TAIWANESE JOURNAL OF MATHEMATICS
Vol. 21, No. 5, pp. 1115–1132, October 2017
DOI: 10.11650/tjm/7951

A Nekhoroshev Type Theorem of Higher Dimensional Nonlinear Schrödinger Equations

Shidi Zhou and Jiansheng Geng*

Abstract. In this paper, we prove a Nekhoroshev type theorem for high dimensional NLS (nonlinear Schrödinger equations):

$$\mathrm{i}\partial_t u - \Delta u + V * u + \partial_{\overline{u}} g(x, u, \overline{u}) = 0, \quad x \in \mathbb{T}^d,\ t \in \mathbb{R}$$

where real-valued function V is sufficiently smooth and g is an analytic function. We prove that, for any given $M \in \mathbb{N}$, there exists an $\varepsilon_0 > 0$, such that for any solution $u = u(t, x)$ with initial data $u_0 = u_0(x)$ whose Sobolev norm $\|u_0\|_s = \varepsilon < \varepsilon_0$, during the time $|t| \leq \varepsilon^{-M}$, its Sobolev norm $\|u(t)\|_s$ remains bounded by $C_s\varepsilon$.

1. Introduction

We consider the Hamiltonian NLS with convolutional type potential:

$$i\partial_t u - \Delta u + V * u + \frac{\partial g(x, u, \overline{u})}{\partial \overline{u}} = 0, \quad x \in \mathbb{T}^d,\ d \geq 2;\ t \in \mathbb{R} \tag{1.1}$$

where the real-valued potential function $V = V(x)$ is smooth on $\mathbb{T}^d$, $g = g(x, a, b)$ is real analytic on $\mathbb{T}^d_\mu$ times a neighborhood of the origin in $\mathbb{C}^2$, where $\mathbb{T}^d_\mu = \{x + iy : x \in \mathbb{T}^d, y \in \mathbb{R}^d, |y_i| \leq \mu\}$. $g = g(x, a, b)$ is real which means that it takes real value when $x \in \mathbb{T}^d$, $b = \overline{a}$. We require that $g = g(x, a, b)$ should be zero of order at least 3 at the origin in $\mathbb{C}^2$, i.e., the Taylor expansion of g with respect to (a, b) at the origin should start from the third order term. Hamiltonian equation (1.1) can be rewritten as

$$i\partial_t u = \frac{\partial H}{\partial \overline{u}}$$

where the Hamiltonian H is

$$H = \int_{\mathbb{T}^d} |\nabla u|^2 + (V * u)\overline{u} + g(x, u, \overline{u})\, dx.$$

Received October 8, 2016; Accepted January 8, 2017.
Communicated by Yingfei Yi.
2010 *Mathematics Subject Classification.* 37K55, 35B10.
Key words and phrases. KAM theory, Hamiltonian systems, Schrödinger equation, Birkhoff normal form.
This work is partially supported by NSFC grant 11271180.
*Corresponding author.

As important models in mathematical physics, higher dimensional Hamiltonian PDEs have attracted a lot of interests. Two aspects are especially interesting: one is the existence of invariant tori, which implies the existence of quasi-periodic solutions, and the other one is the the long-time behavior of solutions.

For the former one, the main tool is the infinite dimensional KAM theory. Geng and You [12] got the existence and linear stability of quasi-periodic solutions of higher dimensional beam equation and nonlocal smooth Schrödinger equation with nonlinearity independent on the spatial variable x in 2006. And one of the most important results, up to now, is [10] by Eliasson and Kuksin, which dealt with nonlinear Schrödinger equation with convolutional type potential. They developed the important "Lipschitz-Domain" property to deal with the measure estimate. Following their idea and method, Geng, Xu and You [13] got the quasi-periodic solutions of two dimensional completely resonant Schrödinger equation by an elaborate choice of tangential sites. Later, C. Procesi and M. Procesi proved the same result in arbitrary dimensional space [21, 22]. Recently Eliasson, Grébert and Kuksin established the KAM theorem for beam equation in higher dimensional space with typical constant potential [9].

For the latter one, it was originated from Nekhoroshev [19], who gave the Nekhoroshev type estimate in the finite dimensional case (See also [6,14,20]). But in the case of PDEs it's much more complex. The trouble mainly comes from working on infinite many frequencies and it's difficult to give the small divisor conditions. Bambusi and Grébert [5] proved a general normal form theorem and applied it to several kinds of Hamiltonian PDEs, including higher dimensional nonlinear Schrödinger equation. In [5], the nonlinearity of equations should satisfy the property "finite-module" which is weaker than analyticity. Later in 2014, Yuan and Zhang extended this result to the derivative nonlinear Schrödinger equation [23]. For the more convenient case that the nonlinearity is analytic, Bambusi [4] proved the Nekhoroshev type estimate of wave equation in one dimension (See also [1–3, 7, 15]). In these known results mentioned above about PDEs, only Sobolev norms was considered and the size of escaping time is up to order ε^{-M}. Apart from these results, an interesting result was obtained by Faou and Grébert [11], which considered the initial data in analytic norm instead of traditional Sobolev norm and got better results: The escaping time could be extended to size $\varepsilon^{-\alpha|\ln\varepsilon|^{\beta}}$, which is much longer than ε^{-M}.

There has also been abundant knowledge about the opposite direction: The fast growth of the solution of NLS on $\mathbb{T}^2$. It was origined from [8], and later, based on the elaborate construction of the "toy-model" in [8], Guardia and Kaloshin got the better escaping time in [18] (Notice that by the result in [11], their estimate of escaping time is sharp). For other important results just refer to [16, 17].

Though there has been rich knowledge about the Nekhoroshev type estimate of NLS in higher dimensional space, most of them depend on a crucial condition: zero momentum condition, which relies on the fact that the nonlinear term should be independent of the spatial variable x. In this paper, we consider the nonlinearity of the form $\partial_{\overline{u}} g(x,u,\overline{u})$. Actually, without zero momentum condition, it would bring some difficulties. It will be found that after a canonical transformation, the Hamiltonian is put into the form H_0+Z+P, where H_0 is integrable, P is the perturbation which is very small, and Z can be written as sums of terms of the form: $I_{j_1}I_{j_2}\cdots I_{j_r}\xi_n\eta_m$, $\|j_1\|\le\|j_2\|\le\cdots\le\|j_r\|\le N<\|n\|=\|m\|$ ($\|\cdot\|$ means the l^2 norm in $\mathbb{Z}^d$), and the parameter N will be chosen appropriately later. Such terms will survive because the corresponding small divisor is $(\omega_{j_1}-\omega_{j_1})+(\omega_{j_2}-\omega_{j_2})+\cdots+(\omega_{j_r}-\omega_{j_r})+(\omega_n-\omega_m)$ with $\omega_n=\|n\|^2+V_n$, to which we can't give a positive lower bound. But if we have the zero momentum condition, such terms are absent except for the case $n=m$, and in this convenient case, it's easy to verify $\{\xi_n\eta_n,Z\}=0$ ($\{\cdot,\cdot\}$ denotes Poisson Bracket), which means that Z has no contribution to the growth of norms of the solution. In our case, it's also crucial to verify that Z contributes nothing to the growth of norms of the solution, but we need to consider "almost action variable" $J_S=\sum_{\|n\|^2=S}\xi_n\eta_n$ instead, and prove that $\{J_S,Z\}=0$. Actually, the energy exchange takes place between different Fourier modes n and m when $\|n\|=\|m\|$, but doesn't exist between n and m when $\|n\|\neq\|m\|$.

Our parameters come from the potential V: For $m>d/2$, $R>0$, let

$$W_m=\left\{V(x)=\sum_{a\in\mathbb{Z}^d}V_ae^{i\langle a,x\rangle}:V_a=\frac{V'_a}{R(1+\|a\|)^m}\in[-1/2,1/2]\right\}. \tag{1.2}$$

Let each V be equivalent to the sequence $\{V'_a\}_{a\in\mathbb{Z}^d}$ and give the latter one product probability measure.

Now we could state our main theorem:

Theorem 1.1 (Nekhoroshev type estimate)**.** *There exists a full measure set $\mathcal{V}\subseteq W_m$, such that $\forall\, V\in\mathcal{V}$, fix $M\ge 4$, there exists s_* large enough, such that $\forall\, s\ge s_*$ there exists constants $C_s,R_s\ge 0$, satisfying the following conditions: For each initial date u_0 with $\varepsilon=\|u_0\|_s\le R_s$, during the time $|t|\le 1/(\mathbb{C}_s\varepsilon^{M-2})$, we have $\|u_{(t)}\|_s\le 4\varepsilon$. Here the norm $\|\cdot\|_s$ means Sobolev norm which will be explained later.*

The proof of this theorem will be delayed to the end of the paper.

Our result is very similar to the part of [5] dealing with higher dimensional Schrödinger equation. In [5] Bambusi and Grébert dealt with Schrödinger equation in higher dimensional space with the nonlinearity having "finite module" which is weaker than analyticity. But in order to overcome the difficulty of absence of analyticity, they gave a very long and complex proof. In our case we only consider analytic nonlinearity, but most technical lemmas could be extended directly from those in [4], and with the help of these lemmas, our proof is more easy to understand. So we simplified the proof in [5] in the analytic case. In addition, when we compute the escaping time, due to the multiplicity of eigenvalues, we used the "almost action variable" instead of traditional action variable. In [5] the "almost action variable" was also mentioned but the concrete method of using it to calculate the escaping time was omitted, and here we give the concrete calculation.

The rest of this paper is organized as follows: In Section 2 we will state some important concepts of function space and give the nonresonant conditions and verify the nonlinearity satisfying our assumptions. In Section 3 we will state a list of technical lemmas without proof, which could be extended directly from those results in [4] without any difficulty. At last, in Section 4 we state and prove the important normal form theorem, and as a corollary, we prove Theorem 1.1.

2. Preliminaries

2.1. Hamiltonian formalism

Now let us consider equation (1.1). Let u be a solution of equation (1.1), and expand it into Fourier series on $\mathbb{T}^d$:

$$u = \sum_{j\in\mathbb{Z}^d} \xi_j e^{i\langle j,x\rangle}, \quad \overline{u} = \sum_{j\in\mathbb{Z}^d} \eta_j e^{-i\langle j,x\rangle}. \tag{2.1}$$

For convenience, we denote the sequence $\{\xi_n\}_{n\in\mathbb{Z}^d}$ by ξ, and $\{\eta_n\}_{n\in\mathbb{Z}^d}$ by η. Direct calculation shows that ξ, η satisfy the equations

$$\partial_t\xi_j = i\omega_j\xi_j + \frac{\partial P}{\partial\eta_j} \quad \text{and} \quad \partial_t\eta_j = -i\omega_j\xi_j - i\frac{\partial P}{\partial\xi_j} \tag{2.2}$$

for each $j \in \mathbb{Z}^d$. In 7951e2.2 the frequencies $\omega_j = V_j + \|j\|^2$, $V = \sum_{j\in\mathbb{Z}^d} V_j e^{i\langle j,x\rangle}$. The nonlinear term

$$f = \frac{1}{(2\pi)^d}\int_{\mathbb{T}^d} g\left(x, \sum_{j\in\mathbb{Z}^d}\xi_j e^{i\langle j,x\rangle}, \sum_{j\in\mathbb{Z}^d}\eta_j e^{-i\langle j,x\rangle}\right) dx.$$

This Hamiltonian system is endowed with the symplectic structure i $\sum_{j\in\mathbb{Z}^d} d\xi_j\wedge d\eta_j$, and for two polynomial functions $F = F(\xi,\eta)$, $G = G(\xi,\eta)$, we define their Poisson Bracket as

$$\{F,G\} = i\,\langle\nabla F, J\nabla G\rangle = i\sum_{j\in\mathbb{Z}^d}\left(\frac{\partial F}{\partial\xi_j}\frac{\partial G}{\partial\eta_j} - \frac{\partial G}{\partial\xi_j}\frac{\partial F}{\partial\eta_j}\right). \tag{2.3}$$

We say a function $F = F(\xi,\eta)$ is real if it takes real value when $\xi_j = \overline{\eta}_j$, $\forall\, j\in\mathbb{Z}^d$. We know that if the initial data is real, then the solution would be real.

To describe the size of u, we define the Sobolev norm for functions defined on $\mathbb{T}^d$: $\|u\|_s = \left(\sum_{|\alpha|\le s} \|\partial^\alpha u\|^2_{L^2(\mathbb{T}^d)}\right)^{1/2}$. It's easy to see that this norm is equivalent to the Sobolev norm of the sequence (ξ,η) (defined in (2.1)):

$$\|(\xi,\eta)\|_s = \sum_{j\in\mathbb{Z}^d} \left(\left(|\xi_j|^2 + |\eta_j|^2\right)|j|^{2s}\right)^{1/2}$$

and we define the space of (ξ,η) with such norm by $\mathcal{P}_s$. In the following we do not separate them again and we mainly use the Sobolev norms of sequences.

2.2. Polynomial space and momentum

In this subsection, we introduce polynomial space and related norms, as well as some useful notations such as momentum, small divisors and so on. First, in order to simplify the multi-indices, for a given sequence $(j_1, j_2, \ldots, j_l) \in (\mathbb{Z}^d)^l$, $l \ge 2$, we denote this sequence by $\boldsymbol{j}$. And for convenience, we denote the length of this sequence by $\#\boldsymbol{j}$, which is equal to l here. Now let us define some related concepts: Assume $\boldsymbol{j} = (j_1, j_2, \ldots, j_l)$, $\boldsymbol{k} = (k_1, k_2, \ldots, k_m)$, we define

(1) the monomial associated with $(\boldsymbol{j},\boldsymbol{k})$

$$\xi_{\boldsymbol{j}}\eta_{\boldsymbol{k}} = \xi_{j_1}\xi_{j_2}\cdots\xi_{j_l}\eta_{k_1}\eta_{k_2}\cdots\eta_{k_m};$$

(2) the momentum of $(\boldsymbol{j},\boldsymbol{k})$

$$M(\boldsymbol{j},\boldsymbol{k}) = j_1 + j_2 + \cdots + j_l - k_1 - k_2 - \cdots - k_m;$$

(3) small divisor of $(\boldsymbol{j},\boldsymbol{k})$

$$\Omega(\boldsymbol{j},\boldsymbol{k}) = \omega_{j_1} + \omega_{j_2} + \cdots + \omega_{j_l} - \omega_{k_1} - \omega_{k_2} - \cdots - \omega_{k_m},$$

where the frequencies $\{\omega_j\}_{j\in\mathbb{Z}^d}$ is defined as $\omega_j = \|j\|^2 + V_n$, and V_n is defined in (1.2).

Based on these notations, we could introduce the space of polynomials:

Definition 2.1. For a polynomial $h = h(\xi,\eta)$ which is homogeneous of degree n_1 in ξ and n_2 in η respectively, namely

$$h(\xi,\eta) = \sum_{\boldsymbol{j},\boldsymbol{k}} h_{\boldsymbol{j},\boldsymbol{k}}\xi_{\boldsymbol{j}}\eta_{\boldsymbol{k}}, \quad \boldsymbol{j} = (j_1,\ldots,j_{n_1}),\ \boldsymbol{k} = (k_1,\ldots,k_{n_2}) \tag{2.4}$$

we define

$$|h|_\mu = \sum_{l\in\mathbb{Z}^d} e^{\mu|l|} \sup_{M(\boldsymbol{j},\boldsymbol{k})=l} |h_{\boldsymbol{j},\boldsymbol{k}}| \tag{2.5}$$

and if it is finite, we call it the μ-module of h. Here in (2.5) the norm $|\cdot|$ means the l^1 norm in $\mathbb{Z}^d$.

For a polynomial h, we define a class of polynomials:

Definition 2.2. For a polynomial $h = \sum_{n_1,n_2} h^{n_1,n_2}$ where h^{n_1,n_2} is homogeneous of degree n_1 in ξ and n_2 in η respectively, if the quantity

$$\langle h\rangle_R^\mu = \sum_{n_1,n_2} \left|h^{n_1,n_2}\right|_\mu R^{n_1+n_2} < \infty$$

then we say h belongs to a class M_R^μ.

2.3. Nonlinear perturbation

The nonlinear part $g = g(x,a,b)$ is assumed to be analytic in a neighborhood of the origin in $\mathbb{T}^d_\mu \times \mathbb{C} \times \mathbb{C}$, so there exists $R_0 > 0$, such that for any $|a|, |b| < 2R_0$, we have that

$$\begin{aligned} g(x,a,b) &= \sum_{k_2\geq 0,k_3\geq 0} \frac{1}{k_2!k_3!}\partial_2^{k_2}\partial_3^{k_3} g(x,0,0)a^{k_2}b^{k_3} \\ &= \sum_{\substack{k_2\geq 0,k_3\geq 0\\ l\in\mathbb{Z}^d}} \frac{1}{k_2!k_3!}\partial_2^{k_2}\partial_3^{k_3}\widehat{g}(-l,0,0)e^{-i\langle l,x\rangle}a^{k_2}b^{k_3}. \end{aligned}$$

Now we substitute a by u and b by $\overline{u}$, and expand them into Fourier series, we get

$$\begin{aligned} f(\xi,\eta) &= \int_{\mathbb{T}^d} g(x,u,\overline{u})\,dx \\ &= \int_{\mathbb{T}^d} \sum_{k_2\geq 0,k_3\geq 0} \frac{1}{k_2!k_3!} \sum_{l\in\mathbb{Z}^d} \left(\partial_2^{k_2}\partial_3^{k_3}\widehat{g}(-l,0,0)\right) e^{-i\langle l,x\rangle} \left(\sum_{a\in\mathbb{Z}^d} \xi_a e^{i\langle a,x\rangle}\right)^{k_2} \\ &\quad \times \left(\sum_{b\in\mathbb{Z}^d} \eta_b e^{-i\langle b,x\rangle}\right)^{k_3} dx \\ &= \int_{\mathbb{T}^d} \sum_{k_2\geq 0,k_3\geq 0} \frac{1}{k_2!k_3!} \sum_{\substack{l\in\mathbb{Z}^d\\ \#\boldsymbol{j}=k_2,\#\boldsymbol{k}=k_3}} \left(\partial_2^{k_2}\partial_3^{k_3}\widehat{g}(-l,0,0)\right) e^{i(\langle M(\boldsymbol{j},\boldsymbol{k}),x\rangle - l)}\xi_{\boldsymbol{j}}\eta_{\boldsymbol{k}}\,dx \\ &= \sum_{k_2\geq 0,k_3\geq 0} \frac{1}{k_2!k_3!} \sum_{l\in\mathbb{Z}^d, M(\boldsymbol{j},\boldsymbol{k})=l} \left(\partial_2^{k_2}\partial_3^{k_3}\widehat{g}(-l,0,0)\right)\xi_{\boldsymbol{j}}\eta_{\boldsymbol{k}}. \end{aligned}$$

We denote $f_{\boldsymbol{j},\boldsymbol{k}}^{k_2,k_3} = \partial_2^{k_2}\partial_3^{k_3}\widehat{g}(-l,0,0)$, where $l = M(\boldsymbol{j},\boldsymbol{k})$.

Due to Cauchy estimate and the fact that the Taylor expansion of g with respect to the latter two variable starts from order 3, it's easy to verify that

Proposition 2.3.

$$\forall\, k_2, k_3,\ \left| f^{k_2,k_3} \right|_{\mu} \le C \frac{M}{(2R_0)^{k_2+k_3}},$$
$$\langle f \rangle_R^{\mu} \le AR^3,$$

here C, M, A are some constants and $R \le \overline{R}_1$, where $\overline{R}_1$ is a fixed constant. If g is independent of x, the momentum l would be restricted to 0, and terms with other momentums will be absent.

2.4. Nonresonance conditions and normal form

The key step in our proof is the control of the small divisors, so we need to impose a so called N-nonresonance condition on the frequencies.

Definition 2.4. For $N \in \mathbb{N}$, $N \ge 3$, we say $\boldsymbol{j} = (j_1, \dots, j_r) \in (\mathbb{Z}^d)^r$ is $N - (\gamma, \alpha)$-nonresonant if it satisfies the following conditions:

(1) $r \ge 3$, the sequence $\boldsymbol{j}$ could be ordered as

$$\|i_1\| \le \|i_2\| \le \cdots \le \|i_{r-2}\| \le N < \|k\|, \|l\|;$$

(2) the corresponding frequencies $\{\omega_a\}_{a \in \mathbb{Z}^d}$ satisfy:

$$\left|\omega_{i_1} \pm \omega_{i_2} \pm \cdots \pm \omega_{i_{r-2}} \pm \omega_k \pm \omega_l\right| \ge \frac{\gamma}{N^{\alpha}} \tag{2.6}$$

except for the case when the lower-indices could be ordered as

$$a_1, a_1, a_2, a_2, \dots, a_m, a_m, k, l, \quad \|a_i\| \le N < \|k\| = \|l\|$$

and (2.6) is written as $(\omega_{a_1} - \omega_{a_1}) + (\omega_{a_2} - \omega_{a_2}) + \cdots + (\omega_{a_m} - \omega_{a_m}) + \omega_k - \omega_l$, $m = (r-2)/2$.

Now we give the notation of N-normal form.

Definition 2.5. Consider a polynomial $Z = \sum_{n_1,n_2} Z^{n_1,n_2}$, where Z^{n_1,n_2} is homogeneous of degree n_1 in ξ and n_2 in η. Set $Z^{n_1,n_2} = \sum_{\boldsymbol{j},\boldsymbol{k}} Z^{n_1,n_2}_{\boldsymbol{j},\boldsymbol{k}} \xi_{\boldsymbol{j}} \eta_{\boldsymbol{k}}$, we assume

$$\boldsymbol{j} = (j_1, j_2, \dots, j_{n_1}), \qquad \|j_1\| \le \|j_2\| \le \cdots \le \|j_{n_1}\|$$
$$\boldsymbol{k} = (k_1, k_2, \dots, k_{n_2}), \qquad \|k_1\| \le \|k_2\| \le \cdots \le \|k_{n_2}\|.$$

We say Z is in N-normal form if $Z^{n_1,n_2}_{\boldsymbol{j},\boldsymbol{k}} = 0$ for each $n_1 \ne n_2$, and the lower-indices satisfy either one of the following two conditions:

$$j_1 = k_1, j_2 = k_2, \dots, j_{n_1} = k_{n_2}, \quad \|j_i\| \le N,\ 1 \le i \le n_1,$$
$$j_1 = k_1, j_2 = k_2, \dots, j_{n_1-1} = k_{n_2-1}, \quad \|j_{n_1-1}\| \le N < \|j_{n_1}\| = \|k_{n_2}\|.$$

We state a very important property working on measure estimate which was proved in [5] :

Proposition 2.6. (Lemma 5.22 in [5]) *Fix $r \geq 0$ and $\gamma > 0$ small enough. There exist positive constants $C = C_r$, $\beta = \beta(r, \gamma)$ and a set $S_\gamma \subseteq \mathcal{V}$ with* $\operatorname{meas}(\mathcal{V} \setminus S_\gamma) \to 0$ *when $\gamma \to 0$ such that, if $V \in S_\gamma$ then for any $N \geq 1$ and any $n \in \mathbb{Z}$, one has*

$$\left|\left\langle \omega^{(N)}, k \right\rangle + n\right| \geq \frac{\gamma}{N^\beta}$$

for any $k \in \mathbb{Z}^{\mathbb{Z}^d}$ with $0 < |k| \leq r$. Here $\omega^{(N)} = (\omega_j)_{\|j\| \leq N}$.

By Proposition 2.6, we could conclude that for $V \in S_\gamma$, each $j \in (\mathbb{Z}^d)^r$ is $N - (\gamma, \alpha)$ nonresonant if it has only at most two components larger than N.

Now we could state the normal form theorem.

Theorem 2.7. *For fixed integer $M \geq 4$, $\forall\, V \in \mathcal{V}$, there exists $\gamma, \alpha > 0$, and a sufficiently large $s_* > 0$, $R_* > 0$, such that $\forall\, 0 < R^{(s_*)} < R_*$, there exists an analytic canonical transformation $\Gamma_R \colon B_{s_*}(R^{(s_*)}/3) \to B_{s_*}(R^{(s_*)})$ which put the Hamiltonian $H_0 + P$ into the main part plus higher order terms, i.e.,*

$$(H_0 + P) \circ \Gamma_R = H_0 + Z + \mathcal{R}$$

where Z is in $N - (\gamma, \alpha)$ normal form and the parameter N will be chosen appropriately later. And the vector field of the remaining term $\mathcal{R}$ is an analytic map from $\mathcal{P}_{s_*}$ to itself. Moreover, $\forall\, s > s_*$, $\exists\, C_s > 0$, $\forall\, R < R_*/C_s$, the following estimates hold:

$$\sup_{\|(\xi,\eta)\|_s < R} \|(\xi, \eta) - \Gamma_R(\xi, \eta)\|_s < C_s R^2,$$
$$\sup_{\|(\xi,\eta)\|_s < R} \|X_{\mathcal{R}}(\xi, \eta)\|_s < C_s R^M. \tag{2.7}$$

The proof of this theorem will be delayed to the end of Section 4.

3. A list of technical lemmas

In this part, we will give a list of lemmas describing properties of functions in M_R^μ, and besides we will also describe the vector field of $f \circ \Phi$ where $f \in M_R^\mu$, Φ is a Lie transformation. Most of these lemmas could be extended from [4] by Bambusi directly without any difficulty and we just omit the proofs.

For any analytic function f that has an analytic vector field, we write $\|X_f\|_s^R = \sup_{\|(\xi,\eta)\|_s < R} \|X_f(\xi, \eta)\|_s < \infty$. And for a given homogeneous polynomial h, which is of degree n_1 in ξ and n_2 in η respectively, namely $h = \sum_{\#\boldsymbol{j} = n_1, \#\boldsymbol{k} = n_2} h_{\boldsymbol{j},\boldsymbol{k}} \xi_{\boldsymbol{j}} \eta_{\boldsymbol{k}}$, we let $\breve{h} = \sum_{\boldsymbol{j},\boldsymbol{k}} e^{-\mu |M(\boldsymbol{j},\boldsymbol{k})|} \xi_{\boldsymbol{j}} \eta_{\boldsymbol{k}}$. Then we have the following estimate:

Lemma 3.1. *There exists a constant Σ_s, $\rho_s > 1$, which only depends on the Sobolev index s and spatial dimension d, for h defined above and $n = n_1 + n_2$ one has*

$$\|X_h\|_s^R \le |h|_\mu \left\|X_{\breve{h}}\right\|_s^R, \tag{3.1}$$
$$\left\|X_{\breve{h}}\right\|_s^R \le n\Sigma_s(\rho_s R)^{n-1}.$$

In the following text, Lemma 3.1 will be used to bound the size of the vector field $\|X_h\|_s^R$ by $\langle h\rangle_R^\mu$. The constants Σ_s, ρ_s in Lemma 3.1 will be fixed throughout this paper.

Lemma 3.2. *Let $h \in M_R^\mu$. For any $R^{(s)} > 0$, $\rho_s > 1$, $\rho_s R^{(s)} < R$, $0 < \delta^{(s)} < R^{(s)}$, we have the estimate $\|X_h\|_s^{R^{(s)}-\delta^{(s)}} \le \frac{\Sigma_s}{\rho_s\delta^{(s)}} \langle h\rangle_{\rho_s R^{(s)}}^\mu$.*

Now we split (ξ, η) into p, q, P, Q, defined as

$$\begin{aligned} p_i &= \xi_i, \quad |i| \le N, \qquad P_i = \xi_i, \quad |i| > N,\\ q_i &= \eta_i, \quad |i| \le N, \qquad Q_i = \eta_i, \quad |i| > N, \end{aligned} \tag{3.2}$$

where N will be specified later. Similar as before, we need to consider a polynomial h which is homogeneous of degree m_1, m_2, m_3, m_4 in p, q, P, Q respectively, namely

$$h(p,q,P,Q) = \sum_{\boldsymbol{k},\boldsymbol{i},\boldsymbol{j},\boldsymbol{l}} h_{\boldsymbol{k},\boldsymbol{i},\boldsymbol{j},\boldsymbol{l}} p_{\boldsymbol{k}} q_{\boldsymbol{i}} P_{\boldsymbol{j}} Q_{\boldsymbol{l}}. \tag{3.3}$$

We let

$$\breve{\breve{h}} = \sum_{\boldsymbol{k},\boldsymbol{i},\boldsymbol{j},\boldsymbol{l}} e^{-\mu|M(\boldsymbol{k},\boldsymbol{i},\boldsymbol{j},\boldsymbol{l})|} p_{\boldsymbol{k}} q_{\boldsymbol{i}} P_{\boldsymbol{j}} Q_{\boldsymbol{l}}.$$

Different from (2.4), here the definition is based on the parameter N. We have the following estimates:

Lemma 3.3. *Let h be a homogeneous polynomial of degree m_1, m_2, m_3, m_4 in p, q, P, Q respectively, $m_3 + m_4 \ge 3$, then we have*

$$\|X_h\|_s^R \le |h|_\mu \left\|X_{\breve{\breve{h}}}\right\|_s^R,$$
$$\left\|X_{\breve{\breve{h}}}\right\|_s^R \le \frac{n\Sigma_s}{N^{s-1}}(\rho_s R)^{n-1}, \quad n = m_1 + m_2 + m_3 + m_4.$$

Lemma 3.4. *Let $h \in M_R^\mu$ be a polynomial of degree $\le r$, and we assume that in the expression (3.3), $m_3 + m_4 \ge 3$, then for any $R^{(s)} > 0$, $\rho_s > 0$, $\rho_s R^{(s)} < R$, $0 < \delta^{(s)} < R^{(s)}$, we have*

$$\|X_h\|_s^{R^{(s)}-\delta^{(s)}} \le \frac{c\Sigma_s}{\rho_s \delta^{(s)} N^{s-1}} \langle h\rangle_{\rho_s R^{(s)}}^M \tag{3.4}$$

where the constant c is related to s and r.

Now we consider the Lie-transformation. For a Hamiltonian χ, the induced time t-map $T = X_\chi^t$ is called the Lie-transformation generated by χ. For an analytic function h, we have

$$h \circ X_\chi^1 = \sum_{l \geq 0} h_l$$

where

$$h_0 = h, \quad h_l = \frac{1}{l} \{h_{l-1}, \chi\}, \quad l \geq 1 \tag{3.5}$$

and

$$\frac{d}{dt}(h \circ X_\chi^t) = \{h, \chi\} \circ X_\chi^t, \quad |t| \leq 1.$$

In the following text we will give a list of estimates concerning the M_{R-d}^μ norm of $h \circ X_\chi^1$ for $h \in M_R^\mu$.

Lemma 3.5. *Let h be homogeneous of degree m_1 in ξ and m_2 in η respectively, and g be homogeneous of degree n_1 in ξ and n_2 in η respectively. Then we have*

$$|\{h, g\}|_\mu \leq mn \, |h|_\mu \, |g|_\mu, \quad m = m_1 + m_2, \ n = n_1 + n_2.$$

Lemma 3.6. *Let $h \in M_R^\mu$, $g \in M_{R-d}^\mu$, $0 < d < R$. For any $0 < d' < R - d$, one has $\{h, g\} \in M_{R-d-d'}^\mu$ and*

$$\langle \{h, g\} \rangle_{R-d-d'}^\mu \leq \frac{1}{d'(d + d')} \langle h \rangle_R^\mu \langle g \rangle_{R-d}^\mu.$$

Based on these lemmas, we give the estimate after Poisson Bracket and after the Lie-transformation:

Lemma 3.7. *Let $h \in M_R^\mu$, $\chi \in M_R^\mu$ are analytic, let h_n be defined as* (3.5), *then for any $0 < d < R$, one has $h_n \in M_{R-d}^\mu$ and*

$$\langle h_n \rangle_{R-d}^\mu \leq \langle h \rangle_R^\mu \left(\frac{e^2}{d^2} \langle \chi \rangle_R^\mu \right)^n.$$

Lemma 3.8. *Let χ be an analytic Hamiltonian, $0 < \delta^{(s)} < R^{(s)}$, if $\|X_\chi\|_s^{R^{(s)}} < \delta^{(s)}$, then for any $|t| \leq 1$, one has*

$$\sup_{\|(\xi,\eta)\|_s < R^{(s)} - \delta^{(s)}} \left\| T^t(\xi, \eta) - (\xi, \eta) \right\|_s \leq \|\chi\|_s^{R^{(s)}}.$$

Lemma 3.9. *Let χ be above and $h \colon B_s(R^{(s)}) \to \mathbb{C}$ be analytic with analytic vector field in $B_s(R^{(s)})$, let $0 < \delta^{(s)} < R^{(s)}$, $\|X_\chi\|_s^{R^{(s)}} < \delta^{(s)}/3$, then for $|t| \leq 1$, one has*

$$\left\| X_{h \circ X_\chi^t} \right\|_s^{R^{(s)} - \delta^{(s)}} \leq \left(1 + \frac{3}{\delta^{(s)}} \|X_\chi\|_s^{R^{(s)}} \right) \|X_h\|_s^{R^{(s)}}. \tag{3.6}$$

4. Proof of main theorem

For a given $R > 0$, $r > 0$, we define $\delta = R/(2r)$, $R_k = R - k\delta$. And for $R^{(s)} > 0$, let $\delta^{(s)} = R^{(s)}/(2r)$, $R_k^{(s)} = R^{(s)} - k\delta^{(s)}$.

Lemma 4.1 (Iterative Lemma). *Fix $r \geq 4$, assume $\langle f\rangle_R^\mu \leq AR^3$, for $R < R_1$, $R < R_*/2$ with $R_* = \gamma/(48e^2AN^\alpha r^2)$. Then for $k \leq r-4$, there exists a canonical transformation $\Gamma^{(k)}$ which puts $H_0 + f$ into*

$$H^{(k)} = (H_0 + f) \circ \Gamma^{(k)} = H_0 + Z^{(k)} + f^{(k)} + \mathcal{R}_N^{(k)} + \mathcal{R}_T^{(k)}$$

where $Z^{(k)}$ is in $N-(\gamma,\alpha)$ normal form, and we have the following estimates:

$$\left\langle Z^{(k)}\right\rangle_{R_k}^{\mu} \leq AR^3 \sum_{l=0}^{k-1} \left(\frac{R}{R_*}\right)^l, \quad \left\langle f^{(k)}\right\rangle_{R_k}^{\mu} \leq AR^3 \left(\frac{R}{R_*}\right)^k. \tag{4.1}$$

And for $\forall\, s \geq 1$, let $R_*^{(s)} = R_*/\rho_s$, then for any $R^{(s)}$ satisfying

$$\frac{R^{(s)}}{R_*^{(s)}} \leq \min\left\{\frac{1}{2}, \frac{2e^2}{\rho_s\Sigma_s}\right\} \tag{4.2}$$

we have that $\Gamma^{(k)}$ is an analytic map from $B_s(R_{k+1}^{(s)})$ to $B_s(R_1^{(s)})$, and

$$\sup_{\|(\xi,\eta)\|_s < R_{k+1}^{(s)}} \left\|\Gamma^{(k)}(\xi,\eta) - (\xi,\eta)\right\|_s < \frac{R^{(s)}\rho_s\Sigma_s}{16e^2 r} \sum_{l=1}^{k} \left(\frac{R^{(s)}}{R_*^{(s)}}\right)^l. \tag{4.3}$$

For the remaining part $\mathcal{R}_N^{(k)}$ and $\mathcal{R}_T^{(k)}$, we have

$$\left\|X_{\mathcal{R}_N^{(k)}}\right\|_s^{R_{k+1}^{(s)}} \leq \frac{1}{N^{s-1}} 2rA\Sigma_s(\rho_s R^{(s)})^2 \left(\sum_{l=0}^{k-1}\left(\frac{R^{(s)}}{R_*^{(s)}}\right)^l\right) \times \prod_{l=0}^{k-1}\left(1 + \frac{\rho_s\Sigma_s}{4e^2}\left(\frac{R^{(s)}}{R_*^{(s)}}\right)^{l+1}\right) \tag{4.4}$$

and

$$\left\|X_{\mathcal{R}_T^{(k)}}\right\|_s^{R_{k+1}^{(s)}} \leq \left(\frac{R^{(s)}}{R_*^{(s)}}\right)^{r-1} \Sigma_s r 2^r A R_*^2 \left(2 - \frac{1}{2^{k-1}}\right) \times \prod_{l=0}^{k-1}\left(1 + \frac{\rho_s\Sigma_s}{14e^2}\left(\frac{R^{(s)}}{R_*^{(s)}}\right)^{l+1}\right). \tag{4.5}$$

Proof. We prove it by an iterative procedure.

At the beginning, $H = H_0 + f$, just let $Z^{(0)} = 0$, $\mathcal{R}_N^{(0)} = \mathcal{R}_T^{(0)} = 0$. Assume that we have arrived at the k-th step, then split $f^{(k)} = f_0^{(k)} + f_T^{(k)}$, where $f_0^{(k)}$ is the part of the expansion of $f^{(k)}$ with order $\leq r-1$ and $f_T^{(k)}$ is the remainder. Expand $f_0^{(k)}$ into Taylor series with respect to (P,Q) only (defined in (3.2)). Let $\check{f}^{(k)}$ be the part containing terms at most quadratic in (P,Q), and $f_N^{(k)}$ be the remaining part. It's easy to see that $\langle \check{f}^{(k)}\rangle_R^\mu$, $\left\langle f_N^{(k)}\right\rangle_R^\mu$, $\left\langle f_T^{(k)}\right\rangle_R^\mu \leq \langle f^{(k)}\rangle_R^\mu$, and we rewrite the Hamiltonian as

$$\begin{aligned} H^{(k)} &= H_0 + Z^{(k)} + f^{(k)} + \mathcal{R}_N^{(k)} + \mathcal{R}_T^{(k)} \\ &= H_0 + Z^{(k)} + \check{f}^{(k)} + (f_N^{(k)} + \mathcal{R}_N^{(k)}) + (f_T^{(k)} + \mathcal{R}_T^{(k)}). \end{aligned}$$

After one step of Lie-transformation induced by χ_k, we get

$$\begin{aligned} H^{(k+1)} &= H^{(k)} \circ X_{\chi_k}^1 \\ &= H_0 \\ &\quad + Z^{(k)} + \{H_0, \chi_k\} + \check{f}^{(k)} \\ &\quad + \sum_{l\geq 2} H_{0l} + \sum_{l\geq 1} Z_l^{(k)} + \sum_{l\geq 1} \check{f}_l^{(k)} \qquad (4.6) \\ &\quad + (f_N^{(k)} + \mathcal{R}_N^{(k)}) \circ X_{\chi_k}^1 \qquad (4.7) \\ &\quad + (f_T^{(k)} + \mathcal{R}_t^{(k)}) \circ X_{\chi_k}^1. \qquad (4.8) \end{aligned}$$

We let $Z^{(k+1)} = Z^{(k)} + Z_k$, where $Z_k = \{H_0, \chi_k\} + \check{f}^{(k)}$, χ_k will be chosen appropriately later to ensure $Z^{(k+1)}$ is still in N-normal form. Let $f^{(k+1)} = (4.6)$, $\mathcal{R}_N^{(k+1)} = (4.7)$, $\mathcal{R}_T^{(k+1)} = (4.8)$ and $\Gamma^{(k+1)} = \Gamma^{(k)} \circ \Gamma_k$.

At first we need to estimate the norm of χ_k. According to Proposition 2.6, $\exists\, \gamma, \alpha > 0$, such that the frequencies satisfy the $N - (\gamma, \alpha)$ nonresonant condition. We let

$$\mathcal{A} = \left\{ (\boldsymbol{j}, \boldsymbol{k}) : |\Omega(\boldsymbol{j}, \boldsymbol{k})| \geq \frac{\gamma}{N^\alpha} \right\}$$

and let $\mathcal{B}$ be its complementary set. Then we set

$$Z_k = \sum_{(\boldsymbol{j},\boldsymbol{k})\in\mathcal{B}} \check{f}_{\boldsymbol{j},\boldsymbol{k}}^{(k)} \xi_{\boldsymbol{j}} \eta_{\boldsymbol{k}},$$

$$\chi_k = \sum_{(\boldsymbol{j},\boldsymbol{k})\in\mathcal{A}} \frac{\check{f}_{\boldsymbol{j},\boldsymbol{k}}^{(k)}}{i\Omega(\boldsymbol{j}, \boldsymbol{k})}$$

which are solutions of equation $Z_k = \{H_0, \chi_k\} + \check{f}^{(k)}$. It's easy to see that Z_k is still in N-normal form so $Z^{(k+1)}$ will be in N-normal form. And the norms of Z_k, χ_k could be bounded by:

$$\langle Z_k \rangle_R^\mu \leq \left\langle \check{f}^{(k)} \right\rangle_R^\mu, \quad \langle \chi_k \rangle_R^\mu \leq \frac{N^\alpha}{\gamma} \left\langle \check{f}^{(k)} \right\rangle_R^\mu.$$

Using (4.1) it's easy to get $\langle \check{f}^{(k)} \rangle_{R_k}^\mu \leq AR^3 (R/R_*)^k$ and so we have

$$\left\langle Z^{(k+1)} \right\rangle_{R_{k+1}}^\mu \leq AR^3 \sum_{l=0}^{k} \left(\frac{R}{R_*} \right)^l.$$

Thus we proved the first part of (4.1).

In order to estimate $f^{(k+1)}$, we consider $Z_l^{(k)}$, $\check{f}_l^{(k)}$, H_{0l} separately. At first

$$\langle \chi_k \rangle_{R_k}^{\mu} \leq \frac{N^\alpha}{\gamma} \left\langle \check{f}^{(k)} \right\rangle_{R_k}^{\mu} \leq AR^3 \frac{N^\alpha}{\gamma} \left(\frac{R}{R_*} \right)^k < \delta^{(s)}.$$

The last inequality comes from the definition of R_*. Let $\kappa = \frac{e^2}{\delta^2} \langle \chi_k \rangle_{R_k}^{\mu} < 1/2$. For $Z_l^{(k)}$, by (3.4), we get

$$\left\langle Z_l^{(k)} \right\rangle_{R_{k+1}}^{\mu} \leq \left\langle Z^{(k)} \right\rangle_{R_k}^{\mu} \kappa^l$$

and so

$$\sum_{l \geq 1} \left\langle Z_l^{(k)} \right\rangle_{R_{k+1}}^{\mu} \leq 4\kappa AR^3 \leq \frac{1}{3} AR^3 \left(\frac{R}{R_*} \right)^{k+1}$$

by the definition of R_*.

For $\check{f}_l^{(k)}$, according to $\langle \check{f}^{(k)} \rangle_{R_k}^{\mu} \leq \langle f^{(k)} \rangle_{R_k}^{\mu} \leq \langle Z^{(k)} \rangle_{R_k}^{\mu}$, we get to know that the estimate of this part is better than that of $\left\langle Z_l^{(k)} \right\rangle_{R_{k+1}}^{\mu}$. The same method applies to H_{0l} and we get the same estimate. Add these three parts up and we get the second part of (4.1).

Now we turn to the proof of the second part. By (3.1), (4.2) we get

$$\begin{aligned} \|X_{\chi_k}\|_s^{R_{k+1}^{(s)}} &\leq \frac{\Sigma_s}{\rho_s \delta^{(s)}} \langle \chi_k \rangle_{\rho_s R_k^{(s)}}^{\mu} \\ &\leq \frac{\Sigma_s}{\rho_s \delta^{(s)}} \frac{N^\alpha}{\gamma} A(\rho_s R^{(s)})^3 \left(\frac{R^{(s)}}{R_*^{(s)}} \right)^k \\ &\leq \frac{\rho_s \Sigma_s R^{(s)}}{24 e^2 r} \left(\frac{R^{(s)}}{R_*^{(s)}} \right)^{k+1} \\ &< \delta^{(s)}. \end{aligned} \tag{4.9}$$

So $\Gamma_k \colon B_s(R_{k+2}^{(s)}) \to B_s(R_{k+1}^{(s)})$ and

$$\Gamma^{(k+1)} = \Gamma^{(k)} \circ \Gamma_k \colon B_s(R_{k+2}^{(s)}) \to B_s(R_1^{(s)})$$

and by using (4.9) and induction, we get (4.3).

At last, we only need to verify the estimate of the remaining part (4.4), (4.5).

By (3.4), we get

$$\begin{aligned} \left\| X_{f_N^{(k)}} \right\|_s^{R_{k+1}^{(s)}} &\leq \frac{\Sigma_s}{\rho_s \delta^{(s)}} \frac{1}{N^{s-1}} \left\langle f_N^{(k)} \right\rangle_{\rho_s R_k^{(s)}}^{\mu} \\ &\leq \frac{\Sigma_s}{\rho_s \delta^{(s)}} \frac{1}{N^{s-1}} A(\rho_s R^{(s)})^3 \left(\frac{R^{(s)}}{R_*^{(s)}} \right)^k. \end{aligned}$$

So by (3.6), we have

$$\left\| X_{f_N^{(k)}\circ\Gamma_k} \right\|_s^{R_{k+2}^{(s)}} \leq \left(1 + \frac{3}{\delta^{(s)}} \left\| X_{\chi_k} \right\|_s^{R_{k+1}^{(s)}}\right) \left\| X_{f_N^{(k)}} \right\|_s^{R_{k+1}^{(s)}}$$
$$\leq \left(1 + \frac{\rho_s \Sigma_s}{4e^2} \left(\frac{R^{(s)}}{R_*^{(s)}}\right)^{k+1}\right) (\rho_s R^{(s)})^2 \frac{2rA\Sigma_s}{N^{s-1}} \left(\frac{R^{(s)}}{R_*^{(s)}}\right)^k$$

and

$$\left\| X_{\mathcal{R}_N^{(k)}\circ\Gamma_k} \right\|_s^{R_{k+2}^{(s)}} \leq \left(1 + \frac{\rho_s \Sigma_s}{4e^2} \left(\frac{R^{(s)}}{R_*^{(s)}}\right)^{k+1}\right) \left\| X_{\mathcal{R}_N^{(k)}} \right\|_s^{R_{k+1}^{(s)}}$$
$$\leq \frac{2rA\Sigma_s}{N^{s-1}} (\rho_s R^{(s)})^2 \left(\sum_{l=0}^{k-1} \left(\frac{R^{(s)}}{R_*^{(s)}}\right)^l\right) \left(\prod_{l=0}^{k-1} \left(1 + \frac{\rho_s \Sigma_s}{4e^2} \left(\frac{R^{(s)}}{R_*^{(s)}}\right)^{l+1}\right)\right).$$

Adding these two parts up, we get (4.4).

For the part $\left\| X_{\mathcal{R}_T^{(k+1)}} \right\|_s^{R_{k+2}^{(s)}}$, we first estimate $\left\langle f_T^{(k)} \right\rangle_{R_k}^{\mu}$. Notice that $\left\langle f_T^{(k)} \right\rangle_R^{\mu}$ is an analytic function of R with positive Taylor coefficients and start from order r, so one has

$$\left\langle f_T^{(k)} \right\rangle_{R_k}^{\mu} \leq \left(\frac{2R}{R_*}\right)^r \left\langle f^{(k)} \right\rangle_{\frac{1}{2}R_{*k}}^{\mu} \leq 2^{r-k-3} A R_*^3 \left(\frac{R}{R_*}\right)^3.$$

By (3.1), we get

$$\left\| X_{f_T^{(k)}} \right\|_s^{R_{k+1}^{(s)}} \leq 2^{r-k-2} r A \Sigma_s R_*^2 \left(\frac{R^{(s)}}{R_*^{(s)}}\right)^{r-1}.$$

The method of estimating $\left\| \mathcal{R}_T^{(k)} \right\|_s^{R_{k+1}^{(s)}}$ is similar to that of $\left\| \mathcal{R}_N^{(k)} \right\|_s^{R_{k+1}^{(s)}}$ and at last we get (4.5). □

Proof of Theorem 2.7. According to the definition of R_*, we get that $\exists\, c_s > 0$, $C_s > 0$, such that $R^{(s)} < c_s R_*$. Let $k = r-3$, we get $\Gamma^{(k)}\colon B_s(R^{(s)}/3) \to B_s(R^{(s)})$ and

$$\left\| \Gamma^{(k)}(\xi,\eta) - (\xi,\eta) \right\|_s < C_s (R^{(s)})^3 N^{2\alpha}.$$

And by (4.4), (4.5) and the definition of R_*, we get

$$\| X_{\mathcal{R}_N} \|_s^{R^{(s)}/3} < C_s \frac{R^{(s)}}{N^{s-1}},$$
$$\left\| X_{\mathcal{R}_T + f^{(k)}} \right\|_s^{R^{(s)}/3} < C_s (R^{(s)} N^{\alpha})^{r-1}.$$

Now we set $N = R^{-1/(2\alpha)}$, then one has

$$\left\| X_{\mathcal{R}_N} + X_{\mathcal{R}_T + f^{(k)}} \right\|_s^{R^{(s)}/3} < C_s \left((R^{(s)})^{(s-1)/(2\alpha)} + (R^{(s)})^{(r-1)/2} \right).$$

For the given M, we choose $r = 2M+1$, $s \geq s_* = 4\alpha M$, and then we get Theorem 2.7. □

Now we could prove Theorem 1.1 as a corollary of Theorem 2.7.

Proof of Theorem 1.1. Let $\{\xi_n\}_{n\in\mathbb{Z}^d}$ be Fourier series of u, $\{\eta_n\}_{n\in\mathbb{Z}^d}$ be Fourier series of $\overline{u}$, we could identify the solution $u=u(t,x)$ of (1.1) with the sequence $(\xi,\eta)=\{(\xi_n,\eta_n)\}_{n\in\mathbb{Z}^d}$. Notice that here we restrict to the case $\xi_n=\overline{\eta}_n$, $\forall\, n\in\mathbb{Z}^d$. For convenience, we define $\boldsymbol{z}=\{\xi_n\}_{n\in\mathbb{Z}^d}$ as well as $|\boldsymbol{z}|_s^2=\|\xi\|_s^2$ and let $\boldsymbol{z}=\Gamma_R\widehat{\boldsymbol{z}}$. We will study the behavior of $\boldsymbol{z}$ through the behavior of $\widehat{\boldsymbol{z}}$. We have $|\widehat{\boldsymbol{z}}|_s\le|\widehat{\boldsymbol{z}}-\boldsymbol{z}|_s+|\boldsymbol{z}|_s\le\epsilon+C_s\varepsilon^2<2\varepsilon$, if R_s is small enough. Then we need to consider the new system of $\widehat{\boldsymbol{z}}$. Now the Hamiltonian has been put into

$$(H_0+f)\circ\Gamma_R=H_0+Z+\mathcal{R}$$

we need to verify that the part H_0+Z contributes nothing to the growth of the solution. Given one element in Z, it should have a form of $I_{j_1}I_{j_2}\cdots I_{j_r}\xi_n\eta_m$, $\|n\|=\|m\|$, $I_j=\xi_j\eta_j$ and here (n,m) may vanish. For any given $S\in\mathbb{N}$ we show that

$$\{I_{j_1}I_{j_2}\cdots I_{j_r}\xi_n\eta_m,J_S\}=0,\quad J_S=\sum_{\|n\|^2=S}\xi_n\eta_n. \tag{4.10}$$

It's easy to verify that $\{I_{j_i},J_S\}=0$ and by the structure of Poisson Bracket (2.3), we only need to verify

$$\{\xi_n\eta_m,J_S\}=0,\quad \|n\|=\|m\|. \tag{4.11}$$

If $\|n\|^2\ne S$, then (4.11) holds automatically, otherwise

$$\{\xi_n\eta_m,J_S\}=\{\xi_n\eta_m,\xi_n\eta_n+\xi_m\eta_m\}=i(\eta_m\xi_n-\xi_n\eta_m)=0.$$

At last (4.10) holds. $\{H_0,J_S\}=0$ is obvious and we get

$$\left\{H_0+Z,\sum_{n\in\mathbb{Z}^d}\|n\|^{2s}\xi_n\eta_n\right\}=\sum_{S\ge0}\{H_0+Z,S^sJ_S\}=0.$$

So we have

$$\frac{d}{dt}|\widehat{\boldsymbol{z}}|_s^2=\left\{|\widehat{\boldsymbol{z}}|_s^2,H_0+Z+\mathcal{R}\right\}=\left\{|\widehat{\boldsymbol{z}}|_s^2,\mathcal{R}\right\} \tag{4.12}$$

and by (2.7), we know that the absolute value of (4.12) can be bounded by $C'\varepsilon^M$ in $B_s(3\varepsilon)$. Denote by T the escape time of $\widehat{\boldsymbol{z}}$ from $B_s(2\varepsilon)$ to outside of $B_s(3\varepsilon)$, and using $|\widehat{\boldsymbol{z}}(t)|_s^2\le|\widehat{\boldsymbol{z}}(0)|_s^2+C'\varepsilon^MT$ we get that $T\ge C\varepsilon^{2-M}$.

At last we go back to $\boldsymbol{z}$,

$$|\boldsymbol{z}|_s\le|\widehat{\boldsymbol{z}}-\boldsymbol{z}|_s+|\widehat{\boldsymbol{z}}|_s\le3\varepsilon+C_s\varepsilon^2\le4\varepsilon$$

so we finished the proof of Theorem 1.1. □

Acknowledgments

We would like to thank the anonymous referee for giving many valuable suggestions.

References

[1] D. Bambusi, *Nekhoroshev theorem for small amplitude solutions in nonlinear Schrödinger equations*, Math. Z. **230** (1999), no. 2, 345–387. https://doi.org/10.1007/pl00004696

[2] ______, *On long time stability in Hamiltonian perturbations of non-resonant linear PDEs*, Nonlinearity **12** (1999), no. 4, 823–850. https://doi.org/10.1088/0951-7715/12/4/305

[3] ______, *An averaging theorem for quasilinear Hamiltonian PDEs*, Ann. Henri Poincaré, **4** (2003), no. 4, 685–712. https://doi.org/10.1007/s00023-003-0144-6

[4] ______, *Birkhoff normal form for some nonlinear PDEs*, Comm. Math. Phys. **234** (2003), no. 2, 253–285. https://doi.org/10.1007/s00220-002-0774-4

[5] D. Bambusi and B. Grébert, *Birkhoff normal form for partial differential equations with tame modulus*, Duke Math. J. **135** (2006), no. 3, 507–567. https://doi.org/10.1215/s0012-7094-06-13534-2

[6] G. Benettin, L. Galgani and A. Giorgilli, *A proof of Nekhoroshev's theorem for the stability times in nearly integrable Hamiltonian systems*, Celestial Mech. **37** (1985), no. 1, 1–25. https://doi.org/10.1007/bf01230338

[7] J. Bourgain, *On diffusion in high-dimensional Hamiltonian systems and PDE*, J. Anal. Math. **80** (2000), 1–35. https://doi.org/10.1007/bf02791532

[8] J. Colliander, M. Keel, G. Staffilani, H. Takaoka and T. Tao, *Transfer of energy to high frequencies in the cubic defocusing nonlinear Schrödinger equation*, Invent. Math. **181** (2010), no. 1, 39–113. https://doi.org/10.1007/s00222-010-0242-2

[9] L. H. Eliasson, B. Grébert and S. B. Kuksin, *KAM for the nonlinear beam equation*, Geom. Funct. Anal. **26** (2016), no. 6, 1588–1715. https://doi.org/10.1007/s00039-016-0390-7

[10] L. H. Eliasson and S. B. Kuksin, *KAM for the nonlinear Schrödinger equation*, Ann. of Math. (2) **172** (2010), no. 1, 371–435. https://doi.org/10.4007/annals.2010.172.371

[11] E. Faou and B. Grébert, *A Nekhoroshev-type theorem for the nonlinear Schrödinger equation on the torus*, Anal. PDE **6** (2013), no. 6, 1243–1262. https://doi.org/10.2140/apde.2013.6.1243

[12] J. Geng and J. You, *A KAM theorem for Hamiltonian partial differential equations in higher dimensional spaces*, Comm. Math. Phys. **262** (2006), no. 2, 343–372. https://doi.org/10.1007/s00220-005-1497-0

[13] J. Geng, X. Xu and J. You, *An infinite dimensional KAM theorem and its application to the two dimensional cubic Schrödinger equation*, Adv. Math. **226** (2011), no. 6, 5361–5402. https://doi.org/10.1016/j.aim.2011.01.013

[14] A. Giorgilli and L. Galgani, *Rigorous estimates for the series expansions of Hamiltonian perturbation theory*, Celestial Mech. **37** (1985), no. 2, 95–112. https://doi.org/10.1007/bf01230921

[15] B. Grébert, *Birkhoff normal form and Hamiltonian PDEs*, in *Partial Differential Equations and Applications*, 1–46, Sémin. Congr. **15**, Soc. Math. France, Paris, 2007.

[16] M. Guardia, *Growth of Sobolev norms in the cubic nonlinear Schrödinger equation with a convolution potential*, Comm. Math. Phys. **329** (2014), no. 1, 405–434. https://doi.org/10.1007/s00220-014-1977-1

[17] M. Guardia, E. Haus and M. Procesi, *Growth of Sobolev norms for the analytic NLS on* $\mathbb{T}^2$, Adv. Math. **301** (2016), 615–692. https://doi.org/10.1016/j.aim.2016.06.018

[18] M. Guardia and V. Kaloshin, *Growth of Sobolev norms in the cubic defocusing nonlinear Schrödinger equation*, J. Eur. Math. Soc. (JEMS) **17** (2015), no. 1, 71–149. https://doi.org/10.4171/jems/499

[19] N. N. Nekhoroshev, *An exponential estimate of the time of stability of nearly integrable Hamiltonian systems*, Uspehi Mat. Nauk **32** (1977), no. 6, 1–65. https://doi.org/10.1070/rm1977v032n06abeh003859

[20] J. Pöschel, *Nekhoroshev estimates for quasi-convex Hamiltonian systems*, Math. Z. **213** (1993), no. 1, 187–216. https://doi.org/10.1007/bf03025718

[21] M. Procesi and C. Procesi, *A normal form for the Schrödinger equation with analytic non-linearities*, Comm. Math. Phys. **312** (2012), no. 2, 501–557. https://doi.org/10.1007/s00220-012-1483-2

[22] C. Procesi and M. Procesi, *A KAM algorithm for the resonant non-linear Schrödinger equation*, Adv. Math. **272** (2015), 399–470.
https://doi.org/10.1016/j.aim.2014.12.004

[23] X. Yuan and J. Zhang, *Long time stability of Hamiltonian partial differential equations*, SIAM J. Math. Anal. **46** (2014), no. 5, 3176–3222.
https://doi.org/10.1137/120900976

Shidi Zhou and Jiansheng Geng
Department of Mathematics, Nanjing University, Nanjing 210093, P. R. China
E-mail address: mathsdzhou@126.com, jgeng@nju.edu.cn