Pure and Applied Mathematics Quarterly
Volume 16, Number 5, 1739–1765, 2020

A Nekhoroshev type theorem for the nonlinear wave equation on the torus

Lufang Mi*, Chunyong Liu, Guanghua Shi†, and Rong Zhao

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

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

on the finite x-interval $[0, \pi]$. The parameter m is real and positive, and the nonlinearity f is assumed to be real analytic in u. More precisely, we prove that if the initial datum is analytic in a district of width $2\rho > 0$ whose norm on this district is equal to ϵ, then if ϵ is small enough, the solution of the nonlinear wave equation above remains analytic in a district of width $\rho/2$, with norm bounded on this district by $C\epsilon$ over a very long time interval of order $\epsilon^{-\sigma|\ln\epsilon|^{\beta}}$, where $0 < \beta < 1/7$ is arbitrary and $C > 0$ and $\sigma > 0$ are positive constants depending on β and ρ.

Keywords: Wave equation, Birkhoff normal form, long time stability.

1. Introduction

We consider the nonlinear wave equation

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

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

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

Received August 28, 2017.
2010 Mathematics Subject Classification: Primary 37K55, 37J40; secondary 35B35, 35Q35.
*The first author is supported by Shandong Provincial Natural Science Foundation No. ZR2019MA062 and Binzhou University (BZXYL1402).
†Corresponding author.

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

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

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

$$H = \frac{1}{2} < v, v > + \frac{1}{2} < Au, u > + \int_0^\pi g(u)dx,$$

where $A := -d^2/dx^2 + m$, $g = \int_0 f(s)ds$ and $< \cdot, \cdot >$ denotes the usual scalar product in L^2, which is well studied by many authors, such as the existence of invariant tori by Kolmogorov-Arnold-Moser theory (see [5], [9] and [13]–[20]) and the long time stability result (see [1]–[4], [7] and [8]).

In [3], Bambusi proved a Birkhoff normal form theorem which is applied to equation (1.1) and studied the dynamical consequences on the long time behavior of the solutions with small initial Cauchy data in Sobolev spaces. Afterwards, Bambusi-Grébert [6] proved that for s sufficiently large, if the Sobolev norm of index s of the initial datum u_0 is sufficiently small (of order ϵ), then the Sobolev norm of index s of the solution is bounded by 2ϵ during a very long time (of order ϵ^{-r} with r arbitrary).

In [6], Bambusi-Grébert exploited the tame property of the nonlinear term for some semilinear Hamiltonian PDEs, such as nonlinear Schrödinger equation and nonlinear wave equation to prove the long time stability of the origin by constructing a partial normal form of high order. Later, Cong-Liu-Yuan and Cong-Gao-Liu generalized the method in [6] to prove the KAM tori is stable in a polynomial long time for nonlinear Schrödinger equation and nonlinear wave equation respectively, see [10] and [11].

Recently, a Nekhoroshev type theorem for the nonlinear Schrödinger equation

$$\sqrt{-1}u_t = -\triangle u + V * u + f(|u|^2)u, \qquad x \in \mathbb{T}^d$$

is given by Faou-Grébert in [12] in an analytic space. The authors prove that if the initial datum is analytic in a strip of width $\rho > 0$ with a bound on this strip equals to ε then, if ε is small enough, the solution of the nonlinear Schrödinger equation above remains analytic in a strip of width $\rho/2$ and bounded on this strip by $C\varepsilon$ during very long time of order $\varepsilon^{-\alpha|\ln \varepsilon|^\delta}$ for some constants $C > 0, \alpha > 0$ and $0 < \delta < 1$.

In our paper, we would like to generalize the method in [12] to nonlinear wave equation (1.1). However, as we all know, the frequencies of nonlinear wave equation have worse approximations than nonlinear Schrödinger equation. Fortunately, inspired by the method in [6], we successfully estimate the measure of nonresonance set (see section 4 for the details). In addition, The definition of zero momentum (see (3.2)) is a little different from the one in [12] since the Dirichlet conditions is considered here, which also leads to some worse estimates (see Proposition 3.1 and Proposition 3.2 for details).

To state our result, we will introduce the analytic function space. For $\rho > 0$, we denote by $\mathcal{A}_\rho \equiv \mathcal{A}_\rho([0,\pi];\mathbb{C})$ the space of functions ϕ that are analytic on the complex neighborhood of x-interval $[0,\pi]$ given by $I_\rho = \{x + \mathrm{i}y \mid x \in [0,\pi], y \in \mathbb{R}^1 \text{ and } |y| < \rho\}$ and continuous on the closure of this district. We then denote by $|\cdot|_\rho$ the usual norm on $\mathcal{A}_\rho$:

$$|\phi|_\rho = \sup_{z \in I_\rho} |\phi(z)|.$$

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

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

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

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

$$|u(t)|_{\rho/2} \le C\epsilon \quad \textit{for } |t| \le \epsilon^{-\sigma_\rho |\ln \epsilon|^\beta}, \tag{1.3}$$

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

2. Hamiltonian system

We study the equation (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 $v = u_t$. The Hamiltonian is then

$$H = \frac{1}{2} < v, v > + \frac{1}{2} < Au, u > + \int_0^\pi g(u)dx, \tag{2.1}$$

where $A := -d^2/dx^2 + m$, $g = \int_0 f(s)ds$ and $< \cdot, \cdot >$ denotes the usual scalar product in L^2. The Hamiltonian equations of motions are

$$u_t = \frac{\partial H}{\partial v} = v, \quad v_t = -\frac{\partial H}{\partial u} = -Au - f(u). \tag{2.2}$$

To rewrite it as a Hamiltonian in infinitely many coordinates we make the ansatz

$$u=\sum_{j\geq 1}\frac{q_j}{\sqrt{\lambda_j}}\phi_j,\qquad v=\sum_{j\geq 1}\sqrt{\lambda_j}p_j\phi_j, \tag{2.3}$$

where $\phi_j=\sqrt{2/\pi}\sin jx$ 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

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

Then the Hamiltonian (2.1) turns into

$$H=\frac{1}{2}\sum_{j\geq 1}\lambda_j(p_j^2+q_j^2)+\int_0^\pi g\left(\sum_{j\geq 1}\frac{q_j}{\sqrt{\lambda_j}}\phi_j\right)dx \tag{2.4}$$

with equations of motions

$$\dot{q}_j=\frac{\partial H}{\partial p_j}=\lambda_j p_j,\quad \dot{p}_j=-\frac{\partial H}{\partial q_j}=-\lambda_j q_j-\frac{\partial F}{\partial q_j},\quad j\geq 1, \tag{2.5}$$

where

$$F(q)=\int_0^\pi g\left(\sum_{j\geq 1}\frac{q_j}{\sqrt{\lambda_j}}\phi_j\right)dx. \tag{2.6}$$

These are the Hamiltonian equations of motion with respect to the standard symplectic structure $\sum_{j\geq 1}dq_j\wedge dp_j$ 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

$$g(u)=\sum_{k=4}^{+\infty}\frac{g^{(k)}(0)}{k!}u^k. \tag{2.7}$$

and $g^{(2l+1)}(0)=0,\ l=2,3,\ldots$. Then

$$F(q)=\int_0^\pi\sum_{k=4}^{+\infty}\frac{g^{(k)}(0)}{k!}(\sum_{j\geq 1}\frac{q_j}{\sqrt{\lambda_j}}\phi_j)^k dx=\sum_{k=4}^{+\infty}\sum_{\substack{j_1\pm\cdots\pm j_k=0\\ j_1,\cdots,j_k>0}}F_{j_1,\cdots,j_k}q_{j_1}\cdots q_{j_k}, \tag{2.8}$$

where

$$F_{j_1,\cdots,j_k} = \frac{g^{(k)}(0)}{k!} \cdot \frac{1}{\sqrt{\lambda_{j_1}\cdots\lambda_{j_k}}} \int_0^\pi \phi_{j_1}\cdots\phi_{j_k} dx. \tag{2.9}$$

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

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

and the corresponding symplectic structure is $\mathrm{i}\sum_{j\geq 1} dw_j \wedge dw_{-j}$. 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 $\mathcal{L}_{\rho,b}$ induced by the symplectic structure. For two functions P and Q, the Poisson Bracket is defined as

$$\{P, Q\} = \nabla P^T J \nabla Q = \mathrm{i} \sum_{j\geq 1} \frac{\partial P}{\partial w_{-j}} \frac{\partial Q}{\partial w_j} - \frac{\partial P}{\partial w_j} \frac{\partial Q}{\partial w_{-j}}. \tag{2.10}$$

We say that $w \in \mathcal{L}_{\rho,b}$ is real if $\bar{w}_j = w_{-j}$ and that a Hamiltonian H is real if $H(w)$ is real for all real $w \in \mathcal{L}_{\rho,b}$.

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

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

Clearly, for P and Q in $\mathcal{H}_\rho$ the formula (2.10) is well defined. With 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

$$\dot{w}_j = -\mathrm{i}\frac{\partial H}{\partial w_{-j}} \quad \text{and} \quad \dot{w}_{-j} = \mathrm{i}\frac{\partial H}{\partial w_j}, \quad j \geq 1. \tag{2.11}$$

We define the local flow $\Phi_H^t(w)$ associated with 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.

Now we introduce the complex coordinates

$$z_j = \frac{1}{\sqrt{2}}(q_j + \mathrm{i}p_j), \quad \bar{z}_j = \frac{1}{\sqrt{2}}(q_j - \mathrm{i}p_j). \tag{2.12}$$

It is convenient to introduce another set of coordinates $(\cdots, w_{-2}, w_{-1}, w_1, w_2, \cdots)$ in $\mathcal{L}_{\rho,b}$ by setting

$$z_j = w_j, \qquad \bar{z}_j = w_{-j} \quad \text{for} \quad j \geq 1. \tag{2.13}$$

Then the system (2.5) is changed into

$$\dot{w}_j \;=\; -\mathrm{i}\lambda_j w_j - \mathrm{i}\frac{\partial F}{\partial w_{-j}}, \quad j \neq 0 \tag{2.14}$$

with Hamiltonian

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

where $\lambda_j = sgnj \cdot \sqrt{j^2+m}$ and $F(w)$ is given by

$$\begin{aligned} F &= \sum_{k=4}^{+\infty} \sum_{\substack{j_1\pm\cdots\pm j_k=0\\ j_1,\cdots,j_k>0}} F_{j_1,\cdots,j_k} \frac{z_{j_1}+\bar{z}_{j_1}}{\sqrt{2}} \cdots \frac{z_{j_k}+\bar{z}_{j_k}}{\sqrt{2}} \\ &= \sum_{k=4}^{+\infty} \sum_{\substack{j_1\pm\cdots\pm j_k=0\\ j_1,\cdots,j_k\neq 0}} \frac{1}{\left(\sqrt{2}\right)^k} F_{j_1,\cdots,j_k} w_{j_1}\cdots w_{j_k}. \end{aligned} \tag{2.15}$$

Notice that $F_{j_1,\cdots,j_k} = F_{|j_1|,\cdots,|j_k|}$.

Finally, we give a lemma showing the relation between the space $\mathcal{A}_\rho$ and the space $\mathcal{L}_{\rho,b}$.

Lemma 2.1. *Let u, v be complex valued function analytic on a complex neighborhood of the x-interval $[0,\pi]$, and let $(w_j)_{j\in\mathcal{Z}}$ be the sequence of its coordinates defined by (2.3), (2.12) and (2.13). Then for all $\mu < \rho$, we have*

$$\text{if} \quad u, v \in \mathcal{A}_\rho \quad \text{then} \quad w \in \mathcal{L}_{\mu,b} \quad \text{and} \quad \|w\|_\mu \leq c_{\mu,\rho}(|u|_\rho + |v|_\rho), \tag{2.16}$$

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

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

Proof. Due to (2.3), it is clear to know that

$$\begin{aligned} u(x) &= \sqrt{\frac{2}{\pi}} \sum_{j\geq 1} \frac{q_j}{\sqrt{\lambda_j}} \sin jx = \sqrt{\frac{2}{\pi}} \sum_{j\geq 1} \frac{q_j}{\sqrt{\lambda_j}} \frac{e^{\mathrm{i}jx} - e^{-\mathrm{i}jx}}{2\mathrm{i}} \\ &= \frac{1}{2\mathrm{i}} \sqrt{\frac{2}{\pi}} \sum_{j\neq 0} \frac{\mathrm{sgn}j \cdot q_{|j|}}{\sqrt{\lambda_{|j|}}} e^{\mathrm{i}jx}, \end{aligned}$$

$$v(x) = \sqrt{\frac{2}{\pi}}\sum_{j\geq 1}\sqrt{\lambda_j}p_j \sin jx = \sqrt{\frac{2}{\pi}}\sum_{j\geq 1}\sqrt{\lambda_j}p_j\frac{e^{\mathrm{i}jx}-e^{-\mathrm{i}jx}}{2\mathrm{i}}$$
$$= \frac{1}{2\mathrm{i}}\sqrt{\frac{2}{\pi}}\sum_{j\neq 0}\sqrt{\lambda_{|j|}}\mathrm{sgn}j\cdot p_{|j|}e^{\mathrm{i}jx}.$$

Then combining the transformation (2.12), (2.13) with the fact that $\lambda_{|j|}\sim|j|$ for $|j|$ large enough, we can prove this lemma by a small change of the proof of Lemma 2.1 in [12]. □

3. Space of polynomial and some properties

In the beginning of this section, we introduce some terminology about the polynomial on $\mathbb{C}^{\mathcal{Z}}$. Let $\ell\geq 2$ and $\mathbf{j}=(j_1,j_2,...,j_\ell)\in\mathcal{Z}^\ell$, we define

- the monomial associated with $\mathbf{j}$

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

- the divisor associated with $\mathbf{j}$

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

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

Besides, we also define the set of indices with zero momentum by

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

On the other hand, we say that $\mathbf{j}=(j_1,j_2,...,j_\ell)\in\mathcal{Z}^\ell$ is resonant, and we write $\mathbf{j}\in\mathcal{N}_\ell$, 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. In particular, if $\mathbf{j}$ is resonant then its associated divisor vanishes, i.e., $\Omega(\mathbf{j})=0$, and its associated monomials depends only on the actions

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

where for all $j\geq 1$, $I_j=w_jw_{-j}$ denotes the action associated with the index j. Finally, we note that if w is real, then $I_j=|w_j|^2$ and we remark that for odd ℓ the resonant set $\mathcal{N}_\ell$ is the empty set.

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

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

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

with the relation $a_{\boldsymbol{j}} = a_{|\boldsymbol{j}|}, |\boldsymbol{j}| = (|j_1|, \cdots, |j_\ell|)$,
- *The coefficients $a_{\boldsymbol{j}}$ are bounded, i.e. $\sup_{\mathbf{j} \in \mathcal{I}_\ell} |a_{\boldsymbol{j}}| < +\infty$ for all $\ell = 2, ..., k$.*

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

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

The nonlinearity f in (1.1) is assumed to be complex analytic in a neighbourhood of zero in $\mathbb{C}$. So there exist positive constants M and R_0 such that the Taylor expansion of its primary function

$$g(u) = \sum_{k=4}^{+\infty} \frac{g^{(k)}(0)}{k!} u^k$$

is uniformly convergent and bounded by M on the disc $|u| \leq R_0$ of $\mathbb{C}$. Hence formula (2.8) defines an analytic function on the ball $\|w\|_\rho \leq R_0$ of $\mathcal{L}_{\rho,b}$ and we have

$$F(w) = \sum_{k \geq 4} P_k,$$

where $P_k \in \mathcal{P}_k$ is homogeneous polynomial of degree k. Due to (1.2), we have $P_{2l+1} = 0, \quad l = 2, 3, \ldots$. Using Cauchy integral formula and by (2.9), (2.15), we obtain

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

At last, in the polynomial space we will give some useful estimates in which the zero momentum plays an important role.

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

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

and

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

Proof. Set

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

we have

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

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

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

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

Therefore,

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

But if $j_1 \pm j_2 \pm \cdots \pm j_{k-1} = \pm\ell$, then

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

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

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

which yields (3.8). □

Proposition 3.2. *For F a homogeneous polynomial of degree k in $\mathcal{P}_k$ and G a homogeneous polynomial of degree ℓ in $\mathcal{P}_\ell$, then $\{F, G\} \in \mathcal{P}_{k+\ell-2}$ and we have the estimate*

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

Remark 3.3. *The zero momentum (3.2) is a little different to the one in [12]. So we get the conclusion that have a small change. This will influence the last result.*

Proof. Now we assume that F and G are homogeneous polynomial 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

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

and such that if for instance $j = k_1$ and $-j = \ell_1$, we necessarily have $(k_2, ..., k_k, \ell_2, ..., \ell_\ell) = \mathbf{j}$. 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 value 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. Nonresonance condition

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

Recall that $\Omega(\mathbf{j}) = \mathrm{sgn} j_1 \cdot \lambda_{|j_1|} + \mathrm{sgn} j_2 \cdot \lambda_{|j_2|} + \cdots + \mathrm{sgn} j_r \cdot \lambda_{|j_r|}$, we define a set

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

and let

$$k_\ell = \begin{cases} 0 & \text{if } S_\ell = \varnothing, \\ \sum_{s \in S_\ell} \mathrm{sgn} j_s & \text{if } S_\ell \neq \varnothing, \end{cases}$$

and $k = (k_\ell)_{\ell \in \mathbb{N}}$. Then $\Omega(\mathbf{j}) = \sum_{\ell \geq 1} k_\ell \lambda_\ell$ and $|k| \leq r$. In the following we set $k = (\tilde{k}, \hat{k})$, where $\tilde{k} = (k_1, \ldots, k_N) \in \mathbb{Z}^N$, $\hat{k} = (k_{N+1}, \ldots) \in \mathbb{Z}^{\mathbb{N}}$ and assume that $|\hat{k}| \leq 2$. 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 $\mu(\mathbf{j})$, then we have the following proposition.

Proposition 4.1. *For any $\gamma > 0$ small enough, there exist a set $\mathcal{J}_\gamma$ satisfying $Meas([m_0, \Delta] - \mathcal{J}_\gamma) \to 0$ as $\gamma \to 0$, and a positive number ν such that for any $m \in \mathcal{J}_\gamma$ one has for any $N \geq 1$*

$$|\Omega(\boldsymbol{j})| \geq \frac{\gamma}{N^{\nu r^6}} \tag{4.1}$$

for any $\boldsymbol{j} \in \mathcal{Z}^r$ and $\mu(\boldsymbol{j}) < N$.

Proof. Given $\gamma > 0$ small enough, and a positive number ν, define the resonant sets $\mathcal{R}_{\tilde{k}\hat{k}}$ by

$$\mathcal{R}_{\tilde{k}\hat{k}} = \{m \in [m_0, \Delta] : |\sum_{\ell \geq 1} k_\ell \lambda_\ell| < \frac{\gamma}{N^{\nu r^6}}\}. \tag{4.2}$$

By combining Lemma A.3 and Lemma A.4, we can get

$$|\mathcal{R}_{\tilde{k}\hat{k}}| \leq \begin{cases} C\frac{\gamma^{1/r}}{N^{\nu r^6}}, & |\hat{k}| = 0; \\ C\frac{\gamma^{1/r} j^{2r}}{N^{\nu r^6}}, & |\hat{k}| = 1; \\ C\frac{\gamma^{1/r} j^{2r} i^{2r}}{N^{\nu r^6}}, & |\hat{k}| = 2, \end{cases} \tag{4.3}$$

where $i, j \geq N+1$, $|\cdot|$ denotes the Lebesgue measure of set and C is a suitable constant.

Let

$$\mathcal{R} = \bigcup_{|\tilde{k}|+|\hat{k}|\neq 0, |\tilde{k}|+|\hat{k}|\leq r, |\hat{k}|\leq 2} \mathcal{R}_{\tilde{k}\hat{k}}. \tag{4.4}$$

Now we would like to estimate the measure of $\mathcal{R}$. We split

$$\mathcal{K} := \{\hat{k} \in \mathbb{Z}^{\mathbb{N}} : |\hat{k}| \leq 2\}$$

as the union of the following four disjoint sets

$$\begin{aligned} \mathcal{K}_0 &= \{\hat{k} = 0\}, \\ \mathcal{K}_1 &= \{\hat{k} = e_i\}, \\ \mathcal{K}_{2+} &= \{\hat{k} = e_i + e_j\}, \\ \mathcal{K}_{2-} &= \{\hat{k} = e_i - e_j, i \neq j\}, \end{aligned}$$

where

$$e_i = (0, \ldots, 0, \overbrace{1}^{i\text{-th}}, 0, \ldots)$$

and $i, j \geq N+1$.

Let $|\hat{k}| = 2$ and $\hat{k} = e_i + e_j \in \mathcal{K}_{2+}$ for some $i, j \geq N+1$. If

$$\min\{i, j\} \geq 2(r-2)N + 1,$$

then it is easy to see that

$$\left| \sum_{\ell=1}^{N} k_\ell \lambda_\ell + \lambda_i + \lambda_j \right| \geq 1,$$

which is not small. Namely, the resonant set $\mathcal{R}_{\tilde{k}\hat{k}}$ is empty. So it is sufficient to consider

$$\max\{i, j\} < 2(r-2)N + 1,$$

when the estimate (4.5) is given below. In fact, we obtain

$$\begin{aligned} Meas & \bigcup_{(\tilde{k},\hat{k}) \in \mathbb{Z}^N \times \mathbb{Z}^{\mathbb{N}} \cap \mathcal{K}_{2+}} \mathcal{R}_{\tilde{k}\hat{k}} \\ \leq & \sum_{(\tilde{k},\hat{k}) \in \mathbb{Z}^N \times \mathbb{Z}^{\mathbb{N}} \cap \mathcal{K}_{2+}} Meas \mathcal{R}_{\tilde{k}\hat{k}} \\ \leq & \sum_{(\tilde{k},\hat{k}) \in \mathbb{Z}^N \times \mathbb{Z}^{\mathbb{N}} \cap \mathcal{K}_{2+}} C \frac{\gamma^{1/r}}{N^{\nu r^6}} \\ \leq & c_1 \gamma, \end{aligned} \tag{4.5}$$

where $c_1 > 0$ is a constant.

Similarly we obtain

$$Meas \bigcup_{(\tilde{k},\hat{k}) \in \mathbb{Z}^N \times \mathbb{Z}^{\mathbb{N}} \cap \mathcal{K}_0} \mathcal{R}_{\tilde{k}\hat{k}} \leq c_2 \gamma \tag{4.6}$$

and

$$Meas \bigcup_{(\tilde{k},\hat{k}) \in \mathbb{Z}^N \times \mathbb{Z}^{\mathbb{N}} \cap \mathcal{K}_1} \mathcal{R}_{\tilde{k}\hat{k}} \leq c_3 \gamma, \tag{4.7}$$

where $c_2, c_3 > 0$ is a constant. Now let

$$(\tilde{k}, \hat{k}) \in \mathbb{Z}^N \times \mathbb{Z}^{\mathbb{N}} \cap \mathcal{K}_{2-},$$

and assume $i > j$ without loss of generality. From $\lambda_i = \sqrt{i^2+m}, \lambda_j = \sqrt{j^2+m}$, then we can obtain that there is a constant $C > 0$ such that

$$\left|\frac{\lambda_i - \lambda_j}{i-j} - 1\right| \leq \frac{C}{j}.$$

Hence,

$$\lambda_i - \lambda_j = i - j + r_{ij},$$

with

$$|r_{ij}| \leq \frac{Ca}{j}$$

and $a = i - j$. Then we have

$$\left|\sum_{\ell=1}^{N} k_\ell \lambda_\ell + \lambda_i - \lambda_j\right| \geq \left|\sum_{\ell=1}^{N} k_\ell \lambda_\ell + a\right| - |r_{ij}|.$$

Therefore,

$$\mathcal{R}_{\tilde{k}\hat{k}} \subset \mathcal{Q}_{\tilde{k}aj} := \left\{\left|\sum_{\ell=1}^{N} k_\ell \lambda_\ell + a\right| \leq \frac{\gamma}{N^{\nu r^6}} + \frac{Ca}{j}\right\}.$$

If $j > j_0$, we have

$$\mathcal{Q}_{\tilde{k}aj} \subset \mathcal{Q}_{\tilde{k}aj_0}.$$

Then it is sufficient to consider

$$a \leq 2(r-2)N + 1,$$

and let

$$j_0 = \gamma^{-1/2} N^{\nu r^3/2}.$$

By similar proof of Theorem 6.19 in [3], we obtain

$$Meas \bigcup_{(\tilde{k},\hat{k})\in\mathbb{Z}^N\times\mathbb{Z}^{\mathbb{N}}\cap\mathcal{K}_{2-}} \mathcal{R}_{\tilde{k}\hat{k}} \leq c_4\sqrt{\gamma}, \tag{4.8}$$

where $c_4 > 0$ is a constant.

In view of (4.4)–(4.8), we let $\mathcal{J}_\gamma = [m_0, \Delta] - \mathcal{R}$, then the proposition is proved. □

5. Recursive equation and normal form results

Now, we define the N-normal form. Fix $N \geq 1$ and $k \geq 4$. Recalling the definition of $\mu(\mathbf{j})$ in the Section 4, we set

$$\mathcal{J}_k(N) = \{\mathbf{j} \in \mathcal{I}_k \mid \mu(\mathbf{j}) > N\}.$$

Definition 5.1 (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*

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

In other words, W contains either monomials depending only on the actions or monomials whose indices $\boldsymbol{j}$ satisfies $\mu(\boldsymbol{j}) > N$, that is, monomials involving at least three modes with index greater than N.

5.1. Recursive equation

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

Lemma 5.1. *Assume that the nonresonance condition (4.1) is satisfied, and let N be fixed. Also assume that $H_0 := \sum_{j\geq 1} \lambda_j w_j w_{-j}$ is the integrable part of Hamiltonian (2.4) and Q is a homogenous polynomial of degree n. Then the homological equation*

$$\{\chi, H_0\} - W = Q \tag{5.1}$$

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

$$\|W\| \leq \|Q\| \quad \textit{and} \quad \|\chi\| \leq \frac{N^{\nu n^6}}{\gamma}\|Q\|. \tag{5.2}$$

Proof. Assume that $Q = \sum_{\mathbf{j}\in\mathcal{I}_n} Q_{\mathbf{j}} w_{\mathbf{j}}$ and find $W = \sum_{\mathbf{j}\in\mathcal{I}_n} W_{\mathbf{j}} w_{\mathbf{j}}$ and $\chi = \sum_{\mathbf{j}\in\mathcal{I}_n} \chi_{\mathbf{j}} w_{\mathbf{j}}$ such that (5.1) is satisfied. Equation (5.1) can be written in term of polynomial coefficients

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

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

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

In view of (4.1), this leads to (5.2). □

In the following part, 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 problem below.

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

$$(H_0 + F) \circ \Phi_\chi^1 = H_0 + W + R. \tag{5.3}$$

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

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

where $\mathrm{ad}_\chi K = \{\chi, K\}$. Also, if K and L are homogeneous polynomials of degree respectively n and ℓ then $\{K, L\}$ is a homogeneous polynomial of degree $n + \ell - 2$. Therefore, we obtain by using the Taylor formula

$$(H_0 + F) \circ \Phi_\chi^1 - (H_0 + F) = \sum_{k=0}^{r/2-2} \frac{1}{(k+1)!} \mathrm{ad}_\chi^k(\{\chi, H_0 + F\}) + \mathcal{O}_r, \tag{5.4}$$

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:

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

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

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

we get

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

where C_r is a differential operator satisfying

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

Applying B_r to the two sides of equation (5.4), 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

$$\{\chi_n, H_0\} - W_n = Q_n, \quad n = 4, \cdots, r, \tag{5.5}$$

where

$$\begin{aligned} Q_n = &- P_n + \sum_{k=4}^{n-2} \{P_{n+2-k}, \chi_k\} \\ &+ \sum_{k=1}^{n/2-2} \frac{B_k}{k!} \sum_{\substack{\ell_1+\cdots+\ell_{k+1}=n+2k \\ 4\le \ell_i \le n-2k}} \mathrm{ad}_{\chi_{\ell_1}} \cdots \mathrm{ad}_{\chi_{\ell_k}} (W_{\ell_{k+1}} - P_{\ell_{k+1}}). \end{aligned} \tag{5.6}$$

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

$$\begin{aligned} R = &\sum_{n\ge r+1} \sum_{k=2}^{n/2-1} \frac{1}{k!} \sum_{\substack{\ell_1+\cdots+\ell_k=n+2k-2 \\ 4\le \ell_i\le r}} \mathrm{ad}_{\chi_{\ell_1}} \cdots \mathrm{ad}_{\chi_{\ell_k}} H_0 \\ &+ \sum_{n\ge r+1} \sum_{k=0}^{n/2-2} \frac{1}{k!} \sum_{\substack{\ell_1+\cdots+\ell_{k+1}=n+2k \\ 4\le \ell_1,\cdots,\ell_k\le r \\ 4\le \ell_{k+1}}} \mathrm{ad}_{\chi_{\ell_1}} \cdots \mathrm{ad}_{\chi_{\ell_k}} P_{\ell_{k+1}}. \end{aligned} \tag{5.7}$$

Lemma 5.2. *Assume that the nonresonance condition (4.1) 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 homogeneous polynomials χ_n and W_n of degree*

n, with W_n in N-normal form, which are solutions of the recursive equation (5.5) and satisfy

$$\|\chi_n\| + \|W_n\| \le (C4^n n N^\nu)^{n^7}. \tag{5.8}$$

Proof. We define χ_n and W_n by induction using Lemma 5.1. Note that (5.8) is clearly satisfied for $n = 4$, provided C big enough. Estimate (5.2) yields

$$\gamma N^{-\nu n^6}\|\chi_n\| + \|W_n\| \le \|Q_n\|. \tag{5.9}$$

Using the definition (5.6) of the term Q_n and the estimate on the Bernoulli numbers, $|B_k| \le k! c^k$ for some $c > 0$, together with (3.9), which implies that for all $\ell \ge 4$, $\|\mathrm{ad}_{\chi_\ell} R\| \le 2^{\min\{n,\ell\}-1} n\ell \|R\|\,\|\chi_\ell\|$ for any polynomial R of degree less than n, we have for all $n \ge 4$

(5.10)

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

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 $C(n,k) \le 4^n$.

We set $\beta_n = n(\|\chi_n\| + \|W_n\|)$. Equation (5.9) implies that

$$\beta_n \le (CN^\nu)^{n^6} n \|Q_n\|, \tag{5.11}$$

for some constant C independent of n.

By the fact that $\|P_n\| \le M R_0^{-n}$ (see (3.6)), we obtain

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

where

$$\beta_n^{(1)} = (CN^\nu)^{n^6} 2^n n^3 \sum_{k=4}^{n-2} \beta_k \tag{5.12}$$

and

$$\beta_n^{(2)} = N^{\nu n^6}(Cn)^{n-1}4^n \sum_{k=1}^{n/2-2} \sum_{\substack{\ell_1+\cdots+\ell_{k+1}=n+2k \\ 4\le \ell_i \le n-2k}} \beta_{\ell_1}\cdots\beta_{\ell_k}(\beta_{\ell_{k+1}} + \|P_{\ell_{k+1}}\|), \tag{5.13}$$

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

$$(C4^n nN^\nu)^{n^7} \ge 1 \quad \text{for all } \ n \ge 4, \tag{5.14}$$

so we get

$$\beta_n^{(1)} \le (CN^\nu)^{n^6} 2^n n^4 (C4^n nN^\nu)^{(n-1)^7} \le \frac{1}{2}(C4^n nN^\nu)^{n^7}$$

for $n \ge 4$ and provided $C > 2$.

Using (5.14) again and the induction hypothesis, we obtain

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

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 $(n-2k)^7 + 4^7 k$. 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

$$\begin{aligned} \beta_n^{(2)} &\le \max_{k=\{1,\cdots,n/2-2\}} N^{\nu n^6}(Cn)^{n-1} Cn^{k+2} 4^n (CN^\nu 4^n (n-2k))^{(n-2k)^7+4^7k} \\ &\le \frac{1}{2}(C4^n nN^\nu)^{n^7} \end{aligned}$$

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

5.2. 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 5.3. *Assume that F is analytic on a ball $B_\rho(R_1)$ for some $R_1 > 0$ and $\rho > 0$. Assume that the nonresonance condition (4.1) is satisfied, and let $\beta < 1/7$ and $M > 1$ be fixed. Then there exist constants $\epsilon_0 > 0$ and $\sigma > 0$ 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*

$$(H_0 + F) \circ \Phi_\chi^1 = H_0 + W + R. \tag{5.15}$$

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

$$\|X_W(w)\|_\rho + \|X_\chi(w)\|_\rho \le 2\epsilon^{3/2} \quad \text{and} \quad \|X_R(w)\|_\rho \le \epsilon^2 e^{-\frac{1}{4}|\ln\epsilon|^{1+\beta}}. \tag{5.16}$$

Remark 5.4. *In this theorem, we let $\beta < 1/7$ (see (5.17)) rather than $\beta < 1$, because the nonresonance condition (4.1) is a little different to the one in [12]. Then we get the stability of the solutions for times that is shorter than [12].*

Proof. Using Lemma 5.2, 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 (5.15) holds with $R = \mathcal{O}_r$. Now for fixed $\epsilon > 0$, we choose

$$N = |\ln\epsilon|^{1+\beta} \quad \text{and} \quad r = |\ln\epsilon|^{\beta}.$$

This choice is motivated by the necessity of balance between W and R in (5.15): The error induced by W is controlled as in Lemma 6.2, while the error induced by R is controlled by Lemma 5.2. By (5.8), we have

(5.17)

$$\begin{aligned}
\|\chi_k\| \le& (C4^k k N^\nu)^{k^7} \le \exp\Big(k\big(\nu k^6(1+\beta)\ln|\ln\epsilon| + k^7\ln 4 + k^6 \ln Ck\big)\Big)\\
\le& \exp\Big(k\big(\nu r^6(1+\beta)\ln|\ln\epsilon| + r^7\ln 4 + r^6 \ln Cr\big)\Big)\\
\le& \exp\Big(k|\ln\epsilon|\big(\nu|\ln\epsilon|^{6\beta-1}(1+\beta)\ln|\ln\epsilon| + |\ln\epsilon|^{7\beta-1}\ln 4\\
&+ |\ln\epsilon|^{6\beta-1}\ln C|\ln\epsilon|^{\beta}\big)\Big)\\
\le& \epsilon^{-k/8}
\end{aligned}$$

as $\beta < 1/7$, and for $\epsilon < \epsilon_0$ sufficiently small. Therefore using Proposition 3.1, we obtain for $w \in B_\rho(M\epsilon)$

$$|\chi_k(w)| \le \epsilon^{-k/8}(M\epsilon)^k \le M^k\epsilon^{7k/8}$$

and thus

$$|\chi(w)| \leq \sum_{k\geq 4} M^k \epsilon^{7k/8} \leq \epsilon^{3/2}$$

for ϵ small enough. Similarly, we have for all $k \leq r$,

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

and

$$\|X_\chi(w)\|_\rho \leq \sum_{k\geq 4} k(2M)^{k-1} \epsilon^{7k/8-1} \leq C\epsilon^{-1}\epsilon^{28/8} \leq \epsilon^{3/2}$$

for ϵ small enough. Similar bounds clearly hold for $W = \sum_{k=4}^r W_k$, which shows the first estimate in (5.16).

On the other hand, using $\mathrm{ad}_{\chi_{\ell_k}} H_0 = W_{\ell_k} + Q_{\ell_k}$ (see (5.5)) and then combining Lemma 5.2 with the definition of Q_n, we get

$$\|\mathrm{ad}_{\chi_{\ell_k}} H_0\| \leq (C 4^{\ell_k} \ell_k N^\nu)^{\ell_k^7} \leq \epsilon^{-\ell_k/8},$$

where the last inequality proceeds as in (5.17). Thus, due to (5.7), (5.17) and $\|P_{\ell_{k+1}}\| \leq M R_0^{-\ell_{k+1}}$, we obtain by Proposition 3.1 that for $w \in B_\rho(M\epsilon)$

$$\begin{aligned}\|X_R(w)\|_\rho &\leq \sum_{n\geq r+1} \sum_{k=0}^{n/2-2} 4^n n (Cr)^{3n} \epsilon^{-\frac{n+2k}{8}} \epsilon^{n-1} \leq \sum_{n\geq r+1} n^2 (4Cr)^{3n} \epsilon^{n/2} \\ &\leq (4Cr)^{3r} \epsilon^{r/2}.\end{aligned}$$

Since $r = |\ln \epsilon|^\beta > 2$ (ϵ small enough), we get $\|X_R(w)\|_\rho \leq \epsilon^2 e^{-\frac{1}{4}|\ln \epsilon|^{1+\beta}}$ for $w \in B_\rho(M\epsilon)$. □

6. Proof of the main result

Before giving the proof, we will introduce two important lemmata.

Lemma 6.1. *Let $f : \mathbb{R} \to \mathbb{R}_+$ a continuous function and $y : \mathbb{R} \to \mathbb{R}_+$ a differentiable function satisfying the inequality*

$$\frac{d}{dt} y(t) \leq 2 f(t) \sqrt{y(t)}, \quad \forall\, t \in \mathbb{R}.$$

Then we have the estimate

$$\sqrt{y(t)} \leq \sqrt{y(0)} + \int_0^t f(s)ds, \quad \forall t \in \mathbb{R}.$$

Proof. The proof can be found in [12]. □

Fix $N > 1$, for any $w \in \mathcal{L}_{\rho,b}$, we define

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

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

$$R_\rho^N(w) \leq e^{-\mu N}\|w\|_{\rho+\mu}. \tag{6.1}$$

Lemma 6.2. *Let $N \in \mathbb{N}$ and $k \geq 4$. 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*

$$R_\rho^N(w(t)) \leq R_\rho^N(w(0)) + 4k^3 2^{k-1}\|W\| \int_0^t R_\rho^N(w(s))^2 \|w(s)\|_\rho^{k-3} ds \tag{6.2}$$

and

$$\|w(t)\|_\rho \leq \|w(0)\|_\rho + 4k^3 2^{k-1}\|W\| \int_0^t R_\rho^N(w(s))^2 \|w(s)\|_\rho^{k-3} ds. \tag{6.3}$$

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 have

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

Then using the Lemma 6.1, we get

$$\begin{aligned} &e^{\rho|j|}\sqrt{I_j(t)} \\ &\quad\leq e^{\rho|j|}\sqrt{I_j(0)} + 2^{k-1}k\|W\| \int_0^t \Big(\sum_{\substack{j_1 \pm \cdots \pm j_{k-1} = \pm j \\ 2 \text{ indices} > N}} e^{\rho|j_1|}|w_{j_1}|\cdots e^{\rho|j_{k-1}|}|w_{j_{k-1}}|\Big)ds. \end{aligned} \tag{6.4}$$

Ordering the multi-indices such way $|j_1|$ 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 obtain, after summation in $|j| > N$,

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

Inequality (6.3) can be proved in the same way. □

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

Proof of the main theorem 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, by Lemma 2.1, to in $\mathcal{L}_{\frac{3}{2}\rho,b}$ with $\|w(0)\|_{\frac{3}{2}\rho} \le \frac{c_\rho}{4}\epsilon$. 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 5.3 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 $\|y(0)\|_{\mathcal{L}_{\frac{3}{2}\rho}} \le \frac{c_\rho}{2}\epsilon$. In particular, notice the facts that

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

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

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

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

So using (6.2) for the first vector field and (5.16) for the second one, we get for $|t| \le T_\epsilon$,

$$
\begin{aligned}
R_\rho^N(y(t)) &\le \frac{1}{2}c_\rho\epsilon e^{-\sigma N} + 4|t|\sum_{k=4}^r \|W_k\| k^3 (2c_\rho\epsilon)^{k-1} e^{-2\sigma N} + |t|\epsilon e^{-\frac{1}{4}|ln\epsilon|^{1+\beta}} \\
&\le \Big(\frac{1}{2} + 4|t|\sum_{k=4}^r \|W_k\| k^3 (2c_\rho\epsilon)^{k-2} e^{-\sigma N} + |t|\epsilon e^{-\frac{1}{8}|ln\epsilon|^{1+\beta}}\Big) c_\rho \epsilon e^{-\sigma N}, \qquad (6.5)
\end{aligned}
$$

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

Using Lemma 5.2, we then verify

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

and thus, for ϵ small enough,

$$R_\rho^N(y(t)) \le c_\rho \epsilon e^{-\sigma N} \text{ for all } |t| \le \min\{T_\epsilon, e^{\sigma N}\}. \tag{6.6}$$

Similarly, we obtain

$$\|y(t)\|_\rho \le c_\rho \epsilon \text{ for all } |t| \le \min\{T_\epsilon, e^{\sigma N}\}. \tag{6.7}$$

In view of the definition of T_ϵ, inequalities (6.6) and (6.7) imply $T_\epsilon \ge e^{\sigma N}$. In particular $\|w(t)\|_\rho \le 2c_\rho\epsilon$ for $|t| \le e^{\sigma N} = \epsilon^{-\sigma|\ln\epsilon|^\beta}$ and using (2.17), we finally obtain (1.3).

Appendix A

In this section, we will give some technical lemmas.

Lemma A.1. *For any $K \le r$, consider K indexes $j_1 < \ldots < j_K \le N$; consider the determinant*

$$D := \begin{vmatrix} \lambda_{j_1} & \lambda_{j_2} & \cdots & \lambda_{j_K} \\ \frac{d\lambda_{j_1}}{dm} & \frac{d\lambda_{j_2}}{dm} & \cdots & \frac{d\lambda_{j_K}}{dm} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{d^{K-1}\lambda_{j_1}}{dm^{K-1}} & \frac{d^{K-1}\lambda_{j_2}}{dm^{K-1}} & \cdots & \frac{d^{K-1}\lambda_{j_K}}{dm^{K-1}} \end{vmatrix}. \tag{A.1}$$

One has

$$D = \Big(\prod_{l=1}^{K} \lambda_{j_l}^{-2K+1}\Big)\Big(\prod_{1\le l<k\le K} (j_l)^2 - (j_k)^2\Big) \ge \frac{C}{N^{2K^2}}. \tag{A.2}$$

Proof. By explicit computation, one has

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

Substituting (A.3) in the right hand site of (A.1) we get the determinant to be estimated. To obtain the estimate factorize from the j-th column the term

$\lambda_j = (j^2+m)^{\frac{1}{2}}$, and from the n-th row the 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

$$\left[\prod_{l=1}^{K}\lambda_{j_l}\right]\left[\frac{1}{2}\prod_{n=2}^{K-1}\frac{(2n-3)!}{2^{n-2}(n-1)!2^n}\right]\begin{vmatrix} 1 & 1 & \cdots & 1 \\ x_{j_1} & x_{j_2} & \cdots & x_{j_K} \\ \vdots & \vdots & \ddots & \vdots \\ x_{j_1}^{K-1} & x_{j_2}^{K-1} & \cdots & x_{j_K}^{K-1} \end{vmatrix}, \tag{A.4}$$

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

$$\prod_{1\le l<n\le K}(x_{j_l}-x_{j_n}). \tag{A.5}$$

Now we have

$$|x_{j_l}-x_{j_n}| = \left|\frac{1}{j_l^2+m}-\frac{1}{j_n^2+m}\right| = \frac{|j_n^2-j_l^2|}{(j_l^2+m)(j_n^2+m)} \ge C x_{j_l}x_{j_n},$$

with a suitable C. So (A.5) is estimated by

$$\prod_{l=1}^{K-1}\prod_{n=l+1}^{K} C x_{j_l}x_{j_n} = C^{\sum_{n=2}^{K}(n-1)}\prod_{l=1}^{K-1}\left(x_{j_l}^{K-l}\prod_{n=l+1}^{K}x_{j_n}\right) = C\prod_{l=1}^{K}x_{j_l}^{K-1},$$

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

Next we need the lemma from appendix B of [8], namely

Lemma A.2. *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 $i\in[1,\ldots,K]$, such that*

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

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 proposition of appendix B in [8]. □

Combining Lemma A.1 and Lemma A.2, we deduce the following lemma.

Lemma A.3. *Let $w \in \mathbb{Z}^\infty$ be a vector with K component different from zero, namely those with index $j_1, \cdots, j_K$; assume that $K \leq r$, and assume that $j_1 < \cdots < j_K \leq N$. Then for any $m \in [m_0, \Delta]$, there exists an index $j \in [0, \cdots, K-1]$ such that*

$$\left| w \cdot \frac{d^j \lambda}{dm^j}(m) \right| \geq C \frac{\|w\|_{l^1}}{N^{2K^2+2}}, \tag{A.6}$$

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

From [18] we learn the following lemma.

Lemma A.4. *Suppose that $g(m)$ is r times differentiable on an interval $J \subset \mathbb{R}$. Let $J_\gamma := \{m \in J : |g(m)| < \gamma\}$, $\gamma > 0$. If $|g^{(r)}(m)| \geq d > 0$ on J, then $|J_\gamma| \leq M\gamma^{1/r}$, where $M := 2(2 + 3 + \cdots + r + d^{-1})$.*

Proof. The proof can be found in Lemma 2.1 of [18]. □

Acknowledgement

The authors would like to thank the referees for their valuable comments and suggestions which have helped to improve the quality of this paper. The authors are very grateful to Professor Hongzi Cong for his invaluable discussions and suggestions.

References

[1] D. Bambusi. On long time stability in Hamiltonian perturbations of non-resonant linear PDEs. Nonlinearity, **12**, 823–850 (1999). MR1709838

[2] D. Bambusi. Nekhoroshev theorem for small amplitude solutions in nonlinear Schrödinger equations. Math. Z., **230**, 345–387 (1999). MR1676714

[3] D. Bambusi. Birkhoff normal form for some nonlinear PDEs. Comm. Math. Phys., **234**, 253–285 (2003). MR1962462

[4] D. Bambusi. A Birkhoff normal form theorem for some semilinear PDEs. Hamiltonian dynamical systems and applications, 213–247 (2008). MR2446257

[5] D. Bambusi, M. Berti and E. Magistrelli. Degenerate KAM theory for partial differential equations. J. Differential Equations, **250**, 3379–3397 (2011). MR2772395

[6] D. Bambusi and B. Grébert. Birkhoff normal form for partial differential equations with tame modulus. Duke Math. J., **135**, 507–567 (2006). MR2272975

[7] D. Bambusi and N. N. Nekhoroshev. A property of exponential stability in nonlinear wave equations near the fundamental linear mode. Phys. D, **122**, 73–104 (1998). MR1650123

[8] 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**, 1–25 (1985). MR0830795

[9] J. Bourgain. Construction of approximative and almost periodic solutions of perturbed linear Schrödinger and wave equations. Gemo. Funct. Anal., **6**, 201–230 (1996). MR1384610

[10] H. Cong, J. Liu and X. Yuan. Stability of KAM tori for nonlinear Schrödinger equation. Mem. Amer. Math. Soc., No.**1134**, 239 (2016). MR3430765

[11] H. Cong, M. Gao and J. Liu. Long time stability of KAM tori for nonlinear wave equation. J. Differential Equations, **258**, 2823–2846 (2015). MR3312647

[12] E. Faou and B. Grébert. A Nekhoroshev-type theorem for the nonlinear Schrödinger equation on the torus. Analysis and PDE, **6**, 1243–1262 (2013). MR3148054

[13] S. B. Kuksin. Hamiltonian perturbations of infinite-dimensional linear systems with an imaginary spectrum. Funct. Anal. Appl., **21**, 192–205 (1987). MR0911772

[14] S. B. Kuksin. Perturbations of quasiperiodic solutions of infinite-dimensional Hamiltonian systems. Math. USSR Izv., **32**, 39–62 (1989). MR0936522

[15] S. B. Kuksin. Nearly integrable infinite-dimensional Hamiltonian systems. Springer-Verlag, Berlin (1993). MR1290785

[16] J. Pöschel. Quasi-periodic solutions for nonlinear wave equation. Comm. Math. Helv., **71**, 269–296 (1996). MR1396676

[17] C. E. Wayne. Periodic and quasi-periodic solutions of nonlinear wave equations via KAM theory. Comm. Math. Phys., **127**, 479–528 (1990). MR1040892

[18] J. Xu, J. You and Q. Qiu. Invariant tori for nearly integrable Hamtonian systems with degenercy. Math. Z., **226**, 375–387 (1997). MR1483538

[19] X. Yuan. Quasi-periodic solutions of nonlinear wave equations with a prescribed potential. Discrete Contin. Dyn. Syst., **16**, 615–634 (2006). MR2257152

[20] X. Yuan. Quasi-periodic solutions of completely resonant nonlinear wave equations. J. Differential Equations, **230**, 213–274 (2006). MR2270553

Lufang Mi
College of Science
Binzhou University
Shandong
P R China
E-mail: milufang@126.com

Chunyong Liu
Shandong Chemical Engineering and Vocational College
Weifang, Shandong
P R China
E-mail: 18340811460@163.com

Guanghua Shi
College of Mathematics and Statistics
Hunan Normal University
Changsha, Hunan
P R China
E-mail: shiguanghua12@fudan.edu.cn

Rong Zhao
Dalian Airforce Communication Non-Commissioned Officer Academy
Liaoning
P R China
E-mail: pd81520@sina.com