Long-Time Anderson Localization for the Nonlinear Random Schrödinger Equation on $\mathbb{Z}^d$

Hongzi Cong[1] · Yunfeng Shi[2] · Xiaoqing Wu[1]

Abstract

In this paper, we prove the long-time Anderson localization for the nonlinear random Schrödinger equation on $\mathbb{Z}^d$ by using the Birkhoff normal form technique.

1 Introduction and Main Results

We study d-dimensional discrete nonlinear random Schrödinger equation (DRNLS)

$$\mathrm{i}\frac{\mathrm{d}q_j}{\mathrm{d}t} = v_j q_j + \epsilon_1 \sum_{|i-j|_1 = 1} q_i + \epsilon_2 \left|q_j\right|^2 q_j, \ i, j \in \mathbb{Z}^d, \tag{1.1}$$

where $\boldsymbol{v} = \left\{v_j\right\}_{j \in \mathbb{Z}^d}$ is a sequence of independent identically distributed (i.i.d.) random variables in $\mathcal{W} = [0, 1]^{\mathbb{Z}^d}$ (denote by mes $(\cdot)$ the standard probability measure on $\mathcal{W}$) with the uniform distribution, $\epsilon_1, \epsilon_2 \geq 0$ and $j = (j_1, \cdots, j_d)$ with $|j|_1 = \sum_{i=1}^{d} |j_i|$. In the case of $\epsilon_2 = 0$, the Eq. (1.1) becomes the celebrated Anderson model, which was first introduced by Anderson [2] to describe the motion of non-interacting electrons in crystal with impurities. The Eq. (1.1) with the nonlinearity is due to interactions of quantum particles which plays an essential role in both classical optics [20] and Bose–Einstein condensate [21] in the presence of disorder. In fact, it is one of the center themes in Anderson localization (namely, absence of diffusion) transitions theory that how does the interplay between the disorder and nonlinearity affect the above transitions. While the study of Anderson localization for the linear case has

✉ Yunfeng Shi
  yunfengshi@scu.edu.cn

  Hongzi Cong
  conghongzi@dlut.edu.cn

  Xiaoqing Wu
  wxq0216@mail.dlut.edu.cn

[1] School of Mathematical Sciences, Dalian University of Technology, Dalian 116024, China

[2] School of Mathematics, Sichuan University, Chengdu 610064, China

 Springer

attracted great attention over the years since the fundamental works of Fröhlich–Spencer [16] and Aizenman–Molchanov [1], much less is known bout the Anderson localization for DRNLS, particularly in multidimensional case (i.e., $d > 1$). In this paper, we prove the first long-time localization result for DRNLS on $\mathbb{Z}^d$ for arbitrary $d \geq 1$ by using the Birkhoff normal form technique.

We state our result more precisely. To this end, we need to give some notations. Consider the phase space

$$\ell_s^2(\mathbb{Z}^d, \mathbb{C}) = \left\{ q = (q_j)_{j \in \mathbb{Z}^d} \in \mathbb{C}^{\mathbb{Z}^d} : \|q\|_s^2 := \sum_{j \in \mathbb{Z}^d} \langle j \rangle^{2s} |q_j|^2 < \infty \right\}, \tag{1.2}$$

where $\langle j \rangle^2 := 1 + |j|_2^2$ and $|j|_2 = \sqrt{\sum_{i=1}^d |j_i|^2}$. Moreover, for a vector $w = (w_j)_{j \in \mathbb{Z}^d} \in \mathbb{C}^{\mathbb{Z}^d}$ and a given large $\mathcal{N} > 0$, we define the projection

$$\widetilde{\Pi}_{\mathcal{N}} w := \widetilde{w} = (\widetilde{w}_j)_{j \in \mathbb{Z}^d} \tag{1.3}$$

by $\widetilde{w}_j = w_j$ when $|j|_2 \leq \mathcal{N}$ and otherwise $\widetilde{w}_j = 0$. Let

$$\widehat{\Pi}_{\mathcal{N}} w := \widehat{w} = w - \widetilde{w}. \tag{1.4}$$

Now we give the following $(r, \mathcal{N})$-nonresonant conditions: for any positive integer $r \geq 2$ and a large $\mathcal{N} > 0$ depending on r and d,

$$|\langle \mathbf{k}, \mathbf{v} \rangle| \geq \frac{1}{\mathcal{N}^{20dr}}, \qquad 0 < |\mathbf{k}|_1 \leq r \tag{1.5}$$

and $\mathbf{k}$ satisfies

$$\left| \widehat{\mathbf{k}} \right|_1 \leq 1 \quad \text{or} \quad \left| \widehat{\mathbf{k}} \right|_1 = 2 \text{ with } \sum_{|j|_2 > \mathcal{N}} k_j = \pm 2. \tag{1.6}$$

Furthermore, define the resonant set $\mathcal{R}_{\mathbf{k}}$ by

$$\mathcal{R}_{\mathbf{k}} := \left\{ \mathbf{v} \in \mathcal{W} \,\middle|\, |\langle \mathbf{k}, \mathbf{v} \rangle| < \frac{1}{\mathcal{N}^{20dr}} \right\}. \tag{1.7}$$

Let

$$\mathcal{R} = \bigcup_{\substack{0 \neq \mathbf{k} \in \mathbb{Z}^{\mathbb{Z}^d}, |\mathbf{k}|_1 \leq r \\ \mathbf{k} \text{ satisfy } (1.6)}} \mathcal{R}_{\mathbf{k}} \tag{1.8}$$

and

$$\mathcal{R}^c = \mathcal{W} \setminus \mathcal{R}.$$

Then it is easy to see that any frequency $\mathbf{v} \in \mathcal{R}^c$ is $(r, \mathcal{N})$-nonresonant.

Let $\epsilon = \epsilon_1 + \epsilon_2$. We have

Theorem 1.1 *Let $r > 1, s - 320dr^2 \geq 2s_0 > d + 1$. Then there is some $\epsilon_* = \epsilon_*(r, s, s_0, d) > 0$ such that, for $0 < \epsilon \leq \epsilon_*$, there exists the nonresonant set $\mathcal{R}^c \subset \mathcal{W}$ satisfying*

$$mes\, (\mathcal{W} \setminus \mathcal{R}^c) \leq \epsilon^{\frac{2r}{s - 2s_0}}$$

with the following properties: For $\mathbf{v} \in \mathcal{R}^c$, if the initial state $q(0)$ satisfies $\|q(0)\|_s \leq \epsilon$, then

$$\|\widetilde{q}(t)\|_{\frac{s}{2}} \leq 4\epsilon, \quad \|\widehat{q}(t)\|_{s_0} \leq 4\epsilon^{r+1}$$

for any $t \leq \epsilon^{-\frac{r}{2s_0+1}}$ *with* $\mathcal{N} = \epsilon^{-\frac{2r}{s-2s_0}}$ *in the truncation* $\widetilde{q}(t)$.

Remark 1.1 Specially, the above result holds true for $q(0)$ with compact support. So for this initial data, the long-time Anderson localization occurs. Moreover, Faou–Grébert in [13] proved a subexponential long stability time interval for d-dimensional NLS in the analytic space by construct a partial normal form, where the terms with three high modes and above can not be eliminated and they are not small. In order to make these terms small, different weights have to be chosen for $q(0)$ and $q(t)$. Similarly, in this paper, one can not eliminate the terms with two high modes in opposite signs, to make these terms small, we had to choose different weights s, $\frac{s}{2}$ and s_0 for $q(0)$, $\widetilde{q}(t)$ and $\widehat{q}(t)$. However, the terms with three high modes and above or two high modes in opposite signs are all eliminated in [7], which means that the desired long time stability results can be achieved by choosing the same weight for $q(0)$ and $q(t)$.

Remark 1.2 We want to mention that the nonresonant set $\mathcal{R}^c$ of random variables does not depend on the choice of initial data $q(0)$.

Remark 1.3 In contrast, we would like to introduce the remarkable work of Bourgain–Wang [10], in which the quasi-periodic solutions were constructed for multi-dimensional DRNLS in the case of both large disorder and weak nonlinearity. Their proof is based on deep Craig–Wayne–Bourgain (CWB) method together with delicate spectral properties of the Anderson model. Very recently, Shi–Wang [22, 23] generalized the result of [10] to the multi-dimensional nonlinear quasi-periodic Schrödinger and wave equations via introducing new ideas of Diophantine approximation on manifolds [18] and Bourgain's geometric lemma [8]. A Birkhoff normal form type method for proving long-time Anderson localization was developed in the more recent work of Cong–Shi–Wang [11], in dealing with multi-dimensional nonlinear quasi-periodic NLS.

1.1 Previous Related Works

In the investigations of localization for nonlinear discrete models, Fröhlich–Spencer–Wayne [17] first showed that, with high probability and weak nonlinearity, any sup-exponentially localized initial state always stayed in a full dimensional KAM tori. Their proof is based on an extension of the KAM techniques. Later, if the initial state is small and is polynomially localized, Benettin–Fröhlich–Giorgilli [3] proved the propagation remains localized in a very long-time scale for some d-dimensional discrete nonlinear oscillation equation with i.i.d. Gaussian random potential by using Birkhoff normal form method. We would like to mention that Yuan [25] first applied the traditional KAM method based on normal form transformations to establish the existence of quasi-periodic solutions for some discrete nonlinear model in the absence of randomness. *These models of* [3, 17, 25] *do not contain the Laplacian term (namely, the third term in* (1.1)*), which is the main obstacle to achieve the localization for* (1.1) *via the KAM or normal form techniques. One of the main reasons may be that the discrete Laplacian has pure absolutely continuous spectrum, thus exhibits delocalization.* In this sense, Bourgain–Wang [10] made the first breakthrough and proved the existence of quasi-periodic solutions for some multi-dimensional DRNLS by using the CWB method based on the Lyapunov–Schmidt decomposition, the Nash–Moser iteration, multi-scale analysis of [16] an the semi-algebraic geometry theory of [6]. For the long-time localization, the most important result applying to very rough initial states was due to Wang–Zhang [24], in which they proved the first "truly" long-time localization for the 1-dimensional DRNLS by using the

Birkhoff normal form type method of Bourgain–Wang [9]. It is well-know that the Anderson localization holds for 1-dimensional Anderson model with arbitrary non-zero disorder. So along this line, Fishman–Krivolapov–Soffer [14] obtained the $O(\epsilon_2^{-2})$ long-time localization (concerning the second moment of the norm) under *only* weak nonlinearity assumption. Liu–Wang [19] proved the existence of quasi-periodic solutions for 1-dimensional DRNLS without requiring the large disorder condition, and the proof also works for multi-dimensional DRNLS in the regime of localization. Some results of [14] have been improved to a time scale of $O(\epsilon_2^{-A})$ for any $A \geq 2$ [15], but the proof is partly rigorous: in some parts it relies on conjectures tested numerically. Very recently, the result of [24] has been generalized to an exponential time scale by Cong–Shi–Zhang [12].

1.2 The Strategy of Proof

While we employ the Birkhoff normal form type methods in this paper, the technical details are different from that of [24], where a finite barrier is constructed to avoid the transfer energies between Fourier models. Instead, our method here is inspired by [5] in the PDE case, thus all Fourier models are involved. In fact, we also obtain a partial normal form around the origin. Compared with [5], one can not eliminate the terms with two high modes in opposite signs which are generated by the discrete Laplacian [see **W** in (2.11)]. Fortunately, we can use some ideas of Bernier–Faou–Grébert in [4] and the mass conservation to deal with these remainder terms. It turns out the frequency of the linear wave operator in [4] is unbounded. In contrast, the frequency considered in this paper is bounded, which leads to the need for a better estimation of the coefficients. As a result, the estimate (3.17) is of vital importance, which together with the so-called short-range property of the DRNLS guarantees the nonresonant conditions.

1.3 Structure of the Paper

In Sect. 2, we will introduce some notations and the normal form theory is stated. In Sect. 3, we will prove the normal form theorem. The proof of Theorem 1.1 is given in Sect. 4. The measure estimate concerning the nonresonant frequencies is established in Sect. 5. Two technical lemmas are given in Sect. 6.

2 Birkhoff Normal Form Theorem

2.1 Some Notations

In this subsection, we will first introduce some notations.

Define the scale of phase space

$$\mathcal{P}_s(\mathbb{C}) := l_s^2(\mathbb{Z}^d, \mathbb{C}) \oplus l_s^2(\mathbb{Z}^d, \mathbb{C}) \ni (q, \bar{q}) = \left(\left(q_j \right)_{j \in \mathbb{Z}^d}, \left(\bar{q}_j \right)_{j \in \mathbb{Z}^d} \right)$$

is endowed by the standard symplectic structure $-2\mathrm{i} \sum_{j \in \mathbb{Z}^d} \mathrm{d}q_j \wedge \mathrm{d}\bar{q}_j$, where the phase space $l_s^2(\mathbb{Z}^d, \mathbb{C})$ is defined in (1.2). Denote by $B_{\mathbb{C},s}(\delta)$ the open ball centered at the origin and of radius δ in $\mathcal{P}_s(\mathbb{C})$.

Let $H : \mathcal{P}_s \to \mathbb{C}$ be a Hamiltonian function given by

$$H(q, \bar{q}) = \sum_{\boldsymbol{m} \in \mathbb{N}^{\mathbb{Z}^d} \times \mathbb{N}^{\mathbb{Z}^d}} H(\boldsymbol{m}) q^{\boldsymbol{\beta}} \bar{q}^{\boldsymbol{\gamma}}, \tag{2.1}$$

where $\boldsymbol{\beta} = (\beta_j)_{j \in \mathbb{Z}^d} \in \mathbb{N}^{\mathbb{Z}^d}$, $\boldsymbol{\gamma} = (\gamma_j)_{j \in \mathbb{Z}^d} \in \mathbb{N}^{\mathbb{Z}^d}$, $\boldsymbol{m} = (\boldsymbol{\beta}, \boldsymbol{\gamma}) \in \mathbb{N}^{\mathbb{Z}^d} \times \mathbb{N}^{\mathbb{Z}^d}$ and $H(\boldsymbol{m})$ is the coefficient of the monomial

$$q^{\boldsymbol{\beta}} \bar{q}^{\boldsymbol{\gamma}} := \prod_{j \in \mathbb{Z}^d} q_j^{\beta_j} \bar{q}_j^{\gamma_j}. \tag{2.2}$$

Define the modulus and vector field of the Hamiltonian function H by

$$\|H\| = \sup_{\boldsymbol{m} \in \mathbb{N}^{\mathbb{Z}^d} \times \mathbb{N}^{\mathbb{Z}^d}} |H(\boldsymbol{m})|$$

and

$$X_H(q, \bar{q}) = -2\mathrm{i} \left(\frac{\partial H}{\partial \bar{q}}, -\frac{\partial H}{\partial q} \right)$$

respectively. Next, we introduce the notation support, diameter and degree:

$$\mathrm{supp}(\boldsymbol{\beta}) = \left\{ j \mid \beta_j \neq 0 \right\}, \ \mathrm{supp}(\boldsymbol{m}) = \left\{ j \mid \beta_j \neq 0 \text{ or } \gamma_j \neq 0 \right\},$$
$$\Delta(\boldsymbol{\beta}) = \mathrm{diam}\{\mathrm{supp}(\boldsymbol{\beta})\}, \ \Delta(\boldsymbol{m}) = \Delta(\boldsymbol{\beta} + \boldsymbol{\gamma}),$$
$$|\boldsymbol{\beta}|_1 = \sum_{j \in \mathrm{supp}(\boldsymbol{\beta})} \beta_j, \ |\boldsymbol{m}|_1 = |\boldsymbol{\beta} + \boldsymbol{\gamma}|_1.$$

If $\beta_j = \gamma_j$ for all $j \in \mathrm{supp}(\boldsymbol{m})$, then the monomial (2.2) is called *resonant*. Otherwise, it is called *nonresonant*. Moreover, we consider the Hamiltonian functions satisfying mass conservation in this paper, i.e., the Hamiltonian functions satisfy $|\boldsymbol{\beta}|_1 = |\boldsymbol{\gamma}|_1$.

Finally, for two Hamiltonian functions $H(q, \bar{q})$ and $F(q, \bar{q})$, we define their Poisson bracket by

$$\{H, F\} := -2\mathrm{i} \sum_{k \in \mathbb{Z}^d} \left(\frac{\partial H}{\partial q_k} \frac{\partial F}{\partial \bar{q}_k} - \frac{\partial H}{\partial \bar{q}_k} \frac{\partial F}{\partial q_k} \right).$$

Precisely, let

$$F(q, \bar{q}) = \sum_{\boldsymbol{n} \in \mathbb{N}^{\mathbb{Z}^d} \times \mathbb{N}^{\mathbb{Z}^d}} F(\boldsymbol{n}) q^{\boldsymbol{\eta}} \bar{q}^{\boldsymbol{\xi}},$$

where $\boldsymbol{n} = (\boldsymbol{\eta}, \boldsymbol{\xi})$. Thus, $\{H, F\}$ can be rewritten as

$$\{H, F\} = \sum_{\boldsymbol{\mu} \in \mathbb{N}^{\mathbb{Z}^d} \times \mathbb{N}^{\mathbb{Z}^d}} \{H, F\}(\boldsymbol{\mu}) q^{\boldsymbol{\alpha}} \bar{q}^{\boldsymbol{\zeta}},$$

where H is defined in (2.1), $\boldsymbol{\mu} = (\boldsymbol{\alpha}, \boldsymbol{\zeta})$ and

$$\{H, F\}(\boldsymbol{\mu}) = -2\mathrm{i} \sum_{k \in \mathbb{Z}^d} \left(\sum_{\boldsymbol{m}, \boldsymbol{n} \in \mathbb{N}^{\mathbb{Z}^d} \times \mathbb{N}^{\mathbb{Z}^d}}^{*} H(\boldsymbol{m}) F(\boldsymbol{n}) \left(\gamma_k \eta_k - \beta_k \xi_k \right) \right) \tag{2.3}$$

with $*$ is taken as

$$\alpha_j = \beta_j + \eta_j - 1, \ \zeta_j = \gamma_j + \xi_j - 1 \ \text{ for } \ j = k,$$
$$\alpha_j = \beta_j + \eta_j, \ \zeta_j = \gamma_j + \xi_j \ \text{ for } \ j \neq k.$$

For simplicity, we often write

$$\mathcal{P}_s \equiv \mathcal{P}_s(\mathbb{C}), \quad B_s(\delta) \equiv B_{\mathbb{C},s}(\delta) \quad \text{and} \quad H \equiv H(q, \bar{q}).$$

2.2 The Birkhoff Normal Form Theorem

In order to prove Theorem 1.1, we will construct a partial normal form in the case that the frequency v is nonresonant.

Next, we give the following Birkhoff normal form theorem.

Theorem 2.1 *Consider the Hamiltonian function*

$$H(q, \bar{q}) = N(q, \bar{q}) + P(q, \bar{q}) + Z(q, \bar{q}) \tag{2.4}$$

with the following form of

$$N(q, \bar{q}) = \frac{1}{2} \sum_{j \in \mathbb{Z}^d} v_j \left| q_j \right|^2, \tag{2.5}$$

$$P(q, \bar{q}) = \sum_{\substack{m \in \mathbb{N}^{\mathbb{Z}^d} \times \mathbb{N}^{\mathbb{Z}^d} \\ |m|_1 = 2, \Delta(m) = 1 \\ |\beta|_1 = |\gamma|_1}} P(m) q^\beta \bar{q}^\gamma \tag{2.6}$$

and

$$Z(q, \bar{q}) = \sum_{\substack{m \in \mathbb{N}^{\mathbb{Z}^d} \times \mathbb{N}^{\mathbb{Z}^d} \\ |m|_1 = 4, \Delta(m) = 0 \\ \beta = \gamma}} Z(m) q^\beta \bar{q}^\gamma.$$

Assume that P and Z satisfy

$$|P(m)|, |Z(m)| \le \frac{1}{2} \epsilon^{\frac{1}{4}(\Delta(m) + |m|_1 - 1)}, \quad \forall m \in \mathbb{N}^{\mathbb{Z}^d} \times \mathbb{N}^{\mathbb{Z}^d} \tag{2.7}$$

and

$$\|P\| \le \frac{\epsilon}{2}, \quad \|Z\| \le \frac{\epsilon}{4}. \tag{2.8}$$

Given any $r > 1$, $s - 320dr^2 \ge 2s_0 > d + 1$, there exists $\epsilon_* = \epsilon_*(r, s, s_0, d) > 0$. Then for each $0 < \epsilon < \epsilon_*$ and each $v \in \mathcal{R}^c$ with

$$\mathcal{N} = \epsilon^{-\frac{2r}{s - 2s_0}}, \tag{2.9}$$

there exists a symplectic map

$$\Phi : B_s(4\epsilon) \to B_s(5\epsilon),$$

such that

$$H \circ \Phi = N + W + V + Z + R, \tag{2.10}$$

where

$$W = \sum_{\substack{m \in \mathbb{N}^{\mathbb{Z}^d} \times \mathbb{N}^{\mathbb{Z}^d}, |m|_1 \le 4r + 1 \\ |\widehat{m}|_1 = 2, |\widehat{\beta}|_1 = |\widehat{\gamma}|_1 = 1}} W(m) q^\beta \bar{q}^\gamma, \tag{2.11}$$

$$V = \sum_{\substack{\boldsymbol{m} \in \mathbb{N}^{\mathbb{Z}^d} \times \mathbb{N}^{\mathbb{Z}^d}, |\boldsymbol{m}|_1 \leq 4r+1 \\ |\widehat{\boldsymbol{m}}|_1 > 2}} V(\boldsymbol{m}) q^{\beta} \bar{q}^{\gamma}, \tag{2.12}$$

$$Z = \sum_{\substack{\boldsymbol{m} \in \mathbb{N}^{\mathbb{Z}^d} \times \mathbb{N}^{\mathbb{Z}^d} \\ \beta = \gamma}} Z(\boldsymbol{m}) q^{\beta} \bar{q}^{\gamma}$$

and

$$\boldsymbol{R} = \mathcal{O}\left(\epsilon^{r+1}\right).$$

Moreover, the following estimates hold:

(i) the symplectic map Φ fulfills

$$\sup_{q \in B_s(4\epsilon)} \|q - \Phi(q)\|_s \leq \epsilon^{\frac{3}{2}}; \tag{2.13}$$

(ii) the Hamiltonian functions $\boldsymbol{W}$ and $\boldsymbol{V}$ satisfy

$$|\boldsymbol{W}(\boldsymbol{m})|, |\boldsymbol{V}(\boldsymbol{m})| \leq \epsilon^{\frac{1}{4}(\Delta(\boldsymbol{m}) + |\boldsymbol{m}|_1 - 1)}, \quad \forall \boldsymbol{m} \in \mathbb{N}^{\mathbb{Z}^d} \times \mathbb{N}^{\mathbb{Z}^d} \tag{2.14}$$

and

$$\|\boldsymbol{W}\|, \|\boldsymbol{V}\| \leq C(r)\epsilon; \tag{2.15}$$

(iii) the Hamiltonian function $\boldsymbol{R}$ satisfies

$$\|\boldsymbol{R}\| \leq C(r)\epsilon^{r+1} \mathcal{N}^{20dr^2}, \tag{2.16}$$

where $C(r)$ is a constant depending on r.

Proof The details of proof for the Theorem 2.1 will be given in the Sect. 3. □

3 The Proof of Birkhoff Normal Form Theorem

In this section, we will prove Theorem 2.1 by a finite step induction.

3.1 The First Step of Birkhoff Normal Form

At the first step $h = 1$, in view of (2.4), we rewrite H as

$$H_1 = N + P_1 + Z_1,$$

where

$$P_1 = P \quad \text{and} \quad Z_1 = Z. \tag{3.1}$$

Let

$$R_1 = \sum_{\substack{\boldsymbol{m} \in \mathbb{N}^{\mathbb{Z}^d} \times \mathbb{N}^{\mathbb{Z}^d} \\ |\boldsymbol{m}|_1 = 2, \Delta(\boldsymbol{m}) = 1 \\ \boldsymbol{m} \in \mathcal{M}}} P_1(\boldsymbol{m}) q^{\beta} \bar{q}^{\gamma}, \tag{3.2}$$

where

$$\mathcal{M} = \left\{ \boldsymbol{m} \in \mathbb{N}^{\mathbb{Z}^d} \times \mathbb{N}^{\mathbb{Z}^d} \, \Big| \, |\widehat{\boldsymbol{m}}|_1 \leq 1 \text{ or } |\widehat{\boldsymbol{m}}|_1 = |\widehat{\boldsymbol{\beta}}|_1 = 2 \text{ or } |\widehat{\boldsymbol{m}}|_1 = |\widehat{\boldsymbol{\gamma}}|_1 = 2 \right\}.$$

Next, we will desire to eliminate the terms R_1. Rewrite R_1 as

$$R_1 = \sum_{\boldsymbol{m} \in \mathbb{N}^{\mathbb{Z}^d} \times \mathbb{N}^{\mathbb{Z}^d}} R_1(\boldsymbol{m}) q^\beta \bar{q}^\gamma.$$

Following the standard Birkhoff normal form approach, we define

$$H_2 = H_1 \circ \Phi_{\chi_1}^1,$$

where

$$\Phi_{\chi_1}^1 : B_s(\delta_1) \to B_s(\delta_2)$$

is the time-1 map of χ_1 with

$$\delta_1 = 4\epsilon, \tag{3.3}$$

$$\chi_1 = \sum_{\boldsymbol{m} \in \mathcal{M}} \frac{R_1(\boldsymbol{m})}{-2\mathrm{i}\langle \boldsymbol{k}, \boldsymbol{v} \rangle} q^\beta \bar{q}^\gamma \tag{3.4}$$

and $\boldsymbol{k} = \boldsymbol{\gamma} - \boldsymbol{\beta}$. Then in view of (2.7), it suffices to eliminate the terms satisfying

$$\Delta(\boldsymbol{m}) + |\boldsymbol{m}|_1 \leq 4r + 1. \tag{3.5}$$

Since the frequency $\boldsymbol{v} \in \mathcal{R}^c$ satisfies $(r, \mathcal{N})$-nonresonant conditions (1.5). Thus in view of (1.5), (2.7), (2.8), (3.1), (3.2) and (3.4), we can obtain

$$\left| \frac{R_1(\boldsymbol{m})}{-2\mathrm{i}\langle \boldsymbol{k}, \boldsymbol{v} \rangle} \right| \leq \frac{1}{2} \epsilon^{\frac{1}{4}(\Delta(\boldsymbol{m}) + |\boldsymbol{m}|_1 - 1)} \mathcal{N}^{20dr}, \quad \forall \boldsymbol{m} \in \mathbb{N}^{\mathbb{Z}^d} \times \mathbb{N}^{\mathbb{Z}^d}$$

and

$$\|\chi_1\| \leq \frac{1}{2} \epsilon \mathcal{N}^{20dr}. \tag{3.6}$$

Moreover, in view of (2.9), (3.6) and Lemma 6.1, we have

$$\sup_{q \in B_s(4\epsilon)} \left\| q - \Phi_{\chi_1}^1(q) \right\|_s \leq \epsilon^{\frac{7}{4}}. \tag{3.7}$$

By (3.3) and (3.7), we have the radius

$$\delta_2 = \delta_1 + \epsilon^{\frac{7}{4}} = 4\epsilon + \epsilon^{\frac{7}{4}}.$$

Employing the Taylor series expansion, we get

$$H_2 = H_1 \circ \Phi_{\chi_1}^1 = N + \{N, \chi_1\} + P_1 + Z_1$$
$$+ \sum_{n=2}^{r-1} \frac{1}{n!} N^{(n)} + \sum_{n=1}^{r-1} \frac{1}{n!} P_1^{(n)} + \sum_{n=1}^{r-1} \frac{1}{n!} Z_1^{(n)} + \mathbf{R}_{2,r},$$

where

$$\mathbf{R}_{2,r} = \sum_{n=r}^{\infty} \frac{1}{n!} N^{(n)} + \sum_{n=r}^{\infty} \frac{1}{n!} P_1^{(n)} + \sum_{n=r}^{\infty} \frac{1}{n!} Z_1^{(n)} \tag{3.8}$$

and

$$\mathcal{Y}^{(0)} = \mathcal{Y}, \quad \mathcal{Y}^{(n)} = \left\{ \mathcal{Y}^{(n-1)}, \chi_1 \right\} \tag{3.9}$$

with $n \geq 1$, $\mathcal{Y} = N, P_1, Z_1$.

Note that χ_1 solves the so-called homological equation

$$R_1 + \{N, \chi_1\} = 0.$$

Now we turn to estimate the Poisson brackets $\mathcal{Y}^{(n)}$. Rewrite $\mathcal{Y}^{(n)}$ as

$$\mathcal{Y}^{(n)} = \sum_{\boldsymbol{\mu} \in \mathbb{N}^{\mathbb{Z}^d} \times \mathbb{N}^{\mathbb{Z}^d}} \mathcal{Y}^{(n)}(\boldsymbol{\mu}) q^{\boldsymbol{\alpha}} \bar{q}^{\boldsymbol{\xi}}.$$

In fact, we only need to estimate Poisson brackets $P_1^{(n)}$. We estimate the Poisson brackets $P_1^{(1)}$ firstly. In view of (2.3), for any $\boldsymbol{\mu} \in \mathbb{N}^{\mathbb{Z}^d} \times \mathbb{N}^{\mathbb{Z}^d}$, we know that

$$P_1^{(1)}(\boldsymbol{\mu}) = -2\mathrm{i} \sum_{k \in \mathbb{Z}^d} \left(\sum_{\boldsymbol{m},\boldsymbol{n} \in \mathbb{N}^{\mathbb{Z}^d} \times \mathbb{N}^{\mathbb{Z}^d}}^{*} \frac{P_1(\boldsymbol{m}) R_1(\boldsymbol{n})}{-2\mathrm{i}\langle \boldsymbol{k}, \boldsymbol{v} \rangle} (\gamma_k \eta_k - \beta_k \xi_k) \right). \tag{3.10}$$

Moreover, we note that the monomials corresponding to multi-index $\boldsymbol{\mu}$ satisfy

$$\Delta(\boldsymbol{\mu}) \leq \Delta(\boldsymbol{m}) + \Delta(\boldsymbol{n}), \ |\boldsymbol{\mu}|_1 = |\boldsymbol{m}|_1 + |\boldsymbol{n}|_1 - 2 \tag{3.11}$$

and the number of realizations of a fixed monomials $q^{\boldsymbol{\alpha}} \bar{q}^{\boldsymbol{\xi}}$ is bounded by

$$2^{|\boldsymbol{\mu}|_1}(\Delta(\boldsymbol{m}) \wedge \Delta(\boldsymbol{n})) < 2^{|\boldsymbol{\mu}|_1}(\Delta(\boldsymbol{m}) + \Delta(\boldsymbol{n})). \tag{3.12}$$

Then we have

$$\left| P_1^{(1)}(\boldsymbol{\mu}) \right| \leq \sum_{\boldsymbol{m},\boldsymbol{n} \in \mathbb{N}^{\mathbb{Z}^d} \times \mathbb{N}^{\mathbb{Z}^d}}^{*} \left(|P_1(\boldsymbol{m})| \, |R_1(\boldsymbol{n})| \, \mathcal{N}^{20dr} \left(\sum_{k \in \mathbb{Z}^d} |\gamma_k \eta_k - \beta_k \xi_k| \right) \right)$$

(in view of (1.5) and (3.10))

$$\leq \sum_{\boldsymbol{m},\boldsymbol{n} \in \mathbb{N}^{\mathbb{Z}^d} \times \mathbb{N}^{\mathbb{Z}^d}}^{*} \epsilon^{\frac{1}{4}(\Delta(\boldsymbol{m})+|\boldsymbol{m}|_1+\Delta(\boldsymbol{n})+|\boldsymbol{n}|_1-2)} \mathcal{N}^{20dr} (|\boldsymbol{m}|_1 + |\boldsymbol{n}|_1)^2$$

(in view of (2.7) and (3.1))

$$\leq \sum_{\boldsymbol{m},\boldsymbol{n} \in \mathbb{N}^{\mathbb{Z}^d} \times \mathbb{N}^{\mathbb{Z}^d}}^{*} C(r) \epsilon^{\frac{1}{8}} \epsilon^{\frac{1}{4}(\Delta(\boldsymbol{\mu})+|\boldsymbol{\mu}|_1-1)}$$

$$\left(\text{in view of (3.5), (3.11) and } \epsilon^{\frac{1}{8}} \mathcal{N}^{20dr} \leq 1 \right)$$

$$\leq C(r) 2^{|\boldsymbol{\mu}|_1}(\Delta(\boldsymbol{m}) + \Delta(\boldsymbol{n})) \epsilon^{\frac{1}{8}} \epsilon^{\frac{1}{4}(\Delta(\boldsymbol{\mu})+|\boldsymbol{\mu}|_1-1)} \quad \text{(in view of (3.12))}$$

$$\leq \epsilon^{\frac{1}{4}(\Delta(\boldsymbol{\mu})+|\boldsymbol{\mu}|_1-1)}$$

$$\left(\text{in view of } 2^{|\boldsymbol{\mu}|_1}(\Delta(\boldsymbol{m}) + \Delta(\boldsymbol{n})) \leq C(r) \text{ and } C(r) \epsilon^{\frac{1}{8}} \leq 1 \right),$$

where $C(r)$ is a constant depending on r. Similarly for $2 \leq n \leq r-1$, by induction we obtain

$$\left| \mathcal{Y}^{(n)}(\boldsymbol{\mu}) \right| \leq \epsilon^{\frac{1}{4}(\Delta(\boldsymbol{\mu})+|\boldsymbol{\mu}|_1-1)}, \quad \forall \boldsymbol{\mu} \in \mathbb{N}^{\mathbb{Z}^d} \times \mathbb{N}^{\mathbb{Z}^d}. \tag{3.13}$$

Furthermore, for any $1 \leq n \leq r - 1$ one also has

$$\left\| \mathcal{Y}^{(n)} \right\| \leq C(r) \epsilon^{n+1} \mathcal{N}^{20dnr}. \tag{3.14}$$

Finally we can rewrite H_2 as

$$H_2 = N + P_2 + \mathbf{Z}_2 + \mathbf{R}_{2,r},$$

where

$$P_2 = P_1 - R_1 + \sum_{n=2}^{r-1} \frac{1}{n!} N^{(n)} + \sum_{n=1}^{r-1} \frac{1}{n!} P_1^{(n)} + \sum_{n=1}^{r-1} \frac{1}{n!} Z_1^{(n)} \tag{3.15}$$

and

$$\mathbf{Z}_2 = Z_1. \tag{3.16}$$

In view of (2.7), (2.8), (3.1), (3.2), (3.8), (3.13)–(3.16), we know that Hamiltonian functions P_2, $\mathbf{Z}_2$ and $\mathbf{R}_{2,r}$ satisfy

$$|P_2(\boldsymbol{\mu})|, |\mathbf{Z}_2(\boldsymbol{\mu})| \leq \epsilon^{\frac{1}{4}(\Delta(\boldsymbol{\mu}) + |\boldsymbol{\mu}|_1 - 1)}, \quad \forall \boldsymbol{\mu} \in \mathbb{N}^{\mathbb{Z}^d} \times \mathbb{N}^{\mathbb{Z}^d},$$

$$\|P_2\|, \|\mathbf{Z}_2\| \leq C(r)\epsilon$$

and

$$\left\| \mathbf{R}_{2,r} \right\| \leq C(r) \epsilon^{r+1} \mathcal{N}^{20dr^2}.$$

3.2 Iterative Lemma

Now, we turn to the general iteration steps.

Lemma 3.1 (Iterative lemma) *For $2 \leq h \leq r$, consider the Hamiltonian function*

$$H_h = N + P_h + \mathbf{Z}_h + \mathbf{R}_{h,r},$$

where

$$P_h = \sum_{\boldsymbol{m} \in \mathbb{N}^{\mathbb{Z}^d} \times \mathbb{N}^{\mathbb{Z}^d}} P_h(\boldsymbol{m}) q^{\beta} \bar{q}^{\gamma},$$

$$\mathbf{Z}_h = \sum_{\substack{\boldsymbol{m} \in \mathbb{N}^{\mathbb{Z}^d} \times \mathbb{N}^{\mathbb{Z}^d} \\ \beta = \gamma}} \mathbf{Z}_h(\boldsymbol{m}) q^{\beta} \bar{q}^{\gamma}$$

and

$$\mathbf{R}_{h,r} = \mathcal{O}\left(\epsilon^{r+1}\right).$$

Suppose that the Hamiltonian functions P_h, $\mathbf{Z}_h$ and $\mathbf{R}_{h,r}$ satisfy the following estimates,

$$|P_h(\boldsymbol{m})|, |\mathbf{Z}_h(\boldsymbol{m})| \leq \epsilon^{\frac{1}{4}(\Delta(\boldsymbol{m}) + |\boldsymbol{m}|_1 - 1)}, \quad \forall \boldsymbol{m} \in \mathbb{N}^{\mathbb{Z}^d} \times \mathbb{N}^{\mathbb{Z}^d}, \tag{3.17}$$

$$\|P_h\|, \|\mathbf{Z}_h\| \leq C(r)\epsilon \tag{3.18}$$

and

$$\|\boldsymbol{R}_{h,r}\| \le C(r)\epsilon^{r+1}\mathcal{N}^{20dr^2},$$

where $C(r)$ is a constant depending on r, and assume that

$$R_h = \sum_{\substack{\boldsymbol{m}\in\mathbb{N}^{\mathbb{Z}^d}\times\mathbb{N}^{\mathbb{Z}^d} \\ \boldsymbol{m}\in\mathcal{M}}} P_h(\boldsymbol{m})q^\beta\bar{q}^\gamma \tag{3.19}$$

satisfying

$$\|R_h\| \le C(r)\epsilon^h\mathcal{N}^{20d(h-1)r}. \tag{3.20}$$

Let the radius δ_h as

$$\delta_h = 4\epsilon + \sum_{i=1}^{h-1}\epsilon^{1+\frac{3i}{4}}.$$

Then for each $\boldsymbol{v}\in\mathcal{R}^c$, there exists a symplectic transformation

$$\Phi^1_{\chi_h} : B_s(\delta_h) \to B_s(\delta_{h+1})$$

as the time-1 map of some Hamiltonian function χ_h such that

$$H_{h+1} = H_h \circ \Phi^1_{\chi_h} = N + P_{h+1} + Z_{h+1} + \boldsymbol{R}_{h+1,r},$$

where

$$P_{h+1} = \sum_{\boldsymbol{m}\in\mathbb{N}^{\mathbb{Z}^d}\times\mathbb{N}^{\mathbb{Z}^d}} P_{h+1}(\boldsymbol{m})q^\beta\bar{q}^\gamma,$$

$$\boldsymbol{Z}_{h+1} = \sum_{\substack{\boldsymbol{m}\in\mathbb{N}^{\mathbb{Z}^d}\times\mathbb{N}^{\mathbb{Z}^d} \\ \beta=\gamma}} \boldsymbol{Z}_{h+1}(\boldsymbol{m})q^\beta\bar{q}^\gamma$$

and

$$\boldsymbol{R}_{h+1,r} = \mathcal{O}\left(\epsilon^{r+1}\right).$$

Moreover, the following estimates hold:

(i) the symplectic transformation $\Phi^1_{\chi_h}$ fulfills

$$\sup_{q\in B_s(\delta_h)}\left\|q - \Phi^1_{\chi_h}(q)\right\|_s \le \epsilon^{1+\frac{3}{4}h} \tag{3.21}$$

and the radius δ_{h+1} satisfies

$$\delta_{h+1} \le 4\epsilon + \sum_{i=1}^{h}\epsilon^{1+\frac{3i}{4}};$$

(ii) the Hamiltonian functions P_{h+1} and Z_{h+1} satisfy

$$|P_{h+1}(\boldsymbol{m})|, |Z_{h+1}(\boldsymbol{m})| \le \epsilon^{\frac{1}{4}(\Delta(\boldsymbol{m})+|\boldsymbol{m}|_1-1)}, \quad \forall\boldsymbol{m}\in\mathbb{N}^{\mathbb{Z}^d}\times\mathbb{N}^{\mathbb{Z}^d} \tag{3.22}$$

and

$$\|P_{h+1}\|, \|Z_{h+1}\| \le C(r)\epsilon; \tag{3.23}$$

(iii) the Hamiltonian function $\boldsymbol{R}_{h+1,r}$ satisfies

$$\left\| \boldsymbol{R}_{h+1,r} \right\| \leq C(r) \epsilon^{r+1} \mathcal{N}^{20dr^2}. \tag{3.24}$$

Proof In view of (3.19), we rewrite R_h as

$$R_h = \sum_{\boldsymbol{m} \in \mathbb{N}^{\mathbb{Z}^d} \times \mathbb{N}^{\mathbb{Z}^d}} R_h(\boldsymbol{m}) q^{\beta} \bar{q}^{\gamma}. \tag{3.25}$$

Following the standard Birkhoff normal form approach, we define

$$H_{h+1} = H_h \circ \Phi^1_{\chi_h},$$

where $\Phi^1_{\chi_h}$ is the time-1 map of χ_h with

$$\chi_h = \sum_{\boldsymbol{m} \in \mathcal{M}} \frac{R_h(\boldsymbol{m})}{-2\mathrm{i}\langle \boldsymbol{k}, \boldsymbol{v} \rangle} q^{\beta} \bar{q}^{\gamma}. \tag{3.26}$$

Note that $\boldsymbol{k}$ satisfies (3.5) and the frequency $\boldsymbol{v}$ satisfies $(r, \mathcal{N})$-nonresonant conditions (1.5). Thus in view of (1.5), (3.17)–(3.20), (3.26) and following from the proof of (3.6) and (3.7) respectively, we can obtain

$$\|\chi_h\| \leq C(r) \epsilon^h \mathcal{N}^{20dhr}$$

and the estimate (3.21). Furthermore one has

$$\delta_{h+1} = 4\epsilon + \sum_{i=1}^{h} \epsilon^{1+\frac{3i}{4}}.$$

By Taylor series, we obtain

$$\begin{aligned}
H_{h+1} = H_h \circ \Phi^1_{\chi_h} &= N + \{N, \chi_h\} + P_h + \mathbf{Z}_h \\
&+ \sum_{n=2}^{r-1} \frac{1}{n!} N^{(n)} + \sum_{n=1}^{r-1} \frac{1}{n!} P_h^{(n)} + \sum_{n=1}^{r-1} \frac{1}{n!} \mathbf{Z}_h^{(n)} + \boldsymbol{R}_{h+1,r},
\end{aligned}$$

where

$$\boldsymbol{R}_{h+1,r} = \sum_{n=r}^{\infty} \frac{1}{n!} N^{(n)} + \sum_{n=r}^{\infty} \frac{1}{n!} P_h^{(n)} + \sum_{n=r}^{\infty} \frac{1}{n!} \mathbf{Z}_h^{(n)} \tag{3.27}$$

and here

$$\mathcal{Y}^{(0)} = \mathcal{Y}, \quad \mathcal{Y}^{(n)} = \left\{ \mathcal{Y}^{(n-1)}, \chi_h \right\}$$

with $n \geq 1$, $\mathcal{Y} = N, P_h, \mathbf{Z}_h$.

For any $1 \leq n \leq r-1$, we can get

$$\left| \mathcal{Y}^{(n)}(\boldsymbol{\mu}) \right| \leq \epsilon^{\frac{1}{4}(\Delta(\boldsymbol{\mu}) + |\boldsymbol{\mu}|_1 - 1)}, \quad \forall \boldsymbol{\mu} \in \mathbb{N}^{\mathbb{Z}^d} \times \mathbb{N}^{\mathbb{Z}^d} \tag{3.28}$$

and

$$\left\| \mathcal{Y}^{(n)} \right\| \leq C(r) \epsilon^{1+nh} \mathcal{N}^{20dnhr}, \tag{3.29}$$

which follows the proof of (3.13) and (3.14) respectively.

Finally, write

$$H_{h+1} = N + P_{h+1} + \mathbf{Z}_{h+1} + \mathbf{R}_{h+1,r},$$

where

$$P_{h+1} = P_h - R_h - Z_h + \sum_{n=2}^{r-1} \frac{1}{n!} N^{(n)} + \sum_{n=1}^{r-1} \frac{1}{n!} P_h^{(n)} + \sum_{n=1}^{r-1} \frac{1}{n!} \mathbf{Z}_h^{(n)} \qquad (3.30)$$

with

$$Z_h = \sum_{\substack{\boldsymbol{m} \in \mathbb{N}^{\mathbb{Z}^d} \times \mathbb{N}^{\mathbb{Z}^d} \\ \boldsymbol{\beta} = \boldsymbol{\gamma}}} P_h(\boldsymbol{m}) q^{\boldsymbol{\beta}} \bar{q}^{\boldsymbol{\gamma}} \qquad (3.31)$$

and

$$\mathbf{Z}_{h+1} = \mathbf{Z}_h + Z_h. \qquad (3.32)$$

Rewrite P_{h+1} and $\mathbf{Z}_{h+1}$ as

$$P_{h+1} = \sum_{\boldsymbol{m} \in \mathbb{N}^{\mathbb{Z}^d} \times \mathbb{N}^{\mathbb{Z}^d}} P_{h+1}(\boldsymbol{m}) q^{\boldsymbol{\beta}} \bar{q}^{\boldsymbol{\gamma}}$$

and

$$\mathbf{Z}_{h+1} = \sum_{\substack{\boldsymbol{m} \in \mathbb{N}^{\mathbb{Z}^d} \times \mathbb{N}^{\mathbb{Z}^d} \\ \boldsymbol{\beta} = \boldsymbol{\gamma}}} \mathbf{Z}_{h+1}(\boldsymbol{m}) q^{\boldsymbol{\beta}} \bar{q}^{\boldsymbol{\gamma}}.$$

In view of (3.17), (3.18), (3.25), (3.27)–(3.32), we know that the Hamiltonian function P_{h+1}, $\mathbf{Z}_{h+1}$ and $\mathbf{R}_{h+1,r}$ satisfy (3.22)–(3.24). Then we finish the proof of Lemma 3.1. $\qquad\square$

3.3 The Proof of Theorem 2.1

Proof We will finish the proof of Theorem 2.1 by using Iterative Lemma 3.1. Let

$$\Phi = \Phi_{\chi_1}^1 \circ \Phi_{\chi_2}^1 \circ \cdots \circ \Phi_{\chi_r}^1. \qquad (3.33)$$

Then based on Lemma 3.1, $H_{r+1} := H \circ \Phi$ has the form of

$$H_{r+1} = N + P_{r+1} + \mathbf{Z}_{r+1} + \mathbf{R}_{r+1,r},$$

where

$$P_{r+1} = \sum_{\boldsymbol{m} \in \mathbb{N}^{\mathbb{Z}^d} \times \mathbb{N}^{\mathbb{Z}^d}} P_{r+1}(\boldsymbol{m}) q^{\boldsymbol{\beta}} \bar{q}^{\boldsymbol{\gamma}},$$

$$\mathbf{Z}_{r+1} = \sum_{\substack{\boldsymbol{m} \in \mathbb{N}^{\mathbb{Z}^d} \times \mathbb{N}^{\mathbb{Z}^d} \\ \boldsymbol{\beta} = \boldsymbol{\gamma}}} \mathbf{Z}_{r+1}(\boldsymbol{m}) q^{\boldsymbol{\beta}} \bar{q}^{\boldsymbol{\gamma}}$$

and

$$\mathbf{R}_{r+1,r} = \mathcal{O}\left(\epsilon^{r+1}\right).$$

In fact, P_{r+1} can be rewritten as

$$P_{r+1} = \mathbf{W} + \mathbf{V} + Z_r + \mathbf{R}_r,$$

where

$$\mathbf{W} = \sum_{\substack{\boldsymbol{m}\in\mathbb{N}^{\mathbb{Z}^d}\times\mathbb{N}^{\mathbb{Z}^d},|\boldsymbol{m}|_1\leq 4r+1 \\ |\widehat{\boldsymbol{m}}|_1=2,|\widehat{\boldsymbol{\beta}}|_1=|\widehat{\boldsymbol{\gamma}}|_1=1}} P_{r+1}(\boldsymbol{m})q^{\beta}\bar{q}^{\gamma},$$

$$\mathbf{V} = \sum_{\substack{\boldsymbol{m}\in\mathbb{N}^{\mathbb{Z}^d}\times\mathbb{N}^{\mathbb{Z}^d},|\boldsymbol{m}|_1\leq 4r+1 \\ |\widehat{\boldsymbol{m}}|_1>2}} P_{r+1}(\boldsymbol{m})q^{\beta}\bar{q}^{\gamma},$$

$$Z_r = \sum_{\substack{\boldsymbol{m}\in\mathbb{N}^{\mathbb{Z}^d}\times\mathbb{N}^{\mathbb{Z}^d} \\ \beta=\gamma}} P_{r+1}(\boldsymbol{m})q^{\beta}\bar{q}^{\gamma}$$

and

$$\mathbf{R}_r = \mathcal{O}\left(\epsilon^{r+1}\right)$$

with

$$\|\mathbf{R}_r\| \leq C(r)\epsilon^{r+1}\mathcal{N}^{20dr^2}.$$

Finally, let

$$\mathbf{Z} = \mathbf{Z}_{r+1} + Z_r \quad \text{and} \quad \mathbf{R} = \mathbf{R}_{r+1,r} + \mathbf{R}_r.$$

Then we obtain that

$$H_{r+1} = N + \mathbf{W} + \mathbf{V} + \mathbf{Z} + \mathbf{R},$$

where $\mathbf{W}$, $\mathbf{V}$ and $\mathbf{R}$ satisfy (2.14)–(2.16) by using Lemma 3.1. Moreover, recalling the definition of Φ [see (3.33)] and by the inequality (3.21) for $h \geq 1$, we have

$$\sup_{q\in B_s(4\epsilon)} \|q - \Phi(q)\|_s \leq \epsilon^{\frac{3}{2}}.$$

$\square$

4 Proof the Main Theorem

Firstly, we write (1.1) as a Hamiltonian equation

$$i\dot{q}_j = 2\frac{\partial H}{\partial \bar{q}_j},$$

where

$$H(q,\bar{q}) = \frac{1}{2}\sum_{j\in\mathbb{Z}^d} v_j |q_j|^2 + \frac{\epsilon}{2}\sum_{j\in\mathbb{Z}^d}\sum_{|i-j|_1=1} q_i\bar{q}_j + \frac{\epsilon}{4}\sum_{j\in\mathbb{Z}^d} |q_j|^4. \tag{4.1}$$

We rewrite H as

$$H = N + P + Z,$$

where

$$N = \frac{1}{2} \sum_{j \in \mathbb{Z}^d} v_j \left| q_j \right|^2,$$

$$P = \sum_{\substack{m \in \mathbb{N}^{\mathbb{Z}^d} \times \mathbb{N}^{\mathbb{Z}^d} \\ |m|_1 = 2, \Delta(m) = 1 \\ |\beta|_1 = |\gamma|_1}} P(m) q^\beta \bar{q}^\gamma \tag{4.2}$$

and

$$Z = \sum_{\substack{m \in \mathbb{N}^{\mathbb{Z}^d} \times \mathbb{N}^{\mathbb{Z}^d} \\ |m|_1 = 4, \Delta(m) = 0 \\ \beta = \gamma}} Z(m) q^\beta \bar{q}^\gamma \tag{4.3}$$

with

$$|P(m)| \leq \frac{\epsilon}{2} \quad \text{and} \quad |Z(m)| \leq \frac{\epsilon}{4}. \tag{4.4}$$

In view of (4.2)–(4.4), we can obtain the Hamiltonian functions P and Z satisfy

$$|P(m)|, |Z(m)| \leq \frac{1}{2} \epsilon^{\frac{1}{4}(\Delta(m) + |m|_1 - 1)}, \quad \forall m \in \mathbb{N}^{\mathbb{Z}^d} \times \mathbb{N}^{\mathbb{Z}^d} \tag{4.5}$$

and

$$\|P\| \leq \frac{\epsilon}{2}, \quad \|Z\| \leq \frac{\epsilon}{4}. \tag{4.6}$$

Based on the formulas (4.5) and (4.6), we know that the Hamiltonian H in (4.1) satisfies assumptions conditions (2.7) and (2.8) in Theorem 2.1. Moreover, in view of Sect. 5, we know that $(r, \mathcal{N})$-nonresonant conditions (1.5) hold. Thus, all assumptions in Theorem 2.1 hold.

Next, we prove the main Theorem 1.1 by applying Theorem 2.1.

By Theorem 2.1, we know that for any $v \in \mathcal{R}^c$, there exists a normalizing transformation

$$\Phi : B_s(4\epsilon) \rightarrow B_s(5\epsilon)$$

such that $\Phi(q(t)) = q'(t)$. By the initial state $\|q(0)\|_s \leq \epsilon$ in Theorem 1.1 and (2.13), for any small enough $\epsilon > 0$, one has

$$\left\| q'(0) \right\|_s \leq \|q(0)\|_s + \|q(0) - \Phi(q(0))\|_s \leq 2\epsilon.$$

Thus we have

$$\left\| \widetilde{q'}(0) \right\|_{\frac{s}{2}} \leq \left\| q'(0) \right\|_s \leq 2\epsilon \tag{4.7}$$

and

$$\left\| \widehat{q'}(0) \right\|_{s_0} \leq \mathcal{N}^{s_0 - s} \left\| \widehat{q'}(0) \right\|_s < 2\epsilon^{2r+1} \tag{4.8}$$

by using (2.9) and (6.2) in Lemma 6.2.

Define

$$t^* = \inf \left\{ t \geq 0 : \left\| \widetilde{q'}(t) \right\|_{\frac{s}{2}} = 3\epsilon \text{ or } \left\| \widehat{q'}(t) \right\|_{s_0} = 3\epsilon^{r+1} \right\}, \tag{4.9}$$

then for all $0 \leq t \leq t^*$ one has

$$\left\| \widetilde{q'}(t) \right\|_{\frac{s}{2}} \leq 3\epsilon \quad \text{and} \quad \left\| \widehat{q'}(t) \right\|_{s_0} \leq 3\epsilon^{r+1}. \tag{4.10}$$

Next, we will prove that

$$t^* \geq \epsilon^{-\frac{r}{2s_0+1}}. \tag{4.11}$$

On the contrary, assume that $t^* < \epsilon^{-\frac{r}{2s_0+1}}$, we will construct the contradiction in two parts.

(1) Estimate $\left\|\widetilde{q'}(t^*)\right\|_{\frac{s}{2}} < 3\epsilon$ when $t^* < \epsilon^{-\frac{r}{2s_0+1}}$.

For any $0 \leq t \leq t^*$, define

$$\widetilde{F}(q'(t)) = \left\|\widetilde{q'}(t)\right\|_{\frac{s}{2}}^2.$$

By Newton–Leibniz formula, we have

$$\widetilde{F}\left(q'(t)\right) - \widetilde{F}\left(q'(0)\right) = \int_0^t \dot{\widetilde{F}}\left(q'(z)\right) dz. \tag{4.12}$$

Note that

$$\dot{\widetilde{F}}\left(q'(t)\right) = \left\{H \circ \Phi, \widetilde{F}\right\}\left(q'(t)\right) = \left\{\mathbf{W}, \widetilde{F}\right\}\left(q'(t)\right) + \left\{\mathbf{V}, \widetilde{F}\right\}\left(q'(t)\right) + \left\{\mathbf{R}, \widetilde{F}\right\}\left(q'(t)\right),$$

where the last equality uses (2.10) in Theorem 2.1 and $\left\{N + \mathbf{Z}, \widetilde{F}\right\}\left(q'(t)\right) = 0$. Then we have

$$\left|\int_0^{t^*} \dot{\widetilde{F}}\left(q'(t)\right) dt\right|$$

$$\leq \int_0^{t^*} \left(\left|\left\{\mathbf{W}, \widetilde{F}\right\}\left(q'(t)\right)\right| + \left|\left\{\mathbf{V}, \widetilde{F}\right\}\left(q'(t)\right)\right| + \left|\left\{\mathbf{R}, \widetilde{F}\right\}\left(q'(t)\right)\right|\right) dt. \tag{4.13}$$

In view of (4.7), (4.12), (4.13) and the Triangle inequality, we obtain

$$\left|\widetilde{F}\left(q'(t^*)\right)\right| \leq 4\epsilon^2 + 3\epsilon \int_0^{t^*} \left(\left\|\widetilde{X}_{\mathbf{W}}(q'(t))\right\|_{\frac{s}{2}} + \left\|\widetilde{X}_{\mathbf{V}}(q'(t))\right\|_{\frac{s}{2}} + \left\|\widetilde{X}_{\mathbf{R}}(q'(t))\right\|_{\frac{s}{2}}\right) dt \tag{4.14}$$

by using $\left\|\widetilde{q'}(t)\right\|_{\frac{s}{2}} \leq 3\epsilon$ in (4.10). Hence we only need to estimate $\left\|\widetilde{X}_{\mathbf{W}}(q'(t))\right\|_{\frac{s}{2}}$, $\left\|\widetilde{X}_{\mathbf{V}}(q'(t))\right\|_{\frac{s}{2}}$ and $\left\|\widetilde{X}_{\mathbf{R}}(q'(t))\right\|_{\frac{s}{2}}$.

In fact, we have

$$\left\|\widetilde{X}_{\mathbf{W}}\left(q'(t)\right)\right\|_{\frac{s}{2}} \leq \mathcal{N}^{\frac{s-2s_0}{2}} \left\|\widetilde{X}_{\mathbf{W}}\left(q'(t)\right)\right\|_{s_0} \quad \text{(in view of (6.1) in Lemma 6.2)}$$

$$\leq C(r, s_0)\mathcal{N}^{\frac{s-2s_0}{2}}\epsilon \left\|\widetilde{q'}(t)\right\|_{s_0} \left\|\widehat{q'}(t)\right\|_{s_0}^2$$

$$\text{(in view of (2.11), (2.15) and Lemma 6.1)}$$

$$\leq C(r, s_0)\mathcal{N}^{\frac{s-2s_0}{2}}\epsilon \left\|\widetilde{q'}(t)\right\|_{\frac{s}{2}} \left\|\widehat{q'}(t)\right\|_{s_0}^2$$

$$\leq \epsilon^{r+\frac{3}{2}}, \tag{4.15}$$

where the last inequality uses (2.9), (4.10) and $C(r, s_0)\epsilon^{\frac{1}{2}} \leq 1$. We also have

$$\left\|\widetilde{X}_{\mathbf{V}}\left(q'(t)\right)\right\|_{\frac{s}{2}} \leq \mathcal{N}^{\frac{s-2s_0}{2}} \left\|\widetilde{X}_{\mathbf{V}}\left(q'(t)\right)\right\|_{s_0} \quad \text{(in view of (6.1) in Lemma 6.2)}$$

$$\leq C(r, s_0)\mathcal{N}^{\frac{s-2s_0}{2}}\epsilon \left\|\widehat{q'}(t)\right\|_{s_0}^3$$

$$\text{(in view of (2.12), (2.15) and Lemma 6.1)}$$

$$\leq \epsilon^{r+\frac{3}{2}}, \tag{4.16}$$

where the last inequality uses (2.9), $\left\|\widehat{q'}(t)\right\|_{s_0} \leq 3\epsilon^{r+1}$ in (4.10) and $C(r, s_0)\epsilon^{\frac{1}{2}} \leq 1$. Moreover, we have

$$\left\|\widetilde{X}_{\mathbf{R}}(q'(t))\right\|_{\frac{s}{2}} \leq \epsilon^{r+\frac{3}{2}}, \tag{4.17}$$

where one applies $\mathbf{R} = \mathcal{O}\left(\epsilon^{r+1}\right)$ in Theorem 2.1 with $r^* := r + 1$ in place of r. Then in view of (4.15)–(4.17), we get

$$\left\|\widetilde{X}_{\mathbf{W}}(q'(t))\right\|_{\frac{s}{2}} + \left\|\widetilde{X}_{\mathbf{V}}(q'(t))\right\|_{\frac{s}{2}} + \left\|\widetilde{X}_{\mathbf{R}}(q'(t))\right\|_{\frac{s}{2}} \leq 3\epsilon^{r+\frac{3}{2}}. \tag{4.18}$$

Thus, in view of (4.14) and (4.18), when $t^* < \epsilon^{-\frac{r}{2s_0+1}}$, one has

$$\left|\widetilde{F}\left(q'(t^*)\right)\right| < 4\epsilon^2 + 9\epsilon^{-\frac{r}{2s_0+1}}\epsilon^{r+\frac{5}{2}} \leq 9\epsilon^2,$$

so we get

$$\left\|\widetilde{q'}(t^*)\right\|_{\frac{s}{2}} < 3\epsilon. \tag{4.19}$$

(2) Estimate $\left\|\widehat{q'}(t^*)\right\|_{s_0} < 3\epsilon^{r+1}$ when $t^* < \epsilon^{-\frac{r}{2s_0+1}}$.
For any $0 \leq t \leq t^*$, define

$$\widehat{F}(q'(t)) = \left\|\widehat{q'}(t)\right\|_{s_0}^2. \tag{4.20}$$

Similar to the proof process in the first part, we only need to estimate $\left|\left\{\mathbf{W}, \widehat{F}\right\}\left(q'(t)\right)\right|$, $\left|\left\{\mathbf{V}, \widehat{F}\right\}\left(q'(t)\right)\right|$ and $\left|\left\{\mathbf{R}, \widehat{F}\right\}\left(q'(t)\right)\right|$.

Firstly, for any $0 \leq t \leq t^*$, we estimate $\left|\left\{\mathbf{V}, \widehat{F}\right\}\left(q'(t)\right)\right|$. Note that

$$\left|\left\{\mathbf{V}, \widehat{F}\right\}\left(q'(t)\right)\right| \leq \left\|\widehat{X}_{\mathbf{V}}\left(q'(t)\right)\right\|_{s_0} \left\|\widehat{q'}(t)\right\|_{s_0} \leq 3\epsilon^{r+1}\left\|\widehat{X}_{\mathbf{V}}\left(q'(t)\right)\right\|_{s_0},$$

where the last inequality uses $\left\|\widehat{q'}(t)\right\|_{s_0} \leq 3\epsilon^{r+1}$ in (4.10). Since

$$\begin{aligned}
\left\|\widehat{X}_{\mathbf{V}}\left(q'(t)\right)\right\|_{s_0} &\leq C(r, s_0)\epsilon\left\|\widetilde{q'}(t)\right\|_{s_0}\left\|\widehat{q'}(t)\right\|_{s_0}^2 \\
&\quad \text{(in view of (2.12), (2.15) and Lemma 6.1)} \\
&\leq C(r, s_0)\epsilon\left\|\widetilde{q'}(t)\right\|_{\frac{s}{2}}\left\|\widehat{q'}(t)\right\|_{s_0}^2 \\
&\leq \epsilon^{2r+\frac{3}{2}} \quad \text{(by using (4.10) and } C(r, s_0)\epsilon^{\frac{1}{2}} \leq 1\text{)}.
\end{aligned} \tag{4.21}$$

Hence, we have

$$\left|\left\{\mathbf{V}, \widehat{F}\right\}\left(q'(t)\right)\right| \leq 3\epsilon^{3r+\frac{5}{2}}. \tag{4.22}$$

Similarly, we have

$$\left|\left\{\mathbf{R}, \widehat{F}\right\}\left(q'(t)\right)\right| \leq \left\|\widehat{X}_{\mathbf{R}}\left(q'(t)\right)\right\|_{s_0}\left\|\widehat{q'}(t)\right\|_{s_0} \leq 3\epsilon^{3r+\frac{5}{2}}, \tag{4.23}$$

where one applies $\mathbf{R} = \mathcal{O}\left(\epsilon^{r+1}\right)$ in Theorem 2.1 with $r^* := 2r + 2$ in place of r.

Next, for any $0 \leq t \leq t^*$, we estimate $\left|\left\{\mathbf{W}, \widehat{F}\right\}\left(q'(t)\right)\right|$. In view of (2.11) in Theorem 2.1, we have

$$\left|\left\{\mathbf{W}, \widehat{F}\right\}(q'(t))\right| \leq \sum_{i=2}^{4r+1}\left|\left\{\mathbf{W}_i, \widehat{F}\right\}(q'(t))\right|, \tag{4.24}$$

where

$$\mathbf{W}_i = \sum_{\substack{\boldsymbol{m} \in \mathbb{N}^{\mathbb{Z}^d} \times \mathbb{N}^{\mathbb{Z}^d}, |\boldsymbol{m}|_1 = i \\ |\widehat{\boldsymbol{m}}|_1 = 2, |\widehat{\boldsymbol{\beta}}|_1 = |\widehat{\boldsymbol{\gamma}}|_1 = 1}} \mathbf{W}_i(\boldsymbol{m}) q^{\boldsymbol{\beta}} \bar{q}^{\boldsymbol{\gamma}}. \tag{4.25}$$

Hence, we only need to estimate $\left| \left\{ \mathbf{W}_i, \widehat{F} \right\} (q'(t)) \right|$. By (4.25), we know

$$\left| \widehat{\boldsymbol{\beta}} \right|_1 = |\widehat{\boldsymbol{\gamma}}|_1 = 1.$$

Then there exist $a_1, a_2 \in \mathbb{Z}^d$ and $|a_1|_2, |a_2|_2 > \mathcal{N}$ such that

$$\left| \beta_{a_1} \right| = \left| \gamma_{a_2} \right| = 1.$$

Thus the Hamiltonian $\mathbf{W}_i$ can be rewritten as

$$\mathbf{W}_i = \sum_{\substack{a_1, a_2 \in \mathbb{Z}^d \\ |a_1|_2, |a_2|_2 > \mathcal{N}}} B_{a_1 a_2} \left(\widetilde{q}' \right) q'_{a_1} \bar{q}'_{a_2}, \tag{4.26}$$

where

$$B_{a_1 a_2} \left(\widetilde{q}' \right) = \sum_{\substack{\widetilde{\boldsymbol{m}} \in \mathbb{N}^{\mathbb{Z}^d} \times \mathbb{N}^{\mathbb{Z}^d} \\ |\widetilde{\boldsymbol{m}}|_1 = i - 2}} B_{a_1 a_2}(\widetilde{\boldsymbol{m}}) \widetilde{q}'^{\widetilde{\boldsymbol{\beta}}} \widetilde{\bar{q}}'^{\widetilde{\boldsymbol{\gamma}}}$$

and

$$\left| B_{a_1 a_2}(\widetilde{\boldsymbol{m}}) \right| \leq C(r)\epsilon$$

by using (2.15). Hence, for any $|a_1|_2, |a_2|_2 > \mathcal{N}$, we obtain

$$\left| B_{a_1 a_2} \left(\widetilde{q}' \right) \right| \leq C(r, s)\epsilon \left\| \widetilde{q}'(t) \right\|_{\frac{s}{2}}^{i-2} \leq C(r, s)\epsilon, \tag{4.27}$$

the last inequality is based on $2 \leq i \leq 4r + 1$ and $\left\| \widetilde{q}'(t) \right\|_{\frac{s}{2}} \leq 3\epsilon$ in (4.10), where constant $C(r, s)$ depends on r and s. Moreover, in view of (3.5), we have

$$\left| |a_1|_2 - |a_2|_2 \right| \leq \sqrt{d(4r + 1)^2}. \tag{4.28}$$

Then one has

$$
\left|\left\{\mathbf{W}_i, \widehat{F}\right\}(q'(t))\right|
$$

$$
= \left|\sum_{\substack{a_1,a_2\in\mathbb{Z}^d \\ |a_1|_2,|a_2|_2>\mathcal{N}}} B_{a_1a_2}\left(\widetilde{q'}\right)\left(\langle a_1\rangle^{2s_0}-\langle a_2\rangle^{2s_0}\right)q'_{a_1}\bar{q}'_{a_2}\right| \quad \text{(in view of (4.26))}
$$

$$
\leq C(r,s)\epsilon \sum_{\substack{a_1,a_2\in\mathbb{Z}^d \\ |a_1|_2,|a_2|_2>\mathcal{N}}} 2s_0\left|\langle a_1\rangle-\langle a_2\rangle\right|\left(\langle a_1\rangle^{2s_0-1}+\langle a_2\rangle^{2s_0-1}\right)\left|q'_{a_1}\bar{q}'_{a_2}\right|
$$

(in view of (4.27))

$$
\leq C(r,s,s_0,d)\epsilon \sum_{\substack{a_1,a_2\in\mathbb{Z}^d \\ |a_1|_2,|a_2|_2>\mathcal{N}}} \langle a_1\rangle^{s_0-1}\langle a_2\rangle^{s_0}\left|q'_{a_1}\bar{q}'_{a_2}\right| \quad \text{(in view of (4.28))}
$$

$$
\leq \left\|\widehat{q'}(t)\right\|_{s_0-1}\left\|\widehat{q'}(t)\right\|_{s_0}
$$

(by using Cauchy–Schwarz inequality and $C(r,s,s_0,d)\epsilon \leq 1$)

$$
\leq \left\|\widehat{q'}(t)\right\|_0^{\frac{1}{s_0}}\left\|\widehat{q'}(t)\right\|_{s_0}^{2-\frac{1}{s_0}} \quad \text{(by using Hölder inequality)}, \tag{4.29}
$$

where $C(r,s,s_0,d)$ is a constant depending on r, s, s_0 and d. Now we estimate $\left\|\widehat{q'}(t)\right\|_0$ and $\left\|\widehat{q'}(t)\right\|_{s_0}$. On one hand, we have

$$
\left\|\widehat{q'}(t)\right\|_{s_0}^{2-\frac{1}{s_0}} \leq \left(3\epsilon^{r+1}\right)^{\left(2-\frac{1}{s_0}\right)} \tag{4.30}
$$

by using $\left\|\widehat{q'}(t)\right\|_{s_0} \leq 3\epsilon^{r+1}$ in (4.10). On the other hand, note that the Hamiltonian $\mathbf{W}_i$ satisfies $\left|\widehat{\boldsymbol{\beta}}\right|_1 = \left|\widehat{\boldsymbol{\gamma}}\right|_1 = 1$, then we have

$$
\left\{\mathbf{W}_i, \left\|\widehat{q'}(t)\right\|_0^2\right\} = 0. \tag{4.31}
$$

Hence, in view of (4.20) with $s_0 = 0$, (4.31) and the Triangle inequality, we can obtain

$$
\left\|\widehat{q'}(t)\right\|_0^2 \leq \left\|\widehat{q'}(0)\right\|_0^2 + \int_0^t \left(\left|\left\{\mathbf{V}, \left\|\widehat{q'}(z)\right\|_0^2\right\}(q'(z))\right| + \left|\left\{\mathbf{R}, \left\|\widehat{q'}(z)\right\|_0^2\right\}(q'(z))\right|\right) \mathrm{d}z
$$

$$
\leq \left\|\widehat{q'}(0)\right\|_{s_0}^2 + \int_0^t \left(\left\|\widehat{X}_\mathbf{V}(q'(z))\right\|_{s_0} + \left\|\widehat{X}_\mathbf{R}(q'(z))\right\|_{s_0}\right)\left\|\widehat{q'}(z)\right\|_{s_0} \mathrm{d}z
$$

$$
\leq 6(1+t)\epsilon^{3r+\frac{5}{2}} \quad \text{(in view of (4.8), (4.10), (4.21) and (4.23))}. \tag{4.32}
$$

Thus, in view of (4.29), (4.30) and (4.32), we obtain

$$
\left|\left\{\mathbf{W}_i, \widehat{F}\right\}(q'(t))\right| \leq \left(6(1+t)\epsilon^{3r+\frac{5}{2}}\right)^{\frac{1}{2s_0}}\left(3\epsilon^{r+1}\right)^{\left(2-\frac{1}{s_0}\right)}, \tag{4.33}
$$

then in view of (4.24) and (4.33), we obtain

$$
\left|\left\{\mathbf{W}, \widehat{F}\right\}(q'(t))\right| \leq C(r)\left(6(1+t)\epsilon^{3r+\frac{5}{2}}\right)^{\frac{1}{2s_0}}\left(3\epsilon^{r+1}\right)^{\left(2-\frac{1}{s_0}\right)}. \tag{4.34}
$$

Consequently, it is easy to get

$$\left| \int_0^{t^*} \hat{F}\left(q'(t)\right) dt \right| \leq \int_0^{t^*} \left(\left| \{\mathbf{W}, \widehat{F}\}\left(q'(t)\right) \right| + \left| \{\mathbf{V}, \widehat{F}\}\left(q'(t)\right) \right| + \left| \{\mathbf{R}, \widehat{F}\}\left(q'(t)\right) \right| \right) dt$$

$$\leq 6t^* \epsilon^{3r+\frac{5}{2}} + C(r)t^* \left(6(1+t^*)\epsilon^{3r+\frac{5}{2}} \right)^{\frac{1}{2s_0}} \left(3\epsilon^{r+1} \right)^{\left(2-\frac{1}{s_0}\right)}$$

(by using (4.22), (4.23) and (4.34))

$$< \frac{1}{2}\epsilon^{2r+2}, \tag{4.35}$$

where the last equality uses $t^* < \epsilon^{-\frac{r}{2s_0+1}}$.

Thus, in view of (4.8), (4.35) and the Triangle inequality, when $t^* < \epsilon^{-\frac{r}{2s_0+1}}$ we obtain

$$\left| \widehat{F}(q'(t^*)) \right| < 9\epsilon^{2r+2},$$

i.e.

$$\left\| \widehat{q'}(t^*) \right\|_{s_0} < 3\epsilon^{r+1}. \tag{4.36}$$

To sum up, we know that (4.19), (4.36) and (4.9) are contradictory. Hence we can conclude that (4.11) holds.

Finally, going back to the original variables $q(t)$, for any $0 \leq t \leq t^*$, we have

$$\|\widetilde{q}(t)\|_{\frac{s}{2}} \leq \left\| \widetilde{q'}(t) \right\|_{\frac{s}{2}} + \|\widetilde{q}(t) - \Phi(\widetilde{q}(t))\|_{\frac{s}{2}}$$

$$\leq 3\epsilon + \|q(t) - \Phi(q(t))\|_s \quad \left(\text{in view of } \left\| \widetilde{q'}(t) \right\|_{\frac{s}{2}} \leq 3\epsilon \text{ in (4.10)} \right)$$

$$\leq 4\epsilon \quad \text{(in view of (2.13))}$$

and

$$\|\widehat{q}(t)\|_{s_0} \leq \left\| \widehat{q'}(t) \right\|_{s_0} + \|\widehat{q}(t) - \Phi(\widehat{q}(t))\|_{s_0}$$

$$\leq 3\epsilon^{r+1} + \mathcal{N}^{\frac{2s_0-s}{2}} \|\widehat{q}(t) - \Phi(\widehat{q}(t))\|_{\frac{s}{2}}$$

$$\text{(in view of } \|\widehat{q}(t)\|_{s_0} \leq 3\epsilon^{r+1} \text{ in (4.10) and (6.2) in Lemma 6.2)}$$

$$\leq 3\epsilon^{r+1} + \mathcal{N}^{\frac{2s_0-s}{2}} \|q(t) - \Phi(q(t))\|_s$$

$$\leq 4\epsilon^{r+1} \quad \text{(in view of (2.9) and (2.13))}.$$

5 Measure Estimate

In this section, we will estimate the measure of the resonant set $\mathcal{R}$ defined in (1.8), i.e., we will show

$$\text{mes}\,(\mathcal{R}) \leq \epsilon^{\frac{2r}{s-2s_0}}. \tag{5.1}$$

Proof For the convenience of calculation, rewritten the resonant set $\mathcal{R}$ defined in (1.8) as

$$\mathcal{R} = \bigcup_{i=0}^{2} \mathcal{R}_i,$$

where

$$\mathcal{R}_i = \bigcup_{\substack{0 \neq \mathbf{k} \in \mathbb{Z}^{\mathbb{Z}^d} \\ |\widehat{\mathbf{k}}|_1 = i}} \mathcal{R}_{\mathbf{k}}, \quad i = 0, 1$$

and

$$\mathcal{R}_2 = \bigcup_{\substack{0 \neq \mathbf{k} \in \mathbb{Z}^{\mathbb{Z}^d} \\ |\widehat{\mathbf{k}}|_1 = 2 \text{ and } \sum_{|j|_2 > \mathcal{N}} k_j = \pm 2}} \mathcal{R}_{\mathbf{k}}.$$

In view of (3.5), we obtain

$$\Delta(\mathbf{k}) + |\mathbf{k}|_1 \leq 4r + 1. \tag{5.2}$$

Now it suffices to prove that the estimate (5.1) holds, which will be given in the following three cases.

Case 1: $|\widehat{\mathbf{k}}|_1 = 0$.
In view of (1.7), one has

$$\text{mes }(\mathcal{R}_{\mathbf{k}}) \leq \frac{1}{\mathcal{N}^{20dr}}.$$

Hence, we have

$$\text{mes }(\mathcal{R}_0) \leq \sum_{\substack{0 \neq \mathbf{k} \in \mathbb{Z}^{\mathbb{Z}^d} \\ |\widehat{\mathbf{k}}|_1 = 0}} \text{mes }(\mathcal{R}_{\mathbf{k}})$$

$$\leq \frac{1}{\mathcal{N}^{20dr}} \cdot (2\mathcal{N} + 1)^{d(4r+1)} \quad \text{(in view of (5.2))}$$

$$\leq \frac{1}{3\mathcal{N}}. \tag{5.3}$$

Case 2: $|\widehat{\mathbf{k}}|_1 = 1$.
Then one has

$$\widetilde{\mathbf{k}} \neq \mathbf{0}$$

by using the mass conservation, and

$$\langle \mathbf{k}, \mathbf{v} \rangle = \langle \widetilde{\mathbf{k}}, \mathbf{v} \rangle \pm v_{a_1},$$

where $a_1 \in \mathbb{Z}^d$ and $|a_1|_2 > \mathcal{N}$.
In view of (5.2), we know that

$$|a_1|_2 \leq \mathcal{N} + \sqrt{d(4r+1)^2}.$$

Then following the proof of (5.3), we get

$$\text{mes }(\mathcal{R}_1) \leq \frac{1}{3\mathcal{N}}. \tag{5.4}$$

Case 3: $|\widehat{\mathbf{k}}|_1 = 2$ **and** $\sum_{|j|_2 > \mathcal{N}} k_j = \pm 2$.
Then we have

$$\widetilde{\mathbf{k}} \neq \mathbf{0}$$

by using the mass conservation, and $\langle \mathbf{k}, \mathbf{v} \rangle$ is either equal to

$$\langle \widetilde{\mathbf{k}}, \mathbf{v} \rangle + v_{a_1} + v_{a_2}$$

or

$$\langle \widetilde{\boldsymbol{k}}, \boldsymbol{v} \rangle - v_{a_1} - v_{a_2},$$

where $a_1, a_2 \in \mathbb{Z}^d$ and $|a_1|_2, |a_2|_2 \geq \mathcal{N}$.

In view of (5.2), we know that

$$|a_1|_2, |a_2|_2 \leq \mathcal{N} + \sqrt{d(4r+1)^2}.$$

Then following the proof of (5.3), we get

$$\text{mes}\,(\mathcal{R}_2) \leq \frac{1}{3\mathcal{N}}. \qquad (5.5)$$

Finally, in view of (5.3)–(5.5), we obtain

$$\text{mes}\,(\mathcal{R}) \leq \frac{1}{\mathcal{N}}.$$

In view of (2.9), we have

$$\text{mes}\,(\mathcal{R}) \leq \epsilon^{\frac{2r}{s-2s_0}}.$$

Thus, one finishes the proof of (5.1). $\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\square$

6 Appendix

Lemma 6.1 *Assume that $s > d + 1$ and let $\chi = \sum_{i=2}^{4r+1} \chi_i$ be a polynomial of order at least 2 at the origin, then for all $\delta \leq \max\{\|\chi_1\|, \cdots, \|\chi_{4r+1}\|\}^{-1}$, the time 1-flow Φ_χ^1 of X_χ is well defined on a ball $B_s(\delta)$ with values in $B_s(4\delta)$ and we have*

$$\left\| \Phi_\chi^1(q) - q \right\|_s \leq \left\| X_\chi \right\|_s \leq C(r,s) \max_{2 \leq i \leq 4r+1} \|\chi_i\| \|q\|_s \leq C(r,s) \max_{2 \leq i \leq 4r+1} \|\chi_i\| \delta,$$

where $C(r,s)$ is a constant depending on r and s.

Proof The proof of this lemma is similar to the proof of Lemma 1 in [4]. $\qquad\qquad\square$

Lemma 6.2 *For any a vector $w = (w_j)_{j \in \mathbb{Z}^d} \in \mathbb{C}^{\mathbb{Z}^d}$, a given large $\mathcal{N} > 0$ and $s - 2s_0 > d+1$, we have*

$$\|\widetilde{w}\|_{\frac{s}{2}} \leq \mathcal{N}^{\frac{s-2s_0}{2}} \|\widetilde{w}\|_{s_0} \qquad (6.1)$$

and

$$\|\widehat{w}\|_{s_0} \leq \mathcal{N}^{\frac{2s_0-s}{2}} \|\widehat{w}\|_{\frac{s}{2}}. \qquad (6.2)$$

Proof Firstly, we prove that (6.1) holds. In view of definition (1.2) and (1.3), we have

$$\|\widetilde{w}\|_{\frac{s}{2}} = \left(\sum_{\substack{j \in \mathbb{Z}^d \\ |j|_2 \leq \mathcal{N}}} \langle j \rangle^s |w_j|^2 \right)^{\frac{1}{2}}$$

$$= \left(\sum_{\substack{j \in \mathbb{Z}^d \\ |j|_2 \leq \mathcal{N}}} \langle j \rangle^{s - 2s_0 + 2s_0} |w_j|^2 \right)^{\frac{1}{2}}$$

$$\leq \mathcal{N}^{\frac{s - 2s_0}{2}} \|\widetilde{w}\|_{s_0}.$$

Similarly, in view of definition (1.2) and (1.4), we can also prove that (6.2) holds. □

Data Availability The manuscript has no associated data.

Declarations

Conflict of interest The authors state that there is no conflict of interest.

References

1. Aizenman, M., Molchanov, S.: Localization at large disorder and at extreme energies: an elementary derivation. Commun. Math. Phys. **157**(2), 245–278 (1993)
2. Anderson, P.-W.: Absence of diffusion in certain random lattices. Phys. Rev. **109**(5), 1492–1505 (1958)
3. Benettin, G., Fröhlich, J., Giorgilli, A.: A Nekhoroshev-type theorem for Hamiltonian systems with infinitely many degrees of freedom. Commun. Math. Phys. **119**(1), 95–108 (1988)
4. Bernier, J., Faou, E., Grébert, B.: Long time behavior of the solutions of NLW on the d-dimensional torus. Forum Math. Sigma **8**, Paper No. e12, 26 (2020)
5. Bambusi, D., Grébert, B.: Birkhoff normal form for partial differential equations with tame modulus. Duke Math. J. **135**(3), 507–567 (2006)
6. Bourgain, J., Goldstein, M., Schlag, W.: Anderson localization for Schrödinger operators on Z^2 with quasi-periodic potential. Acta Math. **188**(1), 41–86 (2002)
7. Biasco, L., Massetti, J.E., Procesi, M.: An abstract Birkhoff normal form theorem and exponential type stability of the 1d NLS. Commun. Math. Phys. **375**(3), 2089–2153 (2020)
8. Bourgain, J.: Anderson localization for quasi-periodic lattice Schrödinger operators on $\mathbb{Z}^d$, d arbitrary. Geom. Funct. Anal. **17**(3), 682–706 (2007)
9. Bourgain, J., Wang, W.-M.: Diffusion bound for a nonlinear Schrödinger equation. In: Mathematical Aspects of Nonlinear Dispersive Equations, Volume 163 of Annals of Mathematics Studies, pp. 21–42. Princeton University Press, Princeton, NJ (2007)
10. Bourgain, J., Wang, W.-M.: Quasi-periodic solutions of nonlinear random Schrödinger equations. J. Eur. Math. Soc. **10**(1), 1–45 (2008)
11. Cong, H., Shi, Y., Wang, W.-M.: Long-time Anderson localization for the nonlinear quasi-periodic Schrödinger equation on $\mathbb{Z}^d$. arXiv:2309.15706 (2023)
12. Cong, H., Shi, Y., Zhang, Z.: Long-time Anderson localization for the nonlinear Schrödinger equation revisited. J. Stat. Phys. **182**(1), Paper No. 10, 22 (2021)
13. Faou, E., Grébert, B.: A Nekhoroshev-type theorem for the nonlinear Schrödinger equation on the torus. Anal. PDE **6**(6), 1243–1262 (2013)

14. Fishman, S., Krivolapov, Y., Soffer, A.: On the problem of dynamical localization in the nonlinear Schrödinger equation with a random potential. J. Stat. Phys. **131**(5), 843–865 (2008)
15. Fishman, S., Krivolapov, Y., Soffer, A.: Perturbation theory for the nonlinear Schrödinger equation with a random potential. Nonlinearity **22**(12), 2861–2887 (2009)
16. Fröhlich, J., Spencer, T.: Absence of diffusion in the Anderson tight binding model for large disorder or low energy. Commun. Math. Phys. **88**(2), 151–184 (1983)
17. Fröhlich, J., Spencer, T., Wayne, C.E.: Localization in disordered, nonlinear dynamical systems. J. Stat. Phys. **42**(3–4), 247–274 (1986)
18. Kleinbock, D.Y., Margulis, G.A.: Flows on homogeneous spaces and Diophantine approximation on manifolds. Ann. Math. (2) **148**(1), 339–360 (1998)
19. Liu, W., Wang, W.-M.: Nonlinear Anderson localized states at arbitrary disorder. arXiv:2201.00173 (2022)
20. Schwartz, T., Bartal, G., Fishman, S., Segev, M.: Transport and Anderson localization in disordered two-dimensional photonic lattices. Nature **446**(7131), 52–55 (2007)
21. Shapiro, B.: Expansion of a Bose-Einstein condensate in the presence of disorder. Phys. Rev. Lett. **99**(6), 060602, 1–4 (2007)
22. Shi, Y., Wang, W.-M.: Anderson localized states for the quasi-periodic nonlinear wave equation on $\mathbb{Z}^d$. arXiv:2306.00513 (2023)
23. Shi, Y., Wang, W.-M.: Anderson localized states for the quasi-periodic nonlinear Schrödinger equation on $\mathbb{Z}^d$. arXiv:2405.17513 (2024)
24. Wang, W.-M., Zhang, Z.: Long time Anderson localization for the nonlinear random Schrödinger equation. J. Stat. Phys. **134**(5–6), 953–968 (2009)
25. Yuan, X.: Construction of quasi-periodic breathers via KAM technique. Commun. Math. Phys. **226**(1), 61–100 (2002)