lemma 3.1 *For $\epsilon>0$ sufficiently small and $s=\frac{1}{8}\epsilon^{\frac{1}{4}}$, if $|I|<s^2$ and $\|q\|_{d,\rho}<s$, then*

$$\|X_{\breve{P}+P}\|_{\mathcal{D}_{d,\rho}(r,s),\mathcal{O}} \le \epsilon^{\frac{1}{4}} = \varepsilon.$$

We need to show that the Hamiltonian $H=\mathcal{N}+\breve{P}+P$ satisfies the assumptions (**A1**) − (**A5**) of the KAM theorem, in which **(A3)** and **(A5)** are obviously satisfied.

(**A1**)**:** Since $\{V(x+n\tilde{\alpha})\}_{n\in\mathbb{Z}}$ is independent of ξ, we have that $\frac{\partial\omega}{\partial\xi}\equiv I_{\mathcal{J}}$ in view of (3.11). Thus (**A1**) holds.

(**A2**)**:** Here, $W\equiv 0$. Then, by (3.13), it is evident that (**A2**) holds with $\hat{N}=\frac{1}{4}|\ln\epsilon|$.

(**A4**)**:** Note that terms of P merely correspond to the normal variables q_n, $\bar{q}_n$, $n\notin\mathcal{J}$, $n-1$ or $n+1\in\mathcal{J}$, with the coefficients no more than ϵ, and $\mathcal{J}\subset[-\hat{N},\hat{N}]=[-\frac{1}{4}|\ln\epsilon|,\ \frac{1}{4}|\ln\epsilon|]$. Then, with $\rho\le\frac{1}{6}\hat{N}^{-1}$, (3.8) is verified since

$$c\epsilon^{1-\frac{1}{24}} \le \epsilon^{\frac{1}{4}}e^{-\rho\hat{N}}.$$

Hence, Theorem 3.1 is a corollary of Theorem 3.2.

§3.3 KAM step

To start the KAM iteration for the Hamiltonian (3.2), let $\mathcal{D}_0=\mathcal{D}_{d,\rho_0}(r_0,s_0)$, $\mathcal{O}_0$, H_0, P_0, $\varepsilon_0=\epsilon^{\frac{1}{4}}$, $\mathcal{N}_0$(including e_0, ω_0, W_0, p_0, σ_0, N_0), denote the initial quantities given in the assumptions $(\mathbf{A1})-(\mathbf{A5})$ respectively, and require that ϵ smaller than the ϵ_0 given in Theorem 1.6.

Suppose we have arrived at the ν^{th} step of the KAM iteration, $\nu = 0, 1, 2, \cdots$, recalling that several sequences have been given in (1.10). We consider the Hamiltonian on $\mathcal{D}_\nu := \mathcal{D}_{d,\rho_\nu}(r_\nu, s_\nu)$ and $\mathcal{O}_\nu$,

$$\begin{aligned} H_\nu &= \mathcal{N}_\nu + \breve{P} + P_\nu \\ &= e_\nu + \langle \omega_\nu, I\rangle + \langle \Omega_\nu q, \bar{q}\rangle + \breve{P} + P_\nu, \end{aligned} \tag{3.14}$$

where $\Omega_\nu = T + A + W_\nu$, and $(\mathbf{A1})-(\mathbf{A5})$ are satisfied, including (3.3), (3.6), (3.7) and

$$(\Omega_\nu)_{mn} \equiv 0, \quad \{m, n\} \cap \mathcal{J} \neq \emptyset, \tag{3.15}$$

$$|(W_\nu)_{mn}|_{\mathcal{O}_\nu} \leq \begin{cases} p_\nu e^{-\sigma_\nu \max\{|m|,|n|\}}, & |m|, |n| \leq N_\nu \\ 0, & \text{otherwise} \end{cases}, \tag{3.16}$$

$$\|(P_\nu)_{\alpha\beta}\|_{\mathcal{D}_\nu, \mathcal{O}_\nu} \leq \begin{cases} \varepsilon_\nu e^{-\rho_\nu n^*_{\alpha\beta}}, & |\alpha| + |\beta| \leq 2 \\ e^{-\rho_\nu n^*_{\alpha\beta}}, & |\alpha| + |\beta| \geq 3 \end{cases}, \tag{3.17}$$

$$\partial_{q_n} P_\nu = \partial_{\bar{q}_n} P_\nu \equiv 0, \quad n \in \mathcal{J}. \tag{3.18}$$

Moreover, $\|X_{\breve{P}+P_\nu}\|_{\mathcal{D}_\nu, \mathcal{O}_\nu} \leq \varepsilon_\nu$.

Choose some $r_{\nu+1}$ such that $0 < r_{\nu+1} < r_\nu$, and let $J_\nu := \left[\frac{5}{2}\varepsilon_\nu^{-\frac{a_\nu}{2}}\right]$. For $j = 0, 1, \cdots, J_\nu$, we define the quantities at each KAM sub-step as

$$\rho_\nu^{(j)} = (1 - \frac{j}{2J_\nu})\rho_\nu, \quad r_\nu^{(j)} = r_\nu - \frac{j(r_\nu - r_{\nu+1})}{J_\nu}, \quad s_\nu^{(j)} = 2^{-3j}\varepsilon_\nu^{\frac{j}{5}} s_\nu,$$

and $\mathcal{D}_\nu^{(j)} = \mathcal{D}_{d,\rho_{\nu+1}}(r_\nu^{(j)}, s_\nu^{(j)})$, $\varepsilon_\nu^{(j)} = \varepsilon_\nu^{\frac{j}{5}+1}$, $j = 0, 1, \cdots, J_\nu$. Our goal is to construct a set $\mathcal{O}_{\nu+1} \subset \mathcal{O}_\nu$ and a finite sequence of maps

$$\Phi_\nu^{(j)} : \mathcal{D}_\nu^{(j)} \to \mathcal{D}_\nu^{(j-1)}, \quad j = 1, 2, \cdots, J_\nu,$$

so that the Hamiltonian transformed into the $(\nu+1)^{\text{th}}$ KAM cycle

$$\begin{aligned} H_{\nu+1} &= H_\nu \circ \Phi_\nu^{(1)} \circ \cdots \circ \Phi_\nu^{(J_\nu)} \\ &= \mathcal{N}_{\nu+1} + \breve{P} + P_{\nu+1} \\ &= e_{\nu+1} + \langle \omega_{\nu+1}, I\rangle + \langle \Omega_{\nu+1} q, \bar{q}\rangle + \breve{P} + P_{\nu+1} \end{aligned}$$

satisfies all the above iterative assumptions $(\mathbf{A1})-(\mathbf{A5})$ on $\mathcal{D}_{\nu+1}=\mathcal{D}_{\nu}^{(J_\nu)}$ and C_W^1 parametrized by $\xi\in\mathcal{O}_{\nu+1}$, with new suitable parameters. Moreover,

$$\|X_{\breve{P}+P_{\nu+1}}\|_{\mathcal{D}_{\nu+1},\mathcal{O}_{\nu+1}}\le\varepsilon_\nu^{(J_\nu)}\le\varepsilon_\nu^{\frac{1}{2}\varepsilon_\nu^{-a_\nu/2}}=\varepsilon_{\nu+1}.$$

§3.3.1 Construction of $\mathcal{O}_{\nu+1}$

As described in Theorem 1.6, there exists an orthogonal matrix U_ν with

$$|(U_\nu-I_{\mathbb{Z}})_{mn}|\le\varepsilon_0^{\frac{1}{2}}e^{-\frac{3}{2}\sigma_\nu|m-n|},\tag{3.19}$$

such that $U_\nu^*TU_\nu=D_\nu+Z_\nu$, where Z_ν is a symmetric matrix satisfying

$$|(Z_\nu)_{mn}|\le\varepsilon_\nu e^{-\rho_\nu|m-n|},\tag{3.20}$$

and D_ν is a symmetric matrix which can be block-diagonalized via an orthogonal matrix Q_ν with

$$(Q_\nu)_{mn}=0,\quad |m-n|>N_\nu.\tag{3.21}$$

More precisely,

$$\tilde{D}_\nu=Q_\nu^*D_\nu Q_\nu=\prod_j\tilde{D}^\nu_{\Lambda_j^\nu},\quad \sharp\Lambda_j^\nu\le M_\nu,\ \ \mathrm{diam}\Lambda_j^\nu\le M_\nu N_\nu,\ \ \forall j.$$

To describe $U_\nu^*\Omega_\nu U_\nu$, we need furthermore to consider $U_\nu^*AU_\nu$ and $U_\nu^*W_\nu U_\nu$. In view of (3.3), (3.16) and (3.19), there exists a constant $c_1>0$ such that

$$|\,(U_\nu^*(A+W_\nu)U_\nu)_{mn}\,|_{\mathcal{O}_\nu}\le c_1\max\{\hat{N}^2e^{3\sigma_\nu\hat{N}},\,p_\nu\sigma_\nu^{-2}\}\cdot e^{-\sigma_\nu\cdot\max\{|m|,|n|\}},$$

by a simple application of Lemma B.1. Define the truncation $\hat{A}_\nu$ as

$$(\hat{A}_\nu)_{mn}:=\begin{cases}(U_\nu^*(A+W_\nu)U_\nu)_{mn}\,, & |m|,|n|\le N_\nu\\ \quad 0, & \text{otherwise}\end{cases}.\tag{3.22}$$

It follows that

$$\left|\left(U_\nu^*(A+W_\nu)U_\nu-\hat{A}_\nu\right)_{mn}\right|_{\mathcal{O}_\nu}\le\varepsilon_\nu e^{-\rho_\nu\max\{|m|,|n|\}},\tag{3.23}$$

under the assumption

(C1): $c_1\max\{\hat{N}^2e^{\frac{3}{2}\sigma_\nu\hat{N}},\,p_\nu\sigma_\nu^{-2}\}\cdot e^{-(\sigma_\nu-\rho_\nu)N_\nu}\le\varepsilon_\nu.$

Let $K_{\nu+1} := N_{\nu+1} - (M_\nu + 1)N_\nu$ with the sequences M_ν, N_ν, $\nu = 0, 1, \cdots$, defined in (1.10) and

$$\tilde{D}^\nu_{\Lambda^\nu} := \prod_{\Lambda^\nu_j \subset \Lambda^\nu} \tilde{D}^\nu_{\Lambda^\nu_j}, \quad \tilde{A}_\nu := Q^*_\nu \hat{A}_\nu Q_\nu, \tag{3.24}$$

where $\Lambda^\nu := \bigcup\{\Lambda^\nu_j : \Lambda^\nu_j \cap [-(K_{\nu+1} + N_\nu),\ K_{\nu+1} + N_\nu] \neq \emptyset\} \subset [-N_{\nu+1},\ N_{\nu+1}]$. In view of (3.21) and (3.22), we have

$$(\tilde{A}_\nu)_{mn} \equiv 0, \quad \max\{|m|, |n|\} > 2N_\nu.$$

Since both of $\tilde{D}^\nu_{\Lambda^\nu}$ and $\tilde{A}_\nu$ are Hermitian, there is an orthogonal matrix O_ν such that

$$O^*_\nu(\tilde{D}^\nu_{\Lambda^\nu} + \tilde{A}_\nu)O_\nu = \mathrm{diag}\{\mu^\nu_j\}_{j\in\Lambda^\nu},$$

where $\{\mu^\nu_j\}_{j\in\Lambda^\nu}$ are eigenvalues of $\tilde{D}^\nu_{\Lambda^\nu} + \tilde{A}_\nu$. Due to the block-diagonal structure of $\tilde{D}^\nu_{\Lambda^\nu} + \tilde{A}_\nu$, we also have

$$(O_\nu)_{mn} \equiv 0, \quad |m - n| > 2(M_\nu + 2)N_\nu. \tag{3.25}$$

Indeed, $\tilde{D}^\nu_{\Lambda^\nu} + \tilde{A}_\nu$ can be expressed as

$$\tilde{D}^\nu_{\Lambda^\nu} + \tilde{A}_\nu = (\tilde{D}^\nu_{\Lambda'_\nu} + \tilde{A}_\nu) \cdot \prod_{\Lambda^\nu_j \cap [-2N_\nu, 2N_\nu] = \emptyset} \tilde{D}^\nu_{\Lambda^\nu_j}$$

where $\Lambda'_\nu := \bigcup\{\Lambda^\nu_j : \Lambda^\nu_j \cap [-2N_\nu, 2N_\nu] \neq \emptyset,\ \Lambda^\nu_j \subset \Lambda^\nu\}$, with $\mathrm{diam}\Lambda'_\nu \leq 2(M_\nu + 2)N_\nu$. The diagonalization of $\tilde{D}^\nu_{\Lambda^\nu} + \tilde{A}_\nu$ is just the diagonalization of blocks $(\tilde{D}^\nu_{\Lambda'_\nu} + \tilde{A}_\nu)$ and $\tilde{D}^\nu_{\Lambda^\nu_j}$.

As for the eigenvalues of $\tilde{D}^\nu_{\Lambda^\nu} + \tilde{A}_\nu$, it is well-known that $\{\mu^\nu_n\}_{n\in\Lambda^\nu}$ C^1_W-smoothly depend on ξ and there exist orthonormal eigenvectors ψ^ν_n corresponding to μ^ν_n, C^1_W-smoothly depending on ξ (see e.g. [13]). In fact, $\mu^\nu_n = \langle(\tilde{D}_{\Lambda^\nu} + \tilde{A}_\nu)\psi^\nu_n, \bar{\psi}^\nu_n\rangle$ and

$$\partial_{\xi_j}\mu^\nu_n = \langle(\partial_{\xi_j}(\tilde{D}_{\Lambda^\nu} + \tilde{A}_\nu))\psi^\nu_n, \bar{\psi}^\nu_n\rangle, \quad j = 1, \cdots, b.$$

By the construction of $\tilde{A}_\nu$, we have $\partial_{\xi_j}\tilde{A}_\nu = Q^*_\nu(\partial_{\xi_j}\hat{A}_\nu)Q_\nu$, with $\hat{A}_\nu$ the truncation of $U^*_\nu(A + W_\nu(\xi))U_\nu$. Since D_ν, A, U_ν and Q_ν are all independent of ξ,

$$\sup_{\xi\in\mathcal{O}_\nu} |\partial_\xi \mu^\nu_n| \leq c \sup_{\substack{\xi\in\mathcal{O}_\nu \\ m,n}} |\partial_\xi (W_\nu)_{mn}| \leq cp_\nu. \tag{3.26}$$

We defined the new parameter set $\mathcal{O}_{\nu+1} \subset \mathcal{O}_\nu$ as

$$\mathcal{O}_{\nu+1} := \left\{ \xi \in \mathcal{O}_\nu : \begin{array}{l} |\langle k, \omega_\nu \rangle| > \frac{\gamma_\nu}{|k|^\tau}, \quad k \neq 0, \\ |\langle k, \omega_\nu \rangle + \mu_n^\nu| > \frac{\gamma_\nu}{|k|^\tau N_{\nu+1}^2}, \quad k \neq 0, \;\; n \in \Lambda^\nu, \\ |\langle k, \omega_\nu \rangle + \mu_m^\nu \pm \mu_n^\nu| > \frac{\gamma_\nu}{|k|^\tau N_{\nu+1}^4}, \quad k \neq 0, \;\; m, n \in \Lambda^\nu. \end{array} \right\} \tag{3.27}$$

for some $0 < \gamma_\nu \ll 1$, $\tau \geq b$. These inequalities are famous small-divisor conditions for controlling the solutions of the linearized equations.

From now on, to simplify notations, the subscripts (or superscripts) "ν" of quantities at the ν^{th} step are neglected, and the corresponding quantities at the $(\nu+1)^{\text{th}}$ step are labeled with "+". In addition, we still use the superscript (j) to distinguish quantities at various sub-steps.

§3.3.2 Homological equation and its approximate solution

For $P = \sum_{k,l,\alpha,\beta} P_{kl\alpha\beta}(\xi) I^l e^{\mathrm{i}\langle k,\theta\rangle} q^\alpha \bar{q}^\beta$, according to (3.8) and the definition of norm in subsection §2.2.1, we have

$$|P_{kl\alpha\beta}|_{\mathcal{O}} \leq \varepsilon e^{-\rho n^*_{\alpha\beta}} e^{-|k| r}, \quad \forall k \in \mathbb{Z}^b, \;\; 2|l| + |\alpha| + |\beta| \leq 2. \tag{3.28}$$

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

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

It follows that $\|X_R\|_{\mathcal{D},\mathcal{O}} \leq \|X_P\|_{\mathcal{D},\mathcal{O}} \leq \varepsilon$. Recalling that $\breve{P}(q, \bar{q})$ is a sum of high-order terms, then for $\eta := \varepsilon^{\frac{1}{5}}$, there exists a constant $c_2 > 0$ such that

$$\|X_{\breve{P}}\|_{\mathcal{D}_{d,\rho}(r,\eta s),\mathcal{O}}, \; \|X_{P-R}\|_{\mathcal{D}_{d,\rho}(r,\eta s),\mathcal{O}} \leq c_2 \eta s \leq \frac{1}{8} \varepsilon^{\frac{6}{5}}, \tag{3.29}$$

provided that

(C2): $c_2 s \leq \frac{1}{8}\varepsilon$.

Let $e' := P_{0000}$ and $\omega' := \displaystyle\int \frac{\partial P}{\partial I} d\theta|_{q=\bar{q}=0, I=0}$. With $\mathcal{O}_+$ defined as in (3.27), we have

Proposition 3.1 *There exist two real-analytic Hamiltonians*

$$F = \sum_{\substack{k \neq 0 \\ 1 \leq 2|l|+|\alpha|+|\beta| \leq 2}} F_{kl\alpha\beta} q^\alpha \bar{q}^\beta I^l e^{\mathrm{i}\langle k,\theta\rangle}, \quad \grave{P} = \sum_{\substack{k \\ 1 \leq |\alpha|+|\beta| \leq 2}} \grave{P}_{k0\alpha\beta} q^\alpha \bar{q}^\beta e^{\mathrm{i}\langle k,\theta\rangle},$$

and a Hermitian matrix W', all of which are C^1_W parametrized by $\xi \in \mathcal{O}_+$, such that

$$\{\mathcal{N}, F\} + R = e' + \langle \omega', I\rangle + \langle W'q, \bar{q}\rangle + \grave{P}. \tag{3.30}$$

Moreover, both of F and $\grave{P}$ have gauge invariance, and for ε sufficiently small,

$$|F_{kl\alpha\beta}|_{\mathcal{O}_+} \leq \varepsilon^{\frac{4}{5}}|k|^{2\tau+1}e^{-|k|r}e^{-\rho n^*_{\alpha\beta}}, \tag{3.31}$$

$$|\grave{P}_{k0\alpha\beta}|_{\mathcal{O}_+} \leq \varepsilon^{\frac{7}{5}}|k|^{2\tau+1}e^{-|k|r}e^{-\rho^{(1)} n^*_{\alpha\beta}}, \tag{3.32}$$

$$|W'_{mn}|_{\mathcal{O}_+} \leq \begin{cases} \varepsilon e^{-\rho\max\{|m|,|n|\}}, & |m|,|n| \leq N_+, \quad m,n \notin \mathcal{J} \\ 0, & otherwise \end{cases}, \tag{3.33}$$

$$\partial_{q_n} F = \partial_{\bar{q}_n} F = \partial_{q_n}\grave{P} = \partial_{\bar{q}_n}\grave{P} \equiv 0, \quad n \in \mathcal{J}. \tag{3.34}$$

Proof of Proposition 3.1: We decompose the proof into the following parts.

- **Truncation and approximate linearized equations**

At first, we rewrite R as

$$\begin{aligned} R \;=\; & \sum_{\substack{k \\ |l|\leq 1}} P_{kl00}e^{\mathrm{i}\langle k,\theta\rangle}I^l + \sum_k (\langle P^{k10}, q\rangle + \langle P^{k01}, \bar{q}\rangle)e^{\mathrm{i}\langle k,\theta\rangle} \\ & + \sum_k (\langle P^{k20}q, q\rangle + \langle P^{k11}q, \bar{q}\rangle + \langle P^{k02}\bar{q}, \bar{q}\rangle)e^{\mathrm{i}\langle k,\theta\rangle}, \end{aligned}$$

where P^{k10}, P^{k01}, P^{k20}, P^{k11}, P^{k02} respectively denote

$$\begin{aligned} &\left(P^{k10}_n\right) := \left(P_{k0e_n0}\right), \qquad \left(P^{k01}_n\right) := \left(P_{k00e_n}\right), \\ &\left(P^{k20}_{mn}\right) := \left(P_{kl(e_m+e_n)0}\right), \quad \left(P^{k11}_{mn}\right) := \left(P_{kle_me_n}\right), \quad \left(P^{k02}_{mn}\right) := \left(P_{kl0(e_m+e_n)}\right). \end{aligned}$$

The gauge invariance of P implies that P^{010}, P^{001}, P^{020}, $P^{002} \equiv 0$.

We try to construct a Hamiltonian F, of the same form as R, such that

$$\{\mathcal{N}, F\} + R = e' + \langle \omega', I\rangle + \langle P^{011}q, \bar{q}\rangle. \tag{3.35}$$

By a straightforward calculation and simple comparison of coefficients, Eq. (3.35) is equivalent to the following equations for $k \neq 0$ and $|l| \leq 1$,

$$\langle k, \omega\rangle F_{kl00} = \mathrm{i}P_{kl00}, \tag{3.36}$$

$$(\langle k, \omega\rangle I_{\mathbb{Z}} - \Omega)F^{k10} = \mathrm{i}P^{k10}, \tag{3.37}$$

$$(\langle k, \omega\rangle I_{\mathbb{Z}} + \Omega)F^{k01} = \mathrm{i}P^{k01}, \tag{3.38}$$

$$(\langle k, \omega\rangle I_{\mathbb{Z}} - \Omega)F^{k20} - F^{k20}\Omega = \mathrm{i}P^{k20}, \tag{3.39}$$

$$(\langle k, \omega\rangle I_{\mathbb{Z}} - \Omega)F^{k11} + F^{k11}\Omega = \mathrm{i}P^{k11}, \tag{3.40}$$

$$(\langle k, \omega\rangle I_{\mathbb{Z}} + \Omega)F^{k02} + F^{k02}\Omega = \mathrm{i}P^{k02}. \tag{3.41}$$

In view of the definition of $\mathcal{O}_+$, we know that (3.36) is solved on $\mathcal{O}_+$, with

$$|F_{kl00}|_{\mathcal{O}_+} \leq \gamma^{-2}|k|^{2\tau+1}\varepsilon e^{-|k|r}. \tag{3.42}$$

As for (3.37) – (3.41), consider equations

$$((\langle k,\omega\rangle I_{\mathbb{Z}} - (D+\hat{A}))\hat{F}^{k10} = \mathrm{i}\hat{R}^{k10}, \tag{3.43}$$

$$((\langle k,\omega\rangle I_{\mathbb{Z}} + (D+\hat{A}))\hat{F}^{k01} = \mathrm{i}\hat{R}^{k01}, \tag{3.44}$$

$$((\langle k,\omega\rangle I_{\mathbb{Z}} - (D+\hat{A}))\hat{F}^{k20} - \hat{F}^{k20}(D+\hat{A}) = \mathrm{i}\hat{R}^{k20}, \tag{3.45}$$

$$((\langle k,\omega\rangle I_{\mathbb{Z}} - (D+\hat{A}))\hat{F}^{k11} + \hat{F}^{k11}(D+\hat{A}) = \mathrm{i}\hat{R}^{k11}, \tag{3.46}$$

$$((\langle k,\omega\rangle I_{\mathbb{Z}} + (D+\hat{A}))\hat{F}^{k02} + \hat{F}^{k02}(D+\hat{A}) = \mathrm{i}\hat{R}^{k02}, \tag{3.47}$$

instead, where D and $\hat{A}$ are defined in the previous subsection, and for $k \neq 0$,

$$\hat{R}^{kx}_n = \begin{cases} (U^*P^{kx})_n, & |n| \leq K_+ \\ 0, & \text{otherwise} \end{cases}, \quad x = \text{“10”},\ \text{“01”}, \tag{3.48}$$

$$\hat{R}^{kx}_{mn} = \begin{cases} (U^*P^{kx}U)_{mn}, & |m|,|n| \leq K_+ \\ 0, & \text{otherwise} \end{cases}, \quad x = \text{“20”},\ \text{“11”},\ \text{“02”}. \tag{3.49}$$

By (3.19) and (3.28), combining with Lemma B.1, there exists $c_3 > 0$ such that

$$|(U^*P^{kx})_n|_{\mathcal{O}} \leq c_3(\sigma-\rho)^{-1}\varepsilon e^{-\rho|n|}e^{-|k|r}, \tag{3.50}$$

$$|(U^*P^{kx}U)_{mn}|_{\mathcal{O}} \leq c_3(\sigma-\rho)^{-2}\varepsilon e^{-\rho\max\{|m|,|n|\}}e^{-|k|r}. \tag{3.51}$$

This means

$$|(U^*P^{kx} - \hat{R}^{kx})_n|_{\mathcal{O}} \leq \frac{1}{4}\varepsilon^{\frac{7}{5}}e^{-\rho^{(1)}|n|}e^{-|k|r}, \tag{3.52}$$

$$|(U^*P^{kx}U - \hat{R}^{kx})_{mn}|_{\mathcal{O}} \leq \frac{1}{4}\varepsilon^{\frac{7}{5}}e^{-\rho^{(1)}\max\{|m|,|n|\}}e^{-|k|r}, \tag{3.53}$$

under the assumption that
(C3): $c_3(\sigma-\rho)^{-4}e^{-(\rho-\rho^{(1)})K_+} \leq \frac{1}{4}\varepsilon^{\frac{2}{5}}$.

Equations (3.43) – (3.47) provide us with approximate solutions to (3.37) – (3.41), with the error estimated later.

- **Block-diagonalization and construction of F**

Consider the equations

$$(\langle k,\omega\rangle I_\Lambda - (\tilde{D}_\Lambda + \tilde{A}))\tilde{F}^{k10} = \mathrm{i}\tilde{R}^{k10}, \tag{3.54}$$

$$(\langle k,\omega\rangle I_\Lambda + (\tilde{D}_\Lambda + \tilde{A}))\tilde{F}^{k01} = \mathrm{i}\tilde{R}^{k01}, \tag{3.55}$$

$$(\langle k,\omega\rangle I_\Lambda - (\tilde{D}_\Lambda + \tilde{A}))\tilde{F}^{k20} - \tilde{F}^{k20}(\tilde{D}_\Lambda + \tilde{A}) = \mathrm{i}\tilde{R}^{k20}, \tag{3.56}$$

$$(\langle k,\omega\rangle I_\Lambda - (\tilde{D}_\Lambda + \tilde{A}))\tilde{F}^{k11} + \tilde{F}^{k11}(\tilde{D}_\Lambda + \tilde{A}) = \mathrm{i}\tilde{R}^{k11}, \tag{3.57}$$

$$(\langle k,\omega\rangle I_\Lambda + (\tilde{D}_\Lambda + \tilde{A}))\tilde{F}^{k02} + \tilde{F}^{k02}(\tilde{D}_\Lambda + \tilde{A}) = \mathrm{i}\tilde{R}^{k02}, \tag{3.58}$$

where $\tilde{D}_\Lambda$, $\tilde{A}$ are defined as in (3.24) via the orthogonal matrix Q, and

$$\tilde{R}^{kx} := \begin{cases} Q^*\hat{R}^{kx}, & x = \text{“10”},\ \text{“01”} \\ Q^*\hat{R}^{kx}Q, & x = \text{“20”},\ \text{“11”},\ \text{“02”} \end{cases}.$$

Note that $Q_{mn} = 0$ if $|m-n| > N$, then by (3.48) and (3.49), we have

$$\begin{aligned} &\tilde{R}^{kx}_n \equiv 0, \quad |n| > K_+ + N, \quad x = \text{“10”},\ \text{“01”}, \\ &\tilde{R}^{kx}_{mn} \equiv 0, \quad \max\{|m|,|n|\} > K_+ + N, \quad x = \text{“20”},\ \text{“11”},\ \text{“02”}. \end{aligned}$$

Thus, recalling that $\Lambda := \bigcup\{\Lambda_j : \Lambda_j \cap [-(K_+ + N),\ K_+ + N] \neq \emptyset\}$, solutions of these finite-dimensional equations satisfy

$$\begin{aligned} &\tilde{F}^{kx}_n \equiv 0, \quad n \notin \Lambda, \quad x = \text{“10”},\ \text{“01”}, \\ &\tilde{F}^{kx}_{mn} \equiv 0, \quad \{m,n\} \cap \Lambda = \emptyset, \quad x = \text{“20”},\ \text{“11”},\ \text{“02”}. \end{aligned}$$

Then, in view of the facts

$$\begin{aligned} &\left(\langle k,\omega\rangle I_{\mathbb{Z}} \pm (\tilde{D} + \tilde{A})\right)\tilde{F}^{kx} = \left(\langle k,\omega\rangle I_\Lambda \pm (\tilde{D}_\Lambda + \tilde{A})\right)\tilde{F}^{kx}, \quad x = \text{“10”},\ \text{“01”}, \\ &\left(\langle k,\omega\rangle I_{\mathbb{Z}} \pm (\tilde{D} + \tilde{A})\right)\tilde{F}^{kx} = \left(\langle k,\omega\rangle I_\Lambda \pm (\tilde{D}_\Lambda + \tilde{A})\right)\tilde{F}^{kx}, \quad x = \text{“20”},\ \text{“11”},\ \text{“02”}, \\ &\tilde{F}^{kx}(\tilde{D} + \tilde{A}) = \tilde{F}^{kx}(\tilde{D}_\Lambda + \tilde{A}), \quad x = \text{“20”},\ \text{“11”},\ \text{“02”}, \end{aligned}$$

they are also solutions of

$$\begin{aligned} &\left(\langle k,\omega\rangle I_{\mathbb{Z}} - (\tilde{D} + \tilde{A})\right)\tilde{F}^{k10} = \mathrm{i}\tilde{R}^{k10}, \\ &\left(\langle k,\omega\rangle I_{\mathbb{Z}} + (\tilde{D} + \tilde{A})\right)\tilde{F}^{k01} = \mathrm{i}\tilde{R}^{k01}, \\ &\left(\langle k,\omega\rangle I_{\mathbb{Z}} - (\tilde{D} + \tilde{A})\right)\tilde{F}^{k20} - \tilde{F}^{k20}(\tilde{D} + \tilde{A}) = \mathrm{i}\tilde{R}^{k20}, \\ &\left(\langle k,\omega\rangle I_{\mathbb{Z}} - (\tilde{D} + \tilde{A})\right)\tilde{F}^{k11} + \tilde{F}^{k11}(\tilde{D} + \tilde{A}) = \mathrm{i}\tilde{R}^{k11}, \\ &\left(\langle k,\omega\rangle I_{\mathbb{Z}} + (\tilde{D} + \tilde{A})\right)\tilde{F}^{k02} + \tilde{F}^{k02}(\tilde{D} + \tilde{A}) = \mathrm{i}\tilde{R}^{k02}, \end{aligned}$$

which are respectively equivalent to Equation (3.43) – (3.47) since D can be block-diagonalized by the orthogonal matrix Q.

Now we focus on the following equations

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

for $k\neq 0$ and $m,n\in\Lambda$, which is transformed from (3.54) – (3.58) by diagonalizing $\tilde{D}_\Lambda+\tilde{A}$ via the orthogonal matrix O. Obviously, these equations can be solved in $\mathcal{O}_+$. Hence, (3.43) – (3.47) are solved with

$$\hat{F}^{kx}=\begin{cases} QO\check{F}^{kx}, & x=\text{“10”},\ \text{“01”}\\ QO\check{F}^{kx}O^*Q^*, & x=\text{“20”},\ \text{“11”},\ \text{“02”}\end{cases}.$$

Let

$$F^{kx}:=\begin{cases} U\hat{F}^{kx}, & x=\text{“10”},\ \text{“01”}\\ U\hat{F}^{kx}U^*, & x=\text{“20”},\ \text{“11”},\ \text{“02”}\end{cases},$$

then we obtain a Hamiltonian

$$\begin{aligned}
F &= \sum_{\substack{k\neq 0\\ |l|\le 1}} F_{kl00}e^{\mathrm{i}\langle k,\theta\rangle}I^l+\sum_{k\neq 0}(\langle F^{k10},q\rangle+\langle F^{k01},\bar{q}\rangle)e^{\mathrm{i}\langle k,\theta\rangle}\\
&\quad +\sum_{k\neq 0}(\langle F^{k20}q,q\rangle+\langle F^{k11}q,\bar{q}\rangle+\langle F^{k02}\bar{q},\bar{q}\rangle)e^{\mathrm{i}\langle k,\theta\rangle}.
\end{aligned}$$

It is easy to see that $\bar{F}=F$, by noting

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

- **Estimates for coefficients of F**

Apart from F_{kl00} which has been estimate in (3.42), we still need to handle F^{k10}_n, F^{k01}_n, F^{k20}_{mn}, F^{k11}_{mn}, F^{k02}_{mn}. Let us consider F^{k20}_{mn} for instance, and the other terms can be treated in an analogous way.

By the construction above, one sees that

$$F_{mn}^{k20} = \mathrm{i}\sum_{\mathcal{F}_0} \frac{U_{mn_1}Q_{n_1n_2}O_{n_2n_3}O^*_{n_3n_4}Q^*_{n_4n_5}\hat{R}^{k20}_{n_5n_6}Q_{n_6n_7}O_{n_7n_8}O^*_{n_8n_9}Q^*_{n_9n_{10}}U^*_{n_{10}n}}{\langle k,\omega\rangle - \mu_{n_3} - \mu_{n_8}}, \qquad (3.59)$$

where the summation notation $\mathcal{F}_0$ denotes

$$\left\{ \begin{array}{c} n_1 \in \mathbb{Z}, \quad |n_2 - n_1| \le N, \quad |n_3 - n_2|, |n_4 - n_3| \le 2(M+2)N, \quad |n_5 - n_4| \le N, \\ n_{10} \in \mathbb{Z}, \quad |n_9 - n_{10}| \le N, \quad |n_8 - n_9|, |n_7 - n_8| \le 2(M+2)N, \quad |n_6 - n_7| \le N \end{array} \right\}$$

by virtue of the structure of Q and O, i.e, (3.21) and (3.25). Then, by (3.51) and Lemma B.1,

$$\sup_{\xi\in\mathcal{O}_+} |F^{k20}_{mn}(\xi)| \le c(\gamma^{-1}|k|^\tau N_+^4)(\sigma-\rho)^{-4}M^4N^8 e^{(4M+10)N\rho}\varepsilon e^{-|k|r}e^{-\rho\max\{|m|,|n|\}}.$$

Here we have applied the property of the orthogonal matrices Q and O, and used the factor $e^{(4M+10)N\rho}$ to recover the exponential decay.

To estimate $|\partial_{\xi_j}F^{k20}_{mn}|$, we need to differentiate both sides of (3.56) with respect to ξ_j, $j = 1, 2, \cdots, b$. Then we obtain the equation about $\partial_{\xi_j}\tilde{F}^{k20}$:

$$(\langle k,\omega\rangle I_\Lambda - (\tilde{D}_\Lambda + \tilde{A}))(\partial_{\xi_j}\tilde{F}^{k20}) - (\partial_{\xi_j}\tilde{F}^{k20})(\tilde{D}_\Lambda + \tilde{A}) = \breve{R}^{k20}_{\xi_j},$$

which can also be solved by diagonalizing $\tilde{D}_\Lambda + \tilde{A}$ via O as above, where

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

We get the formulation

$$\partial_{\xi_j}F^{k20}_{mn} = \sum_{\mathcal{F}_1} \frac{U_{mn_1}Q_{n_1n_2}O_{n_2n_3}O^*_{n_3n_4}(\breve{R}^{k20}_{\xi_j})_{n_4n_5}O_{n_5n_6}O^*_{n_6n_7}Q^*_{n_7n_8}U^*_{n_8n}}{\langle k,\omega\rangle - \mu_{n_3} - \mu_{n_6}},$$

with $\mathcal{F}_1$ denotes

$$\left\{ \begin{array}{c} n_1 \in \mathbb{Z}, \quad |n_2 - n_1| \le N, \quad |n_3 - n_2|, |n_4 - n_3| \le 2(M+2)N, \\ n_8 \in \mathbb{Z}, \quad |n_7 - n_8| \le N, \quad |n_6 - n_7|, |n_5 - n_6| \le 2(M+2)N \end{array} \right\}.$$

By the decay property of $\hat{R}^{k20}$ and $\partial_{\xi_j}\hat{A}$, we have that

$$\sup_{\xi\in\mathcal{O}_+} |(\breve{R}^{k20}_{\xi_j})_{mn}| \le c(\gamma^{-1}|k|^{\tau+1} N_+^4)(\sigma-\rho)^{-4}M^4N^8 e^{(4M+11)N\rho}\varepsilon e^{-|k|r}e^{-\rho\max\{|m|,|n|\}}.$$

Thus there exists $c_4 > 0$ such that

$$\begin{aligned}
&\sup_{\xi\in\mathcal{O}_+}(|F^{k20}_{mn}| + |\partial_\xi F^{k20}_{mn}|) \\
\leq\ & c_4(\gamma^{-2}|k|^{2\tau+1}N^8_+)(\sigma-\rho)^{-6}M^8N^{14}e^{(8M+20)N\rho}\varepsilon e^{-\rho\max\{|m|,|n|\}}e^{-|k|r} \\
\leq\ & \varepsilon^{\frac{4}{5}}|k|^{2\tau+1}e^{-|k|r}e^{-\rho\max\{|m|,|n|\}},
\end{aligned}$$

under the assumption
(C4): $c_4\gamma^{-2}(\sigma-\rho)^{-6}N^8_+M^8N^{14}e^{(8M+20)N\rho}\varepsilon^{\frac{1}{5}} \leq 1.$

Suppose that $\sum_{i=1}^b k_i + 2 \neq 0$, which means $P^{k20} \equiv 0$. Then $\hat{R}^{k20} \equiv 0$, since it is a truncation of $U^*P^{k20}U$. By the formulation of F^{k20}_{mn} in (3.59), $F^{k20} \equiv 0$.

From **(A5)** we see

$$P^{k20} \equiv 0, \quad \sum_{i=1}^b k_i + 2 \neq 0.$$

Since $\hat{R}^{k20}$ is a truncation of $U^*P^{k20}U$, we have

$$\hat{R}^{k20} \equiv 0, \quad \sum_{i=1}^b k_i + 2 \neq 0.$$

By the formulation of F^{k20}_{mn} in (3.59), $F^{k20} \equiv 0$.

Doing the same thing for F^{k11}, F^{k02}, F^{k10}, F^{k01} as above, we obtain the gauge invariance of F and the inequality (3.31).

- **Estimates for coefficients of $\grave{P}$**

Let W' be the truncation of P^{011}, satisfying

$$W'_{mn} = \begin{cases} P^{011}_{mn}, & |m|,|n| \leq N_+ \\ 0, & \text{otherwise} \end{cases},$$

and

$$\grave{P} = \langle \grave{P}^{011}q,\bar{q}\rangle + \sum_{k\neq 0}(\langle \grave{P}^{k10},q\rangle + \langle \grave{P}^{k01},\bar{q}\rangle + \langle \grave{P}^{k20}q,q\rangle + \langle \grave{P}^{k11}q,\bar{q}\rangle + \langle \grave{P}^{k02}\bar{q},\bar{q}\rangle)e^{\mathrm{i}\langle k,\theta\rangle}$$

with

$$\begin{aligned}
\grave{P}^{011} &:= P^{011} - W',\\
\grave{P}^{k10} &:= (P^{k10} - U\hat{R}^{k10}) - \mathrm{i}(\grave{A} + \grave{Z})F^{k10},\\
\grave{P}^{k01} &:= (P^{k01} - U\hat{R}^{k01}) + \mathrm{i}(\grave{A} + \grave{Z})F^{k01},\\
\grave{P}^{k20} &:= (P^{k20} - U\hat{R}^{k20}U^*) - \mathrm{i}(\grave{A} + \grave{Z})F^{k20} - \mathrm{i}F^{k20}(\grave{A} + \grave{Z}),\\
\grave{P}^{k11} &:= (P^{k11} - U\hat{R}^{k11}U^*) - \mathrm{i}(\grave{A} + \grave{Z})F^{k11} + \mathrm{i}F^{k11}(\grave{A} + \grave{Z}),\\
\grave{P}^{k02} &:= (P^{k02} - U\hat{R}^{k02}U^*) + \mathrm{i}(\grave{A} + \grave{Z})F^{k02} + \mathrm{i}F^{k02}(\grave{A} + \grave{Z}),
\end{aligned}$$

where $\grave{A} := (A + W) - U\hat{A}U^*$, $\grave{Z} := UZU^*$. Then we obtain

$$\{\mathcal{N}, F\} + R = e' + \langle\omega', I\rangle + \langle W'q, \bar{q}\rangle + \grave{P}. \tag{3.60}$$

By (3.17) and (3.18), we have (3.33) holds and

$$|\grave{P}^{011}_{mn}|_{\mathcal{O}_+} \le \varepsilon e^{-\rho\max\{|m|,\,|n|\}} \le \varepsilon^{\frac{7}{5}} e^{-\rho^{(1)}\max\{|m|,\,|n|\}},$$

under the assumption
(C5): $e^{-(\rho-\rho^{(1)})N_+} \le \varepsilon^{\frac{2}{5}}$.

As for the case $k \neq 0$ in (3.32), we only estimate $\grave{P}^{k20}$, with the others entirely analogous. By (3.53) and **(C3)**, combining with Lemma B.1, we have

$$\begin{aligned}
\left|\left(P^{k20} - U\hat{R}^{k20}U^*\right)_{mn}\right|_{\mathcal{O}} &= \left|\left(U(U^*P^{k20}U - \hat{R}^{k20})U^*\right)_{mn}\right|_{\mathcal{O}}\\
&\le \frac{1}{4}\varepsilon^{\frac{7}{5}} e^{-\rho^{(1)}\max\{|m|,\,|n|\}} e^{-|k|r}.
\end{aligned} \tag{3.61}$$

In view of (3.20) and (3.23),

$$|\grave{A}_{mn}|_{\mathcal{O}} \le c(\sigma-\rho)^{-2}\varepsilon e^{-\rho\max\{|m|,\,|n|\}}, \quad |\grave{Z}_{mn}| \le c(\sigma-\rho)^{-2}\varepsilon e^{-\rho|m-n|}.$$

Then, by applying Lemma B.1 again, there exists $c_6 > 0$ such that

$$\begin{aligned}
&\left|\left(F^{k20}(\grave{A} + \grave{Z})\right)_{mn}\right|_{\mathcal{O}_+}, \quad \left|\left((\grave{A} + \grave{Z})F^{k20}\right)_{mn}\right|_{\mathcal{O}_+}\\
\le\ & c_6(\sigma-\rho)^{-2}(\rho-\rho^{(1)})^{-1}\varepsilon^{\frac{9}{5}}|k|^{2\tau+1}e^{-|k|r}e^{-\rho^{(1)}\max\{|m|,\,|n|\}}\\
\le\ & \frac{1}{4}\varepsilon^{\frac{7}{5}}|k|^{2\tau+1}e^{-|k|r}e^{-\rho^{(1)}\max\{|m|,\,|n|\}},
\end{aligned} \tag{3.62}$$

provided that
(C6): $c_6(\sigma-\rho)^{-2}(\rho-\rho^{(1)})^{-1}\varepsilon^{\frac{2}{5}} \le \frac{1}{4}$.

Thus, we can obtain the estimate for $\grave{P}^{k20}$ by putting (3.61) and (3.62) together.

By the construction of $\grave{P}$, the gauge invariance is easily verified.

- **Verification of (3.34)**

In view of the construction of R and W' above, the objects in (3.60) that may depend on the variables $(q_n, \bar{q}_n)_{n\in\mathcal{J}}$ are F and $\grave{P}$. Let

$$
\begin{aligned}
\acute{F} &= \sum_{k\neq 0}\left(\sum_{n\in\mathcal{J}}(F_n^{k10}q_n + F_n^{k01}\bar{q}_n)\right)e^{\mathrm{i}\langle k,\theta\rangle} \\
&\quad + \sum_{k\neq 0}\left(\sum_{\{m,n\}\cap\mathcal{J}\neq\emptyset}(F_{mn}^{k20}q_mq_n + F_{mn}^{k11}q_m\bar{q}_n + F_{mn}^{k02}\bar{q}_m\bar{q}_n)\right)e^{\mathrm{i}\langle k,\theta\rangle} \\
&=: \sum_{k\neq 0}\left(\langle \acute{F}^{k10}, q\rangle + \langle \acute{F}^{k01}, \bar{q}\rangle + \langle \acute{F}^{k20}q, q\rangle + \langle \acute{F}^{k11}q, \bar{q}\rangle + \langle \acute{F}^{k02}\bar{q}, \bar{q}\rangle\right)e^{\mathrm{i}\langle k,\theta\rangle}.
\end{aligned}
$$

For m or $n \in \mathcal{J}$, by (3.15), we have

$$
\begin{aligned}
&\left((\langle k,\omega\rangle I_{\mathbb{Z}} - \Omega)\acute{F}^{k20} - \acute{F}^{k20}\Omega\right)_{mn} \\
= &\ \langle k,\omega\rangle \acute{F}_{mn}^{k20} - \sum_{l\notin\mathcal{J}}\Omega_{ml}\acute{F}_{ln}^{k20} - \sum_{l\notin\mathcal{J}}\acute{F}_{ml}^{k20}\Omega_{ln} \\
= &\ \begin{cases} \langle k,\omega\rangle F_{mn}^{k20}, & m,n\in\mathcal{J} \\ \langle k,\omega\rangle F_{mn}^{k20} - \sum_{l\notin\mathcal{J}}\Omega_{ml}F_{ln}^{k20}, & m\notin\mathcal{J},\quad n\in\mathcal{J} \\ \langle k,\omega\rangle F_{mn}^{k20} - \sum_{l\notin\mathcal{J}}F_{ml}^{k20}\Omega_{ln}, & m\in\mathcal{J},\quad n\notin\mathcal{J} \end{cases} \\
= &\ \left((\langle k,\omega\rangle I_{\mathbb{Z}} - \Omega)F^{k20} - F^{k20}\Omega\right)_{mn}.
\end{aligned}
$$

This means, by comparing the coefficients in both side of Equation (3.60),

$$
\left((\langle k,\omega\rangle I_{\mathbb{Z}} - \Omega)\acute{F}^{k20} - \acute{F}^{k20}\Omega\right)_{mn} = -\mathrm{i}\grave{P}_{mn}^{k20}, \quad \{m,n\}\cap\mathcal{J}\neq\emptyset.
$$

Similarly,

$$
\begin{aligned}
&\left((\langle k,\omega\rangle I_{\mathbb{Z}} - \Omega)\acute{F}^{k10}\right)_n = -\mathrm{i}\grave{P}_n^{k10}, \quad n\in\mathcal{J}, \\
&\left((\langle k,\omega\rangle I_{\mathbb{Z}} + \Omega)\acute{F}^{k01}\right)_n = -\mathrm{i}\grave{P}_n^{k01}, \quad n\in\mathcal{J}, \\
&\left((\langle k,\omega\rangle I_{\mathbb{Z}} - \Omega)\acute{F}^{k11} + \acute{F}^{k11}\Omega\right)_{mn} = -\mathrm{i}\grave{P}_{mn}^{k11}, \quad \{m,n\}\cap\mathcal{J}\neq\emptyset, \\
&\left((\langle k,\omega\rangle I_{\mathbb{Z}} + \Omega)\acute{F}^{k02} + \acute{F}^{k02}\Omega\right)_{mn} = -\mathrm{i}\grave{P}_{mn}^{k02}, \quad \{m,n\}\cap\mathcal{J}\neq\emptyset.
\end{aligned}
$$

Thus, $\{\mathcal{N}, \acute{F}\}$ equals to

$$\sum_{k\neq 0}\left(\sum_{n\in\mathcal{J}}(\grave{P}^{k10}_{n} q_n + \grave{P}^{k01}_{n}\bar{q}_n) + \sum_{\{m,n\}\cap\mathcal{J}\neq\emptyset}(\grave{P}^{k20}_{mn} q_m q_n + \grave{P}^{k11}_{mn} q_m\bar{q}_n + \grave{P}^{k02}_{mn}\bar{q}_m\bar{q}_n)\right) e^{\mathrm{i}\langle k,\theta\rangle}.$$

Hence, if we substitute F with $F-\acute{F}$, which is independent of the variables $(q_n, \bar{q}_n)_{n\in\mathcal{J}}$, then the system will keep independent of the variables $(q_n, \bar{q}_n)_{n\in\mathcal{J}}$. (3.34) is satisfied.□

§3.3.3 Verification of assumptions after one sub-step

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

Lemma 3.2 *For ε sufficiently small, we have $\|X_F\|_{\mathcal{D}_3,\mathcal{O}_+} \leq \varepsilon^{\frac{3}{4}}$ and $\|X_{\grave{P}}\|_{\mathcal{D}_3,\mathcal{O}_+} \leq \varepsilon^{\frac{5}{4}}$.*

Proof. In view of the decay property of F in Proposition 3.1, it follows that

$$\frac{1}{s^2}\|\partial_\theta F\|_{\mathcal{D}_3,\mathcal{O}_+},\ \|\partial_I F\|_{\mathcal{D}_3,\mathcal{O}_+} \leq c(r - r^{(1)})^{-(2\tau+b+1)}\varepsilon^{\frac{4}{5}},$$

and

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

Putting together the estimates above, there is a constant $c_7 > 0$ such that

$$\|X_F\|_{\mathcal{D}_3,\mathcal{O}_+} \leq c_7(r - r^{(1)})^{-(2\tau+b+1)}(\rho - \rho_+)^{-2}\varepsilon^{\frac{4}{5}}.$$

In an entirely analogous way, we have

$$\|X_{\grave{P}}\|_{\mathcal{D}_3,\mathcal{O}_+} \leq c_7(r - r^{(1)})^{-(2\tau+b+1)}(\rho^{(1)} - \rho_+)^{-2}\varepsilon^{\frac{7}{5}}.$$

Moreover, if
(C7): $c_7(r - r^{(1)})^{-(2\tau+b+1)}(\rho^{(1)} - \rho_+)^{-2}\varepsilon^{\frac{1}{20}} \leq \frac{1}{3}$,
then Lemma 3.2 follows. □

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

Lemma 3.3 *For ε sufficiently small, we have $\Phi_F^t : \mathcal{D}_{2\eta} \to \mathcal{D}_{3\eta}$, $-1 \le t \le 1$ and*

$$\|D\Phi_F^t - I\|_{\mathcal{D}_{1\eta}} < 2\varepsilon^{\frac{3}{4}}.$$

Let $F^{(1)}$, $e^{(1)}$, $\omega^{(1)}$, $W^{(1)}$, $\grave{P}^{(1)}$ be the corresponding quantities in (3.30) respectively, which means that we are in the 1st sub-step. Define $H^{(1)}$ as

$$\begin{aligned}
H^{(1)} &:= H \circ \Phi^1_{F^{(1)}} \\
&= (\mathcal{N} + \breve{P} + R) \circ \Phi^1_{F^{(1)}} + (P - R) \circ \Phi^1_{F^{(1)}} \\
&= \mathcal{N} + \breve{P} + \{\mathcal{N}, F^{(1)}\} + R + \int_0^1 (1-t)\{\{\mathcal{N}, F^{(1)}\}, F^{(1)}\} \circ \Phi^t_{F^{(1)}}\, dt \\
&\quad + \int_0^1 \{\breve{P} + R, F^{(1)}\} \circ \Phi^t_{F^{(1)}}\, dt + (P - R) \circ \Phi^1_{F^{(1)}} \\
&= \mathcal{N} + \breve{P} + e^{(1)} + \langle \omega^{(1)}, I\rangle + \langle W^{(1)} q, \bar{q}\rangle + P^{(1)},
\end{aligned}$$

where

$$P^{(1)} := \grave{P}^{(1)} + \int_0^1 \{(1-t)\{\mathcal{N}, F^{(1)}\} + \breve{P} + R, F^{(1)}\} \circ \Phi^t_{F^{(1)}}\, dt + (P - R) \circ \Phi^1_{F^{(1)}}.$$

Let $R(t) := (1-t)(e^{(1)} + \langle \omega^{(1)}, I\rangle + \langle W^{(1)} q, \bar{q}\rangle + \grave{P}^{(1)}) + tR$, which satisfies $\|X_{R(t)}\|_{\mathcal{D}_3} \le c\varepsilon$. Then $P^{(1)}$ can be written as

$$P^{(1)} = \grave{P}^{(1)} + \int_0^1 \{R(t) + \breve{P},\ F^{(1)}\} \circ \Phi^t_{F^{(1)}}\, dt + (P - R) \circ \Phi^1_{F^{(1)}}.$$

Hence,

$$X_{P^{(1)} - \grave{P}^{(1)}} \;=\; \int_0^1 (\Phi^t_{F^{(1)}})^* X_{\{R(t) + \breve{P},\, F^{(1)}\}}\, dt \;+\; (\Phi^1_{F^{(1)}})^* X_{(P-R)}.$$

By Lemma A.3,

$$\|X_{\{R(t) + \breve{P},\, F^{(1)}\}}\|_{\mathcal{D}_{2\eta}} \le c\eta^{-2}\varepsilon^{\frac{7}{4}} = \varepsilon^{\frac{27}{20}}.$$

Then, combining with (3.29), recalling the conclusion of Lemma 3.2 and 3.3,

$$\|X_{P^{(1)}}\|_{\mathcal{D}^{(1)}, \mathcal{O}_+} \le \frac{1}{2}\varepsilon^{\frac{6}{5}} + 2\varepsilon^{\frac{5}{4}} + 2c\varepsilon^{\frac{27}{20}} \le \varepsilon^{\frac{6}{5}} = \varepsilon^{(1)}.$$

Now we need to show $P^{(1)}$ satisfies assumptions **(A4)** and **(A5)**. Note that

$$\begin{aligned} P^{(1)} &= \grave{P}^{(1)} + P - R + \{\breve{P}, F^{(1)}\} + \{P, F^{(1)}\} \\ &+\frac{1}{2!}\{\{\mathcal{N}, F^{(1)}\}, F^{(1)}\} + \frac{1}{2!}\{\{\breve{P}, F^{(1)}\}, F^{(1)}\} + \frac{1}{2!}\{\{P, F^{(1)}\}, F^{(1)}\} + \cdots \\ &+\frac{1}{n!}\{\cdots\{\mathcal{N}, \underbrace{F^{(1)}\}\cdots, F^{(1)}}_{n}\} + \frac{1}{n!}\{\cdots\{\breve{P}, \underbrace{F^{(1)}\}\cdots, F^{(1)}}_{n}\} \\ &+\frac{1}{n!}\{\cdots\{P, \underbrace{F^{(1)}\}\cdots, F^{(1)}}_{n}\} + \cdots . \end{aligned}$$

Since all of $\mathcal{N}$, $\breve{P}$, P, $F^{(1)}$, $\grave{P}^{(1)}$ have gauge invariance, independent of variables $(q_n, \bar{q}_n)_{n\in\mathcal{J}}$, so does $P^{(1)}$ due to Lemma A.4 and A.5 in Appendix.

For $P - R = \sum_{2|l|+|\alpha|+|\beta|\geq 3} P_{kl\alpha\beta} e^{\mathrm{i}\langle k,\theta\rangle} I^l q^\alpha \bar{q}^\beta$, we have

$$\|P_{\alpha\beta}\|_{\mathcal{D}^{(1)}} \leq \begin{cases} \frac{1}{4}\varepsilon^{(2)} e^{-\rho n^*_{\alpha\beta}}, & |\alpha|+|\beta|\leq 2 \\ e^{-\rho n^*_{\alpha\beta}}, & |\alpha|+|\beta|\geq 3 \end{cases} .$$

Here we applied the estimate $|I| \leq s^{(1)} \leq \frac{1}{4}\varepsilon^{(1)}$ to handle the case that $|\alpha|+|\beta|\leq 2$ and $2|l|+|\alpha|+|\beta|\geq 3$.

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

Lemma 3.4 *For ε sufficiently small, $\{P, F^{(1)}\}$ satisfies*

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

Proof. A straightforward calculation yields that

$$\{P, F^{(1)}\}_{\alpha\beta} = \mathrm{i} \sum_{\substack{n\in\mathbb{Z} \\ (\breve{\alpha},\breve{\beta})+(\hat{\alpha},\hat{\beta})=(\alpha,\beta)}} \left(P_{\breve{\alpha}+e_n,\breve{\beta}} F^{(1)}_{\hat{\alpha},\hat{\beta}+e_n} - P_{\breve{\alpha},\breve{\beta}+e_n} F^{(1)}_{\hat{\alpha}+e_n,\hat{\beta}}\right) \tag{3.63}$$

$$+ \sum_{(\breve{\alpha},\breve{\beta})+(\hat{\alpha},\hat{\beta})=(\alpha,\beta)} \left\{P_{\breve{\alpha}\breve{\beta}}, F^{(1)}_{\hat{\alpha}\hat{\beta}}\right\}. \tag{3.64}$$

- Terms in (3.63)

  Let us consider terms $P_{\breve{\alpha}+e_n,\breve{\beta}} F^{(1)}_{\hat{\alpha},\hat{\beta}+e_n}$.

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

Since $|\hat\alpha|+|\hat\beta+e_n|=1$ or 2 in view of the construction of $F^{(1)}$, we have that

$$|\check\alpha+e_n|+|\check\beta|=|\alpha|+|\beta|+1-(|\hat\alpha|+|\hat\beta|)\le 3. \tag{3.65}$$

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

$$\|P_{\check\alpha+e_n,\check\beta}F^{(1)}_{\hat\alpha,\hat\beta+e_n}\|_{\mathcal{D}_3,\mathcal{O}_+}\le\varepsilon e^{-\rho n^*_{\check\alpha+e_n,\check\beta}}\cdot\varepsilon^{\frac34}e^{-\rho n^*_{\hat\alpha,\hat\beta+e_n}}\le\varepsilon^{\frac74}e^{-\rho n^*_{\alpha\beta}}. \tag{3.66}$$

If $|\check\alpha+e_n|+|\check\beta|=3$, then, by (3.65), $(\hat\alpha,\hat\beta)=(0,0)$, $(\check\alpha,\check\beta)=(\alpha,\beta)$. By the definition of norm $\|X_P\|_{\mathcal{D},\mathcal{O}}$ and the construction of $F^{(1)}$,

$$\|P_{\alpha+e_n,\beta}\|_{\mathcal{D}_3,\mathcal{O}}\le e^{-\rho n^*_{\alpha+e_n,\beta}},\quad \|F^{(1)}_{0,e_n}\|_{\mathcal{D}_3,\mathcal{O}_+}\le s\varepsilon^{\frac34}e^{-\rho|n|}.$$

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

$$\|P_{\alpha+e_n,\beta}F^{(1)}_{0,e_n}\|_{\mathcal{D}_3,\mathcal{O}_+}\le s\varepsilon^{\frac34}e^{-\rho n^*_{\alpha\beta}}\le\frac14\varepsilon^{\frac74}e^{-\rho n^*_{\alpha\beta}}. \tag{3.67}$$

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

By the same argument as above,

$$\|P_{\check\alpha+e_n,\check\beta}F^{(1)}_{\hat\alpha,\hat\beta+e_n}\|_{\mathcal{D}_3,\mathcal{O}_+}\le e^{-\rho n^*_{\check\alpha+e_n,\check\beta}}\cdot\varepsilon^{\frac34}e^{-\rho n^*_{\hat\alpha,\hat\beta+e_n}}\le\varepsilon^{\frac34}e^{-\rho n^*_{\alpha\beta}}. \tag{3.68}$$

Doing the same for $P_{\check\alpha,\check\beta+e_n}F^{(1)}_{\hat\alpha+e_n,\hat\beta}$, we finish estimates for terms in (3.63).

- Terms in (3.64)

By Lemma A.2 and the inequality $n^*_{\alpha\beta}\le\max\{n^*_{\check\alpha\check\beta},\, n^*_{\hat\alpha\hat\beta}\}$, we have

$$\|\{P_{\check\alpha\check\beta},F^{(1)}_{\hat\alpha\hat\beta}\}\|_{\mathcal{D}_{3\eta}}\le c(r-r^{(1)})^{-1}\eta^{-2}\begin{cases}\varepsilon^{\frac74}e^{-\rho n^*_{\alpha\beta}}, & |\alpha|+|\beta|\le 2\\ \varepsilon^{\frac34}e^{-\rho n^*_{\alpha\beta}}, & |\alpha|+|\beta|\ge 3\end{cases}. \tag{3.69}$$

Combining (3.66) – (3.69), there exists $c_8>0$ such that

$$\|\{P,F^{(1)}\}_{\alpha\beta}\|_{\mathcal{D}_{3\eta}}\le c_8(r-r^{(1)})^{-1}\eta^{-2}(\rho-\rho^{(1)})^{-2}\begin{cases}\varepsilon^{\frac74}e^{-\rho^{(1)}n^*_{\alpha\beta}}, & |\alpha|+|\beta|\le 2\\ \varepsilon^{\frac34}e^{-\rho^{(1)}n^*_{\alpha\beta}}, & |\alpha|+|\beta|\ge 3\end{cases},$$

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

(C8): $c_8(r-r^{(1)})^{-1}\eta^{-2}(\rho-\rho^{(1)})^{-2}\varepsilon^{\frac12}\le\frac14$,

Lemma 3.4 is proved. □

By (3.28), (3.32) and (3.33), it is evident that the coefficients of

$$\{\mathcal{N}, F^{(1)}\} = e^{(1)} + \langle \omega^{(1)}, I\rangle + \langle W^{(1)} q, \bar{q}\rangle + \grave{P}^{(1)} - R$$

satisfies $\|\{\mathcal{N}, F^{(1)}\}_{\alpha\beta}\|_{\mathcal{D}_3,\mathcal{O}_+} \leq c\varepsilon e^{-\rho^{(1)} n^*_{\alpha\beta}}$. Then we have the following lemma, whose proof is analogous to that of Lemma 3.4.

Lemma 3.5 *For ε sufficiently small, $\{\{\mathcal{N}, F^{(1)}\}, F^{(1)}\}$ satisfies*

$$\|\{\{\mathcal{N}, F^{(1)}\}, F^{(1)}\}_{\alpha\beta}\|_{\mathcal{D}_{3\eta},\mathcal{O}_+} \leq \frac{1}{4}\varepsilon^{\frac{6}{5}} e^{-\rho^{(1)} n^*_{\alpha\beta}}.$$

Lemma 3.6 *For ε sufficiently small, $\{\breve{P}, F^{(1)}\}$ satisfies*

$$\|\{\breve{P}, F^{(1)}\}_{\alpha\beta}\|_{\mathcal{D}_3,\mathcal{O}_+} \leq \varepsilon^{\frac{1}{4}} e^{-\rho^{(1)} n^*_{\alpha\beta}}, \quad |\alpha| + |\beta| \geq 3.$$

Proof. It can be calculated that

$$\{\breve{P}, F^{(1)}\}_{\alpha\beta} = \mathrm{i} \sum_{\substack{n\in\mathbb{Z} \\ (\breve{\alpha},\breve{\beta})+(\hat{\alpha},\hat{\beta})=(\alpha,\beta)}} \left(\breve{P}_{\breve{\alpha}+e_n,\breve{\beta}} F^{(1)}_{\hat{\alpha},\hat{\beta}+e_n} - \breve{P}_{\breve{\alpha},\breve{\beta}+e_n} F^{(1)}_{\hat{\alpha}+e_n,\hat{\beta}} \right). \tag{3.70}$$

For $\breve{P}_{\breve{\alpha}+e_n,\breve{\beta}} F^{(1)}_{\hat{\alpha},\hat{\beta}+e_n}$ in (3.70), since $|\hat{\alpha}| + |\hat{\beta} + e_n| = 1$ or 2 and $|\breve{\alpha} + e_n| + |\breve{\beta}| \geq 4$ here, it is obvious that $|\alpha| + |\beta| = |\breve{\alpha}| + |\breve{\beta}| + |\hat{\alpha}| + |\hat{\beta}| \leq 3$.

Note that $n^*_{\alpha\beta} \leq \max\{n^*_{\breve{\alpha}+e_n,\breve{\beta}}\, n^*_{\hat{\alpha},\hat{\beta}+e_n}\}$, and

$$n^*_{\breve{\alpha}+e_n,\breve{\beta}} = \max\{n^+_{\breve{\alpha}+e_n,\breve{\beta}}, -n^-_{\breve{\alpha}+e_n,\breve{\beta}}\}, \quad n^*_{\hat{\alpha},\hat{\beta}+e_n} = \max\{n^+_{\hat{\alpha},\hat{\beta}+e_n}, -n^-_{\hat{\alpha},\hat{\beta}+e_n}\}.$$

Then $n^+_{\breve{\alpha}+e_n,\breve{\beta}} - n^-_{\breve{\alpha}+e_n,\breve{\beta}} + n^*_{\hat{\alpha},\hat{\beta}+e_n} \geq n^*_{\alpha\beta}$, and hence

$$\left\| \breve{P}_{\breve{\alpha}+e_n,\breve{\beta}} F^{(1)}_{\hat{\alpha},\hat{\beta}+e_n} \right\|_{\mathcal{D}_3} \leq e^{-\rho(n^+_{\breve{\alpha}+e_n,\breve{\beta}} - n^-_{\breve{\alpha}+e_n,\breve{\beta}})} \cdot \varepsilon^{\frac{3}{4}} e^{-\rho n^*_{\hat{\alpha},\hat{\beta}+e_n}} \leq \varepsilon^{\frac{3}{4}} e^{-\rho n^*_{\alpha\beta}}.$$

Doing the estimate for $\breve{P}_{\breve{\alpha},\breve{\beta}+e_n} F^{(1)}_{\hat{\alpha}+e_n,\hat{\beta}}$ in (3.70) similarly, we have that

$$\|\{\breve{P}, F^{(1)}\}_{\alpha\beta}\|_{\mathcal{D}_3} \leq c_8(\rho - \rho^{(1)})^{-2} \varepsilon^{\frac{3}{4}} e^{-\rho^{(1)} n^*_{\alpha\beta}} \leq \varepsilon^{\frac{1}{4}} e^{-\rho^{(1)} n^*_{\alpha\beta}}, \quad |\alpha| + |\beta| \geq 3,$$

if **(C8)** holds. □

Summarize the analysis above, then the decay property for $P^{(1)}$ can be expressed as

Proposition 3.2 *For ε sufficiently small, $P^{(1)} = \sum_{\alpha,\beta} P^{(1)}_{\alpha\beta}(\theta, I; \xi) q^\alpha \bar{q}^\beta$ satisfies*

$$\|P^{(1)}_{\alpha\beta}\|_{\mathcal{D}^{(1)},\mathcal{O}_+} \leq \begin{cases} \varepsilon^{(1)} e^{-\rho^{(1)} n^*_{\alpha\beta}}, & |\alpha| + |\beta| \leq 2 \\ e^{-\rho^{(1)} n^*_{\alpha\beta}}, & |\alpha| + |\beta| \geq 3 \end{cases}.$$

§3.3.4 A succession of symplectic transformations

With the verification of assumptions **(A4)** and **(A5)** completed, we finish one sub-step of KAM iteration. Suppose that we have arrived at the j^{th} sub-step, $j = 1, \cdots, J$, with $J = \left[\frac{5}{2}\varepsilon^{\frac{a}{2}}\right]$, then we encounter the Hamiltonian

$$\begin{aligned} H^{(j-1)} &= H \circ \Phi^1_{F^{(1)}} \circ \cdots \circ \Phi^1_{F^{(j-1)}} \\ &= \mathcal{N} + \breve{P} + \sum_{i=1}^{j-1} \left(e^{(i)} + \langle \omega^{(i)}, I \rangle + \langle W^{(i)} q, \bar{q} \rangle \right) + P^{(j-1)}, \end{aligned}$$

with the superscript "(0)" labeling quantities before the 1^{st} sub-step in particular. Let

$$R^{(j-1)} := \sum_{\substack{k \\ 2|l|+|\alpha|+|\beta| \leq 2}} P^{(j-1)}_{kl\alpha\beta} e^{\mathrm{i}\langle k, \theta \rangle} I^l q^\alpha \bar{q}^\beta. \tag{3.71}$$

As demonstrated in Proposition 3.1, on $\mathcal{O}_+$, the following homological equation

$$\{\mathcal{N}, F^{(j)}\} + R^{(j-1)} = e^{(j)} + \langle \omega^{(j)}, I \rangle + \langle W^{(j)} q, \bar{q} \rangle + \grave{P}^{(j)}, \tag{3.72}$$

can be solved, with $F^{(j)}$, $e^{(j)}$, $\omega^{(j)}$, $W^{(j)}$, $\grave{P}^{(j)}$ having properties similar to $F^{(1)}$, $e^{(1)}$, $\omega^{(1)}$, $W^{(1)}$, $\grave{P}^{(1)}$ respectively. Then we obtain

$$H^{(j)} = H^{(j-1)} \circ \Phi^1_{F^{(j)}} = \mathcal{N} + \breve{P} + \sum_{i=1}^{j} \left(e^{(i)} + \langle \omega^{(i)}, I \rangle + \langle W^{(i)} q, \bar{q} \rangle \right) + P^{(j)}.$$

The estimates for $F^{(j)}$ and the verification of assumptions for $P^{(j)}$ can be done similarly as in subsection §3.3.3.

Proposition 3.3 *Consider the Hamiltonian H in (3.14). There exist J symplectic transformations $\Phi^{(j)} = \Phi^1_{F^{(j)}}$, $j = 1, \cdots, J$, generated by the corresponding real-analytic Hamiltonians $F^{(j)}$ respectively, such that*

$$H^{(j)} = H \circ \Phi^{(1)} \circ \cdots \circ \Phi^{(j)} = \mathcal{N} + \breve{P} + G_j + P^{(j)}, \quad j = 1, \cdots, J,$$

is real-analytic on $\mathcal{D}^{(j)} = \mathcal{D}_{d,\rho_+}(r^{(j)}, s^{(j)})$, with $G_j = \sum_{i=1}^{j} \left(e^{(i)} + \langle \omega^{(i)}, I \rangle + \langle W^{(i)} q, \bar{q} \rangle \right)$. For $i = 1, 2, 3, 4$, $\eta = \varepsilon^{\frac{1}{5}}$, let

$$\begin{aligned} \mathcal{D}^{(j)}_i &= \mathcal{D}_{d,\rho_+}(r^{(j+1)} + \frac{i}{4}(r^{(j)} - r^{(j+1)}), \frac{i}{4} s^{(j)}), \\ \mathcal{D}^{(j)}_{i\eta} &= \mathcal{D}_{d,\rho_+}(r^{(j+1)} + \frac{i}{4}(r^{(j)} - r^{(j+1)}), \frac{i}{4} \eta s^{(j)}). \end{aligned}$$

(a) *With* $R^{(j-1)}$ *defined in (3.71),* $F^{(j)}$ *satisfies the homological equation (3.72) on* $\mathcal{O}_+$*, and*

$$\begin{aligned}
&\|X_{F^{(j)}}\|_{\mathcal{D}_3^{(j-1)},\mathcal{O}_+} \le \varepsilon^{-\frac{1}{4}}\varepsilon^{(j-1)},\\
&\Phi^t_{F^{(j)}}\ :\ \mathcal{D}^{(j-1)}_{2\eta} \to \mathcal{D}^{(j-1)}_{3\eta}, \quad -1 \le t \le 1,\\
&\|D\Phi^t_{F^{(j)}} - I\|_{\mathcal{D}^{(j-1)}_{1\eta}} < 2\varepsilon^{-\frac{1}{4}}\varepsilon^{(j-1)},\\
&\|F^{(j)}_{\alpha\beta}\|_{\mathcal{D}_3^{(j-1)},\mathcal{O}_+} \le \begin{cases} \varepsilon^{-\frac{1}{4}}\varepsilon^{(j-1)}e^{-\rho^{(j-1)}n^*_{\alpha\beta}}, & |\alpha|+|\beta|\le 2\\ 0, & |\alpha|+|\beta|\ge 3\end{cases},\\
&\partial_{q_n}F^{(j)} = \partial_{\bar{q}_n}F^{(j)} \equiv 0, \quad \forall n \in \mathcal{J}.
\end{aligned}$$

(b) G_j *satisfies that* $\|X_{G_j}\|_{\mathcal{D}_3^{(j)},\mathcal{O}_+} \le c\varepsilon$ *and for* $i = 1, 2, \cdots, j$*,*

$$\begin{aligned}
&|\omega^{(i)}|_{\mathcal{O}_+} \le \varepsilon^{(i-1)},\\
&|W^{(i)}_{mn}|_{\mathcal{O}_+} \le \begin{cases} \varepsilon^{(i-1)}e^{-\rho^{(i-1)}\max\{|m|,|n|\}}, & |m|,|n| \le N_+,\quad m,n \notin \mathcal{J}\\ 0, & \textit{otherwise}\end{cases}.
\end{aligned}$$

(c) $\|X_{\breve{P}+P^{(j)}}\|_{\mathcal{D}^{(j)},\mathcal{O}_+} \le \varepsilon^{(j)}$ *and* $P^{(j)}$ *satisfies assumptions* **(A4)**, **(A5)**, *which include*

$$\begin{aligned}
&\|P^{(j)}_{\alpha\beta}\|_{\mathcal{D}^{(j)},\mathcal{O}_+} \le \begin{cases} \varepsilon^{(j)}e^{-\rho^{(j)}n^*_{\alpha\beta}}, & |\alpha|+|\beta|\le 2\\ e^{-\rho^{(j)}n^*_{\alpha\beta}}, & |\alpha|+|\beta|\ge 3\end{cases}\\
&\partial_{q_n}P^{(j)} = \partial_{\bar{q}_n}P^{(j)} \equiv 0, \quad \forall n \in \mathcal{J}.
\end{aligned}$$

Let $s_+ = s^{(J)} = 2^{-3J}\varepsilon^{\frac{J}{5}}s$, $\Phi = \Phi^{(1)} \circ \cdots \circ \Phi^{(J)}$, and

$$\mathcal{N}_+ = e_+ + \langle \omega_+, I\rangle + \langle Tq, \bar{q}\rangle + \langle (A+W_+)q, \bar{q}\rangle,$$

with $\Omega_+ = T + A + W_+$, and

$$e_+ = e + \sum_{j=1}^{J} e^{(j)}, \quad \omega_+ = \omega + \sum_{j=1}^{J}\omega^{(j)}, \quad W_+ = W + \sum_{j=1}^{J} W^{(j)}.$$

Then $\Phi : \mathcal{D}_+ \to \mathcal{D}$. From the estimates of $\omega^{(j)}$ and $W^{(j)}$, we have

$$|\omega_+ - \omega|_{\mathcal{O}_+} \le c\varepsilon, \tag{3.73}$$

$$|(W_+ - W)_{mn}|_{\mathcal{O}_+} \le \begin{cases} \varepsilon^{\frac{1}{2}}e^{-\frac{\rho}{2}\max\{|m|,|n|\}}, & |m|,|n| \le N_+,\quad m,n \notin \mathcal{J}\\ 0, & \text{otherwise}\end{cases}. \tag{3.74}$$

Since $W^* = W$ and $(W^{(i)})^* = W^{(i)}$, W_+ is still a Hermitian matrix. So (**A1**) and (**A2**) hold with $p_+ = p + \varepsilon^{\frac{1}{2}}$, $\sigma_+ := \frac{1}{3}\rho$.

Let $P_+ = P^{(J)}$. It can be verified that the assumptions **(A4)** and **(A5)** for $P^{(J)}$ hold, which is an analogue to the process in subsection §3.3.3.

This completes one step of KAM iterations.

§3.4 Proof of Theorem 3.2

With $\varepsilon_0 = \epsilon^{\frac{1}{4}}$, $\sigma_0 = 1$, $\hat{N} = |\ln \varepsilon_0|$, and$N_0 = 6|\ln \varepsilon_0|$, $\rho_0 = N_0^{-1}$,

$$M_0 = \max\left\{2^{\tilde{s}+4}C\frac{L^{\tilde{s}+1}((\tilde{s}+1)!)^2}{\tilde{\xi}},\ 2\tilde{\tau},\ 8,\ \frac{12(2\tau+b+3)}{\tilde{\tau}}\right\},$$

one can define the following sequences as in [16],

$$\begin{aligned} &M_{\nu+1} = M_\nu^{\tilde{s}M_\nu^3}, \quad a_\nu = \frac{1}{\tilde{\tau}}M_\nu^{-3\tilde{s}M_\nu^3}, \quad \varepsilon_{\nu+1} = \varepsilon_\nu^{\frac{1}{2}\varepsilon_\nu^{-a_\nu/2}}, \\ &N_{\nu+1} = \varepsilon_\nu^{-a_\nu}, \qquad \rho_{\nu+1} = \varepsilon_\nu^{a_\nu}, \qquad \sigma_{\nu+1} = \frac{1}{3}\rho_\nu. \end{aligned}$$

Given $p_0 = \varepsilon_0^{\frac{1}{2}}$, $r_0 = r$, $s_0 = s$, 可定义序列the other sequences are defined as

$$\begin{aligned} &p_{\nu+1} = p_\nu + \varepsilon_\nu^{\frac{1}{2}}, \qquad K_{\nu+1} = N_{\nu+1} - (M_\nu+1)N_\nu, \quad J_\nu = \left[\frac{5}{2}\varepsilon_\nu^{-\frac{a_\nu}{2}}\right], \\ &r_\nu = r_0\left(1 - \sum_{i=2}^{\nu+1} 2^{-i}\right), \quad s_{\nu+1} = 2^{-3J_\nu}\varepsilon_\nu^{\frac{J_\nu}{5}} s_\nu, \qquad \gamma_\nu = \varepsilon_\nu^{\frac{1}{80}}. \end{aligned}$$

Let $\mathcal{D}_\nu$ and $\mathcal{O}_\nu$ be as defined in Section §3.3.

§3.4.1 Iteration lemma

The preceding analysis can be summarized as follows.

Lemma 3.7 *There exists ε_0 sufficiently small such that the following holds for all $\nu = 0, 1, \cdots$.*

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

$$\begin{aligned} \mathcal{N}_\nu &= e_\nu + \langle \omega_\nu, I\rangle + \langle \Omega_\nu q, \bar{q}\rangle \\ &= e_\nu + \langle \omega_\nu, I\rangle + \langle (T + A + W_\nu)q, \bar{q}\rangle \\ P_\nu &= \sum_{\alpha,\beta} (P_\nu)_{\alpha\beta}(\theta, I; \xi) q^\alpha \bar{q}^\beta, \end{aligned}$$

satisfying

$$
\begin{aligned}
&(\Omega_\nu)_{mn} \equiv 0, \quad \{m,n\} \cap \mathcal{J} \neq \emptyset, \\
&|(W_\nu)_{mn}|_{\mathcal{O}_\nu} \leq \begin{cases} p_\nu e^{-\sigma_\nu \max\{|m|,|n|\}}, & |m|,|n| \leq N_\nu \\ \quad 0, & otherwise \end{cases}, \\
&|\omega_{\nu+1} - \omega_\nu|_{\mathcal{O}_{\nu+1}} \leq \varepsilon_\nu, \\
&|(W_{\nu+1} - W_\nu)_{mn}|_{\mathcal{O}_{\nu+1}} \leq \begin{cases} \varepsilon_\nu^{\frac{1}{2}} e^{-\frac{\rho_\nu}{2}\max\{|m|,|n|\}}, & |m|,|n| \leq N_{\nu+1}, \quad m,n \notin \mathcal{J} \\ \quad 0, & otherwise \end{cases}.
\end{aligned}
$$

Moreover, P_ν *has gauge invariance and* $\|X_{\check{P}+P_\nu}\|_{\mathcal{D}_\nu,\mathcal{O}_\nu} \leq \varepsilon_\nu$,

$$
\begin{aligned}
&\|(P_\nu)_{\alpha\beta}\|_{\mathcal{D}_\nu,\mathcal{O}_\nu} \leq \begin{cases} \varepsilon_\nu e^{-\rho_\nu n^*_{\alpha\beta}}, & |\alpha|+|\beta| \leq 2 \\ e^{-\rho_\nu n^*_{\alpha\beta}}, & |\alpha|+|\beta| \geq 3 \end{cases}, \\
&\partial_{q_n} P_\nu = \partial_{\bar{q}_n} P_\nu \equiv 0, \quad \forall n \in \mathcal{J}.
\end{aligned}
$$

(b) *For each* ν, *there is a symplectic transformation* $\Phi_\nu : \mathcal{D}_{\nu+1} \to \mathcal{D}_\nu$ *with*

$$
\|D\Phi_\nu - Id\|_{\mathcal{D}_{\nu+1},\mathcal{O}_{\nu+1}} \leq \varepsilon_\nu^{\frac{1}{2}},
$$

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

Proof. Let $c_0 := 8e^{20}\max\{c_1, \cdots, c_8\}$. We need to verify the assumptions (**C1**)$-$(**C8**) for $\nu = 0, 1, \cdots$. By noting that

$$
N_{\nu+1} = \varepsilon_\nu^{a_\nu} = \rho_{\nu+1}^{-1}, \quad \sigma_{\nu+1} = \frac{1}{3}\rho_\nu, \quad r_\nu^{(j)} - r_\nu^{(j+1)} = \frac{r_\nu - r_{\nu+1}}{2J_\nu}, \quad \rho_\nu^{(j)} - \rho_\nu^{(j+1)} = \frac{\rho_\nu - \rho_{\nu+1}}{2J_\nu},
$$

it is sufficient for us to check:

(D1) $c_0 s_\nu \leq \varepsilon_\nu$,

(D2) $c_0 \left(\frac{r_\nu - r_{\nu+1}}{2J_\nu}\right)^{-(2\tau+b+1)} \left(\frac{\rho_\nu - \rho_{\nu+1}}{2J_\nu}\right)^{-2} \leq \varepsilon_\nu^{-\frac{1}{20}}$,

(D3) $c_0 N_{\nu+1}^8 M_\nu^8 N_\nu^{20} e^{8M_\nu N_\nu \rho_\nu} \leq \varepsilon_\nu^{-\frac{7}{40}}$,

(D4) $e^{-\frac{\rho_\nu K_{\nu+1}}{2J_\nu}} \leq \varepsilon_\nu^{\frac{2}{5}}$,

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

By the choice of s_0, the condition (**D1**) clearly holds for $\nu = 0$. Suppose that it holds for some ν, then it is easy to see that

$$
c_0 s_{\nu+1} = 2^{-3J_\nu} \varepsilon_\nu^{\frac{J_\nu}{5}} \cdot c_0 s_\nu < 2^{-3J_\nu} \varepsilon_\nu^{\frac{J_\nu}{5}} \cdot \varepsilon_\nu < \varepsilon_{\nu+1}.
$$

Hence (**D1**) holds for all ν.

Let us first take ε_0 sufficiently small such that

$$\varepsilon_0^{\frac{1}{20}-\frac{1}{2}a_0(2\tau+b+3)} \leq \frac{1}{c_0}\left(\frac{r_0}{20}\right)^{2\tau+b+1}\left(\frac{1-\varepsilon_0^{a_0}}{5}\right)^2.$$

Here we have applied $M_0 \geq \frac{12}{\tilde{\tau}}(2\tau+b+3)$ and $a_0 = M_0^{-3\tilde{s}M_0^3}$ such that $\frac{1}{20}-\frac{1}{2}a_0(2\tau+b+3) > 0$. Then, recalling that $r_\nu - r_{\nu+1} = \frac{r_0}{2^{2+\nu}}$和$J_\nu = \left[\frac{5}{2}\varepsilon_\nu^{-\frac{a_\nu}{2}}\right]$, 可得

$$c_0\left(\frac{r_0-r_1}{2J_0}\right)^{-(2\tau+b+1)}\left(\frac{\rho_0-\rho_1}{2J_0}\right)^{-2} \leq \varepsilon_0^{-\frac{1}{20}},$$

i.e., (**D2**) holds for $\nu = 0$. Since for $\nu \geq 1$ and for ε_0 sufficiently small,

$$\varepsilon_\nu^{\frac{1}{40}-\frac{1}{2}a_\nu(2\tau+b+3)} \ll \varepsilon_0^{\left(\frac{6}{5}\right)^\nu} \ll \frac{1}{2^{\nu(2\tau+b+1)}c_0}\left(\frac{r_0}{20}\right)^{2\tau+b+1}, \quad \varepsilon_\nu^{\frac{1}{40}} \ll \left(\frac{\varepsilon_{\nu-1}^{a_{\nu-1}}-\varepsilon_\nu^{a_\nu}}{5}\right)^2,$$

we have

$$c_0\left(\frac{r_\nu-r_{\nu+1}}{2J_\nu}\right)^{-(2\tau+b+1)}\left(\frac{\rho_\nu-\rho_{\nu+1}}{2J_\nu}\right)^{-2} \leq \varepsilon_\nu^{-\frac{1}{20}}.$$

Thus, (**D2**) holds true.

In Section 6 of [16], the basic smallness assumption of ε_ν, i.e., the inequality (D.3) in Lemma D.1, has been verified, then all other assumptions are immediate, including the inequality

$$\Gamma_\nu N_\nu^2 e^{6M_\nu N_\nu \rho_\nu} \leq \varepsilon_\nu^{-\frac{1}{8}},$$

where Γ_ν increases superexponentially in M_ν. Since all of M_ν, N_ν, ρ_ν and ε_ν here are defined in the same way as [16], we can apply this inequality. So (**D3**) has been verified.

By the definition of ρ_ν, a_ν and ε_ν, we have

$$\rho_\nu \varepsilon_\nu^{-\frac{1}{2}a_\nu} > \ln\frac{1}{\varepsilon_\nu}.$$

Then we see that (**D4**) holds for $\nu = 0, 1, \cdots$. □

§3.4.2 Convergence

Now we fix $x \in \tilde{\mathcal{X}}$, with $\tilde{\mathcal{X}}$ defined as in Proposition 1.6. This means that the blocks mentioned in Proposition 1.6 are eventually stationary after some step, i.e., for each $n \in \mathbb{Z}$, there is a $\nu_0(n)$ such that

$$\Lambda^{\nu+1}(n) = \Lambda^\nu(n), \quad \forall \nu \geq \nu_0(n).$$

In this case, the local decay rate for n may not shrink with ν necessarily(ρ_ν is the global upper bound of the rates for all $n \in \mathbb{Z}$).

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

$$H_0 \circ \Psi^\nu = H_\nu = \mathcal{N}_\nu + \breve{P} + P_\nu.$$

Let $\mathcal{O}_{\varepsilon_0} = \cap_{\nu=0}^\infty \mathcal{O}_\nu$. As in standard arguments, thanks to Lemma 3.3, it concludes that H_ν, $\mathcal{N}_\nu$, P_ν, Ψ^ν, e_ν, ω_ν and W_ν converge uniformly on $\mathcal{D}_{d,0}(\frac{1}{2}r_0, 0) \times \mathcal{O}_{\varepsilon_0}$ to, say, H_∞, $\mathcal{N}_\infty$, P_∞, Ψ^∞, e_∞, ω_∞ and W_∞ respectively, in which case it is clear that

$$\mathcal{N}_\infty = e_\infty + \langle \omega_\infty, I \rangle + \langle (T + A + W_\infty) q, \bar{q} \rangle,$$

with $\Omega_\infty = T + A + W_\infty$ satisfying $(\Omega_\infty)_{mn} \equiv 0$ if m or $n \in \mathcal{J}$. Since $\|X_{P_\nu}\|_{\mathcal{D}_\nu, \mathcal{O}_\nu} \leq \varepsilon_\nu$ with $\varepsilon_\nu \to 0$, it follows that $\|X_{P_\infty}\|_{\mathcal{D}_{d,0}(\frac{1}{2}r_0,0), \mathcal{O}_{\varepsilon_0}} = 0$.

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

$$\Phi^t_{H_0} \circ \Psi^\infty = \Psi^\infty \circ \Phi^t_{H_\infty}, \quad \Psi^\infty : \mathcal{D}_{d,0}(\frac{1}{2}r_0, 0) \to \mathcal{D}_0.$$

Hence,

$$\Phi^t_{H_0}(\Psi^\infty(\mathbb{T}^b \times \{\xi\})) = \Psi^\infty \Phi^t_{\mathcal{N}_\infty}(\mathbb{T}^b \times \{\xi\}) = \Psi^\infty(\mathbb{T}^b \times \{\xi\}), \quad \forall \xi \in \mathcal{O}_{\varepsilon_0}.$$

This means that $\Psi^\infty(\mathbb{T}^b \times \{\xi\})$ is an embedded invariant torus of the original perturbed Hamiltonian system at $\xi \in \mathcal{O}_{\varepsilon_0}$. Moreover, the frequencies $\omega_\infty(\xi)$ associated with $\Psi^\infty(\mathbb{T}^b \times \{\xi\})$ are slightly deformed from the unperturbed ones, $\omega(\xi)$.

§3.4.3 Measure estimate

At the ν^{th} step of KAM iteration, we need to exclude the following resonant parameter set

$$\mathcal{R}^\nu_k := \mathcal{R}^{\nu 1}_k \bigcup \left(\bigcup_{n \in \Lambda^\nu} \mathcal{R}^{\nu 2}_{kn} \right) \bigcup \left(\bigcup_{m,n \in \Lambda^\nu} \mathcal{R}^{\nu 3}_{kmn} \right) \bigcup \left(\bigcup_{m,n \in \Lambda^\nu} \mathcal{R}^{\nu 4}_{kmn} \right), \quad k \neq 0$$

for any fixed $x \in \tilde{\mathcal{X}}$, where

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

with $\{\mu_j^\nu\}_{j\in\Lambda^\nu}$ eigenvalues of $\tilde{D}_{\Lambda^\nu}^\nu + \tilde{A}_\nu$. It is clear that $\mathcal{O} \setminus \mathcal{O}_\varepsilon \subseteq \bigcup_{\nu\geq 0}\bigcup_{k\neq 0}\mathcal{R}_k^\nu$.

Recalling that ω_0 is a diffeomorphism of ξ, together with the estimates in (3.26), (3.73) and (3.74), we have

$$|\partial_\xi(\langle k, \omega_\nu\rangle + \mu_m^\nu - \mu_n^\nu)| \geq |\partial_\xi\langle k, \omega_0\rangle| - \varepsilon_0^{\frac{1}{4}}|k| - p = O(|k|)$$

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

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

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

$$|\mathcal{O}_0 \setminus \mathcal{O}_\varepsilon| \leq \left|\bigcup_{\nu\geq 0}\bigcup_{k\neq 0}\mathcal{R}_k^\nu\right| \leq c\sum_{\nu\geq 0}\sum_{k\neq 0}\frac{\gamma_\nu}{|k|^{\tau+1}} = c\sum_{\nu\geq 0}\gamma_\nu \sim \gamma_0 = \varepsilon_0^{\frac{1}{80}}.$$

附录一 Hamiltonian vector field and Poisson bracket

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 A.1 *The norm* $\|\cdot\|_{\mathcal{D},\mathcal{O}}$ *has the Banach algebraic property, i.e.,*

$$\|FG\|_{\mathcal{D}}\leq\|F\|_{\mathcal{D}}\|G\|_{\mathcal{D}}.$$

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

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

□

Lemma A.2 (Generalized Cauchy Inequalities) *The various components of the Hamiltonian vector field* X_F *satisfy: for any* $0<r'<r,\ 0<\rho'<\rho$,

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

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

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

by the usual Cauchy inequality. As a linear operator on $\ell^1_{d,\rho}(\mathbb{Z})$, $\partial_q F$ 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 $0 < \rho' < \rho$,

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

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

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

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. □

Lemma A.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* $0 < \eta \ll 1$.

For the proof, refer to [21].

Lemma A.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 $\sum_{j=1}^{b} k_i + |\alpha| - |\beta| \neq 0$. By a simple calculation, we have

$$\{F, G\}_{k\alpha\beta} = \mathrm{i} \sum_{\substack{\check{k}+\hat{k}=k \\ \check{\alpha}+\hat{\alpha}=\alpha \\ \check{\beta}+\hat{\beta}=\beta}} \left(\langle \partial_I F_{\check{k}\check{\alpha}\check{\beta}}, \hat{k} \rangle G_{\hat{k}\hat{\alpha}\hat{\beta}} - \langle \check{k}, \partial_I G_{\hat{k}\hat{\alpha}\hat{\beta}} \rangle F_{\check{k}\check{\alpha}\check{\beta}} \right) \tag{A.1}$$

$$+\mathrm{i} \sum_{\substack{\check{k}+\hat{k}=k \\ \check{\alpha}+\hat{\alpha}=\alpha \\ \check{\beta}+\hat{\beta}=\beta}} \sum_{m\in\mathbb{Z}} \left(F_{\check{k}(\check{\alpha}+e_m)\check{\beta}} G_{\hat{k}\hat{\alpha}(\hat{\beta}+e_m)} - F_{\check{k}\check{\alpha}(\check{\beta}+e_m)} G_{\hat{k}(\hat{\alpha}+e_m)\hat{\beta}} \right). \tag{A.2}$$

Assume $\sum_{j=1}^{b} k_i + |\alpha| - |\beta| \neq 0$. 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 (A.1) and (A.2), each term $\equiv 0$. Thus Lemma A.4 is obtained. □

Lemma A.5 *If there exists $n_* \in \mathbb{Z}$ such that*

$$\partial_{q_{n_*}} F = \partial_{\bar{q}_{n_*}} F = \partial_{q_{n_*}} G = \partial_{\bar{q}_{n_*}} G \equiv 0,$$

then $\partial_{q_{n_*}} \{F, G\} = \partial_{\bar{q}_{n_*}} \{F, G\} \equiv 0$.

Proof. Since

$$\begin{aligned} \partial_{q_{n_*}} \{F, G\} &= \partial_{q_{n_*}} \left(\langle \partial_I F, \partial_\theta G \rangle - \langle \partial_\theta F, \partial_I G \rangle + \mathrm{i} \sum_{m\in\mathbb{Z}} (\partial_{q_m} F \cdot \partial_{\bar{q}_m} G - \partial_{\bar{q}_m} F \cdot \partial_{q_m} G) \right) \\ &= \big\langle \partial_I (\partial_{q_{n_*}} F), \partial_\theta (\partial_{q_{n_*}} G) \big\rangle - \big\langle \partial_\theta (\partial_{q_{n_*}} F), \partial_I (\partial_{q_{n_*}} G) \big\rangle \\ &\quad + \mathrm{i} \sum_{m\in\mathbb{Z}} \big(\partial_{q_m} (\partial_{q_{n_*}} F) \cdot \partial_{\bar{q}_m} (\partial_{q_{n_*}} G) - \partial_{\bar{q}_m} (\partial_{q_{n_*}} F) \cdot \partial_{q_m} (\partial_{q_{n_*}} G) \big) \\ &\equiv 0 \end{aligned}$$

and similarly, $\partial_{\bar{q}_{n_*}} \{F, G\} \equiv 0$, Lemma A.5 is proven. □

附录二 Decay property of matrices

Lemma B.1 *Given two matrices* $G = (G_{mn})_{m,n\in\mathbb{Z}}$ *and* $F = (F_{mn})_{m,n\in\mathbb{Z}}$*. Let* $K = GF$*.*

(1) *If* $|G_{mn}| \le c_G e^{-\sigma_G|m-n|}$, $|F_{mn}| \le c_F e^{-\sigma_F|m-n|}$ *for some positive* c_G, c_F, σ_G, $\sigma_F > 0$, *then we have*

$$|K_{mn}| \le c_K e^{-\sigma_K|m-n|}$$

for any $0 < \sigma_K < \min\{\sigma_G,\, \sigma_F\}$ *and* $c_K = c\cdot c_G c_F(\min\{\sigma_G,\, \sigma_F\} - \sigma_K)^{-1}$.

(2) *If* $|G_{mn}| \le c_G e^{-\sigma_G \max\{|m|,\,|n|\}}$, $|F_{mn}| \le c_F e^{-\sigma_F|m-n|}$, *then*

$$|K_{mn}| \le c_K e^{-\sigma_K \max\{|m|,\,|n|\}}.$$

(3) *If* $|G_{mn}| \le c_G e^{-\sigma_G|m-n|}$, $|F_{mn}| \le c_F e^{-\sigma_F \max\{|m|,\,|n|\}}$, *then*

$$|K_{mn}| \le c_K e^{-\sigma_K \max\{|m|,\,|n|\}}.$$

(4) *If* $|G_{mn}| \le c_G e^{-\sigma_G \max\{|m|,\,|n|\}}$, $|F_{mn}| \le c_F e^{-\sigma_F \max\{|m|,\,|n|\}}$, *then*

$$|K_{mn}| \le c_K e^{-\sigma_K \max\{|m|,\,|n|\}}.$$

In particular, if $\sigma_G \ne \sigma_F$*, then the conclusions above hold with* $\sigma_K = \min\{\sigma_G,\, \sigma_F\}$ *and* $c_K = c\cdot c_G c_F|\sigma_G - \sigma_F|^{-1}$.

Proof. Since the matrix element of $K = GF$ can be formulated as $K_{mn} = \sum_{l\in\mathbb{Z}} G_{ml}F_{ln}$, we have that, in Case (1), for any $0 < \sigma_K < \min\{\sigma_G, \sigma_F\}$,

$$\begin{aligned}
|(GF)_{mn}| &\le \sum_{l\in\mathbb{Z}} |G_{ml}||F_{ln}| \\
&\le c_G c_F \sum_{l\in\mathbb{Z}} e^{-\sigma_G|m-l|} e^{-\sigma_F|l-n|} \\
&\le c_G c_F e^{-\sigma_K|m-n|} \sum_{l\in\mathbb{Z}} e^{-(\sigma_G-\sigma_K)|m-l|} e^{-(\sigma_F-\sigma_K)|l-n|} \\
&\le c\cdot c_G c_F(\min\{\sigma_G,\, \sigma_F\} - \sigma_K)^{-1} e^{-\sigma_K|m-n|}.
\end{aligned}$$

Here we have applied the basic triangular inequality $|m-l| + |l-n| \ge |m-n|$.

Moreover, if $\sigma_G \neq \sigma_F$, assume that $0 < \sigma_G < \sigma_F$ without loss of generality, then

$$\begin{aligned}|(GF)_{mn}| &\leq c_G c_F \sum_{l\in\mathbb{Z}} e^{-\sigma_G|m-l|}e^{-\sigma_F|l-n|} \\ &\leq c_G c_F e^{-\sigma_G|m-n|}\sum_{l\in\mathbb{Z}} e^{-(\sigma_F-\sigma_G)|l-n|} \\ &\leq c\cdot c_G c_F(\sigma_F-\sigma_G)^{-1}e^{-\sigma_G|m-n|}.\end{aligned}$$

As for Case (2)−(4), the corresponding conclusions can also be obtained by using the trivial facts

$$|m-l|+\max\{|l|,\,|n|\} \geq \max\{|m|,\,|n|\},\quad \max\{|m|,\,|l|\}+\max\{|l|,\,|n|\} \geq \max\{|m|,\,|n|\}.$$

Thus Lemma B.1 has been proved. □

Remark B.1 *If we replace the matrix* F *satisfying* $|F_{mn}| \leq c_F e^{-\sigma_F \max\{|m|,\,|n|\}}$ *with a vector* $f = (f_n)_{n\in\mathbb{Z}}$ *satisfying* $|f_n| \leq c_f e^{-\sigma_f|n|}$ *in Case (3) and (4), then for the vector* Gf*, we can obtain also the conclusion that* $|(Gf)_n| \leq c_K e^{-\sigma_K|n|}$.

附录三 Proof of Theorem 1.4

For $R > 0$, let $\mathcal{A}_R$ denote the set of period 1 holomorphic bounded functions on

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

equiped with the sup-norm

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

Assume that V is a period 1 meromorphic function on $\mathcal{S}_R$, and there exists some $C > 0$ such that

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

$V(x) = \tan \pi x$ is a typical example, with any C any value between 0 and π. Such function has the following stability.

Lemma C.1 *Given any $g \in \mathcal{A}_R$, satisfying $\|g\|_R < \varrho C$. If $0 < \varrho < R$, then $\tilde{V} := V + g$ is a meromorphic function on $\mathcal{S}_{R-\varrho}$, satisfying*

$$|\tilde{V}(z) - \tilde{V}(z-a)| \geq \left(C - \frac{1}{\varrho}\|g\|_R\right)|a|_1, \quad \forall a \in \mathbb{R}, \quad z \in \mathcal{S}_{R-\varrho}.$$

Moreover, $z \in \mathcal{S}_{R-\varrho}$ is the pole of $\tilde{V}$ if and only if it is the pole of V.

Proof. Since $g \in \mathcal{A}_R$, using the Cauchy formula,

$$\left|\frac{dg}{dz}(z_*)\right| \leq \left|\oint_\gamma \frac{1}{2\pi \mathrm{i}} \frac{g(z)}{(z-z_*)^2} dz\right|, \quad \forall z_* \in \mathcal{S}_{R-\varrho},$$

where γ is any path contained in $\mathcal{S}_R$ and enclosing z_*. Thus

$$\left|\frac{dg}{dz}(z_*)\right| \leq \frac{|\gamma|}{2\pi\, \mathrm{dist}(z,\gamma)^2}\|g\|_R.$$

According to the fact that $z_* \in \mathcal{S}_{R-\varrho}$, we can choose γ as as the circle of radius ϱ around z_*. Then

$$\left\|\frac{dg}{dz}\right\|_{R-\varrho} \leq \frac{1}{\varrho}\|g\|_R.$$

Thus, if $\|g\|_R < \varrho C$, from the inequality

$$\frac{|g(z) - g(z-a)|}{|a|_1} \leq \left\| \frac{dg}{dz} \right\|_{R-\varrho}, \quad \forall a \in \mathbb{R}, \quad z \in \mathcal{S}_{R-\varrho},$$

we can see

$$|\tilde{V}(z) - \tilde{V}(z-a)| \geq |V(z) - V(z-a)| - |g(z) - g(z-a)| \geq \left(C - \frac{1}{\varrho}\|g\|_R \right) |a|_1.$$

For the invariance of poles, it is evident. □

We are going to analyze the linear operator L on $\ell^2(\mathbb{Z}^d)$, by applying the KAM iteration with the normal form

$$D(x) = \mathrm{diag}\{V(x + \langle n, \tilde{\alpha} \rangle)\}_{n \in \mathbb{Z}^d},$$

where $\tilde{\alpha} \in \mathbb{R}^d$ satisfies the Diophantine condition (1.3) and $x \in R/\mathbb{Z}$ such that $x + n\tilde{\alpha}$ is not the pole of V. Defined by (1.4), L is a sum of two infinite-dimensional matrix

$$L = D_0 + Z_0 = \mathrm{diag}\{\tan \pi(x + \langle n, \tilde{\alpha} \rangle)\}_{n \in \mathbb{Z}^d} + \epsilon \Delta.$$

Consider the symmetric matrix $Z = (Z_{mn})_{m,n \in \mathbb{Z}^d}$, with $Z_{mn} \in \mathcal{A}_R$, real-analytic on $\mathbb{R}/\mathbb{Z}$, and satisfying the shift condition(with respect to $\tilde{\alpha}$), i.e.,

$$Z_{m+k,n+k}(x) = Z_{mn}(x + \langle k, \tilde{\alpha} \rangle), \quad x \in \mathbb{R}/\mathbb{Z},$$ [1]

and there exists $\varepsilon > 0$ such that

$$\|Z_{mn}\|_R \leq \varepsilon e^{-\rho|m-n|}. \tag{C.2}$$

Lemma C.2 *There exists un anti-symmetric matric*$F = (F_{mn})_{m,n \in \mathbb{Z}^d}$*, with* $F_{mn} \in \mathcal{A}_R$ *real-analytic on* $\mathbb{R}/\mathbb{Z}$*, satisfying the shift condition, such that*

$$[D, F] + Z = \mathrm{diag}\{Z_{nn}\}_{n \in \mathbb{Z}^d}. \tag{C.3}$$

Moreover,

$$\|F_{mn}\|_R \leq C^{-1}\gamma^{-1}|m-n|^{\tilde{\tau}} \cdot \|Z_{mn}\|_R.$$

[1] It is easy to verify that this property is conserved under the product of matrix.

Proof. By a straightforward calculation,

$$[D, F]_{mn} = (V(x + \langle m, \tilde{\alpha}\rangle) - V(x + \langle n, \tilde{\alpha}\rangle)F_{mn}.$$

So, define F as

$$F_{mn} = \begin{cases} \dfrac{Z_{mn}}{V(x + \langle n, \tilde{\alpha}\rangle) - V(x + \langle m, \tilde{\alpha}\rangle)}, & m \neq n \\ 0, & m = n \end{cases},$$

then we get the equality (C.3). It is obvious that F is anti-symmetric and satisfies the shift condition.

Since V satisfies the condition (C.1) and $\tilde{\alpha}$ satisfies the Diophantine condition (1.3), we have

$$|V(x + \langle n, \tilde{\alpha}\rangle) - V(x + \langle m, \tilde{\alpha}\rangle)| \geq C|(m-n)\tilde{\alpha}|_1 \geq C\tilde{\gamma}|m-n|^{-\tilde{\tau}}.$$

Hence,

$$\|F_{mn}\|_R \leq C^{-1}\tilde{\gamma}^{-1}|m-n|^{\tilde{\tau}} \cdot \|Z_{mn}\|_R.$$

□

Corollary C.1 *There exists an orthogonal matrix U with $U_{mn} \in \mathcal{A}_R$ real-analytic on $\mathbb{R}/\mathbb{Z}$, satisfying the shift condition, such that*

$$U^*(D + Z)U = D + \mathrm{diag}\{Z_{nn}\}_{n\in\mathbb{Z}^d} + Z_+,$$

where Z_+ is a symmetric matrix with $(Z_+)_{mn} \in \mathcal{A}_R$ real-analytic on $\mathbb{R}/\mathbb{Z}$, satisfying the shift condition. Moreover, for any $0 < \rho_+ < \rho$, if ε is sufficiently small,

$$\|(U - I_{\mathbb{Z}^d})_{mn}\|_R \leq cC^{-1}\tilde{\gamma}^{-1}(\rho - \rho_+)^{-(\tilde{\tau}+1)}\varepsilon e^{-\rho_+|m-n|},$$
$$\|(Z_+)_{mn}\|_R \leq \varepsilon^{\frac{3}{2}} e^{-\rho_+|m-n|}.$$

Proof. Let $U = e^F$. For $k \geq 1$, in view of

$$(F^k)_{mn} = \sum_{\substack{l_j\in\mathbb{Z}^d \\ j=1,\cdots,k-1}} F_{ml_1}F_{l_1l_2}\cdots F_{l_{k-1}n},$$

if ε is sufficiently small,

$$\|(F^k)_{mn}\|_R \le c(C^{-1}\tilde{\gamma}^{-1}(\rho-\rho_+)^{-(\tilde{\tau}+1)}\varepsilon)^k e^{-\rho_+|m-n|} \le \varepsilon^{\frac{2k}{3}} e^{-\rho_+|m-n|}.$$

Expand $e^{\pm F}$ as a power series, we can see

$$\|(e^{\pm F} - I_{\mathbb{Z}^d})_{mn}\|_R \le \varepsilon^{\frac{1}{2}} e^{-\rho_+|m-n|}.$$

Noting that

$$\begin{aligned} Z_+ &= e^{-F}(D+Z)e^F - D - \operatorname{diag}\{Z_{nn}\}_{n\in\mathbb{Z}^d} \\ &= \sum_{k=2}^{\infty} \frac{1}{k!}[\cdots[D, \underbrace{F]\cdots, F}_{k}] + \sum_{k=1}^{\infty} \frac{1}{k!}[\cdots[Z, \underbrace{F]\cdots, F}_{k}], \end{aligned}$$

we get

$$\|(Z_+)_{mn}\|_R \le \varepsilon^{\frac{3}{2}} e^{-\rho_+|m-n|}.$$

□

Let $V_+ = V + Z_{00}$. According to Lemma C.1, for any $0 < R_+ < R$, V_+ is a meromorphic function on $\mathcal{S}_{R_+}$, satisfying

$$|V_+(z) - V_+(z-a)| \ge \left(C - \frac{\varepsilon}{R - R_+}\right)|a|_1, \quad \forall a \in \mathbb{R}, \quad z \in \mathcal{S}_{R_+}.$$

Moreover, $z \in \mathcal{S}_{R_+}$ is the pole of V_+ if and only if it is the pole of V.

Back to the analyse to the operator L, let $V_0(z) = \tan \pi z$, $\varepsilon_0 = e^4\epsilon$, $\rho_0 = 4$, and choose any $R_0 > 0$, $1 < C_0 < \pi$. For $\nu = 1, 2, \cdots$, define the sequences:

$$\varepsilon_\nu = \varepsilon_0^{\left(\frac{3}{2}\right)^\nu}, \quad \rho_\nu = \frac{\rho_0}{2} + \frac{\rho_0}{2^{\nu+1}}, \quad R_\nu = \frac{R_0}{2} + \frac{R_0}{2^{\nu+1}}, \quad C_\nu = C - \sum_{j=0}^{\nu} \frac{\varepsilon_j}{R_j - R_{j+1}}.$$

According to Lemma C.2 and Corollary C.1, we can get the following iteration lemma:

Proposition C.1 *There exists $\epsilon_0 = \epsilon_0(\tilde{\alpha})$, such that for the linear operator L, the following holds if $0 < \epsilon < \epsilon_0$.*

For $\nu = 1, 2, \cdots$, there exists a meromorphic function V_ν on $\mathcal{S}_{R_\nu}$, satisfying

$$|V_\nu(z) - V_\nu(z-a)| \ge C_\nu |a|_1, \quad \forall a \in \mathbb{R}, \quad z \in \mathcal{S}_{R_\nu},$$

and an orthogonal matrix U_ν, with $(U_\nu)_{mn} \in \mathcal{A}_{R_\nu}$ real-analytic on $\mathbb{R}/\mathbb{Z}$, satisfying the shift condition, and

$$\|(U_\nu - I_{\mathbb{Z}})_{mn}\|_{R_\nu} \leq cC_{\nu-1}^{-1}\tilde{\gamma}^{-1}(\rho_{\nu-1} - \rho_\nu)^{-(\tilde{\tau}+1)}\varepsilon_{\nu-1}e^{-\rho_\nu|m-n|},$$

such that

$$U_\nu^* \cdots U_1^* L U_1 \cdots U_\nu = \operatorname{diag}\{V_\nu(x + \langle n, \tilde{\alpha}\rangle)\}_{n\in\mathbb{Z}^d} + Z_\nu, \quad \forall x \in \mathcal{X}.$$

Here Z_ν is symmetric, with $(Z_\nu)_{mn} \in \mathcal{A}_{R_\nu}$ real-analytic on $\mathbb{R}/\mathbb{Z}$, satisfying the shift condition, and

$$\|(Z_\nu)_{mn}\|_{R_\nu} \leq \varepsilon_\nu e^{-\rho_\nu|m-n|}.$$

According the iteration lemma above, when $\nu \to \infty$, $V_\nu \to \hat{V}$ holds on $\mathcal{S}_{\frac{R}{2}}$, and, in the sense of $\|\cdot\|_{\frac{R}{2}}$,

$$U_1 \cdots U_\nu \to U, \quad Z_\nu \to 0.$$

By a direct calculation,

$$\|(U - I_{\mathbb{Z}^d})_{mn}\|_{\frac{R}{2}} \leq cC^{-1}\tilde{\gamma}^{-1}\epsilon e^{-2|m-n|}.$$

附录四　Outline of the proof of Proposition 1.6

For any smooth function f defined on $\mathcal{I} \subset \mathbb{R}/\mathbb{Z}$, let $|f|_{C^j} := \max_{0\le k\le j} \sup_{x\in\mathcal{I}} \frac{1}{k!}|\partial_x^k f(x)|$.

The operator T in (1.6) can be viewed as a sum of two infinite-dimensional matrices, i.e.,diag$\{V(x+n\tilde{\alpha})\}_{n\in\mathbb{Z}} + \epsilon\Delta$ with Δ denoting the discrete Laplacian. It is natural to define an abstract normal form containing diag$\{V(x+n\tilde{\alpha})\}_{n\in\mathbb{Z}}$.

Definition D.1 *Given a symmetric matrix D, smoothly parametrized by $x \in \mathbb{R}/\mathbb{Z}$ and satisfying the shift condition*

$$D_{m+k,n+k}(x) = D_{mn}(x + k\tilde{\alpha}), \quad \forall k \in \mathbb{Z}, \tag{D.1}$$

where $\tilde{\alpha}$ is a Diophantine number, i.e., for some $\tilde{\gamma} > 0$ and $\tilde{\tau} > 1$,

$$|n\tilde{\alpha}|_1 \ge \frac{\tilde{\gamma}}{|n|^{\tilde{\tau}}}, \quad n \ne 0.$$

We say that D is in **normal form** *if the following conditions hold.*

(a) Short-range.

$$|D_{mn}|_{C^k} \le \begin{cases} Ce^{-\rho|m-n|}L^k, & |m-n| \le N \\ 0, & |m-n| > N \end{cases}, \quad \forall k \ge 0.$$

(b) Block diagonalization. *Fix any $x_* \in \mathbb{R}/\mathbb{Z}$. There exist an interval $\mathcal{I}$ centered in x_*, a disjoint decomposition $\bigcup_j \Lambda_j = \mathbb{Z}$ and a smooth orthogonal matrix Q on $\mathcal{I}$ such that*

- (b1) $\sharp\Lambda_j \le M$ *and* diam$\Lambda_j \le MN$ *for each j;*
- (b2) $\tilde{D} = Q^*DQ = \prod_j \tilde{D}_{\Lambda_j}(x)$, $\forall x \in \mathcal{I}$*;*
- (b3) $Q_{mn} \equiv 0$ *if $|m-n| > N$. Moreover, for all m, $Q_{mn} \not\equiv 0$ for at most M different n;*
- (b4) $|Q|_{C^k} \le L^k$, $\forall k \ge 0$.

(c) Eigenvalues. *There is a piecewise smooth function $E(x)$ such that for each j,*

$$\{E(x_* + n\tilde{\alpha})\}_{n\in\Lambda_j} \ \textit{are the eigenvalues of} \ \tilde{D}_{\Lambda_j}(x_*),$$

and there are sets $\Omega_j \supset \Lambda_j$ such that (c1) *for each* n*, if* $\inf_{l\in\Lambda_j} |E(x_* + l\tilde{\alpha}) - E(x_* + n\tilde{\alpha})| < \kappa$*, then*

$$x_* + n\tilde{\alpha} \in x_* + m\tilde{\alpha} + \frac{1}{2}(\mathcal{I} - x_*) \ \textit{for some}\ m \in \Omega_j,$$

$$Q(x)(\mathbb{R}^{\Lambda(n)}) \subset \mathbb{R}^{\Omega_j+n-m}, \quad \forall x \in \mathcal{I};$$

(c2) *the resultant*

$$u_{\Omega_j}(\varphi, x) = \mathrm{Res}\left(\det(D(x+\varphi)_{\Omega_j} - tI_{\Omega_j}), \det(D(x)_{\Omega_j} - tI_{\Omega_j})\right)^{1}$$

satisfies

$$|u_{\Omega_j}|_{C^k} < (4MC)^{2M^2} B^k, \quad \forall k \leq \tilde{s}M^2 + 1,^{2}$$

$$\max_{0\leq k\leq \tilde{s}M^2} \left| \frac{1}{\nu! B^k} \partial_\varphi^k u_{\Omega_j}(\varphi, x) \right| \geq \vartheta, \quad \forall \varphi, \forall x \in \mathbb{R}/\mathbb{Z};$$

(c3) $\sharp\Omega_j \leq M$ *and* $\mathrm{diam}\Omega_j \leq \left(\frac{1}{\lambda}\right)^{\tilde{\tau}+2}$*;*

(c4) *the intervals* $\{n\tilde{\alpha} + \mathcal{I}\}_{\mathrm{dist}(n,\Omega_j)<N}$ *are pairwise disjoint;*

(c5) *for each* $\varphi \in \mathcal{I}$*,* $u_{\Omega_j}(\varphi, x)$ *satisfies*

$$|u_{\Omega_j}|_{C^k} < (2MC)^{2M^2} B^k, \quad \forall k \leq \tilde{s}M^2 + 1,^{3}$$

$$\max_{0\leq k\leq \tilde{s}M^2} \left| \frac{1}{\nu! B^k} \partial_x^k u_{\Omega_j}(\varphi, x) \right| \geq \vartheta \left(\prod_{m,n\in\Omega_j} |\varphi + (m-n)\tilde{\alpha}|_1 \right), \quad \forall x \in \mathbb{R}/\mathbb{Z}.$$

Remark D.1 *Condition* (a) *implies an estimate of* D *in the operator norm on* $\ell^2(\mathbb{Z})$*:*

$$\|D\|_{C^k} \leq C\frac{e^\rho + 1}{e^\rho - 1} L^k \leq C\frac{4}{\rho} L^k, \quad \forall k \geq 0.$$

Consider the symmetric matrix $Z(x)$, smoothly parametrized by $x \in \mathbb{R}/\mathbb{Z}$, satisfying the shift condition, and

$$|Z_{mn}|_{C^k} < \varepsilon e^{-\varrho|m-n|} L^k, \quad \forall k \geq 0. \tag{D.2}$$

[1]The resultant of two monic polynomials P and Q is defined as the product $\mathrm{Res}(P, Q) = \prod_{\substack{P(x)=0\\Q(y)=0}} (x - y)$.

[2]The norm is with respect to the variable φ.

[3]The norm is with respect to the variable x.

Lemma D.1 (The inductive lemma of [16]) *Let D be in normal form on an interval $\mathcal{I} \subset \mathbb{R}/\mathbb{Z}$ with parameters C, L, ρ, M, N, κ, B, ϑ, λ, and let $a < g < h$ be numbers restricted by*

$$\frac{1}{\tilde{\tau} M^{3\tilde{s}M^3}} \leq a < \frac{g}{20\tilde{s}\tilde{\tau}M^4} < \frac{h}{100\tilde{s}^2\tilde{\tau}M^8}, \quad h \leq \frac{1}{5\tilde{s}M^{2\tilde{s}M^3}}.$$

Assume, as simplification, that

$$1 \leq B \leq L, \quad M \geq 8, \quad 1 < C < 2, \quad \rho, \kappa, \vartheta \leq 1.$$

Let Z be a symmetric matrix, smoothly parametrized on $\mathbb{R}/\mathbb{Z}$, satisfying the shift condition. Assume that

$$\lambda \leq |\mathcal{I}| \leq \vartheta/B,$$

$$|Z_{mn}|_{C^k} < \varepsilon e^{-\rho|m-n|} L^k, \quad k \geq 0.$$

If there is a constant $\Gamma = \Gamma(\tilde{\gamma}, \tilde{\tau}, \tilde{s}, M)$, super-exponentially decaying in M, such that

$$|\varepsilon| < \Gamma \left[\frac{\rho^{\tilde{\tau}} \kappa \vartheta \lambda^{\tilde{\tau}^2}}{L N^{\tilde{\tau}}} e^{-N\rho} \right]^{e^{e^{\tilde{s}M^4}}}, \tag{D.3}$$

then there is a smooth orthogonal matrix $\tilde{U}$, satisfying the shift condition, such that

$$|(\tilde{U} - I)_{mn}|_{C^k} < \varepsilon^{\frac{1}{2}} e^{-\rho'|m-n|} L'^k$$

and

$$\tilde{U}^*(D + Z)\tilde{U} = D' + Z',$$

with Z' a symmetric matrix, smoothly parametrized on $\mathbb{R}/\mathbb{Z}$, satisfying the shift condition, and D' in normal form on an interval $\mathcal{I}' \subset \mathcal{I}$, with parameters

$$\begin{array}{lll} C' = (1 + \varepsilon^{\frac{1}{2}})C, & L' = \varepsilon^{-h} L, & \rho' = \frac{1}{2}\rho, \\ \lambda' = 9^{-M'}\lambda, & M' = M^{\tilde{s}M^3}, & N' = \varepsilon^{-a}, \\ \kappa' = \varepsilon^h, & B' = L, & \vartheta' = \varepsilon^g L, \end{array}$$

and

$$2\lambda' \leq |\mathcal{I}'| \leq \varepsilon^g,$$

$$|Z'_{mn}|_{C^k} < \varepsilon^{\frac{1}{2}\varepsilon^{-a/2}} e^{-\rho'|m-n|} L'^k.$$

In addition,

$$|E(x_* + m\tilde{\alpha}) - E(x_* + n\tilde{\alpha})| < M'\frac{L}{\rho}\varepsilon^g, \quad \forall m \in \Lambda'(n),$$

$$Q'(x)(\mathbb{R}^{\Lambda'(n)}) \subset \sum_{m\in\Lambda'(n)} Q(x)(\mathbb{R}^{\Lambda(m)}), \quad \forall x \in \mathcal{I}',$$

D' is in normal form with the same parameters also on $x_* + \frac{1}{2}(\mathcal{I}' - x_*)$.

Finally, if $M \geq 2\tilde{\tau}$ then the closure of the sets

$$\{x_* + m\tilde{\alpha} : |E(x_* + m\tilde{\alpha}) - E(x_* + (m+n)\tilde{\alpha})| < 2M'\frac{L}{\rho}\varepsilon^g\}, \ \ \forall\, 4(1/\lambda)^{\tilde{\tau}+2} < |n| < M'N',$$

$$\{x_* + m\tilde{\alpha} : |E(x_* + m\tilde{\alpha}) - E(x_* + (m+n)\tilde{\alpha})| < 2\varepsilon^{\frac{1}{8}}\}, \ \ \forall\, M'N' < |n| < 4(1/\lambda')^{\tilde{\tau}+2}$$

are unions of, respectively, at most $\varepsilon^{-\frac{g}{5\tilde{s}M^2}}$ and $\varepsilon^{-M^4 g}$ many components, each component being of length, respectively, at most $\varepsilon^{\frac{g}{4\tilde{s}M^2}}$ and $\varepsilon^{2M^4 g}$.

For the detail of proof, which contains the construction of new blocks Λ'_i, i.e., the new equivalence relation on $\mathbb{Z}$, and the new orthogonal transformation Q', see Section 5 of Reference [16].

For $Z_0 = \epsilon\Delta$, we have

$$|(Z_0)_{mn}|_{C^k} < \varepsilon_0 e^{-\rho_0|m-n|} L_0^k.$$

with $\varepsilon_0 = e\epsilon$, $\rho_0 = 1$ and $L_0 = L$(see (1.7)). It has been proven by Eliasson in Section 6 of [16] that $D_0 = \mathrm{diag}\{V(x + n\tilde{\alpha})\}_{n\in\mathbb{Z}}$ is in normal form with $C_0 = C$, $L_0 = L$, any

$$M_0 \geq \max\left\{2^{\tilde{s}+4}C\frac{L^{\tilde{s}+1}((\tilde{s}+1)!)^2}{\tilde{\xi}},\ 2\tilde{\tau},\ 8\right\}, \quad N_0 \geq 1, \quad \rho = N_0^{-1},$$

and other suitable parameters κ_0, B_0, λ_0, ϑ_0.

For $\nu = 0, 1, 2, \cdots$, let $M_{\nu+1} = M_\nu^{\tilde{s}M_\nu^3}$, and

$$a_\nu = \frac{1}{\tilde{\tau}}\left(\frac{1}{M_\nu}\right)^{3\tilde{s}M_\nu^3}, \quad g_\nu = 20\tilde{s}\tilde{\tau}M_\nu^4 a_\nu, \quad h_\nu = \frac{1}{5\tilde{s}}\left(\frac{1}{M_\nu}\right)^{2\tilde{s}M_\nu^3}.$$

The other sequences can be defined as

$$\begin{aligned}
&\varepsilon_{\nu+1} = \varepsilon_\nu^{\frac{1}{2}\varepsilon_\nu^{-a_\nu/2}}, && C_{\nu+1} = (1+\varepsilon_\nu^{1/2})C_\nu, && L_{\nu+1} = \varepsilon_\nu^{-h_\nu}L_\nu,\\
&N_{\nu+1} = \varepsilon_\nu^{-a_\nu}, && \rho_{\nu+1} = \varepsilon_\nu^{a_\nu}, && \kappa_{\nu+1} = \varepsilon_\nu^{h_\nu},\\
&B_{\nu+1} = L_\nu, && \lambda_{\nu+1} = 9^{-M_\nu}\varepsilon_\nu^{g_\nu}, && \vartheta_{\nu+1} = \varepsilon_\nu^{g_\nu}L_\nu.
\end{aligned}$$

The inequality (D.3), about parameters at the ν^{th} step, has been verified in Section 6 of [16], so we can apply Lemma D.1 iteratively. For each $\nu \geq 0$, there is an orthogonal matrix $\tilde{U}_\nu$ satisfying the shift condition, such that

$$|(\tilde{U}_\nu - I_{\mathbb{Z}})_{mn}|_{C^k} < \varepsilon_\nu^{\frac{1}{2}} e^{-\frac{\rho_\nu}{2}|m-n|} L_{\nu+1}^k$$

and

$$(\tilde{U}_0 \cdots \tilde{U}_\nu)^*(D_0 + Z_0)(\tilde{U}_0 \cdots \tilde{U}_\nu) = D_{\nu+1} + Z_{\nu+1},$$

where $D_{\nu+1}$ is in normal form with parameters $C_{\nu+1}$, $L_{\nu+1}$, $\rho_{\nu+1}$, $M_{\nu+1}$, $N_{\nu+1}$, $\kappa_{\nu+1}$, $B_{\nu+1}$, $\vartheta_{\nu+1}$, $\lambda_{\nu+1}$, and

$$|(Z_{\nu+1})_{mn}|_{C^k} \leq \varepsilon_{\nu+1} e^{-\rho_{\nu+1}|m-n|} L_{\nu+1}^k.$$

Hence, in the operator norm $\|\cdot\|_{C^k}$,

$$\tilde{U}_0 \cdots \tilde{U}_\nu \to U, \quad Z_\nu \to 0, \quad D_\nu \to D_\infty.$$

Let $U_{\nu+1} = \tilde{U}_0 \cdots \tilde{U}_\nu$, by a simple calculation, we have

$$|(U_{\nu+1} - I_{\mathbb{Z}})_{mn}|_{C^k} < \varepsilon_0^{\frac{1}{2}} e^{-\frac{\rho_\nu}{2}|m-n|} L_{\nu+1}^k.$$

Clearly there is a uniform limit $E^\nu(x) \to E^\infty(x)$ which describes the spectrum of $D_\infty(x)$–it is the closure of the image of E^∞. Consider now the closure S_ν of the set of all x such that

$$|E_\infty(x) - E_\infty(x + n\tilde{\alpha})| < \frac{3}{2} M_{\nu+1} \frac{L_\nu}{\rho_\nu} \varepsilon_\nu^{g_\nu} \quad \text{for some} \;\; 4(1/\lambda_\nu)^{\tilde{\tau}+2} < |n| < M_{\nu+1}N_{\nu+1}$$

or

$$|E_\infty(x) - E_\infty(x + n\tilde{\alpha})| < \frac{3}{2} \varepsilon_\nu^{\frac{1}{8}} \quad \text{for some} \;\; M_{\nu+1}N_{\nu+1} < |n| < 4(1/\lambda_{\nu+1})^{\tilde{\tau}+2}.$$

According to the final statement of Lemma D.1, this set is of measure less than $c\varepsilon_\nu^{g_\nu/20\tilde{s}M_\nu^2}$. By Borel-Cantelli Lemma, we conclude that there is a full-measure subset $\tilde{\mathcal{X}}$ of $\mathbb{R}/\mathbb{Z}$ such that for any $x \in \tilde{\mathcal{X}}$, each $x + n\tilde{\alpha}$ will belong to only finitely many S_ν's. Choose $x = x_*$ of this sort, i.e., for all $n \in \mathbb{Z}$ there is a $\nu_0(n)$ such that $x_* + n\tilde{\alpha} \notin S_\nu$ for $\nu \geq \nu_0(n)$. Hence for such ν's,

$$|E^\nu(x_* + n\tilde{\alpha}) - E^\nu(x_* + n\tilde{\alpha} + m\tilde{\alpha})| \geq 2M_{\nu+1} \frac{L_\nu}{\rho_\nu} \varepsilon_\nu^{g_\nu}, \quad \forall\, 4(1/\lambda_\nu)^{\tilde{\tau}+2} < |m| < M_{\nu+1}N_{\nu+1},$$

$$|E^{\nu}(x_* + n\tilde{\alpha}) - E^{\nu}(x_* + n\tilde{\alpha} + m\tilde{\alpha})| \geq 2\varepsilon_{\nu}^{\frac{1}{8}}, \quad \forall\, M_{\nu+1}N_{\nu+1} < |m| < 4(1/\lambda_{\nu+1})^{\tilde{\tau}+2}.$$

This implies that $\Lambda^{\nu}(n) \subset [n - 4(1/\lambda_{\nu_0(n)})^{\tilde{\tau}+2},\ n + 4(1/\lambda_{\nu_0(n)})^{\tilde{\tau}+2}]$ for $\nu \geq \nu_0(n)$. The blocks $\Lambda^{\nu}(n)$ therefore become eventually stationary:

$$\Lambda^{\nu+1}(n) = \Lambda^{\nu}(n), \quad \forall \nu \geq \nu_0(n).$$

附录五　Proof of Lemma 2.1

For $|i|, |j|, |n|, |m| \le \kappa |\ln \epsilon|$, 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 (2.24), it is sufficient to show that

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

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

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

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

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

$$\epsilon^{\frac{1}{1200}} \le |\tan \pi(x + n\tilde{\alpha})| \le \left|\tan \pi \left(\frac{1}{2} - \epsilon^{\frac{1}{1200}}\right)\right| = \left|\tan \epsilon^{\frac{1}{1200}}\pi\right|^{-1} \le c\epsilon^{-\frac{1}{1200}}, \tag{E.1}$$

if ϵ 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

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

for ϵ sufficiently small.

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

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

Thus, we have

$$|V^0_{i,j,n,m}(x)| \ge \pi |(n - m)\tilde{\alpha}|_1 \ge \frac{\pi\tilde{\gamma}}{(2\kappa)^{\tilde{\tau}} |\ln \epsilon|^{\tilde{\tau}}} \ge \epsilon^{\frac{1}{1200}}. \tag{E.3}$$

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 (E.2). 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 E.1 *Given $|i|,|j|,|n|,|m|\le\kappa|\ln\epsilon|$. If ϵ 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 $|\det(B_1(x))| \geq \epsilon^{\frac{1}{120}}$, by (E.1) and (E.3), 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 $|\det(B_2(x))|, |\det(B_3(x))| \geq \epsilon^{\frac{1}{200}}$, which can be proved in the same way as above. □

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

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

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

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

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

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

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

$$\left|V^{(s)}(x)\right| \leq c\epsilon^{-\frac{s+1}{1200}}, \tag{E.4}$$

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

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

Thus the determinant of the 4×4 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 E.1.

We need to arrive at some transversality conditions, which are elaborated in Corollary $E.1$, by virtue of the following lemma .

Lemma E.2 (Proposition of appendix B in [6]) *Let $u^{(0)},\cdots,u^{(L-1)}$ be L independent vectors in $\mathbb{R}^L$ with $\|u^{(s)}\|_{\ell^1} \leq 1$. Let $v \in \mathbb{R}^L$ be an arbitrary vector, then there exists $s \in \{0, \cdots, L-1\}$, such that*

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

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 [6].

Corollary E.1 *Given $|i|, |j|, |n|, |m| \leq \kappa |\ln \epsilon|$, and $\{i, n\} \cap \{j, m\} = \emptyset$. If ϵ 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*

$$\left| V^{0(s)}_{i,j,n,m}(x) \right| \geq c\epsilon^{\frac{1}{60}}; \tag{E.5}$$

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

$$\left| V^{0(s)}_{i,j,n,m}(x) \right| \geq c\epsilon^{\frac{1}{100}}. \tag{E.6}$$

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 (E.4),

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

for $x \in \mathcal{X}_0$. Apply Lemma E.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 E.2 with $v = (2, -1, -1)$ and $v = (1, 1, -2)$ respectively, together with the corresponding conclusion Lemma E.1. □

From now on, we set the constant $c = 1$ in (E.5) and (E.6) for convenience. The proof of Lemma 2.1 ends with the following lemma.

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

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

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

$$|V^0_{i,j,n,m}(x)| \geq 2\epsilon^{\frac{1}{4}}, \quad x \in \mathcal{X}_\epsilon. \tag{E.7}$$

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

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

We only deal with the subcase (**S1**), with the others done similarly. By Corollary E.1, 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 (E.4), $A \leq c\epsilon^{-\frac{1}{240}}$.

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 $1 \leq s \leq 3$, $\left|V^{0(s)}_{i,j,n,m}(x_0)\right| \geq \epsilon^{\frac{1}{60}}$ by Corollary E.1. Let us say $s = 3$, which is considered as the most complex case, so $\left|V^{0(3)}_{i,j,n,m}(x_0)\right| \geq \epsilon^{\frac{1}{60}}$. Since for $x \in I$,

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

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

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

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

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

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

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

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

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

$$|V^{0}_{i,j,n,m}(x) - V^{0}_{i,j,n,m}(x_1')| = \left|V^{0(1)}_{i,j,n,m}(y)\right| \cdot |x - x_1'| \geq \epsilon^{\frac{1}{6}} \cdot 4\epsilon^{\frac{1}{12}} = 4\epsilon^{\frac{1}{4}}.$$

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 $|I_1'| \leq 4\epsilon^{\frac{1}{12}}$, so that if $x \in J_1' \setminus I_1'$, then $|V^{0}_{i,j,n,m}(x)| \geq 2\epsilon^{\frac{1}{4}}$. 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 $|I_k'| \leq 4\epsilon^{\frac{1}{12}}$, $k = 2, 3, 4$, such that if $x \in \bigcup_{k=1}^{4} (J_k' \setminus I_k')$, then

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

Hence, (E.7) holds on I after excluding a subset with measure less than $5\epsilon^{\frac{1}{15}}$ since ϵ 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 $|i|, |j|, |n|, |m| \leq \kappa|\ln\epsilon|$, the measure of the subset of parameters we exclude is less than $c\kappa^4|\ln\epsilon|^4 \cdot \epsilon^{\frac{1}{45}} < \epsilon^{\frac{1}{50}}$. □

附录六　Productions and publications

1、Zhiyan Zhao, Jiansheng Geng: Linearly stable quasi-periodic breathers in a class of random Hamiltonian systems. J. Dyn. Diff. Eqs, **23**, 961–997(2011).

2、Shiwen Zhang, Zhiyan Zhao: Diffusion bound and reducibility for discrete Schrödinger equations with tangent potential. Front. Math. China, **7(6)**, 1213–1235(2012).

3、Jiansheng Geng, Zhiyan Zhao: Quasi-periodic solutions for one-dimensional discrete nonlinear Schrödinger equations with tangent potential. Preprint.[1]

4、Jiansheng Geng, Jiangong You, Zhiyan Zhao: Localization in one-dimensional quasi-periodic nonlinear systems. Preprint.[2]

[1] It has been published: Siam. J. Math. Anal. **45(6)**, 3651–3689(2013).

[2] It has been published: Geom. And Func. Anal. **24**, 116–158(2014).

References

[1] Aizenman, M., Frierich, R., Hundertmark, D., Shankar, S.: Constructive fractional-moment criteria for localization in random operators. Physica A. **279**, 369–377(2000).

[2] Aizenman, M., Molchanov, S.: Localization at large disorder and at extreme energies: An elementary derivations. Commun. Math. Phys. **157**, 245–278(1993).

[3] Anderson, P.: Absence of diffusion in certain random lattices. Phys. Rev. **109**, 1492(1958).

[4] Aubry, S., André, G.: Analyticity breaking and Anderson localization in incommensurate lattices. Ann. Israel. Phys. Soc. **3**, 33(1980).

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

[6] 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)

[7] Bourgain, J.: Anderson localization for quasi-periodic lattice Schrödinger operators on $\mathbb{Z}^d$, d arbitrary. Geom. Funct. Anal. **17**, 682–706(2007).

[8] Bourgain, J., Goldstein, M.: On nonperturbative localization with quasi-periodic potential. Ann. Math. **152**, 835–879(2000).

[9] Bourgain, J., Goldstein, M., Schlag, W.: Anderson localization for Schrödinger operators on $\mathbb{Z}^2$ with quasi-periodic potential. Acta Math. **188**, 41–86(2002).

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

[11] Chulaevsky, V. A., Dinaburg, E. I.: Methods of KAM-theory for long-range quasiperiodic operators on $\mathbb{Z}^\nu$. Pure point spectrum. Commun. Math. Phys. **153**, 559–577(1993).

[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.: Discrete one-dimensional quasi-periodic Schrödinger operators with pure point spectrum. Acta. Math. **179**, 153–196(1997).

[17] Fröhlich, J., Martinelli, F., Scoppola, E., Spencer, T.: Constructive proof of localization in the Anderson tight binding model. Commun. Math. Phys. **101**, 21–46(1985).

[18] Fröhlich, J., Spencer, T.: Absence of diffusion in the Anderson tight binding model for large disorder or low energy. Commun. Math. Phys. **88**, 151–184(1983).

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

[20] Fröhlich, J., Spencer, T., Wittwer, P.: Localization for a class of one dimensional quasi-periodic Schrödinger operators. Commun. Math. Phys. **132**, 5–25(1990).

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

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

[23] Germinet, F.: Dynamical localization II with an application to the almost Mathieu operator. J. Stat. Phys. **95**, No.1-2, 273–286(1999).

[24] Germinet, F., De Bièvre, S.: Dynamical localization for discrete and continuous random Schrödinger operators. Commun. Math. Phys. **194**, 323–341(1998).

[25] Germinet, F., Jitomirskaya, S. Ya.: Strong dynamical localization for the almost Mathieu model. Rev. Math. Phys. **13**, 755–765(2001).

[26] Germinet, F., Klein, A.: Bootstrap multiscale analysis and localization in random media. Commun. Math. Phys. **222**, 415–448(2001).

[27] Gol'dsheid, Ya., Molchanov, S., Pastur, L.: A pure point spectrum of the stochastic one-dimensional Schrödinger operators. Funct. Anal. Appl. **11**, No.1, 1–8(1977).

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

[29] Harper, P. G.: Single band motion of conduction electrons in a uniform magnetic field. Proc. Phys. Soc. London Sect. A **68**, 874–878(1955).

[30] Hiramoto, H., Abe, S.: Dynamics of an Electron in Quasiperiodic Systems. II. Harper's Model. J. Phys. Soc. Jpn. **57** 1365–1371(1988).

[31] Jitomirskaya, S. Ya.: Anderson localization for the almost Mathieu equation: a non-perturbative proof. Commun. Math. Phys. **165**, 49–57(1994).

[32] Jitomirskaya, S. Ya.: Metal-insulator transition for the almost Mathieu operator. Annal. Math., **150**, 1159–1175(1999).

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

[35] Larcher, M., Dalfovo, F., Modugno, M.: Effects of interaction on the diffusion of atomic matter waves in one-dimensional quasiperiodic potentials. Phys. Rev. A **80**, 053606(2009).

[36] Pastur, L. A.: Spectral properties of disordered systems in the one-body approximation. Commun. Math. Phys. **75**, 179–196(1980).

[37] Pastur, L., Figotin, A.: Spectra of Random and Almost Periodic Operators. Springer(1992).

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

[39] Pöschel, J.: Small divisors with spatial structrue in infinite dimensional Hamiltonian systems. Comment. Math. Phys. **127**, 351–393(1990).

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

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

[42] Roati, G., D'Errico, C., Fallani, L., Fattori, M., Fort, C., Zaccanti, M., Modugno, G., Modugno, M., Inguscio, M,: Anderson localization of a non-interacting Bose-Einstein condensate. Nature **453**, 895–898(2008).

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

[44] Sinai, Ya. G.: Anderson localization for the one-dimensional difference Schrödinger operator with a quasi-periodic potential. J. Statist. Phys. **46**, 861–909(1987).

[45] Tcheremchantsev, S.: How to prove dynamical localization. Commun. Math. Phys. **221**, 27–56(2001).

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

[47] Vittot, M., Bellissard, J.: Invariant tori for an infinite lattice of coupled classical rotators. CPT-Marseille(1985).

[48] von Dreifus, H., Klein, A.: A new proof of localization in the Anderson tight binding model. Commun. Math. Phys. **124**, 285–299(1989).

E-mail address: zyqiao1985@gmail.com