

# LONG TIME BEHAVIOR OF THE SOLUTIONS OF NLW ON THE $d$ -DIMENSIONAL TORUS

JOACKIM BERNIER<sup>1</sup>, ERWAN FAOU<sup>2</sup> and BENOÎT GRÉBERT<sup>3</sup>

<sup>1</sup> Institut de Mathématiques de Toulouse, UMR5219, Université de Toulouse, CNRS,  
Université Paul Sabatier, F-31062 Toulouse Cedex 9, France;

email: joackim.bernier@math.univ-toulouse.fr

<sup>2</sup> Univ Rennes, INRIA, CNRS, IRMAR - UMR 6625, F-35000 Rennes, France;

email: Erwan.Faou@inria.fr

<sup>3</sup> Laboratoire de Mathématiques Jean Leray, Université de Nantes, UMR CNRS 6629, 2,  
rue de la Houssinière, 44322 Nantes Cedex 03, France;

email: benoit.grebert@univ-nantes.fr

Received 26 September 2019; accepted 22 January 2020

## Abstract

We consider the nonlinear wave equation (NLW) on the  $d$ -dimensional torus  $\mathbb{T}^d$  with a smooth nonlinearity of order at least 2 at the origin. We prove that, for almost any mass, small and smooth solutions of high Sobolev indices are stable up to arbitrary long times with respect to the size of the initial data. To prove this result, we use a normal form transformation decomposing the dynamics into low and high frequencies with weak interactions. While the low part of the dynamics can be put under classical Birkhoff normal form, the high modes evolve according to a time-dependent linear Hamiltonian system. We then control the global dynamics by using polynomial growth estimates for high modes and the preservation of Sobolev norms for the low modes. Our general strategy applies to any semilinear Hamiltonian Partial Differential Equations (PDEs) whose linear frequencies satisfy a very general nonresonance condition. The (NLW) equation on  $\mathbb{T}^d$  is a good example since the standard Birkhoff normal form applies only when  $d = 1$  while our strategy applies in any dimension.

2010 Mathematics Subject Classification: 35L05, 37K55 (primary); 35B40, 35Q55 (secondary)

## 1. Introduction

Let us consider the nonlinear wave equation (NLW) set on the  $d$ -dimensional torus  $\mathbb{T}^d = (\mathbb{R}/2\pi\mathbb{Z})^d$  with  $d \geq 1$ ,

$$u_{tt} - \Delta u + mu + f(u) = 0, \quad x \in \mathbb{T}^d, \quad (\text{NLW})$$

© The Author(s) 2020. This is an Open Access article, distributed under the terms of the Creative Commons Attribution licence (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted re-use, distribution, and reproduction in any medium, provided the original work is properly cited.

satisfied by a real valued function  $u(t, x)$  with given initial data  $u(0) \equiv u(0, \cdot)$  and  $\dot{u}(0) \equiv \dot{u}(0, \cdot) = \partial_t u(0, \cdot)$ . The function  $f \in C^\infty(\mathbb{R}, \mathbb{R})$  is at least of order 2 at the origin, that is,  $f(0) = f'(0) = 0$ . For small and smooth initial data  $u(0) \in H^s(\mathbb{T}^d)$  and  $\dot{u}(0) \in H^{s-1}(\mathbb{T}^d)$  with large  $s$ , we are interested in a description of the long time behavior of  $u(t) \equiv u(t, \cdot)$  solution of (NLW).

In dimension  $d = 1$ , it is known that if  $\varepsilon$  measures the size of the initial data, the solution is controlled for arbitrary polynomial times with respect to  $\varepsilon$  and for almost all  $m$  away from zero. More precisely, the following result has been proved (see, for instance, [BG06] and also [Bam03]).

**THEOREM 1 [BG06].** *Let  $d = 1$  and  $r \geq 2$ . For almost any  $m > 0$ , there exists  $s_*(r)$  such that for  $s > s_*(r)$  there exists  $\varepsilon_0(r, s) > 0$  such that for all  $\varepsilon < \varepsilon_0$*

$$\|(u(0), \dot{u}(0))\|_{H^s \times H^{s-1}} \leq \varepsilon \implies \|(u(t), \dot{u}(t))\|_{H^s \times H^{s-1}} \leq 2\varepsilon, \quad t \leq \varepsilon^{-r}. \quad (1)$$

The crucial tool to obtain this result is to show that for a large set of parameters  $m$ , the frequencies  $\omega_j = \sqrt{|j|^2 + m}$  of the linear wave operator satisfy a nonresonance condition of the following form.

Fix  $r \geq 3$ , there exists  $\gamma > 0$  such that for  $\mathbf{k} = (k_1, \dots, k_p) \in (\mathbb{Z}^d)^p$ ,  $\boldsymbol{\ell} = (\ell_1, \dots, \ell_q) \in (\mathbb{Z}^d)^q$  with  $p + q \leq r$ , we have

$$|\omega_{k_1} + \dots + \omega_{k_p} - \omega_{\ell_1} - \dots - \omega_{\ell_q}| \geq \frac{\gamma}{\mu_3(\mathbf{k}, \boldsymbol{\ell})^\alpha}, \quad (\mu_3)$$

unless  $(|k_1|, \dots, |k_p|)$  and  $(|\ell_1|, \dots, |\ell_q|)$  are equal up to a permutation, and where  $\mu_3(\mathbf{k}, \boldsymbol{\ell})$  denotes the third largest number among the collection  $(|k_i|, |\ell_j|)_{i,j}$  and  $\alpha$  depends on  $r$ .

This condition, introduced in [BG06], allows to eliminate (or normalize) all the terms in the Hamiltonian of the perturbation (depending on  $F$ , a primitive of  $f$  in (NLW)) involving at most two high Fourier modes via a Birkhoff normal form procedure. On the other hand, it is known since [Bam03] that we can neglect all the monomials involving more than three high modes (see, for instance, [Gre07] or [Bam07] for a simple presentation of these two facts). So once we have  $(\mu_3)$ , we can expect a control of the Sobolev norms similar to (1).

Note that  $(\mu_3)$  is close from the so-called *second-order* Melnikov nonresonance condition that says in a formulation allowing comparison with  $(\mu_3)$ .

Fix  $n \geq 2$ , there exists  $\gamma > 0$  such that for  $\mathbf{k} = (k_1, \dots, k_p) \in (\mathbb{Z}^d)^p$ ,  $\boldsymbol{\ell} = (\ell_1, \dots, \ell_q) \in (\mathbb{Z}^d)^q$  with  $|k_i|, |\ell_i| \leq n$  and for  $(j_1, j_2) \in (\mathbb{Z}^d)^2$  with  $|j_1|, |j_2| > n$ ,

$$|\omega_{k_1} + \dots + \omega_{k_p} - \omega_{\ell_1} - \dots - \omega_{\ell_q} + \omega_{j_1} - \omega_{j_2}| \geq \frac{\gamma}{r^\alpha},$$

unless  $(|k_1|, \dots, |k_p|)$  and  $(|\ell_1|, \dots, |\ell_q|)$  are equal up to a permutation, where  $r = p + q$ . This terminology refers to the original papers [Mel65, Mel68]

where similar conditions were introduced for proving the stability of low-dimensional invariant tori in Hamiltonian dynamics and popularized in the KAM literature [Mos67, Eli88, Bou97, XY01] and later while extending these results to Hamiltonian PDEs [Kuk93, Pös96].

We note that in the Melnikov case, the ‘length’ of the resonance  $(r + 2)$  is free, but the number of ‘interior’ modes (here  $(2n + 1)^d$ ) is fixed while it is exactly the converse in  $(\mu_3)$ . So the two conditions are not equivalent but similar.

The condition  $(\mu_3)$  applies to many situations, including the one-dimensional wave equations, the one-dimensional nonlinear Schrödinger equation with an external potential [Bam03, Bam07, BG06], the  $d$ -dimensional nonlinear Schrödinger equation with a convolution potential [BG03, BG06], plane waves stability for nonlinear Schrödinger equation [FGL13], wave equations on Zoll manifolds [BDGS07] or quantum harmonic oscillator on  $\mathbb{R}^d$  [GIP09].

The main difficulty of the higher dimensional case for (NLW) is that the frequencies do not satisfy the second-order Melnikov condition for a large set of parameters  $m$ , as already noted for instance in [Bou95, Del09]. In fact, in dimension  $d \geq 2$ ,  $(\mu_3)$  is not satisfied either: let, for instance,  $|k_1| \geq |\ell_1| \geq \mu_3(\mathbf{k}, \ell)$ . Then in the left-hand side of  $(\mu_3)$ , the quantity  $\omega_{k_1} - \omega_{\ell_1}$  depends on indices that are not controlled by  $\mu_3(\mathbf{k}, \ell)$ . For  $d = 1$ , if  $k_1$  and  $\ell_1$  are large, the difference  $\omega_{k_1} - \omega_{\ell_1}$  is close to be an integer and the parameter  $m$  can be chosen in a large set so that  $(\mu_3)$  holds. However, in dimension  $d \geq 2$ ,  $\omega_{k_1} - \omega_{\ell_1}$  describes a dense set at the scale  $|k_1|$  which prevents  $(\mu_3)$  to hold. Actually, we can only prove that for a large set of parameters  $m$ , the following condition (which is related to the so-called *first-order* Melnikov condition) holds:

$$|\omega_{k_1} + \dots + \omega_{k_p} - \omega_{\ell_1} - \dots - \omega_{\ell_q}| \geq \frac{\gamma}{\mu_2(\mathbf{k}, \ell)^\alpha} \quad (\mu_2)$$

unless  $(|k_1|, \dots, |k_p|)$  and  $(|\ell_1|, \dots, |\ell_q|)$  are equal up to a permutation where  $\mu_2(\mathbf{k}, \ell)$  is the second largest index in the multi-index  $(\mathbf{k}, \ell)$ .

Despite this known problem, some results have been proved concerning the existence of quasiperiodic solutions, where the loss of derivative is controlled by the use of KAM–Newton schemes; see [Bou95]. Concerning the control of large open sets of solutions, some results can be found in [DS06, Del09], but the time control depends on the shape on the nonlinearity inducing restriction on the index  $r$  (essentially driven by the annulation degree of the nonlinearity in 0).

Another situation where  $(\mu_2)$  appears in a natural way is given by numerical discretization of Hamiltonian PDEs. For example, standard splitting methods applied to wave equations in dimension  $d = 1$  induces numerical resonances destroying the property  $(\mu_3)$  and degenerating to  $(\mu_2)$  even for generic time discretization parameters. (The fundamental reason is that time discrete numerical

schemes require the control of small divisors of the form  $e^{ih\Omega(k,\ell)} - 1$  instead of  $\Omega(k, \ell)$  as defined in  $(\mu_3)$ , where  $h$  is the time discretization parameter. Hence, numerical resonances can occur when, for instance,  $h(\omega_{k_1} - \omega_{\ell_1})$  is close to an arbitrary large multiple of  $2\pi$  (see, for instance, [Sha00, HLW06] for a finite-dimensional analysis of symplectic integrators.) In this case, it is however possible to control the solution by playing with the time integrator or with the space discretization; see [CHL08, FGP10a, FGP10b, Fao12, FGL14].

In this paper, we propose a new method to overcome this difficulty by a careful examination of the normal form induced by  $(\mu_2)$  and a control in mixed Sobolev norm inspired by some tools used in numerical analysis. In particular, our result can be seen as a nonlinear extension of [DF07], here in a continuous time setting.

As a byproduct of this method, we prove the following.

**THEOREM 2.** *Let  $d \geq 2$ ,  $r \geq 2$  and  $s > d/2$ . For almost any  $m > 0$ , there exist  $s_* = s_*(r, s)$  and  $\varepsilon_0(r, s) > 0$  such that for all  $\varepsilon < \varepsilon_0$ , the solution to (NLW) satisfies*

$$\|(u(0), \dot{u}(0))\|_{H^{s_*} \times H^{s_*-1}} \leq \varepsilon \implies \|(u(t), \dot{u}(t))\|_{H^s \times H^{s-1}} \leq 2\varepsilon, \quad t \leq \varepsilon^{-r}. \quad (2)$$

In other words, (1) holds up to a finite loss of derivative in the initial condition. (Historically, a similar result with loss of derivative was first given under the condition  $(\mu_3)$  in [Bou96].) Concretely, we can prove that  $s_*(r, s)$  is of order less than  $\mathcal{O}(r^5 s^5)$ , but this is certainly not optimal.

The previous property (2) is in fact a corollary of a stronger abstract result proved in Theorem 3. The main idea is to decompose the dynamics into low and high frequencies according to some large threshold depending on  $\varepsilon$  and then to try to conjugate the system to a normal form whose dynamics can be described and controlled. We note that our abstract result applies to more general (nonlocal) wave equations where the nonlinearity is of order 1 instead of just 0 (see Remark 2).

When the  $(\mu_3)$  condition holds, this normal form approach allows to conjugate all the (NLW) flow to a flow preserving the  $H^s$  norm up to terms that are arbitrarily small. When only  $(\mu_2)$  is satisfied, this cannot be done, and linear terms remain in the dynamics of high modes, coupled with the low modes. Fortunately, these terms can be put under a *symmetric* Hamiltonian quadratic form. Hence, despite the linear nature of their dynamics, the  $L^2$  norm of the high modes can be proved to be preserved over long times in the normal form analysis. This crucial information allows to initiate an effective decomposition between low and high modes. To prove the preservation of higher Sobolev norms of the high modes, we use a sort of pseudodifferential argument (or commutator lemma) that allows to gain one derivative (see (49) and (50)). Then for given indices  $s$  and  $s_0$  with  $s \gg s_0 > d/2$ ,

we can control the low modes in a Sobolev norm  $H^s$  and we show a polynomial growth in time of a Sobolev norm  $H^{s_0}$  of the high modes of order  $\mathcal{O}(t^{s_0})$ . By choosing a smoother initial condition such that  $\|(u(0), \dot{u}(0))\|_{H^{2s} \times H^{2s-1}} \leq \varepsilon$ , we then obtain (see Theorem 4)

$$\|u(t)_{\leq N_\varepsilon}\|_{H^s} \leq 2\varepsilon \quad \text{and} \quad \|u(t)_{> N_\varepsilon}\|_{H^{s_0}} \leq \varepsilon^r, \quad t \leq \varepsilon^{-\frac{r}{s_0+1}}, \quad (3)$$

in a regime where  $s - s_0$  is large with respect to  $r$ . Here  $u(t)_{\leq N_\varepsilon}$  and  $u(t)_{> N_\varepsilon}$  denote the low and high modes parts, according to the threshold  $N_\varepsilon = \varepsilon^{-\frac{r}{s-s_0}}$ . When  $s$  is large, we obtain (2) after a change of indices, but Theorem 4 (see also Theorem 3) gives more precise information. It shows that the dynamics of the low modes preserves the superactions (Already used in [BG06] for (NLW) in dimension  $d = 1$  in a periodic setting.), that is, the quantities  $J_n(t) = \sum_{|k|=n} |u_k(t)|^2$ ,  $n \leq N_\varepsilon$  over very long times, where  $u_k(t)$ ,  $k \in \mathbb{Z}^d$  denotes the Fourier coefficients of  $u(t)$ . Such a result does not hold for high modes since the interaction between two close large modes cannot be eliminated but only controlled. Theorem 3 expresses the fact that the condition  $(\mu_2)$  – much more general than  $(\mu_3)$  – is enough to ensure a decoupling of the dynamics of low and high modes for very long times.

In the previous estimate,  $s_0$  has only to satisfy the condition  $s_0 > d/2$ . (Condition that could probably be refined in the critical case using Gagliardo–Nirenberg inequality.) It also typically corresponds to what is numerically observed for initial data taken as trigonometric polynomials for which the dynamics does not exhibit energy exchanges over long times. (Note that in the present result, all the constants depend on the Sobolev indices  $s$  and  $s_0$ . An optimization of these constants could be done in an analytic context, following the techniques used in [FG13].) Such two-stage norms with different Sobolev scaling were previously used in the context of numerical analysis of splitting methods for Schrödinger equations in the linear case (see [DF07, DF09]) where again the preservation of the  $L^2$  norm of the high modes was crucial to obtain a global control of the dynamics.

We also believe that results of the form (2) or (3) involving mixed Sobolev norms provide a natural setting for numerical discretization, for which  $(\mu_2)$  is the generic control of nonresonance condition. It might, for instance, allow to weaken the usual Courant-Friedrichs-Lewy (CFL) conditions required or to derive low-order integrators following the analysis of [HS17, OS18].

## 2. Abstract statement

**2.1. Hamiltonian formalism.** We recall in this subsection the formalism used in [Gre07, FG13, BFG] to deal with infinite-dimensional Hamiltonians and flows

depending on an infinite set of symplectic variables  $(q, p) = (q_a, p_a)_{a \in \mathbb{Z}^d} \in \mathbb{R}^{\mathbb{Z}^d} \times \mathbb{R}^{\mathbb{Z}^d}$  equipped with the usual  $\ell_s^2(\mathbb{Z}^d, \mathbb{R}^2)$  norm defined as

$$\|(q, p)\|_s^2 = \sum_{a \in \mathbb{Z}^d} \langle a \rangle^{2s} (p_a^2 + q_a^2), \quad \langle a \rangle^2 = 1 + |a|^2.$$

(With the usual notation  $|a|^2 = a_1^2 + \dots + a_d^2$  for  $a = (a_1, \dots, a_d) \in \mathbb{Z}^d$ .) As explained in [Gre07], for  $\mathcal{U}$  an open set of  $\ell_s^2$ , for a function  $H(q, p)$  such that  $H \in C^\infty(\mathcal{U}, \mathbb{R})$ , with  $\ell^2$  gradient  $\nabla_{(q,p)} H \in C^\infty(\mathcal{U}, \ell_s^2)$ , we can define the flow of a Hamiltonian system

$$\forall a \in \mathbb{Z}^d, \quad \dot{q}_a = \frac{\partial H}{\partial p_a}(q, p), \quad \dot{p}_a = -\frac{\partial H}{\partial q_a}(q, p). \tag{4}$$

To easily deal with normal form transformations, it is convenient to use the complex representation  $(\xi_a)_{a \in \mathbb{Z}^d} = (\frac{1}{\sqrt{2}}(q_a + ip_a))_{a \in \mathbb{Z}^d}$  in  $\mathbb{C}^{\mathbb{Z}^d}$  equipped with the  $\ell_s^2(\mathbb{Z}^d, \mathbb{C})$  norm. Then with the notations

$$\frac{\partial}{\partial \xi_a} = \frac{1}{\sqrt{2}} \left( \frac{\partial}{\partial q_a} - i \frac{\partial}{\partial p_a} \right) \quad \text{and} \quad \frac{\partial}{\partial \bar{\xi}_a} = \frac{1}{\sqrt{2}} \left( \frac{\partial}{\partial q_a} + i \frac{\partial}{\partial p_a} \right),$$

the real Hamiltonian system is equivalent to the complex system,

$$\forall a \in \mathbb{Z}^d, \quad \dot{\xi}_a = -i \frac{\partial H}{\partial \bar{\xi}_a}(\xi, \bar{\xi}) =: (X_H(\xi, \bar{\xi}))_a, \tag{5}$$

where  $H(q, p) = H(\xi, \bar{\xi}) \in \mathbb{R}$  is called a *real* Hamiltonian. The notation  $X_H(\xi, \bar{\xi}) = (X_H(\xi, \bar{\xi}))_{a \in \mathbb{Z}^d}$  thus denotes the Hamiltonian vector field associated with the Hamiltonian  $H$ . If we associate with  $(\xi, \bar{\xi})$  a complex function  $\psi$  on  $\mathbb{T}^d$ , through the formula

$$\psi(x) = \sum_{a \in \mathbb{Z}^d} \xi_a e^{iax}, \tag{6}$$

then the Sobolev norm  $\|\psi\|_{H^s}$  is equivalent to the norm

$$\|\xi\|_s^2 = \sum_{a \in \mathbb{Z}^d} \langle a \rangle^{2s} |\xi_a|^2.$$

The symplectic structure is given by

$$\sum_{a \in \mathbb{Z}^d} dq_a \wedge dp_a = -i \sum_{a \in \mathbb{Z}^d} d\xi_a \wedge d\bar{\xi}_a \tag{7}$$

and the Poisson bracket in complex notation reads

$$\{F, G\} = -i \sum_{a \in \mathbb{Z}^d} \frac{\partial F}{\partial \xi_a} \frac{\partial G}{\partial \bar{\xi}_a} - \frac{\partial F}{\partial \bar{\xi}_a} \frac{\partial G}{\partial \xi_a}. \tag{8}$$

DEFINITION 1. For a given  $s > d/2$  and a domain  $\mathcal{U}$  containing 0 in  $\ell_s^2 := \ell_s^2(\mathbb{Z}^d, \mathbb{C})$ , we denote by  $\mathcal{H}_s(\mathcal{U})$  the space of real Hamiltonians  $P(\xi, \bar{\xi}) \in \mathbb{R}$  satisfying

$$P \in C^\infty(\mathcal{U}, \mathbb{R}) \quad \text{and} \quad X_P \in C^\infty(\mathcal{U}, \ell_s^2),$$

where  $X_P$  is defined in (5). We will use the shortcut  $F \in \mathcal{H}_s$  to indicate that there exists a domain  $\mathcal{U}$  containing 0 in  $\ell_s^2$  such that  $F \in \mathcal{H}_s(\mathcal{U})$ .

Note that for  $F$  and  $G$  in  $\mathcal{H}_s$ , the formula (8) is well defined in a neighborhood of 0. With a given Hamiltonian function  $H \in \mathcal{H}_s$ , we associate the Hamiltonian system (5), and we can naturally define its flow.

PROPOSITION 1. Let  $s > d/2$ . Any Hamiltonian in  $\mathcal{H}_s$  defines a local flow in  $\ell_s^2$  which is a symplectic transformation.

**2.2. Polynomial Hamiltonians.** To algebraically deal with polynomials depending on  $(\xi, \bar{\xi})$ , we identify  $\mathbb{C}^{\mathbb{Z}^d} \times \mathbb{C}^{\mathbb{Z}^d} \simeq \mathbb{C}^{\mathbb{U}_2 \times \mathbb{Z}^d}$  where  $\mathbb{U}_2 = \{\pm 1\}$  and use the convenient notation  $(\xi, \bar{\xi}) = z = (z_j)_{j \in \mathbb{U}_2 \times \mathbb{Z}^d} \in \mathbb{C}^{\mathbb{U}_2 \times \mathbb{Z}^d}$  where

$$j = (\delta, a) \in \mathbb{U}_2 \times \mathbb{Z}^d \implies \begin{cases} z_j = \xi_a & \text{if } \delta = 1, \\ z_j = \bar{\xi}_a & \text{if } \delta = -1. \end{cases} \tag{9}$$

We define the  $\ell_s^2$  norm of an element  $z = (\xi, \bar{\xi})$  to be

$$\|z\|_s^2 := \sum_{j \in \mathbb{U}_2 \times \mathbb{Z}^d} \langle j \rangle^{2s} |z_j|^2 = 2 \sum_{a \in \mathbb{Z}^d} \langle a \rangle^{2s} |\xi_a|^2 = 2\|\xi\|_s^2,$$

where for  $j = (\delta, a) \in \mathbb{U}_2 \times \mathbb{Z}^d$ , we set  $\langle j \rangle^2 = \langle a \rangle^2$ . With this notation, the Hamiltonian system (5) can be written

$$\dot{z} = X_H(z) \quad \text{where } (X_H(z))_{(\delta,a)} := \delta(X_H(\xi, \bar{\xi}))_a, \quad z = (\xi, \bar{\xi}).$$

Here  $X_H(\xi, \bar{\xi})$  denote the vector field in (5). Another way to formulate this notation is to say that with the identification (9), the vector field  $X_H(\xi, \bar{\xi})$  is naturally extended as  $(X_H(z))_j = -i\delta \frac{\partial H}{\partial z_{\bar{j}}}(z)$ , for  $j = (\delta, a) \in \mathbb{U}_2 \times \mathbb{Z}^d$ , where  $\bar{j} = (-\delta, a)$ .

For  $\mathbf{k} = (k_1, \dots, k_m) = (\delta_i, a_i)_{i=1}^m \in (\mathbb{U}_2 \times \mathbb{Z}^d)^m$ , we denote the momentum

$$\mathcal{M}(\mathbf{k}) = \sum_{i=1}^m \delta_i a_i,$$

and we define the conjugate  $\bar{\mathbf{k}} = (-\delta_i, a_i)_{i=1}^m \in (\mathbb{U}_2 \times \mathbb{Z}^d)^m$ . We set

$$\mathcal{M}_m = \{\mathbf{k} \in (\mathbb{U}_2 \times \mathbb{Z}^d)^m \mid \mathcal{M}(\mathbf{k}) = 0\} \quad (10)$$

the set of zero-momentum multi-indices.

For a given  $\mathbf{k} \in \mathcal{M}_m$ , we write

$$z_{\mathbf{k}} = z_{k_1} \cdots z_{k_m}.$$

We also define

$$\begin{aligned} \mathcal{R}_m &= \{(\delta_j, a_j)_{j=1}^m \in \mathcal{M}_m \mid \exists \sigma \in \mathfrak{S}_m, \forall j = 1, \dots, m, \delta_j = -\delta_{\sigma_j} \\ &\text{and } \langle a_j \rangle = \langle a_{\sigma_j} \rangle\} \end{aligned} \quad (11)$$

the set of *resonant* multi-indices. Note that by construction if  $m$  is odd, then  $\mathcal{R}_m$  is empty and that if  $\mathbf{k} = (\delta_j, a_j)_{j=1}^m \in \mathcal{R}_m$  is associated with a permutation  $\sigma$ , then we have

$$z_{\mathbf{k}} = \prod_{\substack{1 \leq j \leq m \\ \delta_j = 1}} \xi_{a_j} \bar{\xi}_{a_{\sigma_j}}. \quad (12)$$

**DEFINITION 2.** We say that  $P(z)$  is a homogeneous polynomial of order  $m$  if it can be written as

$$P(z) = P[c](z) = \sum_{j \in \mathcal{M}_m} c_j z_j \quad \text{with } c = (c_j)_{j \in \mathcal{M}_m} \in \ell^\infty(\mathcal{M}_m), \quad (13)$$

and such that the coefficients  $c_j$  satisfy  $c_{\bar{j}} = \bar{c}_j$ .

Note that the last condition ensures that  $P$  is real valued as the set of indices are invariant by the application  $j \mapsto \bar{j}$ . Following [FG13, BFG] but in a  $\ell_s^2$  framework, we get the following proposition. It turns out to be a consequence of the more general Lemma 2 proved below.

**PROPOSITION 2.** *Let  $s > d/2$ .*

- (i) *A homogeneous polynomial,  $P[c]$ , of degree  $m \geq 2$  belongs to  $\mathcal{H}_s(\ell_s^2)$  and we have*

$$\|X_{P[c]}(z)\|_s \leq (C_s)^m \|c\|_{\ell^\infty} \|z\|_s^{m-1}, \quad z = (\xi, \bar{\xi}), \quad (14)$$

*for some constant  $C_s$  depending only on  $s$ .*

- (ii) *For two homogeneous polynomials  $P[c]$  and  $P[c']$  of degree, respectively,  $m$  and  $n$ , the Poisson bracket is a homogeneous polynomial of degree  $m+n-2$ ,*



$\{P[c], P[c']\} = P[c'']$  and we have the estimate

$$\|c''\|_{\ell^\infty} \leq 2mn \|c\|_{\ell^\infty} \|c'\|_{\ell^\infty}. \quad (15)$$

We end this section with a result concerning Lie transformations. We recall that to a Hamiltonian function  $F$ , we associate, if it exists, the Lie transformation  $\Phi_F^1$  which is the time one flow generated by  $F$ . This transformation is automatically symplectic.

LEMMA 1. Assume that  $s > d/2$  and let  $P = P[c_3] + \dots + P[c_r]$ ,  $c_m \in \mathcal{M}_m$ , be a polynomial of order at least 3 at the origin, then for all  $\nu \leq \kappa_{r,s} (\max(\|c_3\|_{\ell^\infty}, \dots, \|c_r\|_{\ell^\infty}))^{-1}$ , the Lie transformation  $\tau = \Phi_P^1$  is well defined on a ball  $B_s(0, \nu)$  of  $\ell_s^2$  with values in  $B_s(0, 2\nu)$  and we have

$$\|\tau(z) - z\|_s \leq C_{r,s} \max_m \|c_m\|_{\ell^\infty} \nu^2, \quad \forall z \in B_s(0, \nu), \quad (16)$$

where  $C_{r,s}$  and  $\kappa_{r,s}$  are some positive constants depending only on  $r$  and  $s$ .

*Proof.* In view of Proposition 2 assertion (i), the vector field  $X_{P[c]}(z)$  is locally Lipschitz in  $\ell_s^2$  and we have an estimate of the form

$$\|X_{P[c]}(z) - X_{P[c]}(z')\|_s \leq C_{s,m} \|c\|_{\ell^\infty} \max(\|z\|_s^{m-2}, \|z'\|_s^{m-2}) \|z - z'\|_s$$

for some constant  $C_{s,m}$  and for  $z$  and  $z'$  in  $\ell_s^2$ . We deduce from the Cauchy–Lipschitz theorem that the flow  $\Phi_P^t$  is locally well defined on  $\ell_s^2$ . Furthermore, we have for  $z \in B_s(0, \nu)$  and as long as  $z(t) = \Phi_P^t(z) \in B(0, 2\nu)$ ,

$$\begin{aligned} \|\Phi_P^t(z) - z\|_s &\leq \left| \int_0^t \|X_P(z(w))\|_s dw \right| \leq \sum_{m=3}^r C_{m,s} \|c_m\|_{\ell^\infty} |t| (2\nu)^{m-1} \\ &\leq C_{r,s} |t| \max_m \|c_m\|_{\ell^\infty} \nu^2. \end{aligned}$$

Thus, we conclude by a bootstrap argument that, taking  $\nu$  small enough, the flow is defined for all  $z \in B(0, \nu)$  up to time  $t = 1$  and that  $z(t)$  remains in  $B(0, 2\nu)$  for  $t \leq 1$  and satisfies (16).  $\square$

### 2.3. Nonresonance condition.

DEFINITION 3. A family of frequencies  $\omega = \{\omega_a, a \in \mathbb{Z}^d\}$  is *nonresonant* if there exist nondecreasing sequences  $(\alpha(r))_{r \geq 1} \in (\mathbb{R}_+^*)^{\mathbb{N}}$  and  $(\gamma_r)_{r \geq 1} \in (\mathbb{R}_+^*)^{\mathbb{N}}$  such that

for all  $r \geq 1$ , all  $N \geq 1$  and all  $\mathbf{k} = (\delta_i, a_i)_{i=1}^r \in (\mathbb{U}_2 \times \mathbb{Z}^d)^r$  satisfying  $\langle a_i \rangle \leq N$ , for  $i = 1, \dots, r$ , we have

$$|\delta_1 \omega_{a_1} + \dots + \delta_r \omega_{a_r}| \geq \gamma_r N^{-\alpha(r)} \quad \text{when } \mathbf{k} \notin \mathcal{R}_r, \tag{H1}$$

$$|\delta_1 \omega_{a_1} + \dots + \delta_r \omega_{a_r} + \omega_b| \geq \gamma_{r+1} N^{-\alpha(r+1)}, \quad \forall \langle b \rangle > N \text{ with } \mathcal{M}(\mathbf{k}) + b = 0, \tag{H2}$$

$$|\delta_1 \omega_{a_1} + \dots + \delta_r \omega_{a_r} + \omega_{b_1} + \omega_{b_2}| \geq \gamma_{r+2} N^{-\alpha(r+2)}, \quad \forall \langle b_1 \rangle, \langle b_2 \rangle > N \tag{H3}$$

with  $\mathcal{M}(\mathbf{k}) + b_1 + b_2 = 0$ .

We note that conditions (H1)–(H2) are equivalent to condition  $(\mu_2)$  introduced in Section 1, while conditions (H1)–(H2)–(H3) are not equivalent to  $(\mu_3)$  since in (H3), we are not considering the case where the two high frequencies have opposite sign.

REMARK 1. Note that in (H2) using the zero-momentum condition (see (10)),  $\langle b \rangle$  is in fact bounded by  $rN$ . Hence, (H2) is a trivial consequence of (H1). Similarly, as in many applications  $\omega_a \sim |a|^v$  when  $a \rightarrow \infty$ , for some  $v > 0$ , (H3) is also a consequence of (H1) as we can restrict (H3) to a set of  $(b_1, b_2)$  that are bounded by  $C(r)N$ .

**2.4. Statement.** Let us start with the following notation: for  $\xi \in \ell_s^2$  and a given number  $N$ , we decompose  $\xi = \xi_{\leq N} + \xi_{> N}$  where for all  $j \in \mathbb{U}_2 \times \mathbb{Z}^d$ ,

$$(\xi_{\leq N})_j = \begin{cases} \xi_j & \text{for } \langle j \rangle \leq N, \\ 0 & \text{for } \langle j \rangle > N, \end{cases} \quad \text{and} \quad \xi_{> N} = \xi - \xi_{\leq N}.$$

Given a function  $\psi \in H^s(\mathbb{T}^d)$  with Fourier coefficients  $\xi_a, a \in \mathbb{Z}^d$  and a number  $N \geq 1$ , the previous decomposition induces naturally the decomposition  $\psi = \psi_{\leq N} + \psi_{> N}$  with

$$\psi_{\leq N}(x) = \sum_{\langle a \rangle \leq N} \xi_a e^{ia \cdot x} \quad \text{and} \quad \psi_{> N}(x) = \sum_{\langle a \rangle > N} \xi_a e^{ia \cdot x}.$$

Similarly, we note by  $z_{\geq N}$  and  $z_{> N}$  the decomposition induced by the notation (9).

We now make our hypothesis on the Hamiltonian  $H$  that we will consider.

HYPOTHESIS 1. *The Hamiltonian  $H$  can be written as*

$$H = H_2 + P = \sum_{a \in \mathbb{Z}^d} \omega_a |\xi_a|^2 + P, \tag{17}$$

where  $\omega = (\omega_a)_{a \in \mathbb{Z}^d}$  is a nonresonant family of real numbers in the sense of Definition 3,  $P$  belongs to  $\mathcal{H}_s$  for some  $s > d/2$  and  $P$  is of order at least 3 at the origin (which means that  $P$  and its differentials up to the order 2 vanish at 0).

Note that in general the frequencies  $\omega_a$  will not be uniformly bounded with respect to  $a$ , and hence the quadratic part  $\sum_{a \in \mathbb{Z}^d} \omega_a |\xi_a|^2$  does not belong to  $\mathcal{H}_s$ . Nevertheless, it generates a continuous flow which maps  $\ell_s^2$  into  $\ell_s^2$  explicitly given for all time  $t$  and for all indices  $a$  by  $\xi_a(t) = e^{-i\omega_a t} \xi_a(0)$ . Furthermore, this flow has the group property. By standard arguments (see, for instance, [Caz03, BFG] in a similar framework), this is enough to define the local symplectic flow,  $\dot{z} = X_H(z)$ , in  $\ell_s^2$ .

**THEOREM 3.** *Let  $H$  be a Hamiltonian satisfying Hypothesis 1. Then for all  $r \geq 2$  and all  $s > s_0 > d/2$  satisfying*

$$s - s_0 \geq s_*(r) := 6r^2\alpha(3r) + 2dr, \tag{18}$$

there exists  $\varepsilon_0(r, s, s_0, \omega) > 0$  such that for all  $\varepsilon < \varepsilon_0(r, s, s_0, \omega)$  the solution  $\xi(t)$ , generated by the flow of  $H$  issued from an initial datum  $\xi(0) \in \ell_{2s}^2$ , satisfying  $\|\xi(0)\|_{2s} \leq \varepsilon$ , exists for all time  $t \leq \varepsilon^{-\frac{r}{s_0+1}}$  and satisfies

$$\begin{aligned} \forall \langle a \rangle \leq N_\varepsilon &:= \varepsilon^{-\frac{r}{s-s_0}}, \\ \forall t \leq \varepsilon^{-\frac{r}{s_0+1}}, & \left| \sum_{\substack{b \in \mathbb{Z}^d \\ \langle b \rangle = \langle a \rangle}} |\xi_b(t)|^2 - \sum_{\substack{b \in \mathbb{Z}^d \\ \langle b \rangle = \langle a \rangle}} |\xi_b(0)|^2 \right| \leq \varepsilon^3 \langle a \rangle^{-2s} \end{aligned} \tag{19}$$

and

$$\begin{cases} \|\xi(t)_{\leq N_\varepsilon}\|_s^2 = \sum_{\langle a \rangle \leq N_\varepsilon} \langle a \rangle^{2s} |\xi_a(t)|^2 \leq 4\varepsilon^2, \\ \|\xi(t)_{> N_\varepsilon}\|_{s_0}^2 = \sum_{\langle a \rangle > N_\varepsilon} \langle a \rangle^{2s_0} |\xi_a(t)|^2 \leq \varepsilon^{2r}, \end{cases} \quad \text{for } t \leq \varepsilon^{-\frac{r}{s_0+1}}.$$

Note that by playing with the indices, we can obtain the following corollary (in the same vein as Theorem 2) which essentially proves that arbitrary high regularity small solutions to (5) are controlled over arbitrary long times.

**COROLLARY 1.** *Let  $H$  be a Hamiltonian satisfying Hypothesis 1. For all  $r \geq 2$  and  $s > d/2$ , there exist  $s_*(r, s)$  and  $\varepsilon_0(r, s, \omega) > 0$  such that for all  $\varepsilon < \varepsilon_0(r, s, \omega)$ ,*

$$\|\xi(0)\|_{s_*} \leq \varepsilon \implies \|\xi(t)\|_s \leq 2\varepsilon \quad \text{for } t \leq \varepsilon^{-r}.$$

*Proof.* It is a consequence of the previous theorem by replacing  $r$  by  $r(s_0 + 1)$ ,  $s_0$  by  $s$  and assuming that  $s_*$  is large enough with respect to  $r$  and  $s$ .  $\square$

### 3. Application to (NLW) on $\mathbb{T}^d$

Introducing  $v = u_t \equiv \dot{u}$ , we rewrite (NLW) as

$$\begin{cases} \dot{u} = v, \\ \dot{v} = -\Lambda^2 u - f(u), \end{cases} \quad (20)$$

where  $\Lambda = (-\Delta + m)^{1/2}$  and  $f \in C^\infty$  having a zero of order at least 2 at the origin. When  $m > 0$ , we can define

$$\psi = \frac{1}{\sqrt{2}}(\Lambda^{1/2}u + i\Lambda^{-1/2}v), \quad (21)$$

and we get that  $(u, v) \in H^s(\mathbb{T}^d, \mathbb{R}) \times H^{s-1}(\mathbb{T}^d, \mathbb{R})$  is a solution of (20) if and only if  $\psi \in H^{s+1/2}(\mathbb{T}^d, \mathbb{C})$  is a solution of

$$i\dot{\psi} = \Lambda\psi + \frac{1}{\sqrt{2}}\Lambda^{-1/2}f\left(\Lambda^{-1/2}\left(\frac{\psi + \bar{\psi}}{\sqrt{2}}\right)\right). \quad (22)$$

Then, endowing the space  $L^2(\mathbb{T}^d, \mathbb{C})$  with the standard (real) symplectic structure  $-id\psi \wedge d\bar{\psi}$ , equation (22) reads as a Hamiltonian equation

$$i\dot{\psi} = \frac{\partial H}{\partial \bar{\psi}},$$

where  $H$  is the Hamiltonian function

$$H(\psi, \bar{\psi}) = \frac{1}{(2\pi)^d} \int_{\mathbb{T}^d} (\Lambda\psi)\bar{\psi} \, dx + \frac{1}{(2\pi)^d} \int_{\mathbb{T}^d} F\left(\Lambda^{-1/2}\left(\frac{\psi + \bar{\psi}}{\sqrt{2}}\right)\right) \, dx,$$

and  $F$  is a primitive of  $f$  with respect to the variable  $u$ , that is,  $f = \partial_u F$ . The linear operator  $\Lambda$  is diagonal in the complex Fourier basis  $\{e^{ia \cdot x}\}_{a \in \mathbb{Z}^d}$ , with eigenvalues

$$\omega_a = \sqrt{|a|^2 + m}, \quad a = (a_1, \dots, a_d) \in \mathbb{Z}^d, \quad |a|^2 = a_1^2 + \dots + a_d^2. \quad (23)$$

(Here for  $a = (a_1, \dots, a_d) \in \mathbb{Z}^d$  and  $x = (x_1, \dots, x_d) \in \mathbb{T}^d$ , we set  $a \cdot x = a_1x_1 + \dots + a_dx_d$ .) Decomposing  $\psi$  in Fourier variables with Fourier coefficients

$(\xi_a)_{a \in \mathbb{Z}^d}$  as in (6), (22) takes the form (5) where the Hamiltonian function  $H$  is given by  $H = H_2 + P$  with

$$H_2(\xi, \bar{\xi}) = \sum_{a \in \mathbb{Z}^d} \omega_a |\xi_a|^2,$$

$$P(\xi, \bar{\xi}) = \frac{1}{(2\pi)^d} \int_{\mathbb{T}^d} F \left( \sum_{a \in \mathbb{Z}^d} \frac{\xi_a e^{ia \cdot x} + \bar{\xi}_a e^{-ia \cdot x}}{\sqrt{2\omega_a}} \right) dx.$$

As  $F$  is smooth, the function  $P$  is in  $\mathcal{H}_s$  and we can define its flow. Finally,  $(u, v) \in H^s(\mathbb{T}^d, \mathbb{R}) \times H^{s-1}(\mathbb{T}^d, \mathbb{R})$  is a solution of (20) if and only if  $\xi \in \ell^2_{s-1/2}$  is a solution of the Hamiltonian system associated with the Hamiltonian  $H$ .

In order to apply Theorem 3, we need the following result.

**PROPOSITION 3** [Del09, Theorem 2.1.1]. *For almost all  $m \in (0, +\infty)$ , the family of frequencies*

$$\omega_a(m) = \sqrt{|a|^2 + m}, \quad a \in \mathbb{Z}^d$$

*associated with (NLW) is nonresonant in the sense of Definition 3.*

A direct proof of this proposition can also be done by using the arguments given in [Bam03, EGK16] or [FGL13]. By following these proofs, one can verify that  $\alpha(r)$  is of order  $\mathcal{O}(r^3)$ .

As a consequence, Theorem 3 applies. By scaling back to the variable  $(u, v)$ , we obtain the following.

**THEOREM 4.** *Let  $f$  be a  $C^\infty$  function with zero of order at least 2 at the origin. For almost all  $m \in (0, +\infty)$ , for all  $r \geq 2$  and all  $s_0 > (d+1)/2$ , there exists  $s_1(r, s_0)$  such that for all  $s \geq s_1(r, s_0)$ , there exists  $\varepsilon_0(r, s_0, s, m)$  and for all  $\varepsilon < \varepsilon_0(r, s_0, s, m)$ , if  $(u(0), \dot{u}(0)) \in H^{2s} \times H^{2s-1}(\mathbb{T}^d)$  satisfies  $\|(u(0), \dot{u}(0))\|_{H^{2s} \times H^{2s-1}} \leq \varepsilon$ , then the system (NLW) admits a solution over a time  $T \geq \varepsilon^{-\frac{r}{s_0+1}}$ , and we have*

$$\begin{cases} \|u(t)_{\leq N_\varepsilon}\|_s \leq 2\varepsilon, \\ \|u(t)_{> N_\varepsilon}\|_{s_0} \leq \varepsilon^r, \end{cases} \quad \text{for } t \leq \varepsilon^{-\frac{r}{s_0+1}}$$

where  $N_\varepsilon := \varepsilon^{-\frac{r}{s-s_0}}$ .

Using that  $\alpha(r) = \mathcal{O}(r^3)$ , we can verify that  $s_1(r, s) \simeq (rs)^5$ . Then Theorem 2 is just a reformulation of Corollary 1.

REMARK 2. The previous theorem also straightforwardly apply to nonlocal wave equations of the form

$$\ddot{u} = -\Lambda^2 u - \Lambda^\beta f(\Lambda^\beta u), \quad 0 \leq \beta \leq \frac{1}{2},$$

instead of (20). Indeed, in this case, (22) shows that the nonlinearity is well defined in  $\ell_s^2$  and Theorem (2) can be applied.

#### 4. Normal form

The strategy used to prove Theorem 3 is to put the original Hamiltonian (17) into normal form eliminating most of the interactions between the low and high frequencies. By using Taylor expansion at the origin, the Hamiltonian  $H$  can be written as

$$H = H_2 + \sum_{m=3}^r P_m + R_{r+1} \quad \text{with } H_2 = \sum_{a \in \mathbb{Z}^d} \omega_a |\xi_a|^2, \quad (24)$$

where  $P_m$  is a homogeneous polynomial of degree  $m$  and where  $R_{r+1} \in \mathcal{H}_s$  is of order  $r+1$  which means that its differentials vanish up to the order  $r$ . In particular, it is small in the sense that we have

$$\|X_{R_{r+1}}(z)\|_s \leq (C_s)^r \|z\|_s^r \quad (25)$$

for some constant depending only on  $s$  and for  $z$  small enough in  $\ell_s^2$ .

For  $\mathbf{j} \in \mathcal{M}_m$ , let us denote by  $\mu_n(\mathbf{j})$  the  $n$ th largest number among the collection  $\langle j_i \rangle_{i=1}^m$ :

$$\mu_1(\mathbf{j}) \geq \mu_2(\mathbf{j}) \geq \dots \geq \mu_m(\mathbf{j}).$$

By convention, we will also set  $\mu_0(\mathbf{j}) = +\infty$ .

Let  $N$  be a fixed number and let  $H$  be a Hamiltonian satisfying Hypothesis 1. For a given  $r$ , we decompose the Hamiltonian  $H$  in (24) as follows:

$$H = H_2 + \sum_{m=3}^r (P_m^{(\circ)} + P_m^{(i)} + P_m^{(ii)} + P_m^{(iii)}) + R_{r+1},$$

where for all  $m$ ,

- the polynomial  $P_m^{(\circ)} = P[c_m^{(\circ)}]$  depends only on low modes:

$$\forall \mathbf{j} \in \mathcal{M}_m, \quad (c_m^{(\circ)})_j \neq 0 \implies \mu_1(\mathbf{j}) \leq N. \quad (26)$$

- $P_m^{(i)} = P[c_m^{(i)}]$  contains only one high mode:

$$\forall \mathbf{j} \in \mathcal{M}_m, \quad (c_m^{(i)})_{\mathbf{j}} \neq 0 \implies \mu_1(\mathbf{j}) > N \geq \mu_2(\mathbf{j}). \quad (27)$$

- $P_m^{(ii)} = P[c_m^{(ii)}]$  contains only two high modes:

$$\forall \mathbf{j} \in \mathcal{M}_m, \quad (c_m^{(ii)})_{\mathbf{j}} \neq 0 \implies \mu_2(\mathbf{j}) > N \geq \mu_3(\mathbf{j}). \quad (28)$$

- $P_m^{(iii)} = P[c_m^{(iii)}]$  contains at least three high modes:

$$\forall \mathbf{j} \in \mathcal{M}_m, \quad (c_m^{(iii)})_{\mathbf{j}} \neq 0 \implies \mu_3(\mathbf{j}) > N. \quad (29)$$

Following [BG06], we know that polynomials of the form  $P_m^{(iii)} = P[c_m^{(iii)}]$  are already small in the sense that

$$\|X_{P_m^{(iii)}}\|_s \leq CN^{-s} \|z\|_s^{m-1}.$$

On the other hand, thanks to our nonresonance condition, polynomials of the form  $P_m^{(i)} = P[c_m^{(i)}]$  can be killed by a symplectic change of variables since it cannot be resonant (only one high mode) and polynomials of the form  $P_m^{(o)} = P[c_m^{(o)}]$  can be normalized by a standard Birkhoff normal form procedure. In addition to these two known facts, the following theorem says that we can also symmetrize polynomials of the form  $P_m^{(ii)} = P[c_m^{(ii)}]$ .

**THEOREM 5.** *Assume that the frequencies  $\omega = (\omega_a)_{a \in \mathbb{Z}^d}$  are nonresonant in the sense of Definition 3, and let  $r \geq 2$  be given. There exists a constant  $C$  depending on  $r$  such that for all  $N \geq 1$ , there exists a polynomial Hamiltonian*

$$\chi = \sum_{m=3}^r (\chi_m^{(o)} + \chi_m^{(i)} + \chi_m^{(ii)}) \quad (30)$$

such that  $\chi_m^{(o)} = P[a_m^{(o)}]$ ,  $\chi_m^{(i)} = P[a_m^{(i)}]$  and  $\chi_m^{(ii)} = P[a_m^{(ii)}]$  contain zero, one and two high modes, respectively (that is, satisfy (26), (27) and (28) respectively), with coefficients satisfying

$$\|a_m^{(o)}\|_{\ell^\infty} + \|a_m^{(i)}\|_{\ell^\infty} + \|a_m^{(ii)}\|_{\ell^\infty} \leq CN^{r\alpha(r)}, \quad (31)$$

and such that the Lie transformation  $\Phi_\chi^1$ , whose existence is guaranteed in a neighborhood of the origin of  $\ell_s^2$  for all  $s > d/2$  by Lemma 1, puts  $H$  in normal form:

$$H \circ \Phi_\chi^1 = H_2 + \sum_{m=3}^r (Z_m^{(o)} + S_m^{(ii)} + \tilde{P}_m^{(iii)}) + \tilde{R}_{r+1}, \quad (32)$$

where  $Z_m^{(\circ)} = P[b_m^{(\circ)}]$ ,  $S_m^{(ii)} = P[b_m^{(ii)}]$  and  $\tilde{P}_m^{(iii)} = P[b_m^{(iii)}]$  contain zero, two and at least three high modes, respectively (that is, satisfy (26), (28), (29)), with coefficients satisfying

$$\|b_m^{(\circ)}\|_{\ell^\infty} + \|b_m^{(ii)}\|_{\ell^\infty} + \|b_m^{(iii)}\|_{\ell^\infty} \leq CN^{r\alpha(r)}.$$

Moreover,

- $Z_m^{(\circ)}$  contains only resonant monomials, which means that

$$\forall j \in \mathcal{M}_m, \quad j \notin \mathcal{R}_m \implies (b_m^{(\circ)})_j = 0.$$

- $S_m^{(ii)}$  contains terms that are symmetric in the high modes which means that if  $b_j^{(ii)} \neq 0$  for  $j \in \mathcal{M}_m$ , the two highest modes are of opposite signs: they are of the form  $(\delta, a)$  and  $(-\delta, b)$  for some  $a$  and  $b \in \mathbb{Z}^d$ .
- The remainder term  $\tilde{R}_{r+1}$  is of the form

$$\tilde{R}_{r+1} = R_{r+1} \circ \Phi_\chi^1 + \int_0^1 (1-s)^{r+1} P[b_{r+1}] \circ \Phi_\chi^s ds, \tag{33}$$

where  $P[b_{r+1}]$  defines a homogeneous polynomial of order  $r + 1$  with coefficients bounded by  $\|b_{r+1}\|_{\ell^\infty} \leq CN^{r\alpha(r)}$ .

*Proof.* The proof is standard and uses the nonresonant Birkhoff normal form procedure (see [BG06, BDGS07, Gre07]). We follow here the construction made in [FG13]. By using the formal series expansions  $H = H_2 + \sum_{m \geq 3} P_m$  and  $\chi = \sum_{m \geq 3} \chi_m$  in homogeneous polynomials, the formal normal form problem is to find  $\chi$  and  $X$  (under normal form) such that

$$H \circ \Phi_\chi^1 = \sum_{k \geq 0} \text{ad}_\chi^k (H_2 + P) = H_2 + X.$$

In the formal series algebra, this problem is equivalent to a sequence of homological equation of the form

$$\forall m \geq 3, \quad \{H_2, \chi_m\} = Q_m - X_m,$$

where  $Q_m$  depends on the function  $P_k, \chi_k$  and  $X_k, k < m$  previously constructed. It is obtained by iterated Poisson brackets preserving the homogeneity of polynomial and boundedness of coefficients (see (15)). Formula for  $Q_m$  can be found in [FG13, Equation (3.4)].

Now assume that  $Q_m = \sum_{j \in \mathcal{M}_m} q_j z_j$  is given. For a given  $N$ , we can decompose it into terms containing zero, one, two and at least three high modes:



$Q_m = Q_m^{(o)} + Q_m^{(i)} + Q_m^{(ii)} + Q_m^{(iii)}$  and a similar decomposition for the coefficients  $q_j$ . The normal form term  $X_m$  is then the sum of the resonant terms in  $Q_m^{(o)}$  (contributing to the term  $Z_m^{(o)}$ ), the symmetric part of  $Q_m^{(ii)}$  (contributing to the term  $S_m^{(ii)}$ ) and the term  $Q_m^{(iii)}$  (contributing to the term  $\tilde{P}_m^{(iii)}$ ). By noting that

$$\{H_2, z_k\} = i\Omega(k)z_k \quad \text{with } \Omega(k) = -\sum_{i=1}^m \delta_i \omega_{a_i}, \quad k = ((\delta_i, a_i))_{i=1}^m \in \mathcal{M}_m,$$

we then solve the other terms by setting

$$\chi_m = \sum_{j \in \mathcal{M}_m \setminus \mathcal{R}_m} \frac{q_j^{(o)}}{i\Omega(j)} z_j + \sum_{j \in \mathcal{M}_m} \frac{q_j^{(i)}}{i\Omega(j)} z_j + \sum_{j \in \mathcal{M}_m \setminus \mathcal{S}_m} \frac{q_j^{(ii)}}{i\Omega(j)} z_j,$$

where  $\mathcal{S}_m$  denote the set of indices with two symmetric high modes. Note that when  $j \in \mathcal{M}_m \setminus \mathcal{S}_m$ , the two highest modes (larger than  $N$ ) have the same sign and the denominator  $\Omega(j)$  is controlled by (H3). Similarly, the first term can be controlled using (H1) and the second using (H2). We then observe that we loose a factor  $N^{\alpha(m)}$  after each solution of the homological equation, yielding a bound of order  $N^{r\alpha(r)}$  after  $r$  iterations. Note also that all the operations (solution of the homological equation and Poisson brackets) preserve homogeneity and the reality of the global Hamiltonians.

It is easy to see that for all  $s > s_0$  for  $z$  small enough (such that  $C_{r,s} N^{r\alpha(r)} \|z\|_s \leq 1$  for some constant  $C_{r,s}$  by using (14)), the flow  $\Phi_\chi^1(z)$  is well defined and locally invertible in  $\ell_s^2$  (its inverse being  $\Phi_\chi^{-1}$ ); see Lemma 1.

Finally, to obtain (32), we use a Taylor expansion of the term  $H \circ \Phi_\chi^t$  for  $t \in (0, 1)$ . □

For  $\xi \in \ell_s^2$ , we define the *pseudoactions*:

$$J_a(\xi, \bar{\xi}) = \sum_{\substack{b \in \mathbb{Z}^d \\ \langle a \rangle = \langle b \rangle}} |\xi_b|^2 \quad \left( \text{and } J_j = \frac{1}{2} \sum_{\substack{\ell \in \mathbb{U}_2 \times \mathbb{Z}^d \\ \langle \ell \rangle = \langle j \rangle}} |z_\ell|^2, j \in \mathbb{U}_2 \times \mathbb{Z}^d \right). \quad (34)$$

By definition of the resonant set  $\mathcal{R}_m$  (see (11)) and the corresponding resonant monomials (see (12)), we see that for all  $m$ , the normal form terms  $Z_m^{(o)}$  can be written as

$$Z_m^{(o)}(\xi, \bar{\xi}) = \sum_{\substack{k \in \mathcal{R}_m \\ \forall i, \langle k_i \rangle \leq N}} c_k z_k = \sum_{\substack{a, b \in (\mathbb{Z}^d)^{\frac{m}{2}} \\ \forall i, \langle a_i \rangle = \langle b_i \rangle \leq N}} c_{ab} \prod_{i=1}^{\frac{m}{2}} \xi_{a_i} \bar{\xi}_{b_i}, \quad z = (\xi, \bar{\xi}).$$

(Recall that  $Z_m^{(\circ)} = 0$  when  $m$  is odd.) We note that a polynomial in normal form commutes with all the *pseudoactions*:

$$\{Z_m^{(\circ)}, J_a\} = 0, \quad \forall a \in \mathbb{Z}, \tag{35}$$

in such a way that the flow generated by  $Z_m^{(\circ)}$  will not modify the  $\ell_s^2$  norms since  $\|z\|_s^2 = \sum (1 + a^2)^{2s} J_a$ .

### 5. Proof of the main theorem

To prove the main theorem, we will need the following lemma which controls polynomial vector fields in mixed Sobolev norms, as high and low modes are not controlled at the same Sobolev scale. The following lemma is an extension of Proposition 2. It will be used to control the dynamics in the new variables and the effect of normal form transformation.

LEMMA 2. *For all  $s \geq s_0 > d/2$ , there exists a constant  $C$  such that for all  $m \geq 3$ , for all  $N \geq 1$  and for all  $c \in \ell^\infty(\mathcal{M}_m)$ , if the polynomial  $P[c]$  contains at least  $n$  high modes,  $n = 0, 1, 2, 3$  (that is,  $c_j \neq 0 \implies \mu_n(j) > N$ ), then we have for  $z = (\xi, \bar{\xi})$  with  $z_{\leq N} \in \ell_s^2$  and  $z_{>N} \in \ell_{s_0}^2$*

$$\|X_{P[c]}(z)_{\leq N}\|_s \leq \begin{cases} C^m \|c\|_{\ell^\infty} \|z_{\leq N}\|_s^{m-1} & \text{if } n = 0, \\ C^m \|c\|_{\ell^\infty} N^{n(s-s_0)} \|z_{\leq N}\|_s^{m-n-1} \|z_{>N}\|_{s_0}^n & \text{if } n \geq 1, \end{cases} \tag{36}$$

and

$$\|X_{P[c]}(z)_{>N}\|_{s_0} \leq \begin{cases} 0 & \text{if } n = 0, \\ C^m \|c\|_{\ell^\infty} N^{s_0-s} \|z_{\leq N}\|_s^{m-n-1} & \text{if } n = 1, \\ C^m \|c\|_{\ell^\infty} \|z_{\leq N}\|_s^{m-n} \|z_{>N}\|_{s_0}^{n-1} & \text{if } n \geq 2. \end{cases} \tag{37}$$

*Proof.* For  $z = (\xi, \bar{\xi}) \in \ell_s^2$ , we have

$$\|X_{P[c]}(z)_{\leq N}\|_s^2 \leq \sum_{\langle \ell \rangle \leq N} \langle \ell \rangle^{2s} \left| \frac{\partial P[c]}{\partial z_\ell}(z) \right|^2 \leq m^2 \|c\|_{\ell^\infty}^2 \sum_{\langle \ell \rangle \leq N} \langle \ell \rangle^{2s_0} P_\ell(z)^2,$$

where

$$P_\ell(z) = \langle \ell \rangle^{s-s_0} \sum_{\substack{\mathcal{M}(j)=0 \\ \mu_n(j, \ell) > N}} |z_{j_1}| \cdots |z_{j_{m-1}}|. \tag{38}$$

Let  $j = (j_1, \dots, j_m)$  be given such that  $\mathcal{M}(j, \ell) = 0$ . We have

$$\langle \ell \rangle^{s-s_0} \leq (\langle j_1 \rangle + \cdots + \langle j_{m-1} \rangle)^{s-s_0} \leq m^{s-s_0} \langle j_1 \rangle^{s-s_0} \cdots \langle j_{m-1} \rangle^{s-s_0}.$$

If  $P[c]$  contains no high modes, that is, if  $n = 0$ , then we can define  $\tilde{z}_j = \langle j \rangle^{s-s_0} z_j$  and we obtain

$$P_\ell(z) \leq \sum_{\substack{\mathcal{M}(j,\ell)=0 \\ \mu_1(j) \leq N}} |\tilde{z}_{j_1}| \cdots |\tilde{z}_{j_{m-1}}|.$$

For a function  $g$  on the torus, we will denote  $\hat{g}_b = (1/2\pi)^d \int_{\mathbb{T}^d} g(x) e^{-ib \cdot x} dx$ ,  $b \in \mathbb{Z}^d$ , the Fourier coefficients of  $g$ . Then denoting  $f(x) = \sum_{\langle \ell \rangle \leq N} |\tilde{z}_\ell| e^{i\delta a \cdot x}$ , we have

$$\sum_{\substack{\mathcal{M}(j,\ell)=0 \\ \mu_1(j) \leq N}} |\tilde{z}_{j_1}| \cdots |\tilde{z}_{j_{m-1}}| = \widehat{(f^{m-1})}_{-\delta a}, \quad \ell = (\delta, a) \in \mathbb{U}_2 \times \mathbb{Z}^d.$$

So we get for some generic constant  $A$  depending on  $s$  and  $s_0$  but independent on  $m$

$$\begin{aligned} \|X_{P[c]}(z)_{\leq N}\|_s^2 &\leq Am^{s-s_0+2} \|c\|_{\ell^\infty}^2 \sum_{a \in \mathbb{Z}^d} \langle a \rangle^{2s_0} |(f^{m-1})_a|^2 \\ &= Am^{s-s_0+2} \|c\|_{\ell^\infty}^2 \|f^{m-1}\|_{H^{s_0}}^2 \leq AC^m \|c\|_{\ell^\infty}^2 \|f\|_{H^{s_0}}^{2m-2}, \end{aligned}$$

where  $\|g\|_{H^{s_0}}$ ,  $s_0 > d/2$  is the usual Sobolev norm on  $\mathbb{T}^d$  equivalent to the  $\ell_{s_0}^2$  norm of the Fourier coefficients of  $g$ . Here  $C$  depends on  $s$  and  $s_0$  but not on  $m$ . We then note that  $\|f\|_{H^{s_0}} = \|z_{\leq N}\|_s$  which shows the first equation in (36).

To prove the second line in the case  $n \geq 1$ , we simply bound  $\langle \ell \rangle^{s-s_0}$  by  $N^{s-s_0}$  and we obtain

$$P_\ell(y) \leq N^{s-s_0} \sum_{\substack{\mathcal{M}(j,\ell)=0 \\ \mu_n(j) > N}} |z_{j_1}| \cdots |z_{j_{m-1}}| = N^{s-s_0} \widehat{(f^{m-n} g^n)}_{-\delta a}$$

for  $\ell = -\delta a$  with the same notation as before, where

$$f(x) = \sum_{\langle \ell \rangle \leq N} |z_\ell| e^{i\delta a \cdot x} \quad \text{and} \quad g(x) = \sum_{\langle \ell \rangle > N} |z_\ell| e^{i\delta a \cdot x},$$

and we conclude as in the previous case.

To show (37), we use

$$\|X_{P[c]}(z)_{> N}\|_{s_0}^2 \leq m^2 \|c\|_{\ell^\infty}^2 \sum_{\langle \ell \rangle > N} \langle \ell \rangle^{2s_0} P_\ell(z)^2,$$

where

$$P_\ell(z) = \sum_{\substack{\mathcal{M}(j)=0 \\ \mu_n(j,\ell) > N}} |z_{j_1}| \cdots |z_{j_{m-1}}|. \tag{39}$$

In the case  $n = 1$ , we have no high mode in the sum. However, due to the zero-momentum condition, there exists one mode greater than  $N/(2m)$ ; hence, with the same notation as before, we have

$$P_\ell(z) \leq AN^{s_0-s} \sum_{\substack{\mathcal{M}(j)=0 \\ \mu_n(j,\ell) > N}} |\tilde{z}_{j_1}| \cdots |\tilde{z}_{j_{m-1}}|,$$

and we conclude as before. The proof of the last estimate in (37) is similar as the proof of (36) for  $n \geq 1$ . □

*Proof of Theorem 3.* Let  $\varepsilon > 0, s \geq s_0 > d/2$  and  $r$  be given, and  $z(0) \in \ell_{2s}^2$  such that  $\|z(0)\|_{2s} \leq \varepsilon$ . We apply the normal form Theorem (5) at the order  $3r$  and set

$$N = N_\varepsilon = \varepsilon^{-\frac{r}{s-s_0}}.$$

Under the hypothesis (18), we have the bound

$$N^{3r\alpha(3r)+d} \leq \varepsilon^{-1/2}. \tag{40}$$

In particular, all the coefficients of the polynomial Hamiltonian  $\chi, Z_m^o, S_m^{(ii)}$  and  $\tilde{P}_m^{(iii)}$  in (32) are bounded by  $CN^{3r\alpha(3r)} \leq C\varepsilon^{-\frac{1}{2}}$ . Hence, using Lemma 1, we have for  $\varepsilon$  small enough (in a way depending on  $r$  and  $s$ ) that  $y(0) = \Phi_\chi^{-1}(z(0))$  is in  $\ell_{2s}^2$  and satisfies  $\|y(0)\|_{2s} \leq \frac{5}{4}\varepsilon$ .

This implies in particular that  $\|y(0)\|_{\leq N} \leq \frac{5}{4}\varepsilon$  and

$$\begin{aligned} \|y(0)_{>N}\|_{s_0}^2 &= \sum_{\langle j \rangle > N} \langle j \rangle^{2s_0} |y_j|^2 \leq N^{2s_0-4s} \sum_{\langle j \rangle > N} \langle j \rangle^{4s} |y_j|^2 \\ &\leq 4\varepsilon^2 N^{4(s_0-s)} \leq 4\varepsilon^{4r+2}. \end{aligned}$$

Now we need to control the dynamics of  $y(t)$ , the solution of the Hamiltonian system associated with the Hamiltonian  $H \circ \Phi_\chi^1$  given by (32). We define

$$T := \sup\{t > 0 \mid \|y(t)\|_{\leq N} \leq \frac{3}{2}\varepsilon \text{ and } \|y(t)_{>N}\|_{s_0} \leq \varepsilon^{r+1}\}. \tag{41}$$

We note that  $\|y(0)\|_{2s} \leq \frac{5}{4}\varepsilon < \frac{3}{2}\varepsilon$  and  $\|y(0)_{>N}\|_{s_0} \leq 2\varepsilon^{2r+1} < \varepsilon^{r+1}$ . Therefore, by classical results for the definition of mild solutions of semilinear problems in Sobolev spaces with index greater than  $d/2$ ,  $T$  is strictly positive.

Let us prove that if  $t \leq \min(T, \varepsilon^{-\frac{r}{s_0+1}})$ , then  $\|y(t)\|_{\leq N} \leq \frac{11}{8}\varepsilon$  and  $\|y(t)_{>N}\|_{s_0} \leq \frac{1}{2}\varepsilon^{r+1}$ . We will then conclude by a continuity argument that  $T \geq \varepsilon^{-\frac{r}{s_0+1}}$ .

*Control of the transformation.* In view of Lemma 2, under bootstrap hypothesis, the vector field  $X_\chi(y(t))$  satisfies the estimates

$$\|X_\chi(y)_{\leq N}\|_s \leq CN^{3r\alpha(3r)}(\varepsilon^2 + N^{s-s_0}\varepsilon^{r+1}) \leq C\varepsilon^{\frac{3}{2}}$$

where we used  $N^{s-s_0} = \varepsilon^{-r}$ , and

$$\|X_\chi(y)_{>N}\|_{s_0} \leq CN^{3r\alpha(3r)}(N^{s_0-s}\varepsilon^2 + \varepsilon^{r+1}\varepsilon) \leq C\varepsilon^{r+\frac{3}{2}}.$$

Hence, this shows that as soon as  $y$  satisfies the bootstrap hypothesis, that is,

$$\|y(t)_{\leq N}\|_s \leq \frac{3}{2}\varepsilon \quad \text{and} \quad \|y(t)_{>N}\|_{s_0} \leq \varepsilon^{r+1}, \quad (42)$$

we have for  $\varepsilon$  small enough and for all  $w \in (0, 1)$ ,

$$\|\Phi_\chi^w(y)_{\leq N}\|_s \leq 2\varepsilon \quad \text{and} \quad \|\Phi_\chi^w(y)_{>N}\|_{s_0} \leq 2\varepsilon^{r+1}. \quad (43)$$

Now let us write the Hamiltonian (32) as  $H \circ \Phi_\chi^1 = H_2 + Z^{(o)} + S^{(ii)} + \tilde{P}^{(iii)} + \tilde{R}_{3r+1}$  by gathering together the terms with different homogeneity.

*Control of the low modes*  $y(t)_{\leq N}$ . For  $j \in \mathbb{U}_2 \times \mathbb{Z}^d$ , let  $J_j(t) = J_j(y(t))$  with the definition (34). As  $Z^{(o)}$  and  $H_2$  commute with  $J_a$  for all  $a \in \mathbb{Z}^d$ , we have for  $t \leq \min(T, \varepsilon^{-\frac{r}{s_0+1}})$  and  $\langle j \rangle \leq N$ ,

$$\begin{aligned} \langle j \rangle^{2s} |\dot{J}_j| &= \langle j \rangle^{2s} |\{J_j, H \circ \Phi_\chi^1\}| \\ &\leq \sum_{\langle \ell \rangle = \langle j \rangle} \langle \ell \rangle^s \left| \frac{\partial(S^{(ii)} + \tilde{P}^{(iii)} + \tilde{R}_{3r+1})}{\partial y_{\bar{\ell}}}(z) \right| \langle \ell \rangle^s |y_{\bar{\ell}}|. \end{aligned}$$

By summing in  $\langle j \rangle \leq N$  and using the Cauchy–Schwarz inequality, we obtain

$$\sum_{\langle j \rangle \leq N} \langle j \rangle^{2s} |\dot{J}_j| \leq \|X_{S^{(ii)} + \tilde{P}^{(iii)} + \tilde{R}_{3r+1}}(y)_{\leq N}\|_s \|y_{\leq N}\|_s.$$

Now under bootstrap hypothesis (42), we conclude using Lemma 2

$$\|X_{S^{(ii)} + \tilde{P}^{(iii)}}(y)_{\leq N}\|_s \leq CN^{3r\alpha(3r)}N^{s-s_0}\varepsilon^{2r+2} \leq C\varepsilon^{r+\frac{3}{2}}.$$

On the other hand, since  $R_{3r+1}$  is a Hamiltonian in  $\mathcal{H}_s$  of order at least  $3r + 1$ , we have using (25), (33) and Lemma 1,

$$\|X_{\tilde{R}_{3r+1}}(y)_{\leq N}\|_s \leq CN^{3r\alpha(3r)}\varepsilon^{3r}.$$

Hence, we obtain

$$\sum_{\langle j \rangle \leq N} \langle j \rangle^{2s} |\dot{J}_j| \leq C(\varepsilon^{3r+1} + \varepsilon^{r+\frac{5}{2}}).$$

Therefore, for  $t \leq \min(T, \varepsilon^{-r+1})$ , we conclude

$$\sum_{\langle j \rangle \leq N} \langle j \rangle^{2s} |J_j(t) - J_j(0)| \leq \varepsilon^{\frac{7}{2}}$$

for  $\varepsilon$  small enough, which in turn implies that

$$\|y(t)_{\leq N}\|_s \leq \frac{11}{8}\varepsilon. \tag{44}$$

*Control of the high modes.* The Hamiltonian  $S^{(ii)}$  can be written as

$$S_m^{(ii)} = \sum_{\substack{a, b \in \mathbb{Z}^d \\ \langle a \rangle > N, \langle b \rangle > N}} B_{ab}(y_{\leq N}) \xi_a \bar{\xi}_b, \quad y = (\xi, \bar{\xi}).$$

As the Hamiltonian is real, we have  $\overline{B_{ab}(y_{\leq N})} = B_{ba}(y_{\leq N})$ , that is, the operator  $A = (B_{ab}(y_{\leq N}))_{\langle a \rangle, \langle b \rangle > N}$  acting on  $\ell_s^2(\mathbb{Z}_{>N}^d)$  is Hermitian. Moreover, we have

$$B_{ab}(y_{\leq N}) = \sum_{m=1}^{2r-2} \sum_{\substack{j \in \mathcal{M}_m \\ \mu_1(j) \leq N \\ \mathcal{M}(j) = a-b}} b_{ab,j} y_j, \tag{45}$$

where the coefficients  $b_{ab,j}$  are uniformly bounded by  $CN^{3r\alpha(3r)}$ . Hence, for  $s > d/2$  and as soon as  $y$  satisfies the bootstrap assumption (42), we obtain that

$$\forall \langle a \rangle, \langle b \rangle > N \quad |B_{ab}(y_{\leq N})| \leq CN^{3r\alpha(3r)} \|y_{\leq N}\|_s. \tag{46}$$

When writing the dynamics of  $y = (\xi, \bar{\xi})$ , we get

$$\dot{\xi}_a = -i\omega_a \xi_a - i \sum_{b \in \mathbb{Z}^d} B_{ab}(y_{\leq N}) \xi_b - iQ_a(y), \quad \langle a \rangle > N$$

where  $Q_a = \frac{\partial}{\partial \xi_a} (\tilde{P}^{(iii)} + \tilde{R}_{3r+1})$ . Using (37), we have

$$\|Q(y)_{>N}\|_{s_0} \leq CN^{3r\alpha(3r)} (\varepsilon^{2r+2} + \varepsilon^{3r}). \tag{47}$$

So using the fact that the operator  $A$  is Hermitian, we get for  $t \leq T$  (recall that we assume that  $\|y(0)_{>N}\|_0 \leq \varepsilon^{2r+1}$ ),

$$\|y(t)_{>N}\|_0 = 2\|\xi(t)_{>N}\|_0 \leq 2\|\xi(0)_{>N}\|_0 + tC\varepsilon^{2r+\frac{3}{2}} \leq C(1+t)\varepsilon^{2r+1}. \tag{48}$$

Let us define  $D^{s_0}$  the diagonal operator from  $\ell_s^2(\mathbb{Z}_{>N}^d)$  into  $\ell_{s-2s_0}^2(\mathbb{Z}_{>N}^d)$  given by

$$D^{s_0} = \text{diag}(\langle a \rangle^{2s_0}, \langle a \rangle > N).$$

With the notation  $(\eta, \xi)_{>N} = \sum_{\langle a \rangle > N} \bar{\eta}_a \xi_a$ , we have

$$\frac{d}{dt} (\xi, D^{s_0} \xi)_{>N} = -i(\xi, [D^{s_0}, A(t)]\xi)_{>N} + \text{Im}(Q(y), D^{s_0} \xi)_{>N},$$

where  $[A, B]$  denotes the commutator of  $A$  and  $B$ :  $[A, B] = AB - BA$ . Hence, by bootstrap hypothesis (and using (47)),

$$\left| \frac{d}{dt} \langle \xi, D^{s_0} \xi \rangle_{>N} \right| \leq |\langle \xi, [D^{s_0}, A(t)] \xi \rangle_{>N}| + C \varepsilon^{3r + \frac{5}{2}}.$$

Then we note that using the zero-momentum condition (45), we have  $B_{ab}(y) = 0$  when  $|a - b| > 3rN$ . Hence, since  $s_0 > 1/2$ , we have

$$\begin{aligned} |\langle \xi, [D^{s_0}, A(t)] \xi \rangle_{>N}| &= \left| \sum_{\substack{\langle a \rangle, \langle b \rangle \geq N \\ |a-b| \leq 3rN}} B_{ab}(y) (\langle a \rangle^{2s_0} - \langle b \rangle^{2s_0}) \bar{\xi}_a \xi_b \right| \\ &\leq \sum_{\substack{\langle a \rangle, \langle b \rangle \geq N \\ |a-b| \leq 3rN}} 2s_0 |B_{ab}(y)| |\langle a \rangle - \langle b \rangle| (\langle a \rangle^{2s_0-1} + \langle b \rangle^{2s_0-1}) |\bar{\xi}_a \xi_b|. \end{aligned}$$

Using the bound (46) and the fact that  $|\langle a \rangle - \langle b \rangle| \leq \langle a - b \rangle \leq CN$ , we get

$$|\langle \xi, [D^{s_0}, A] \xi \rangle| \leq CN^{3r\alpha(3r)} \varepsilon \sum_{\substack{\langle a \rangle, \langle b \rangle \geq N \\ |a-b| \leq 3rN}} \langle a \rangle^{s_0-1} \langle b \rangle^{s_0} |\bar{\xi}_a \xi_b|, \tag{49}$$

where we used that for  $|a - b| \leq 3rN$ , we have  $\langle a \rangle \leq \langle b \rangle + 3rN$ . Now we apply the Cauchy–Schwarz inequality to get

$$\begin{aligned} &\sum_{\substack{\langle a \rangle, \langle b \rangle \geq N \\ |a-b| \leq 3rN}} \langle a \rangle^{s_0-1} \langle b \rangle^{s_0} |\bar{\xi}_a \xi_b| \\ &\leq \left( \sum_{\substack{\langle a \rangle, \langle b \rangle \geq N \\ |a-b| \leq 3rN}} \langle a \rangle^{2s_0-2} |\bar{\xi}_a|^2 \right)^{\frac{1}{2}} \left( \sum_{\substack{\langle a \rangle, \langle b \rangle \geq N \\ |a-b| \leq 3rN}} \langle b \rangle^{2s_0} |\xi_b|^2 \right)^{\frac{1}{2}} \\ &\leq CN^d \|y_{>N}\|_{s_0-1} \|y_{>N}\|_{s_0} \end{aligned}$$

where we used again that for  $|a - b| \leq 3rN$ , we have  $\langle a \rangle \leq \langle b \rangle + 3rN$  with  $a \in \mathbb{Z}^d$ . Thus, using (40), we obtain

$$\frac{d}{dt} \|y(t)_{>N}\|_{s_0}^2 \leq C \|y(t)_{>N}\|_{s_0-1} \|y(t)_{>N}\|_{s_0} + C \varepsilon^{3r + \frac{5}{2}}. \tag{50}$$

Then, since  $s_0 > 1$ , applying the Hölder inequality in (50), we get

$$\frac{d}{dt} \|y(t)_{>N}\|_{s_0}^2 \leq C \|y(t)_{>N}\|_0^{1/s_0} \|y(t)_{>N}\|_{s_0}^{2(1-1/2s_0)} + C \varepsilon^{3r + \frac{5}{2}}. \tag{51}$$

Using, in (51), the bound on  $\|y(t)_{>N}\|_0$  obtained in (48), we get

$$\frac{d}{dt} \|y(t)_{>N}\|_{s_0}^2 \leq C(1+t)^{1/s_0} \varepsilon^{\frac{2r+1}{s_0}} \|y(t)_{>N}\|_{s_0}^{2(1-\frac{1}{2s_0})} + C\varepsilon^{3r+\frac{5}{2}}. \quad (52)$$

We are going to apply the following elementary lemma.

LEMMA 3. Let  $\alpha \in (0, 1)$ ,  $f : \mathbb{R} \rightarrow \mathbb{R}_+$  be a continuous function and  $x : \mathbb{R} \rightarrow \mathbb{R}_+$  be a differentiable function satisfying the inequality

$$\forall t \in \mathbb{R}, \quad \frac{d}{dt} x(t) \leq \frac{1}{1-\alpha} f(t)(x(t))^\alpha.$$

Then we have the estimate

$$\forall t \in \mathbb{R}, \quad x(t)^{1-\alpha} \leq x(0)^{1-\alpha} + \int_0^t f(s) ds.$$

Considering (52), we are going to apply Lemma 3 with  $\alpha = 1 - 1/2s_0$  and  $x(t) = \|y(t)_{>N}\|_{s_0}^2 + \varepsilon^{\iota(s_0, r)}$  where  $\iota(s_0, r) = (3r + \frac{5}{2} - \frac{2r+1}{s_0})(1 - \frac{1}{2s_0})^{-1}$ . Indeed,  $x$  naturally satisfies the estimate

$$\frac{d}{dt} x(t) \leq C(1+t)^{\frac{1}{s_0}} \varepsilon^{\frac{2r+1}{s_0}} x(t)^{1-\frac{1}{2s_0}}.$$

Now applying Lemma (3), we get

$$x(t) \leq Cx(0) + C\varepsilon^{2(2r+1)}(1+t)^{2(s_0+1)}.$$

Since we have assumed that  $\|y(0)_{>N}\|_{s_0} \leq \varepsilon^{2r+1}$  and since a straightforward estimate proves that  $\iota(s_0, r) \geq 2r + 1$ , we deduce that

$$\|y(t)_{>N}\|_{s_0} \leq C\varepsilon^{2r+1}(1+t)^{s_0+1},$$

and, hence, for  $t \leq \min(T, \varepsilon^{-\frac{r}{(s_0+1)}})$  and  $\varepsilon$  small enough,

$$\|y(t)_{>N}\|_{s_0} \leq \frac{1}{2} \varepsilon^{r+1}. \quad (53)$$

Hence, combining (44) and (53), we conclude by continuity argument that  $T \geq \varepsilon^{-\frac{r}{(s_0+1)}}$  in (41). To finish the proof, we consider  $\tilde{z}(t) = \Phi_\chi^1(y(t))$ . By using Lemma 1, this solution is well defined in  $\ell_{s_0}^2$  and using the estimates of Lemma 2 for the vector field  $X_\chi$ , we deduce that  $\tilde{z}(t)$  satisfies the estimate (19). We conclude by noting that  $\tilde{z}(t) = z(t)$  is the unique mild solution of the partial differential equation associated with the Hamiltonian  $H$ .  $\square$



## Acknowledgements

During the preparation of this work, the three authors benefited from the support of the Centre Henri Lebesgue ANR-11-LABX-0020-01 and BG was supported by ANR-15-CE40-0001-02 'BEKAM' and ANR-16-CE40-0013 'ISDEEC' of the Agence Nationale de la Recherche.

## Conflict of Interest:

## References

- [BFG] J. Bernier, E. Faou and B. Grébert, 'Rational normal forms and stability of small solutions to nonlinear Schrödinger equations'. Preprint, 2018, [arXiv:1812.11414](https://arxiv.org/abs/1812.11414).
- [Bam03] D. Bambusi, 'Birkhoff normal form for some nonlinear PDEs', *Comm. Math. Phys.* **234** (2003), 253–283.
- [Bam07] D. Bambusi, 'A Birkhoff normal form theorem for some semilinear PDEs', in *Hamiltonian Dynamical Systems and Applications* (Springer, 2007), 213–247.
- [BDGS07] D. Bambusi, J.-M. Delort, B. Grébert and J. Szeftel, 'Almost global existence for Hamiltonian semilinear Klein–Gordon equations with small Cauchy data on Zoll manifolds', *Comm. Pure Appl. Math.* **60**(11) (2007), 1665–1690.
- [BG03] D. Bambusi and B. Grébert, 'Forme normale pour NLS en dimension quelconque', *C. R. Acad. Sci. Paris, Ser. I* **337** (2003), 409–414.
- [BG06] D. Bambusi and B. Grébert, 'Birkhoff normal form for PDE's with tame modulus', *Duke Math. J.* **135**(3) (2006), 507–567.
- [Bou95] J. Bourgain, 'Construction of periodic solutions of nonlinear wave equations in higher dimension', *Geom. Funct. Anal.* **5** (1995), 629–639.
- [Bou96] J. Bourgain, 'Construction of approximative and almost-periodic solutions of perturbed linear Schrödinger and wave equations', *Geom. Funct. Anal.* **6** (1996), 201–230.
- [Bou97] J. Bourgain, 'On Melnikov's persistency problem', *Math. Res. Lett.* **4** (1997), 445–458.
- [Caz03] T. Cazenave, *Semilinear Schrödinger Equations*, Courant Lecture Notes in Mathematics, 10 (New York University, Courant Institute of Mathematical Sciences, New York; American Mathematical Society, Providence, RI, 2003).
- [CHL08] D. Cohen, E. Hairer and C. Lubich, 'Conservation of energy, momentum and actions in numerical discretizations of nonlinear wave equations', *Numer. Math.* **110** (2008), 113–143.
- [DF09] E. Faou and A. Debussche, 'Modified energy for split-step methods applied to the linear Schrödinger equation', *SIAM J. Numer. Anal.* **47** (2009), 3705–3719.
- [Del09] J.-M. Delort, 'On long time existence for small solutions of semi-linear Klein–Gordon equations on the torus', *J. Anal. Math.* **107** 161–194.
- [DS06] J.-M. Delort and J. Szeftel, 'Long-time existence for semi-linear Klein–Gordon equations with small Cauchy data on Zoll manifolds', *Amer. J. Math.* **128**(5) (2006), 1187–1218.
- [DF07] G. Dujardin and E. Faou, 'Normal form and long time analysis of splitting schemes for the linear Schrödinger equation with small potential', *Numer. Math.* **106** (2007), 223–262.

- [Eli88] L. H. Eliasson, ‘Perturbations of stable invariant tori for Hamiltonian systems’, *Ann. Sc. Norm. Super. Pisa Cl. Sci.* **15** (1988), 115–147.
- [EGK16] L. H. Eliasson, B. Grébert and S. B. Kuksin, ‘KAM for non-linear beam equation’, *Geom. Funct. Anal.* **26** (2016), 1588–1715.
- [Fao12] E. Faou, *Geometric Numerical Integration and Schrödinger Equations*, Zurich Lectures in Advanced Mathematics (European Mathematical Society, 2012), 146.
- [FG13] E. Faou and B. Grébert, ‘A Nekhoroshev-type theorem for the nonlinear Schrödinger equation on the torus’, *Anal. PDE* **6** (2013), 1243–1262.
- [FGP10a] E. Faou, B. Grébert and E. Patrel, ‘Birkhoff normal form for splitting methods applied to semi linear Hamiltonian PDEs. Part I: finite dimensional discretization’, *Numer. Math.* **114** (2010), 429–458.
- [FGP10b] E. Faou, B. Grébert and E. Patrel, ‘Birkhoff normal form for splitting methods applied to semi linear Hamiltonian PDEs. Part II: abstract splitting’, *Numer. Math.* **114** (2010), 459–490.
- [FGL13] E. Faou, L. Gauckler and C. Lubich, ‘Sobolev stability of plane wave solutions to the cubic nonlinear Schrödinger equation on a torus’, *Comm. PDE* **38** (2013), 1123–1140.
- [FGL14] E. Faou, L. Gauckler and C. Lubich, ‘Plane wave stability of the split-step Fourier method for the nonlinear Schrödinger equation’, *Forum Math. Sigma* **2** (2014), 45.
- [Gre07] B. Grébert, ‘Birkhoff normal form and Hamiltonian PDEs’, *Sémin. Congr.* **15** (2007), 1–46.
- [GIP09] B. Grébert, R. Imekraz and É. Patrel, ‘Normal forms for semilinear quantum harmonic oscillators’, *Commun. Math. Phys.* **291** (2009), 763–798.
- [HLW06] E. Hairer, C. Lubich and G. Wanner, *Geometric Numerical Integration, Structure-Preserving Algorithms for Ordinary Differential Equations*, (Springer, 2006).
- [HS17] M. Hofmanová and K. Schratz, ‘An exponential-type integrator for the KdV equation’, *Numer. Math.* **136** (2017), 1117–1137.
- [Kuk93] S. B. Kuksin, *Nearly Integrable Infinite Dimensional Hamiltonian Systems*, Lecture Notes in Mathematics, 1556 (Springer, Berlin, 1993).
- [Mel65] V. K. Mel’nikov, ‘On some cases of conservation of conditionally periodic motions under a small change of the Hamiltonian function’, *Soviet Math. Dokl.* **6** (1965), 1592–1596.
- [Mel68] V. K. Mel’nikov, ‘A family of conditionally periodic solutions of a Hamiltonian system’, *Soviet Math. Dokl.* **9**(4) (1968), 882–885.
- [Mos67] J. Moser, ‘Convergent series expansions for quasiperiodic motions’, *Math. Ann.* **169** (1967), 136–176.
- [OS18] A. Ostermann and K. Schratz, ‘Low regularity exponential-type integrators for semilinear Schrödinger equations’, *Found. Comput. Math.* **18** (2018), 731–755.
- [Pös96] J. Pöschel, ‘A KAM-theorem for some nonlinear partial differential equations’, *Ann. Sc. Norm. Super. Pisa Cl. Sci.* **23** (1996), 119–148.
- [Sha00] Z. Shang, ‘Resonant and Diophantine step sizes in computing invariant tori of Hamiltonian systems’, *Nonlinearity* **13** (2000), 299–308.
- [XY01] J. Xu and J. You, ‘Persistence of lower-dimensional tori under the first Melnikov’s non-resonance condition’, *J. Math. Pures Appl.* **80** (2001), 1047–1067.