Chin. Ann. Math. Ser. B
40(3), 2019, 389–410
DOI: 10.1007/s11401-019-0140-x

Chinese Annals of Mathematics, Series B

A Nekhoroshev Type Theorem for the Nonlinear Wave Equation in Gevrey Space*

Chunyong LIU[1] Huayong LIU[2] Rong ZHAO[3]

Abstract In this paper, the authors prove a Nekhoroshev type theorem for the nonlinear wave equation

$$u_{tt} = u_{xx} - mu - f(u), \quad x \in [0, \pi]$$

in Gevrey space.

Keywords Gevrey space, Nonlinear wave equation, Normal form, Stability
2000 MR Subject Classification 35B40, 35Q55, 37K55

1 Introduction

We consider the nonlinear wave (NLW for short) equation

$$u_{tt} = u_{xx} - mu - f(u) \tag{1.1}$$

on the finite x-interval $[0, \pi]$ with Dirichlet boundary conditions

$$u(t, 0) = u(t, \pi) = 0, \quad t \in \mathbb{R}.$$

The parameter m is real and positive, and the nonlinearity f is assumed to be real analytic in u and of the form

$$f(u) = au^3 + \sum_{k \geq 2} f_k u^{2k+1}, \quad a \neq 0. \tag{1.2}$$

(1.1) is a typical model of infinite-dimensional Hamiltonian system associated with the Hamiltonian function

$$H = \frac{1}{2}\langle v, v\rangle + \frac{1}{2}\langle Au, u\rangle + \int_0^\pi g(u)\mathrm{d}x,$$

Manuscript received April 19, 2016. Revised August 21, 2017.
[1]School of Mathematical Sciences, Dalian University of Technology, Dalian 116024, Liaoning, China; Shangdong Chemical Engineering & Vocational College, Weifang 261108, Shandong, China.
E-mail: 403020931@mail.dlut.edu.cn
[2]School of Mathematical Sciences, Dalian University of Technology, Dalian 116024, Liaoning, China.
E-mail: 1039822663@qq.com
[3]Dalian Airforce Communication Non-Commissioned Officer Academy, Dalian 116600, Liaoning, China.
E-mail: pd81520@sina.com
*This work was supported by the National Natural Science Foundation of China (No. 11671066) and the Fundamental Research Funds for the Central Universities (No. DUT15LK05).

where $v = u_t$, $A := -\frac{\mathrm{d}^2}{\mathrm{d}x^2} + m$, $g = \int f(s)\mathrm{d}s$ and $\langle\cdot,\cdot\rangle$ denotes the usual scalar product in L^2. (1.1) is well studied by many authors as an infinite dimensional Hamiltonian system, such as the long time stability result (see [1–4, 7–10, 19]) and the existence of invariant tori by Kolmogorov-Arnold-Moser (KAM for short) theory (see [5, 11, 16–18, 21–26]).

In [3], Bambusi proved a Birkhoff normal form theorem which is applied to (1.1) with Dirichlet boundary conditions and obtained the dynamical consequences on the long time behavior of the solutions with small initial Cauchy data in Sobolev spaces H^s. Afterwards, Bambusi and Grébert [6] gave an abstract Birkhoff normal form theorem for Hamiltonian partial differential equations, which is applied to semi-linear equations with nonlinearity satisfying tame modulus. They got that for s sufficiently large, the Sobolev norm of index s of the solution is bounded by 2ϵ during a polynomial long time (of order ϵ^{-r} with r arbitrary). Later, Cong-Liu-Yuan [14] and Cong-Gao-Liu [13] generalized the method in [6] and proved the KAM tori are stable in a polynomial long time for nonlinear Schrödinger equation and nonlinear wave equation.

Recently, in [15], Faou and Grébert proved a Nekhoroshev type theorem for the nonlinear Schrödinger equation

$$\mathrm{i}u_t = -\Delta u + V * u + f(|u|^2)u, \quad x \in \mathbb{T}^d$$

in an analytic space. The authors proved that if the initial datum is analytic in a strip of width $\rho > 0$ with a bound on this strip equal to ϵ then, if ϵ is small enough, the solution of the nonlinear Schrödinger equation above remains analytic in a strip of width $\frac{\rho}{2}$ and bounded on this strip by $C\epsilon$ during very long time of order $\epsilon^{-\alpha|\ln\epsilon|^\delta}$ for some constants $C > 0, \alpha > 0$ and $0 < \delta < 1$. We should point out that there is no so-called tame property in analytic space compared to that in Sobolev space. Later, Mi-Liu-Shi-Zhao [20] generalized the method in [15] to prove the similar result for (1.1) with Dirichlet boundary.

As Bourgain [12] said, the topology is very important for studying the long time stability result. On the other hand, the Gevrey space is a phase space which is well studied for infinite-dimensional Hamiltonian system (see [9–10]). It is natural to ask that whether such a stability result holds in Gevrey space. In this paper, we will prove that if the Gevrey norm of initial datum is small in a strip of width $2\rho > 0$ with a bound on this strip equal to ϵ, then, if ϵ is small enough, the Gevrey norm of the solution of the nonlinear wave equation above remains small in a strip of width $\frac{\rho}{2}$ and bounded on this strip by $C\epsilon$ during very long time of order $\epsilon^{-\sigma|\ln\epsilon|^\beta}$ for some constants $C > 0, \sigma > 0$ and $0 < \beta < \frac{1}{7}$.

Before stating the result, we firstly introduce the Gevrey space. Suppose a function $u : [0,\pi] \to \mathbb{C}$ that can be expressed as

$$u(x) = \sum_{j\in\mathbb{Z}} \widehat{u}_j \mathrm{e}^{\mathrm{i}jx}.$$

For $\rho > 0$, we denote

$$\mathcal{A}_\rho \equiv \mathcal{A}_\rho([0,\pi];\mathbb{C}) = \Big\{u \mid |u|_\rho := \sum_{j\in\mathbb{Z}} |\widehat{u}_j| \mathrm{e}^{\rho\sqrt{|j|}} < +\infty\Big\}. \tag{1.3}$$

Note that $(\mathcal{A}_\rho, |\cdot|_\rho)$ is a Banach space. Then our main result is as follows.

Theorem 1.1 *There exist* $0 < \beta < \frac{1}{7}$ *and* $\rho > 0$, *and the following holds*: *There exist constants* $C > 0$ *and* $\epsilon_0 > 0$ *such that if*

$$u_0, v_0 \in \mathcal{A}_{2\rho} \quad \textit{and} \quad |u_0|_{2\rho} + |v_0|_{2\rho} = \epsilon \leq \epsilon_0,$$

then the solution of (1.1) *with initial datum* u_0 *and* v_0 *exists in* $\mathcal{A}_{\frac{\rho}{2}}$ *for times* $|t| \leq \epsilon^{-\sigma_\rho |\ln \epsilon|^\beta}$ *and satisfies*

$$|u(t)|_{\frac{\rho}{2}} \leq C\epsilon \quad \textit{for } |t| \leq \epsilon^{-\sigma_\rho |\ln \epsilon|^\beta} \tag{1.4}$$

with $\sigma_\rho = \min\left\{\frac{1}{10}, \frac{\rho}{2}\right\}$.

The rest of the paper is organized as follow. In Section 2, (1.1) is turned into an infinite dimensional Hamiltonian system in complex coordinates. And we give an important lemma which will be used to prove the main result. In Section 3, some important definitions are given. Then, we obtain estimate of the nonlinearity, vector field and Passion bracket. In Section 4, we introduce recursive equation and give two lemmas which we will use to get the normal form result (i.e., Theorem 4.1). In Section 5, two important lemmas are given. Theorem 1.1 is proved and the long time stability of solutions to NLW equation is obtained in Gevrey space.

2 Hamiltonian System

We study (1.1) as an infinite dimensional Hamiltonian system. As the phase space one may take, for example, the product of the usual Sobolev space $H_0^1([0,\pi]) \times L^2([0,\pi])$ with coordinates u and $v = u_t$, i.e.,

$$u_t = v = \frac{\partial H}{\partial v}, \quad v_t = u_{xx} - mu - f(u) = -\frac{\partial H}{\partial u}, \tag{2.1}$$

then the Hamiltonian is

$$H = \frac{1}{2}\int_0^\pi (v^2 - |u_x|^2 + mu^2)\mathrm{d}x + \int_0^\pi g(u)\mathrm{d}x, \tag{2.2}$$

where $g = \int f(s)\mathrm{d}s$. Let $A := -\frac{\mathrm{d}^2}{\mathrm{d}x^2} + m$. Then the second equation of (2.1) is changed into

$$v_t = -\frac{\partial H}{\partial u} = -Au - f(u),$$

and the Hamiltonian equations (2.2) is changed into

$$H = \frac{1}{2}\langle v, v\rangle + \frac{1}{2}\langle Au, u\rangle + \int_0^\pi g(u)\mathrm{d}x, \tag{2.3}$$

where $\langle \cdot, \cdot \rangle$ denotes the usual scalar product in L^2. To simplify calculation, we changed it into a Hamiltonian in infinitely many coordinates. Therefore, suppose

$$u = \sum_{j\geq 1} \frac{q_j}{\sqrt{\lambda_j}}\phi_j, \quad v = \sum_{j \geq 1} \sqrt{\lambda_j} p_j \phi_j, \tag{2.4}$$

where $\phi_j = \sqrt{\frac{2}{\pi}}\sin jx$ for $j = 1, 2, \cdots$ are the normalized Dirichlet eigenfunctions of the operator A with eigenvalues $\lambda_j^2 = j^2 + m$. The coordinates are taken from some Banach space

$\mathcal{L}_\rho$ ($\rho > 0$) of all real valued sequences $w = (w_1, w_2, \cdots)$ with finite norm

$$\|w\|_\rho := \sum_{j\geq 1} |w_j| \mathrm{e}^{\sqrt{j}\rho}.$$

Then the Hamiltonian (2.3) turns into

$$H = \frac{1}{2}\sum_{j\geq 1} \lambda_j (p_j^2 + q_j^2) + \int_0^\pi g\Big(\sum_{j\geq 1} \frac{q_j}{\sqrt{\lambda_j}}\phi_j\Big)\mathrm{d}x \tag{2.5}$$

with equations of motions turned into

$$\dot{q}_j = \frac{\partial H}{\partial p_j} = \lambda_j p_j, \quad \dot{p}_j = -\frac{\partial H}{\partial q_j} = -\lambda_j q_j - \frac{\partial F}{\partial q_j}, \quad j \geq 1, \tag{2.6}$$

where

$$F(q) = \int_0^\pi g\Big(\sum_{j\geq 1} \frac{q_j}{\sqrt{\lambda_j}}\phi_j\Big)\mathrm{d}x. \tag{2.7}$$

These are Hamiltonian equations of motion with respect to the standard symplectic structure $\sum\limits_{j\geq 1} \mathrm{d}q_j \wedge \mathrm{d}p_j$ on $\mathcal{L}_\rho \times \mathcal{L}_\rho$. Since the nonlinearity f in (1.1) is real analytic in a neighborhood of zero and of the form (1.2), we have

$$g(u) = \sum_{k=2}^{+\infty} \frac{g^{(2k)}(0)}{(2k)!} u^{2k}. \tag{2.8}$$

According to (2.7)–(2.8), we get

$$F(q) = \int_0^\pi \sum_{k=2}^{+\infty} \frac{g^{(2k)}(0)}{(2k)!}\Big(\sum_{j\geq 1} \frac{q_j}{\sqrt{\lambda_j}}\phi_j\Big)^{2k}\mathrm{d}x = \sum_{k=2}^{+\infty} \sum_{\substack{j_1\pm\cdots\pm j_{2k}=0\\ j_1,\cdots,j_{2k}>0}} F_{j_1,\cdots,j_{2k}} q_{j_1}\cdots q_{j_{2k}}, \tag{2.9}$$

where

$$F_{j_1,\cdots,j_{2k}} = \frac{g^{(2k)}(0)}{(2k)!} \frac{1}{\sqrt{\lambda_{j_1}\cdots\lambda_{j_{2k}}}} \int_0^\pi \phi_{j_1}\cdots\phi_{j_{2k}}\mathrm{d}x. \tag{2.10}$$

Let $\mathcal{Z} := \mathbb{Z}^1 \setminus \{0\}$. If the Hamiltonian is defined on the complex Banach space $\mathcal{L}_{\rho,b}$ collecting all the two-side sequences with norm

$$\|w\|_\rho := \sum_{j\in\mathcal{Z}} |w_j| \mathrm{e}^{\sqrt{|j|}\rho},$$

the corresponding symplectic structure is $\mathrm{i}\sum\limits_{j\geq 1} \mathrm{d}w_j \wedge \mathrm{d}w_{-j}$. Then for a function P of $\mathcal{C}^1(\mathcal{L}_{\rho,b}, \mathbb{C})$, we define its Hamiltonian vector field by $X_P = J\nabla P$, where J is the symplectic operator on $\mathcal{L}_{\rho,b}$. For two functions P and Q, we define the Poisson bracket by

$$\{P, Q\} = \nabla P^T J \nabla Q = \mathrm{i}\sum_{j\geq 1} \frac{\partial P}{\partial w_{-j}}\frac{\partial Q}{\partial w_j} - \frac{\partial P}{\partial w_j}\frac{\partial Q}{\partial w_{-j}}. \tag{2.11}$$

If $\overline{w}_j = w_{-j}$, we say that $w \in \mathcal{L}_{\rho,b}$ is real. Similarly, if $H(w)$ is real for all real $w \in \mathcal{L}_{\rho,b}$, we also say that the Hamiltonian H is real.

Definition 2.1 *For a given $\rho > 0$, we define the space of real Hamiltonian P by $\mathcal{H}_\rho$ satisfying*

$$P \in \mathcal{C}^1(\mathcal{L}_{\rho,b}, \mathbb{C}) \quad \textit{and} \quad X_P \in \mathcal{C}^1(\mathcal{L}_{\rho,b}, \mathcal{L}_{\rho,b}).$$

Obviously, for P and Q in $\mathcal{H}_\rho$, the formula (2.11) is well defined. For a given Hamiltonian function $H \in \mathcal{H}_\rho$, we associate the Hamiltonian system

$$\dot{w} = X_H(w) = J\nabla H(w),$$

which is equivalent to

$$\dot{w}_j = -\mathrm{i}\frac{\partial H}{\partial w_{-j}}, \quad \dot{w}_{-j} = \mathrm{i}\frac{\partial H}{\partial w_j}, \quad j \geq 1. \tag{2.12}$$

We define the local flow $\Phi_H^t(w)$ associated with the above system. Note that if both w and H are real, the flow is also real, i.e., $\Phi_H^t(w)$ is real for all t.

Then, let us introduce the complex coordinates

$$z_j = \frac{1}{\sqrt{2}}(q_j + \mathrm{i}p_j), \quad \overline{z}_j = \frac{1}{\sqrt{2}}(q_j - \mathrm{i}p_j), \quad j \geq 1. \tag{2.13}$$

To simplify calculation, we introduce another set of coordinates $(\cdots, w_{-2}, w_{-1}, w_1, w_2, \cdots)$ in $\mathcal{L}_{\rho,b}$ by setting

$$w_j = z_j, \quad w_{-j} = \overline{z}_j \quad \text{for } j \geq 1. \tag{2.14}$$

Therefore, the system (2.6) is turned into

$$\dot{w}_j = -\mathrm{i}\lambda_j w_j - \mathrm{i}\,\mathrm{sgn}\, j\frac{\partial F}{\partial w_{-j}}, \quad j \neq 0 \tag{2.15}$$

with Hamiltonian (2.5) changed into

$$H(w) = \sum_{j\geq 1} \lambda_j w_j w_{-j} + F(w),$$

where $\lambda_j = \mathrm{sgn}\, j\sqrt{j^2 + m}$ and

$$\begin{aligned} F(w) &= \sum_{k=2}^{+\infty} \sum_{\substack{j_1 \pm \cdots \pm j_{2k} = 0 \\ j_1, \cdots, j_{2k} > 0}} F_{j_1,\cdots,j_{2k}} \frac{z_{j_1} + \overline{z}_{j_1}}{\sqrt{2}} \cdots \frac{z_{j_{2k}} + \overline{z}_{j_{2k}}}{\sqrt{2}} \\ &= \sum_{k=2}^{+\infty} \sum_{\substack{j_1 \pm \cdots \pm j_{2k} = 0 \\ j_1, \cdots, j_{2k} \neq 0}} \frac{1}{(\sqrt{2})^{2k}} F_{j_1,\cdots,j_{2k}} w_{j_1} \cdots w_{j_{2k}}. \end{aligned} \tag{2.16}$$

Notice that $F_{j_1,\cdots,j_{2k}} = F_{|j_1|,\cdots,|j_{2k}|}$.

In the end of the section, we give a lemma that shows the relation between the space $\mathcal{A}_\rho$ and the space $\mathcal{L}_{\rho,b}$, which will be used in the proof of the main result.

Lemma 2.1 *Let u, v be valued infinite differentiable functions on the closure of the x-interval $[0, \pi]$, and let $(w_j)_{j\in\mathcal{Z}}$ be the sequence of its coordinates defined by* (2.4) *and* (2.13)–(2.14). *Then for any $\mu < \rho$, we have*

$$\textit{if } u, v \in \mathcal{A}_\rho, \ \textit{then } w \in \mathcal{L}_{\mu,b} \ \textit{and} \ \ \|w\|_\mu \leq c_{\mu,\rho}(|u|_\rho + |v|_\rho), \tag{2.17}$$

$$\textit{if } w \in \mathcal{L}_{\rho,b}, \ \textit{then } u, v \in \mathcal{A}_\mu \ \textit{and} \ \ |u|_\mu, |v|_\mu \leq c_{\mu,\rho}\|w\|_\rho, \tag{2.18}$$

where $c_{\mu,\rho}$ is a constant depending on μ and ρ.

Proof Due to (2.13)–(2.14), we know

$$|w_j| \le \frac{1}{\sqrt{2}}(|p_{|j|}| + |q_{|j|}|) \quad \text{for } j \in \mathcal{Z}$$

and

$$p_j = \frac{1}{\sqrt{2}i}(w_j - w_{-j}), \quad q_j = \frac{1}{\sqrt{2}}(w_j + w_{-j}), \quad j \ge 1.$$

Due to (2.4), it is clear to know that

$$\begin{aligned} u(x) &= \sqrt{\frac{2}{\pi}} \sum_{j\ge 1} \frac{q_j}{\sqrt{\lambda_j}} \sin jx \\ &= \sqrt{\frac{2}{\pi}} \sum_{j\ge 1} \frac{q_j}{\sqrt{\lambda_j}} \frac{\mathrm{e}^{\mathrm{i}jx} - \mathrm{e}^{-\mathrm{i}jx}}{2\mathrm{i}} \\ &= \frac{1}{2\mathrm{i}} \sqrt{\frac{2}{\pi}} \sum_{j\ne 0} \frac{\operatorname{sgn} j q_{|j|}}{\sqrt{\lambda_{|j|}}} \mathrm{e}^{\mathrm{i}jx} = \sum_{j\in\mathbb{Z}} \widehat{u}_j \mathrm{e}^{\mathrm{i}jx} \end{aligned}$$

and

$$\begin{aligned} v(x) &= \sqrt{\frac{2}{\pi}} \sum_{j\ge 1} \sqrt{\lambda_j} p_j \sin jx \\ &= \sqrt{\frac{2}{\pi}} \sum_{j\ge 1} \sqrt{\lambda_j} p_j \frac{\mathrm{e}^{\mathrm{i}jx} - \mathrm{e}^{-\mathrm{i}jx}}{2\mathrm{i}} \\ &= \frac{1}{2\mathrm{i}} \sqrt{\frac{2}{\pi}} \sum_{j\ne 0} \sqrt{\lambda_{|j|}} \operatorname{sgn} j p_{|j|} \mathrm{e}^{\mathrm{i}jx} = \sum_{j\in\mathbb{Z}} \widehat{v}_j \mathrm{e}^{\mathrm{i}jx}. \end{aligned}$$

So, for $\mu < \rho$ and for any $u, v \in \mathcal{A}_\rho$, we have the estimation

$$\begin{aligned} \|w\|_\mu &= \sum_{j\in\mathcal{Z}} |w_j| \mathrm{e}^{\mu\sqrt{|j|}} \\ &\le \frac{1}{\sqrt{2}} \sum_{j\in\mathcal{Z}} (|p_{|j|}| + |q_{|j|}|) \mathrm{e}^{\mu\sqrt{|j|}} \\ &\le \sqrt{\pi} \sum_{j\in\mathcal{Z}} \Big(\sqrt{\lambda_{|j|}} |\widehat{u}_j| + \frac{1}{\sqrt{\lambda_{|j|}}} |\widehat{v}_j| \Big) \mathrm{e}^{\mu\sqrt{|j|}} \\ &\le \sqrt{\pi} \sum_{j\in\mathbb{Z}} \sqrt{\lambda_{|j|}} (|\widehat{u}_j| + |\widehat{v}_j|) \mathrm{e}^{\mu\sqrt{|j|}} \\ &\le \sqrt{\pi} \sum_{j\in\mathbb{Z}} (|\widehat{u}_j| + |\widehat{v}_j|) \mathrm{e}^{\rho\sqrt{|j|}} \sqrt{\lambda_{|j|}} \mathrm{e}^{(\mu-\rho)\sqrt{|j|}} \\ &\le c(|u|_\rho + |v|_\rho) \sum_{j\in\mathbb{Z}} \sqrt{\lambda_{|j|}} \mathrm{e}^{(\mu-\rho)\sqrt{|j|}} \le c_{\mu,\rho}(|u|_\rho + |v|_\rho) \end{aligned}$$

with $\sum\limits_{j\in\mathbb{Z}} \sqrt{\lambda_{|j|}}\mathrm{e}^{(\mu-\rho)\sqrt{|j|}}$ convergent. Namely, (2.17) is proved. Conversely,

$$\begin{aligned}|u|_\mu &= \sum_{j\in\mathbb{Z}} |\widehat{u}_j|\mathrm{e}^{\mu\sqrt{|j|}}\\ &\le \frac{1}{\sqrt{2\pi}}\sum_{j\in\mathcal{Z}}\Big|\frac{q_{|j|}}{\sqrt{\lambda_{|j|}}}\Big|\mathrm{e}^{\mu\sqrt{|j|}}\\ &\le c\sum_{j\in\mathcal{Z}}\Big|\frac{w_j}{\sqrt{\lambda_{|j|}}}\Big|\mathrm{e}^{\mu\sqrt{|j|}}\\ &\le c\sum_{j\in\mathcal{Z}}|w_j|\mathrm{e}^{\rho\sqrt{|j|}}\frac{1}{\sqrt{\lambda_{|j|}}}\mathrm{e}^{(\mu-\rho)\sqrt{|j|}}\\ &\le c\|w\|_\rho\sum_{j\in\mathcal{Z}}\frac{1}{\sqrt{\lambda_{|j|}}}\mathrm{e}^{(\mu-\rho)\sqrt{|j|}} \le c_{\mu,\rho}\|w\|_\rho\end{aligned}$$

with $\sum\limits_{j\in\mathcal{Z}} \frac{1}{\sqrt{\lambda_{|j|}}}\mathrm{e}^{(\mu-\rho)\sqrt{|j|}}$ convergent. Similarly, we have

$$|v|_\mu \le c_{\mu,\rho}\|w\|_\rho.$$

Therefore, (2.18) is also proved.

3 Space of Polynomial and Some Properties

Firstly, let us introduce some terminologies and notions about the polynomial on $\mathbb{C}^{\mathcal{Z}}$. Let $\ell \ge 2$ and $\mathbf{j} = (j_1, j_2, \cdots, j_\ell) \in \mathcal{Z}^\ell$. We define

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

$$w_{\mathbf{j}} = w_{j_1}\cdots w_{j_\ell},$$

(2) the divisor associated with $\mathbf{j}$,

$$\Omega(\mathbf{j}) = \lambda_{j_1} + \cdots + \lambda_{j_\ell}, \tag{3.1}$$

where for $j_i \in \mathcal{Z},\ \lambda_{j_i} = \operatorname{sgn} j_i\sqrt{j_i^2 + m},\ i = 1, 2, \cdots, \ell$.

In addition, we also denote the set of indices with zero momentum by

$$\mathcal{I}_\ell = \{\mathbf{j} = (j_1, j_2, \cdots, j_\ell) \in \mathcal{Z}^\ell \mid j_1 \pm j_2 \pm \cdots \pm j_\ell = 0\}. \tag{3.2}$$

Furthermore, if ℓ is even and $\mathbf{j}$ is of the form $(j_1, -j_1, \cdots, j_{\frac{\ell}{2}}, -j_{\frac{\ell}{2}})$ or some permutation of it, we say that $\mathbf{j} = (j_1, j_2, \cdots, j_\ell) \in \mathcal{Z}^\ell$ is resonant writing $\mathbf{j}\in \mathcal{N}_\ell$. Specially, if $\mathbf{j}$ is resonant, its associated divisor vanishes, i.e., $\Omega(\mathbf{j}) = 0$, and its associated monomials depend only on the actions

$$w_{\mathbf{j}} = w_{j_1}\cdots w_{j_\ell} = w_{j_1}w_{-j_1}\cdots w_{j_{\frac{\ell}{2}}}w_{-j_{\frac{\ell}{2}}} = I_{j_1}\cdots I_{j_{\frac{\ell}{2}}}, \tag{3.3}$$

where for $j \ge 1,\ I_j = w_j w_{-j}$ denotes the action associated with the index j.

We note that if w is real, then $I_j = |w_j|^2$ and if ℓ is odd, the resonant set $\mathcal{N}_\ell$ is the empty set. So, we know that the orders of monomials in the nonlinearity F are all even.

Definition 3.1 *If P is real, then for $k \geq 2$, a formal polynomial $P(w) = \Sigma a_{\mathbf{j}} w_{\mathbf{j}} \in \mathcal{P}_k$, of degree k, has a zero of order at least 2 in $w = 0$, and satisfies the following conditions*:

(1) *P contains only monomials having zero momentum (i.e., such that $\mathbf{j} \in \mathcal{I}_\ell$ for some ℓ, when $a_{\mathbf{j}} \neq 0$), and P denotes*

$$P(w) = \sum_{\ell=2}^{k} \sum_{\mathbf{j} \in \mathcal{I}_\ell} a_{\mathbf{j}} w_{\mathbf{j}}, \tag{3.4}$$

where $a_{\mathbf{j}} = a_{|\mathbf{j}|}$, $|\mathbf{j}| = (|j_1|, \cdots, |j_\ell|)$.

(2) *The coefficients $a_{\mathbf{j}}$ are bounded, i.e., $\sup\limits_{\mathbf{j} \in \mathcal{I}_\ell} |a_{\mathbf{j}}| < +\infty$ for all $\ell = 2, \cdots, k$.*

We endow $\mathcal{P}_k$ with the norm

$$\|P\| = \sum_{\ell=2}^{k} \sup_{\mathbf{j} \in \mathcal{I}_\ell} |a_{\mathbf{j}}|. \tag{3.5}$$

Recall that we assume that the nonlinearity f in (1.1) is complex analytic in a neighbourhood of zero in $\mathbb{C}$. Therefore, there exist two positive constants M and R_0 such that the Taylor expansion of its primary function (2.8) is uniformly convergent and bounded by M on the district of $|u| \leq R_0$ of $\mathbb{C}$. So the formula (2.9) defines an analytic function on the ball $\|w\|_\rho \leq R_0$ of $\mathcal{L}_{\rho,b}$ and we have

$$F(w) = \sum_{k \geq 2} P_{2k},$$

where $P_{2k} \in \mathcal{P}_{2k}$ is a homogeneous polynomial of degree $2k$. By Cauchy integral formula, (2.10) and (2.16), we have

$$\|P_{2k}\| = \sup_{\mathbf{j} \in \mathcal{I}_{2k}} \frac{|F_{j_1, \cdots, j_{2k}}|}{(\sqrt{2})^{2k}} \leq \frac{|g^{(2k)}(0)|}{(2k)!(\sqrt{\pi})^{2k-2}} \leq M R_0^{-2k}. \tag{3.6}$$

Finally, we will give some useful estimates of the polynomial space which are used in the following section. And we note that the zero momentum will play an important role in the estimates.

Proposition 3.1 *For $k \geq 2$ and $\rho > 0$, we obtain $\mathcal{P}_k \subset \mathcal{H}_\rho$. Moreover, for any homogeneous polynomial $F \in \mathcal{P}_k$, of degree k, we get the following two estimates*:

$$|F(w)| \leq \|F\| \, \|w\|_\rho^k \tag{3.7}$$

and

$$\|X_F(w)\|_\rho \leq 2^{k-1} k \|F\| \, \|w\|_\rho^{k-1} \quad \text{for all } w \in \mathcal{L}_{\rho,b}. \tag{3.8}$$

Proof Let

$$F(w) = \sum_{\mathbf{j} \in \mathcal{I}_k} a_{\mathbf{j}} w_{\mathbf{j}}.$$

We obtain

$$|F(w)| \le \|F\| \sum_{\mathbf{j}\in\mathcal{Z}^k} |w_{j_1}| \cdots |w_{j_k}| \le \|F\| \, \|w\|_{l^1}^k \le \|F\| \, \|w\|_\rho^k,$$

where $\|\cdot\|_{l^1}$ denotes the l^1-norm of vector. Thus (3.7) is proved.

To prove the second estimate, let us take $\ell \in \mathcal{Z}$, by using the zero momentum condition, we have

$$\Big|\frac{\partial F}{\partial w_\ell}\Big| \le k\|F\| \sum_{\substack{\mathbf{j}\in\mathcal{Z}^{k-1}\\ j_1\pm j_2\pm\cdots\pm j_{k-1}=\pm\ell}} |w_{j_1}\cdots w_{j_{k-1}}|.$$

So we have the estimate

$$\|X_F(w)\|_\rho = \sum_{\ell\in\mathcal{Z}} \mathrm{e}^{\rho\sqrt{|\ell|}}\Big|\frac{\partial F}{\partial w_\ell}\Big| \le k\|F\| \sum_{\ell\in\mathcal{Z}} \sum_{\substack{\mathbf{j}\in\mathcal{Z}^{k-1}\\ j_1\pm j_2\pm\cdots\pm j_{k-1}=\pm\ell}} \mathrm{e}^{\rho\sqrt{|\ell|}} |w_{j_1}\cdots w_{j_{k-1}}|.$$

Due to $j_1 \pm j_2 \pm \cdots \pm j_{k-1} = \pm\ell$, then we have the inequality

$$\mathrm{e}^{\rho\sqrt{|\ell|}} \le \exp(\rho(\sqrt{|j_1|} + \cdots + \sqrt{|j_{k-1}|})) \le \prod_{n=1}^{k-1} \mathrm{e}^{\rho\sqrt{|j_n|}}.$$

Therefore, after summing in $j_1, \cdots, j_{k-1}$ and ℓ, we have

$$\|X_F(z)\|_\rho \le 2^{k-1}k\|F\| \sum_{\mathbf{j}\in\mathcal{Z}^{k-1}} \mathrm{e}^{\rho\sqrt{|j_1|}}|w_{j_1}| \cdots \mathrm{e}^{\rho\sqrt{|j_{k-1}|}}|w_{j_{k-1}}| \le 2^{k-1}k\|F\| \, \|w\|_\rho^{k-1}.$$

So (3.8) is also proved.

Proposition 3.2 *If F and G are homogeneous polynomials of degree k and ℓ respectively in $\mathcal{P}_k$ and $\mathcal{P}_\ell$, then $\{F, G\} \in \mathcal{P}_{k+\ell-2}$ and we have the estimate*

$$\|\{F, G\}\| \le 2^{\min\{k,\ell\}-1} k\ell \|F\| \, \|G\|. \tag{3.9}$$

Proof Now we assume that F and G are homogeneous polynomials of degrees k and ℓ respectively and with coefficients $a_{\mathbf{k}}$, $\mathbf{k} \in \mathcal{I}_k$ and $b_{\mathbf{l}}$, $\mathbf{l} \in \mathcal{I}_\ell$. It is clear that $\{F, G\}$ is a homogeneous polynomial of degree $k + \ell - 2$ satisfying the zero momentum condition. Furthermore, we can write

$$\{F, G\}(w) = \sum_{\mathbf{j}\in\mathcal{I}_{k+\ell-2}} c_{\mathbf{j}} w_{\mathbf{j}},$$

where $c_{\mathbf{j}}$ is expressed as a sum of coefficients $a_{\mathbf{k}} b_{\mathbf{l}}$ for which there exists a $j \in \mathcal{Z}$ such that

$$j \subset \mathbf{k} \in \mathcal{I}_k, \quad -j \subset \mathbf{l} \in \mathcal{I}_\ell,$$

and such that if for instance $j = k_1$ and $-j = \ell_1$, we necessarily have $(k_2, \cdots, k_k, \ell_2, \cdots, \ell_\ell) = \mathbf{j}$. Hence, for a given $\mathbf{j}$, the zero momentum condition on $\mathbf{k}$ and on $\mathbf{l}$ determines the value of j which in turn determines $2^{\min\{k,\ell\}-1}$ possible values of j.

This proves (3.9) for monomials. The extension to polynomials follows from the definition of the norm (3.5).

The last assertion and the fact that the Poisson bracket of two real Hamiltonian is real follow immediately from the definitions.

4 Recursive Equation and Normal Form Results

In this section, let us firstly introduce the definition of the N-normal form. For $r \geq 4$ and $\mathbf{j} = (j_1, \cdots, j_r) \in \mathcal{Z}^r$, we denote the third largest integer amongst $|j_1|, \cdots, |j_r|$ by $\mu(\mathbf{j})$, and set

$$\mathcal{J}_k(N) = \{\mathbf{j} \in \mathcal{I}_k \mid \mu(\mathbf{j}) > N\} \quad \text{for } N \geq 1 \text{ and } k \geq 4.$$

Definition 4.1 (N-Normal Form) *Let N be an integer. We say that a polynomial $W \in \mathcal{P}_k$ is in N-normal form if it can be written as*

$$W = \sum_{\ell=4}^{k} \sum_{\mathbf{j} \in \mathcal{N}_\ell \cup \mathcal{J}_\ell(\mathbf{N})} a_{\mathbf{j}} w_{\mathbf{j}}.$$

Namely, W contains either monomials depending only on the actions or monomials whose indices $\mathbf{j}$ satisfy $\mu(\mathbf{j}) > N$, i.e., monomials involving at least three modes with index greater than N.

4.1 Recursive equation

Firstly, we give a lemma which is an easy consequence of the nonresonance condition and the definition of the normal forms.

Lemma 4.1 *Firstly, we suppose that the nonresonance condition* (6.6) *is satisfied, and let N be fixed. Also suppose that $H_0 := \sum\limits_{j \geq 1} \lambda_j w_j w_{-j}$ is the integrable part of Hamiltonian* (2.5) *and Q is a homogenous polynomial of degree n. Then the homological equation*

$$\{\chi, H_0\} - W = Q \tag{4.1}$$

admits a polynomial solution (χ, W) homogenous of degree n such that W is in N-normal form, and such that

$$\|W\| \leq \|Q\| \quad \text{and} \quad \|\chi\| \leq N^{16n^6} \|Q\|. \tag{4.2}$$

Proof Assume that $Q = \sum\limits_{\mathbf{j} \in \mathcal{I}_n} Q_{\mathbf{j}} w_{\mathbf{j}}$ and find $W = \sum\limits_{\mathbf{j} \in \mathcal{I}_n} W_{\mathbf{j}} w_{\mathbf{j}}$ and $\chi = \sum\limits_{\mathbf{j} \in \mathcal{I}_n} \chi_{\mathbf{j}} w_{\mathbf{j}}$ such that (4.1) is satisfied. (4.1) can be written in term of polynomial coefficients

$$-\mathrm{i}\Omega(\mathbf{j})\chi_{\mathbf{j}} - W_{\mathbf{j}} = Q_{\mathbf{j}}, \quad \mathbf{j} \in \mathcal{I}_n,$$

where $\Omega(\mathbf{j})$ is given in (3.1). We then define

(1) $W_{\mathbf{j}} = -Q_{\mathbf{j}},\ \chi_{\mathbf{j}} = 0$ if $\mathbf{j} \in \mathcal{N}_n$ or $\mu(\mathbf{j}) > N$,

(2) $W_{\mathbf{j}} = 0,\ \chi_{\mathbf{j}} = -\frac{Q_{\mathbf{j}}}{\mathrm{i}\Omega(\mathbf{j})}$ if $\mathbf{j} \notin \mathcal{N}_n$ and $\mu(\mathbf{j}) \leq N$.

In view of (6.6), this leads to (4.2).

In the following section, we will introduce the recursive equation. The solutions of recursive equation can generate a canonical transformation Φ such that in the new variables, the Hamiltonian $H_0 + F$ is in normal form modulo a small remainder term. To obtain the recursive equation, we consider the following problem.

Seek polynomials $\chi = \sum\limits_{n=4}^{r} \chi_n$ and $W = \sum\limits_{n=4}^{r} W_n$ in normal form and a smooth Hamiltonian R satisfying $\partial^\alpha R(0) = 0$ for all $\alpha \in \mathbb{N}^{\mathcal{Z}}$ with $|\alpha| \leq r$, such that

$$(H_0 + F) \circ \Phi_\chi^1 = H_0 + W + R. \tag{4.3}$$

Recall that for two Hamiltonian functions χ and K, we have for all $k \geq 0$,

$$\frac{\mathrm{d}^k}{\mathrm{d}t^k}(K \circ \Phi_\chi^t) = \{\chi, \{\cdots\{\chi, K\}\cdots\}\}(\Phi_\chi^t) = (\mathrm{ad}_\chi^k K)(\Phi_\chi^t),$$

where $\mathrm{ad}_\chi K = \{\chi, K\}$. Moreover, if K and L are homogeneous polynomials of degree respectively n and ℓ, then $\{K, L\}$ is a homogeneous polynomial of degree $n + \ell - 2$. Hence, we obtain

$$(H_0 + F) \circ \Phi_\chi^1 - (H_0 + F) = \sum_{k=0}^{\frac{r}{2}-2} \frac{1}{(k+1)!} \mathrm{ad}_\chi^k(\{\chi, H_0 + F\}) + \mathcal{O}_r \tag{4.4}$$

by using the Taylor formula, where $\mathcal{O}_r$ stands for any smooth function R satisfying $\partial^\alpha R(0) = 0$ for all $\alpha \in \mathbb{N}^{\mathcal{Z}}$ with $|\alpha| \leq r$. On the other hand, we know that for $\zeta \in \mathbb{C}$, the following relation holds:

$$\Big(\sum_{k=0}^{\frac{r}{2}-2} \frac{B_k}{k!}\zeta^k\Big)\Big(\sum_{k=0}^{\frac{r}{2}-2} \frac{1}{(k+1)!}\zeta^k\Big) = 1 + O(|\zeta|^{\frac{r}{2}-1}),$$

where B_k are the Bernoulli numbers defined by the expansion of the generating function $\frac{z}{\mathrm{e}^z - 1}$. Hence, by defining the two differential operators

$$A_r = \sum_{k=0}^{\frac{r}{2}-2} \frac{1}{(k+1)!} \mathrm{ad}_\chi^k, \quad B_r = \sum_{k=0}^{\frac{r}{2}-2} \frac{B_k}{k!} \mathrm{ad}_\chi^k,$$

we get

$$B_r A_r = \mathrm{Id} + C_r,$$

where C_r is a differential operator satisfying

$$C_r \mathcal{O}_{\frac{r}{2}+2} = \mathcal{O}_r.$$

Applying B_r to the two sides of (4.4), we obtain

$$\{\chi, H_0 + F\} = B_r(W - F) + \mathcal{O}_r.$$

Plugging the decompositions in homogeneous polynomials of χ, W and F in the last equation and equating the terms of same degree, after a straightforward calculation, we obtain the recursive equations

$$\{\chi_n, H_0\} - W_n = Q_n, \quad n = 4, \cdots, r, \tag{4.5}$$

where

$$\begin{aligned} Q_n = & -P_n + \sum_{k=4}^{n-2} \{P_{n+2-k}, \chi_k\} \\ & + \sum_{k=1}^{\frac{n}{2}-2} \frac{B_k}{k!} \sum_{\substack{\ell_1 + \cdots + \ell_{k+1} = n+2k \\ 4 \leq \ell_i \leq n-2k}} \mathrm{ad}_{\chi_{\ell_1}} \cdots \mathrm{ad}_{\chi_{\ell_k}} (W_{\ell_{k+1}} - P_{\ell_{k+1}}). \end{aligned} \tag{4.6}$$

In the last sum, $\ell_i \leq n-2k$ appears as a consequence of $\ell_i \geq 4$ and $\ell_1+\cdots+\ell_{k+1} = n+2k$. Once these recursive equations solved, we define the remainder term as $R = (H_0+F)\circ\Phi_\chi^1 - H_0 - W$. By construction, R is analytic on a neighborhood of the origin in $\mathcal{L}_{\rho,b}$ and $R = \mathcal{O}_r$. So, by the Taylor's formula,

$$\begin{aligned} R = &\sum_{n\geq r+1}\sum_{k=2}^{\frac{n}{2}-1}\frac{1}{k!}\sum_{\substack{\ell_1+\cdots+\ell_k=n+2k-2\\ 4\leq \ell_i\leq r}} \mathrm{ad}_{\chi_{\ell_1}}\cdots \mathrm{ad}_{\chi_{\ell_k}} H_0 \\ &+ \sum_{n\geq r+1}\sum_{k=0}^{\frac{n}{2}-2}\frac{1}{k!}\sum_{\substack{\ell_1+\cdots+\ell_{k+1}=n+2k\\ 4\leq \ell_1,\cdots,\ell_k\leq r\\ 4\leq \ell_{k+1}}} \mathrm{ad}_{\chi_{\ell_1}}\cdots \mathrm{ad}_{\chi_{\ell_k}} P_{\ell_{k+1}}. \end{aligned} \tag{4.7}$$

Lemma 4.2 *Suppose that the nonresonance condition* (6.6) *is fulfilled for some constants* γ,ν. *Then there exists* $C>0$ *such that for all* r, N, *and for* $n = 4,\cdots,r$, *there exist two homogeneous polynomials* χ_n *and* W_n *of degree* n, *with* W_n *in* N*-normal form, which are solutions of the recursive equation* (4.5) *and satisfy*

$$\|\chi_n\| + \|W_n\| \leq (C4^n nN^{16})^{n^7}. \tag{4.8}$$

Proof We define χ_n and W_n by induction using Lemma 4.1. Note that (4.8) is clearly satisfied for $n=4$, provided C big enough. Estimate (4.2) yields

$$N^{-16n^6}\|\chi_n\| + \|W_n\| \leq \|Q_n\|. \tag{4.9}$$

Using the definition (4.6) of the term Q_n and the estimate on the Bernoulli numbers, $|B_k| \leq k!C^k$ for some $C>0$, together with (3.9), which implies that for all $\ell \geq 4$, $\|\mathrm{ad}_{\chi_\ell} R\| \leq 2^{\min\{n,\ell\}-1} n\ell\|R\|\,\|\chi_\ell\|$ for any polynomial R of degree less than n, we have for all $n\geq 4$,

$$\begin{aligned} \|Q_n\| \leq\; &\|P_n\| + 2^n\sum_{k=4}^{n-2} k(n+2-k)\|P_{n+2-k}\|\,\|\chi_k\| \\ &+ 2\sum_{k=1}^{\frac{n}{2}-2}(Cn)^k \sum_{\substack{\ell_1+\cdots+\ell_{k+1}=n+2k\\ 4\leq \ell_i\leq n-2k}} C(n,k)\ell_1\|\chi_{\ell_1}\|\cdots\ell_k\|\chi_{\ell_k}\|\,\|W_{\ell_{k+1}} - P_{\ell_{k+1}}\|, \end{aligned} \tag{4.10}$$

where

$$C(n,k) = 2^{\min\{\ell_1,n+2-\ell_1\}-1}2^{\min\{\ell_2,n+2\cdot 2-\ell_1-\ell_2\}-1}\cdots 2^{\min\{\ell_k,\ell_{k+1}\}-1}$$

and C is a constant. It is easy to know $C(n,k)\leq 4^n$.

We set $\beta_n = n(\|\chi_n\| + \|W_n\|)$. Equation (4.9) implies that

$$\beta_n \leq (CN^\nu)^{n^6} n\|Q_n\| \tag{4.11}$$

for some constant C independent of n.

By the fact that $\|P_n\| \leq MR_0^{-n}$ (see (3.6)), we obtain

$$\beta_n \leq \beta_n^{(1)} + \beta_n^{(2)},$$

where

$$\beta_n^{(1)} = (CN^\nu)^{n^6} 2^n n^3 \sum_{k=4}^{n-2} \beta_k \tag{4.12}$$

and

$$\beta_n^{(2)} = N^{\nu n^6}(Cn)^{n-1}4^n \sum_{k=1}^{\frac{n}{2}-2} \sum_{\substack{\ell_1+\cdots+\ell_{k+1}=n+2k \\ 4\le \ell_i \le n-2k}} \beta_{\ell_1}\cdots\beta_{\ell_k}(\beta_{\ell_{k+1}} + \|P_{\ell_{k+1}}\|), \tag{4.13}$$

where C depends on M and R_0. It remains to prove by induction that $\beta_n \le (C4^n nN^{16})^{n^7}$. Assume that $\beta_j \le (C4^j jN^{16})^{j^7},\ j = 4, \cdots, n-1$. Then for $C > 1$, we have

$$(C4^n nN^{16})^{n^7} \ge 1 \quad \text{for all } \ n \ge 4, \tag{4.14}$$

so we get

$$\beta_n^{(1)} \le (CN^{16})^{n^6} 2^n n^4 (C4^n nN^{16})^{(n-1)^7} \le \frac{1}{2}(C4^n nN^{16})^{n^7} \quad \text{for } n \ge 4$$

provided $C > 2$.

Using (4.14) again and the induction hypothesis, we obtain

$$\beta_n^{(2)} \le N^{16n^6}(Cn)^{n-1}4^n \sum_{k=1}^{\frac{n}{2}-2} \sum_{\substack{\ell_1+\cdots+\ell_{k+1}=n+2k \\ 4\le \ell_i \le n-2k}} (CN^{16}4^{n-1}(n-2k))^{\ell_1^7+\cdots+\ell_{k+1}^7}.$$

Notice that the maximum of $\ell_1^7+\cdots+\ell_{k+1}^7$ when $\ell_1+\cdots+\ell_{k+1} = n+2k$ and $4 \le \ell_i \le n-2k$ is obtained for $\ell_1 = \cdots = \ell_k = 4$ and $\ell_{k+1} = n-2k$ and its value is $(n-2k)^7 + 4^7k$. Furthermore, the cardinality of $\{\ell_1 + \cdots + \ell_{k+1} = n+2k,\ 4 \le \ell_i \le n-2k\}$ is smaller than n^{k+1}, and hence we obtain

$$\beta_n^{(2)} \le \max_{k=\left\{1,\cdots,\frac{n}{2}-2\right\}} N^{16n^6}(Cn)^{n-1}Cn^{k+2}4^n(CN^{16}4^n(n-2k))^{(n-2k)^7+4^7k} \le \frac{1}{2}(C4^n nN^{16})^{n^7}$$

for $n \ge 5$ and after adapting C if necessary.

4.2 Normal form result

For any $R_1 > 0$, we set $B_\rho(R_1) = \{w \in \mathcal{L}_{\rho,b} \mid \|w\|_\rho < R_1\}$.

Theorem 4.1 *Suppose that F is analytic on a ball $B_\rho(R_1)$ for some $R_1 > 0$ and $\rho > 0$. Also suppose that the nonresonance condition* (6.6) *is satisfied, and let $\beta < \frac{1}{7}$ and $M > 1$ be fixed. Then there exists a constant $\epsilon_0 > 0$ such that for all $\epsilon < \epsilon_0$, there exist a polynomial χ, a polynomial W in N-normal form, and a Hamiltonian R analytic on $B_\rho(M\epsilon)$, such that*

$$(H_0 + F) \circ \Phi_\chi^1 = H_0 + W + R. \tag{4.15}$$

Furthermore, for all $w \in B_\rho(M\epsilon)$,

$$\|X_W(w)\|_\rho + \|X_\chi(w)\|_\rho \le 2\epsilon^{\frac{3}{2}}, \quad \|X_R(w)\|_\rho \le \epsilon^2 \mathrm{e}^{-\frac{1}{4}|\ln\epsilon|^{1+\beta}}. \tag{4.16}$$

Proof Using Lemma 4.2, for all N and r, we can construct polynomial Hamiltonians

$$\chi(w)=\sum_{k=4}^{r}\chi_k(w) \quad \text{and} \quad W(w)=\sum_{k=4}^{r}W_k(w),$$

with W in N-normal form, such that (4.15) holds with $R=\mathcal{O}_r$. Now for fixed $\epsilon>0$, we choose

$$N=|\ln\epsilon|^{2+2\beta} \quad \text{and} \quad r=|\ln\epsilon|^{\beta}.$$

This choice is motivated by the necessity of balance between W and R in (4.15). The error induced by W is controlled as in Lemma 5.2, while the error induced by R is controlled by Lemma 4.2. By (4.8), we get

$$\begin{aligned}\|\chi_k\| &\le (C4^k kN^{16})^{k^7}\le \exp(k(16k^6(2+2\beta)\ln|\ln\epsilon|+k^7\ln 4+k^6\ln Ck))\\ &\le \exp(k(\nu r^6(2+2\beta)\ln|\ln\epsilon|+r^7\ln 4+r^6\ln Cr))\\ &\le \exp(k|\ln\epsilon|(16|\ln\epsilon|^{6\beta-1}(2+2\beta)\ln|\ln\epsilon|+|\ln\epsilon|^{7\beta-1}\ln 4+|\ln\epsilon|^{6\beta-1}\ln C|\ln\epsilon|^{\beta}))\\ &\le \epsilon^{-\frac{k}{8}} \end{aligned}\tag{4.17}$$

as $\beta<\frac17$ (we take $\beta<\frac17$ such that $\max\{6\beta-1,7\beta-1\}<0$), and for $\epsilon<\epsilon_0$ sufficiently small. So using Proposition 3.1, we have

$$|\chi_k(w)|\le \epsilon^{-\frac{k}{8}}(M\epsilon)^k\le M^k\epsilon^{7\frac{k}{8}} \quad \text{for } w\in B_\rho(M\epsilon)$$

and thus

$$|\chi(w)|\le\sum_{k\ge4}M^k\epsilon^{\frac{7k}{8}}\le\epsilon^{\frac32}$$

for ϵ small enough.

Similarly, we have for all $k\le r$,

$$\|X_{\chi_k}(w)\|_\rho\le 2^{k-1}k\epsilon^{-\frac{k}{8}}(M\epsilon)^{k-1}\le k(2M)^{k-1}\epsilon^{\frac{7k}{8}-1}$$

and

$$\|X_\chi(w)\|_\rho\le\sum_{k\ge4}k(2M)^{k-1}\epsilon^{\frac{7k}{8}-1}\le C\epsilon^{-1}\epsilon^{\frac{28}{8}}\le\epsilon^{\frac32}$$

for ϵ small enough. Similar bounds clearly hold for $W=\sum\limits_{k=4}^{r}W_k$, which shows the first estimate in (4.16).

On the other hand, using $\mathrm{ad}_{\chi_{\ell_k}}H_0=W_{\ell_k}+Q_{\ell_k}$ (see (4.5)) and then combining Lemma 4.2 with the definition of Q_n (see (4.6)), we can obtain

$$\|\mathrm{ad}_{\chi_{\ell_k}}H_0\|\le(C4^{\ell_k}\ell_kN^{16})^{\ell_k^7}\le\epsilon^{-\frac{\ell_k}{8}},$$

where the last inequality proceeds as in (4.17). Therefore, due to (4.7), (4.17) and $\|P_{\ell_{k+1}}\|\le MR_0^{-\ell_{k+1}}$, we obtain by Proposition 3.1 that for $w\in B_\rho(M\epsilon)$,

$$\|X_R(w)\|_\rho\le\sum_{n\ge r+1}\sum_{k=0}^{\frac{n}{2}-2}4^nn(Cr)^{3n}\epsilon^{-\frac{n+2k}{8}}\epsilon^{n-1}\le\sum_{n\ge r+1}n^2(4Cr)^{3n}\epsilon^{\frac{n}{2}}\le(4Cr)^{3r}\epsilon^{\frac{r}{2}}.$$

Since $r=|\ln\epsilon|^\beta>2$ (ϵ small enough), we get $\|X_R(w)\|_\rho\le\epsilon^2\mathrm{e}^{-\frac14|\ln\epsilon|^{1+\beta}}$ for $w\in B_\rho(M\epsilon)$.

5 Proof of the Main Result

Firstly, before giving the proof of the main theorem, we will introduce two important lemmas.

Lemma 5.1 (see [15]) *Let $f:\mathbb{R}\to\mathbb{R}_+$ be a continuous function and $y:\mathbb{R}\to\mathbb{R}_+$ be a differentiable function satisfying the inequality*

$$\frac{\mathrm{d}}{\mathrm{d}t}y(t)\le 2f(t)\sqrt{y(t)}\quad \textit{for all } t\in\mathbb{R}.$$

Then we have the estimate

$$\sqrt{y(t)}\le\sqrt{y(0)}+\int_0^t f(s)\mathrm{d}s\qquad \textit{for all } t\in\mathbb{R}.$$

Fix $N>1$. For all $w\in\mathcal{L}_{\rho,b}$, we set

$$R_\rho^N(w)=\sum_{|j|>N}\mathrm{e}^{\rho\sqrt{|j|}}|w_j|.$$

Notice that if $w\in\mathcal{L}_{\rho+\mu,b}$, then

$$R_\rho^N(w)\le\mathrm{e}^{-\mu\sqrt{N}}\|w\|_{\rho+\mu}. \tag{5.1}$$

Lemma 5.2 *Let $N\in\mathbb{N}$ and $k\ge4$. Suppose that W is a homogeneous polynomial of degree k in N-normal form. Let $w(t)$ be a real solution of the flow generated by the Hamiltonian H_0+W. Then we have*

$$R_\rho^N(w(t))\le R_\rho^N(w(0))+4k^3 2^{k-1}\|W\|\int_0^t R_\rho^N(w(s))^2\|w(s)\|_\rho^{k-3}\mathrm{d}s \tag{5.2}$$

and

$$\|w(t)\|_\rho\le\|w(0)\|_\rho+4k^3 2^{k-1}\|W\|\int_0^t R_\rho^N(w(s))^2\|w(s)\|_\rho^{k-3}\mathrm{d}s. \tag{5.3}$$

Proof Fix $j\in\mathcal{Z}$ and let $I_j(t)=w_j(t)w_{-j}(t)$ be the actions associated with the solution of the Hamiltonian system generated by H_0+W. Due to (3.9), we can obtain

$$|\mathrm{e}^{2\rho\sqrt{|j|}}\dot{I}_j|=|\mathrm{e}^{2\rho\sqrt{|j|}}\{I_j,W\}|\le 2^{k-1}k\|W\|\,|\,\mathrm{e}^{\rho\sqrt{|j|}}\sqrt{I_j}\,|\,\Big(\sum_{\substack{j_1\pm\cdots\pm j_{k-1}=\pm j\\ 2\text{ indices}>N}}\mathrm{e}^{\rho\sqrt{|j|}}|w_{j_1}\cdots w_{j_{k-1}}|\Big).$$

Then using Lemma 5.1, we can get

$$\begin{aligned}&\mathrm{e}^{\rho\sqrt{|j|}}\sqrt{I_j(t)}\\ &\le\mathrm{e}^{\rho\sqrt{|j|}}\sqrt{I_j(0)}+2^{k-1}k\|W\|\int_0^t\Big(\sum_{\substack{j_1\pm\cdots\pm j_{k-1}=\pm j\\ 2\text{ indices}>N}}\mathrm{e}^{\rho\sqrt{|j_1|}}|w_{j_1}|\cdots\mathrm{e}^{\rho\sqrt{|j_{k-1}|}}|w_{j_{k-1}}|\Big)\mathrm{d}s.\end{aligned} \tag{5.4}$$

Ordering the multi-indices in such a way that $|j_1|$ and $|j_2|$ are the largest, and making use of the fact that $w(t)$ is real (and thus $|w_j|=\sqrt{I_j}$), we have, after summation in $|j|>N$,

$$\begin{aligned}R_\rho^N(w(t))&\le R_\rho^N(w(0))+4k^3 2^{k-1}\|W\|\int_0^t\Big(\sum_{\substack{|j_1|,|j_2|\ge N\\ j_3,\cdots,j_{k-1}\in\mathcal{Z}}}\mathrm{e}^{\rho\sqrt{|j_1|}}|w_{j_1}|\cdots\mathrm{e}^{\rho\sqrt{|j_{k-1}|}}|w_{j_{k-1}}|\Big)\mathrm{d}s\\ &\le R_\rho^N(w(0))+4k^3 2^{k-1}\|W\|\int_0^t R_\rho^N(w(s))^2\|w(s)\|_\rho^{k-3}\mathrm{d}s.\end{aligned}$$

Inequality (5.3) can be proved in the same way.

Now we are in position to prove the main theorem in Section 1 in which we will take advantage of the bootstrap argument.

Proof of Theorem 1.1 Let $u_0, v_0 \in \mathcal{A}_{2\rho}$ with $|u_0|_{2\rho} + |v_0|_{2\rho} = \epsilon$, and denotes by $w(0)$ the corresponding sequence of its Fourier coefficients which belongs to $\mathcal{L}_{\frac{3}{2}\rho,b}$ with $\|w(0)\|_{\frac{3}{2}\rho} \le \frac{c_\rho}{4}\epsilon$ by Lemma 2.1. Let $w(t)$ be the local solution in $\mathcal{L}_{\rho,b}$ of the Hamiltonian system associated with $H = H_0 + F$.

Let χ, W and R given by Theorem 4.1 with $M = c_\rho$ and let $y(t) = \Phi_\chi^1(w(t))$. We recall that since $\chi(w) = O(\|w\|^4)$, the transformation Φ_χ^1 is close to the identity, $\Phi_\chi^1(w) = w + O(\|w\|^3)$ and thus, for ϵ small enough, we have $\|y(0)\|_{\mathcal{L}_{\frac{3}{2}\rho}} \le \frac{c_\rho}{2}\epsilon$. Specially, note that

$$R_\rho^N(y(0)) \le \frac{c_\rho}{2}\epsilon \mathrm{e}^{-\frac{\rho}{2}\sqrt{N}} \le \frac{c_\rho}{2}\epsilon \mathrm{e}^{-\sigma\sqrt{N}},$$

where $\sigma = \sigma_\rho \le \frac{\rho}{2}$.

Let T_ϵ be the maximum of time T such that $R_\rho^N(y(t)) \le c_\rho\epsilon\mathrm{e}^{-\sigma\sqrt{N}}$ and $\|y(t)\|_\rho \le c_\rho\epsilon$ for all $|t| \le T_\epsilon$. By construction, we have

$$y(t) = y(0) + \int_0^t X_{H_0+W}(y(s))\mathrm{d}s + \int_0^t X_R(y(s))\mathrm{d}s.$$

So using (5.2) for the first vector field and (4.16) for the second one, we get for $|t| \le T_\epsilon$,

$$\begin{aligned} R_\rho^N(y(t)) &\le \frac{1}{2}c_\rho\epsilon\mathrm{e}^{-\sigma\sqrt{N}} + 4|t|\sum_{k=4}^r \|W_k\|k^3(2c_\rho\epsilon)^{k-1}\mathrm{e}^{-2\sigma\sqrt{N}} + |t|\epsilon^2\mathrm{e}^{-\frac{1}{4}|\ln\epsilon|^{1+\beta}} \\ &\le \Big(\frac{1}{2} + 4|t|\sum_{k=4}^r \|W_k\|k^3(2c_\rho\epsilon)^{k-2}\mathrm{e}^{-\sigma\sqrt{N}} + |t|\epsilon\mathrm{e}^{-\frac{1}{8}\sqrt{|\ln\epsilon|^{1+\beta}}}\Big)c_\rho\epsilon\mathrm{e}^{-\sigma\sqrt{N}}, \end{aligned} \tag{5.5}$$

where in the last inequality we have used $\sigma = \min\left\{\frac{1}{10}, \frac{\rho}{2}\right\}$ and $N = |\ln\epsilon|^{2+2\beta}$.

Using Lemma 4.2, we then verify

$$R_\rho^N(y(t)) \le \Big(\frac{1}{2} + C|t|\epsilon\mathrm{e}^{-\sigma\sqrt{N}}\Big)c_\rho\epsilon\mathrm{e}^{-\sigma\sqrt{N}}$$

and thus, for ϵ small enough,

$$R_\rho^N(y(t)) \le c_\rho\epsilon\mathrm{e}^{-\sigma\sqrt{N}} \quad \text{for all } |t| \le \min\{T_\epsilon, \mathrm{e}^{\sigma\sqrt{N}}\}. \tag{5.6}$$

Similarly, we obtain

$$\|y(t)\|_\rho \le c_\rho\epsilon \quad \text{for all } |t| \le \min\{T_\epsilon, \mathrm{e}^{\sigma\sqrt{N}}\}. \tag{5.7}$$

In view of the definition of T_ϵ, (5.6)–(5.7) imply $T_\epsilon \ge \mathrm{e}^{\sigma\sqrt{N}}$. Specially, $\|w(t)\|_\rho \le 2c_\rho\epsilon$ for $|t| \le \mathrm{e}^{\sigma\sqrt{N}} = \epsilon^{-\sigma|\ln\epsilon|^\beta}$. Using (2.18), we finally obtain (1.4).

6 Appendix

In this section, we will give some technical lemmas and nonresonance condition. This section can be also find in [20].

Lemma 6.1 *For any $K \le r$, consider K indexes $j_1 < \cdots < j_K \le N$, and consider the determinant*

$$D := \begin{vmatrix} \lambda_{j_1} & \lambda_{j_2} & \cdots & \lambda_{j_K} \\ \dfrac{\mathrm{d}\lambda_{j_1}}{\mathrm{d}m} & \dfrac{\mathrm{d}\lambda_{j_2}}{\mathrm{d}m} & \cdots & \dfrac{\mathrm{d}\lambda_{j_K}}{\mathrm{d}m} \\ \vdots & \vdots & & \vdots \\ \dfrac{\mathrm{d}^{K-1}\lambda_{j_1}}{\mathrm{d}m^{K-1}} & \dfrac{\mathrm{d}^{K-1}\lambda_{j_2}}{\mathrm{d}m^{K-1}} & \cdots & \dfrac{\mathrm{d}^{K-1}\lambda_{j_K}}{\mathrm{d}m^{K-1}} \end{vmatrix}. \tag{6.1}$$

It holds

$$|D| = \Big[\prod_{j=1}^{K-1} \frac{(2j-3)!}{2^{j-2}(j-1)!2^j}\Big] = \Big(\prod_{l=1}^{K} \lambda_{j_l}^{-2K+1}\Big)\Big(\prod_{1\le l<k\le K} (j_l)^2 - (j_k)^2\Big) \ge \frac{C}{N^{2K^2}}. \tag{6.2}$$

Proof By explicit computation, one has

$$\frac{\mathrm{d}^n\lambda_j}{\mathrm{d}m^n} = \begin{cases} \dfrac{1}{2^n}(j^2+m)^{\frac{1}{2}-n}, & 0 \le n \le 1, \\ \dfrac{(2n-3)!}{2^{n-2}(n-1)!2^n}\dfrac{(-1)^n}{(j^2+m)^{n-\frac{1}{2}}}, & 2 \le n \le K-1. \end{cases} \tag{6.3}$$

Substituting (6.3) into the right hand site of (6.1), we get the determinant to be estimated. To obtain the estimate factorize from the j-th column term $\lambda_j = (j^2+m)^{\frac{1}{2}}$, and from the n-th row term $\frac{(2n-3)!}{2^{n-2}(n-1)!2^n}$. Forgetting the essential power of -1, we obtain that the determinant to be estimated is given by

$$\Big[\prod_{l=1}^{K}\lambda_{j_l}\Big]\Big[\frac{1}{2}\prod_{n=2}^{K-1}\frac{(2n-3)!}{2^{n-2}(n-1)!2^n}\Big]\begin{vmatrix} 1 & 1 & \cdots & 1 \\ x_{j_1} & x_{j_2} & \cdots & x_{j_K} \\ \vdots & \vdots & & \vdots \\ x_{j_1}^{K-1} & x_{j_2}^{K-1} & \cdots & x_{j_K}^{K-1} \end{vmatrix},$$

where we denoted by $x_j = (j^2+m)^{-1}$. The last determinant is a Vandermond determinant whose value is given by

$$\prod_{1\le l<n\le K}(x_{j_l} - x_{j_n}). \tag{6.4}$$

Now we have

$$|x_{j_l} - x_{j_n}| = \Big|\frac{1}{j_l^2+m} - \frac{1}{j_n^2+m}\Big| = \frac{|j_n^2 - j_l^2|}{(j_l^2+m)(j_n^2+m)} \ge Cx_{j_l}x_{j_n}$$

with a suitable C. So (6.4) is estimated by

$$\prod_{l=1}^{K-1}\prod_{n=l+1}^{K} Cx_{j_l}x_{j_n} = C^{\sum\limits_{n=2}^{K}(n-1)}\prod_{l=1}^{K-1}\Big(x_{j_l}^{K-l}\prod_{n=l+1}^{K}x_{j_n}\Big) = C\prod_{l=1}^{K}x_{j_l}^{K-1},$$

from which, using the asymptotics of the frequencies, the thesis immediately follows.

Next we need the lemma from [8, Appendix B].

Lemma 6.2 (see [8]) *Let $u^{(1)},\cdots,u^{(K)}$ be K independent vectors with $\|u^{(i)}\|_{l^1}\le 1$. Let $w\in\mathbb{R}^K$ be an arbitrary vector. Then there exists $i\in[1,\cdots,K]$, such that*

$$|u^{(i)}w|\ge\frac{\|w\|_{l^1}\det(u^{(i)})}{K^{\frac{3}{2}}},$$

where $\det(u^{(i)})$ is the determinant of the matrix formed by the components of the vectors $u^{(i)}$.

Proof The proof can be found in the proposition of Appendix B in [8].

Combining Lemmas 6.1 and 6.2, we deduce the following lemma.

Lemma 6.3 *Let $w\in\mathbb{Z}^\infty$ be a vector with K components different from zero, namely those with indices $j_1,\cdots,j_K$. Assume that $K\le r$ and $j_1<\cdots<j_K\le N$. Then for any $m\in[m_0,\Delta]$, there exists an index $j\in[0,\cdots,K-1]$ such that*

$$\Big|w\frac{d^j\lambda}{\mathrm{d}m^j}(m)\Big|\ge C\frac{\|w\|_{l^1}}{N^{2K^2+2}},\tag{6.5}$$

where $\lambda=(\lambda_{j_1},\lambda_{j_2},\cdots,\lambda_{j_K})$ is the frequency vector.

From [24] we learn the following lemma.

Lemma 6.4 *Suppose that $g(m)$ is r times differentiable on an interval $J\subset\mathbb{R}$. Let $J_\gamma:=\{m\in J\mid |g(m)|<\gamma\}$, $\gamma>0$. If $|g^{(r)}(m)|\ge d>0$ on J, then $|J_\gamma|\le M\gamma^{\frac{1}{r}}$, where $M:=2(2+3+\cdots+r+d^{-1})$.*

Proof The proof can be found in [24, Lemma 2.1].

Nonresonance condition In order to control the divisors (3.1), we need to impose the nonresonance condition on the linear frequencies $\lambda_j,\ j\in\mathcal{Z}$.

Recall that $\Omega(\mathbf{j})=\operatorname{sgn}j_1\lambda_{|j_1|}+\operatorname{sgn}j_2\lambda_{|j_2|}+\cdots+\operatorname{sgn}j_r\lambda_{|j_r|}$. We define a set

$$S_\ell=\{s:\ |j_s|=\ell\}$$

and let $k=(k_\ell)_{\ell\in\mathbb{N}}$, where

$$k_\ell=\begin{cases}0, & \text{if } S_\ell=\emptyset,\\ \displaystyle\sum_{s\in S_\ell}\mathrm{sgn}j_s, & \text{if } S_\ell\ne\emptyset.\end{cases}$$

Then $\Omega(\mathbf{j})=\sum\limits_{\ell\ge1}k_\ell\lambda_\ell$ and $|k|\le r$. In the following section, we set $k=(\widetilde{k},\widehat{k})$, where $\widetilde{k}=(k_1,\cdots,k_N)\in\mathbb{Z}^N$, $\widehat{k}=(k_{N+1},\cdots)\in\mathbb{Z}^{\mathbb{N}}$ and we assume that $|\widehat{k}|\le 2$.

Recalling the definition of $\mu(\mathbf{j})$ in Section 4, then we have the following proposition.

Proposition 6.1 *For a given positive number N, there exists a set $\mathcal{J}$ satisfying* $\mathrm{Meas}([m_0,\Delta]-\mathcal{J})\to0$ *as $N\to+\infty$, such that for any $m\in\mathcal{J}$,*

$$|\langle k,\omega^{(N)}\rangle+\varepsilon_1\omega_{j_1}+\varepsilon_2\omega_{j_2}|\ge\frac{\gamma}{N^{16r^6}},\tag{6.6}$$

where $|k|\le r+2$, $\varepsilon_1,\varepsilon_2\in\{-1,0,1\}$, $|j_1|,|j_2|>N$ and $\mu(\mathbf{j})<N$.

Proof For a given positive number N, we define the resonant set $\widetilde{\mathcal{R}}_{kj_1j_2}$ by

$$\widetilde{\mathcal{R}}_{kj_1j_2} = \Big\{m \in [m_0, \Delta] \mid |\langle k, \omega^{(N)}\rangle + \varepsilon_1\omega_{j_1} + \varepsilon_2\omega_{j_2}| < \frac{1}{N^{16r^6}}\Big\}, \tag{6.7}$$

where $|k| \le r+2$, $\varepsilon_1, \varepsilon_2 \in \{-1, 0, 1\}$, $|j_1|, |j_2| > N$.

Case 1 $\varepsilon_1 = \varepsilon_2 = 0$.

Denote the resonant $\mathcal{R}_k = \big\{m \in [m_0, \Delta] \mid |\langle k, \omega^{(N)}\rangle| < \frac{1}{N^{4r^3}}\big\}$. By combining Lemmas 6.3–6.4, we can get

$$\begin{aligned} |\mathcal{R}_k| &\le 2(2+3+\cdots+r+1+C^{-1}N^{2r^2+2})\frac{1}{N^{\frac{4r^3}{r+1}}} \\ &\le \frac{3}{N^{\frac{4r^3}{r+1}-2r^2-1}}, \end{aligned} \tag{6.8}$$

where the last inequality is based on $N > r$ and $C^{-1}N < 1$ if N is large enough. Here $|\cdot|$ denotes the Lebesgue measure of set and C is a constant in Lemma 6.3. We set $\widetilde{R}_1 = \bigcup\limits_{|k|<r+2} \mathcal{R}_k$. Then we have

$$\begin{aligned} |\widetilde{R}_1| &\le \sum_{|k|<r+2} |\mathcal{R}_k| = \sum_{|k|<r+2} \frac{3}{N^{\frac{4r^3}{r+1}-2r^2-1}} \\ &\le \frac{3}{N^{\frac{4r^3}{r+1}-2r^2-1}} N^{r+2} \\ &= \frac{3}{N^{\frac{4r^3}{r+1}-2r^2-r-3}} \\ &< \frac{3}{N^{\frac{2r^3-3r^2-4r-3}{r+1}}} < \frac{1}{9N} \quad (r \text{ is large enough}). \end{aligned} \tag{6.9}$$

Case 2 $\varepsilon_1 = \pm1$, $\varepsilon_2 = 0$ or $\varepsilon_1 = 0$, $\varepsilon_2 = \pm 1$ or $\varepsilon_1\varepsilon_2 = 1$.

In this case, we take $\varepsilon_1 = \pm1$, $\varepsilon_2 = 0$ without loss of generality. Denote the resonant $\mathcal{R}_{kj_1} = \big\{m \in [m_0, \Delta] \mid |\langle k, \omega^{(N)}\rangle + \omega_{j_1}| < \frac{1}{N^{4r^3}}\big\}$. Due to $\omega_{j_1} = \sqrt{j_1^2 + m}$, one has

$$|\langle k, \omega^{(N)}\rangle + \omega_{j_1}| \ge |\omega_{j_1}| - |\langle k, \omega^{(N)}\rangle| \ge 1,$$

when $|j_1| \ge 2(r+2)N+1$. Then the resonant $\mathcal{R}_{kj_1}$ is empty. So we only consider $|j_1| < 2(r+2)N+1$. Setting $\langle \widetilde{k}, \omega^{(\widetilde{N})}\rangle = \langle k, \omega^{(N)}\rangle + \omega_{j_1}$ in place of $\langle k, \omega^{(N)}\rangle$ and $\widetilde{N} = 2(r+2)N+1$ in place of N, then according to Case 1, we have

$$\begin{aligned} |\mathcal{R}_{kj_1}| &\le \frac{3}{(2(r+2)N+1)^{\frac{4r^3}{r+2}-2(r+1)^2-1}} \\ &\le \frac{3}{N^{\frac{2r^3-12r^2-11r-6}{r+2}}}. \end{aligned}$$

Setting $\widetilde{R}_2 = \bigcup_{|k|<r+2} \bigcup_{|j_1|<2(r+2)N+1} \mathcal{R}_{kj_1}$, then we have

$$
\begin{aligned}
|\widetilde{R}_2| &\le \sum_{|k|<r+2} \sum_{|j_1|<2(r+2)N+1} |\mathcal{R}_{kj_1}| \\
&= \sum_{|k|<r+2} \sum_{|j_1|<2(r+2)N+1} \frac{3}{N^{\frac{2r^3-12r^2-11r-6}{r+2}}} \\
&\le \frac{3}{N^{\frac{2r^3-12r^2-11r-6}{r+2}}} N^{r+2}(2(r+2)N+1) \\
&\le \frac{9}{N^{\frac{2r^3-13r^2-17r-14}{r+2}}} < \frac{1}{9N} \quad (r \text{ is large enough}).
\end{aligned}
\tag{6.10}
$$

Case 3 $\varepsilon_1\varepsilon_2 = -1$.

In this case, we take $\varepsilon_1 = 1$, $\varepsilon_2 = -1$ without loss of generality.

Denote the resonant $\mathcal{R}_{kj_1j_2} = \left\{m \in [m_0, \Delta] \mid |\langle k, \omega^{(N)}\rangle + \omega_{j_1} - \omega_{j_2}| < \frac{1}{N^{4r^3}}\right\}$. Assume $j_1 > j_2$ without loss of generality. Because of $\omega_{j_1} = \sqrt{j_1^2+m}, \omega_{j_2} = \sqrt{j_2^2+m}$, then there is a constant $C > 0$ such that

$$
\Big|\frac{\omega_{j_1}-\omega_{j_2}}{j_1-j_2} - 1\Big| \le \frac{C}{j_2}.
$$

Therefore,

$$
\omega_{j_1} - \omega_{j_2} = j_1 - j_2 + r_{j_1j_2}
$$

with

$$
|r_{j_1j_2}| \le \frac{Ca}{j_2}
$$

and $a = j_1 - j_2$. Then we get

$$
|\langle k, \omega^{(N)}\rangle + \omega_{j_1} - \omega_{j_2}| \ge |\langle k, \omega^{(N)}\rangle + a| - |r_{ij}|.
$$

Hence,

$$
\mathcal{R}_{kj_1j_2} \subset \mathcal{R}_{kaj_2} := \Big\{m \in [m_0, \Delta] \mid |\langle k, \omega^{(N)}\rangle + a| \le \frac{1}{N^{4r^3}} + \frac{Ca}{j_2}\Big\}.
$$

If $j > j_0$, we get

$$
\mathcal{R}_{kaj} \subset \mathcal{R}_{kaj_0}.
$$

Then it is sufficient to consider

$$
a \le 2(r+2)N + 1,
$$

and let

$$
j_0 = 2N^{4r^3}.
$$

Setting $\langle \widetilde{k}, \omega^{(\widetilde{N})}\rangle = \langle k, \omega^{(N)}\rangle + \omega_{j_1} - \omega_{j_2}$ in place of $\langle k, \omega^{(N)}\rangle$ and $\widetilde{N} = 2N^{4r^3+1}$ in place of N, then according to Case 1, we have

$$\begin{aligned}|\mathcal{R}_{kj_1j_2}| &\le \frac{3}{(2N^{4r^3+1})^{\frac{4r^3}{r+2}-2(r+1)^2-1}}\\ &\le \frac{3}{N^{\frac{16r^6+2r^3-12r^2-11r-6}{r+2}}}.\end{aligned}$$

Set $\widetilde{R}_2 = \bigcup\limits_{|k|<r+2} \bigcup\limits_{|j_1|<2N^{4r^3+1}} \bigcup\limits_{|j_2|<2N^{4r^3}} \mathcal{R}_{kj_1j_2}$. Then we have

$$\begin{aligned}|\widetilde{R}_3| &\le \sum_{|k|<r+2} \sum_{|j_1|<2N^{4r^3+1}} \sum_{|j_2|<2N^{4r^3}} |\mathcal{R}_{kj_1}|\\ &= \sum_{|k|<r+2} \sum_{|j_1|<2N^{4r^3+1}} \sum_{|j_2|<2N^{4r^3}} \frac{3}{N^{\frac{16r^6+2r^3-12r^2-11r-6}{r+2}}}\\ &\le \frac{3}{N^{\frac{16r^6+2r^3-12r^2-11r-6}{r+2}}} N^{r+2}(2N^{4r^3+1})(2N^{4r^3})\\ &\le \frac{36}{N^{\frac{16r^6-8r^4-6r^3-13r^2-18r-16}{r+2}}} < \frac{1}{9N} \quad (r \text{ is large enough}).\end{aligned} \tag{6.11}$$

In view of (6.9)–(6.11), we obtain

$$|\mathcal{R}| \le \sum_{k,j_1,j_2} \widetilde{\mathcal{R}}_{kj_1j_2} = \widetilde{R}_1 + 6\widetilde{R}_2 + 2\widetilde{R}_3 < \frac{1}{N}.$$

Let $\mathcal{J} = [m_0, \Delta] - \mathcal{R}$. Then the proposition is proved.

Acknowledgements The authors would like to thank Hongzi Cong for his invaluable discussions and suggestions. The authors would also like to thank the anonymous referee for helpful comments and suggestions.

References

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

[2] Bambusi, D., Nekhoroshev theorem for small amplitude solutions in nonlinear Schrödinger equations, *Math. Z.*, **230**, 1999, 345–387.

[3] Bambusi, D., Birkhoff normal form for some nonlinear PDEs, *Comm. Math. Phy.*, **234**, 2003, 253–285.

[4] Bambusi, D., A Birkhoff normal form theorem for some semilinear PDEs, *Hamiltonian dynamical systems and applications*, 2008, 213–247.

[5] Bambusi, D., Berti, M. and Magistrelli, E., Degenerate KAM theory for partial differential equations, *J. Diff. Eqs.*, **250**, 2011, 3379–3397.

[6] Bambusi, D. and Grébert, B., Birkhoff normal form for partial differential equations with tame modulus, *Duke Math. J.*, **135**, 2006, 507–567.

[7] Bambusi, D. and Nekhoroshev, N. N., A property of exponential stability in nonlinear wave equations near the fundamental linear mode, *Phys. D.*, **122**, 1998, 73–104.

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

[9] Bounemoura, A., Nekhoroshev estimates for finitely differentiable quasi-convex Hamiltonians, *J. Diff. Eqs.*, **249**, 2000, 2905–2920.

[10] Bounemoura, A., Normal forms, stability and splitting of invariant manifolds I, Gevrey Hamiltonians, *Regular and Chaotic Dynamics*, **18**, 2013, 237–260.

[11] Bourgain, J., Construction of approximative and almost periodic solutions of perturbed linear Schrödinger and wave equations, *Gemo. Funct. Anal.*, **6**, 1996, 201–230.

[12] Bourgain, J., Remarks on stability and diffusion in high-dimensional Hamiltonian systems and partial differential equations, *Ergodic Theory and Dynamical Systems*, **24**(5), 2004, 1331–1357.

[13] Cong, H., Gao, M. and Liu, J., Long time stability of KAM tori for nonlinear wave equation, *J. Diff. Eqs.*, **258**, 2015, 2823–2846.

[14] Cong, H., Liu, J. and Yuan, X., Stability of KAM tori for nonlinear Schrödinger equation, *Mem. Amer. Math. Soc.*, **239**(1134), 2016, 1–85.

[15] Faou, E. and Grébert, B., A Nekhoroshev-type theorem for the nonlinear Schrödinger equation on the torus, *Analysis and PDE.*, **6**, 2013, 1243–1262.

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

[17] Kuksin, S. B., Perturbations of quasiperiodic solutions of infinite-dimensional Hamiltonian systems, *Math. USSR Izv.*, **32**, 1989, 39–62.

[18] Kuksin, S. B., Nearly Integrable Infinite-Dimensional Hamiltonian Systems, Springer-Verlag, Berlin, 1993.

[19] Marco, J. P. and Sauzin, D., Stability and instability for Gevrey quasi-convex near-integrable Hamiltonian systems, *Publications Mathématiques de l'IHÉS*, **96**, 2003, 199–275.

[20] Mi, L., Liu, C., Shi, G. and Zhao, R., A Nekhoroshev type theorem for the nonlinear wave equation, *Pure and Applied Mathematics Quarterly,* preprint.

[21] Popov, G., KAM theorem for Gevrey Hamiltonians, *Ergodic Theory Dynam. Systems*, **24**, 2004, 1753–1786.

[22] Pöschel, J., Quasi-periodic solutions for nonlinear wave equation, *Comm. Math. Helv.*, **71**, 1996, 269–296.

[23] Wayne, C. E., Periodic and quasi-periodic solutions of nonlinear wave equations via KAM theory, *Comm. Math. Phys.*, **127**, 1990, 479–528.

[24] Xu, J., You, J. and Qiu, Q., Invariant tori for nearly integrable Hamtonian systems with degenercy, *Math. Z.*, **226**, 1997, 375–387.

[25] Yuan, X., Quasi-periodic solutions of nonlinear wave equations with a prescribed potential, *Discrete Contin. Dyn. Syst.*, **16**, 2006, 615–634.

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