




# Long-Time Existence for Semi-linear Beam Equations on Irrational Tori

Joackim Bernier<sup>1</sup> · Roberto Feola<sup>1</sup> · Benoît Grébert<sup>1</sup>  · Felice Iandoli<sup>2</sup>

Received: 31 October 2020 / Revised: 18 January 2021 / Accepted: 30 January 2021 /  
Published online: 1 March 2021

© The Author(s), under exclusive licence to Springer Science+Business Media, LLC part of Springer Nature 2021

## Abstract

We consider the semi-linear beam equation on the  $d$  dimensional irrational torus with smooth nonlinearity of order  $n - 1$  with  $n \geq 3$  and  $d \geq 2$ . If  $\varepsilon \ll 1$  is the size of the initial datum, we prove that the lifespan  $T_\varepsilon$  of solutions is  $O(\varepsilon^{-A(n-2)^-})$  where  $A \equiv A(d, n) = 1 + \frac{3}{d-1}$  when  $n$  is even and  $A = 1 + \frac{3}{d-1} + \max(\frac{4-d}{d-1}, 0)$  when  $n$  is odd. For instance for  $d = 2$  and  $n = 3$  (quadratic nonlinearity) we obtain  $T_\varepsilon = O(\varepsilon^{-6^-})$ , much better than  $O(\varepsilon^{-1})$ , the time given by the local existence theory. The irrationality of the torus makes the set of differences between two eigenvalues of  $\sqrt{\Delta^2 + 1}$  accumulate to zero, facilitating the exchange between the high Fourier modes and complicating the control of the solutions over long times. Our result is obtained by combining a Birkhoff normal form step and a modified energy step.

**Keywords** Lifespan for semi-linear PDEs · Birkhoff normal forms · Modified energy · Irrational torus

---

Communicated by Yingfei Yi.

---

In memory of Walter Craig whose beautiful voice, always relevant and friendly, we miss.

---

Felice Iandoli has been supported by ERC grant ANADEL 757996. Roberto Feola, Joackim Bernier and Benoît Grébert have been supported by the Centre Henri Lebesgue ANR-11-LABX-0020-01 and by ANR-15-CE40-0001-02 “BEKAM” of the ANR.

---

✉ Benoît Grébert  
benoit.grebert@univ-nantes.fr

Joackim Bernier  
joackim.bernier@univ-nantes.fr

Roberto Feola  
roberto.feola@univ-nantes.fr

Felice Iandoli  
felice.iandoli@sorbonne-universite.fr

<sup>1</sup> Laboratoire de Mathématiques Jean Leray, Université de Nantes, UMR CNRS 6629, 2, rue de la Houssinière, 44322 Nantes Cedex 03, France

<sup>2</sup> Laboratoire Jacques-Louis Lions, Sorbonne Université, UMR CNRS 7598, 4, Place Jussieu, 75005 Paris Cedex 05, France

**Mathematics Subject Classification** 35Q35 · 35Q53 · 37K55

### 1 Introduction

In this article we consider the beam equation on an irrational torus

$$\begin{cases} \partial_{tt}\psi + \Delta^2\psi + \psi + f(\psi) = 0, \\ \psi(0, y) = \psi_0, \\ \partial_t\psi(0, y) = \psi_1, \end{cases} \tag{1.1}$$

where  $f \in C^\infty(\mathbb{R}, \mathbb{R})$ ,  $\psi = \psi(t, y)$ ,  $y \in \mathbb{T}_\nu^d$ , with  $\nu = (\nu_1, \dots, \nu_d) \in [1, 2]^d$  and

$$\mathbb{T}_\nu^d := (\mathbb{R}/2\pi\nu_1\mathbb{Z}) \times \dots \times (\mathbb{R}/2\pi\nu_d\mathbb{Z}). \tag{1.2}$$

The initial data  $(\psi_0, \psi_1)$  have small size  $\varepsilon$  in the standard Sobolev space  $H^{s+1}(\mathbb{T}_\nu^d) \times H^{s-1}(\mathbb{T}_\nu^d)$  for some  $s \gg 1$ . The nonlinearity  $f(\psi)$  has the form

$$f(\psi) := (\partial_\psi F)(\psi) \tag{1.3}$$

for some smooth function  $F \in C^\infty(\mathbb{R}, \mathbb{R})$  having a zero of order at least  $n \geq 3$  at the origin. Local existence theory implies that (1.1) admits, for small  $\varepsilon > 0$ , a unique smooth solution defined on an interval of length  $O(\varepsilon^{-n+2})$ . Our goal is to prove that, generically with respect to the irrationality of the torus (i.e. generically with respect to the parameter  $\nu$ ), the solution actually extends to a larger interval.

Our main theorem is the following.

**Theorem 1** *Let  $d \geq 2$ . There exists  $s_0 \equiv s_0(n, d) \in \mathbb{R}$  such that for almost all  $\nu \in [1, 2]^d$ , for any  $\delta > 0$  and for any  $s \geq s_0$  there exists  $\varepsilon_0 > 0$  such that for any  $0 < \varepsilon \leq \varepsilon_0$  we have the following. For any initial data  $(\psi_0, \psi_1) \in H^{s+1}(\mathbb{T}_\nu^d) \times H^{s-1}(\mathbb{T}_\nu^d)$  such that*

$$\|\psi_0\|_{H^{s+1}} + \|\psi_1\|_{H^{s-1}} \leq \varepsilon, \tag{1.4}$$

*there exists a unique solution of the Cauchy problem (1.1) such that*

$$\begin{aligned} \psi(t, x) \in C^0([0, T_\varepsilon]; H^{s+1}(\mathbb{T}_\nu^d)) \cap C^1([0, T_\varepsilon]; H^{s-1}(\mathbb{T}_\nu^d)), \\ \sup_{t \in [0, T_\varepsilon]} \left( \|\psi(t, \cdot)\|_{H^{s+1}} + \|\partial_t\psi(t, \cdot)\|_{H^{s-1}} \right) \leq 2\varepsilon, \quad T_\varepsilon \geq \varepsilon^{-a+\delta}, \end{aligned} \tag{1.5}$$

where  $a = a(d, n)$  has the form

$$a(d, n) := \begin{cases} (n-2) \left( 1 + \frac{3}{d-1} \right), & n \text{ even} \\ (n-2) \left( 1 + \frac{3}{d-1} \right) + \frac{\max\{4-d, 0\}}{d-1}, & n \text{ odd.} \end{cases} \tag{1.6}$$

Originally, the beam equation has been introduced in physics to model the oscillations of a uniform beam, so in a one dimensional context. In dimension 2, similar equations can be used to model the motion of a clamped plate (see for instance the introduction of [28]). In larger dimension ( $d \geq 3$ ) we do not claim that the beam Eq. (1.1) has a physical interpretation but nevertheless remains an interesting mathematical model of dispersive PDE. We note that when the equation is posed on a torus, there is no physical reason to assume the torus to be rational.

This problem of extending solutions of semi-linear PDEs beyond the time given by local existence theory has been considered many times in the past, starting with Bourgain [11],

Bambusi [1] and Bambusi–Grébert [2] in which the authors prove the almost global existence for the Klein Gordon equation:

$$\begin{cases} \partial_{tt}\psi - \Delta\psi + m\psi + f(\psi) = 0, \\ \psi(0, x) = \psi_0, \\ \partial_t\psi(0, x) = \psi_1, \end{cases} \quad (1.7)$$

on a one dimensional torus. Precisely, they proved that, given  $N \geq 1$ , if the initial datum has a size  $\varepsilon$  small enough in  $H^s(\mathbb{T}) \times H^{s-1}(\mathbb{T})$ , and if the mass stays outside an exceptional subset of zero measure, the solution of (1.7) exists at least on an interval of length  $O(\varepsilon^{-N})$ . This result has been extended to Eq. (1.7) on Zoll manifolds (in particular spheres) by Bambusi–Delort–Grébert–Szeftel [3] but also for the nonlinear Schrödinger equation posed on  $\mathbb{T}^d$  (the square torus of dimension  $d$ ) [2, 19] or on  $\mathbb{R}^d$  with a harmonic potential [24]. What all these examples have in common is that the spectrum of the linear part of the equation can be divided into clusters that are well separated from each other. Actually if you considered (1.1) with a generic mass  $m$  on the square torus  $\mathbb{T}^d$  then the spectrum of  $\sqrt{\Delta^2 + m}$  (the square root comes from the fact that the equation is of order two in time) is given by  $\{\sqrt{|j|^4 + m} \mid j \in \mathbb{Z}^d\}$  which can be divided in clusters around each integers  $n$  whose diameter decreases with  $|n|$ . Thus for  $n$  large enough these clusters are separated by  $1/2$ . So in this case also we could easily prove, following [2], the almost global existence of the solution.

On the contrary when the equation is posed on an irrational torus, the nature of the spectrum drastically changes: the differences between couples of eigenvalues accumulate to zero. Even for the Klein Gordon Eq. (1.7) posed on  $\mathbb{T}^d$  for  $d \geq 2$  the linear spectrum is not well separated. In both cases we could expect exchange of energy between high Fourier modes and thus the almost global existence in the sense described above is not reachable (at least up to now!). Nevertheless it is possible to go beyond the time given by the local existence theory. In the case of (1.7) on  $\mathbb{T}^d$  for  $d \geq 2$ , this local time has been extended by Delort [13] and then improved in different ways by Fang and Zhang [18], Zhang [29] and Feola et al. [20] (in this last case a quasi linear Klein Gordon equation is considered). We quote also the remarkable work on multidimensional periodic water wave by Ionescu and Pusateri [26].

The beam equation has already been considered on irrational torus in dimension 2 by Imekraz [25]. In the case he considered, the irrationality parameter  $\nu$  was diophantine and fixed, but a mass  $m$  was added in the game (for us  $m$  is fixed and for convenience we chose  $m = 1$ ). For almost all mass, Imekraz obtained a lifespan  $T_\varepsilon = O(\varepsilon^{-\frac{5}{4}(n-2)^+})$  while we obtain, for almost all  $\nu$ ,  $T_\varepsilon = O(\varepsilon^{-4(n-2)^+})$  when  $n$  is even and  $T_\varepsilon = O(\varepsilon^{-4(n-2)^-2^+})$  when  $n$  is odd.

We notice that applying the Theorem 3 of [6] (and its Corollary 1) we obtain the almost global existence for (1.1) on irrational tori up to a large but finite loss of derivatives.

Let us also mention some recent results about the longtime existence for periodic water waves [7–10]. In the same spirit we quote the long time existence for a general class of quasi-linear *Hamiltonian* equations [21] and quasi-linear *reversible* Schrödinger equations [22] on the circle. The main theorem in [21] applies also for quasi-linear perturbations of the beam equation. We mention also [16], here the authors study the lifespan of small solutions of the semi-linear Klein–Gordon equation posed on a general compact boundary-less Riemannian manifold.

All previous results [13, 18, 20, 25, 29] have been obtained by a modified energy procedure. Such procedure partially destroys the algebraic structure of the equation and, thus, it makes

more involved to iterate the procedure.<sup>1</sup> On the contrary, in this paper, we begin by a Birkhoff normal form procedure (when  $d = 2, 3$ ) before applying a modified energy step. Further in dimension 2 we can iterate two steps of Birkhoff normal form and therefore we get a much better time. The other key tool that allows us to go further in time is an estimate of small divisors that we have tried to optimize to the maximum: essentially small divisors make us lose  $(d - 1)$  derivatives (see Proposition 2.2) which explains the strong dependence of our result on the dimension  $d$  of the torus and also explains why we obtain a better result than [25]. In Sect. 1.2 we detail the scheme of the proof of Theorem 1.

### 1.1 Hamiltonian Formalism

We denote by  $H^s(\mathbb{T}^d; \mathbb{C})$  the usual Sobolev space of functions  $\mathbb{T}^d \ni x \mapsto u(x) \in \mathbb{C}$ . We expand a function  $u(x)$ ,  $x \in \mathbb{T}^d$ , in Fourier series as

$$u(x) = \frac{1}{(2\pi)^{d/2}} \sum_{n \in \mathbb{Z}^d} u_n e^{in \cdot x}, \quad u_n := \frac{1}{(2\pi)^{d/2}} \int_{\mathbb{T}^d} u(x) e^{-in \cdot x} dx. \tag{1.8}$$

We also use the notation

$$u_n^{+1} := u_n \quad \text{and} \quad u_n^{-1} := \overline{u_n}. \tag{1.9}$$

We set  $\langle j \rangle := \sqrt{1 + |j|^2}$  for  $j \in \mathbb{Z}^d$ . We endow  $H^s(\mathbb{T}^d; \mathbb{C})$  with the norm

$$\|u(\cdot)\|_{H^s}^2 := \sum_{j \in \mathbb{Z}^d} \langle j \rangle^{2s} |u_j|^2. \tag{1.10}$$

Moreover, for  $r \in \mathbb{R}^+$ , we denote by  $B_r(H^s(\mathbb{T}^d; \mathbb{C}))$  the ball of  $H^s(\mathbb{T}^d; \mathbb{C})$  with radius  $r$  centered at the origin. We shall also write the norm in (1.10) as  $\|u\|_{H^s}^2 = (\langle D \rangle^s u, \langle D \rangle^s u)_{L^2}$ , where  $\langle D \rangle e^{ij \cdot x} = \langle j \rangle e^{ij \cdot x}$ , for any  $j \in \mathbb{Z}^d$ .

In the following it will be more convenient to rescale the Eq. (1.1) and work on squared tori  $\mathbb{T}^d$ . For any  $y \in \mathbb{T}_v^d$  we write  $\psi(y) = \phi(x)$  with  $y = (x_1 v_1, \dots, x_d v_d)$  and  $x = (x_1, \dots, x_d) \in \mathbb{T}^d$ . The beam equation in (1.1) reads

$$\partial_{tt} \phi + \Omega^2 \phi + f(\phi) = 0 \tag{1.11}$$

where  $\Omega$  is the Fourier multiplier defined by linearity as

$$\Omega e^{ij \cdot x} = \omega_j e^{ij \cdot x}, \quad \omega_j := \sqrt{|j|_a^4 + 1}, \quad |j|_a^2 := \sum_{i=1}^d a_i |j_i|^2, \quad a_i := v_i^2, \quad \forall j \in \mathbb{Z}^d. \tag{1.12}$$

Introducing the variable  $v = \dot{\phi} = \partial_t \phi$  we can rewrite Eq. (1.11) as

$$\dot{\phi} = -v, \quad \dot{v} = \Omega^2 \phi + f(\phi). \tag{1.13}$$

By (1.3) we note that (1.13) can be written in the Hamiltonian form

$$\partial_t \begin{bmatrix} \phi \\ v \end{bmatrix} = X_{H_{\mathbb{R}}}(\phi, v) = J \begin{pmatrix} \partial_{\phi} H_{\mathbb{R}}(\phi, v) \\ \partial_v H_{\mathbb{R}}(\phi, v) \end{pmatrix}, \quad J = \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}$$

<sup>1</sup> Actually there are papers in which such procedure is iterated. We quote for instance [15] and reference therein.

where  $\partial$  denotes the  $L^2$ -gradient of the Hamiltonian function

$$H_{\mathbb{R}}(\phi, v) = \int_{\mathbb{T}^d} \left( \frac{1}{2}v^2 + \frac{1}{2}(\Omega^2\phi)\phi + F(\phi) \right) dx, \tag{1.14}$$

on the phase space  $H^2(\mathbb{T}^d; \mathbb{R}) \times L^2(\mathbb{T}^d; \mathbb{R})$ . Indeed we have

$$dH_{\mathbb{R}}(\phi, v) \left[ \begin{smallmatrix} \hat{\phi} \\ \hat{v} \end{smallmatrix} \right] = -\lambda_{\mathbb{R}}(X_{H_{\mathbb{R}}}(\phi, v), \left[ \begin{smallmatrix} \hat{\phi} \\ \hat{v} \end{smallmatrix} \right]) \tag{1.15}$$

for any  $(\phi, v), (\hat{\phi}, \hat{v})$  in  $H^2(\mathbb{T}^d; \mathbb{R}) \times L^2(\mathbb{T}^d; \mathbb{R})$ , where  $\lambda_{\mathbb{R}}$  is the non-degenerate symplectic form

$$\lambda_{\mathbb{R}}(W_1, W_2) := \int_{\mathbb{T}^d} (\phi_1 v_2 - v_1 \phi_2) dx, \quad W_1 := \begin{bmatrix} \phi_1 \\ v_1 \end{bmatrix}, W_2 := \begin{bmatrix} \phi_2 \\ v_2 \end{bmatrix}.$$

The Poisson bracket between two Hamiltonian  $H_{\mathbb{R}}, G_{\mathbb{R}} : H^2(\mathbb{T}^d; \mathbb{R}) \times L^2(\mathbb{T}^d; \mathbb{R}) \rightarrow \mathbb{R}$  are defined as

$$\{H_{\mathbb{R}}, G_{\mathbb{R}}\} = \lambda_{\mathbb{R}}(X_{H_{\mathbb{R}}}, X_{G_{\mathbb{R}}}). \tag{1.16}$$

We define the complex variables

$$\begin{bmatrix} u \\ \bar{u} \end{bmatrix} := \mathcal{C} \begin{bmatrix} \phi \\ v \end{bmatrix}, \quad \mathcal{C} := \frac{1}{\sqrt{2}} \begin{pmatrix} \Omega^{\frac{1}{2}} & i\Omega^{-\frac{1}{2}} \\ \Omega^{\frac{1}{2}} & -i\Omega^{-\frac{1}{2}} \end{pmatrix}, \tag{1.17}$$

where  $\Omega$  is the Fourier multiplier defined in (1.12). Then the system (1.13) reads

$$\dot{u} = i\Omega u + \frac{i}{\sqrt{2}}\Omega^{-1/2} f \left( \Omega^{-1/2} \begin{pmatrix} u + \bar{u} \\ \sqrt{2} \end{pmatrix} \right). \tag{1.18}$$

Notice that (1.18) can be written in the Hamiltonian form

$$\partial_t \begin{bmatrix} u \\ \bar{u} \end{bmatrix} = X_H(u) = iJ \begin{pmatrix} \partial_u H(u) \\ \partial_{\bar{u}} H(u) \end{pmatrix} = \begin{pmatrix} i\partial_{\bar{u}} H(u) \\ -i\partial_u H(u) \end{pmatrix}, \quad J = \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix} \tag{1.19}$$

with Hamiltonian function (see (1.14))

$$H(u) = H_{\mathbb{R}}(\mathcal{C}^{-1} \begin{bmatrix} u \\ \bar{u} \end{bmatrix}) = \int_{\mathbb{T}^d} \bar{u}\Omega u \, dx + \int_{\mathbb{T}^d} F \left( \frac{\Omega^{-1/2}(u + \bar{u})}{\sqrt{2}} \right) \, dx \tag{1.20}$$

and where  $\partial_{\bar{u}} = (\partial_{\Re u} + i\partial_{\Im u})/2, \partial_u = (\partial_{\Re u} - i\partial_{\Im u})/2$ . Notice that

$$X_H = \mathcal{C} \circ X_{H_{\mathbb{R}}} \circ \mathcal{C}^{-1} \tag{1.21}$$

and that (using (1.17))

$$dH(u) \left[ \begin{smallmatrix} h \\ \bar{h} \end{smallmatrix} \right] = (dH_{\mathbb{R}})(\phi, v) \left[ \mathcal{C}^{-1} \begin{bmatrix} h \\ \bar{h} \end{bmatrix} \right] \stackrel{(1.15), (1.21)}{=} -\lambda \left( X_H(u), \begin{bmatrix} h \\ \bar{h} \end{bmatrix} \right) \tag{1.22}$$

for any  $h \in H^2(\mathbb{T}^d; \mathbb{C})$  and where the two form  $\lambda$  is given by the push-forward  $\lambda = \lambda_{\mathbb{R}} \circ \mathcal{C}^{-1}$ . In complex variables the Poisson bracket in (1.16) reads

$$\{H, G\} := \lambda(X_H, X_G) = i \int_{\mathbb{T}^d} \partial_u G \partial_{\bar{u}} H - \partial_{\bar{u}} G \partial_u H \, dx, \tag{1.23}$$

where we set  $H = H_{\mathbb{R}} \circ \mathcal{C}^{-1}, G = G_{\mathbb{R}} \circ \mathcal{C}^{-1}$ . Let us introduce an additional notation:

**Definition 1.1** If  $j \in (\mathbb{Z}^d)^r$  for some  $r \geq k$  then  $\mu_k(j)$  denotes the  $k^{st}$  largest number among  $|j_1|, \dots, |j_r|$  (multiplicities being taken into account). If there is no ambiguity we denote it only with  $\mu_k$ .

Let  $r \in \mathbb{N}, r \geq n$ . A Taylor expansion of the Hamiltonian  $H$  in (1.20) leads to

$$H = Z_2 + \sum_{k=n}^{r-1} H_k + R_r \tag{1.24}$$

where

$$Z_2 := \int_{\mathbb{T}^d} \bar{u} \Omega u \, dx \stackrel{(1.12)}{=} \sum_{j \in \mathbb{Z}^d} \omega_j |u_j|^2 \tag{1.25}$$

and  $H_k, k = n, \dots, r - 1$ , is an homogeneous polynomial of order  $k$  of the form

$$H_k = \sum_{\substack{\sigma \in \{-1, 1\}^k, j \in (\mathbb{Z}^d)^k \\ \sum_{i=1}^k \sigma_i j_i = 0}} (H_k)_{\sigma, j} u_{j_1}^{\sigma_1} \cdots u_{j_k}^{\sigma_k} \tag{1.26}$$

with (noticing that the zero momentum condition  $\sum_{i=1}^k \sigma_i j_i = 0$  implies  $\mu_1(j) \lesssim \mu_2(j)$ )

$$|(H_k)_{\sigma, j}| \lesssim_k \frac{1}{\mu_1(j)^2}, \quad \forall \sigma \in \{-1, 1\}^k, j \in (\mathbb{Z}^d)^k \tag{1.27}$$

and

$$\|X_{R_r}(u)\|_{H^{s+2}} \lesssim_s \|u\|_{H^s}^{r-1}, \quad \forall u \in B_1(H^s(\mathbb{T}^d; \mathbb{C})). \tag{1.28}$$

The estimate above follows by Moser’s composition theorem in [27], section 2. Estimates (1.27) and (1.28) express the regularizing effect of the semi-linear nonlinearity in the Hamiltonian writing of (1.11).

### 1.2 Scheme of the Proof of Theorem 1

As usual Theorem 1 will be proved by a bootstrap argument and thus we want to control,  $N_s(u(t)) := \|u(t)\|_{H^s}^2$ , for  $t \mapsto u(t, \cdot)$  a small solution (whose local existence is given by the standard theory for semi-linear PDEs) of the Hamiltonian system generated by  $H$  given by (1.24) for the longest time possible (and at least longer than the existence time given by the local theory). So we want to control its derivative with respect to  $t$ . We have

$$\frac{d}{dt} N_s(u) = \{N_s, H\} = \sum_{k=n}^{r-1} \{N_s, H_k\} + \{N_s, R_r\}. \tag{1.29}$$

By (1.28) we have  $\{N_s, R_r\} \lesssim \|u\|_{H^s}^{r-1}$  and thus we can neglect this term choosing  $r$  large enough. Then we define  $H_k^{\leq N}$  the truncation of  $H_k$  at order  $N$ :

$$H_k^{\leq N} = \sum_{\substack{\sigma \in \{-1, 1\}^k, j \in (\mathbb{Z}^d)^k \\ \sum_{i=1}^k \sigma_i j_i = 0, \mu_2(j) \leq N}} (H_k)_{\sigma, j} u_{j_1}^{\sigma_1} \cdots u_{j_k}^{\sigma_k}$$

and we set  $H_k^{> N} = H_k - H_k^{\leq N}$ . As a consequence of (1.27) we have  $\{N_s, H_k^{> N}\} \lesssim N^{-2} \|u\|_{H^s}^{k-1}$  and thus we can neglect these terms choosing  $N$  large enough. So it remains to take care of  $\sum_{k=n}^{r-1} \{N_s, H_k^{\leq N}\}$ .

The natural idea to eliminate  $H_k^{\leq N}$  consists in using a Birkhoff normal form procedure (see [2,23]). In order to do that, we have first to solve the homological equation

$$\{\chi_k, Z_2\} + H_k^{\leq N} = Z_k.$$

This is achieved in Lemma 3.6 and, thanks to the control of the small divisors given by Proposition 2.2, we get that there exists  $\alpha \equiv \alpha(d, k) > 0$  such that for any  $\delta > 0$

$$|(\chi_k)_{\sigma,j}| \lesssim_{\delta} \mu_1(j)^{d-3+\delta} \mu_3(j)^{\alpha}, \quad \forall \sigma \in \{-1, 1\}^k, \quad j \in (\mathbb{Z}^d)^k. \tag{1.30}$$

From [2] we learn that the positive power of  $\mu_3(j)$  appearing in the right hand side of (1.30) is not dangerous<sup>2</sup> (taking  $s$  large enough) but the positive power of  $\mu_1(j)$  implies a loss of derivatives. So this step can be achieved only assuming  $d \leq 3$  and in that case the corresponding flow is well defined in  $H^s$  (with  $s$  large enough) and is controlled by  $N^\delta$  (see Lemma 3.7). In other words, this step is performed only when  $d = 2, 3$ , when  $d \geq 4$  we directly go to the modified energy step.

For  $d = 2, 3$ , let us focus on  $n = 3$ . After this Birkhoff normal form step, we are left with

$$H \circ \Phi_{\chi_3} = Z_2 + Z_3 + Q_4 + \text{negligible terms}$$

where  $Q_4$  is a Hamiltonian of order 4 whose coefficients are bounded by  $\mu_1(j)^{d-3+\delta}$  (see Lemma 3.5, estimate (3.15)) and  $Z_3$  is a Hamiltonian of order 3 which is resonant:  $\{Z_2, Z_3\} = 0$ . Actually, as consequence Proposition 2.2,  $Z_3 = 0$  and thus we have eliminated all the terms of order 3 in (1.29).

In the case  $d = 2$ ,  $Q_4^{\leq N}$  is still  $(1 - \delta)$ -regularizing and we can perform a second Birkhoff normal form. Actually, since in eliminating  $Q_4^{\leq N}$  we create terms of order at least 6, we can eliminate both  $Q_4^{\leq N}$  and  $Q_5^{\leq N}$ . So, for  $d = 2$ , we are left with

$$\tilde{H} = H \circ \Phi_{\chi_3} \circ \Phi_{\chi_4+\chi_5} = Z_2 + Z_4 + Q_6 + \text{negligible terms}$$

where  $Z_4$  is Hamiltonian of order 4 which is resonant,<sup>3</sup>  $\{Z_2, Z_4\} = 0$ , and  $Q_6$  is a Hamiltonian of order 6 whose coefficients are bounded by  $N^{2\delta}$ . Since resonant Hamiltonians commute with  $N_s$ , the first contribution in (1.29) is  $\{N_s, Q_6\}$ . This is essentially the statement of Theorem 2 (which will be stated in Sect. 3) in the case  $d = 2$  and  $n = 3$  and this achieves the Birkhoff normal forms step.

Let us describe the modified energy step only in the case  $d = 2$  and  $n = 3$  and let us focus on the worst term in  $\{N_s, \tilde{H}\}$ , i.e.  $\{N_s, Q_6\}$ . Let us write

$$Q_6 = \sum_{\substack{\sigma \in \{-1, 1\}^6, \quad j \in (\mathbb{Z}^d)^k \\ |j_1| \geq \dots \geq |j_6| \\ \sum_{i=1}^6 \sigma_i j_i = 0}} (Q_6)_{\sigma,j} u_{j_1}^{\sigma_1} \cdots u_{j_6}^{\sigma_6}.$$

<sup>2</sup> When you have a control of the small divisors involving only  $\mu_3(j)$  then you can solve the homological equation at any order and you obtain an almost global existence result in the spirit of [2]. This would be the case if we consider the semi-linear beam equation on the squared torus  $\mathbb{T}^d$ .

<sup>3</sup> Notice that there is no resonant term of odd order by Proposition 2.2, in other words  $Z_3 = Z_5 = 0$ .

From Proposition 2.2 we learn that if  $\sigma_1\sigma_2 = 1$  then the small divisor associated with  $(j, \sigma)$  is controlled by  $\mu_3(j)$  and thus we can eliminate the corresponding monomial by one more Birkhoff normal forms step.<sup>4</sup> Now if we assume  $\sigma_1\sigma_2 = -1$  we have

$$\begin{aligned} |\{N_s, u_{j_1}^{\sigma_1} \cdots u_{j_6}^{\sigma_6}\}| &= \left| \sum_{i=1}^6 \sigma_{j_i} \langle j_i \rangle^{2s} \right| |u_{j_1}^{\sigma_1} \cdots u_{j_6}^{\sigma_6}| \\ &\leq (\langle j_1 \rangle^{2s} - \langle j_2 \rangle^{2s} + 4\langle j_3 \rangle^{2s}) |u_{j_1}^{\sigma_1} \cdots u_{j_6}^{\sigma_6}| \\ &\leq (s(\langle j_1 \rangle^2 - \langle j_2 \rangle^2) \langle j_1 \rangle^{2(s-1)} + 4\langle j_3 \rangle^{2s}) |u_{j_1}^{\sigma_1} \cdots u_{j_6}^{\sigma_6}| \\ &\lesssim_s (\langle j_1 \rangle^{2s-1} \langle j_3 \rangle + 4\langle j_3 \rangle^{2s}) |u_{j_1}^{\sigma_1} \cdots u_{j_6}^{\sigma_6}| \\ &\lesssim_s \mu_1^{-1} \|u\|_{H^s}^6 \end{aligned}$$

where we used the zero momentum condition,  $\sum_{i=1}^6 \sigma_i j_i = 0$ , to obtain  $|j_1 - j_2| \leq 4|j_3|$ . This gain of one derivative, also known as the commutator trick, is central in a lot of results about modified energy [6,13] or growth of Sobolev norms [4,5,12,14].

So if  $Q_6^-$  denotes the restriction of  $Q_6$  to monomials satisfying  $\sigma_1\sigma_2 = -1$  we have essentially proved that

$$|\{N_s, Q_6^{-, > N_1}\}| \lesssim N_1^{-1} \|u\|_{H^s}^6.$$

Then we can consider the modified energy  $N_s + E_6$  with  $E_6$  solving

$$\{E_6, Z_2\} = -\{N_s, Q_6^{-, \leq N_1}\}$$

in such a way that

$$\{N_s + E_6, \tilde{H}\} = \{N_s, Q_6^{-, > N_1}\} + \{N_s, \tilde{H}_7\} + \{E_6, Z_4\} + \text{negligible terms}.$$

Since this modified energy will not produce new terms of order 7, we can in the same time eliminate  $Q_7^{-, \leq N_1}$ . Thus we obtain a new energy,  $N_s + E_6 + E_7$ , which is equivalent to  $N_s$  in a neighborhood of the origin, and such that, by neglecting all the powers of  $N^\delta$  and  $N_1^\delta$  which appear when we work carefully (see (4.6) for a precise estimate),

$$|\{N_s + E_6 + E_7, \tilde{H}\}| \lesssim_s N_1^{-1} \|u\|_{H^s}^6 + \|u\|_{H^s}^8 + N^{-1} \|u\|_{H^s}^3.$$

Then, a suitable choice of  $N$  and  $N_1$  and a standard bootstrap argument lead to,  $T_\varepsilon = O(\varepsilon^{-6})$  by using this rough estimate, and  $T_\varepsilon = O(\varepsilon^{-6^-})$  by using the precise estimate (see Sect. 5).

**Remark 1.2** In principle a Birkhoff normal form procedure gives more than just the control of  $H^s$  norm of the solutions, it gives an equivalent Hamiltonian system and therefore potentially more information about the dynamics of the solutions. However, if one wants to control only the solution in  $H^s$  norm, the modified energy method is sufficient and simpler. One could therefore imagine applying this last method from the beginning. However, when we iterate it, the modified energy method brings up terms that, when we apply a Birkhoff procedure, turn out to be zero. Unfortunately we have not been able to prove the cancellation of these terms directly by the modified energy method, that is why we use successively a Birkhoff normal form procedure and a modified energy procedure.

**Notation** We shall use the notation  $A \lesssim B$  to denote  $A \leq CB$  where  $C$  is a positive constant depending on parameters fixed once for all, for instance  $d, n$ . We will emphasize by writing  $\lesssim_q$  when the constant  $C$  depends on some other parameter  $q$ .

<sup>4</sup> In fact in Sect. 4, for the sake of simplicity, we prefer to apply a modified energy strategy to all the terms of  $Q_6$  (see also Remark 1.2).



## 2 Small Divisors

As already remarked in the introduction, the proof of Theorem 1 is based on a normal form approach. In particular we have to deal with a small divisors problem involving linear combination of linear frequencies  $\omega_j$  in (1.12).

This section is devoted to establish suitable lower bounds for *generic* (in a probabilistic way) choices of the parameters  $\nu$  excepted for exceptional indices for which the small divisor is identically zero. According to the following definition such indices are called *resonant*.

**Definition 2.1** (*Resonant indices*) Being given  $r \geq 3$ ,  $j_1, \dots, j_r \in \mathbb{Z}^d$  and  $\sigma_1, \dots, \sigma_r \in \{-1, 1\}$ , the couple  $(\sigma, j)$  is resonant if  $r$  is even and there exists a permutation  $\rho \in \mathfrak{S}_r$  such that

$$\forall k \in \llbracket 1, r/2 \rrbracket, \begin{pmatrix} |j_{\rho_{2k-1}, 1}| \\ \vdots \\ |j_{\rho_{2k-1}, d}| \end{pmatrix} = \begin{pmatrix} |j_{\rho_{2k}, 1}| \\ \vdots \\ |j_{\rho_{2k}, d}| \end{pmatrix} \quad \text{and} \quad \sigma_{\rho_{2k-1}} = -\sigma_{\rho_{2k}}.$$

In this section we aim at proving the following proposition whose proof is postponed to the end of this section (see Sect. 2.3). We recall that  $a$  is defined with respect to the length,  $\nu$ , of the torus by the relation  $a_i = \nu_i^2$  (see (1.12)).

**Proposition 2.2** *For almost all  $a \in (1, 4)^d$ , there exists  $\gamma > 0$  such that for all  $\delta > 0$ ,  $r \geq 3$ ,  $\sigma_1, \dots, \sigma_r \in \{-1, 1\}$ ,  $j_1, \dots, j_r \in \mathbb{Z}^d$  satisfying  $\sigma_1 j_1 + \dots + \sigma_r j_r = 0$  and  $|j_1| \geq \dots \geq |j_r|$  at least one of the following assertion holds*

- (i)  $(\sigma, j)$  is resonant (see Definition 2.1)
- (ii)  $\sigma_1 \sigma_2 = 1$  and

$$\left| \sum_{k=1}^r \sigma_k \sqrt{1 + |j_k|_a^4} \right| \gtrsim_r \gamma \langle j_3 \rangle \dots \langle j_r \rangle^{-9dr^2},$$

- (iii)  $\sigma_1 \sigma_2 = -1$  and

$$\left| \sum_{k=1}^r \sigma_k \sqrt{1 + |j_k|_a^4} \right| \gtrsim_{r, \delta} \gamma \langle j_1 \rangle^{-(d-1+\delta)} \langle j_3 \rangle \dots \langle j_r \rangle^{-44dr^4}.$$

We refer the reader to Lemma 2.9 and its corollary to understand how we get this degeneracy with respect to  $j_1$ .

### 2.1 A Weak Non-resonance Estimate

In this subsection we aim at proving the following technical lemma.

**Lemma 2.3** *If  $r \geq 1$ ,  $(j_1, \dots, j_r) \in (\mathbb{N}^d)^r$  is injective,<sup>5</sup>  $n \in (\mathbb{Z}^*)^r$  and  $\kappa \in \mathbb{R}^d$  satisfies  $\kappa_{i_*} = 0$  for some  $i_* \in \llbracket 1, d \rrbracket$  then we have*

$$\forall \gamma > 0, \left\{ a \in (1, 4)^d : \left| \kappa \cdot a + \sum_{k=1}^r n_k \sqrt{1 + |j_k|_a^4} \right| < \gamma \right\} \lesssim_{r, d} \gamma^{\frac{1}{r(r+1)}} \langle j_1 \rangle \dots \langle j_r \rangle^{\frac{12}{r+1}}.$$

Its proof (postponed to the end of this subsection) rely essentially on the following lemma.

<sup>5</sup> i.e.  $\forall k, \ell \in \llbracket 1, r \rrbracket, k \neq \ell \Rightarrow j_k \neq j_\ell$ .

**Lemma 2.4** *If  $I, J$  are two bounded intervals of  $\mathbb{R}_+^*$ ,  $r \geq 1$ ,  $(j_1, \dots, j_r) \in (\mathbb{N}^d)^r$  is injective,  $n \in (\mathbb{Z}^*)^r$  and  $h : J^{d-1} \rightarrow \mathbb{R}$  is measurable then for all  $\gamma > 0$  we have*

$$\left| \left\{ (m, b) \in I \times J^{d-1} : \left| h(b) + \sum_{k=1}^r n_k \sqrt{m + |j_k|_{(1,b)}^4} \right| < \gamma \right\} \right| \lesssim_{r,d,I,J} \gamma^{\frac{1}{r(r+1)}} (\langle j_1 \rangle \dots \langle j_r \rangle)^{\frac{12}{r+1}}$$

where  $(1, b) := (1, b_1, \dots, b_{d-1}) \in \mathbb{R}^d$ .

**Proof of Lemma 2.4** The proof of this lemma is classical and follows the lines of [1].

Without loss of generality, we assume that  $\gamma \in (0, 1)$ . Let  $\eta \in (0, 1)$  be a positive number which will be optimized later with respect to  $\gamma$ . If  $1 \leq i < k \leq r$  then we have

$$|j_i|_{1,b}^2 - |j_k|_{1,b}^2 = (j_{i,1}^2 - j_{k,1}^2) + b_1(j_{i,2}^2 - j_{k,2}^2) + \dots + b_{d-1}(j_{i,d}^2 - j_{k,d}^2).$$

Since, by assumption,  $(j_1, \dots, j_r)$  is injective, either there exists  $\ell \in \llbracket 2, d \rrbracket$  such that  $j_{i,\ell} \neq j_{k,\ell}$  or  $j_{i,1} \neq j_{k,1}$  and  $j_{i,\ell} = j_{k,\ell}$  for  $\ell = 2, \dots, d$ . Note that in this second case, we have  $\|j_i|_{1,b}^2 - |j_k|_{1,b}^2\| \geq 1$ . In any case, since the dependency with respect to  $b$  is affine the set

$$\mathcal{P}_\eta^{(i,k)} = \{b \in J^{d-1} \mid |j_i|_{1,b}^2 - |j_k|_{1,b}^2 < \eta\} \text{ satisfies } |\mathcal{P}_\eta^{(i,k)}| < \eta(1 + |J|^{d-1}).$$

Therefore, we have

$$\left| \left\{ (m, b) \in I \times J^{d-1} : \left| h(b) + \sum_{k=1}^r n_k \sqrt{m + |j_k|_{(1,b)}^4} \right| < \gamma \right\} \right| \leq \frac{r(r-1)}{2} |I| \eta (1 + |J|^{d-1}) + |J|^{d-1} \sup_{\forall i < k, b \notin \mathcal{P}_\eta^{(i,k)}} \left| \left\{ m \in I : \left| h(b) + \sum_{k=1}^r n_k \sqrt{m + |j_k|_{(1,b)}^4} \right| < \gamma \right\} \right|. \tag{2.1}$$

In order to estimate this last measure we fix  $b \in J^{d-1} \setminus \bigcup_{i < k} \mathcal{P}_\eta^{(i,k)}$  and we define  $g : I \rightarrow \mathbb{R}$  by

$$g(m) = h(b) + \sum_{k=1}^r n_k \sqrt{m + |j_k|_{(1,b)}^4}.$$

By a straightforward calculation, for  $\ell \geq 1$ , we have

$$\partial_m^\ell g(m) = c_\ell \sum_{k=1}^r n_k (m + |j_k|_{(1,b)}^4)^{\frac{1}{2} - \ell} \quad \text{where} \quad c_\ell = \prod_{i=0}^{\ell-1} \frac{1}{2} - i. \tag{2.2}$$

Therefore, we have

$$\begin{pmatrix} c_1^{-1} \partial_m^1 g \\ \vdots \\ c_r^{-1} \partial_m^r g \end{pmatrix} = \begin{pmatrix} (m + |j_1|_{(1,b)}^4)^0 & \dots & (m + |j_r|_{(1,b)}^4)^0 \\ \vdots & & \vdots \\ (m + |j_1|_{(1,b)}^4)^{-(r-1)} & \dots & (m + |j_r|_{(1,b)}^4)^{-(r-1)} \end{pmatrix} \begin{pmatrix} n_1 \sqrt{m + |j_1|_{(1,b)}^4}^{-1} \\ \vdots \\ n_r \sqrt{m + |j_r|_{(1,b)}^4}^{-1} \end{pmatrix}.$$

Denoting by  $V$  this Vandermonde matrix, by  $\|x\|_\infty := \max |x_i|$  for  $x \in \mathbb{R}^d$  and also by  $\|\cdot\|_\infty$  the associated matrix norm, we deduce that

$$\max_{i=1}^r c_i^{-1} |\partial_m^i g(m)| \geq |V^{-1}|_\infty^{-1} \max_{i=1}^r |n_i| \sqrt{m + |j_i|_{(1,b)}^4}^{-1}. \tag{2.3}$$

We recall that the inverse of  $V$  is given by

$$(V^{-1})_{i,\ell} = (-1)^{r-\ell} \frac{S_{r-\ell} \left( \left( \frac{1}{m+|j_k|^4_{(1,b)}} \right)_{k \neq i} \right)}{\prod_{k \neq i} \frac{1}{m+|j_i|^4_{(1,b)}} - \frac{1}{m+|j_k|^4_{(1,b)}}} \tag{2.4}$$

(this formula can be easily derived using the Lagrange interpolation polynomials) where  $S_\ell : \mathbb{R}^{r-1} \rightarrow \mathbb{R}$  is the  $\ell^{st}$  elementary symmetric function

$$S_\ell(x) = \sum_{1 \leq k_1 < \dots < k_\ell \leq r-1} x_{k_1} \dots x_{k_\ell} \quad \text{and} \quad S_0(x) := 1.$$

Furthermore, we have

$$|V^{-1}|_\infty = \max_{i=1}^r \sum_{\ell=1}^r |(V^{-1})_{i,\ell}|. \tag{2.5}$$

To estimate  $|V^{-1}|_\infty$  in (2.3), we use the estimates

$$S_{r-\ell} \left( \left( \frac{1}{m+|j_k|^4_{(1,b)}} \right)_{k \neq i} \right) \lesssim_{r,J,I} 1 \quad \text{and} \quad \left| \frac{1}{m+|j_i|^4_{(1,b)}} - \frac{1}{m+|j_k|^4_{(1,b)}} \right| \gtrsim_{J,I} \frac{\eta}{\langle jk \rangle^6}.$$

Indeed, if  $||j_i|^4_{(1,b)} - |j_k|^4_{(1,b)}| \geq \frac{1}{2}|j_i|^4_{(1,b)}$  we have

$$\left| \frac{1}{m+|j_i|^4_{(1,b)}} - \frac{1}{m+|j_k|^4_{(1,b)}} \right| = \left| \frac{|j_i|^4_{(1,b)} - |j_k|^4_{(1,b)}}{(m+|j_i|^4_{(1,b)})(m+|j_k|^4_{(1,b)})} \right| \gtrsim_{I,J} \frac{1}{\langle jk \rangle^4}$$

and conversely, if  $||j_i|^4_{(1,b)} - |j_k|^4_{(1,b)}| \leq \frac{1}{2}|j_i|^4_{(1,b)}$  then  $|j_i|^4_{(1,b)} \leq 2|j_k|^4_{(1,b)}$  and so, since  $b \in J^{d-1} \setminus \bigcup_{i < k} \mathcal{P}_\eta^{(i,k)}$ , we have

$$\left| \frac{1}{m+|j_i|^4_{(1,b)}} - \frac{1}{m+|j_k|^4_{(1,b)}} \right| \gtrsim_{I,J} \frac{(|j_i|^2_{(1,b)} + |j_k|^2_{(1,b)})||j_i|^2_{(1,b)} - |j_k|^2_{(1,b)}|}{\langle jk \rangle^8} \gtrsim_{I,J} \frac{\eta}{\langle jk \rangle^6}.$$

Therefore by (2.5) and (2.4), we have

$$|V^{-1}|_\infty \lesssim_{r,I,J} \eta^{-(r-1)} (\langle j_1 \rangle \dots \langle j_r \rangle)^6$$

Consequently, we deduce from (2.3) that

$$\max_{i=1}^r |\partial_m^i g(m)| \gtrsim_{r,I,J} \eta^{r-1} (\langle j_1 \rangle \dots \langle j_r \rangle)^{-6} |n|_\infty. \tag{2.6}$$

Furthermore, considering (2.2), it is clear that

$$|\partial_m^\ell g(m)| \lesssim_{\ell,I,J} |n|_\infty.$$

As a consequence, being given  $\rho > 0$  (that will be optimized later), applying Lemma B.1. of [17], we get  $N$  sub-intervals of  $I$ , denoted  $\Delta_1, \dots, \Delta_N$  such that

$$N \lesssim_{I,r} (\langle j_1 \rangle \dots \langle j_r \rangle)^6 \eta^{-(r-1)}, \quad \max_{i=1}^N |\Delta_i| \lesssim_{I,r} \left( \frac{\rho (\langle j_1 \rangle \dots \langle j_r \rangle)^6}{\eta^{r-1} |n|_\infty} \right)^{\frac{1}{r-1}},$$

$$|\partial_m g(m)| \geq \rho \quad \forall m \in I \setminus (\Delta_1 \cup \dots \cup \Delta_N).$$

Observing that  $I \setminus (\Delta_1 \cup \dots \cup \Delta_N)$  can be written as the union of  $M$  intervals with  $M \lesssim 1 + N$ , we deduce that

$$\left| \left\{ m \in I : \left| h(b) + \sum_{k=1}^r n_k \sqrt{m + |j_k|_{(1,b)}^4} \right| < \gamma \right\} \right| < M \rho^{-1} \gamma + N \max_{i=1}^N |\Delta_i|$$

$$\lesssim_{I,r} (\langle j_1 \rangle \dots \langle j_r \rangle)^6 \eta^{-(r-1)} \left[ \rho^{-1} \gamma + \left( \frac{\rho (\langle j_1 \rangle \dots \langle j_r \rangle)^6}{\eta^{r-1} |n|_\infty} \right)^{\frac{1}{r-1}} \right].$$

We optimize  $\rho$  to equalize the two terms in this last sum:

$$\rho^{\frac{r}{r-1}} = \gamma \left( \frac{\eta^{r-1} |n|_\infty}{(\langle j_1 \rangle \dots \langle j_r \rangle)^6} \right)^{\frac{1}{r-1}}.$$

This provides the estimate

$$\left| \left\{ m \in I : \left| h(b) + \sum_{k=1}^r n_k \sqrt{m + |j_k|_{(1,b)}^4} \right| < \gamma \right\} \right|$$

$$\lesssim_{I,r} \gamma^{\frac{1}{r}} (\langle j_1 \rangle \dots \langle j_r \rangle)^6 \eta^{-(r-1)} \left( \frac{(\langle j_1 \rangle \dots \langle j_r \rangle)^6}{\eta^{r-1} |n|_\infty} \right)^{\frac{1}{r}}$$

$$\lesssim_{I,r} \left( \frac{\gamma}{|n|_\infty} \right)^{\frac{1}{r}} \eta^{-(r-1+\frac{r-1}{r})} (\langle j_1 \rangle \dots \langle j_r \rangle)^{12}.$$

Finally, we optimize (2.1) by choosing

$$\eta = \gamma^{\frac{1}{r}} \eta^{-(r-1+\frac{r-1}{r})} (\langle j_1 \rangle \dots \langle j_r \rangle)^{12}$$

and, recalling that  $|n|_\infty \geq 1$ , we get

$$\left| \left\{ (m, b) \in I \times J^{d-1} : \left| h(b) + \sum_{k=1}^r n_k \sqrt{m + |j_k|_{(1,b)}^4} \right| < \gamma \right\} \right|$$

$$\lesssim_{r,d,I,J} \left( \gamma^{\frac{1}{r}} (\langle j_1 \rangle \dots \langle j_r \rangle)^{12} \right)^{\frac{1}{r+\frac{r-1}{r}}}.$$

Since this measure is obviously bounded by  $|I||J|^{d-1}$ , the exponent  $r + \frac{r-1}{r}$  can be replaced by  $r + 1$  in the above expression which conclude this proof.  $\square$

Now using Lemma 2.4, we prove Lemma 2.3.

**Proof of Lemma 2.3** Without loss of generality we assume that  $i_\star = 1$ . First, since  $\kappa_1 = 0$ , we note that we have

$$G(a) := \kappa \cdot a + \sum_{k=1}^r n_k \sqrt{1 + |j_k|_a^4} = \frac{1}{\sqrt{m}} (h(b) + \sum_{k=1}^r n_k \sqrt{m + |j_k|_{(1,b)}^4}) =: \frac{1}{\sqrt{m}} F(m, b)$$

where

$$m = \frac{1}{a_1^2}, \quad b = \left( \frac{a_2}{a_1}, \dots, \frac{a_d}{a_1} \right) \quad \text{and} \quad h(b) = \sum_{k=2}^d \kappa_k b_k.$$

Let denote by  $\Psi$  the map  $a \mapsto (m, b)$ . It is clearly smooth and injective. Furthermore, we have

$$\det d\Psi(a) = \begin{vmatrix} -2a_1^{-3} & -a_2a_1^{-2} & \dots & -a_da_1^{-2} \\ & a_1^{-1} & & \\ & & \ddots & \\ & & & a_1^{-1} \end{vmatrix} = 2(-1)a_1^{-d-2}.$$

Consequently,  $\Psi$  is a smooth diffeomorphism onto its image  $\Psi((1, 4)^d)$  which is included in the rectangle  $(\frac{1}{16}, 1) \times (\frac{1}{4}, 4)^{d-1}$ . Therefore, by a change of variable, we have

$$\begin{aligned} |\{a \in (1, 4)^d : |G(a)| < \gamma\}| &= \int_{a \in (1, 4)^d} \mathbb{1}_{|G(a)| < \gamma} da \\ &= \int_{(m, b) \in \Psi((1, 4)^d)} \mathbb{1}_{|F(m, b)| < \sqrt{m}\gamma} (2\sqrt{m}^{-d-2}) d(m, b) \\ &\leq 2^{2d+5} \left| \left\{ (m, b) \in \left(\frac{1}{16}, 1\right) \times \left(\frac{1}{4}, 4\right)^{d-1} : |F(m, b)| < \gamma \right\} \right|. \end{aligned}$$

Finally, by applying Lemma 2.4, we get the expected estimate. □

### 2.2 Non-resonance Estimates for Two Large Modes

In this subsection we consider  $r \geq 3$ ,  $(j_k)_{k \geq 3} \in (\mathbb{Z}^d)^{r-2}$  and  $\sigma \in \{-1, 1\}^r$  such that  $\sigma_1 = -\sigma_2$  as fixed. We define  $j_{\geq 3} \in \mathbb{Z}^d$  by

$$j_{\geq 3} := \sigma_3 j_3 + \dots + \sigma_r j_r. \tag{2.7}$$

Being given  $j_1 \in \mathbb{Z}^d$ , we define implicitly  $j_2 := j_1 + \sigma_1 j_{\geq 3}$  in order to satisfy the *zero momentum condition*

$$\sum_{k=1}^r \sigma_k j_k = 0, \tag{2.8}$$

and we define the function  $g_{j_1} : (1, 4)^d \rightarrow \mathbb{R}$  by

$$g_{j_1}(a) = \sum_{k=1}^r \sigma_k \sqrt{1 + |j_k|_a^4}.$$

Finally, for  $\gamma > 0$ , we introduce the following sets<sup>6</sup>

$$\mathcal{I} = \{i : j_{\geq 3, i} \neq 0\}, \quad C_i = \left\{ j_1 \in \mathbb{Z}^d : |j_{1, i}| \geq 2 \left( 1 + \sum_{k \geq 3} |j_{k, i}|^2 \right) \right\},$$

$$S = \left\{ j_1 \in \mathbb{Z}^d \setminus \bigcup_{i \in \mathcal{I}} C_i : (\sigma, j) \text{ is non-resonant} \right\},$$

and  $R_\gamma = \{j_1 \in S : |j_1| \geq \gamma^{-1/2} (\langle j_3 \rangle \dots \langle j_r \rangle)^{2dr^2}\}.$

<sup>6</sup> see Definition 2.1.

First, we prove the following technical lemma whose Corollary 2.6 allows to deal with the non-degenerated cases.

**Lemma 2.5** *If there exists  $i \in \llbracket 1, d \rrbracket$  such that*

$$|(j_{1,i} + j_{2,i})_{j_{\geq 3,i}}| \geq 2 \left( 1 + \sum_{k=3}^r j_{k,i}^2 \right) \tag{2.9}$$

then for all  $\gamma > 0$

$$\left| \left\{ a \in (1, 4)^d : \left| \sum_{k=1}^r \sigma_k \sqrt{1 + |j_k|_a^4} < \gamma \right| \right\} \right| < \frac{2\gamma}{|j_{1,i} + j_{2,i}|}. \tag{2.10}$$

**Proof** Without loss of generality we assume that  $\sigma_1 = 1$  and  $\sigma_2 = -1$ . We compute the derivative with respect to  $a_1$

$$\partial_{a_1} \sum_{k=1}^r \sigma_k \sqrt{1 + |j_k|_a^4} = \sum_{k=1}^r \sigma_k j_{k,i}^2 \frac{|j_k|_a^2}{\sqrt{1 + |j_k|_a^4}}.$$

Consequently, we have

$$\left| \partial_{a_1} \left( \sum_{k=1}^r \sigma_k \sqrt{1 + |j_k|_a^4} \right) \right| \geq \left| j_{1,i}^2 \frac{|j_1|_a^2}{\sqrt{1 + |j_1|_a^4}} - j_{2,i}^2 \frac{|j_2|_a^2}{\sqrt{1 + |j_2|_a^4}} \right| - \sum_{k=3}^r j_{k,i}^2.$$

Furthermore, we have

$$\left| \frac{|j_1|_a^2}{\sqrt{1 + |j_1|_a^4}} - 1 \right| \leq \frac{1}{2|j_1|_a^2}.$$

Consequently, we get

$$\left| \partial_{a_1} \left( \sum_{k=1}^r \sigma_k \sqrt{1 + |j_k|_a^4} \right) \right| \geq |j_{1,i}^2 - j_{2,i}^2| - 1 - \sum_{k=3}^r j_{k,i}^2.$$

Observing that by definition we have  $j_{1,i}^2 - j_{2,i}^2 = j_{\geq 3,i}(j_{1,i} + j_{2,i})$ , we deduce of the assumption (2.9) that

$$\left| \partial_{a_1} \sum_{k=1}^r \sigma_k \sqrt{1 + |j_k|_a^4} \right| \geq \frac{1}{2} |j_{\geq 3,i}(j_{1,i} + j_{2,i})|$$

Since by (2.9) we know that  $j_{\geq 3,i} \in \mathbb{Z} \setminus \{0\}$ , we deduce that

$$\left| \partial_{a_1} \sum_{k=1}^r \sigma_k \sqrt{1 + |j_k|_a^4} \right| \geq \frac{1}{2} |(j_{1,i} + j_{2,i})|.$$

Therefore  $a_1 \mapsto \sum_{k=1}^r \sigma_k \sqrt{1 + |j_k|_a^4}$  is a diffeomorphism (it is a smooth monotonic function). Consequently, applying this change of coordinate, we get directly (2.10) which conclude this proof. □

**Corollary 2.6** *For all  $\gamma > 0$  we have*

$$\forall i \in \mathcal{I}, \left| \{ a \in (1, 4)^d : \exists j_1 \in C_i, |g_{j_1}(a)| < \gamma |j_1|^{-(d-1)} \log^{-2d}(|j_1|) \} \right| \lesssim_d \gamma \tag{2.11}$$

**Proof of Corollary 2.6** Let  $j_1 \in C_i$ . By definition of  $j_2$ , we have

$$|j_{1,i} + j_{2,i}| \geq 2|j_{1,i}| - \sum_{k=3}^r |j_{k,i}|.$$

Consequently, since  $j_1 \in C_i$ , we have

$$|j_{1,i} + j_{2,i}| \geq 2|j_{1,i}| - \sum_{k=3}^r |j_{k,i}|^2 \geq \frac{3}{2}|j_{1,i}|.$$

Therefore, since  $j_{\geq 3,i} \neq 0$ , we have

$$|j_{\geq 3,i}(j_{1,i} + j_{2,i})| \geq \frac{3}{2}|j_{1,i}| \geq 3 \left( 1 + \sum_{k \geq 3} |j_{k,i}|^2 \right).$$

Applying Lemma 2.5, we deduce that for all  $\gamma > 0$

$$\left| \left\{ a \in (1, 4)^d : \left| \sum_{k=1}^r \sigma_k \sqrt{1 + |j_k|_a^4} < \gamma \right| \right\} \right| < \frac{4\gamma}{3|j_{1,i}|}.$$

Consequently, we have

$$\begin{aligned} & |\{a \in (1, 4)^d : \exists j_1 \in C_i, |g_{j_1}(a)| < \gamma |j_1|^{-(d-1)} \log^{-2d}(|j_1|)\}| \\ &= \left| \bigcup_{j_1 \in C_i} \{a \in (1, 4)^d : |g_{j_1}(a)| < \gamma |j_1|^{-(d-1)} \log^{-2d}(|j_1|)\} \right| \\ &\leq \sum_{j_1 \in C_1} |\{a \in (1, 4)^d : |g_{j_1}(a)| < \gamma |j_1|^{-(d-1)} \log^{-2d}(|j_1|)\}| \\ &\lesssim \gamma \sum_{j_1 \in C_i} \frac{1}{|j_1|^{(d-1)} |j_{1,i}| \log^{2d}(|j_1|)} \lesssim_d \gamma. \end{aligned}$$

□

In the following lemma, we deal with most of the degenerated cases.

**Lemma 2.7** For all  $\gamma > 0$ , we have

$$|\{a \in (1, 4)^d : \exists j_1 \in R_\gamma, |g_{j_1}(a)| < \gamma\}| \lesssim_{r,d} \gamma^{\frac{1}{r^2}} (\langle j_3 \rangle \dots \langle j_r \rangle)^{2d}. \tag{2.12}$$

**Proof** Without loss of generality, we assume that  $\gamma < \min((2r)^{-2}, (36d)^{-1})$ . If  $j_1 \in R_\gamma$ , recalling that for  $x \geq 0$ , we have  $|\sqrt{1+x} - 1| \leq x/2$ , we deduce that

$$\begin{aligned} |g_{j_1}(a)| &\geq |h_{j_1}(a)| - \frac{1}{2|j_1|_a^2} - \frac{1}{2|j_2|_a^2} \\ \text{where } h_{j_1}(a) &:= |j_1|_a^2 - |j_2|_a^2 + \sigma_1 \sum_{k=3}^r \sigma_k \sqrt{1 + |j_k|_a^4}. \end{aligned}$$

However, by definition of  $j_2$  and  $R_\gamma$ , we have

$$|j_2| \geq |j_1| - \sum_{k=3}^r |j_k| \geq \gamma^{-1/2} (\langle j_3 \rangle \dots \langle j_r \rangle)^{2dr^2} - (r - 2) (\langle j_3 \rangle \dots \langle j_r \rangle) \geq \frac{\gamma^{-1/2}}{2} (\langle j_3 \rangle \dots \langle j_r \rangle)^{2dr^2}.$$

Noting that, for  $a \in (1, 4)^d$ , we have  $|\cdot| \leq |\cdot|_a$ , we deduce that

$$|g_{j_1}(a)| \geq |h_{j_1}(a)| - 3\gamma (\langle j_3 \rangle \dots \langle j_r \rangle)^{-4dr^2}.$$

Consequently, it is enough to prove that

$$\{|a \in (1, 4)^d : \exists j_1 \in R_1, |h_{j_1}(a)| < \gamma (\langle j_3 \rangle \dots \langle j_r \rangle)^{-4dr^2}\} \lesssim_{r,d} \gamma^{\frac{1}{(r-1)(r-2)}}. \tag{2.13}$$

To prove this estimate, we have to note the following result whose proof is postponed/to the end of this proof. □

**Lemma 2.8** *If  $j_1 \in R_\gamma$  then there exists  $\kappa_{j_1} \in \mathbb{Z}^d$  such that*

$$|j_1|_a^2 - |j_2|_a^2 = \kappa_{j_1} \cdot a, \quad |\kappa_{j_1}|_\infty \leq 10(\langle j_3 \rangle \dots \langle j_r \rangle)^3 \quad \text{and} \quad \exists i_\star \in \llbracket 1, d \rrbracket, \kappa_{j_1, i_\star} = 0.$$

Now we have to distinguish two cases.

- *Case 1:  $(\sigma_k, j_k)_{k \geq 3}$  is resonant.* If  $j_1 \in R_\gamma$ , let  $\kappa_{j_1} \in \mathbb{Z}^d$  be given by Lemma 2.8. Note that  $\kappa_{j_1} \neq 0$  because else we would have  $j_{1,i}^2 = j_{2,i}^2$  for all  $i \in \llbracket 1, d \rrbracket$  and so  $(\sigma, j)$  would be resonant (which is excluded by definition of  $R_\gamma$ ). Furthermore, here  $h_{j_1}(a) = \kappa_{j_1} \cdot a$  is a linear form. Consequently, for all  $\gamma > 0$ , we have the following estimate which is much stronger than (2.13):

$$\begin{aligned} & \{|a \in (1, 4)^d : \exists j_1 \in R_1, |h_{j_1}(a)| < \gamma\}| \\ & \leq \left| \bigcup_{\substack{\kappa \in \mathbb{Z}^d \setminus \{0\} \\ |\kappa|_\infty \leq 10(\langle j_3 \rangle \dots \langle j_r \rangle)^3}} \{a \in (1, 4)^d : \kappa \cdot a < \gamma\} \right| \\ & \leq \sum_{\substack{\kappa \in \mathbb{Z}^d \setminus \{0\} \\ |\kappa|_\infty \leq 10(\langle j_3 \rangle \dots \langle j_r \rangle)^3}} |\{a \in (1, 4)^d : \kappa \cdot a < \gamma\}| \leq \gamma (20(\langle j_3 \rangle \dots \langle j_r \rangle)^3)^d \end{aligned}$$

- *Case 2:  $(\sigma_k, j_k)_{k \geq 3}$  is non-resonant.* If  $j_1 \in R_\gamma$ ,  $h_{j_1}$  writes

$$h_{j_1}(a) = \kappa_{j_1} \cdot a + \sum_{k=1}^{\tilde{r}} n_k \sqrt{1 + |\tilde{j}_k|_a^4}$$

where  $\kappa_{j_1}$  is given by Lemma 2.8,  $\tilde{r} \leq r - 2$ ,  $(\tilde{j}_1, \dots, \tilde{j}_{\tilde{r}}) \in (\mathbb{N}^d)^{\tilde{r}}$  is injective,  $n_k \in (\mathbb{Z} \setminus \{0\})^d$  is defined by

$$n_k = \sum_{\substack{i \in \llbracket 3, r \rrbracket \\ \forall \ell, |j_{i,\ell}| = |\tilde{j}_{k,\ell}}} \sigma_i \sigma_i.$$



Consequently, by Lemma 2.8, we have

$$\begin{aligned} & |\{a \in (1, 4)^d : \exists j_1 \in R_1, |h_{j_1}(a)| < \gamma\}| \\ & \leq \left| \bigcup_{\substack{\kappa \in \mathbb{Z}^d \\ |\kappa|_\infty \leq 10(\langle j_3 \rangle \dots \langle j_r \rangle)^3 \\ \exists i_*, \kappa_{i_*} = 0}} \left\{ a \in (1, 4)^d : \left| \kappa \cdot a + \sum_{k=1}^{\tilde{r}} n_k \sqrt{1 + |\tilde{j}_k|_a^4} \right| < \gamma \right\} \right| \\ & \leq \sum_{\substack{\kappa \in \mathbb{Z}^d \\ |\kappa|_\infty \leq 10(\langle j_3 \rangle \dots \langle j_r \rangle)^3 \\ \exists i_*, \kappa_{i_*} = 0}} \left| \left\{ a \in (1, 4)^d : \left| \kappa \cdot a + \sum_{k=1}^{\tilde{r}} n_k \sqrt{1 + |\tilde{j}_k|_a^4} \right| < \gamma \right\} \right|. \end{aligned}$$

Finally, by applying Lemma 2.3 we get

$$|\{a \in (1, 4)^d : \exists j_1 \in R_1, |h_{j_1}(a)| < \gamma\}| \lesssim_{r,d} \gamma^{\frac{1}{(r-2)(r-1)}} (\langle j_3 \rangle \dots \langle j_r \rangle)^{\frac{12}{r-1} + 3d},$$

which is also stronger than (2.13). □

**Proof of Lemma 2.8** First let us note that

$$|j_1|_a^2 - |j_2|_a^2 = \kappa_{j_1} \cdot a \quad \text{where} \quad \kappa_{j_1,i} = j_{1,i}^2 - j_{2,i}^2 = \sigma_2 j_{\geq 3,i} (j_{1,i} + j_{2,i}).$$

First we aim at controlling  $|\kappa|_\infty$ . If  $i \notin \mathcal{I}$  then  $j_{\geq 3,i} = 0$  and so  $\kappa_{j_1,i} = 0$ . Else, since  $j_1 \in \mathbb{Z}^d \setminus \bigcup_{i \in \mathcal{I}} C_i$ , we have  $|j_{1,i}| \leq 2(1 + \sum_{k \geq 3} |j_{k,i}|^2)$ . Consequently, we deduce that

$$|\kappa_{j_1,i}| \leq \left( \sum_{k \geq 3} |j_{k,i}| \right) \left( 4 + 4 \sum_{k \geq 3} |j_{k,i}|^2 + \sum_{k \geq 3} |j_{k,i}| \right) \leq 10(\langle j_3 \rangle \dots \langle j_r \rangle)^3.$$

Now we assume by contradiction that  $\kappa_{j_1,i} \neq 0$  for all  $i \in \llbracket 1, d \rrbracket$ . Consequently, we have  $\mathcal{I} = \llbracket 1, d \rrbracket$  and so

$$|j_1|_\infty \leq 2 \left( 1 + \sum_{k \geq 3} |j_k|^2 \right) \leq 6 \langle j_3 \rangle^2 \dots \langle j_r \rangle^2. \tag{2.14}$$

However, since  $j_1 \in R_\gamma$ , we have  $|j_1| \geq \gamma^{-1/2} (\langle j_3 \rangle \dots \langle j_r \rangle)^{2dr^2}$  which is in contradiction with (2.14) because we have assumed that  $\gamma < (36d)^{-1}$ . □

Finally in the following lemma we deal with the general degenerated cases.

**Lemma 2.9** For all  $\gamma > 0$ , we have

$$|\{a \in (1, 4)^d : \exists j_1 \in S, |g_{j_1}(a)| < \gamma\}| \lesssim_{r,d} \gamma^{\frac{1}{8r^4}} (\langle j_3 \rangle \dots \langle j_r \rangle)^{5d}. \tag{2.15}$$

**Proof** Without loss of generality we assume that  $\gamma \in (0, 1)$ . Let  $\eta \in (0, 1)$  be a small number that will be optimized with respect to  $\gamma$  later. From the decomposition  $S = R_\eta \cup (S \setminus R_\eta)$  we get

$$\begin{aligned} & |\{a \in (1, 4)^d : \exists j_1 \in S, |g_{j_1}(a)| < \gamma\}| \leq \sum_{j_1 \in S \setminus R_\eta} |\{a \in (1, 4)^d : |g_{j_1}(a)| < \gamma\}| \\ & + |\{a \in (1, 4)^d : \exists j_1 \in R_\eta, |g_{j_1}(a)| < \eta\}|. \end{aligned} \tag{2.16}$$

To estimate the sum, we apply Lemma 2.3 (with  $\kappa = 0$ ) and we get

$$\begin{aligned} \sum_{j_1 \in S \setminus R_\eta} |\{a \in (1, 4)^d : |g_{j_1}(a)| < \gamma\}| &\leq \sum_{\substack{|j_1| < \eta^{-1/2} \langle j_3 \rangle \dots \langle j_r \rangle \\ (\sigma, j) \text{ is non-resonant}}} |\{a \in (1, 4)^d : |g_{j_1}(a)| < \gamma\}| \\ &\leq \sum_{|j_1| < \eta^{-1/2} \langle j_3 \rangle \dots \langle j_r \rangle} \gamma^{\frac{1}{r(r+1)} \langle j_1 \rangle \dots \langle j_r \rangle^{\frac{12}{r+1}}}. \end{aligned}$$

Furthermore, by the zero momentum condition (2.8), since  $\eta \in (0, 1)$ , we also have

$$|j_2| \lesssim_r \eta^{-1/2} \langle j_3 \rangle \dots \langle j_r \rangle^{2dr^2}.$$

Consequently, we have

$$\begin{aligned} \sum_{j_1 \in S \setminus R_\eta} |\{a \in (1, 4)^d : |g_{j_1}(a)| < \gamma\}| &\lesssim_r \gamma^{\frac{1}{r(r+1)}} \eta^{-\frac{1}{2} - \frac{12}{r+1}} \langle j_3 \rangle \dots \langle j_r \rangle^{2dr^2 + \frac{12}{r+1} + \frac{24}{r+1} 2dr^2} \\ &\lesssim_r \gamma^{\frac{1}{2r^2}} \eta^{-\frac{7}{2}} \langle j_3 \rangle \dots \langle j_r \rangle^{15dr^2}. \end{aligned}$$

Therefore, applying Lemma 2.7, we deduce of (2.16) that

$$\begin{aligned} |\{a \in (1, 4)^d : \exists j_1 \in S, |g_{j_1}(a)| < \gamma\}| &\lesssim_{r,d} \eta^{\frac{1}{r^2}} \langle j_3 \rangle \dots \langle j_r \rangle^{2d} \\ &+ \gamma^{\frac{1}{2r^2}} \eta^{-\frac{7}{2}} \langle j_3 \rangle \dots \langle j_r \rangle^{15dr^2}. \end{aligned}$$

Finally, we get (2.15) by optimizing this last estimate choosing

$$\eta = \gamma^{\frac{1}{7r^2+2}} \langle j_3 \rangle \dots \langle j_r \rangle^{\frac{15dr^2-2d}{7/2+1/r^2}}.$$

□

### 2.3 Proof of Proposition 2.2

For  $r \geq 3$  let  $\mathcal{M}_r$  and  $\mathcal{R}_r$  be the sets defined by

$$\begin{aligned} \mathcal{M}_r &= \{(\sigma, j) \in \{-1, 1\}^r \times (\mathbb{Z}^d)^r : \sum_{k=1}^r \sigma_k j_k = 0\} \text{ and} \\ \mathcal{R}_r &= \{(\sigma, j) \in \{-1, 1\}^r \times (\mathbb{Z}^d)^r : (\sigma, j) \text{ is resonant}\}. \end{aligned}$$

On the one hand, as a direct corollary of Lemma 2.9 and Corollary 2.6, for all  $\gamma > 0$  we have

$$\left| \left\{ a \in (1, 4)^d : \exists r \geq 3, \exists (\sigma, j) \in \mathcal{M}_r \setminus \mathcal{R}_r, \sigma_1 \sigma_2 = -1 \text{ and } \left| \sum_{k=1}^r \sigma_k \sqrt{1 + |j_k|_a^4} \right| < c_{r,d} \gamma^{8r^4} \langle j_1 \rangle^{-(d-1)} \log^{-2d} \langle j_1 \rangle \langle j_3 \rangle \dots \langle j_r \rangle^{-44dr^4} \right\} \right| < \gamma$$

where  $c_{r,d} > 0$  is a constant depending only on  $r$  and  $d$ . Consequently, it is enough to prove that for all  $\gamma \in (0, 1)$ , we have

$$\begin{aligned} I_\gamma &:= \left| \left\{ a \in (1, 4)^d : \exists r \geq 3, \exists (\sigma, j) \in \mathcal{M}_r \setminus \mathcal{R}_r, \sigma_1 \sigma_2 = 1 \text{ and } \left| \sum_{k=1}^r \sigma_k \sqrt{1 + |j_k|_a^4} \right| < \kappa_{r,d} \gamma^{r(r+1)} \langle j_3 \rangle \dots \langle j_r \rangle^{-9dr^2} \right\} \right| < \gamma \end{aligned} \tag{2.17}$$

where  $\kappa_{r,d} \in (0, 1)$  is another constant depending only on  $r$  and  $d$  (and that will be determined later). Indeed, by additivity of the measure, we have

$$I_\gamma \leq \sum_{r \geq 3} \sum_{\substack{(\sigma,j) \in \mathcal{M}_r \setminus \mathcal{R}_r \\ \sigma_1 \sigma_2 = 1}} \left\{ a \in (1, 4)^d : \left| \sum_{k=1}^r \sigma_k \sqrt{1 + |j_k|_a^4} \right| < \kappa_{r,d} \gamma^{r(r+1)} (\langle j_3 \rangle \dots \langle j_r \rangle)^{-9dr^2} \right\}.$$

Note that if  $|j_1| \geq 2\sqrt{r} \langle j_3 \rangle \dots \langle j_r \rangle$  and  $\sigma_1 \sigma_2 = 1$  then

$$\begin{aligned} \left| \sum_{k=1}^r \sigma_k \sqrt{1 + |j_k|_a^4} \right| &\geq \sqrt{1 + |j_1|_a^4} - \sum_{k=3}^r \sqrt{1 + |j_k|_a^4} \geq \sqrt{1 + |j_1|_a^4} - \sum_{k=3}^r \sqrt{1 + 16|j_k|^4} \\ &\geq \sqrt{1 + |j_1|^4} - 4 \sum_{k=3}^r \sqrt{1 + |j_k|^4} \geq |j_1|^2 - 4 \sum_{k=3}^r (1 + |j_k|^2) \geq 4(\langle j_3 \rangle \dots \langle j_r \rangle)^2 > 1 \end{aligned}$$

and so  $\left\{ a \in (1, 4)^d : \left| \sum_{k=1}^r \sigma_k \sqrt{1 + |j_k|_a^4} \right| < \kappa_{r,d} \gamma^{r(r+1)} (\langle j_3 \rangle \dots \langle j_r \rangle)^{-9dr^2} \right\}$  vanishes. Since the same holds if  $j_1$  is replaced by  $j_2$ , consequently, we have that  $I_\gamma$  is bounded from above by

$$\sum_{r \geq 3} \sum_{\substack{(\sigma,j) \in \mathcal{M}_r \setminus \mathcal{R}_r \\ |j_1| \leq 2\sqrt{r} \langle j_3 \rangle \dots \langle j_r \rangle \\ |j_2| \leq 2\sqrt{r} \langle j_3 \rangle \dots \langle j_r \rangle}} \left\{ a \in (1, 4)^d : \left| \sum_{k=1}^r \sigma_k \sqrt{1 + |j_k|_a^4} \right| < \kappa_{r,d} \gamma^{r(r+1)} (\langle j_3 \rangle \dots \langle j_r \rangle)^{-9dr^2} \right\}.$$

Now denoting by  $c_{r,d} > 0$  the constant given by Lemma 2.3, we get

$$I_\gamma \leq \sum_{r \geq 3} c_{r,d} \sum_{\substack{(\sigma,j) \in \mathcal{M}_r \setminus \mathcal{R}_r \\ |j_1| \leq 2\sqrt{r} \langle j_3 \rangle \dots \langle j_r \rangle \\ |j_2| \leq 2\sqrt{r} \langle j_3 \rangle \dots \langle j_r \rangle}} \left( \kappa_{r,d} \gamma^{r(r+1)} (\langle j_3 \rangle \dots \langle j_r \rangle)^{-9dr^2} \right)^{\frac{1}{r(r+1)}} (\langle j_1 \rangle \dots \langle j_r \rangle)^{\frac{12}{r+1}}.$$

Consequently, we get another constant  $\tilde{c}_{r,d} > 0$  such that

$$I_\gamma \leq \gamma \sum_{r \geq 3} \tilde{c}_{r,d} \kappa_{r,d}^{\frac{1}{r(r+1)}} \sum_{j_3, \dots, j_r \in \mathbb{Z}^d} (\langle j_3 \rangle \dots \langle j_r \rangle)^{-9d \frac{r^2}{r(r+1)} + \frac{36}{r+1}}.$$

Noting that  $9d \frac{r^2}{r(r+1)} - \frac{36}{r+1} \geq 2d$ , we deduce that

$$I_\gamma \leq \gamma \sum_{r \geq 3} \tilde{c}_{r,d} \kappa_{r,d}^{\frac{1}{r(r+1)}} \left( \sum_{j \in \mathbb{Z}^d} \langle j \rangle^{-2d} \right)^{r-2}.$$

Consequently, we deduce a natural choice for  $\kappa_{r,d}$  such that  $I_\gamma < \gamma$  which conclude this proof.

### 3 The Birkhoff Normal form Step

In the rest of the paper we shall fix the parameter  $\nu$ , (see (1.2) and (1.12)) defining the irrationality of the torus, in the full Lebesgue measure set given by Proposition 2.2. For

$d \geq 2$  and  $n \in \mathbb{N}$  we define

$$M_{d,n} := \begin{cases} n + 2(n - 2) + 1 & \text{if } d = 2 \text{ and } n \text{ odd} \\ n + 2(n - 2) & \text{if } d = 2 \text{ and } n \text{ even} \\ n + (n - 2) & \text{if } d = 3 \\ n & \text{if } d \geq 4. \end{cases} \tag{3.1}$$

The main result of this section is the following.

**Theorem 2** *Let  $d = 2, 3$  and let  $r \in \mathbb{N}$  such that  $M_{d,n} \leq r \leq 4n$ . There exists  $\beta = \beta(d, r) > 0$  such that for any  $N \geq 1$ , any  $\delta > 0$  and  $s \geq s_0 = s_0(\beta)$ , there exist  $\varepsilon_0 \lesssim_{s,\delta} N^{-\delta}$  and two canonical transformation  $\tau^{(0)}$  and  $\tau^{(1)}$  making the following diagram to commute*

$$\begin{array}{ccc} B_s(0, \varepsilon_0) & \xrightarrow{\tau^{(0)}} & B_s(0, 2\varepsilon_0) & \xrightarrow{\tau^{(1)}} & H^s(\mathbb{T}^d) \\ & \searrow & \text{id}_{H^s} & \swarrow & \end{array} \tag{3.2}$$

and close to the identity

$$\forall \sigma \in \{0, 1\}, \|u\|_{H^s} < 2^\sigma \varepsilon_0 \implies \|\tau^{(\sigma)}(u) - u\|_{H^s} \lesssim_{s,\delta} N^\delta \|u\|_{H^s}^2 \tag{3.3}$$

such that, on  $B_s(0, 2\varepsilon_0)$ ,  $H \circ \tau^{(1)}$  writes

$$H \circ \tau^{(1)} = Z_2 + \sum_{k=n}^{M_{d,n}-1} Z_k^{\leq N} + \sum_{k=M_{d,n}}^{r-1} K_k + K^{>N} + \tilde{R}_r \tag{3.4}$$

where  $M_{d,n}$  is given in (3.1) and where

(i)  $Z_k^{\leq N}$ , for  $k = n, \dots, M_{d,n} - 1$ , are resonant Hamiltonians of order  $k$  given by the formula

$$Z_k^{\leq N} = \sum_{\substack{\sigma \in \{-1,1\}^k, j \in (\mathbb{Z}^d)^k, \mu_2(j) \leq N \\ \sum_{i=1}^k \sigma_i j_i = 0 \\ \sum_{i=1}^k \sigma_i \omega_{j_i} = 0}} (Z_k^{\leq N})_{\sigma,j} u_{j_1}^{\sigma_1} \cdots u_{j_k}^{\sigma_k}, \quad |(Z_k^{\leq N})_{\sigma,j}| \lesssim_\delta N^\delta \frac{\mu_3(j)^\beta}{\mu_1(j)}; \tag{3.5}$$

(ii)  $K_k, k = M_{d,n}, \dots, r - 1$ , are homogeneous polynomials of order  $k$

$$K_k = \sum_{\substack{\sigma \in \{-1,1\}^k, j \in (\mathbb{Z}^d)^k \\ \sum_{i=1}^k \sigma_i j_i = 0}} (K_k)_{\sigma,j} u_{j_1}^{\sigma_1} \cdots u_{j_k}^{\sigma_k}, \quad |(K_k)_{\sigma,j}| \lesssim_\delta N^\delta \mu_3(j)^\beta; \tag{3.6}$$

(iii)  $K^{>N}$  and  $\tilde{R}_r$  are remainders satisfying

$$\|X_{K^{>N}}(u)\|_{H^s} \lesssim_{s,\delta} N^{-1+\delta} \|u\|_{H^s}^{n-1}, \tag{3.7}$$

$$\|X_{\tilde{R}_r}(u)\|_{H^s} \lesssim_{s,\delta} N^\delta \|u\|_{H^s}^{r-1}. \tag{3.8}$$

It is convenient to introduce the following class.

**Definition 3.1 (Formal Hamiltonians)** Let  $N \in \mathbb{R}, k \in \mathbb{N}$  with  $k \geq 3$  and  $N \geq 1$ .

(i) We denote by  $\mathcal{L}_k$  the set of Hamiltonian having homogeneity  $k$  and such that they may be written in the form

$$G_k(u) = \sum_{\substack{\sigma_i \in \{-1, 1\}, j_i \in \mathbb{Z}^d \\ \sum_{i=1}^k \sigma_i j_i = 0}} (G_k)_{\sigma, j} u_{j_1}^{\sigma_1} \cdots u_{j_k}^{\sigma_k}, \quad (G_k)_{\sigma, j} \in \mathbb{C}, \quad \begin{matrix} \sigma := (\sigma_1, \dots, \sigma_k) \\ j := (j_1, \dots, j_k) \end{matrix} \tag{3.9}$$

with symmetric coefficients  $(G_k)_{\sigma, j}$ , i.e. for any  $\rho \in \mathfrak{S}_k$  one has  $(G_k)_{\sigma, j} = (G_k)_{\sigma \circ \rho, j \circ \rho}$ .

(ii) If  $G_k \in \mathcal{L}_k$  then  $G_k^{>N}$  denotes the element of  $\mathcal{L}_k$  defined by

$$(G_k^{>N})_{\sigma, j} := \begin{cases} (G_k)_{\sigma, j}, & \text{if } \mu_2(j) > N, \\ 0, & \text{else.} \end{cases} \tag{3.10}$$

We set  $G_k^{\leq N} := G_k - G_k^{>N}$ .

**Remark 3.2** Consider the Hamiltonian  $H$  in (1.20) and its Taylor expansion in (1.24). One can note that the Hamiltonians  $H_k$  in (1.26) belong to the class  $\mathcal{L}_k$ . This follows from the fact that, without loss of generality, one can substitute the Hamiltonian  $H_k$  with its symmetrization.

We also need the following definition.

**Definition 3.3** Consider the Hamiltonian  $Z_2$  in (1.25) and  $G_k \in \mathcal{L}_k$ .

- (Adjoint action). We define the adjoint action  $\text{ad}_{Z_2} G_k$  in  $\mathcal{L}_k$  by

$$(\text{ad}_{Z_2} G_k)_{\sigma, j} := \left( i \sum_{i=1}^k \sigma_i \omega_{j_i} \right) (G_k)_{\sigma, j}. \tag{3.11}$$

- (Resonant Hamiltonian). We define  $G_k^{\text{res}} \in \mathcal{L}_j$  by

$$(G_k^{\text{res}})_{\sigma, j} := (G_k)_{\sigma, j}, \quad \text{when } \sum_{i=1}^k \sigma_i \omega_{j_i} = 0$$

and  $(G_k^{\text{res}})_{\sigma, j} = 0$  otherwise.

- We define  $G_k^{(+1)} \in \mathcal{L}_k$  by

$$(G_k^{(+1)})_{\sigma, j} := (G_k)_{\sigma, j}, \quad \text{when } \exists i, p = 1, \dots, k \text{ s.t.} \\ \mu_1(j) = |j_i|, \quad \mu_2(j) = |j_p| \quad \text{and } \sigma_i \sigma_p = +1.$$

We define  $G_k^{(-1)} := G_k - G_k^{(+1)}$ .

**Remark 3.4** Notice that, in view of Proposition 2.2, the resonant Hamiltonians given in Definition 3.3 must be supported on indices  $\sigma \in \{-1, 1\}^k, j \in \mathbb{Z}^{kd}$  which are resonant according to Definition 2.1. We remark that  $(G_k)^{\text{res}} \equiv 0$  if  $k$  is odd.

In the following lemma we collect some properties of the Hamiltonians in Definition 3.1.

**Lemma 3.5** Let  $N \geq 1, 0 \leq \delta_i < 1, q_i \in \mathbb{R}, k_i \geq 3$ , consider  $G_{k_i}^i(u)$  in  $\mathcal{L}_{k_i}$  for  $i = 1, 2$ . Assume that the coefficients  $(G_{k_i}^i)_{\sigma, j}$  satisfy

$$|(G_{k_i}^i)_{\sigma, j}| \leq C_i N^{\delta_i} \mu_3(j)^{\beta_i} \mu_1(j)^{-q_i}, \quad \forall \sigma \in \{-1, +1\}^k, j \in \mathbb{Z}^{kd}, \tag{3.12}$$

for some  $\beta_i > 0$  and  $C_i > 0, i = 1, 2$ .

(i) **(Estimates on Sobolev spaces)** Set  $k = k_i$ ,  $\delta = \delta_i$ ,  $q = q_i$ ,  $\beta = \beta_i$ ,  $C = C_i$  and  $G_{k_i}^i = G_k$  for  $i = 1, 2$ . There is  $s_0 = s_0(\beta, d)$  such that for  $s \geq s_0$ ,  $G_k$  defines naturally a smooth function from  $H^s(\mathbb{T}^d)$  to  $\mathbb{R}$ . In particular one has the following estimates:

$$|G_k(u)| \lesssim_s C N^\delta \|u\|_{H^s}^k, \tag{3.13}$$

$$\|X_{G_k}(u)\|_{H^{s+q}} \lesssim_s C N^\delta \|u\|_{H^s}^{k-1}, \tag{3.14}$$

$$\|X_{G_k^{>N}}(u)\|_{H^s} \lesssim_s C N^{-q+\delta} \|u\|_{H^s}^{k-1}, \tag{3.15}$$

for any  $u \in H^s(\mathbb{T}^d)$ .

(ii) **(Poisson bracket)** The Poisson bracket between  $G_{k_1}^1$  and  $G_{k_2}^2$  is an element of  $\mathcal{L}_{k_1+k_2-2}$  and it verifies the estimate

$$|(\{G_{k_1}^1, G_{k_2}^2\})_{\sigma,j}| \lesssim_s C_1 C_2 N^{\delta_1+\delta_2} \mu_3^{\beta_1+\beta_2} \mu_1(j)^{-\min\{q_1, q_2\}}, \tag{3.16}$$

for any  $\sigma \in \{+1, -1\}^{k_1+k_2-1}$  and  $j \in \mathbb{Z}^{d(k_1+k_2-2)}$ .

**Proof** We prove item (i). Concerning the proof of (3.13) it is sufficient to give the proof in the case  $q = 0$ . For convenience, without loss of generality, we assume  $C_i = 1$ ,  $i = 1, 2$ . We have

$$\begin{aligned} |G_k(u)| &\leq k! \sum_{\substack{j_1, \dots, j_k \in \mathbb{Z}^d \\ |j_1| \geq |j_2| \geq |j_3| \geq \dots \geq |j_k|}} |(G_k)_{\sigma,k}| |u_{j_1}^{\sigma_1}| \dots |u_{j_k}^{\sigma_k}| \\ &\lesssim_k N^\delta \sum_{j_3 \in \mathbb{Z}^d} |j_3|^\beta |u_{j_3}^{\sigma_3}| \prod_{3 \neq i=1}^k \sum_{j_i \in \mathbb{Z}^d} |u_{j_i}^{\sigma_i}| \lesssim_{k,\epsilon} N^\delta \|u\|_{H^{d/2+\beta+\epsilon}} \|u\|_{H^{d/2+\epsilon}}^{k-1}, \end{aligned}$$

for any  $\epsilon > 0$ , we proved the (3.13) with  $s_0 = d/2 + \epsilon + \beta$ .

We now prove (3.14). Since the coefficients of  $G_k$  are symmetric, we have

$$\partial_{\bar{u}_n} G_k(u) = k \sum_{\sigma_1 j_1 + \dots + \sigma_{k-1} j_{k-1} = n} (G_k)_{(\sigma,-1),(j,n)} u_{j_1}^{\sigma_1} \dots u_{j_{r-1}}^{\sigma_{r-1}}$$

Therefore, we have

$$\begin{aligned} \langle n \rangle^{s+q} |\partial_{\bar{u}_n} G_k(u)| &\leq k! \sum_{\substack{\sigma_1 j_1 + \dots + \sigma_{k-1} j_{k-1} = n \\ |j_1| \geq \dots \geq |j_{k-1}|}} |(G_k)_{(\sigma,-1),(j,n)}| |u_{j_1}^{\sigma_1}| \dots |u_{j_{r-1}}^{\sigma_{r-1}}| \langle n \rangle^{s+q} \\ &\stackrel{(3.12)}{\lesssim} N^\delta \sum_{\substack{\sigma_1 j_1 + \dots + \sigma_{k-1} j_{k-1} = n \\ |j_1| \geq \dots \geq |j_{k-1}|}} \mu_3(j, n)^\beta \mu_1(j, n)^{-q} |u_{j_1}^{\sigma_1}| \dots |u_{j_{r-1}}^{\sigma_{r-1}}| \langle n \rangle^{s+q}. \end{aligned}$$

We note that in the last sum above, we have  $\langle n \rangle \lesssim \langle j_1 \rangle$ ,  $\mu_1(j, n) \geq \langle j_1 \rangle$  and  $\mu_3(j, n) \leq \langle j_2 \rangle$ . As a consequence, we deduce that

$$\begin{aligned} \langle n \rangle^{s+q} |\partial_{\bar{u}_n} G_k(u)| &\lesssim_s N^\delta \sum_{\substack{\sigma_1 j_1 + \dots + \sigma_{k-1} j_{k-1} = n \\ |j_1| \geq \dots \geq |j_{k-1}|}} \langle j_1 \rangle^s \langle j_2 \rangle^\beta |u_{j_1}^{\sigma_1}| \dots |u_{j_{r-1}}^{\sigma_{r-1}}| \\ &\lesssim_s N^\delta \sum_{j_1 + \dots + j_{k-1} = n} \langle j_1 \rangle^s \langle j_2 \rangle^\beta |u_{j_1}| \dots |u_{j_{r-1}}|. \end{aligned}$$

Consequently, applying the Young convolutional inequality, we get

$$\begin{aligned} \|X_{G_k}(u)\|_{H^{s+q}} &= \|(\langle n \rangle^{s+q} |\partial_{\bar{u}_n} G_k(u)|)_{n \in \mathbb{Z}^d}\|_{\ell^2} \\ &\lesssim_s N^\delta \|u\|_{H^s} \left( \sum_{j \in \mathbb{Z}^d} \langle j \rangle^\beta |u_j| \right) \left( \sum_{j \in \mathbb{Z}^d} |u_j| \right)^{k-3} \\ &\lesssim_s N^\delta \|u\|_{H^s}^{k-1}. \end{aligned}$$

The proof of (3.15) follows the same lines. The proof of item (ii) of the lemma is a direct consequence of the previous computations, definition (1.23) and the momentum condition.  $\square$

We are in position to prove the main Birkhoff result.

**Proof of Theorem 2** In the case  $d = 2$  we perform two steps of Birkhoff normal form procedure, see Lemmata 3.8, 3.12. The case  $d = 3$  is slightly different. Indeed, due to the estimates on the small divisors given in Proposition 2.2, we can note that the Hamiltonian in (3.24) has already the form (3.4) since the coefficients of the Hamiltonians  $\tilde{K}_k$  (see (3.25)) do not decay anymore in the largest index  $\mu_1(j)$ . The proof of Theorem 2 is then concluded after just one step of Birkhoff normal form.

*Step 1 if  $d = 2$  or  $d = 3$ .* We have the following Lemma.

**Lemma 3.6** (Homological equation 1) *Let  $q_d = 3 - d$  for  $d = 2, 3$ . For any  $N \geq 1$  and  $\delta > 0$  there exist multilinear Hamiltonians  $\chi_k^{(1)}$ ,  $k = n, \dots, 2n - 3$  in the class  $\mathcal{L}_k$  with coefficients  $(\chi_k^{(1)})_{\sigma,j}$  satisfying*

$$|(\chi_k^{(1)})_{\sigma,j}| \lesssim_\delta N^\delta \mu_3(j)^\beta \mu_1(j)^{-q_d}, \tag{3.17}$$

such that (recall Definition 3.3)

$$\{\chi_k^{(1)}, Z_2\} + H_k = Z_k + H_k^{>N}, \quad k = n, \dots, 2n - 3, \tag{3.18}$$

where  $Z_2, H_k$  are given in (1.25), (1.26) and  $Z_k$  is the resonant Hamiltonian defined as

$$Z_k := (H_k^{\leq N})^{\text{res}}, \quad k = n, \dots, 2n - 3. \tag{3.19}$$

Moreover  $Z_k$  belongs to  $\mathcal{L}_k$  and has coefficients satisfying (3.5).

**Proof** Consider the Hamiltonians  $H_k$  in (1.26) with coefficients satisfying (1.27). Recalling Definition 3.1 we write

$$H_k = Z_k + (H_k^{\leq N} - Z_k) + H_k^{>N}, \quad k = n, \dots, r - 1,$$

with  $Z_k$  as in (3.19). We define

$$\chi_k^{(1)} := (\text{ad}_{Z_2})^{-1} \left[ H_k^{\leq N} - Z_k \right], \quad k = n, \dots, 2n - 3, \tag{3.20}$$

where  $\text{ad}_{Z_2}$  is given by Definition 3.3. In particular (recall formula (3.11)) their coefficients have the form

$$(\chi_k^{(1)})_{\sigma,j} := (H_k)_{\sigma,j} \left( i \sum_{i=1}^k \sigma_i \omega_{j_i} \right)^{-1} \tag{3.21}$$

for indices  $\sigma \in \{-1, +1\}^k, j \in (\mathbb{Z}^d)^k$  such that

$$\sum_{i=1}^k \sigma_i j_i = 0, \quad \mu_2(j) \leq N \quad \text{and} \quad \sum_{i=1}^k \sigma_i \omega_{j_i} \neq 0.$$

By (1.27) and Proposition 2.2 (with  $d = 2, 3$ ) we deduce the bound (3.17) for some  $\beta > 0$ . The resonant Hamiltonians  $Z_k$  in (3.19) have the form (3.5). One can check by an explicit computation that Eq. (3.18) is verified.  $\square$

We shall use the Hamiltonians  $\chi_k^{(1)}$  given by Lemma 3.6 to generate a symplectic change of coordinates.

**Lemma 3.7** *Let us define*

$$\chi^{(1)} := \sum_{k=n}^{2n-3} \chi_k^{(1)}. \tag{3.22}$$

There is  $s_0 = s_0(d, r)$  such that for any  $\delta > 0$ , for any  $N \geq 1$  and any  $s \geq s_0$ , if  $\varepsilon_0 \lesssim_{s,\delta} N^{-\delta}$ , then the problem

$$\begin{cases} \partial_\tau Z(\tau) = X_{\chi^{(1)}}(Z(\tau)) \\ Z(0) = U = \begin{bmatrix} u \\ \bar{u} \end{bmatrix}, \quad u \in B_s(0, \varepsilon_0) \end{cases} \tag{3.23}$$

has a unique solution  $Z(\tau) = \Phi_{\chi^{(1)}}^\tau(u)$  belonging to  $C^k([-1, 1]; H^s(\mathbb{T}^d))$  for any  $k \in \mathbb{N}$ . Moreover the map  $\Phi_{\chi^{(1)}}^\tau : B_s(0, \varepsilon_0) \rightarrow H^s(\mathbb{T}^d)$  is symplectic. The flow map  $\Phi_{\chi^{(1)}}^\tau$  and its inverse  $\Phi_{\chi^{(1)}}^{-\tau}$  satisfy

$$\begin{aligned} \sup_{\tau \in [0,1]} \|\Phi_{\chi^{(1)}}^{\pm\tau}(u) - u\|_{H^s} &\lesssim_{s,\delta} N^\delta \|u\|_{H^s}^{n-1}, \\ \sup_{\tau \in [0,1]} \|d\Phi_{\chi^{(1)}}^{\pm\tau}(u)[\cdot]\|_{\mathcal{L}(H^s; H^s)} &\leq 2. \end{aligned}$$

**Proof** By estimate (3.17) and Lemma 3.5 we have that the vector field  $X_{\chi^{(1)}}$  is a bounded operator on  $H^s(\mathbb{T}^d)$ . Hence the flow  $\Phi_{\chi^{(1)}}^\tau$  is well-posed by standard theory of Banach space ODE. The estimates of the map and its differential follow by using the equation in (3.23), the fact that  $\chi^{(1)}$  is multilinear and the smallness condition on  $\varepsilon_0$ . Finally the map is symplectic since it is generated by a Hamiltonian vector field.  $\square$

We now study how changes the Hamiltonian  $H$  in (1.24) under the map  $\Phi_{\chi^{(1)}}^\tau$ .

**Lemma 3.8** (The new Hamiltonian 1) *There is  $s_0 = s_0(d, r)$  such that for any  $N \geq 1, \delta > 0$  and any  $s \geq s_0$ , if  $\varepsilon_0 \lesssim_{s,\delta} N^{-\delta}$  then we have that*

$$H \circ \Phi_{\chi^{(1)}} = Z_2 + \sum_{k=n}^{2n-3} Z_k + \tilde{K}^{>N} + \sum_{k=2n-2}^{r-1} \tilde{K}_k + \mathcal{R}_r \tag{3.24}$$

where

- $\Phi_{\chi^{(1)}} := (\Phi_{\chi^{(1)}}^\tau)_{|\tau=1}$  is the flow map given by Lemma 3.7;
- the resonant Hamiltonians  $Z_k$  are defined in (3.19);



- $\tilde{K}_k$  are in  $\mathcal{L}_k$  with coefficients  $(\tilde{K}_k)_{\sigma,j}$  satisfying

$$|(\tilde{K}_k)_{\sigma,j}| \lesssim_{\delta} N^{\delta} \mu_3(j)^{\beta} \mu_1(j)^{-q_d}, \quad k = 2n - 2, \dots, r - 1, \tag{3.25}$$

with  $q_d = 3 - d$  for  $d = 2, 3$ ;

- the Hamiltonian  $\tilde{K}^{>N}$  and the remainder  $\mathcal{R}_r$  satisfy

$$\|X_{\tilde{K}^{>N}}(u)\|_{H^s} \lesssim_{s,\delta} N^{-1} \|u\|_{H^s}^{n-1}, \tag{3.26}$$

$$\|X_{\mathcal{R}_r}(u)\|_{H^s} \lesssim_{s,\delta} N^{\delta} \|u\|_{H^s}^{r-1}, \quad \forall u \in B_S(0, 2\varepsilon_0). \tag{3.27}$$

**Proof** Fix  $\delta > 0$  and  $\varepsilon_0 N^{\delta}$  small enough. We apply Lemma 3.7 with  $\delta \rightsquigarrow \delta'$  to be chosen small enough with respect to  $\delta$  we have fixed (which ensures us that the smallness condition  $\varepsilon_0 N^{\delta'} \lesssim_{s,\delta'} 1$  of Lemma 3.7 is fulfilled). Let  $\Phi_{\chi^{(1)}}^{\tau}$  be the flow at time  $\tau$  of the Hamiltonian  $\chi^{(1)}$ . We note that

$$\partial_{\tau} H \circ \Phi_{\chi^{(1)}}^{\tau} = dH(z)[X_{\chi^{(1)}}(z)]|_{z=\Phi_{\chi^{(1)}}^{\tau}} \stackrel{(1.22),(1.23)}{=} \{\chi^{(1)}, H\} \circ \Phi_{\chi^{(1)}}^{\tau}.$$

Then, for  $L \geq 2$ , we get the Lie series expansion

$$H \circ \Phi_{\chi^{(1)}} = H + \{\chi^{(1)}, H\} + \sum_{p=2}^L \frac{1}{p!} \text{ad}_{\chi^{(1)}}^p [H] + \frac{1}{L!} \int_0^1 (1 - \tau)^L \text{ad}_{\chi^{(1)}}^{L+1} [H] \circ \Phi_{\chi^{(1)}}^{\tau} d\tau$$

where  $\text{ad}_{\chi^{(1)}}^p$  is defined recursively as

$$\text{ad}_{\chi^{(1)}} [H] := \{\chi^{(1)}, H\}, \quad \text{ad}_{\chi^{(1)}}^p [H] := \{\chi^{(1)}, \text{ad}_{\chi^{(1)}}^{p-1} [H]\}, \quad p \geq 2. \tag{3.28}$$

Recalling the Taylor expansion of the Hamiltonian  $H$  in (1.24) we obtain

$$\begin{aligned} H \circ \Phi_{\chi^{(1)}} &= Z_2 + \sum_{k=n}^{2n-3} \left( H_k + \{\chi_k^{(1)}, Z_2\} \right) + \sum_{k=2n-2}^{r-1} H_k \\ &\quad + \sum_{p=2}^L \frac{1}{p!} \text{ad}_{\chi^{(1)}}^p [Z_2] + \sum_{j=n}^{r-1} \sum_{p=1}^L \frac{1}{p!} \text{ad}_{\chi^{(1)}}^p [H_j] \end{aligned} \tag{3.29}$$

$$+ \frac{1}{L!} \int_0^1 (1 - \tau)^L \text{ad}_{\chi^{(1)}}^{L+1} [Z_2 + \sum_{j=n}^{r-1} H_j] \circ \Phi_{\chi^{(1)}}^{\tau} d\tau \tag{3.30}$$

$$+ R_r \circ \Phi_{\chi^{(1)}}. \tag{3.31}$$

We study each summand separately. First of all, by definition of  $\chi_k^{(1)}$  (see (3.18) in Lemma 3.6), we deduce that

$$\sum_{k=n}^{2n-3} (H_k + \{\chi_k^{(1)}, Z_2\}) = \sum_{k=n}^{2n-3} Z_k + \tilde{K}^{>N}, \quad \tilde{K}^{>N} := \sum_{k=n}^{2n-3} H_k^{>N}. \tag{3.32}$$

One can check, using Lemma 3.5 (see (3.15)), that  $\tilde{K}^{>N}$  satisfies (3.26). Consider now the term in (3.29). By definition of  $\chi^{(1)}$  (see (3.18) and (3.22)), we get, for  $p = 2, \dots, L$ ,

$$\text{ad}_{\chi^{(1)}}^p [Z_2] = \text{ad}_{\chi^{(1)}}^{p-1} [\{\chi^{(1)}, Z_2\}] \stackrel{(3.32)}{=} \text{ad}_{\chi^{(1)}}^{p-1} \left[ \sum_{k=n}^{2n-3} (Z_k - H_k^{\leq N}) \right].$$

Therefore, by Lemma 3.5-(ii) and recalling (3.28), we get

$$(3.29) = \sum_{k=2n-2}^{L(2n-3)+r-1-2L} \tilde{K}_k$$

where  $\tilde{K}_k$  are  $k$ -homogeneous Hamiltonians in  $\mathcal{L}_k$ . In particular, by (3.16), (3.17) and (1.27) (with  $\delta \rightsquigarrow \delta'$ ), we have

$$|(\tilde{K}_k)_{\sigma,j}| \lesssim_{\delta'} N^{L\delta'} \mu_3(j)^\beta \mu_1(j)^{-q_d}$$

for some  $\beta > 0$  depending only on  $d, n$ . This implies the estimates (3.25) taking  $L\delta' \leq \delta$ , where  $L$  will be fixed later. Then formula (3.24) follows by setting

$$\mathcal{R}_r := \sum_{k=r}^{L(2n-3)+r-1-2L} \tilde{K}_k + (3.30) + (3.31). \tag{3.33}$$

The estimate (3.27) holds true for  $X_{\tilde{K}_k}$  with  $k = r, \dots, L(2n - 3) + r - 1 - 2L$ , thanks to (3.25) and Lemma 3.5. It remains to study the terms appearing in (3.30), (3.31). We start with the remainder in (3.31). We note that

$$X_{R_r \circ \Phi}(u) = (d\Phi_{\chi^{(1)}})^{-1}(u) \left[ X_{R_r}(\Phi_{\chi^{(1)}}(u)) \right].$$

We obtain the estimate (3.27) on the vector field  $X_{R_r \circ \Phi}$  by using (1.28) and Lemma 3.7. In order to estimate the term in (3.30) we reason as follows. First notice that

$$\text{ad}_{\chi^{(1)}}^{L+1}[Z_2 + H_j] \stackrel{(3.32)}{=} \text{ad}_{\chi^{(1)}}^L \left[ \sum_{k=n}^{2n-3} (Z_k - H_k^{\leq N}) \right] + \text{ad}_{\chi^{(1)}}^{L+1}[H_j] := \mathcal{Q}_j$$

with  $j = n, \dots, r - 1$ . Using Lemma 3.5 we deduce that

$$\|X_{\mathcal{Q}_j}(u)\|_{H^s} \lesssim_{\delta'} N^{(L+1)\delta'} \|u\|_{H^s}^{(Ln+n-2L)-1}.$$

We choose  $L = 9$  which implies  $Ln + n - 2L \geq r$  since  $r \leq 4n$ . Notice also that all the summand in (3.30) are of the form

$$\int_0^1 (1 - \tau)^L \mathcal{Q}_j \circ \Phi_{\chi^{(1)}}^\tau d\tau.$$

Then we can estimate their vector fields by reasoning as done for the Hamiltonian  $R_r \circ \Phi_{\chi^{(1)}}$ . This concludes the proof. □

**Remark 3.9** (Case  $d = 3$ ) We remark that Theorem 2 for  $d = 3$  follows by Lemmata 3.6, 3.7, 3.8, by setting  $\tau^{(1)} := \Phi_{\chi^{(1)}}$  and recalling that (see (3.1))  $M_{d,n} = 2n - 2$  for  $d = 3$ .

*Step 2 if  $d = 2$ .* This step is performed only in the case  $d = 2$ . Consider the Hamiltonian in (3.24). Our aim is to reduce in Birkhoff normal form all the Hamiltonians  $\tilde{K}_k$  of homogeneity  $k = 2n - 2, \dots, M_{2,n} - 1$  where  $M_{2,n}$  is given in (3.1). We follow the same strategy adopted in the previous step.

**Lemma 3.10** (Homological equation 2) *Let  $N \geq 1, \delta > 0$  and consider the Hamiltonian in (3.24). There exist multilinear Hamiltonians  $\chi_k^{(2)}, k = 2n - 2, \dots, M_{2,n} - 1$  in the class  $\mathcal{L}_k$ , with coefficients satisfying*

$$|(\chi_k^{(2)})_{\sigma,j}| \lesssim_\delta N^\delta \mu_3(j)^\beta, \tag{3.34}$$

for some  $\beta > 0$ , such that

$$\{\chi_k^{(2)}, Z_2\} + \tilde{K}_k = Z_k + \tilde{K}_k^{>N}, \quad k = 2n - 2, \dots, M_{2,n} - 1, \tag{3.35}$$

where  $\tilde{K}_k$  are given in Lemma 3.8 and  $Z_k$  is the resonant Hamiltonian defined as

$$Z_k := (\tilde{K}_k^{\leq N})^{\text{res}}, \quad k = 2n - 2, \dots, M_{2,n} - 1. \tag{3.36}$$

Moreover  $Z_k$  belongs to  $\mathcal{L}_k$  and has coefficients satisfying (3.5).

**Proof** Recalling Definitions 3.1, 3.3, we write

$$\tilde{K}_k = Z_k + (\tilde{K}_k^{\leq N} - Z_k) + \tilde{K}_k^{>N},$$

with  $Z_k$  as in (3.36), and we define

$$\chi_k^{(2)} := (\text{ad}_{Z_2})^{-1} [\tilde{K}_k^{\leq N} - Z_k], \quad k = 2n - 2, \dots, M_{2,n} - 1. \tag{3.37}$$

The Hamiltonians  $\chi_k^{(2)}$  have the form (3.9) with coefficients

$$(\chi_k^{(2)})_{\sigma,j} := (\tilde{K}_k)_{\sigma,j} \left( i \sum_{i=1}^k \sigma_i \omega_{j_i} \right)^{-1} \tag{3.38}$$

for indices  $\sigma \in \{-1, +1\}^k, j \in (\mathbb{Z}^d)^k$  such that

$$\sum_{i=1}^k \sigma_i j_i = 0, \quad \mu_2(j) \leq N \quad \text{and} \quad \sum_{i=1}^k \sigma_i \omega_{j_i} \neq 0.$$

Recalling that we are in the case  $d = 2$ , by (3.25) and Proposition 2.2 we deduce (3.34). The resonant Hamiltonians  $Z_k$  in (3.36) have the form (3.5). The (3.35) follows by an explicit computation. □

**Lemma 3.11** *Let us define*

$$\chi^{(2)} := \sum_{k=2n-2}^{M_{2,n}-1} \chi_k^{(2)}. \tag{3.39}$$

There is  $s_0 = s_0(d, r)$  such that for any  $\delta > 0$ , for any  $N \geq 1$  and any  $s \geq s_0$ , if  $\varepsilon_0 \lesssim_{s,\delta} N^{-\delta}$ , then the problem

$$\begin{cases} \partial_\tau Z(\tau) = X_{\chi^{(2)}}(Z(\tau)) \\ Z(0) = U = \begin{bmatrix} u \\ \bar{u} \end{bmatrix}, \quad u \in B_s(0, \varepsilon_0) \end{cases}$$

has a unique solution  $Z(\tau) = \Phi_{\chi^{(2)}}^\tau(u)$  belonging to  $C^k([-1, 1]; H^s(\mathbb{T}^d))$  for any  $k \in \mathbb{N}$ . Moreover the map  $\Phi_{\chi^{(2)}}^\tau : B_s(0, \varepsilon_0) \rightarrow H^s(\mathbb{T}^d)$  is symplectic. The flow map  $\Phi_{\chi^{(2)}}^\tau$ , and its inverse  $\Phi_{\chi^{(2)}}^{-\tau}$ , satisfy

$$\begin{aligned} \sup_{\tau \in [0,1]} \|\Phi_{\chi^{(2)}}^{\pm\tau}(u) - u\|_{H^s} &\lesssim_{s,\delta} N^\delta \|u\|_{H^s}^{n-1}, \\ \sup_{\tau \in [0,1]} \|d\Phi_{\chi^{(2)}}^{\pm\tau}(u)[\cdot]\|_{\mathcal{L}(H^s; H^s)} &\leq 2. \end{aligned}$$

**Proof** It follows reasoning as in the proof of Lemma 3.7. □

We have the following.

**Lemma 3.12** (The new Hamiltonian 2) *There is  $s_0 = s_0(d, r)$  such that for any  $N \geq 1, \delta > 0$  and any  $s \geq s_0$ , if  $\varepsilon_0 \lesssim_{s, \delta} N^{-\delta}$  then we have that  $H \circ \Phi_{\chi^{(1)}} \circ \Phi_{\chi^{(2)}}$  has the form (3.4) and satisfies items (i), (ii), (iii) of Theorem 2.*

**Proof** We fix  $\delta > 0$  and we apply Lemmata 3.8, 3.10 with  $\delta \rightsquigarrow \delta'$  with  $\delta'$  to be chosen small enough with respect to  $\delta$  fixed here.

Reasoning as in the previous step we have (recall (3.1), (3.28) and (3.24))

$$H \circ \Phi_{\chi^{(1)}} \circ \Phi_{\chi^{(2)}} = Z_2 + \sum_{k=n}^{2n-3} Z_k + \sum_{k=2n-2}^{M_{2,d}-1} \left( \tilde{K}_k + \{\chi_k^{(2)}, Z_2\} \right) + \sum_{k=M_{2,n}}^{r-1} \tilde{K}_k \tag{3.40}$$

$$+ \tilde{K}^{>N} \circ \Phi_{\chi^{(2)}} \tag{3.41}$$

$$+ \sum_{p=2}^L \frac{1}{p!} \text{ad}_{\chi^{(2)}}^p [Z_2] + \sum_{p=1}^L \frac{1}{p!} \text{ad}_{\chi^{(2)}}^p \left[ \sum_{k=n}^{2n-3} Z_k + \sum_{k=2n-2}^{r-1} \tilde{K}_k \right] \tag{3.42}$$

$$+ \mathcal{R}_r \circ \Phi_{\chi^{(2)}} + \frac{1}{L!} \int_0^1 (1 - \tau)^L \text{ad}_{\chi^{(2)}}^{L+1} [Z_2] \circ \Phi_{\chi^{(2)}}^\tau d\tau \tag{3.43}$$

$$+ \frac{1}{L!} \int_0^1 (1 - \tau)^L \text{ad}_{\chi^{(2)}}^{L+1} \left[ \sum_{k=n}^{2n-3} Z_k + \sum_{k=2n-2}^{r-1} \tilde{K}_k \right] \circ \Phi_{\chi^{(2)}}^\tau d\tau, \tag{3.44}$$

where  $\Phi_{\chi^{(2)}}^\tau, \tau \in [0, 1]$ , is the flow at time  $\tau$  of the Hamiltonian  $\chi^{(2)}$ . We study each summand separately. First of all, thanks to (3.35), we deduce that

$$\sum_{k=2n-2}^{M_{2,n}-1} \left( \tilde{K}_k + \{\chi_k^{(2)}, Z_2\} \right) = \sum_{k=2n-2}^{M_{2,n}-1} Z_k + \tilde{K}_+^{>N}, \quad \tilde{K}_+^{>N} := \sum_{k=2n-2}^{M_{2,n}-1} \tilde{K}_k^{>N}. \tag{3.45}$$

One can check, using Lemma 3.5, that  $\tilde{K}_+^{>N}$  satisfies

$$\|X_{\tilde{K}_+^{>N}}\|_{H^s} \lesssim_{s, \delta} N^{-1+\delta'} \|u\|_{H^s}^{2n-3}. \tag{3.46}$$

Consider now the terms in (3.42). First of all notice that we have

$$\text{ad}_{\chi^{(2)}}^p [Z_2] \stackrel{(3.45)}{=} \sum_{k=2n-2}^{M_{2,n}-1} \text{ad}_{\chi^{(2)}}^{p-1} [Z_k - \tilde{K}_k^{\leq N}], \quad p = 2, \dots, L.$$

The Hamiltonian above has a homogeneity at least of degree  $4n - 6$  which actually is larger or equal to  $M_{2,n}$  (see (3.1)). The terms with lowest homogeneity in the sum (3.42) have degree exactly  $M_{2,n}$  and come from the term  $\text{ad}_{\chi^{(2)}} \left[ \sum_{k=n}^{2n-3} Z_k \right]$  recalling that (see Remark 3.4) if  $n$  is odd then  $Z_n \equiv 0$ . Then, by (3.34), (3.25) and Lemma 3.5-(ii), we get

$$(3.42) = \sum_{k=M_{2,n}}^{L(M_{2,n}-1)+r-1-2L} \tilde{K}_k^+ \tag{3.47}$$

where  $\tilde{K}_k^+$  are  $k$ -homogeneous Hamiltonians of the form (3.6) with coefficients satisfying

$$|(\tilde{K}_k^+)_{\sigma, j}| \lesssim_{\delta'} N^{(L+1)\delta'} \mu_3(j)^\beta, \tag{3.47}$$

for some  $\beta > 0$ . By the discussion above, using formulæ (3.40)–(3.44), we obtain that the Hamiltonian  $H \circ \Phi_{\chi^{(1)}} \circ \Phi_{\chi^{(2)}}$  has the form (3.4) with (recall (3.19), (3.36), (3.41), (3.45))

$$Z_k^{\leq N} := Z_k, \quad k = n, \dots, M_{2,n-1}, \quad K_k := \tilde{K}_k + \tilde{K}_k^+, \quad k = M_{2,n}, \dots, r - 1, \quad (3.48)$$

$$K^{>N} := \tilde{K}^{>N} \circ \Phi_{\chi^{(2)}} + \tilde{K}_+^{>N} \quad (3.49)$$

and with remainder  $\tilde{R}_r$  defined as

$$\tilde{R}_r := \sum_{k=r}^{L(M_{2,n-1})+r-1-2L} \tilde{K}_k^+ + (3.43) + (3.44). \quad (3.50)$$

Recalling (3.19), (3.36) and the estimates (1.27), (3.25) we have that  $Z_k^{\leq N}$  in (3.48) satisfies the condition of item (i) of Theorem 2. Similarly  $K_k$  in (3.48) satisfies (3.6) thanks to (3.25) and (3.47) as long as  $\delta'$  is sufficiently small. The remainder  $K^{>N}$  in (3.49) satisfies the bound (3.7) using (3.46), (3.26) and Lemma 3.5-(i). It remains to show that the remainder defined in (3.50) satisfies the estimate (3.8). The claim follows for the terms  $\tilde{K}_k^+$  for  $k = r, \dots, L(M_{2,n} - 1) + r - 1 - 2L$  by using (3.47) and Lemma 3.5. For the remainder in (3.43), (3.44) one can reason following almost word by word the proof of the estimate of the vector field of  $\mathcal{R}_r$  in (3.33) in the previous step. In this case we choose  $L + 1 = 8$  which implies  $L + 1 \geq (r + n)/(2n - 4)$ .  $\square$

Theorem 2 follows by Lemmata 3.8, 3.12 setting  $\tau^{(1)} := \Phi_{\chi^{(1)}} \circ \Phi_{\chi^{(2)}}$ . The bound (3.3) follows by Lemmata 3.7 and 3.11.

### 4 The Modified Energy Step

In this section we construct a modified energy which is an approximate constant of motion for the Hamiltonian system of  $H \circ \tau^{(1)}$  in (3.4), when  $d = 2, 3$ , and for the Hamiltonian  $H$  in (1.24) when  $d \geq 4$ . For compactness we shall write, for  $s \in \mathbb{R}$ ,

$$N_s(u) := \|u\|_{H^s}^2 = \sum_{j \in \mathbb{Z}^2} \langle j \rangle^{2s} |u_j|^2, \quad (4.1)$$

for  $u \in H^s(\mathbb{T}^2; \mathbb{C})$ . For  $d \geq 2$  and  $n \in \mathbb{N}$  we define (recall (3.1))

$$\tilde{M}_{d,n} := \begin{cases} M_{d,n} + n - 1 & n \text{ odd} \\ M_{d,n} + n - 2 & n \text{ even.} \end{cases} \quad (4.2)$$

**Proposition 4.1** *There exists  $\beta = \beta(d, n) > 0$  such that for any  $\delta > 0$ , any  $N \geq N_1 > 1$  ( $N = N_1$  if  $d \geq 4$ ) and any  $s \geq \tilde{s}_0$ , for some  $\tilde{s}_0 = \tilde{s}_0(\beta) > 0$ , if  $\varepsilon_0 \lesssim_{s,\delta} N^{-\delta}$ , there are multilinear maps  $E_k, k = M_{d,n}, \dots, \tilde{M}_{d,n} - 1$ , in the class  $\mathcal{L}_k$  such that the following holds:*

- the coefficients  $(E_k)_{\sigma,j}$  satisfies

$$|(E_k)_{\sigma,j}| \lesssim_{s,\delta} N^\delta N_1^{\kappa_d} \mu_3(j)^\beta \mu_1(j)^{2s}, \quad (4.3)$$

for  $\sigma \in \{-1, 1\}^k, j \in (\mathbb{Z}^d)^k, k = M_{d,n}, \dots, \tilde{M}_{d,n} - 1$ , where

$$\kappa_d := 0 \text{ if } d = 2, \quad \kappa_d := 1 \text{ if } d = 3, \quad \kappa_d := d - 4 \text{ if } d \geq 4. \quad (4.4)$$

- for any  $u \in B_s(0, 2\varepsilon_0)$  setting

$$E(u) := \sum_{k=M_{d,n}}^{\tilde{M}_{d,n-1}} E_k(u). \tag{4.5}$$

one has

$$\begin{aligned} |\{N_s + E, H \circ \tau^{(1)}\}| &\lesssim_{s,\delta} N_1^{\kappa_d} N^\delta (\|u\|_{H^s}^{\tilde{M}_{d,n}} + N^{-1} \|u\|_{H^s}^{M_{d,n}+n-2}) \\ &+ N_1^{-\varepsilon_d+\delta} \|u\|_{H^s}^{M_{d,n}} + N^{-\varepsilon_d+\delta} \|u\|_{H^s}^n, \end{aligned} \tag{4.6}$$

where

$$\varepsilon_d := 1, \text{ for } d = 2, 3, \text{ and } \varepsilon_d := 3, \text{ for } d \geq 4. \tag{4.7}$$

**Remark 4.2** We remark that in the proposition above we introduced a second truncation parameter  $N_1$ . This is needed in order to optimize the time of existence that we shall deduce by estimate (4.6). In Sect. 5 we shall choose  $N, N_1$  (depending on  $\varepsilon$ ) is such a way that the last two summands in the r.h.s. of (4.6) are negligible w.r.t. the first two summands. Since for  $d = 3$  the term  $\|u\|_{H^s}^n$  is larger than  $\|u\|_{H^s}^{M_{d,n}}$  it would be convenient to choose  $N \gg N_1$  to make the last summand small enough. This is possible since the factor  $N_1^{\kappa_d} N^\delta$  grows very slowly in  $N$  since  $\delta$  is arbitrary small. Note that in the case  $d \geq 4$  we need just one truncation since no preliminary Birkhoff normal form is performed. In the case  $d = 2$  one could use the same truncation  $N$  since  $\kappa_d = 0$ .

We need the following technical lemma.

**Lemma 4.3** (Energy estimate) *Let  $N \geq 1, 0 \leq \delta < 1, p \in \mathbb{N}, p \geq 3$ . Consider the Hamiltonians  $N_s$  in (4.1),  $G_p \in \mathcal{L}_p$  and write  $G_p = G_p^{(+1)} + G_p^{(-1)}$  (recall Definition 3.3). Assume also that the coefficients of  $G_p$  satisfy*

$$|(G_p^{(\eta)})_{\sigma,j}| \leq CN^\delta \mu_3(j)^\beta \mu_1(j)^{-q}, \quad \forall \sigma \in \{-1, +1\}^p, j \in \mathbb{Z}^d, \eta \in \{-1, +1\}, \tag{4.8}$$

for some  $\beta > 0, C > 0$  and  $q \geq 0$ . We have that the Hamiltonian  $Q_p^{(\eta)} := \{N_s, G_p^{(\eta)}\}, \eta \in \{-1, 1\}$ , belongs to the class  $\mathcal{L}_p$  and has coefficients satisfying

$$|(Q_p^{(\eta)})_{\sigma,j}| \lesssim_s CN^\delta \mu_3(j)^{\beta+1} \mu_1(j)^{2s} \mu_1(j)^{-q-\alpha}, \quad \alpha := \begin{cases} 1 & \text{if } \eta = -1 \\ 0 & \text{if } \eta = +1. \end{cases} \tag{4.9}$$

**Proof** Using formulæ (4.1), (1.23), (3.9) and recalling Definition 3.3 we have that the Hamiltonian  $\{N_s, G_p^{(\eta)}\}$  has coefficients

$$(Q_p^{(\eta)})_{\sigma,j} = (G_p^{(\eta)})_{\sigma,j} i \left( \sum_{i=1}^p \sigma_i \langle j_i \rangle^{2s} \right)$$

for any  $\sigma \in \{-1, +1\}^p, j \in (\mathbb{Z}^d)^p$  satisfying

$$\sum_{i=1}^p \sigma_i j_i = 0, \quad \sigma_i \sigma_k = \eta, \quad \mu_1(j) = |j_i|, \mu_2(j) = |jk|,$$

for some  $i, k = 1, \dots, p$ . Then the bound (4.9) follows by the fact that

$$|\langle j_i \rangle^{2s} + \eta \langle j_k \rangle^{2s}| \lesssim_s \begin{cases} \mu_1(j)^{2s-1} \mu_3(j) & \text{if } \eta = -1 \\ \mu_1(j)^{2s} & \text{if } \eta = +1. \end{cases}$$

and using the assumption (4.8). □

**Proof of Proposition 4.1** *Case  $d = 2, 3$ .* Consider the Hamiltonians  $K_k$  in (3.6) for  $k = M_{d,n}, \dots, \tilde{M}_{n,d} - 1$  where  $\tilde{M}_{n,d}$  is defined in (4.2). Recalling Definition 3.3 we set  $E_k := E_k^{(+1)} + E_k^{(-1)}$ , where

$$E_k^{(+1)} := (\text{ad}_{Z_2})^{-1}\{N_s, K_k^{(+1)}\}, \quad E_k^{(-1)} := (\text{ad}_{Z_2})^{-1}\{N_s, K_k^{(-1, \leq N_1)}\}, \quad (4.10)$$

for  $k = M_{d,n}, \dots, \tilde{M}_{d,n} - 1$ . Notice that formulæ in (4.10) are well-defined since  $\{N_s, K_k^{(+1)}\}$  and  $\{N_s, K_k^{(-1, \leq N_1)}\}$  are in the range of the adjoint action  $\text{ad}_{Z_2}$  thanks to Proposition 2.2. It is easy to note that  $E_k \in \mathcal{L}_k$ . Moreover, using the bounds on the coefficients  $(K_k)_{\sigma,j}$  in (3.6) and item (ii) of Proposition 2.2 (with  $\delta$  therein possibly smaller than the one fixed here), one can check that the coefficients  $(E_k^{(+1)})_{\sigma,j}$  satisfy the (4.3). By (3.6), Lemma 4.3 (in particular formula (4.9) with  $\eta = -1$ ) and item (iii) of Proposition 2.2, one gets that the coefficients  $(E_k^{(-1)})_{\sigma,j}$  satisfy the (4.3) as well. Using (4.10) we notice that

$$\{N_s, K_k\} + \{E_k, Z_2\} = \{N_s, K_k^{(-1, > N_1)}\}, \quad k = M_{d,n}, \dots, \tilde{M}_{d,n} - 1. \quad (4.11)$$

Combining Lemmata 3.5 and 4.3 we deduce

$$|\{N_s, K_k^{(-1, > N_1)}\}(u)| \lesssim_{s,\delta} N_1^{-1+\delta} \|u\|_{H^s}^k, \quad (4.12)$$

for  $s$  large enough with respect to  $\beta$ . We define the energy  $E$  as in (4.5). We are now in position to prove the estimate (4.6).

Using the expansions (3.4) and (4.5) we get

$$\{N_s + E, H \circ \tau^{(1)}\} = \left\{ N_s, Z_2 + \sum_{k=n}^{M_{d,n}-1} Z_k^{\leq N} \right\} \quad (4.13)$$

$$+ \{N_s, K^{>N}\} + \{N_s, \tilde{R}_r\} \quad (4.14)$$

$$+ \sum_{k=M_{d,n}}^{\tilde{M}_{d,n}-1} (\{N_s, K_k\} + \{E_k, Z_2\}) \quad (4.15)$$

$$+ \left\{ E, \sum_{k=n}^{M_{d,n}-1} Z_k^{\leq N} \right\} + \left\{ E, \sum_{k=M_{d,n}}^{r-1} K_k + \tilde{R}_r \right\} \quad (4.16)$$

$$+ \{E, K^{>N}\}. \quad (4.17)$$

We study each summand separately. First of all note that, by item (i) in Theorem 2 and Proposition 2.2 we deduce that the right hand side of (4.13) vanishes. Consider now the term in (4.14). Using the bounds (3.7), (3.8) and recalling (1.23) one can check that, for  $\varepsilon_0 N^\delta \lesssim_{s,\delta} 1$ ,

$$|(4.14)| \lesssim_{s,\delta} N^{-1+\delta} \|u\|_{H^s}^n + N^\delta \|u\|_{H^s}^r. \quad (4.18)$$

By (4.11) and (4.12) we deduce that

$$|(4.15)| \lesssim_{s,\delta} N_1^{-1+\delta} \|u\|_{H^s}^{M_{d,n}}. \quad (4.19)$$

By (4.3), (3.4)–(3.8), Lemma 3.5 (recall also (4.2)) we get

$$|(4.16)| \lesssim_{s,\delta} N_1^{\kappa_d} N^\delta (\|u\|_{H^s}^{\tilde{M}_{d,n}} + \|u\|_{H^s}^r),$$

$$|(4.17)| \lesssim_{s,\delta} N_1^{\kappa_d} N^{-1+\delta} \|u\|_{H^s}^{M_{d,n}+n-2}.$$

The discussion above implies the bound (4.6) using that  $r \geq \tilde{M}_{d,n}$ . This concludes the proof in the case  $d = 2, 3$ .

Case  $d \geq 4$ . In this case we consider the Hamiltonian  $H$  in (1.24). Recalling Definition 3.3 we set

$$E_k := E_k^{(+1)} + E_k^{(-1)}$$

where

$$E_k^{(+1)} := (\text{ad}_{Z_2})^{-1}\{N_s, H_k\}^{(+1)}, \quad E_k^{(-1)} := (\text{ad}_{Z_2})^{-1}\{N_s, H_k^{(-1, \leq N_1)}\}, \quad (4.20)$$

for  $k = M_{d,n}, \dots, \tilde{M}_{d,n} - 1$ . Notice that the energies  $E_k^{(+1)}, E_k^{(-1)}$  are in  $\mathcal{L}_k$  with coefficients

$$(E_k^{(+1)})_{\sigma,j} = \left( \sum_{i=1}^k \sigma_i \langle j_i \rangle^{2s} \right) \left( \sum_{i=1}^k \sigma_i \omega_{j_i} \right)^{-1} (H_k^{(+1)})_{\sigma,j}, \quad \sigma \in \{-1, +1\}^k, \quad j \in (\mathbb{Z}^d)^k,$$

and

$$(E_k^{(-1)})_{\sigma,j} = \left( \sum_{i=1}^k \sigma_i \langle j_i \rangle^{2s} \right) \left( \sum_{i=1}^k \sigma_i \omega_{j_i} \right)^{-1} (H_k^{(-1)})_{\sigma,j}, \quad \mu_2(j) \leq N_1,$$

with  $\sigma \in \{-1, +1\}^k, j \in (\mathbb{Z}^d)^k$ . Recall that in this case  $M_{d,n} = n$  (see (3.1)). Using Proposition 2.2 and reasoning as in the proof of Lemma 4.3 one can check that estimate (4.3) on the coefficients of  $E_k^{(+1)}$  and  $E_k^{(-1)}$  holds true with  $\kappa_d$  as in (4.4). Equation (4.20) implies

$$\{N_s, H_k\} + \{E_k, Z_2\} = \{N_s, H_k^{(-1, > N_1)}\}, \quad k = n, \dots, \tilde{M}_{d,n} - 1, \quad (4.21)$$

where  $\tilde{M}_{d,n} - 1 = 2n - 1$  if  $n$  odd and  $\tilde{M}_{d,n} - 1 = 2n - 2$  if  $n$  even (see (4.2)). Recall that the coefficients of the Hamiltonian  $H_k$  satisfy the bound (1.27). Therefore, combining Lemmata 4.3 and 3.5, we deduce

$$|\{N_s, H_k^{(-1, > N_1)}\}(u)| \lesssim_{s,\delta} N_1^{-3} \|u\|_{H^s}^k, \quad (4.22)$$

for  $s$  large enough with respect to  $\beta$ . Recalling (1.24) we have

$$\begin{aligned} \{N_s + E, H\} &= \{N_s, Z_2\} + \{N_s, R_r\} + \left\{ E, \sum_{k=n}^{r-1} H_k + R_r \right\} \\ &\quad + \sum_{k=n}^{\tilde{M}_{d,n}-1} (\{N_s, K_k\} + \{E_k, Z_2\}). \end{aligned}$$

One can obtain the bound (4.6) by reasoning as in the case  $d = 2, 3$ , using (4.22), (1.28). This concludes the proof. □



### 5 Proof of Theorem 1

In this section we show how to combine the results of Theorem 2 and Proposition 4.1 in order to prove Theorem 1.

Consider  $\psi_0$  and  $\psi_1$  satisfying (1.4) and let  $\psi(t, y), y \in \mathbb{T}_\nu^d$ , be the unique solution of (1.1) with initial conditions  $(\psi_0, \psi_1)$  defined for  $t \in [0, T]$  for some  $T > 0$ . By rescaling the space variable  $y$  and passing to the complex variable in (1.17) we consider the function  $u(t, x), x \in \mathbb{T}^d$  solving the Eq. (1.18). We recall that (1.18) can be written in the Hamiltonian form

$$\partial_t u = i\partial_{\bar{u}} H(u), \tag{5.1}$$

where  $H$  is the Hamiltonian function in (1.20) (see also (1.24)). We have that Theorem 1 is a consequence of the following Lemma.

**Lemma 5.1** (Main bootstrap) *There exists  $s_0 = s_0(n, d)$  such that for any  $\delta > 0, s \geq s_0$ , there exists  $\varepsilon_0 = \varepsilon_0(\delta, s)$  such that the following holds. Let  $u(t, x)$  be a solution of (5.1) with  $t \in [0, T], T > 0$  and initial condition  $u(0, x) = u_0(x) \in H^s(\mathbb{T}^d)$ . For any  $\varepsilon \in (0, \varepsilon_0)$  if*

$$\|u_0\|_{H^s} \leq \varepsilon, \quad \sup_{t \in [0, T]} \|u(t)\|_{H^s} \leq 2\varepsilon, \quad T \leq \varepsilon^{-a+\delta}, \tag{5.2}$$

with  $a = a(d, n)$  in (1.6), then we have the improved bound  $\sup_{t \in [0, T]} \|u(t)\|_{H^s} \leq \frac{3}{2}\varepsilon$ .

In order to prove Lemma 5.1 we first need a preliminary result.

**Lemma 5.2** (Equivalence of the energy norm) *Let  $\delta > 0, N \geq N_1 \geq 1$ . Let  $u(t, x)$  as in (5.2) with  $s \gg 1$  large enough. Then, for any  $0 < c_0 < 1$ , there exists  $C = C(\delta, s, d, n, c_0) > 0$  such that, if we have the smallness condition*

$$\varepsilon C N^\delta N_1^{K_d} \leq 1, \tag{5.3}$$

the following holds true. Define

$$z := \tau^{(0)}(u), \quad u = \tau^{(1)}(z), \quad \mathcal{E}_s(z) := (N_s + E)(z) \tag{5.4}$$

where  $\tau^{(\sigma)}, \sigma = 0, 1$ , are the maps given by Theorem 2 and  $N_s$  is in (4.1),  $E$  is given by Proposition 4.1. We have

$$1/(1 + c_0)\|z\|_{H^s} \leq \|u\|_{H^s} \leq (1 + c_0)\|z\|_{H^s}, \quad \forall t \in [0, T]; \tag{5.5}$$

$$1/(1 + 12c_0)\mathcal{E}_s(z) \leq \|u\|_{H^s}^2 \leq (1 + 12c_0)\mathcal{E}_s(z), \quad \forall t \in [0, T]. \tag{5.6}$$

**Proof** Thanks to (5.3) we have that Theorem 2 and Proposition 4.1 apply. Consider the function  $z = \tau^{(0)}(u)$ . By estimate (3.3) we have

$$\|z\|_{H^s} \leq \|u\|_{H^s} + \tilde{C} N^\delta \|u\|_{H^s}^2 \stackrel{(5.3)}{\leq} \|u\|_{H^s} (1 + c_0),$$

where  $\tilde{C}$  is some constant depending on  $s$  and  $\delta$ . The latter inequality follows by taking  $C$  in (5.3) large enough. Reasoning similarly and using the bound (3.3) on  $\tau^{(1)}$  one gets the (5.5). Let us check the (5.6). First notice that, by (4.3), (4.5) and Lemma 3.5,

$$|E(z)| \leq \tilde{C} \|z\|_{H^s}^{M_{d,n}} N^\delta N_1^{K_d}, \tag{5.7}$$

for some  $\tilde{C} > 0$  depending on  $s$  and  $\delta$ . Then, recalling (5.4), we get

$$|\mathcal{E}_s(z)| \leq \|z\|_{H^s}^2 (1 + \tilde{C} \|z\|_{H^s}^{M_{d,n}-2} N^\delta N_1^{K_d}) \stackrel{(5.5), (5.3)}{\leq} \|u\|_{H^s}^2 (1 + c_0)^3,$$

where we used that  $M_{d,n} - 2 \geq 1$ . This implies the first inequality in (5.6). On the other hand, using (5.5), (5.7) and (5.2), we have

$$\|u\|_{H^s}^2 \leq (1 + c_0)^2 \mathcal{E}_s(z) + (1 + c_0)^{M_{d,n}+2} \tilde{C} N^\delta N_1^{\kappa_d} \varepsilon^{M_{d,n}-2} \|u\|_{H^s}^2.$$

Then, since  $M_{d,n} > 2$  (see (3.1)), taking  $C$  in (5.3) large enough we obtain the second inequality in (5.6). □

**Proof of Lemma 5.1** Assume the (5.2). We study how the Sobolev norm  $\|u(t)\|_{H^s}$  evolves for  $t \in [0, T]$  by inspecting the equivalent energy norm  $\mathcal{E}_s(z)$  defined in (5.4). Notice that

$$\partial_t \mathcal{E}_s(z) = -\{\mathcal{E}_s, H \circ \tau^{(1)}\}(z).$$

Therefore, for any  $t \in [0, T]$ , we have that

$$\begin{aligned} \left| \int_0^T \partial_t \mathcal{E}_s(z) dt \right| &\stackrel{(4.6),(5.2)}{\lesssim_{s,\delta}} T N_1^{\kappa_d} N^\delta (\varepsilon^{\tilde{M}_{d,n}} + N^{-1} \varepsilon^{M_{d,n}+n-2}) \\ &\quad + T N_1^{-s_d+\delta} \varepsilon^{M_{d,n}} + T N^{-s_d+\delta} \varepsilon^n. \end{aligned}$$

We now fix

$$N_1 := \varepsilon^{-\alpha}, \quad N := \varepsilon^{-\gamma},$$

with  $0 < \alpha \leq \gamma$  to be chosen properly. Hence we have

$$\left| \int_0^T \partial_t \mathcal{E}_s(z) dt \right| \lesssim_{s,\delta} \varepsilon^2 T \left( \varepsilon^{M_{d,n}-2+\alpha s_d-\delta\alpha} + \varepsilon^{\tilde{M}_{n,d}-2-\alpha\kappa_d-\delta\gamma} \right) \tag{5.8}$$

$$+ \varepsilon^2 T \left( \varepsilon^{n-2+\gamma s_d-\delta\gamma} + \varepsilon^{M_{n,d}+n-4+\gamma-\alpha\kappa_d-\delta\gamma} \right). \tag{5.9}$$

We choose  $\alpha > 0$  such that

$$M_{d,n} - 2 + \alpha s_d = \tilde{M}_{n,d} - 2 - \alpha\kappa_d, \tag{5.10}$$

i.e.

$$\alpha := \frac{\tilde{M}_{n,d} - M_{d,n}}{s_d + \kappa_d} \stackrel{(4.2),(4.7),(4.4)}{=} \begin{cases} \frac{n-1}{d-1} & \text{if } n \text{ odd} \\ \frac{n-2}{d-1} & \text{if } n \text{ even.} \end{cases} \tag{5.11}$$

We shall choose  $\gamma > 0$  is such a way the terms in (5.9) are negligible with respect to the terms in (5.8). In particular we set (recall (5.11))

$$\gamma \geq \max \left\{ \frac{1}{s_d} (M_{d,n} - n + \frac{\tilde{M}_{d,n} - M_{d,n}}{s_d + \kappa_d} s_d), 2 - n + \tilde{M}_{d,n} - M_{d,n} \right\}. \tag{5.12}$$

Therefore estimates (5.8)–(5.9) become

$$\left| \int_0^T \partial_t \mathcal{E}_s(z) dt \right| \lesssim_{s,\delta} \varepsilon^2 T \varepsilon^a (\varepsilon^{-\delta\alpha} + \varepsilon^{-\delta\gamma})$$

where  $a$  is defined in (1.6) and appears thanks to definitions (3.1), (4.2), (4.4), (4.7) and (5.11). Moreover we define

$$\delta' := 2\delta \max\{\alpha, \gamma\},$$

with  $\alpha, \gamma$  given in (5.11) and (5.12). Notice that, since  $\delta > 0$  is arbitrary small, then  $\delta'$  can be chosen arbitrary small. Since  $\varepsilon$  can be chosen arbitrarily small with respect to  $s$  and  $\delta$ , with this choices we get

$$\left| \int_0^T \partial_t \mathcal{E}_s(z) dt \right| \leq \varepsilon^2/4$$

as long as  $T \leq \varepsilon^{-a+\delta'}$ . Then, using the equivalence of norms (5.6) and choosing  $c_0 > 0$  small enough, we have

$$\begin{aligned} \|u(t)\|_{H^s}^2 &\leq (1 + 12c_0)\mathcal{E}_0(z(t)) \\ &\leq (1 + 12c_0) \left[ \mathcal{E}_s(z(0)) + \left| \int_0^T \partial_t \mathcal{E}_s(z) dt \right| \right] \\ &\leq (1 + 12c_0)^2 \varepsilon^2 + (1 + 12c_0)\varepsilon^2/4 \leq \varepsilon^2 3/2, \end{aligned}$$

for times  $T \leq \varepsilon^{-a+\delta'}$ . This implies the thesis.  $\square$

## References

- Bambusi, D.: Birkhoff normal form for some nonlinear PDEs. *Commun. Math. Phys.* **234**, 253–283 (2003)
- Bambusi, D., Grébert, B.: Birkhoff normal form for partial differential equations with tame modulus. *Duke Math. J.* **135**(3), 507–567 (2006)
- Bambusi, D., Delort, J.M., Grébert, B., Szeftel, J.: Almost global existence for Hamiltonian semi-linear Klein–Gordon equations with small Cauchy data on Zoll manifolds. *Commun. Pure Appl. Math.* **60**, 1665–1690 (2007)
- Bambusi, D., Grébert, B., Maspero, A., Robert, D.: Growth of Sobolev norms for abstract linear Schrödinger equations. To appear in *JEMS*, (2020)
- Bernier, J.: Bounds on the growth of high discrete Sobolev norms for the cubic discrete nonlinear Schrödinger equations on  $h\mathbb{Z}$ . *Discret. Contin. Dyn. Syst. A* **39**(6), 3179–3195 (2019)
- Bernier, J., Faou, E., Grébert, B.: Long time behavior of the solutions of NLW on the  $d$ -dimensional torus. *Forum Math. Sigma* **8**, 12 (2020)
- Berti, M., Delort, J.M.: Almost global solutions of capillary-gravity water waves equations on the circle. In: *UMI Lecture Notes* (2017)
- Berti, M., Feola, R., Pusateri, F.: Birkhoff normal form and long time existence for periodic gravity water waves (2018). Preprint at [arXiv:1810.11549](https://arxiv.org/abs/1810.11549)
- Berti, M., Feola, R., Franzoi, L.: Quadratic life span of periodic gravity-capillary water waves. *Water Waves* (2020). <https://doi.org/10.1007/s42286-020-00036-8>
- Berti, M., Feola, R., Pusateri, F.: Birkhoff normal form for gravity water waves. *Water Waves* (2020). <https://doi.org/10.1007/s42286-020-00024-y>
- Bourgain, J.: Construction of approximative and almost periodic solutions of perturbed linear Schrödinger and wave equations. *Geom. Funct. Anal.* **6**(2), 201–230 (1996)
- Bourgain, J.: Growth of Sobolev norms in linear Schrödinger equations with quasi-periodic potential. *Commun. Math. Phys.* **204**(1), 207–247 (1999)
- Delort, J.M.: On long time existence for small solutions of semi-linear Klein–Gordon equations on the torus. *J. d'Analyse Mathématique* **107**, 161–194 (2009)
- Delort, J.M.: Growth of Sobolev norms of solutions of linear Schrödinger equations on some compact manifolds. *Int. Math. Res. Not.* **12**, 2305–2328 (2010)
- Delort, J.M.: Quasi-linear perturbations of Hamiltonian Klein–Gordon equations on spheres. *Am. Math. Soc.* (2015). <https://doi.org/10.1090/memo/1103>
- Delort, J.M., Imekraz, R.: Long time existence for the semi-linear Klein-Gordon equation on a compact boundaryless Riemannian manifold. *Commun. PDE* **42**, 388–416 (2017)
- Eliasson, H.: Perturbations of linear quasi-periodic systems. In: *Dynamical Systems and Small Divisors* (Cetraro, Italy, 1998). *Lect. Notes Math.*, vol. 1784, pp. 1–60. Springer, Berlin (2002)
- Fang, D., Zhang, Q.: Long-time existence for semi-linear Klein–Gordon equations on tori. *J. Differ. Equ.* **249**, 151–179 (2010)

19. Faou, E., Grébert, B.: A Nekhoroshev-type theorem for the nonlinear Schrödinger equation on the torus. *Anal. PDE* **6**(6) (2013)
20. Feola, R., Grébert, B., Iandoli, F.: Long time solutions for quasi-linear Hamiltonian perturbations of Schrödinger and Klein–Gordon equations on tori. Preprint. [arXiv:2009.07553](https://arxiv.org/abs/2009.07553) (2020)
21. Feola, R., Iandoli, F.: A non-linear Egorov theorem and Poincaré–Birkhoff normal forms for quasi-linear pdes on the circle. Preprint. [arXiv:2002.12448](https://arxiv.org/abs/2002.12448) (2020)
22. Feola, R., Iandoli, F.: Long time existence for fully nonlinear NLS with small Cauchy data on the circle. *Annali della Scuola Normale Superiore di Pisa (Classe di Scienze)*. To appear (2019). <https://doi.org/10.2422/2036-2145.201811-003>
23. Grébert, B.: Birkhoff normal form and Hamiltonian PDEs. *Séminaires et Congrès* **15**, 3067–3102 (2016)
24. Grébert, B., Imekraz, R., Paturel, E.: Normal forms for semilinear quantum harmonic oscillators. *Commun. Math. Phys* **291**, 763–798 (2009)
25. Imekraz, R.: Long time existence for the semi-linear beam equation on irrational tori of dimension two. *Nonlinearity* **29**, 1–46 (2007)
26. Ionescu, A.D., Pusateri, F.: Long-time existence for multi-dimensional periodic water waves. *Geom. Funct. Anal.* **29**, 811–870 (2019)
27. Moser, J.: A rapidly convergent iteration method and non-linear partial differential equations—I. *Ann. Sc. Norm. Sup. Pisa Cl. Sci. III Ser.* **20**(2), 265–315 (1966)
28. Pausader, B.: Scattering and the Levandosky–Strauss conjecture for fourth-order nonlinear wave equations. *J. Differ. Equ.* **241**(2), 237–278 (2007)
29. Zhang, Q.: Long-time existence for semi-linear Klein–Gordon equations with quadratic potential. *Commun. Partial Differ. Equ.* **35**(4), 630–668 (2010)

**Publisher's Note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.