



# *J*ournal de l'École polytechnique *Mathématiques*

Joackim BERNIER & Nicolas CAMPS

Long time stability for cubic nonlinear Schrödinger equations on non-rectangular flat tori

Tome 12 (2025), p. 713-800.

<https://doi.org/10.5802/jep.300>

© Les auteurs, 2025.



Cet article est mis à disposition selon les termes de la licence  
LICENCE INTERNATIONALE D'ATTRIBUTION CREATIVE COMMONS BY 4.0.

<https://creativecommons.org/licenses/by/4.0/>

Publié avec le soutien  
du Centre National de la Recherche Scientifique



Publication membre du  
Centre Mersenne pour l'édition scientifique ouverte

[www.centre-mersenne.org](http://www.centre-mersenne.org)

e-ISSN : 2270-518X

# LONG TIME STABILITY FOR CUBIC NONLINEAR SCHRÖDINGER EQUATIONS ON NON-RECTANGULAR FLAT TORI

BY JOACKIM BERNIER & NICOLAS CAMPS

**ABSTRACT.** — We consider nonlinear Schrödinger equations on flat tori satisfying a simple and explicit Diophantine non-degeneracy condition. Provided that the nonlinearity contains a cubic term, we prove the almost global existence and stability of most of the small solutions in high regularity Sobolev spaces. To this end, we develop a normal form approach designed to handle general resonant Hamiltonian partial differential equations for which it is possible to modulate the frequencies by using the initial data.

**RÉSUMÉ** (Stabilité en temps long des équations de Schrödinger non linéaires cubiques sur des tores plats non rectangulaires)

Nous considérons des équations de Schrödinger non linéaires sur des tores plats satisfaisant à une condition simple et explicite de non-dégénérescence diophantienne. Sous la condition que la non-linéarité contienne un terme cubique, nous prouvons l'existence et la stabilité presque globales de la plupart des petites solutions dans des espaces de Sobolev de forte régularité. À cette fin, nous développons une approche de forme normale conçue pour traiter des équations aux dérivées partielles hamiltoniennes résonnantes générales, pour lesquelles il est possible de moduler les fréquences en utilisant les données initiales.

## CONTENTS

|                                                              |     |
|--------------------------------------------------------------|-----|
| 1. Introduction.....                                         | 714 |
| 2. Low frequency reduction.....                              | 728 |
| 3. Re-centered polynomials.....                              | 745 |
| 4. Functional setting for re-centered polynomials.....       | 750 |
| 5. Small-divisor estimate for the modulated frequencies..... | 762 |
| 6. Normal form with internal parameters.....                 | 765 |
| 7. Proof of the main theorem.....                            | 788 |
| Appendix.....                                                | 794 |
| References.....                                              | 798 |

**MATHEMATICAL SUBJECT CLASSIFICATION (2020).** — 35B34, 35B35, 35Q55, 37K45, 37K55.

**KEYWORDS.** — Almost global existence, flat tori, nonlinear Schrödinger equations, normal forms, internal parameters.

During the preparation of this work the authors benefited from the support of the Centre Henri Lebesgue ANR-11-LABX-0020-0, the region “Pays de la Loire” through the project “MasCan” and the ANR project KEN ANR-22-CE40-0016.

## 1. INTRODUCTION

1.1. CONTEXT. — We consider the nonlinear Schrödinger equation

$$(NLS) \quad i\partial_t u + \operatorname{div} G \nabla u = f(|u|^2)u, \quad (t, x) \in \mathbb{R} \times \mathbb{T}^d,$$

where  $G$  is a real symmetric  $d \times d$  positive matrix and  $f : \mathbb{R} \rightarrow \mathbb{R}$  is a  $C^\infty$  function satisfying  $f'(0) \neq 0$  (to ensure the existence of a cubic nonlinear term in the equation). Without loss of generality, we assume that  $f(0) = 0$ .

This is a convenient way of rewriting nonlinear Schrödinger equations on rescaled tori. Indeed,  $d$ -dimensional flat tori writes  $\mathbb{T}_{\mathcal{L}}^d = \mathbb{R}^d / \mathcal{L}$  where  $\mathcal{L} =: V\mathbb{Z}^d$  is a lattice of  $\mathbb{R}^d$  ( $V$  being a  $d \times d$  real invertible matrix) and so by setting  $G = V^{-1} {}^t(V^{-1})$  and by applying a linear change of coordinate  $x \mapsto Vx$ , (NLS) is equivalent to

$$i\partial_t u + \Delta u = f(|u|^2)u, \quad (t, x) \in \mathbb{R} \times \mathbb{T}_{\mathcal{L}}^d.$$

By considering small solutions, (NLS) can be seen as a perturbation of the linear integrable system

$$(1.1) \quad i\partial_t u + \operatorname{div} G \nabla u = 0$$

whose actions  $I_n(u) = |u_n|^2$  are the square modulus of the Fourier coefficients

$$u_n = \frac{1}{(2\pi)^d} \int_{\mathbb{T}^d} e^{-in \cdot x} u(x) \, dx.$$

Since the actions are constants of motion for the linear evolution (1.1), a natural question in this context of perturbation theory is then to understand how much the nonlinear flow of (NLS) preserves these actions. Or more formally, by setting

$$\|u\|_{H^s} := \left( \sum_{n \in \mathbb{Z}^d} \langle n \rangle^{2s} |u_n|^2 \right)^{1/2}, \quad \langle n \rangle := (1 + |n|^2)^{1/2},$$

for initial data of size  $\varepsilon$  in  $H^s(\mathbb{T}^d)$ , on what time scales  $T(\varepsilon)$  are the integrable dynamics orbitally stable under the flow of (NLS), in the sense that the actions of the solution are slow-variables:

$$(1.2) \quad \sup_{|t| \leq T(\varepsilon)} \sum_{n \in \mathbb{Z}^d} \langle n \rangle^{2s} \left| |u_n(t)|^2 - |u_n(0)|^2 \right| \ll \varepsilon^2 ?$$

A sub-question is the one of the stability of the zero solution  $u(t, x) = 0$ . Or more precisely, for initial data of size  $\varepsilon$  in  $H^s(\mathbb{T}^d)$ , on what time scales  $T(\varepsilon)$  do we have that

$$(1.3) \quad \sup_{|t| \leq T(\varepsilon)} \|u(t)\|_{H^s} \leq C_s \|u(0)\|_{H^s} ?$$

(where  $C_s > 0$  is constant depending only on  $s$ ). The statement (1.2) is stronger than (1.3). Note that the local Cauchy theory ensures that  $T(\varepsilon)$  is larger than or equal to the linear time-scale  $\sim \varepsilon^{-2}$ , at least when  $s > d/2$ .

The above question is a mathematical formulation of the (absence) of transfer of energy from large to small scales of oscillation. Extensive research has been conducted over the past decades to construct energy cascades out of the resonant interactions,

and we refer to [CKS<sup>+</sup>10, GG22, GK15] for partial results about norm amplification for the cubic Schrödinger equation on tori. In this setting, the existence of infinite energy cascades dynamics conjectured by Bourgain [Bou00b] have only been established when the NLS is posed on the waveguide  $\mathbb{R} \times \mathbb{T}^d$  manifold. The unbounded direction provides a stronger dispersion that reduces the dynamic to the effective one of the resonant system [HPTV15]. In contract, the second author and Staffilani [CS24, Cam24] proved that the Sobolev norms remain bounded in infinite time when NLS is posed on a Diophantine waveguide (namely  $\mathbb{R} \times \mathbb{T}_{\mathcal{L}}^d$  for  $\mathcal{L}$  having the same Diophantine property as in the present paper).

In recent decades, important developments have been made in Birkhoff normal form techniques to prove the stability of small solution over longer time scales, for many different Hamiltonian PDEs (see e.g. [BG06, BDGS07, GIP09, Del12, FGL13, BD18, BMM24]). All these results ensure the *almost global existence and stability of the small solutions* in the sense that (1.3) holds for  $T(\varepsilon) \sim \varepsilon^{-r}$  with  $r$  arbitrarily large, provided that  $s \gg r$ . The case of low regularity is still wide open (see e.g. [BG25, BGR23] for results in this direction). However, all these results apply to PDEs associated with non-resonant Hamiltonians: the eigenvalues of the operator associated with the linearized equation (also called the frequencies) must be rationally independent. For (NLS), the frequencies are

$$\lambda_n^2 := g(n, n), \quad \text{where} \quad g(a, b) := {}^t a G b.$$

They are clearly<sup>(1)</sup> not rationally independent, and therefore (NLS) is resonant. A common way to overcome this obstacle is to add a random convolution potential  $V$  (with real Fourier coefficients) to the equation, and the equation (NLS) becomes

$$(NLS^*) \quad i\partial_t u + \operatorname{div} G \nabla u = V * u + f(|u|^2)u.$$

See for example [BG06, YZ14, FI21, BMP20, BG25, BFM24]. In this case, the frequencies are modulated and become  $\omega_n = \lambda_n^2 + V_n$ . They are almost surely rationally independent. For rational tori (i.e.,  $G \in \mathbb{Q}^{d \times d}$  up to multiplication by a scalar), the almost global existence and stability of the small solutions of (NLS\*) was proved by [BG06]. For irrational flat tori, small divisors are degenerate which may generate high frequencies instability. Nevertheless, in [BFG20a], Bernier, Faou and Grébert developed an approach to prove almost global existence of small solutions for semi-linear equations enjoying such small divisor estimates. More recently, Bambusi, Feola and Montalto [BFM24] proved the almost global stability of small solutions of (NLS\*).

In this paper, we aim to extend these results by removing the Fourier multiplier  $V * u$ . Since (NLS) is resonant, the previous results do not hold, i.e., the stability (1.2) or (1.3) with  $T(\varepsilon) \sim \varepsilon^{-r}$  and  $r \gg 1$  arbitrarily large cannot be derived from the Birkhoff normal form theorem. The situation is even worse, and we expect (1.2) and (1.3) to be false after some polynomial time scales  $T(\varepsilon) \sim \varepsilon^{-\alpha}$ . For example, one can deduce from [CKS<sup>+</sup>10] that (1.3) is false on the square torus  $\mathbb{T}^2$  (i.e.,  $G = I_2$ ) for some initial data after  $T(\varepsilon) = \varepsilon^{-2} \log(\varepsilon^{-1})$ .

<sup>(1)</sup>They belong to the finite dimensional  $\mathbb{Q}$  vector space generated by the coefficients of  $G$ .

The only resonant equations for which stability results were proved have a special algebraic property: the four-waves interactions must be trivial or, in other words, the quartic terms of the Birkhoff normal form of the equation must be integrable. This property allows to modulate the frequencies using the initial data as parameters and leads to the almost global existence and stability of most of the small solutions: an exceptional set of initial data has to be excluded. Such results have been proved for (NLS) in dimension  $d = 1$  [Bou00a, BFG20b], for the generalized Korteweg–de Vries and a modified version of Benjamin–Ono equations [BG21] and for the Kirchhoff equation [BH22] but only for  $T(\varepsilon) \sim \varepsilon^{-6}$ . We also mention [LX24, BCGW24], who proved stability in Gevrey spaces over exponentially long times.

Kuksin and Pöschel observed in [KP96] that four-waves interactions are trivial for (NLS) in dimension  $d = 1$ . The point is that this property is also true in any dimension provided that

$$(1.4) \quad g(a, b) \neq 0, \quad \forall a, b \in \mathbb{Z}^d \setminus \{0\}.$$

Indeed, as in dimension 1, if  $n_1 - n_2 + n_3 - n_4 = 0$ , we have

$$\Omega_{\vec{n}} := \lambda_{n_1}^2 - \lambda_{n_2}^2 + \lambda_{n_3}^2 - \lambda_{n_4}^2 = 2g(n_1 - n_2, n_1 - n_4)$$

which ensure that if  $\Omega_{\vec{n}} = 0$  then  $\{n_1, n_3\} = \{n_2, n_4\}$ .

**1.2. MAIN THEOREM AND COMMENTS.** — In order to establish our main result, we have to make the non-degeneracy condition (1.4) quantitative.

**DEFINITION 1.1** (Admissible tori). — The flat tori  $\mathbb{T}_{\mathcal{L}}^d$  is admissible if there exist  $C > 0$  and  $\tau^* > 0$  such that for all  $a, b \in \mathbb{Z}^d \setminus \{0\}$ ,

$$(1.5) \quad |g(a, b)| \geq \frac{C}{\|a\|_2^{\tau^*} \|b\|_2^{\tau^*}}.$$

In particular,  $g(a, b) = 0$  if and only if  $a = 0$  or  $b = 0$ .

**REMARK 1.2.** — If the vector  $(G_{i,j})_{i \leq j}$  of the coefficients of  $G$  is Diophantine then the torus is admissible. Note that the converse is not true (see (1.12)). In any case, as a consequence, since this property is true for almost all matrix  $G$ , we can say that almost all flat torus are admissible (as soon as  $\tau_* > d(d+1)/2$ ).

To state our result, we have to introduce  $\Pi_M$ ,  $M > 0$ , the orthogonal projectors onto frequencies smaller than  $M$ :

$$\Pi_M u = \sum_{|n| \leq M} u_n e^{in \cdot x}.$$

Moreover, by abuse of notation, we denote by  $\text{meas}$  the canonical Lebesgue measure on  $\Pi_M L^2(\mathbb{T}^d; \mathbb{C})$ , for  $M > 0$ .

**THEOREM 1.3.** — *Let  $\mathbb{T}_{\mathcal{L}}^d$  be an admissible torus in the sense of Definition 1.1 and  $r \gtrsim 1$ . There exist  $\mu_d > 0$  and  $\nu \lesssim_{r, \mathcal{L}} 1$  such that for all  $s \gtrsim_r 1$  and for all  $\varepsilon \lesssim_{r, s} 1$ , provided that  $M > 0$  satisfies*

$$(1.6) \quad \varepsilon^{-\mu_d r/s} \leq M \leq \varepsilon^{-\nu},$$

there exists an open set  $\Theta_\varepsilon \subset \Pi_M B_s(\varepsilon)$  for which the following holds: local solutions to (NLS) in  $H^s$  initiated from initial data

$$(1.7) \quad u(0) \in B_s(2\varepsilon) \cap \Pi_M^{-1} \Theta_\varepsilon$$

exist in  $C([-T_\varepsilon, T_\varepsilon], H^s)$  for  $T_\varepsilon = \varepsilon^{-r}$ , and satisfy, for all  $|t| \leq T_\varepsilon$ ,

$$(1.8) \quad \|u(t)\|_{H^s} \leq 2^{s+1} \|u(0)\|_{H^s}.$$

Moreover, these initial data are typical, in the sense that

$$(1.9) \quad \text{meas}(\Theta_\varepsilon) \geq (1 - \varepsilon^{1/40}) \text{meas}(\Pi_M B_s(\varepsilon)).$$

Roughly speaking, for any  $r \gtrsim 1$ ,  $s \gtrsim_r 1$  and  $\varepsilon > 0$  sufficiently small, we prove that, under a generic condition on finitely many low Fourier modes, initial data of size  $\varepsilon$  in  $H^s(\mathbb{T}^d)$  lead to stable solutions over time scales of order  $\varepsilon^{-r}$ .

### 1.2.1. Comments about the literature

– Theorem 1.3 is the first almost global existence result for a resonant PDE in dimension  $d \geq 2$ .

– It extends the results [Bou00a, BFG20b] in the one-dimensional case which is much more favorable. In particular, the almost global preservation of all the actions (i.e., (1.2)) holds when  $d = 1$ . When  $d \geq 2$ , however, we prove the stability of the low-frequency actions and of the high-frequency super-actions, as discussed Section 1.2.2.

– In a similar finite dimensional setting the geometric part of the Nekhoroshev theorem allows to prove the long time stability (in a sense weaker than (1.2) and (1.3)) of any small solution, removing the restriction (1.7). In contrast, [BK05] strongly suggests that such a result should not extend to the infinite dimensional setting. We nevertheless mention [Bam99, BG24] for results in this direction.

### 1.2.2. Technical comments

– We propose a new formulation (1.9) to quantify the proportion of initial data leading to almost global and stable solutions. In this formalism, we only need to impose a condition on the low Fourier coefficients. Moreover, we do not need to assume any additional decay of the Fourier coefficients as it is usually done (e.g. in [Bou00a, BFG20b, BG21, LX24]). We defer comments on this formulation to Section 1.2.3.

– The constraint  $s \gtrsim_r 1$  could be refined in  $s \gtrsim r^2$  (which is a classical constraint for this kind of problems, see e.g. [BG21]). It mainly comes from the existence of a number  $M$  satisfying (1.6) and the fact that in the proof we impose  $\nu \lesssim r^{-1}$ , as specified in (3.1).

– In dimension  $d \geq 2$ , we are not able to control the variation of all the actions (it is an open problem even for (NLS\*)). Actually (in (2.34)), we only control the variation of the low actions

$$\sup_{|t| \leq \varepsilon^{-r}} \sum_{n \in \mathcal{C}_{\text{lo}}} \langle n \rangle^{2s} | |u_n(t)|^2 - |u_n(0)|^2 | \ll \varepsilon^2,$$

where  $\mathcal{C}_{\text{lo}} \subset \mathbb{Z}^d$  is a finite set, with  $\{n \in \mathbb{Z}^d \mid |n| \leq (17r)^{-1}M\} \subset \mathcal{C}_{\text{lo}}$ . For the high modes, however, as [BFM24], we only control the variations of the super-actions

based on a the cluster decomposition of the frequency space due to [Bou98, BM19]. We refer to Proposition 2.18 for a precise estimate (the truncation parameter  $M$  of this proposition is not the same as the one of Theorem 1.3). The unusual exponent  $2^{s+1}$  in (1.8) is a consequence of the dyadic structure of the cluster decomposition.

### 1.2.3. Comments on the set of initial data

– To overcome the possible instabilities generated by the resonances of the equation, we modulate the frequencies by using initial data. This leads (very technical) Diophantine conditions on the set of initial data. Therefore, in order to prove that the stability result is generic (and even not empty), we have to prove that this set of stable initial data is large in some sense. In the finite dimensional setting, we would prove that the set of admissible initial data is open and asymptotically of full Lebesgue measure.

In the infinite dimensional setting, however, there is no Lebesgue measure and therefore no canonical choice to express such a result. Fortunately, the constraints we impose on the initial datum are quite weak, and it is possible to prove that this set is big for any measure that is not too much degenerated (see e.g. [BFG20b, BG21, LX24, BCGW24] for different possible choices and [BG21] for a discussion about the topology of this set). We do not know if there is an intrinsic way to express that this set is large.

– In this paper, as in [BCGW24], we stress out the fact that it suffices to impose a condition on the low Fourier modes of the initial (i.e.,  $u_n$  with  $|n| \leq M$ , with  $M \sim \varepsilon^{-\mu_d r/s}$  with  $s \gtrsim r$ ). This is essentially the meaning of (1.7).

– The dimension of the finite-dimensional space  $\Pi_M H^s(\mathbb{T}^d, \mathbb{C})$ , which is of order  $M^d \geq \varepsilon^{-d\mu_d r/s}$ , is large with respect to the relative measure of the set of the initial data we exclude ( $\varepsilon^{1/40}$ ). Hence, our probability framework is closer to an infinite dimensional one than a finite dimensional one, even if our statement involves the (finite-dimensional) Lebesgue measure. Indeed, in the limit  $\varepsilon \rightarrow 0$  the dimension of  $\Pi_M H^s(\mathbb{T}^d, \mathbb{C})$  goes to infinity, and we can show that the Lebesgue measure on  $\Pi_M B_s(\varepsilon)$  concentrates its mass on the sphere of radius  $\varepsilon$ :

$$\frac{1}{\text{meas}(\Pi_M B_s(\varepsilon))} \text{meas} \{u \in \Pi_M B_s(\varepsilon) \mid \|u\|_{H^s} \leq \varepsilon(1 - M^{-1/2})\} \\ = (1 - M^{-1/2})^{\dim \Pi_M L^2} \ll \varepsilon^{1/40}.$$

Therefore, to prove Theorem 1.3 we could assume for free that  $\|\Pi_M u(0)\|_{H^s} \geq \varepsilon(1 - M^{-1/2})$ . Nevertheless, we do not use such a property in the proof and, to avoid any hidden smallness assumption of this kind on  $\|u(0) - \Pi_M u(0)\|_{H^s}$ , we allow  $\|u(0)\|_{H^s}$  to be as large as  $2\varepsilon$  in (1.7).

– A popular way to draw initial data is to draw the Fourier coefficients independently, see e.g. [Bou00a, BFG20b, BG21, LX24]. For example, here it could consist in considering random initial data of the form<sup>(2)</sup>

$$V^{(\varepsilon)} = c\varepsilon \sum_{k \in \mathbb{Z}^d} V_k \langle k \rangle^{-s-\alpha},$$

<sup>(2)</sup>Note that  $V^{(\varepsilon)}$  also lives on a sphere in the sense that  $\|V^{(\varepsilon)}\|_{\ell_{s+\alpha}^\infty} = c\varepsilon$  almost surely.

where the random variables  $V_k$  are independent and uniformly distributed in the complex unit disk  $\mathbb{D}(0, 1)$ ,  $\alpha > d/2$  and  $c(\alpha) > 0$  is a normalizing constant to ensure that  $\|V^{(\varepsilon)}\|_{H^s} < \varepsilon$  almost surely. With minor changes in the proof,<sup>(3)</sup> we could easily replace the measure estimate (1.9) of Theorem 1.3 by

$$(1.10) \quad \mathbb{P}(\Pi_M V^{(\varepsilon)} \in \Theta_\varepsilon) \geq 1 - \varepsilon^{1/40}.$$

We would deduce from this that the set of the stable initial data  $\bigcup_\varepsilon B_s(2\varepsilon) \cap \Pi_M^{-1}\Theta_\varepsilon$  is asymptotically of full measure:

$$\mathbb{P}(\varepsilon V^{(1)} \in \Pi_M^{-1}\Theta_\varepsilon) \geq 1 - \varepsilon^{1/40}.$$

Approaches of this type have the disadvantage of requiring initial data smoother than necessary. We believe that one of the strengths of our probabilistic formulation in Theorem 1.3 is to describe dynamics in  $H^s(\mathbb{T}^d)$  for solutions no more regular than  $H^s(\mathbb{T}^d)$ .

– In contrast with papers using rational normal forms, the set  $\Theta_\varepsilon$  a priori also depends on the angles of the initial data: we do not prove that it is invariant by rotation of the angles, in the sense that

$$(1.11) \quad u \in \Theta_\varepsilon \iff \sum_{k \in \mathbb{Z}^d} |u_k| e^{ik \cdot x}.$$

That is why contrary to [BFG20b, BG21, BCGW24], we do not only draw the actions of the initial data randomly but also their angles. We point out that nevertheless, here, (1.11) is true up to conjugation by a diffeomorphism  $(\Psi, \text{defined in (7.1)})$ .

#### 1.2.4. Comments about flat tori

– The strength of our method with respect to rational normal form methods is to only require a very weak and explicit condition on the external parameters  $G$  associated to the tori  $\mathbb{T}_{\mathcal{L}}^d$ . Explicit examples with Diophantine numbers can be easily produced, e.g.

$$(1.12) \quad G = \begin{pmatrix} 1 & \sqrt{2} \\ \sqrt{2} & 3 \end{pmatrix} \text{ is admissible.}$$

– Note that admissible tori are not rectangular but can be chosen arbitrarily close to rectangular tori. For rectangular tori, however, the four-wave interaction set is not trivial. It can be decoupled into two one-dimensional systems (see [SW20]). We believe that it is an interesting problem to understand the situation on a Diophantine rectangular tori.

– Most of our analysis can be easily transferred to the case of a hyperbolic Laplacian (when the symmetric matrix  $G$  is indefinite). However, the finite-dimensional reduction, which is based on the separation property of the frequencies (see Lemma 2.15), may no longer be possible.

<sup>(3)</sup>in measure estimates, it suffices to replace round balls  $\Pi_M B_s(0, \varepsilon)$  by rectangular ones  $\Pi_M B_{\ell_{s+\alpha}^\infty}^\infty(0, c\varepsilon)$ .

– Theorem 1.3 is in line with the study of nonlinear Schrödinger equations on irrational tori that has been developed in recent years. The general idea is to mimic the dispersion in compact settings. Under some Diophantine conditions on the torus, the refocusing time for the waves is actually longer [DGG17, DGGRM22] than it is on the square torus, and the four-wave resonant system is smaller [SW20]. On Diophantine rectangular tori, Deng–Germain [DG19] improved polynomial upper bounds for the growth of Sobolev norms, while Deng [Den19] achieved polynomial growth in the energy-critical case for small energy solutions (a challenging open problem on the square torus). In a similar spirit, Colot–Germain [CG22] derived the kinetic wave equation for larger set of scalings by considering dispersion relations associated with non-rectangular Diophantine tori.

– Our result goes in the same direction, showing stronger stability properties on non-rectangular Diophantine tori, in contrast to the norm amplification observed in [CKS<sup>+</sup>10]. We stress out that Guilani–Guardia [GG22] proved that the same norm amplification mechanism occurs on rectangular Diophantine tori as well through quasi-resonant quartets, but after exponential time-scales. This does not contradict our Theorem 1.3, which achieves polynomial time-scales.

**1.3. DISCUSSIONS ON THE PROOF.** — In this paper we develop a normal form approach for proving stability of solutions to resonant Hamiltonian systems whose modulated frequencies are highly degenerate (more than what the rational normal form approach requires). We believe that beyond the almost global existence result stated in Theorem 1.3, this method is one of the main interests of this paper. In the case of (NLS) on flat tori, one advantage of this method is that it results in the simple (and quite minimal) non-degeneracy assumption (1.5) for admissible tori.

Our normal form is inspired by an approach developed by Bourgain in the paper [Bou00a] for (NLS) in dimension  $d = 1$ . Apart from the fact that we overcome the degeneracy of linear frequencies which is specific to the dimension  $d \geq 2$ , the main difference is that we also provide a formalism that allows us to prove and quantify the sense in which most small solutions are stable over very long times.

In his proof, Bourgain describes a generic step of his normal form procedure: after a preliminary change of variable (let us say  $\tau^0$ ) to put the Hamiltonian under resonant normal form, he explains how to construct a new change of variable to get better properties. The point is that (due to the modulated small divisors) the transformation  $\tau^0$  itself depends strongly on  $u(0)$  (let us denote it by  $\tau^0 =: \tau_{u(0)}^0$ ). Therefore, in order to construct the new transformation, he has to assume that the new initial data  $v(0) := \tau_{u(0)}^0(u(0))$  belong to a set  $\Xi_{u(0)}$  encoding some Diophantine conditions and depending on  $u(0)$ . He proves that  $\Xi_{u(0)}$  is asymptotically of full measure, but he does not explain why this implies that  $v(0) = \tau_{u(0)}^0(u(0)) \in \Xi_{u(0)}$  for most real initial data  $u(0)$ .

This type of problem is quite classical in standard KAM theory and is usually solved by introducing Lipschitz norms to track the dependencies with respect to the internal parameters. This is the strategy we implement in this paper. However, due to the small

divisor degeneracy this is far from obvious and has required the introduction of a whole technical framework (and associated technical estimates), which largely explains the length of this paper. It has also required us to refine some of the fundamental estimates of [Bou00a] (see e.g. Remark 4.2 for a more detailed discussion on this point). This difficulty is inherent in modulating the frequencies with internal parameters. It is a well-known obstacle in the open problem of constructing infinite dimensional invariant tori for (NLS)<sup>(4)</sup> in dimension  $d = 1$  (see e.g. [Pös02]).

To continue the discussion of the proof, we need to introduce the Hamiltonian structure of (NLS). More precisely, (NLS) rewrites

$$(1.13) \quad i\partial_t u = \nabla H(u) \quad \text{with} \quad H(u) = Z_2(u) + \frac{1}{2} \int_{\mathbb{T}^d} F(|u(x)|^2) dx,$$

where  $F$  is the primitive of  $f$  vanishing at the origin and

$$Z_2(u) = \frac{1}{2} \sum_{n \in \mathbb{Z}^d} \lambda_n^2 |u_n|^2.$$

The classical notations (like  $\nabla$ ) are defined in Section 1.4 below.

### 1.3.1. *Modulated frequencies, small divisors and comparison with rational normal forms*

First, let us assume that<sup>(5)</sup> (NLS) can be put in Birkhoff normal form (at least up to some high order  $2r \gg 1$ ). This means that there would be a canonical change of variable  $\tau$  close to the identity in  $H^s$  such that

$$(1.14) \quad H \circ \tau^{-1}(u) = Z_2(u) + Q(u) + \mathcal{O}(u^{2r}) \quad \text{and} \quad \{Z_2, Q\} = 0.$$

Then, using the admissibility condition on  $\mathbb{T}_{\mathcal{L}}^d$ , the four-waves interactions are integrable and we deduce that (up to a gauge transform), setting  $v = e^{it \operatorname{div} G \nabla} \tau(u)$ , (NLS) rewrites

$$i\partial_t v_n = \omega_n(v) v_n + \mathcal{O}(v^5), \quad \text{where} \quad \omega_n(v) = -f'(0)|v_n|^2.$$

Since  $f'(0) \neq 0$ , this allows to modulate the frequencies using the initial data as parameters. Note, however, that since  $v \in H^s$  this modulation is highly degenerate in the high frequency regime since  $\omega_n$  decays at least like  $\langle n \rangle^{-2s}$  with  $s \gg 1$ . Under some generic conditions on  $v$ , this provides small divisors estimates of type

$$(1.15) \quad \left| \sum_{n \in \mathbb{Z}^d} h_n \omega_n(v) \right| \gtrsim \|v\|_{H^s}^2 (\mu_{\max}(h))^{-2|h|_{\ell^1}} (\mu_{\min}(h))^{-2s},$$

where  $\mu_{\min}(h) = \min\{\langle n \rangle \mid h_n \neq 0\}$ ,  $\mu_{\max}$  is defined similarly and  $h \in \mathbb{Z}^{(\mathbb{Z}^d)}$  is a family of integers with finite support. We prove such an estimate in Section 5. This type of small divisors seems to be ubiquitous for resonant Hamiltonian PDEs (e.g. it is

<sup>(4)</sup>Note that since in dimension  $d = 1$  we control the variation of all the actions, what we prove in this paper, as well as what is proved in [Bou00a, BFG20b], is somehow not so far from such a result: we prove that most of the small solutions to (NLS) stay very close to some infinite dimensional tori for very long times.

<sup>(5)</sup>Due to the degeneracy of the small divisor associated with the linear frequencies, this is an open problem in dimension  $d \geq 2$  discussed in Section 1.3.3.

also true for KdV, Benjamin–Ono and their generalizations [BG21], Kirchhoff [BH22], the pure-gravity water waves [BFP23]) to the notable exception of the Schrödinger–Poisson equation (see [BFG20b, BCGW24]).

There are three different contributions in the right-hand-side of the small divisor estimate (1.15).

- The first one,  $\|v\|_{H^s}^2$ , comes from homogeneity arguments and is harmless.
- The second one,  $(\mu_{\max}(h))^{-2|h|_{\ell^1}}$ , comes from counting estimates and could be very critical, but truncation arguments (presented in Section 1.3.3) enable to handle these resonances.
- The last contribution  $(\mu_{\min}(h))^{-2s}$  is the most critical. In this sense, we say that the small divisor estimate (1.15) is degenerate. To explain why it is so critical, let us focus on just one aspect that we think is the most important.

To describe the dynamics over longer timescales one has to conjugate (at least locally) (NLS) to an integrable system up to some higher order terms. In other words, up to a new canonical change of variables  $z = \tau^\sharp(v)$ , (NLS) rewrites

$$i\partial_t z_n = \omega_n^\sharp(z) z_n + \mathcal{O}(z^{2q+1}), \quad \text{where} \quad \omega_n^\sharp(z) = \omega_n(z) + \mathcal{O}(z^4),$$

and  $3 \leq q \leq r$ . To transform the remainder terms of order  $\mathcal{O}(z^{2q+1})$  into terms of order  $\mathcal{O}(z^{2q+3})$ ,<sup>(6)</sup> a naive approach (in the spirit of Birkhoff normal forms) would be to average the terms of order  $\mathcal{O}(z^{2q+1})$  by the flow generated by the frequencies  $\omega_n(z)$ <sup>(7)</sup> to make them integrable. The point, identified and discussed in [BFG20b], is that usually, when  $n$  is large enough,  $\omega_n$  is negligible with respect to  $\omega_n^\sharp$ , which makes this naive approach fail. More precisely, due to higher order correction terms  $\omega_n^\sharp$  has no reason to decay like  $\langle n \rangle^{-2s}$  (i.e., like  $\omega_n$ ;<sup>(8)</sup> for example, for (NLS) in dimension  $d = 1$ , explicit computation of the corrections of order 6 in [BFG20b] proves that it actually decays at most like  $\langle n \rangle^{-2}$ ). In any case, more terms of  $\omega_n^\sharp$  have to be considered to average the terms of order  $\mathcal{O}(z^{2q+1})$ .

We discuss two strategies to overcome this problem: either we identify a higher order term correction term allowing to prove that the small divisors are stable by perturbation, or we use that the derivative of the small divisors is stable by perturbation.

- The first strategy, used in the papers following the rational normal form approach [BFG20b, BG21, LX24], relies on an explicit computation of the correction terms of order 4 of  $\omega_n^\sharp$  (coming from the terms of order 6 in the Hamiltonian) and on the hope that they are much less degenerate than  $\omega_n$ . For example, here, we could prove that

<sup>(6)</sup>Of course, in practice, as we will see in the proof, the gain could be much weaker.

<sup>(7)</sup>i.e., those of  $i\partial_t z_n = \omega_n(z) z_n$ .

<sup>(8)</sup>Actually for some equations like Benjamin–Ono [BG21] or the pure gravity water waves [BFP23],  $\omega_n$  contains some terms that do decay like  $\langle n \rangle^{-2s}$  but they do not contribute to some of the small divisors.

they are of the form

$$\tilde{\omega}_n(z) = (f'(0))^2 \sum_{m \neq n} \frac{|z_m|^4}{\lambda_{n-m}^2} + \text{other terms.}$$

In the favorable cases it leads to small divisors estimates of the form<sup>(9)</sup>

$$(1.16) \quad \left| \sum_{n \in \mathbb{Z}^d} h_n(\omega_n + \tilde{\omega}_n)(z) \right| \gtrsim \|z\|_{H^s}^2 (\mu_{\max}(h))^{-C|h|_{\ell^1}} \max(\|z\|_{H^s}^2, (\mu_{\min}(h))^{-2s}),$$

which are much less degenerate than the one associated with  $\omega_n$  (see (1.15)). Since, as before, the factor  $(\mu_{\max}(h))^{-C|h|_{\ell^1}}$  can be overcome by assuming high-regularity ( $s$  large enough), such estimates are stable and imply similar small divisor estimates for the full modulated frequencies  $\omega_n^\sharp$ . It also implies that it suffices to average the terms of order  $\mathcal{O}(z^{2q+1})$  using the flow generated by the modulated frequencies  $\omega_n + \tilde{\omega}_n$ . Moreover, it also implies that it is not necessary to impose new Diophantine conditions on the initial data at each step of the normal form procedure, it suffices to ensure that (1.16) holds for  $z = u(0)$ .

– The second strategy, used in [Bou00a] and which we implement in this paper, is to average the terms of order  $\mathcal{O}(z^{2q+1})$  using the flow generated by all the modulated frequencies  $\omega_n^\sharp$ . To prove that they satisfy acceptable small divisor estimates, it is important to note that even if  $\omega_n$  is not the leading part of  $\omega_n^\sharp$ ,  $\partial_{|z_n|^2} \omega_n = -f'(0)$  is necessarily the leading part of  $\partial_{|z_n|^2} \omega_n^\sharp$ . This allows to prove that if  $z(0)$  lives in a set of asymptotically full measure, the modulated frequencies  $\omega_n^\sharp$  enjoy the same small divisor estimate (1.15) as  $\omega_n$ . As discussed above, a significant part of the proof (not considered in [Bou00a]) then consists in proving that this corresponds to a set of initial data  $u(0)$  of asymptotically full measure.

The first strategy suffers from several limitations compared to the second one, which is more flexible and which we follow in this paper.

– The higher order correction terms (such as  $\tilde{\omega}_n$ ) depend on the linear frequencies (here  $\lambda_n^2$ ). To get non-degenerate small divisor estimates of the type (1.16), it is then necessary to impose additional assumptions on the linear frequencies (e.g. here we would have to impose additional Diophantine conditions on the numbers  $\lambda_n^{-2}$ , which would be much less explicit than the admissibility condition (1.5)).

– For some equations these higher order correction terms may also be degenerate. For the Benjamin-Ono equation they vanish exactly. In [BG21] it was then necessary to assume the existence a term of the form  $\partial_x u^3$  in the perturbed equation in order to follow the rational normal form approach.

– Depending on the equation at hand, and possibly also on preliminary Birkhoff normal form reductions (to remove the non-resonant terms from the Hamiltonian), obtaining explicit expressions for these higher order correction terms requires heavy computation. In [BG21] it was necessary to use of a formal computation software.

<sup>(9)</sup>where  $C_{|h|_{\ell^1}} > 0$  is an explicit function of  $|h|_{\ell^1}$  which plays no role.

### 1.3.2. Algebraic framework: rational fractions versus re-centered polynomials

The algebraic framework of the papers based on rational normal forms is quite different from that of this paper and [Bou00a]. Rational normal forms are quite close to standard Birkhoff normal form since it consists in replacing polynomial expansions by rational fraction expansions (which explains their name). Conversely, here, as in [Bou00a], we consider expansions in re-centered<sup>(10)</sup> polynomial expansions which hold much more locally (more in a KAM spirit: see [Bou05, BMP21] where a similar formalism is used in this context). The second approach offers more flexibility in the construction whereas the first one provides a normal form which is more global. This choice of algebraic framework does not seem to be related to small divisor considerations: it is likely that we could replace rational fractions by re-centered polynomials in [BFG20b, BG21, LX24, BCGW24]. Nevertheless, here and in [Bou00a], to overcome the small divisor degeneracy (i.e., that we do not have (1.16)), the normal form is constructed more locally (around some high-dimensional tori in  $H^s$ ) and it is unclear how a formalism based on rational fractions would allow to recover the smallness estimates provided by the fact that the solution remains very close to high-dimensional tori. We refer to the end of Section 1.3.4 for further discussions about the way we exploit this extra smallness.

**1.3.3. Reduction to a finite dimensional system.** — In order to put (NLS) in Birkhoff normal form, the frequencies of the system must satisfy for all  $q \geq 2$  a Diophantine condition of type

$$|\lambda_{n_1}^2 - \lambda_{n_2}^2 + \cdots - \lambda_{n_{2q}}^2| \neq 0 \implies |\lambda_{n_1}^2 - \lambda_{n_2}^2 + \cdots - \lambda_{n_{2q}}^2| \gtrsim_q \max_3 \langle n \rangle^{-\beta_q},$$

where  $\beta_q > 0$ ,  $n_1, \dots, n_{2q} \in \mathbb{Z}^d$  satisfy the zero momentum condition  $n_1 - n_2 + \cdots - n_{2q} = 0$  and  $\max_3 \langle n \rangle$  denotes the third largest number among  $\langle n_1 \rangle, \dots, \langle n_{2q} \rangle$ . Such small divisor estimates are obvious for integer frequencies, and also holds for generic choices of convolution potential in (NLS\*) on rational tori. In contrast, for Diophantine flat tori this estimate does not hold in general with  $\max_3 \langle n \rangle$  but with  $\max_1 \langle n \rangle$ . For this reason the almost global stability of small solution of (NLS\*) for irrational tori remained an open problem until the recent work of Bambusi–Feola–Montalto [BFM24].

To address this, [BFM24] used cluster decomposition and frequency separation properties between different clusters, initially established by Bourgain [Bou98] for the square torus  $\mathbb{T}^d$  and later generalized by Berti–Maspero [BM19] for any flat tori (see Lemma 2.15). This allows them, while leaving certain quasi-resonant terms in their normal form, to prove the quasi-preservation of the super actions associated with this decomposition and therefore to control the  $H^s$  norm of the solution for very long times.

In Section 2 we detail the finite-dimensional approximation. We use the cluster decomposition to prove that, up to a first change of variable  $\Phi_\chi^1$  and a bootstrap

---

<sup>(10)</sup>around the actions in the final variables, which is quite implicit.

assumption, the super actions are almost preserved and that the low-frequency part of the solution  $\Pi_{\mathcal{E}_{10}} u(t)$  solves an evolution equation associated with a truncated resonant Hamiltonian  $H \circ \Phi_\chi^1 \circ \Pi_{\mathcal{E}_{10}}$ , up to a very small remainder term. However, we have added a number of technical refinements to ensure that the change of variable  $\Phi_\chi^1$  does not destroy our measure estimates (of the type (1.9)). The results of Section 2 reduce the proof of Theorem 1.3 to the almost global stability of a truncated resonant Hamiltonian system, as stated in Theorem 2.7. Since the system is truncated, it can be put on Birkhoff normal form (almost) for free and we can fully exploit the fact that quartic resonant terms of the Hamiltonian are integrable.

**1.3.4. A normal form with a two parameters scale.** — Starting from Section 3, as in [Bou00a], we only use the first correction term to modulate the frequencies. As a consequence, the small divisors estimates are of the form (1.15). Due to the degeneracy of these small divisors, we believe that an approach in the spirit of rational normal forms, in which we would remove the non-integrable terms degree by degree (like in classical Birkhoff normal forms) would fail. Indeed, with such an approach, there does not seem to exist any reasonable way to absorb the losses due to the factor  $(\mu_{\min}(h))^{-2s}$ . Instead, as in [Bou00a], we introduce a two parameters scale to decompose and classify the polynomials appearing in our Taylor expansions. In addition to its presentation given below, we also refer to the Section 3 of the proceeding [Cam24] for further explanations.

More precisely, following [Bou00a] we decompose a polynomial as a sum of monomials of the form

$$(1.17) \quad \prod_{n \in \mathbb{Z}^d} u_n^{k_n} \overline{u_n}^{\ell_n} (|u_n|^2 - \xi_n)^{m_n},$$

where  $k, \ell, m \in \mathbb{N}^{\mathbb{Z}^d}$  are some finitely supported families of indices counting the multiplicities satisfying the condition  $\ell_n k_n = 0$  for all  $n \in \mathbb{Z}^d$  (to ensure the uniqueness of the decomposition). The parameters  $\xi_n \in \mathbb{R}$  modulate the frequencies. They are implicitly the actions of the initial data in the final variables. To each non-integrable monomial of the form (1.17) we associate a frequency scale  $N_\alpha$  such that

$$(1.18) \quad N_\alpha \leq \mathbf{n}_- < N_{\alpha+1}, \quad \text{where} \quad N_\alpha^s = \varepsilon^{-\alpha/200},$$

$\varepsilon$  is the size of the initial data in  $H^s$  and  $\mathbf{n}_-$  is the size of the smallest index appearing in the associated small divisor, i.e.,

$$\mathbf{n}_- = \min\{|n| \mid n \in \mathbb{Z}^d, k_n + \ell_n \geq 1\}.$$

The parameter  $\alpha$  classifies polynomials according to the cost of the associated small divisors. Monomials at scale  $\alpha$  as in (1.18) satisfy two properties:

- They preserve the actions  $|u_n|^2$  with  $|n| < N_\alpha$ . This type of property is also central in papers in low regularity like [BGR23].
- Assuming moreover that they are resonant (or quasi-resonant in some sense to be specified), since they are not integrable and quartic resonant terms are integrable,

we have  $\sum_n k_n + \ell_n \geq 6$ . Consequently, they generate vector fields from  $H^s$  to  $H^s$  of order at most  $\varepsilon^5 N_\alpha^{-4s}$  (see Proposition 4.6).

The cost of the small divisor is at most of order  $\varepsilon^{-2} N_{\alpha+1}^{2s}$  (see (1.15)). Since, by construction, the ratio between two consecutive scales  $(N_{\alpha+1}/N_\alpha)^{2s} \lesssim \varepsilon^{-1/100}$  is not too large, we see that there is room to remove these monomials (they could be associated to quite large coefficients).

The heart of the proof consists in removing, by induction on the frequency scales  $N_\alpha$ , the monomials satisfying  $N_\alpha \leq \mathbf{n}_- < N_{\alpha+1}$  (see Theorem 6.2). For this purpose, we consider for each  $\alpha$  a norm  $Y_\alpha^{\text{sup}}$  well suited to polynomials composed of monomials satisfying  $N_\alpha \leq \mathbf{n}_-$ . Then imposing new Diophantine conditions on  $\xi$ , we construct by induction (in Proposition 6.19) canonical changes of variables to make these norms smaller and smaller in terms of powers of  $\varepsilon$  (this is the second scale of parameters) for polynomials composed of monomials satisfying  $N_\alpha \leq \mathbf{n}_- < N_{\alpha+1}$ . At the end of this second induction, the monomials satisfying  $N_\alpha \leq \mathbf{n}_- < N_{\alpha+1}$  are associated with very small coefficients, and do not contribute to the dynamics over time scales of order  $\varepsilon^{-r}$ .

However, in the second induction, the terms generated by the Poisson bracket with an *integrable monomials with low-indices*  $|n| \ll N_\alpha$ , and only one index at frequency scale  $\sim N_\alpha$ , do not seem to be sufficiently small to compensate for the small divisor loss  $\sim N_\alpha^{-2s}$ . To address these terms, Bourgain made the following observation: re-centered actions are much smaller than the actions themselves (which we prove are expected to be slow variables). This observation provides an extra smallness factor for re-centered actions  $||u_n|^2 - \xi_n|$  with a gain  $N_\alpha^{-2s}$ , even for low indices  $|n| \ll N_\alpha$ . The definition of the non-resonant neighborhood (4.8) encodes this gain, which is a crucial ingredient to address the degenerate small-divisor losses. At this point, we also see that it is natural to introduce the frequency scale to compensate the small divisor losses thanks to this extra smallness.

When we reach a certain frequency scale  $N_\alpha$  large enough (with the notations of the paper, when  $\alpha = \beta$ ), we conclude that the actions are almost preserved over very long times using that the only remaining monomials are either integrable or associated with negligible coefficients: they generate vector fields of order at most  $N_\beta^{-4s} \ll \varepsilon^{-2r}$  (see estimate (7.15)).

#### 1.4. NOTATIONS AND FUNCTIONAL SETTING

**1.4.1. Functional setting.** — We equip the torus  $\mathbb{T}^d = \mathbb{R}^d/2\pi\mathbb{Z}^d$  with the normalized Lebesgue measure  $(2\pi)^{-d}dx$ . Therefore, the Lebesgue norms  $\|\cdot\|_{L^p}$ ,  $1 \leq p < \infty$ , are defined by density through the formula

$$\forall u \in C^0(\mathbb{T}^d), \quad \|u\|_{L^p}^p := (2\pi)^{-d} \int_{\mathbb{T}^d} |u(x)|^p dx.$$

We identify each distribution  $u$  on  $\mathbb{T}^d$  with the sequence of its Fourier coefficients defined by density by

$$u_k = (2\pi)^{-d} \int_{\mathbb{T}^d} u(x) e^{-ikx} dx.$$

Given  $1 \leq p < \infty$  and  $s \in \mathbb{R}$  we set

$$\ell_s^p(\mathbb{Z}^d) := \{u \in \mathbb{C}^{\mathbb{Z}^d} \mid \|u\|_{\ell_s^p}^p := \sum_{k \in \mathbb{Z}^d} \langle k \rangle^{ps} |u_k|^p < \infty\}.$$

As usual, we set

$$H^s(\mathbb{T}^d) = h^s(\mathbb{Z}^d) = \ell_s^2(\mathbb{Z}^d) \quad \text{and} \quad \ell^p(\mathbb{Z}^d) := \ell_0^p(\mathbb{Z}^d).$$

Note that with these conventions, the Fourier–Plancherel isometry writes

$$\forall u \in L^2(\mathbb{T}^d), \quad \|u\|_{L^2}^2 = \|u\|_{\ell^2}^2.$$

Throughout this paper, we consider  $L^2$  and  $\ell^2$  as real Hilbert spaces equipped with the scalar products

$$\forall u, v \in L^2(\mathbb{T}^d), \quad (u, v)_{L^2} := (2\pi)^{-d} \operatorname{Re} \int_{\mathbb{T}^d} u(x) \overline{v(x)} dx = \operatorname{Re} \sum_{k \in \mathbb{Z}^d} u_k \overline{v_k} =: (u, v)_{\ell^2}.$$

Note that if  $s \in \mathbb{N}$  is a nonnegative integer, we have

$$\forall u \in H^s(\mathbb{T}^d), \quad \|u\|_{H^s}^2 = (2\pi)^{-d} \int_{\mathbb{T}^d} |\partial_x^s u(x)|^2 dx.$$

**1.4.2. Differential calculus and Poisson brackets.** — Given  $p \in [1, \infty)$ ,  $s \in \mathbb{R}$ ,  $\mathcal{U}$  an open subset of  $\ell_s^p(\mathbb{Z}^d)$ , a smooth function  $P : \mathcal{U} \rightarrow \mathbb{R}$  and  $u \in \ell_s^p(\mathbb{Z}^d)$ , its gradient  $\nabla P(u)$  is the unique element of  $\ell_{-s}^{p'}(\mathbb{Z}^d)$  satisfying

$$\forall v \in \ell_s^p(\mathbb{Z}^d), \quad (\nabla P(u), v)_{L^2} = dP(u)(v).$$

It can be checked that

$$(1.19) \quad \forall k \in \mathbb{Z}^d, \quad (\nabla P(u))_k = 2\partial_{\overline{u_k}} P(u).$$

We equip  $L^2(\mathbb{T}^d)$  with the usual symplectic form  $(i\cdot, \cdot)_{L^2}$ . Therefore, provided that  $\ell_s^p(\mathbb{Z}^d) \subset \ell^2(\mathbb{Z}^d)$ , a smooth map  $\tau : \mathcal{U} \rightarrow \ell_s^p(\mathbb{Z}^d)$  is *symplectic* if

$$\forall u \in \mathcal{U}, \forall v, w \in \ell_s^p(\mathbb{Z}^d), \quad (iv, w)_{L^2} = (id\tau(u)(v), d\tau(u)(w))_{L^2}.$$

Moreover, if  $H, K : \mathcal{U} \rightarrow \mathbb{R}$  are two smooth functions such that  $\nabla H$  (or  $\nabla K$ ) is  $\ell_s^p(\mathbb{Z}^d)$  valued then the *Poisson bracket* of  $H$  and  $K$  is defined by

$$\{H, K\}(u) := (i\nabla H(u), \nabla K(u))_{L^2}.$$

Note that, as usual, we have

$$(1.20) \quad \{H, K\} = 2i \sum_{k \in \mathbb{Z}^d} \partial_{\overline{u_k}} H \partial_{u_k} K - \partial_{u_k} H \partial_{\overline{u_k}} K.$$

Note that this relation allows to extend the Poisson bracket to the case where  $H$  and  $K$  are not real valued.

1.4.3. *Subspaces and projections.* — Given any subset  $\mathcal{A} \subset \mathbb{Z}^d$ , we always consider  $\mathbb{C}^{\mathcal{A}}$  as the subspace of  $\mathbb{C}^{\mathbb{Z}^d}$  of the sequences supported on  $\mathcal{A}$ . We denote by  $\Pi_{\mathcal{A}} : \mathbb{C}^{\mathbb{Z}^d} \rightarrow \mathbb{C}^{\mathcal{A}}$  the projection defined by restriction. We extend implicitly any function  $F$  on  $\mathbb{C}^{\mathcal{A}}$  to a function on  $\mathbb{C}^{\mathbb{Z}^d}$  by  $F = F \circ \Pi_{\mathcal{A}}$ . As a consequence, any function on  $\ell^2(\mathcal{A}; \mathbb{C}) := \ell^2(\mathbb{Z}^d) \cap \mathbb{C}^{\mathcal{A}}$  can be seen as a smooth function in  $\ell^2(\mathbb{Z}^d)$  and so all the previous definitions make sense. In particular, we note that if  $F : \mathbb{C}^{\mathcal{A}} \rightarrow \mathbb{R}$  is  $C^1$  then  $\nabla F$  is  $\mathbb{C}^{\mathcal{A}}$  valued. Given a positive number  $N$ , we set  $\Pi_{\geq N} := \Pi_{\{n \in \mathbb{Z}^d \mid |n| \geq N\}}$ ,  $\Pi_N := \Pi_{\{n \in \mathbb{Z}^d \mid |n| \leq N\}}$  and  $\mathbb{Z}_N^d = \{n \in \mathbb{Z}^d \mid |n| \leq N\}$ .

*Acknowledgements.* — The authors would like to thank the anonymous referee for the careful reading of the manuscript and for the valuable comments and suggestions. J.B. would like to thank E. Faou and B. Grébert for enthusiastic discussions about this problem many years ago.

## 2. LOW FREQUENCY REDUCTION

In this section, we reduce the proof of Theorem 1.3 to that of Theorem 2.7 on the stability of small solutions of a finite-dimensional dynamical system. First, we introduce some notations about polynomials on  $\ell^1$  in order to formulate Theorem 2.7 in a second subsection. Then we prove some a priori estimates for the time-variation of the high-frequency super-actions, together with some mismatch lemmas to control the remainder terms generated by the truncations. Finally, we prove that Theorem 2.7 implies Theorem 1.3. The dimension  $d \geq 1$ , the admissible torus  $\mathbb{T}_{\mathcal{L}}^d$  and the nonlinearity  $f$  of (NLS) are considered as fixed. With a few exceptions, we do not follow in details the dependencies with respect to these constants.

Many of the proof's ingredients are the same as in [BFM24]. However, we use them differently. The main reason is that changes of variables mixing high modes and low modes are, a priori, not compatible with our measure estimates (and in particular that it suffices to draw the low modes of the initial datum randomly to get the almost global existence of the solution to (NLS)). Thus, contrary to [BFM24], we have to consider two kinds of Birkhoff-like normal forms for (NLS).

– On the one hand, we consider (in the proof of Proposition 2.18) a normal form in which we have removed all the terms whose small divisors are larger or equal to 1. It allows us to prove, under a bootstrap assumption, the almost preservation of the high super-actions in Proposition 2.18. It does not require any non-resonance condition but just small and smooth solutions.

– On the other hand, in Section 2.6, we consider a normal form in which we essentially<sup>(11)</sup> remove all non-resonant terms that do not involve high modes. By construction, the associated change of variable is the identity on high mode which is good

---

<sup>(11)</sup>It would be true if  $\mathbb{T}_{\mathcal{L}}^d$  satisfied a stronger Diophantine condition. Here, since we did not assume non-resonance conditions for terms of degree larger than or equal to 6, some quasi-resonant terms may remain. However, these terms do not pose significant difficulties and can be ignored in a first reading.

for the measure estimates. This normal form allows us to decouple (in the new variables), up to a bootstrap assumption, the low modes dynamics from the high modes dynamics (thanks to the mismatch lemmas 2.22 and 2.23). In other words, it reduces the analysis of the stability of the low modes of (NLS) to the one of a finite (but high) dimensional resonant Hamiltonian system (of which the rest of the paper is devoted).

**2.1. A FIRST HAMILTONIAN FORMALISM.** — We first introduce a set of multi-indices to describe the polynomials:

**DEFINITION 2.1.** — For  $q \geq 1$ , we define  $\mathcal{N}_{2q}$  the set of multi-indices of degree  $2q$  with zero momentum

$$\mathcal{N}_{2q} = \{\vec{n} = (n_1, \dots, n_{2q}) \in (\mathbb{Z}^d)^{2q} \mid \sum_{i=1}^{2q} (-1)^i n_i = 0\}.$$

Given  $\vec{n} \in \mathcal{N}_{2q}$ , we denote the decreasing rearrangement of  $(|n_1|, \dots, |n_{2q}|)$  by

$$n_1^* \geq \dots \geq n_{2q}^*.$$

Now, we introduce our main class of polynomials.

**DEFINITION 2.2** (Real homogeneous polynomial). — Let  $q \geq 2$  and  $\mathcal{A} \subset \mathbb{Z}^d$ . The set  $\mathcal{H}_{2q}(\mathcal{A})$  of real homogeneous polynomials of degree  $2q$  supported on  $\ell^1(\mathcal{A})$  corresponds to the set of functions

$$Q(u) = \sum_{\vec{n} \in \mathcal{N}_{2q}} Q_{\vec{n}} u_{n_1} \overline{u_{n_2}} \cdots \overline{u_{n_{2q}}} =: \sum_{\vec{n} \in \mathcal{N}_{2q}} Q_{\vec{n}} u_{\vec{n}},$$

where  $(Q_{\vec{n}}) \in \mathbb{C}^{\mathcal{N}_{2q}}$  satisfies

(1) (symmetry condition)

$$Q_{\vec{n}} \text{ is symmetric in } (n_1, n_3, \dots, n_{2q-1}) \text{ and in } (n_2, n_4, \dots, n_{2q}).$$

(2) (reality condition)

$$\text{For all } \vec{n} = (n_1, n_2, \dots, n_{2q}) \in \mathcal{N}_{2q}, \quad Q_{(n_1, n_2, \dots, n_{2q})} = \overline{Q_{(n_{2q}, n_{2q-1}, \dots, n_1)}}.$$

(3) (boundedness condition)

$$\|Q\|_{\infty} := \sup_{\vec{n} \in \mathcal{N}_{2q}} |Q_{\vec{n}}| < +\infty.$$

(4) (support)

$$Q_{(n_1, n_2, \dots, n_{2q})} \neq 0 \implies \vec{n} \in \mathcal{A}^{2q}.$$

**REMARK 2.3.** — Thanks to the boundedness condition, these polynomials are smooth functions on  $\ell^1(\mathbb{Z}^d)$  and the support condition only means that  $Q = Q \circ \Pi_{\mathcal{A}}$ .

**REMARK 2.4.** — Note that this definition is well suited to the Hamiltonian  $H$  of (NLS) (defined by (1.13)). More precisely, in Fourier variables, the Taylor expansion of  $H$  writes

$$H(u) \underset{u \rightarrow 0}{=} \frac{1}{2} \sum_{n \in \mathbb{Z}^d} \lambda_n^2 |u_n|^2 + \frac{1}{2} \sum_{q \geq 2} \frac{f^{(q-1)}(0)}{q!} \sum_{\vec{n} \in \mathcal{N}_{2q}} u_{\vec{n}}$$

or more quantitatively, for all  $r \geq 2$ ,

$$(2.1) \quad \begin{aligned} H(u) &= \frac{1}{2} \sum_{n \in \mathbb{Z}^d} \lambda_n^2 |u_n|^2 + \frac{1}{2} \sum_{q=2}^r \frac{f^{(q-1)}(0)}{q!} \sum_{\vec{n} \in \mathcal{N}_{2q}} u_{\vec{n}} + R^{(2r+2)}(u) \\ &=: H^{(\leq 2r)}(u) + R^{(2r+2)}(u), \end{aligned}$$

where  $R^{(2r+2)}$  is a remainder term of order  $2r+2$  in the sens that for all  $u \in B_s(1)$ ,  $s > d/2$

$$(2.2) \quad \|\nabla R^{(2r+2)}(u)\|_{h^s} \lesssim_{r,s} \|u\|_{h^s}^{2r+1}.$$

DEFINITION 2.5 (Resonant function). — Given a multi-index  $\vec{n} \in \mathcal{N}_{2q}$ , the resonant function  $\Omega_{\vec{n}}$  is

$$\Omega_{\vec{n}} = \sum_{n=1}^{2q} (-1)^{n+1} \lambda_n^2.$$

DEFINITION 2.6 (Quasi-resonant Hamiltonian). — Let  $\kappa > 0$ ,  $q \geq 2$ . A real homogeneous polynomial  $Q \in \mathcal{H}_{2q}$  is  $\kappa$ -resonant if for all  $\vec{n} \in \mathcal{N}_{2q}$ ,

$$|\Omega_{\vec{n}}| > \kappa \implies Q_{\vec{n}} = 0.$$

2.2. DYNAMICS OF THE LOW MODES : MAIN RESULT. — We set  $c_* := 2\tau_* + 1$  depending only on the geometry of the torus ( $\tau_*$  is the exponent associated to the admissibility of  $\mathbb{T}_{\mathcal{L}}^d$ , see Definition 1.1).

THEOREM 2.7. — For all  $r \geq 9$ , there exists  $\nu \leq \min((2c_*)^{-1}, (2d)^{-1})$  such that for all  $s > d$ , if  $\varepsilon \lesssim_{r,s} 1$  is small enough,  $M \geq 2$  is a truncation parameter satisfying

$$M \leq \varepsilon^{-\nu},$$

and  $H_{10}$  is a real valued polynomial supported on  $\ell^1(\mathbb{Z}_{\leq M}^d)$  of the form

$$(2.3) \quad H_{10}(u) = \frac{1}{2} \sum_{|n| \leq M} \left( \lambda_n^2 - \frac{f'(0)}{2} |u_n|^2 \right) |u_n|^2 + \sum_{j=3}^{2r} P^{(2j)}(u),$$

where  $P^{(2j)} \in \mathcal{H}_{2j}(\mathbb{Z}_{\leq M}^d)$  is a real homogeneous polynomial of degree  $2j$  which is  $\varepsilon^{c_*\nu}$  resonant and satisfies  $\|P^{(2j)}\|_{\infty} \leq \varepsilon^{-c_*\nu j}$ , then there exists an open set  $\Theta_{\varepsilon}^b \subset \Pi_M B_s(\varepsilon)$  such that

$$(2.4) \quad \text{meas}(\Theta_{\varepsilon}^b) \geq (1 - \varepsilon^{1/39}) \text{meas}(\Pi_M B_s(\varepsilon)),$$

and, if  $u \in C^1([-T_{\varepsilon}, T_{\varepsilon}]; \Pi_M h^s)$  is such that  $u(0) \in \Theta_{\varepsilon}^b$  and

$$\sup_{0 \leq t \leq T_{\varepsilon}} \|i\partial_t u - \nabla H_{10}(u)\|_{h^s} \leq \varepsilon^{3r} \rho$$

for some  $T_{\varepsilon} \leq \varepsilon^{-r}$  and some  $\rho > 0$  satisfying  $\|u(0)\|_{h^s} \leq \rho \leq 2\varepsilon$ , then

$$(2.5) \quad \sup_{0 \leq t \leq T} \sum_{|n| \leq M} \langle n \rangle^{2s} | |u_n(t)|^2 - |u_n(0)|^2 | \leq \frac{\varepsilon}{2} \rho^2.$$

2.3. A BIRKHOFF NORMAL FORM THEOREM AND SOME BASIC ESTIMATES. — In the energy estimates we will need the following estimate, which is a consequence of Cauchy–Schwarz and Young’s convolution inequalities:

LEMMA 2.8. — *Let  $q \geq 1$ ,  $\iota \in \{-1, 1\}^{2q}$  and  $a^{(1)}, \dots, a^{(2q)} \in \ell^1(\mathbb{Z}^d; \mathbb{R}_+)$ . We have*

$$\sum_{\substack{\vec{n}=(n_1, \dots, n_{2q}) \in (\mathbb{Z}^d)^{2q} \\ \iota_1 n_1 + \iota_2 n_2 + \dots + \iota_{2q} n_{2q} = 0}} a_{n_1}^{(1)} \dots a_{n_{2q}}^{(2q)} \leq \min_{\sigma \in \mathfrak{S}_{2q}} \|a^{(\sigma(1))}\|_{\ell^2} \|a^{(\sigma(2))}\|_{\ell^2} \prod_{i=3}^{2q} \|a^{(\sigma(i))}\|_{\ell^1}.$$

As usual, we deduce some useful estimate for the vector fields.

COROLLARY 2.9 (Vector field estimates). — *Given  $q \geq 2$ ,  $P \in \mathcal{H}_{2q}(\mathbb{Z}^d)$ ,  $s \geq 0$ ,  $P$  defines a smooth function on  $h^s \cap \ell^1$ ,  $\nabla P$  is smooth from  $h^s \cap \ell^1$  to  $h^s \cap \ell^1$  and for all  $u \in h^s \cap \ell^1$ , we have*

$$\|\nabla P\|_{h^s}(u) \lesssim_{q,s} \|P\|_{\infty} \|u\|_{h^s} \|u\|_{\ell^1}^{2q-2}.$$

Then, as usual, we deduce the local existence of the Hamiltonian flows of the polynomials of which we summarize the properties that will be useful for us in this paper.

COROLLARY 2.10 (Hamiltonian flows). — *Let  $s > d/2$ ,  $r \geq 2$ ,  $C \geq 1$ ,  $\mathcal{A} \subset \mathbb{Z}^d$ . For all  $\kappa \in (0, 1)$ , there exists  $\varepsilon_* \gtrsim_{r,s,C} \sqrt{\kappa}$  such that for all real polynomial  $\chi$  of degree smaller than or equal  $2r$  of the form*

$$\chi = \chi^{(4)} + \dots + \chi^{(2r)},$$

where

$$\forall j \in \{2, \dots, r\}, \quad \chi^{(2j)} \in \mathcal{H}_{2j}(\mathcal{A}) \text{ satisfies } \|\chi^{(2j)}\|_{\infty} \leq C\kappa^{-j+1},$$

there exists a smooth map

$$\Phi_{\chi} : \begin{cases} [-1, 1] \times \Pi_{\mathcal{A}}^{-1} \Pi_{\mathcal{A}} B_s(\varepsilon_*) \longrightarrow h^s(\mathbb{Z}^d) \\ (t, u) \longmapsto \Phi_{\chi}^t(u) \end{cases}$$

such that, for all  $u \in \Pi_{\mathcal{A}}^{-1} \Pi_{\mathcal{A}} B_s(\varepsilon_*)$ ,

$$i\partial_t \Phi_{\chi}^t(u) = (\nabla \chi) \circ \Phi_{\chi}^t(u) \quad \forall t \in [-1, 1],$$

and for all  $t \in [-1, 1]$ ,  $\Phi_{\chi}^t$  is symplectic,

– close to the identity

$$\|\Phi_{\chi}^t(u) - u\|_{h^s} \leq \left( \frac{\|\Pi_{\mathcal{A}} u\|_{h^s}}{\varepsilon_*} \right)^2 \|\Pi_{\mathcal{A}} u\|_{h^s},$$

– its derivative is not too big, i.e.,

$$(2.6) \quad \|d\Phi_{\chi}^t(u)\|_{h^s \rightarrow h^s} \leq 2$$

– and provided that  $\Phi_{\chi}^t(u) \in \Pi_{\mathcal{A}}^{-1} \Pi_{\mathcal{A}} B_s(\varepsilon_*)$ , we have  $\Phi_{\chi}^{-t}(\Phi_{\chi}^t(u)) = u$ .

REMARK 2.11. — These projections may seem unusual but they are just a consequence of the block diagonal structure of  $\Phi_\chi^t$  :

$$\Phi_\chi^t = \Pi_{\mathcal{A}}(\Phi_\chi^t)|_{h^s(\mathcal{A})} \oplus \text{id}_{h^s(\mathcal{A}^c)}.$$

We also recall the following continuity estimates of the Poisson brackets when applied to polynomials.

LEMMA 2.12. — *Let  $\mathcal{A} \subset \mathbb{Z}^d$  be a subset of  $\mathbb{Z}^d$ ,  $P \in \mathcal{H}_{2p}(\mathcal{A})$  a homogeneous real polynomial of degree  $2p \geq 4$  supported on  $\ell^1(\mathcal{A})$  and  $Q \in \mathcal{H}_{2q}(\mathcal{A})$  a real homogeneous polynomial of degree  $2q \geq 4$  supported on  $\ell^1(\mathcal{A})$  then their Poisson bracket  $\{P, Q\} \in \mathcal{H}_{2(p+q-1)}(\mathcal{A})$  is a real homogeneous polynomial of degree  $2(p+q-1)$  supported on  $\ell^1(\mathcal{A})$  satisfying the bound*

$$\|\{P, Q\}\|_\infty \lesssim pq\|P\|_\infty\|Q\|_\infty.$$

As usual, we deduce of these estimates the following Birkhoff normal form theorem.

THEOREM 2.13 (Quasi-resonant normal form). — *Let  $r \geq 2$ ,  $s > \max(d/2, 1)$ ,  $\mathcal{A} \subset \mathbb{Z}^d$ ,  $\kappa \in (0, 1)$  and  $P = P^{(4)} + \dots + P^{(2r)}$  be a real Hamiltonian of degree  $2r$  supported on  $\ell^1(\mathcal{A})$  with  $P^{(2j)} \in \mathcal{H}_{2j}(\mathcal{A})$  for all  $j \in \{2, \dots, r\}$  and set*

$$H = Z_2^{\mathcal{A}} + P \quad \text{where} \quad Z_2^{\mathcal{A}}(u) = \frac{1}{2} \sum_{n \in \mathcal{A}} \lambda_n^2 |u_n|^2.$$

*There exist some polynomials  $\chi^{(2j)} \in \mathcal{H}_{2j}(\mathcal{A})$ ,  $2 \leq j \leq r$ , satisfying*

$$\|\chi^{(2j)}\|_\infty \lesssim_{C,r} \kappa^{-j+1},$$

*where  $C = \max_{j \geq 2} \kappa^{2-j} \|P^{(2j)}\|_\infty$  such that, on  $\Pi_{\mathcal{A}}^{-1} \Pi_{\mathcal{A}} B_s(\varepsilon_*)$  (where  $\varepsilon_*$  is given by Corollary 2.10),*

$$H \circ \Phi_\chi^1 = Z_2^{\mathcal{A}} + Q^{(4)} + \dots + Q^{(2r)} + \Upsilon,$$

*where  $\chi = \chi^{(4)} + \dots + \chi^{(2r)}$ , the polynomials  $Q^{(2j)} \in \mathcal{H}_{2j}(\mathcal{A})$  are some  $\kappa$  resonant real homogeneous polynomials satisfying for all  $j \in \{2, \dots, r\}$  and all  $\vec{n} \in (\mathbb{Z}^d)^4$*

$$\|Q^{(2j)}\|_\infty \lesssim_{C,r} \kappa^{-j+2} \quad \text{and} \quad Q_{\vec{n}}^{(4)} = \mathbf{1}_{|\Omega_{\vec{n}}| \leq \kappa} P_{\vec{n}}^{(4)},$$

*and  $\Upsilon \in C^1(\Pi_{\mathcal{A}}^{-1} \Pi_{\mathcal{A}} B_s(\varepsilon_*); \mathbb{R})$  is a remainder of order  $2(r+1)$  in the sense that for all  $u \in \Pi_{\mathcal{A}}^{-1} \Pi_{\mathcal{A}} B_s(\varepsilon_*)$ ,*

$$(2.7) \quad \|\nabla \Upsilon(u)\|_{h^s} \lesssim_{C,r,s} \left( \frac{\|\Pi_{\mathcal{A}} u\|_{h^s}}{\sqrt{\kappa}} \right)^{2r} \|\Pi_{\mathcal{A}} u\|_{h^s}.$$

This is a formulation of Birkhoff normal form theorem, which is by now standard. For the proof, see [BG25, Th. 2.15] or [BGR23, Th. 2.12].

**2.4. ALMOST PRESERVATION OF THE HIGH-SUPER ACTIONS.** — In this subsection, we prove Proposition 2.14, in which we show that up to a bootstrap assumption we control the  $H^s$  norm of the high modes of the solution of (NLS) for very long times. Note that we do not need to use any internal or external parameters.

**PROPOSITION 2.14.** — *There exist  $\delta > 0$  and  $N_0 \geq 2$  depending only on  $\mathcal{L}$  such that given  $T > 0$ ,  $c > 1$ ,  $s > d$ ,  $\rho > 0$  and  $u$  a solution to (NLS) in  $C([0, T], h^s)$  with  $\|u(0)\|_{h^s} \leq \rho$  and*

$$(2.8) \quad \sup_{t \in [0, T]} \|u(t)\|_{h^s} \leq c\rho,$$

*then for all  $N \geq N_0$ , for all  $r \geq 2$ , provided that  $\varepsilon \lesssim_{c,r,s} 1$  is small enough and  $t \leq T$ ,*

$$\|\Pi_{\geq N} u(t)\|_{h^s}^2 \leq 2^{2s} \|u(0)\|_{h^s}^2 + \rho^3 + TN^{-(\delta/2)(s-d)} \rho^3 + T\rho^{2r+1}.$$

The proof is based on Birkhoff normal forms, and on the extension of Bourgain's cluster decomposition lemma [Bou98] to any flat tori, which was proved by Berti and Maspero in [BM19, Th. 2.1]. The use of this decomposition in the context of Birkhoff normal forms was recently initiated by Bambusi, Feola and Montalto in [BFM24]. It is a way to overcome the small divisor degeneracy discussed in Section 1.3.3. In the following, we will mainly focus on the proof of Proposition 2.18, from which Proposition 2.14 easily follows.

**LEMMA 2.15** (Clustering of eigenvalues of  $-\Delta_{\mathcal{L}}$ , [BM19, Th. 2.1]). — *There exist constants  $\delta \equiv \delta(d) \in (0, 1)$  and  $C(\mathcal{L}, d) \geq 2$  and a partition of  $\mathbb{Z}^d$*

$$\mathbb{Z}^d = \bigcup_v \mathcal{C}_v,$$

*satisfying the following properties:*

- (1) *The sets  $\mathcal{C}_v$  are finite and, up to a bounded set  $\mathcal{C}_{v_0}$  such that*

$$\max_{n \in \mathcal{C}_{v_0}} |n| \leq C(\mathcal{L}, d) \quad \text{and} \quad |n| < 2 \implies n \in \mathcal{C}_{v_0}.$$

- (2) *They are dyadic in the sense that*

$$(2.9) \quad \max_{n \in \mathcal{C}_v} |n| \leq 2 \min_{n \in \mathcal{C}_v} |n| =: 2m_v.$$

- (3) *If one denotes  $[n_1]$  the class of equivalence of  $n_1 \in \mathbb{Z}^d$ , that is,*

$$[n_1] = \mathcal{C}_v, \quad \text{where } v \text{ is such that } n_1 \in \mathcal{C}_v,$$

*then we have the separation property: for all  $n_1, n_2 \in \mathbb{Z}^d$ ,*

$$[n_1] \neq [n_2] \implies |n_1 - n_2| + |\lambda_{n_1}^2 - \lambda_{n_2}^2| > (|n_1| + |n_2|)^{\delta(d)}.$$

**REMARK 2.16.** — As usual, we identify the set of the indices  $v$  with the set of the equivalence classes.

We stress out that this lemma holds for any flat torus, without any Diophantine assumption on the metric. As in [BFM24], this decomposition allows us to design some almost conserved quantities of (NLS) called *super-action*.

DEFINITION 2.17 (Super actions). — Given a mode  $n_0 \in \mathbb{Z}^d$  the corresponding super-action is defined by

$$S_{n_0}(u) := \sum_{n \in [n_0]} |u_n|^2.$$

The cluster being dyadic (see (2.9)), Proposition 2.14 can easily be deduced from the following:

PROPOSITION 2.18 (A priori estimates for the high super-actions). — *Let  $T, u, c, s, \rho$  be as in Proposition 2.14. For all  $M \geq 2$ , for all  $r \geq 2$ , provided that  $\rho \lesssim_r 1$  is small enough and  $t \leq T$ , we have*

$$(2.10) \quad \sum_{m_v \geq M} m_v^{2s} |S_v(u(t)) - S_v(u(0))| \lesssim_{r,s,c} (\rho^4 + TM^{-(\delta/2)(s-d)} \rho^4 + T\rho^{2(r+1)}),$$

where  $m_v$  was defined in (2.9).

*Proof of Proposition 2.14 from Proposition 2.18.* — Indeed, to deduce Proposition 2.14 from Proposition 2.18, thanks to the dyadicity of the clusters (see (2.9)), it is enough to set  $N_0 = 3C(\mathcal{L}, d)$  and  $M = N/2$  to get that

$$\|\Pi_{\geq N} u\|_{h^s}^2 \leq 2^{2s} \sum_{|n| \geq N} m_{[n]}^{2s} |u_n|^2 \leq 2^{2s} \sum_{m_v \geq M} m_v^{2n} S_v(u),$$

and so, thanks to Proposition 2.18, that

$$\|\Pi_{\geq N} u(t)\|_{h^s}^2 \leq 2^{2s} \sum_{m_v \geq M} m_v^{2n} S_v(u(0)) + C_{r,s,c} \rho (\rho^3 + TN^{-(\delta/2)(s-d)} \rho^3 + T\rho^{2r+1}),$$

where  $C_{r,s,c}$  is a constant depending only on  $(r, s, c)$ . We then conclude the proof of Proposition 2.14 by using the estimate  $\sum_{m_v \geq M} m_v^{2s} S_v(u(0)) \leq \|u(0)\|_{h^s}^2$  and by absorbing the constant  $C_{r,s,c}$  thanks to the smallness of  $\rho$ .  $\square$

Before proving Proposition 2.18, let us consider the following technical lemmas.

LEMMA 2.19. — *Let  $\vec{n} \in \mathcal{N}_{2q}$  for some  $q \geq 2$  be such that  $|\Omega_{\vec{n}}| \leq 1$ . If there exists a frequency cluster  $\mathcal{C}_v$  such that*

$$\{S_v, u_{\vec{n}}\} \neq 0,$$

*then*

$$n_1^* \geq m_v \quad \text{and} \quad n_3^* \gtrsim_q (m_v)^{\delta/2}.$$

*Proof.* — First, we note that

$$\{S_v, u_{\vec{n}}\} = 2i \left( \sum_{j=1}^q \mathbf{1}_{n_{2j} \in \mathcal{C}_v} - \mathbf{1}_{n_{2j-1} \in \mathcal{C}_v} \right) u_{\vec{n}}.$$

As a consequence, if  $\{S_v, u_{\vec{n}}\} \neq 0$ , there exists  $j_0 \in \{1, \dots, 2q\}$  and an odd number  $\theta \in \mathbb{Z}$  such that  $n_{j_0} \in \mathcal{C}_v$  and  $n_{j_0+\theta} \notin \mathcal{C}_v$ . By symmetry, without loss of generality, we assume that  $j_0 = 1$  and  $j_0 + \theta = 2$ . Since  $n_1 \in \mathcal{C}_v$ , we have  $n_1^* \geq m_v$ . Moreover, thanks to the separation property of the clusters, we have

$$|n_1 - n_2| + |\lambda_{n_1}^2 - \lambda_{n_2}^2| > (|n_1| + |n_2|)^{\delta} \geq m_v^{\delta}.$$

As a consequence, either (by the zero momentum condition)

$$(2q-2) \max_{j \geq 3} |n_j| \geq |n_1 - n_2| \geq \frac{m_v^\delta}{2},$$

or (since  $|\Omega_{\vec{n}}| \leq 1$ )

$$(2q-2) \max_{j \geq 3} |n_j|^2 + 1 \gtrsim_{\mathcal{L}} |\lambda_{n_1}^2 - \lambda_{n_2}^2| \geq \frac{m_v^\delta}{2}.$$

In any case, we have

$$n_3^* \geq \max_{j \geq 3} |n_j| \gtrsim_{\mathcal{L}, q} m_v^{\delta/2}. \quad \square$$

LEMMA 2.20. — *Let  $s \geq \sigma \geq 0$ ,  $q \geq 2$  and  $Q \in \mathcal{H}_{2q}$ . If  $Q$  is 1-resonant then for all  $u \in h^s \cap \ell_\sigma^1$ , we have*

$$\sum_v m_v^{2s+\sigma\delta/2} |\{S_v, Q\}(u)| \lesssim_{q,s} \|Q\|_\infty \|u\|_{h^s}^2 \|u\|_{\ell_\sigma^1} \|u\|_{\ell^1}^{2q-3}.$$

*Proof.* — First, we note that if  $v = v_0$  then  $m_v = 0$  so we do not have to consider this class in the sum. Therefore, we consider a class  $v \neq v_0$ . Note that it implies that  $m_v \geq 2$ . Then, we note that

$$\{S_v, Q\} = \sum_{\vec{n} \in \mathcal{N}_{2q}} Q_{\vec{n}} \{S_v, u_{\vec{n}}\}.$$

Using Lemma 2.19, we deduce that we can reduce the sum to the set of the indices satisfying  $n_1^* \geq m_v \geq 2$  and

$$n_3^* \gtrsim_q m_v^{\delta/2}.$$

Moreover, using the zero momentum condition, we note that  $n_2^* \geq (2q)^{-1} n_1^* \geq (2q)^{-1} m_v$ . It follows that

$$\begin{aligned} \sum_v m_v^{2s+\sigma\delta/2} |\{S_v, Q\}(u)| &\lesssim_q \sum_v m_v^{2s+\sigma\delta/2} \sum_{\substack{n_2^* \gtrsim_q m_v \\ n_3^* \gtrsim_q m_v^{\delta/2}}} |Q_{\vec{n}}| |u_{n_1}| \cdots |u_{n_{2q}}| \\ &\lesssim_{q,s} \|Q\|_\infty \sum_{\vec{n} \in \mathcal{N}_{2q}} (n_1^*)^s (n_2^*)^s (n_3^*)^{\delta\sigma/2} |u_{n_1}| \cdots |u_{n_{2q}}| \\ &\lesssim_{q,s} \|Q\|_\infty \sum_{\varphi \in \mathfrak{S}_{2q}} \sum_{\vec{n} \in \mathcal{N}_{2q}} |n_{\varphi_1}|^s |n_{\varphi_2}|^s |n_{\varphi_3}|^{\delta\sigma/2} |u_{n_1}| \cdots |u_{n_{2q}}|. \end{aligned}$$

And so we get the expected estimate by applying the Young estimate of Lemma 2.8.  $\square$

By applying this lemma with  $\sigma = s - d$  and by considering only the classes such that  $m_v$  is large enough we deduce the following estimate.

COROLLARY 2.21. — *Let  $s \geq d$ ,  $q \geq 2$  and  $Q \in \mathcal{H}_{2q}$ . If  $Q$  is 1-resonant, then, for all  $u \in h^s$  and  $M \geq 2$ , we have*

$$(2.11) \quad \sum_{m_v \geq M} m_v^{2s} |\{S_v, Q\}(u)| \lesssim_{q,s} \|Q\|_\infty M^{-(\delta/2)(s-d)} \|u\|_{h^s}^3 \|u\|_{\ell^1}^{2q-3}.$$

Now we start proving Proposition 2.18 about the preservation of the super-actions of (NLS).

*Proof of Proposition 2.18.* — Recall the Hamiltonian formulation (1.13) of (NLS) and the Taylor expansion  $H = H^{(\leq 2r)} + R^{(2r+2)}$  of its Hamiltonian (2.1). We apply the Birkhoff normal form theorem 2.13 with  $\kappa = 1$  and  $\mathcal{A} = \mathbb{Z}^d$  to the Hamiltonian  $H^{(\leq 2r)}$  of degree  $2r$  to put it in 1-resonant normal form up to order  $2r$ . This gives an auxiliary Hamiltonian  $\chi = \chi^{(4)} + \dots + \chi^{(2r)}$  with  $\|\chi^{(2j)}\|_\infty \lesssim_r 1$  and, through Corollary 2.10, a radius  $\varepsilon_* \gtrsim_{r,s} 1$  such that on  $B_s(\varepsilon_*)$

$$\tilde{H} := H^{(\leq 2r)} \circ \Phi_\chi^1 = Z_2 + \sum_{q=2}^r \tilde{Q}^{(2q)} + \Upsilon =: \tilde{H}^{(\leq 2r)} + \Upsilon,$$

where  $\tilde{Q}^{(2q)} \in \mathcal{H}_{2q}$  are 1-resonant homogeneous polynomials satisfying  $\|Q^{(2q)}\|_\infty \lesssim_q 1$ , and the remainder  $\Upsilon$  is of order  $2(r+1)$  in the sense that for all  $w \in B_s(\varepsilon_*)$

$$(2.12) \quad \|\nabla \Upsilon(w)\|_{h^s} \lesssim_{r,s} \|w\|_{h^s}^{2r+1}.$$

From now on, we assume that the solution  $u$  of (NLS) satisfies  $\|u(0)\|_{h^s} \leq \rho \leq 10^{-2} c^{-1} \varepsilon_*$ . It follows that for all  $t \in [0, T]$ , we have  $\|u(t)\|_{h^s} \leq c\rho \leq 10^{-2} \varepsilon_*$  and so that  $v = \Phi_\chi^{-1}(u) \in C^0([0, T]; h^s)$  is well-defined for  $t \in [0, T]$ . Moreover,  $\Phi_\chi^{-1}$  being close to the identity, we have

$$(2.13) \quad \|u(t) - v(t)\|_{h^s} \leq \varepsilon_*^{-2} \|u(t)\|_{h^s}^3 \lesssim_{r,c} \rho^3, \quad \text{and so} \quad \|v(t)\|_{h^s} \lesssim_{r,c} \rho.$$

Without loss of generality, we can assume that<sup>(12)</sup>  $u(0) \in h^{s+2}$ , so that  $u \in C^1([0, T]; h^s)$  and  $v \in C^1([0, T]; h^s)$ . Since  $\Phi_\chi^{-1}$  is symplectic, using the chain rule, we get that

$$i\partial_t v = \nabla \tilde{H}^{(\leq 2q)}(v) + \nabla \Upsilon(v) + d\Phi_\chi^{-1}(u)(\nabla R^{(2r+2)}(u)).$$

As a consequence, one can bound the variation of the super-actions as follows:

$$(2.14) \quad \sum_{m_v \geq M} m_v^{2s} \left| \frac{d}{dt} S_v(v(t)) \right| \leq \sum_{m_v \geq M} m_v^{2s} |\{S_v, \tilde{H}^{(\leq 2r)}\}(v)| + \|\nabla \Upsilon(v)\|_{h^s} \|v\|_{h^s} \\ + \|d\Phi_\chi^{-1}(u)\|_{h^s \rightarrow h^s} \|\nabla R^{(2r+2)}(u)\|_{h^s} \|v\|_{h^s}.$$

According to Corollary 2.21 the first term on the right-hand side is bounded (up to a constant depending only on  $r$  and  $s$ ) by

$$\max_{2 \leq q \leq r} \|\tilde{Q}^{(2q)}\|_\infty M^{-(\delta/2)(s-d)} \|v(t)\|_{h^s}^3 \|v(t)\|_{\ell^1}^{2q-3} \lesssim_{r,s,c} M^{-(\delta/2)(s-d)} \rho^4.$$

Moreover, the estimate (2.6) on  $d\Phi_\chi^t$  together with the estimate (2.2) on  $\nabla R^{(2r+2)}$  yield

$$\|d\Phi_\chi^{-1}(u)\|_{h^s \rightarrow h^s} \|\nabla R^{(2r+2)}(u)\|_{h^s} \lesssim_{r,s,c} \rho^{2r+1}.$$

<sup>(12)</sup>It can be done by approximating any  $u(0)$  by smoother initial data (e.g. trigonometric polynomials) and using the preservation of the regularity for (NLS) (i.e., that while the  $\ell^1$  norm of the solution is bounded, its  $h^{s+2}$  norm does not blow up). Such an approximation is classical and is done in details, for example, in [BG25].

Using also the estimate (2.12) on  $\nabla \Upsilon$ , we conclude that

$$\sum_{m_v \geq M} m_v^{2s} \left| \frac{d}{dt} S_v(v(t)) \right| \lesssim_{c,r,s} M^{-(\delta/2)(s-d)} \rho^4 + \rho^{2r+1}.$$

Finally, using the fundamental theorem of calculus and that  $v$  is close to  $u$  (see (2.13), it provides the term  $\rho^4$  without factor  $T$  in (2.10)), and we get the expected estimate.  $\square$

**2.5. MISMATCH LEMMAS.** — In order to deduce Theorem 1.3 from Theorem 2.7, we will have to consider truncated subsystems. In this subsection, we establish two technical lemmas proving that the associated truncation remainder terms are very small.

**LEMMA 2.22** (Mismatch vector field estimate: I). — *For all  $q \geq 2$ ,  $s > d$ ,  $u \in h^s$ , all  $M, N \geq 2$  with  $N < \frac{1}{2q}M$ , and for all  $Q \in \mathcal{H}_{2q}$  such that  $Q \circ \Pi_M = 0$ , we have*

$$\|\Pi_N \nabla Q(u)\|_{h^s} \lesssim_q M^{-(s-d)} \|Q\|_\infty \|u\|_{h^s}^3 \|u\|_{\ell^1}^{2q-4}.$$

Note that we do not ask for any constraint on the size of the resonance function. The gap between  $N$  and  $M$  is sufficiently large to directly have the mismatch estimate.

*Proof.* — For  $n \in \mathbb{Z}^d$  with  $|n| \leq N$  and  $u \in h^s$ ,

$$\begin{aligned} \langle n \rangle^s |\partial_{u_n} Q(u)| &\leq 2q \langle n \rangle^s \sum_{\substack{\vec{n} \in \mathcal{N}_{2q} \\ n_{2q} = n}} |Q_{\vec{n}}| |u_{n_1}| \cdots |u_{n_{2q-1}}| \\ &\leq 2q \|Q\|_\infty \langle n \rangle^s \sum_{\substack{\vec{n} \in \mathcal{N}_{2q} \\ n_{2q} = n}} \mathbf{1}_{n_1^* \geq M} |u_{n_1}| \cdots |u_{n_{2q-1}}|. \end{aligned}$$

From the zero momentum condition, we have that  $\vec{n} \in \mathcal{N}_{2q}$  contributes only if  $n_1^* \geq M$  and

$$n_2^* \geq \frac{n_1^*}{2q} \geq \frac{M}{2q} > N \geq n.$$

As a consequence, we have  $|n| \leq n_3^*$  and so

$$\begin{aligned} \langle n \rangle^s |\partial_{u_n} Q(u)| &\leq 2q M^{-(s-d)} \|Q\|_\infty \sum_{\substack{\vec{n} \in \mathcal{N}_{2q} \\ n_{2q} = n}} \langle n_1^* \rangle^{s-d} \langle n_3^* \rangle^s |u_{n_1}| \cdots |u_{n_{2q-1}}| \\ &\leq 2q M^{-(s-d)} \|Q\|_\infty \sum_{\substack{\vec{n} \in \mathcal{N}_{2q} \\ n = (\ell, n_{2q})}} \langle \ell_1^* \rangle^{s-d} \langle \ell_2^* \rangle^s |u_{n_1}| \cdots |u_{n_{2q-1}}| \\ &\lesssim_q M^{-(s-d)} \|Q\|_\infty \sum_{\varphi \in \mathfrak{S}_{2q-1}} \sum_{\substack{\vec{n} \in \mathcal{N}_{2q} \\ n_1 = n}} \langle n_{\varphi_1} \rangle^{s-d} \langle n_{\varphi_2} \rangle^s |u_{n_2}| \cdots |u_{n_{2q}}|. \end{aligned}$$

Finally by applying the Young convolution inequality (see Lemma 2.8) we get that

$$\begin{aligned} \|\Pi_N \nabla Q(u)\|_{h^s} &\lesssim_q M^{-(s-d)} \|Q\|_\infty \|u\|_{\ell^1}^{2q-4} \|u\|_{h^s}^2 \|u\|_{\ell_{s-d}^1} \\ &\lesssim_q M^{-(s-d)} \|Q\|_\infty \|u\|_{\ell^1}^{2q-4} \|u\|_{h^s}^3. \end{aligned} \quad \square$$

LEMMA 2.23 (Mismatch vector field estimate: II). — Let  $\mathcal{C}_{\text{lo}} \cup \mathcal{C}_{\text{hi}} \subset \mathbb{Z}^d$  be a partition of  $\mathbb{Z}^d$  compatible with the Bourgain's partition given by Lemma 2.15, i.e.,

$$n \in \mathcal{C}_{\text{lo}} \implies [n] \subset \mathcal{C}_{\text{lo}}$$

such that  $\mathcal{C}_{v_0} \subset \mathcal{C}_{\text{lo}}$  and set

$$N = \min_{n \in \mathcal{C}_{\text{hi}}} |n|.$$

For all  $q \geq 2$ , all 1-resonant homogeneous polynomial  $Q \in \mathcal{H}_{2q}$  of degree  $2q$  satisfying  $Q \circ \Pi_{\mathcal{C}_{\text{lo}}} = 0$ , and all  $u \in h^s$  with  $s > d$ , we have

$$\|\Pi_{\mathcal{C}_{\text{lo}}} \nabla Q(u)\|_{h^s} \lesssim_{q,s} \|Q\|_{\infty} N^{-(\delta/2)(s-d)} \|u\|_{h^s}^3 \|u\|_{\ell^1}^{2q-4}.$$

*Proof.* — For clarity, in this proof, we highlight the dependencies with respect to  $\mathcal{L}$ . For  $n \in \mathcal{C}_{\text{lo}}$ , by symmetry, we have

$$\begin{aligned} \langle n \rangle^s |\partial_{\overline{u_n}} Q(u)| &\leq 2q \langle n \rangle^s \sum_{\substack{\vec{n} \in \mathcal{N}_{2q} \\ n_1 = n}} |Q_{\vec{n}}| |u_{n_2}| \cdots |u_{n_{2q}}| \\ &\lesssim_q \|Q\|_{\infty} \langle n \rangle^s \left( \sum_{\substack{\vec{n} \in \mathcal{N}_{2q} \\ n_1 = n \\ n_2 \in \mathcal{C}_{\text{hi}} \\ |\Omega_{\vec{n}}| \leq 1}} |u_{n_2}| \cdots |u_{n_{2q}}| + \sum_{\substack{\vec{n} \in \mathcal{N}_{2q} \\ n_1 = n \\ n_3 \in \mathcal{C}_{\text{hi}} \\ |\Omega_{\vec{n}}| \leq 1}} |u_{n_2}| \cdots |u_{n_{2q}}| \right) \\ &=: \|Q\|_{\infty} (\mathcal{E}_n^{(2)} + \mathcal{E}_n^{(3)}), \end{aligned}$$

and so

$$\|\Pi_{\mathcal{C}_{\text{lo}}} \nabla Q(u)\|_{h^s} \lesssim_q \|\Pi_{\mathcal{C}_{\text{lo}}} \mathcal{E}^{(2)}\|_{\ell^2} + \|\Pi_{\mathcal{C}_{\text{lo}}} \mathcal{E}^{(3)}\|_{\ell^2}.$$

*Case 1: Suppose that  $n_3 \in \mathcal{C}_{\text{hi}}$ .* — First, we have

$$1 \geq |\Omega_{\vec{n}}| \geq \sum_{j=1}^q \lambda_{n_{2j-1}}^2 - \sum_{j=1}^q \lambda_{n_{2j}}^2 \gtrsim_{\mathcal{L}} |n|^2 - q \max_{1 \leq j \leq q} \lambda_{n_{2j}}^2.$$

As a consequence, we have

$$\max_{1 \leq j \leq q} \langle n_{2j} \rangle^2 \gtrsim_{\mathcal{L}} \langle n \rangle^2$$

and so, we have

$$\mathcal{E}_n^{(3)} \lesssim_{q,\mathcal{L},s} \sum_{\substack{\vec{n} \in \mathcal{N}_{2q} \\ n_1 = n \\ |n_3| \geq N}} \langle n_2 \rangle^s |u_{n_2}| \cdots |u_{n_{2q}}| \lesssim_{q,\mathcal{L}} N^{-(s-d)} \sum_{\substack{\vec{n} \in \mathcal{N}_{2q} \\ n_1 = n}} \langle n_2 \rangle^s \langle n_3 \rangle^{s-d} |u_{n_2}| \cdots |u_{n_{2q}}|.$$

Thus applying the Young convolutional inequality, we get

$$\|\Pi_{\mathcal{C}_{\text{lo}}} \mathcal{E}^{(3)}\|_{\ell^2} \lesssim_{q,\mathcal{L},s} N^{-(s-d)} \|u\|_{h^s}^3 \|u\|_{\ell^1}^{2q-4}.$$

*Case 2: Suppose that  $n_2 \in \mathcal{C}_{\text{hi}}$ .* — Since  $[n_1] \neq [n_2]$ , thanks to the separation property of the clusters, we have

$$|n_1 - n_2| + |\lambda_{n_1}^2 - \lambda_{n_2}^2| > (|n_1| + |n_2|)^{\delta} \geq N^{\delta}.$$

As a consequence, either

$$(2q-2) \max_{j \geq 3} |n_j| \geq |n_1 - n_2| \geq \frac{N^{\delta}}{2},$$

or

$$(2q-2) \max_{j \geq 3} |n_j|^2 + 1 \gtrsim_{\mathcal{L}} |\lambda_{n_1}^2 - \lambda_{n_2}^2| \geq \frac{N^\delta}{2}.$$

In any case, we have

$$\max_{j \geq 3} \langle n_j \rangle \gtrsim_{\mathcal{L}, q} N^{\delta/2}.$$

Moreover, as previously, we have  $\max_{1 \leq j \leq q} \langle n_{2j} \rangle^2 \gtrsim_{\mathcal{L}} \langle n \rangle^2$ . Thus, since  $\delta \leq 1$ , it follows that<sup>(13)</sup>

$$\mathcal{E}_n^{(2)} \lesssim_{q, \mathcal{L}, s} \sum_{\substack{\vec{n}=(n, \ell) \in \mathcal{N}_{2q} \\ \langle \ell_2^* \rangle \gtrsim_{\mathcal{L}, q} N^{\delta/2}}} \langle \ell_1^* \rangle^s |u_{n_2}| \cdots |u_{n_{2q}}|$$

and reasoning as previously, we get that  $\|\Pi_{\mathcal{E}_{10}} \mathcal{E}^{(2)}\|_{\ell^2} \lesssim_{q, \mathcal{L}} N^{-(\delta/2)(s-d)} \|u\|_{h^s}^3 \|u\|_{\ell^1}^{2q-4}$ .  $\square$

**2.6. PROOF THAT THEOREM 2.7 IMPLIES THEOREM 1.3.** — We recall that the flat torus  $\mathbb{T}_{\mathcal{L}}^d$  is fixed and admissible in the sense of the Definition 1.1. It provides an exponent  $\tau_*$  (which can typically be chosen equal to  $d(d+1)/2 + 1$ ). We recall the associated constant  $c_* = 2\tau_* + 1$ . We will also need the Bourgain decomposition of Lemma 2.15 and the corresponding exponent  $\delta \equiv \delta(d)$ . We also recall that the nonlinearity  $f$  of (NLS) is fixed and satisfies the condition  $f'(0) \neq 0$ . Then we divide the proof into three main steps : *preparation*, *dynamics* and *measure estimates*.

*Step 1 : Preparation.* — Let  $r \geq 9$ ,  $\nu \leq \min((2c_*)^{-1}, 1/(2d))$  be the associated constant given by Theorem 2.7, and assume that

$$(2.15) \quad s_0 := d + 20\delta^{-1}\nu^{-1}r.$$

We also set  $\varepsilon \leq \varepsilon_0$  where  $\varepsilon_0$  is a constant depending only on  $r, s$  (and  $\mathcal{L}$  and  $f$ ) that will be chosen small enough. Let  $M > 0$  be a real number such that

$$(2.16) \quad \varepsilon^{-(8/\delta)r/(s-d)} \leq M \leq \varepsilon^{-\nu}$$

and define

$$N = (8r+1)^{-1}M.$$

Note that the existence of such a  $M$  is ensured by construction of  $s_0$ . Then we set a partition  $\mathcal{E}_{10} \cup \mathcal{E}_{hi} \subset \mathbb{Z}^d$  of  $\mathbb{Z}^d$  by

$$(2.17) \quad \mathcal{E}_{10} := \bigcup_{|n| \leq N} [n] \quad \text{and} \quad \mathcal{E}_{hi} := \mathbb{Z}^d \setminus \mathcal{E}_{10}.$$

Note that since the clusters are dyadic (see (2.9)), we have

$$(2.18) \quad \frac{1}{2}N \leq \min_{n \in \mathcal{E}_{hi}} |n| \quad \text{and} \quad \max_{n \in \mathcal{E}_{10}} |n| \leq 2N \leq M \leq \varepsilon^{-\nu}.$$

We recall that the Hamiltonian  $H$  of (NLS) (defined by (1.13)) admits the Taylor expansion  $H = H^{(\leq 4r)} + R^{(4r+2)}$  (see (2.1)). Then, setting

$$\kappa := \varepsilon^{c_*\nu},$$

<sup>(13)</sup>for details it is convenient to distinguish the case  $|n_2|^2 = \max_{1 \leq j \leq q} |n_{2j}|^2$ .

we apply the Birkhoff normal form theorem 2.13 to  $H^{(\leq 4r)} \circ \Pi_M$ . We get a polynomial  $\chi = \chi^{(4)} + \dots + \chi^{(4r)}$  supported on  $\ell^1(\mathbb{Z}_{\leq M}^d)$  (with  $\|\chi^{(2j)}\| \lesssim_r \kappa^{-j+1}$ ) and  $\varepsilon_* \gtrsim_{r,s} \sqrt{\kappa}$  (through Corollary 2.10) such that on  $\Pi_M^{-1} \Pi_M B_s(\varepsilon_*)$

$$(2.19) \quad \begin{aligned} (H^{(\leq 4r)} \circ \Pi_M) \circ \Phi_\chi^1 &= Z_2 \circ \Pi_M + Q^{(4)} + \dots + Q^{(4r)} + \Upsilon \\ &=: Z_2 \circ \Pi_M + Q^{(4)} + Q^{(\geq 6)} + \Upsilon, \end{aligned}$$

where  $Q^{(2j)} \in \mathcal{H}_{2j}(\mathbb{Z}_{\leq M}^d)$  is  $\kappa$  resonant and satisfies  $\|Q^{(2j)}\|_\infty \lesssim_r \kappa^{-j+2}$ , and  $\Upsilon$  is of order  $4r+2$ .

Then we set

$$(2.20) \quad H_{\text{lo}} := Z_2 \circ \Pi_M + Q^{(4)} + Q^{(\geq 6)} \circ \Pi_{\mathcal{C}_{\text{lo}}} - \frac{f'(0)}{2} \|\Pi_M \cdot\|_{\ell^2}^4.$$

As a consequence, if  $H_{\text{lo}}$  satisfies the assumptions of Theorem 2.7, then, provided that  $\varepsilon_0$  is small enough, we get an open set  $\Theta_\varepsilon^b \subset \Pi_M B_s(\varepsilon)$  and we just have to define our set of good initial data by

$$(2.21) \quad \Theta_\varepsilon := (\Pi_M B_s(\varepsilon)) \cap (\Phi_\chi^{-1})^{-1} \Theta_\varepsilon^b.$$

Now, we just have to check that  $H_{\text{lo}}$  satisfies the assumptions of Theorem 2.7. First, it is clear that the homogeneous terms  $P^{(2j)} := Q^{(2j)} \circ \Pi_{\mathcal{C}_{\text{lo}}}$ ,  $j \geq 3$ , of  $H_{\text{lo}}$  are  $\kappa = \varepsilon^{c_*\nu}$  resonant and satisfy  $\|P^{(2j)}\|_\infty \lesssim_r \kappa^{-j+2}$  and so  $\|P^{(2j)}\|_\infty \leq \varepsilon^{-j c_*\nu}$  provided that  $\varepsilon_0$  is small enough. So, the only thing we have to check is the form of the quartic terms of  $H_{\text{lo}}$ . To do this, it is enough to compute  $Q^{(4)}$ . We recall that by the Birkhoff normal form theorem 2.13, we have, for all  $\vec{n} \in \mathcal{N}_4$

$$Q_{\vec{n}}^{(4)} = \frac{f'(0)}{4} \mathbf{1}_{|\Omega_{\vec{n}}| \leq \kappa} \mathbf{1}_{\vec{n} \in (\mathbb{Z}_{\leq M}^d)^4}.$$

We will see that if  $\vec{n} \in (\mathbb{Z}_{\leq M}^d)^4$  satisfies  $|\Omega_{\vec{n}}| \leq \kappa$  then  $\{n_1, n_3\} = \{n_2, n_4\}$ . Note that it directly implies, by the Poincaré formula, that, as expected, for all  $u \in \mathbb{C}^{\mathbb{Z}_{\leq M}^d}$ ,

$$Q^{(4)}(u) = \frac{f'(0)}{2} \left( \sum_{|n| \leq M} |u_n|^2 \right)^2 - \frac{f'(0)}{4} \sum_{|n| \leq M} |u_n|^4.$$

So, let  $\vec{n} \in (\mathbb{Z}_{\leq M}^d)^4$  be such that  $|\Omega_{\vec{n}}| \leq \kappa$ . It means that

$$\begin{aligned} \kappa &\geq |\lambda_{n_1}^2 - \lambda_{n_2}^2 + \lambda_{n_3}^2 - \lambda_{n_4}^2| \\ &= |g(n_1, n_1) - g(n_2, n_2) + g(n_3, n_3) - g(n_4, n_4)| \\ &= |g(n_1, n_1) - g(n_2, n_2) + g(n_3, n_3) - g(n_1 + n_2 - n_3, n_1 + n_2 - n_3)| \\ &= 2|g(n_1 - n_2, n_3 - n_2)|. \end{aligned}$$

Now, assuming by contradiction that  $n_1 \neq n_2$  and  $n_3 \neq n_2$ , since the torus is admissible we get that

$$\varepsilon^{c_*\nu} = \kappa \gtrsim |n_1 - n_2|^{-\tau_*} |n_3 - n_2|^{-\tau_*} \gtrsim M^{-2\tau_*} \gtrsim \varepsilon^{2\tau_*\nu},$$

which is impossible, provided that  $\varepsilon_0$  is small enough, by definition of  $c_* = 2\tau_* + 1$ . As a consequence, we have proved that  $n_1 = n_2$  or  $n_3 = n_2$ , and so, using the zero momentum condition that  $\{n_1, n_3\} = \{n_2, n_4\}$ .

*Step 2 : Dynamics.* — We now prove stability of the low frequency actions and high-frequency super-actions over very long times  $T \leq \varepsilon^{-r}$ .

*Substep 2.1 : Setting of the bootstrap and low modes reduction.* — We consider a maximal solution  $u \in C^0([-T_-, T_+]; h^s)$  of (NLS) such that  $u(0) \in B_s(2\varepsilon) \cap \Pi_M^{-1}\Theta_\varepsilon$ . We denote

$$\rho := \|u(0)\|_{h^s} \leq 2\varepsilon.$$

Without loss of generality, we focus only on positive times. As usual, we proceed by bootstrap: we consider  $0 < T < T_+$  such that

$$T \leq \varepsilon^{-r} \quad \text{and} \quad \sup_{0 \leq t \leq T} \|u(t)\|_{h^s} \leq 2^{s+2}\rho,$$

and we aim at proving that

$$(2.22) \quad \sup_{0 \leq t \leq T} \|u(t)\|_{h^s} \leq 2^{s+1}\rho.$$

It will prove that (2.22) holds for  $T = \varepsilon^{-r}$ . First, since  $\mathcal{C}_{\text{hi}}$  contains only high modes, by applying Proposition 2.14, we know that, provided that  $\varepsilon_0$  is small enough, we have, for  $t \in [0, T]$

$$\|\Pi_{\mathcal{C}_{\text{hi}}} u(t)\|_{h^s}^2 \leq \|\Pi_{\geq N/2} u(t)\|_{h^s}^2 \leq 2^{2s} \|u(0)\|_{h^s}^2 + \rho^3 + T(N/2)^{-(\delta/2)(s-d)} \rho^3 + T\rho^{2r+1}.$$

Then, using the lower bound on  $M$  and the upper bound on  $T$ , we have

$$\|\Pi_{\mathcal{C}_{\text{hi}}} u(t)\|_{h^s}^2 \leq \rho^2(2^{2s} + \rho + C_{r,s}\rho\varepsilon^{-r}\varepsilon^{2r} + \varepsilon^{-r}\rho^{2r-1}),$$

where  $C_{r,s}$  is a constant depending only on  $r$  and  $s$  (and  $d$ ). It follows that, provided that  $\varepsilon_0$  is small enough, we have

$$(2.23) \quad \sup_{0 \leq t \leq T} \|\Pi_{\mathcal{C}_{\text{hi}}} u(t)\|_{h^s}^2 \leq 2^{2s+1}\rho^2.$$

To close the bootstrap, it remains to control  $\|\Pi_{\mathcal{C}_{\text{lo}}} u\|_{h^s}^2$ . To do it, we will prove that

$$(2.24) \quad \sup_{0 \leq t \leq T} \sum_{n \in \mathcal{C}_{\text{lo}}} \langle n \rangle^{2s} ||u_n(t)|^2 - |u_n(0)|^2| \leq \varepsilon\rho^2.$$

The conclusion (2.22) of the bootstrap is just then a consequence of the triangular inequality, the Pythagorean identity and (2.23). So, from now, we focus on proving (2.24).

*Substep 2.2 : Normal form and remainders.* — First, we note that since  $\nu \leq (2c_*)^{-1}$ , we have

$$(2.25) \quad \varepsilon_* \gtrsim_{r,s} \kappa = \varepsilon^{c_*\nu} \geq \sqrt{\varepsilon},$$

where  $\varepsilon_*$  was introduced in Corollary 2.10. It follows that, provided that  $\varepsilon_0$  is small enough, we have  $\varepsilon_* \gg \varepsilon$  and so (thanks to the bootstrap assumption) it makes sense to define

$$z := \Phi_\chi^{-1}(u).$$

As previously, without loss of generality (i.e., up to a standard approximation process), we assume that  $u(0) \in h^{s+2}$ , so that  $u \in C^1([0, T]; h^s)$  and  $z \in C^1([0, T]; h^s)$ . Since  $\Phi_\chi^{-1}$  is symplectic, using the chain rule, we get that

$$i\partial_t z = \nabla(H^{(\leq 4r)} \circ \Phi_\chi^1)(z) + \mathcal{E}^{(1)},$$

where

$$\mathcal{E}^{(1)} := i d\Phi_\chi^{-1}(u)(-i\nabla R^{(4r+2)}(u)).$$

Then, using that  $\chi$  is supported on  $\ell^1(\mathbb{Z}_{\leq M}^d)$ , we deduce that

$$i\partial_t \Pi_M z = \nabla(H^{(\leq 4r)} \circ \Pi_M \circ \Phi_\chi^1)(\Pi_M z) + \Pi_M \mathcal{E}^{(1)} + \Pi_M \mathcal{E}^{(2)},$$

where, setting  $H^{>M} := H^{(\leq 4r)} - H^{(\leq 4r)} \circ \Pi_M$ ,

$$\mathcal{E}^{(2)} := \Pi_M \nabla(H^{>M} \circ \Phi_\chi^1)(z).$$

By definition of  $Q^{(\geq 6)}$  and  $\Upsilon$  (see (2.19)), this evolution equation rewrites

$$i\partial_t \Pi_M z = \nabla(Z_2 + Q^{(4)} + Q^{(\geq 6)})(\Pi_M z) + \nabla \Upsilon(\Pi_M z) + \Pi_M \mathcal{E}^{(1)} + \Pi_M \mathcal{E}^{(2)}.$$

Then, projecting on  $\mathbb{C}^{\mathcal{C}_{\text{lo}}}$ , by definition of  $H_{\text{lo}}$  (see (2.20)), we get

$$i\partial_t \Pi_{\mathcal{C}_{\text{lo}}} z = \nabla H_{\text{lo}}(\Pi_{\mathcal{C}_{\text{lo}}} z) + 2f'(0) \|\Pi_M z\|_{\ell^2}^2 \Pi_{\mathcal{C}_{\text{lo}}} z + \mathcal{E}^{(\text{tot})},$$

where

$$(2.26) \quad \mathcal{E}^{(\text{tot})} := \Pi_{\mathcal{C}_{\text{lo}}} \nabla \Upsilon(\Pi_M z) + \Pi_{\mathcal{C}_{\text{lo}}} \mathcal{E}^{(1)} + \Pi_{\mathcal{C}_{\text{lo}}} \mathcal{E}^{(2)} + \Pi_{\mathcal{C}_{\text{lo}}} \mathcal{E}^{(3)}$$

and, setting  $H^{\text{high}} = Q^{(\geq 6)} - Q^{(\geq 6)} \circ \Pi_{\mathcal{C}_{\text{lo}}}$ ,

$$\mathcal{E}^{(3)} := \nabla H^{\text{high}}(\Pi_M z).$$

Observing that  $H_{\text{lo}}$  is invariant by gauge transform (because it commutes with the  $\ell^2$  norm) and setting

$$w = e^{i2f'(0) \int_0^t \|\Pi_M z(\tau)\|_{\ell^2}^2 d\tau} \Pi_{\mathcal{C}_{\text{lo}}} z,$$

we have

$$i\partial_t w = \nabla H_{\text{lo}}(w) + e^{i2f'(0) \int_0^t \|\Pi_M z(\tau)\|_{\ell^2}^2 d\tau} \mathcal{E}^{(\text{tot})}.$$

Finally, setting

$$(2.27) \quad v_n(t) = \begin{cases} w_n(t) & \text{if } n \in \mathcal{C}_{\text{lo}}, \\ 0 & \text{if } |n| > M, \\ e^{-it\lambda_n^2 + itf'(0)|z_n(0)|^2} z_n(0) & \text{else,} \end{cases}$$

we have  $v(0) = \Pi_M z(0)$  and as previously

$$(2.28) \quad i\partial_t v = \nabla H_{\text{lo}}(v) + e^{i2f'(0) \int_0^t \|\Pi_M z(\tau)\|_{\ell^2}^2 d\tau} \mathcal{E}^{(\text{tot})}.$$

*Substep 2.3: Control of the remainder terms.* — To apply Theorem 2.7 to  $v$ , we have to prove that  $\|\mathcal{E}^{(\text{tot})}(t)\|_{h^s} \leq \varepsilon^{3r} \rho$  for all  $t \in [0, T]$ . First, using the bound (2.6) on  $d\Phi_\chi^1$  and the bound (2.2) on  $\nabla R^{(4r+2)}$  (and the bootstrap assumption), we have

$$(2.29) \quad \|\mathcal{E}^{(1)}(t)\|_{h^s} \lesssim_{r,s} \rho^{4r+1}.$$

To estimate the other remainder terms, we have to estimate  $\|z(t)\|_{h^s}$ . To do it, we use that  $\Phi_\chi^{-1}$  is close to the identity, to get

$$(2.30) \quad \|z(t) - u(t)\|_{h^s} \lesssim_{r,s} \rho^3 \kappa^{-1}$$

and so, recalling that  $\kappa \geq \sqrt{\varepsilon}$  (provided that  $\varepsilon_0$  is small enough),

$$\|z(t)\|_{h^s} \leq 2\|u(t)\|_{h^s} \leq 2^{s+2} \rho.$$

Then, using the estimate (2.7) on  $\nabla \Upsilon$ , we have

$$(2.31) \quad \|\nabla \Upsilon(\Pi_M z(t))\|_{h^s} \lesssim_{r,s} \kappa^{-2r} \rho \varepsilon^{4r} \lesssim_{r,s} \rho \varepsilon^{3r}.$$

To control  $\mathcal{E}^{(2)}$  and  $\mathcal{E}^{(3)}$ , we will apply the mismatch lemmas. First, for  $\mathcal{E}^{(3)}$ , using the definition (2.19) of  $Q^{(\geq 6)}$ , we have

$$\Pi_{\mathcal{C}_{10}} \mathcal{E}^{(3)} = \sum_{j=3}^{2r} \Pi_{\mathcal{C}_{10}} \nabla(Q^{(2j)} - Q^{(2j)} \circ \Pi_{\mathcal{C}_{10}})(\Pi_M(z))$$

and so, since  $\mathcal{C}_{\text{hi}}$  contains only high modes (see (2.18)), by Lemma 2.23, we have

$$(2.32) \quad \begin{aligned} \|\Pi_{\mathcal{C}_{10}} \mathcal{E}^{(3)}(t)\|_{h^s} &\lesssim_{r,s} \sum_{j=3}^{2r} \|Q^{(2j)}\|_{\infty} (N/2)^{-(\delta/2)(s-d)} \rho^{2j-1} \\ &\lesssim_{r,s} \sum_{j=3}^{2r} \kappa^{-j+2} \varepsilon^{4r} \rho^{2j-1} \lesssim_{r,s} \varepsilon^{3r} \rho^5. \end{aligned}$$

The estimate on  $\mathcal{E}^{(2)}$  is more technical. First, we set

$$K^{(2j)} := \frac{f^{(j-1)}}{2j!} \|\cdot\|_{L^{2j}}^{2j} \in \mathcal{H}_{2j}(\mathbb{Z}^d)$$

and we note that the linear terms vanish in the definition of  $\mathcal{E}^{(2)}$ , i.e.,

$$\mathcal{E}^{(2)} = \Pi_M \nabla(K^{>M} \circ \Phi_\chi^1)(z),$$

where  $K^{>M} = K^{(4),>M} + \dots + K^{(4r),>M}$  and  $K^{(2j),>M} = K^{(2j)} - K^{(2j)} \circ \Pi_M$ . Then, doing the Taylor expansion of  $\Phi_\chi^t$  in  $t = 0$ , we get, as usual

$$K^{>M} \circ \Phi_\chi^1 = \sum_{q=0}^{2r} \frac{\text{ad}_\chi^q K^{>M}}{q!} + \int_0^1 \frac{(1-\tau)^{2r}}{(2r)!} \text{ad}_\chi^{2r+1} K^{>M} \circ \Phi_\chi^\tau d\tau =: \sum_{q=0}^{2r} \frac{\text{ad}_\chi^q K^{>M}}{q!} + \Gamma.$$

From the expansions of  $\chi$  and  $K^{>M}$ , we have

$$\text{ad}_\chi^q K^{>M} = \sum_{j=q+2}^{(q+1)(2r-1)+1} L^{q,j},$$

where, by Lemma 4.13,

$$L^{q,j} := \sum_{a_1 + \dots + a_{q+1} = j+q} \text{ad}_{\chi^{(2a_1)}} \dots \text{ad}_{\chi^{(2a_q)}} K^{(2a_{q+1}),>M} \in \mathcal{H}_{2j}(\mathbb{Z}^d).$$

Since  $\|\chi^{(2j)}\|_\infty \lesssim_r \kappa^{-j+1}$  and  $\|K^{(2j),>M}\|_\infty \lesssim_j 1 \lesssim_j \kappa^{-j+2}$ , we get by Lemma 4.13 that

$$\|L^{q,j}\|_\infty \lesssim_r \kappa^{-j+2},$$

and so, by Corollary 2.9,

$$\|\nabla \text{ad}_\chi^q L^{q,j}\|_{h^s} \lesssim_{r,s} \kappa^{-j+2} \cdot \|\cdot\|_{h^s}^{2j-1}.$$

As a consequence, if  $j \geq 2r + 1$ , we have

$$\|\nabla \text{ad}_\chi^q L^{q,j}(z(t))\|_{h^s} \lesssim_{r,s} \rho \varepsilon^{3r}.$$

Similarly, since  $j \geq q + 2$ , using that  $\Phi_\chi^\tau$  is close to the identity, we have that

$$\|\nabla \Gamma(z(t))\|_{h^s} \lesssim_{r,s} \rho \varepsilon^{3r}.$$

Recalling that  $\mathcal{C}_{\text{lo}}$  contains only low modes (see (2.18)), it follows that

$$\|\Pi_{\mathcal{C}_{\text{lo}}} \mathcal{E}^{(2)}(t)\|_{h^s} \lesssim_{r,s} \rho \varepsilon^{3r} + \sum_{q=0}^{2r-2} \sum_{j=q+2}^{2r} \|\Pi_{2N} \nabla L^{q,j}(z(t))\|_{h^s}.$$

Now, we note that since the polynomials  $\chi^{(2a)}$  are supported on  $\ell^1(\mathbb{Z}_{\leq M}^d)$  (i.e., they do not depend on  $(u_a)_{|a|>M}$ ) and the polynomials  $K^{(2a),>M}$  vanish on  $\ell^1(\mathbb{Z}_{\leq M}^d)$  (i.e., they do depend on  $(u_a)_{|a|>M}$ ) then the polynomials  $L^{q,j}$  vanish on  $\ell^1(\mathbb{Z}_{\leq M}^d)$ , i.e.,  $L^{q,j} \circ \Pi_M = 0$ . As a consequence, since  $M = (8r+1)N > 4rN$ , applying the mismatch lemma 2.22 and using the lower bound on  $M$  (see (2.16)), we have

$$(2.33) \quad \|\Pi_{\mathcal{C}_{\text{lo}}} \mathcal{E}^{(2)}(t)\|_{h^s} \lesssim_{r,s} \rho \varepsilon^{3r} + \sum_{j=2}^r M^{-(s-d)} \kappa^{-j+2} \rho^{2j-1} \lesssim_{r,s} \rho \varepsilon^{3r}.$$

Putting together the estimates (2.29), (2.33), (2.32) and (2.31) on the remainder terms forming  $\mathcal{E}^{(\text{tot})}$  (defined by (2.26)) and recalling that  $r \geq 9$ , we have proved that

$$\forall t \in [0, T], \quad \|\mathcal{E}^{(\text{tot})}(t)\|_{h^s} \leq \varepsilon^{3r} \rho.$$

*Substep 2.4: Application of Theorem 2.7 and conclusion.* — Since  $w$  solves the equation (2.28), we have proved that

$$\sup_{0 \leq t \leq T} \|i\partial_t v - \nabla H_{\text{lo}}(v)\|_{h^s} \leq \varepsilon^{3r} \rho.$$

Now, we aim at checking that  $v(0) \in \Theta_\varepsilon^b$ . We recall that, by assumption,

$$\Pi_M u(0) \in \Theta_\varepsilon = (\Pi_M B_s(\varepsilon)) \cap (\Phi_\chi^{-1})^{-1} \Theta_\varepsilon^b,$$

and so, since  $\chi$  is supported on  $\ell^1(\mathbb{Z}_{\leq M}^d)$ ,

$$v(0) = \Pi_M z(0) = \Pi_M \Phi_\chi^{-1} u(0) = \Phi_\chi^{-1} \Pi_M u(0) \in \Theta_\varepsilon^b.$$

Applying Theorem 2.7 and by definition of  $v$  (see (2.27)), we have that

$$\sup_{0 \leq t \leq T} \sum_{n \in \mathcal{C}_{\text{lo}}} \langle n \rangle^{2s} | |z_n(t)|^2 - |z_n(0)|^2 | = \sup_{0 \leq t \leq T} \sum_{|n| \leq M} \langle n \rangle^{2s} | |v_n(t)|^2 - |v_n(0)|^2 | \leq \frac{\varepsilon}{2} \rho^2.$$

Finally, by using that  $\|u(t) - z(t)\|_{h^s} \lesssim_{r,s} \rho^{5/2}$  (see (2.30)), we get, as expected (provided that  $\varepsilon_0$  is small enough),

$$(2.34) \quad \sup_{0 \leq t \leq T} \sum_{n \in \mathcal{C}_{10}} \langle n \rangle^{2s} ||u_n(t)|^2 - |u_n(0)|^2| \leq \varepsilon \rho^2.$$

*Step 3: Measure estimates.* — First, we note that, by time reversibility of  $\Phi_\chi^t$  (see the last property of Corollary 2.10), provided that  $\varepsilon$  is small enough, we have

$$\Phi_\chi^1 \Theta_\varepsilon^b \subset (\Phi_\chi^{-1})^{-1} \Theta_\varepsilon^b, \quad \text{and therefore} \quad \Phi_\chi^1 \Theta_\varepsilon^b \cap \Pi_M B_s(\varepsilon) \subset \Theta_\varepsilon.$$

Then, since  $\Phi_\chi^1$  is symplectic, it is volume preserving and so

$$\text{meas}(\Phi_\chi^1 \Theta_\varepsilon^b) = \text{meas}(\Theta_\varepsilon^b).$$

Now, since  $\Phi_\chi^1$  is close to the identity and  $\kappa \geq \varepsilon$ , we have (provided that  $\varepsilon$  is small enough)

$$\Phi_\chi^1 \Theta_\varepsilon^b \subset \Pi_M B_s(\varepsilon + \varepsilon^2).$$

As a consequence, we have

$$(2.35) \quad \text{meas}(\Theta_\varepsilon) \geq \text{meas}(\Theta_\varepsilon^b) - \text{meas} \mathcal{A}_\varepsilon,$$

where

$$\mathcal{A}_\varepsilon = \{u \in \Pi_M h^s \mid \varepsilon \leq \|u\|_{h^s} \leq \varepsilon + \varepsilon^2\}.$$

On the one hand, the measure estimate of Theorem 2.7, ensures that

$$(2.36) \quad \text{meas}(\Theta_\varepsilon^b) \geq (1 - \varepsilon^{1/39}) \text{meas}(\Pi_M B_s(\varepsilon)).$$

On the other hand, by homogeneity, we have

$$(2.37) \quad \frac{\text{meas} \mathcal{A}_\varepsilon}{\text{meas} \Pi_M B_s(\varepsilon)} = (1 + \varepsilon)^{2\#\mathbb{Z}_\leq^d M} - 1 \leq (1 + \varepsilon)^{(9M)^d} - 1 \leq e^{\varepsilon(9M)^d} - 1 \lesssim \sqrt{\varepsilon},$$

the last estimate coming from the upper bound  $M \leq \varepsilon^{-\nu}$  and assumption  $\nu \leq 1/(2d)$ .

Putting together the three last estimates (2.35), (2.36) and (2.37), we conclude that, as expected,

$$\text{meas}(\Theta_\varepsilon) \geq (1 - \varepsilon^{1/40}) \text{meas}(\Pi_M B_s(\varepsilon)).$$

This completes the proof that Theorem 2.7 implies Theorem 1.3.

### 3. RE-CENTERED POLYNOMIALS

The rest of the paper is devoted to the proof of Theorem 2.7.

**3.1. SET-UP.** — We start by defining some parameters within the framework of Theorem 2.7.

– In the rest of the paper we only need  $s > 0$  (the reason for this is that the reduced system is finite dimensional).

– We can choose (any)  $\nu > 0$  such that

$$(3.1) \quad 10^8 c_* d r \nu \leq 10^{-4}.$$

–  $\varepsilon_* > 0$  is a small parameter given by Corollary 2.10 which depends only on  $r, s$  and on the torus  $\mathbb{T}_{\mathcal{L}}^d$ .

– We fix  $\varepsilon \in (0, \varepsilon_*)$ . It measures the amplitude of the initial data.

- The frequency-truncation parameter  $M$  depends on  $\varepsilon$ :

$$2 \leq M \leq \varepsilon^{-\nu}.$$

- We set

$$\bar{r} = 10^5 r.$$

Then,  $s, \varepsilon, \nu$  and  $M$  are defined once and for all.

REMARK 3.1. — In the Section 2, to prove the main Theorem 1.3 from Theorem 2.7 and in order to control the contribution of high frequencies, we required that

$$s \gtrsim r\nu^{-1} \gtrsim r^2,$$

where the first (resp. second) inequality is a consequence of (2.15) (resp. (3.1)). The implicit constants depend on  $d$  and on the geometry of the torus  $\mathbb{T}_{\mathcal{L}}^d$ .

3.2. MODULATION PARAMETERS AND RE-CENTERED ACTIONS. — In the subsequent analysis, the amplitudes of Fourier coefficients of a function  $\phi$  in (say)  $\Pi_M B_s(20\varepsilon)$ , serve as internal modulation parameters for the frequencies. We set

$$\xi_n(\phi) = |\phi_n|^2, \quad n \in \mathbb{Z}_M^d.$$

We define

$$(3.2) \quad \mathcal{U}_s(\varepsilon) = \{\xi \in \mathbb{R}^{\mathbb{Z}_M^d} \mid \sum_{|n| \leq M} \langle n \rangle^{2s} |\xi_n| < 400\varepsilon^2\},$$

an open set of  $\mathbb{R}^{\mathbb{Z}_M^d}$  endowed with the natural topology of  $\ell_1^{2s}(\mathbb{Z}_M^d; \mathbb{R})$ , made of *internal parameters*. In particular,

$$\phi \in \Pi_M B_s(20\varepsilon) \iff \xi(\phi) \in \mathcal{U}_s(\varepsilon).$$

To modulate the frequencies, we center the actions around the parameters  $\xi = (\xi_n)_{|n| \leq M} \in \mathcal{U}_s(\varepsilon)$  in the Hamiltonian, and define

$$y_n(\xi, u) = |u_n|^2 - \xi_n, \quad n \in \mathbb{Z}_M^d.$$

Then, the Hamiltonian's coefficients depend on the parameters  $\xi$ , and we shall introduce another Hamiltonian formalism more convenient for the analysis.

3.3. PARAMETER DEPENDENT POLYNOMIALS. — In Definition 2.2, we denoted  $\mathcal{H}_{2q}(\mathbb{Z}_M^d)$  a class of real homogeneous polynomials of degree  $2q$ , supported on  $\ell^1(\mathbb{Z}_M^d)$ . We set

$$\mathcal{H}_{\leq 2r}(\mathbb{Z}_M^d) := \bigoplus_{0 \leq q \leq r} \mathcal{H}_{2q}(\mathbb{Z}_M^d), \quad \mathcal{H}(\mathbb{Z}_M^d) := \bigoplus_{0 \leq q} \mathcal{H}_{2q}(\mathbb{Z}_M^d),$$

where  $\mathcal{H}_0(\mathbb{Z}_M^d)$  is the class of constant real functions.

3.3.1. *Extended class of real polynomials.* — We now motivate the subsequent definitions and notations. In the expansions of polynomials in  $\mathcal{H}(\mathbb{Z}_M^d)$  we would like to keep track of the actions and formally see them as new variables:

$$u_n, \quad \overline{u_n}, \quad |u_n|^2.$$

To a multi-index  $\vec{n}$  in the class  $\mathcal{N}$  defined in Definition 2.1 corresponds a sequence  $\mathbf{n} \in (\mathbb{N}^{\mathbb{Z}_M^d})^3$  where

$$\mathbf{n} = (\mathbf{k}, \mathbf{\ell}, \mathbf{m}), \quad \text{with} \quad \mathbf{k} = (k_n)_{|n| \leq M}, \quad \mathbf{\ell} = (\ell_n)_{|n| \leq M}, \quad \mathbf{m} = (m_n)_{|n| \leq M}.$$

For  $n \in \mathbb{Z}_M^d$ , the integer  $m_n$  (resp.  $k_n, \ell_n$ ) is the multiplicity of  $|u_n|^2$  (resp.  $u_n, \overline{u_n}$ ). Such a decomposition only make sense, and is unique, under a non-pairing condition between  $u_n$  and  $\overline{u_n}$ .

DEFINITION 3.2. — We define the class of multi-index with zero momentum condition

$$(3.3) \quad \mathfrak{N} := \{\mathbf{n} = (\mathbf{k}, \ell, \mathbf{m}) \mid \forall n \in \mathbb{Z}_M^d, k_n \ell_n = 0 \text{ and } \sum_{n \in \mathbb{Z}_M^d} k_n - \ell_n = \sum_{n \in \mathbb{Z}_M^d} n(k_n - \ell_n) = 0\}.$$

The degree of a multi-index  $\mathbf{n}$  is its total degree:

$$\deg(\mathbf{n}) := \sum_{n \in \mathbb{Z}_M^d} 2m_n + k_n + \ell_n.$$

REMARK 3.3. — In the above definition,  $k_n \ell_n = 0$  is the non-pairing condition between  $u_n$  and  $\overline{u_n}$ ,  $\sum_{|n| \leq M} k_n - \ell_n = 0$  ensures that  $u$  and  $\overline{u}$  have the same multiplicity and  $\sum_{|n| \leq M} n(k_n - \ell_n) = 0$  is the zero momentum condition.

For further use we also denote by  $\tilde{\mathfrak{N}}$  the set of multi-indices satisfying only the non-pairing condition

$$(3.4) \quad \tilde{\mathfrak{N}} = \{\mathbf{n} = (\mathbf{k}, \ell, \mathbf{m}) \mid k_n \ell_n = 0 \forall n \in \mathbb{Z}_M^d\}.$$

If  $q \geq 0$ ,

$$(3.5) \quad \mathfrak{N}_{2q} = \{\mathbf{n} \in \mathfrak{N} \mid \deg(\mathbf{n}) = 2q\}, \quad \mathfrak{N}_{\leq 2q} = \{\mathbf{n} \in \mathfrak{N} \mid \deg(\mathbf{n}) \leq 2q\},$$

with similar definitions for  $\tilde{\mathfrak{N}}_{2q}$  and  $\tilde{\mathfrak{N}}_{\leq 2q}$ .

NOTATION 3.4. — We denote by  $\mathbf{n}_-$  the size of the smallest frequency that is not completely paired:

$$(3.6) \quad \mathbf{n}_- = \min\{|n| \mid n \in \mathbb{Z}_M^d, k_n + \ell_n \geq 1\}.$$

The class of integrable monomials is denoted by

$$(3.7) \quad \mathfrak{I} = \{\mathbf{n} = (\mathbf{k}, \ell, \mathbf{m}) \in \mathfrak{N} \mid \mathbf{k} = \ell \equiv 0\}.$$

For convenience, given  $n \in \mathbb{Z}_M^d$ , we introduce special multi-indices (that are not in  $\mathfrak{N}$ , but in  $\tilde{\mathfrak{N}}$ ):

$$\mathbf{e}_m(n) := (\delta(n, \cdot), 0, 0), \quad \mathbf{e}_t(n) := (0, \delta(n, \cdot), 0), \quad \mathbf{e}_l(n) := (0, 0, \delta(n, \cdot)),$$

where  $\delta$  is the Kronecker symbol. They correspond to the following monomials:

$$z_{\mathbf{e}_m(n)}(u, y) = y_n, \quad z_{\mathbf{e}_t(n)}(u, y) = u_n, \quad z_{\mathbf{e}_l(n)}(u, y) := \overline{u_n}.$$

3.3.2. *The class of polynomials.* — We now define the corresponding class of extended polynomials  $\mathcal{H}^\sharp(\mathbb{Z}_M^d)$ .

LEMMA 3.5. — *For every  $H \in \mathcal{H}_{\leq 2q}(\mathbb{Z}_M^d)$  there exists a unique sequence of coefficients  $(H_n) \in \mathbb{C}^{\mathfrak{N}_{\leq 2q}}$ , such that*

$$H(u) = \sum_{\mathbf{n} \in \mathfrak{N}} H_{\mathbf{n}} \prod_{n \in \mathbb{Z}_M^d} u_n^{k_n} \overline{u_n}^{\ell_n} |u_n|^{2m_n}.$$

Given a multi-index  $\mathbf{n} = (\mathbf{k}, \ell, \mathbf{m}) \in \mathfrak{N}_{\leq 2q}$  we use the short-hand notation:

$$u^{\mathbf{k}} := \prod_{n \in \mathbb{Z}_M^d} u_n^{k_n}, \quad \overline{u}^{\ell} := \prod_{n \in \mathbb{Z}_M^d} \overline{u_n}^{\ell_n}, \quad I(u)^{\mathbf{m}} := \prod_{n \in \mathbb{Z}_M^d} |u_n|^{2m_n},$$

so that, alternatively,

$$H(u) = \sum_{\mathbf{n} \in \mathfrak{N}} H_{\mathbf{n}} u^{\mathbf{k}} \overline{u}^{\ell} I(u)^{\mathbf{m}}.$$

We can now define the new formalism.

DEFINITION 3.6 (Extended class of real polynomials). — Let  $q \geq 0$ . The set  $\mathcal{H}_{\leq 2q}^\sharp(\mathbb{Z}_M^d)$  of real polynomials is the set of functions defined on  $\mathbb{C}^{\mathbb{Z}_M^d} \times \mathbb{R}^{\mathbb{Z}_M^d}$

$$Q(u, y) = \sum_{\mathbf{n} \in \mathfrak{N}_{\leq 2q}} Q_{\mathbf{n}} z_{\mathbf{n}}(u, y),$$

where the coefficients  $(Q_{\mathbf{n}}) \in \mathbb{C}^{\mathfrak{N}_{\leq 2q}}$  satisfy the reality condition:

$$Q_{\mathbf{k}, \ell, \mathbf{m}} = \overline{Q_{\ell, \mathbf{k}, \mathbf{m}}},$$

and

$$z_{\mathbf{n}}(u, y) := u^{\mathbf{k}} \overline{u}^{\ell} y^{\mathbf{m}} = \prod_{n \in \mathbb{Z}_M^d} u_n^{k_n} \overline{u_n}^{\ell_n} y_n^{m_n}.$$

LEMMA 3.7. — *By associating to  $H \in \mathcal{H}_{\leq 2q}(\mathbb{Z}_M^d)$  its coefficients  $(H_{\mathbf{n}})_{\mathbf{n} \in \mathfrak{N}_{\leq 2q}}$  the following map is an isomorphism*

$$\begin{aligned} A: \mathcal{H}_{\leq 2q}(\mathbb{Z}_M^d) &\longrightarrow \mathcal{H}_{\leq 2q}^\sharp(\mathbb{Z}_M^d) \\ H &\longmapsto \sum_{\mathbf{n} \in \mathfrak{N}} H_{\mathbf{n}} u^{\mathbf{k}} \overline{u}^{\ell} y^{\mathbf{m}}, \end{aligned}$$

the inverse of  $A$  being

$$\begin{aligned} A^{-1}: \mathcal{H}_{\leq 2q}^\sharp(\mathbb{Z}_M^d) &\longrightarrow \mathcal{H}_{\leq 2q}(\mathbb{Z}_M^d) \\ Q(u, y) &\longmapsto Q(u, I(u)), \end{aligned}$$

where  $I(u) = (|u_n|^2)_{n \in \mathbb{Z}_M^d}$ .

In particular, we have

$$H(u) = A(H)(u, I(u)) \quad \forall H \in \mathcal{H}_{\leq 2q}(\mathbb{Z}_M^d).$$

3.3.3. *Centering procedure.* — With this formalism, we can interpret the centering of a polynomial  $H$  with respect to a sequence of parameters  $\xi$ , as a translation of  $\xi$  in the  $y$ -variable of  $A(H)$ , in the space  $\mathcal{H}^\sharp(\mathbb{Z}_M^d)$ :

DEFINITION 3.8. — Let  $\xi \in \mathcal{U}_s(\varepsilon)$ . We define the translation operator

$$T_\xi : \mathcal{H}_{\leq 2q}^\sharp(\mathbb{Z}_M^d) \longrightarrow \mathcal{H}_{\leq 2q}^\sharp(\mathbb{Z}_M^d)$$

such that for all  $H \in \mathcal{H}_{\leq 2q}^\sharp(\mathbb{Z}_M^d)$ ,

$$(3.8) \quad T_\xi Q(u, y) := Q(u, y - \xi).$$

With these operators at hand, for  $H \in \mathcal{H}_{\leq 2q}^\sharp(\mathbb{Z}_M^d)$ , centering  $H$  around  $\xi$  boils down to write<sup>(14)</sup>

$$(3.9) \quad H(u) = A(H)(u, I(u)) = (A^{-1} \circ T_\xi \circ T_{-\xi} \circ A)(H)(u) = (T_{-\xi} \circ A)(H)(u, I(u) - \xi).$$

We stress out that the coefficients of  $(T_{-\xi} \circ A)(H)$  depend polynomially on  $\xi$ . Nevertheless, this dependency is quite artificial: the polynomial function  $H$  does not really depend on  $\xi$ . However, our normal form procedure will naturally generate Hamiltonian functions really depending on  $\xi$  (due to the small divisors). For this reason, we introduce a class of parameter-dependent real polynomials.

DEFINITION 3.9 (Class of parameter-dependent real polynomials). — We define

$$(3.10) \quad X_{2q}(\varepsilon) := C^\infty(\mathcal{U}_s(\varepsilon), \mathcal{H}_{2q}(\mathbb{Z}_M^d)),$$

and

$$X_{\leq 2q}(\varepsilon) := C^\infty(\mathcal{U}_s(\varepsilon), \mathcal{H}_{\leq 2q}(\mathbb{Z}_M^d)), \quad X(\varepsilon) := C^\infty(\mathcal{U}_s(\varepsilon), \mathcal{H}(\mathbb{Z}_M^d)).$$

In addition, we define the class of integrable polynomials of degree less than or equal to  $2q$  by

$$X_{2q, \text{Int}}(\varepsilon) := \{H \in X_{2q}(\varepsilon) \mid \forall \mathbf{n} \in \mathfrak{N}_{\leq 2q}, \forall \xi \in \mathcal{U}_s(\varepsilon), \quad (H_{\mathbf{n}}(\xi) \neq 0 \implies \mathbf{n} \in \mathfrak{J})\}.$$

We can now introduce the main notation.

NOTATION 3.10. — Centering  $H \in X_{\leq 2q}(\varepsilon)$  around  $\xi \in \mathcal{U}_s(\varepsilon)$  gives a function of  $\xi$  valued in  $\mathcal{H}_{\leq 2q}^\sharp(\mathbb{Z}_M^d)$ , denoted  $H[\cdot]$  and defined by

$$\begin{aligned} H[\cdot] : \mathcal{U}_s(\varepsilon) &\longrightarrow \mathcal{H}_{\leq 2q}^\sharp(\mathbb{Z}_M^d) \\ \xi &\longmapsto H[\xi] := [T_{-\xi} \circ A](H(\xi)). \end{aligned}$$

According to (3.9), a parameter-dependent polynomial  $H \in X_{\leq 2q}(\varepsilon)$ ,

$$H : \xi \in \mathcal{U}_s(\varepsilon) \longmapsto H(\xi; \cdot) \in \mathcal{H}_{\leq 2q}(\mathbb{Z}_M^d),$$

once re-centered, can be represented by

$$(3.11) \quad H(\xi; u) = H[\xi](u, I(u) - \xi).$$

<sup>(14)</sup>In the following formula we are just writing  $|u_n|^2 = |u_n|^2 - \xi_n + \xi_n$  and then expanding around  $|u_n|^2 - \xi_n$  in the expansion of  $H$  given by Lemma 3.5.

In the expanded form, we have

$$H(\xi; u) = \sum_{\mathbf{n} \in \mathfrak{N}_{\leq 2q}} (H[\xi])_{\mathbf{n}} z_{\mathbf{n}}(u, I(u) - \xi),$$

where we recall that  $z_{\mathbf{n}}$  is defined by

$$z_{\mathbf{n}}(u, y) = u^{\mathbf{k}} \bar{u}^{\mathbf{\ell}} y^{\mathbf{m}} = \prod_{n \in \mathbb{Z}_M^d} u_n^{k_n} \bar{u}_n^{-\ell_n} y_n^{m_n}.$$

### 3.3.4. Operations for re-centered polynomials

LEMMA 3.11 (Poisson bracket with an action). — *If  $H \in X_{\leq 2q}(\varepsilon)$ , then, for all  $n \in \mathbb{Z}^d$  and  $\xi \in \mathcal{U}_s(\varepsilon)$ ,*

$$(3.12) \quad \{I_n, H(\xi; \cdot)\}(u) = 2i \sum_{\mathbf{n} \in \mathfrak{N}} (k_n - \ell_n) (H[\xi])_{\mathbf{n}} u^{\mathbf{k}} \bar{u}^{\mathbf{\ell}} (I(u) - \xi)^{\mathbf{m}}.$$

By linearity, we deduce the following.

COROLLARY 3.12. — *Given a sequence of real numbers  $(\omega_n)_{n \in \mathbb{Z}_M^d}$  (the frequencies), if*

$$Z_2(\xi; u) := \sum_{n \in \mathbb{Z}_M^d} \omega_n (|u_n|^2 - \xi_n),$$

*and  $H \in X_{\leq 2q}(\varepsilon)$ , then we have*

$$\{Z_2, H\}(u) = 2i \sum_{\mathbf{n} \in \mathfrak{N}} \Omega_{\mathbf{n}}(\omega) (H[\xi])_{\mathbf{n}} u^{\mathbf{k}} \bar{u}^{\mathbf{\ell}} (I(u) - \xi)^{\mathbf{m}},$$

*where  $\Omega_{\mathbf{n}}(\omega)$  is the resonance function:*

$$(3.13) \quad \Omega_{\mathbf{n}}(\omega) := \sum_{n \in \mathbb{Z}_M^d} (k_n - \ell_n) \omega_n.$$

## 4. FUNCTIONAL SETTING FOR RE-CENTERED POLYNOMIALS

4.1. PARAMETERS. — We have recalled the set-up of Theorem 2.7 in Section 3.1. We now introduce the new parameters that play a role in the subsequent finite-dimensional analysis, to control the frequencies of the solution smaller than  $M$ . Recall that the small parameter  $\varepsilon \in (0, \varepsilon_*)$  was fixed in paragraph (3.1).

–  $\eta > 0$  is a large portion of  $\varepsilon$

$$\eta = \varepsilon^{1-1/100}.$$

– For  $\alpha \in \mathbb{N}$ , the frequency scale  $N_\alpha = \varepsilon^{-\alpha/200s}$  (which is not necessarily an integer), satisfies

$$N_0 = 1, \quad (N_{\alpha+1}/N_\alpha)^s = \varepsilon^{-1/200}.$$

– We set  $\beta = 100r$ . In this way,  $N_\beta$  (which will be our largest scale) satisfies

$$(4.1) \quad N_\beta^{-2s} = \varepsilon^{\beta/100} = \varepsilon^r.$$

–  $\tau$  (proportional to  $s$ ) is a small parameter compared to  $s$ :

$$\tau = \frac{s}{10\beta} = \frac{s}{10^3 r}.$$

We ensure that for all  $\alpha \in \{0, \dots, \beta\}$ ,

$$(4.2) \quad (\eta^{-1}\varepsilon)N_\alpha^{5\tau} = \varepsilon^{1/100-5\alpha\tau/200s} \leq \varepsilon^{1/100-5\beta\tau/200s} = \varepsilon^{3/500} \leq \varepsilon^{1/200}.$$

–  $\bar{r}$  (proportional to  $r$ ) is large enough to ensure that

$$(4.3) \quad \bar{r}\tau \geq 6s.$$

In view of the definition of  $\tau$ , we can set

$$\bar{r} = 10^5 r.$$

– Recalling the definition (3.5) of the set of multi-indices  $\mathfrak{N}_{\leq 2q}$  observe that, at least when  $q \leq \bar{r}^2$ ,

$$(4.4) \quad \#\mathfrak{N}_{\leq 2q} \leq \#\tilde{\mathfrak{N}}_{\leq 2q} \leq \#(\mathbb{Z}_M^d)^{2q} \ll \varepsilon^{-1/10^6}.$$

–  $\gamma$  (and  $\gamma(\alpha)$ ) are numbers in  $(0, 1)$ , depending on  $\varepsilon$ , which give the large measure of the non-resonant parameter set, but they also appear in the small divisor estimates. The following choice is acceptable (although other choices are also possible):

$$(4.5) \quad \gamma = \varepsilon^{1/30}, \quad \gamma(\alpha) := 4^\alpha \gamma.$$

We have the following key relationships between the parameters (if  $\varepsilon_*$  is small enough):

$$(4.6) \quad \gamma(\alpha)^{-1}(\varepsilon^{-1}\eta)^2(N_{\alpha+1}/N_\alpha)^{2s} \leq \varepsilon^{-1/15}.$$

#### 4.2. FREQUENCY-SCALES

DEFINITION 4.1 (Weights for the coefficient bounds). — For  $\alpha \in \{0, \dots, \beta\}$  and  $n \in \mathbb{Z}_M^d$  set

$$D(\alpha) := \eta^{-2-1/5} N_\alpha^{2s},$$

$$C_n(\alpha) := \eta^{-1} \min(|n|, N_\alpha)^s N_\alpha^\tau.$$

Given a multi-index  $\mathbf{n} = (\mathbf{k}, \ell, \mathbf{m}) \in \mathfrak{N}$  defined in (3.3) we set

$$\mathbf{w}_{\mathbf{n}}^0(\alpha) := N_\alpha^{-6s} \eta^6 \prod_{n \in \mathbb{Z}_M^d} D(\alpha)^{m_n} C_n(\alpha)^{k_n + \ell_n},$$

and

$$\mathbf{w}_{\mathbf{n}}^1(\alpha) := N_\alpha^{-4s} \eta^4 \prod_{n \in \mathbb{Z}_M^d} D(\alpha)^{m_n} C_n(\alpha)^{k_n + \ell_n}.$$

REMARK 4.2. — Note that our choice of weights  $D(\alpha)$  differs slightly from the choice of Bourgain [Bou00a], where (adapting the notation) we would have the following:

$$D(\alpha) = \eta^{-2-1/10} N_\alpha^{2s} \min(|n|, N_\alpha)^s.$$

The above choice has some advantages when moving from the scale  $\alpha$  to the scale  $\alpha+1$  for the integrable quartic terms generated in the algorithm. In our case, however, we measure these quartic terms in the norm  $Z$  (defined below) that does not depend

on the scales. This allows us to have smaller weights. We emphasize that such a saving is crucial to obtain the bounds on the symplectic transformations (e.g. Lemma 6.16) to measure the size of the non-resonant initial data set (a task not addressed in [Bou00a]).

**4.3. NORMS AND VECTOR FIELD ESTIMATE.** — We now define suitable norms for re-centered parameter-dependent polynomials, which are written as in (3.11). Recall that we defined the set of parameters  $\mathcal{U}_s(\varepsilon)$  in (3.2).

**DEFINITION 4.3** (Norms for the re-centered polynomials). — Given  $\alpha \in \{0, \dots, \beta\}$  and a parameter-dependent polynomial  $H \in X(\varepsilon)$  (defined in (3.10)), we define two norms at scale  $\alpha$ :

$$\begin{aligned} \|H\|_{Y_\alpha^{\text{sup}}} &:= \sup_{\xi \in \mathcal{U}_s(\varepsilon)} \max_{\mathbf{n} \in \mathfrak{N}} w_{\mathbf{n}}^0(\alpha)^{-1} |(H[\xi])_{\mathbf{n}}|, \\ \|H\|_{Y_\alpha^{\text{lip}}} &:= \sup_{\xi \in \mathcal{U}_s(\varepsilon)} \max_{\mathbf{n} \in \mathfrak{N}} \max_{|n| \leq M} w_{\mathbf{n}}^1(\alpha)^{-1} |\partial_{\xi_n}(H[\xi])_{\mathbf{n}}| \end{aligned}$$

and

$$\begin{aligned} \|H\|_{Z^{\text{sup}}} &:= \sup_{\xi \in \mathcal{U}_s(\varepsilon)} \max_{\mathbf{n} \in \mathfrak{N}} |(H[\xi])_{\mathbf{n}}|, \\ \|H\|_{Z^{\text{lip}}} &:= \sup_{\xi \in \mathcal{U}_s(\varepsilon)} \max_{\mathbf{n} \in \mathfrak{N}} \max_{|n| \leq M} |\partial_{\xi_n}(H[\xi])_{\mathbf{n}}|. \end{aligned}$$

The norm  $Y_\alpha^{\text{sup}}$  is nothing but a weighted  $\ell^\infty$ -norm on the coefficients of the re-centered polynomial  $H[\xi]$  (as introduced in Notation 3.10), with scale-dependent weights as in Definition 4.1. We will extensively use that for all  $\mathbf{n} \in \mathfrak{N}$  and  $\xi \in \mathcal{U}_s(\varepsilon)$ ,

$$|(H[\xi])_{\mathbf{n}}| \leq \|H\|_{Y_\alpha^{\text{sup}}} w_{\mathbf{n}}^0(\alpha),$$

and, for all  $|n| \leq M$ ,

$$|\partial_{\xi_n}(H[\xi])_{\mathbf{n}}| \leq \|H\|_{Y_\alpha^{\text{lip}}} w_{\mathbf{n}}^1(\alpha).$$

We say that a Hamiltonian *operates at frequency scale*  $\alpha$  (see Theorem 6.2) when its  $Y_\alpha^{\text{sup}}$ -norm is small compared to  $\varepsilon^{-1/10^3}$ , say.

**REMARK 4.4.** — For all  $H$  of order greater than 6 (in the sense that for all  $\xi \in \mathcal{U}_s(\varepsilon)$ ,  $H_{\mathbf{n}}[\xi] \neq 0$  only if  $\deg(\mathbf{n}) \geq 6$ ), we have that

$$(4.7) \quad \|H\|_{Y_0^{\text{sup}}} \leq \|H\|_{Z^{\text{sup}}}, \quad \|H\|_{Y_0^{\text{lip}}} \leq \|H\|_{Z^{\text{lip}}}.$$

The reason is that  $N_0 = 1$ .

Given  $\xi$  and  $\alpha$  we introduce the annulus

$$(4.8) \quad \mathcal{V}_{\alpha,s}(\varepsilon, \xi) := \{u \in \Pi_M B_s(\varepsilon) \mid \sum_{n \in \mathbb{Z}_M^d} \langle n \rangle^{2s} |u_n|^2 - \xi_n \leq \varepsilon^{2+1/5} N_\alpha^{-2s}\}.$$

We now estimate a multilinear quantity for functions in the above neighborhood. The set  $\tilde{\mathfrak{N}}$  was defined in (3.4).

LEMMA 4.5 (Weighted multilinear estimate). — For all  $\alpha \in \{0, \dots, \beta\}$ ,  $\xi \in \mathcal{U}_s(\varepsilon)$ ,  $\mathbf{n} \in \tilde{\mathfrak{N}}$  and  $u \in \mathcal{V}_{\alpha,s}(20\varepsilon, \xi)$ ,

$$(4.9) \quad w_{\mathbf{n}}^0(\alpha) |z_{\mathbf{n}}(u, I(u) - \xi)| \leq N_{\alpha}^{-6s} \eta^6.$$

The above lemma can be understood as follows. When  $u \in \mathcal{V}_{\alpha,s}(20\varepsilon, \xi)$ , which is an additional smallness condition on the centered actions  $(|u_n|^2 - \xi_n)_{n \in \mathbb{Z}_M^d}$ , we gain a factor  $N_{\alpha}^{-6s} \eta^6$  in the multilinear estimate (4.9). We emphasize that the zero momentum condition is not required, and we assume that  $\mathbf{n}$  is in the set  $\tilde{\mathfrak{N}}$  defined in (3.4).

*Proof.* — Fix  $u \in \mathcal{V}_{\alpha,s}(20\varepsilon, \xi)$  and  $\mathbf{n} \in \tilde{\mathfrak{N}}$ . According to the definition 4.1 of the weights,

$$(4.10) \quad w_{\mathbf{n}}^0(\alpha) z_{\mathbf{n}}(u, I(u) - \xi) = N_{\alpha}^{-6s} \eta^6 \prod_{n \in \mathbb{Z}_M^d} C_n(\alpha)^{k_n + \ell_n} u_n^{k_n} \bar{u}_n^{\ell_n} D(\alpha)^{m_n} (|u_n|^2 - \xi_n)^{m_n}.$$

For all  $n \in \mathbb{Z}_M^d$  and  $\alpha \leq \beta$ ,

$$C_n(\alpha) |u_n| \leq \langle n \rangle^s N_{\alpha}^{\tau} \eta^{-1} |u_n| \leq \|\langle n \rangle^s u_n\|_{\ell^\infty} N_{\beta}^{\tau} \eta^{-1} \leq 20\varepsilon \eta^{-1} N_{\beta}^{\tau} \leq 1,$$

under Condition (4.2). In addition,

$$D(\alpha) ||u_n|^2 - \xi_n| \leq \eta^{-2-1/5} N_{\alpha}^{2s} ||u_n|^2 - \xi_n\|_{\ell^\infty} \lesssim (\varepsilon \eta^{-1})^{2+1/5} \leq 1.$$

Multiplying over the contributions where  $m_n + k_n + \ell_n \geq 1$ , for  $n \in \mathbb{Z}_M^d$ , gives the multilinear estimate (4.9).  $\square$

We can now state the key vector field estimate, which motivates the definition of the norm  $Y_{\alpha}^{\text{sup}}$  and the normal form Theorem 6.2.

PROPOSITION 4.6 (Vector field estimate at scale  $\alpha$ ). — Given  $\alpha \in \{0, \dots, \beta\}$  and  $Q \in X_{\leq \tau^2}(\varepsilon)$ , we have that for all  $\xi \in \mathcal{U}_s(\varepsilon)$  and  $u \in \mathcal{V}_{\alpha,s}(20\varepsilon, \xi)$ ,

$$(4.11) \quad \|\nabla Q(\xi; u)\|_{h^s} \leq \|Q\|_{Y_{\alpha}^{\text{sup}}} N_{\alpha}^{-4s} \varepsilon^{4-1/4} \|u\|_{h^s}.$$

*Proof.* — We fix  $\xi \in \mathcal{U}_s(\varepsilon)$  and, without loss of generality we suppose that  $\|Q\|_{Y_{\alpha}^{\text{sup}}} = 1$ . For  $n \in \mathbb{Z}_M^d$ , we have

$$(\nabla Q(\xi; u))_n = 2 \sum_{\mathbf{n} \in \mathfrak{N}_{\leq cr}} (Q[\xi])_{\mathbf{n}} \partial_{\bar{u}_n} z_{\mathbf{n}}(u, I(u) - \xi).$$

Observe that

$$\partial_{\bar{u}_n} z_{\mathbf{n}}(u, I(u) - \xi) = \ell_n \frac{z_{\mathbf{n}}(u, I(u) - \xi)}{\bar{u}_n} + m_n \frac{u_n z_{\mathbf{n}}(u, I(u) - \xi)}{|u_n|^2 - \xi_n}.$$

Hence,

$$\begin{aligned}
& \|\nabla Q(u)\|_{h^s} \\
&= 2 \left( \sum_{n \in \mathbb{Z}_M^d} \langle n \rangle^{2s} \left| \sum_{\mathbf{n} \in \mathfrak{N}_{\leq 2q}} (Q[\xi])_{\mathbf{n}} \left( \ell_n \frac{z_{\mathbf{n}}(u, I(u) - \xi)}{\bar{u}_n} + m_n \frac{u_n z_{\mathbf{n}}(u, I(u) - \xi)}{|u_n|^2 - \xi_n} \right) \right|^2 \right)^{1/2} \\
&\leq \#(\mathbb{Z}_M^d) \#(\mathfrak{N}_{\leq 2q}) \max_{\substack{n \in \mathbb{Z}_M^d \\ \mathbf{n} \in \mathfrak{N}_{\leq 2q}}} w_{\mathbf{n}}^0(\alpha) \langle n \rangle^s \left( \left| \ell_n \frac{z_{\mathbf{n}}(u, I(u) - \xi)}{\bar{u}_n} \right| + \left| m_n \frac{u_n z_{\mathbf{n}}(u, I(u) - \xi)}{|u_n|^2 - \xi_n} \right| \right) \\
&\leq \varepsilon^{-1/10^4} \max_{n \in \mathbb{Z}_M^d} \max_{\mathbf{n} \in \mathfrak{N}_{\leq 2q}} w_{\mathbf{n}}^0(\alpha) \langle n \rangle^s \left( \left| \ell_n \frac{z_{\mathbf{n}}(u, I(u) - \xi)}{\bar{u}_n} \right| + \left| m_n \frac{u_n z_{\mathbf{n}}(u, I(u) - \xi)}{|u_n|^2 - \xi_n} \right| \right),
\end{aligned}$$

where we used the counting estimate (4.4). Then, for fixed  $n \in \mathbb{Z}_M^d$  and  $\mathbf{n} \in \mathfrak{N}_{\leq 2q}$ , suppose that  $\ell_n \geq 1$ . We have from the zero momentum condition that there exists  $j \in \mathbb{Z}_M^d \setminus \{n\}$  with

$$\max(k_j, \ell_j) \geq 1, \quad |j| \gtrsim \frac{|n|}{r^2}.$$

We only consider the situation when  $k_j \geq 1$  (the situation when  $\ell_j \geq 1$  is analogous), in which case we have

$$\begin{aligned}
w_{\mathbf{n}}^0(\alpha) \langle n \rangle^s \ell_n \left| \frac{z_{\mathbf{n}}(u, I(u) - \xi)}{\bar{u}_n} \right| &\lesssim_{r,s} C_n(\alpha) C_j(\alpha) \langle j \rangle^s |u_j| \left| \frac{w_{\mathbf{n}}^0(\alpha)}{C_n(\alpha) C_j(\alpha)} \right| \left| \frac{z_{\mathbf{n}}(u, I(u) - \xi)}{u_j \bar{u}_n} \right| \\
&\lesssim_{r,s} C_n(\alpha) C_j(\alpha) \|u\|_{h^s} w_{\mathbf{n}'}^0(\alpha) |z_{\mathbf{n}'}(u, I(u) - \xi)|,
\end{aligned}$$

where

$$\mathbf{n}' = \mathbf{n} - \mathbf{e}_{\mathbf{k}}(j) - \mathbf{e}_{\mathbf{l}}(n) \in \tilde{\mathfrak{N}}.$$

Note that

$$w_{\mathbf{n}'}^0(\alpha) = \frac{w_{\mathbf{n}}^0(\alpha)}{C_n(\alpha) C_j(\alpha)}.$$

It follows from Lemma 4.5 that

$$w_{\mathbf{n}'}^0(\alpha) |z_{\mathbf{n}'}(u, I(u) - \xi)| \leq N_{\alpha}^{-6s} \eta,$$

and therefore

$$\begin{aligned}
w_{\mathbf{n}}^0(\alpha) \langle n \rangle^s \ell_n \left| \frac{z_{\mathbf{n}}(u, I(u) - \xi)}{\bar{u}_n} \right| &\lesssim_{r,s} C_n(\alpha) C_j(\alpha) N_{\alpha}^{-6s} \eta^6 \|u\|_{h^s} \\
&\lesssim_{r,s} N_{\alpha}^{2\tau} \eta^4 N_{\alpha}^{-4s} \|u\|_{h^s} \\
&\lesssim_{r,s} \eta^{4 - \frac{1}{10}} N_{\alpha}^{-4s} \|u\|_{h^s},
\end{aligned}$$

according to the constraint (4.2) on the parameters. Similarly, if  $m_n \geq 1$  then

$$\begin{aligned}
w_{\mathbf{n}}^0(\alpha) \langle n \rangle^s m_n \left| \frac{u_n z_{\mathbf{n}}(u, I(u) - \xi)}{|u_n|^2 - \xi_n} \right| &\lesssim_{r,s} D(\alpha) \langle n \rangle^s |u_n| \left| \frac{w_{\mathbf{n}}^0(\alpha)}{D(\alpha)} \right| \left| \frac{z_{\mathbf{n}}(u, I(u) - \xi)}{|u_n|^2 - \xi_n} \right| \\
&\lesssim_{r,s} D(\alpha) \|u\|_{h^s} w_{\mathbf{n}'}^0(\alpha) |z_{\mathbf{n}'}(u, I(u) - \xi)|,
\end{aligned}$$

where

$$\mathbf{n}' = \mathbf{n} - \mathbf{e}_{\mathbf{m}}(n) \in \mathfrak{N}, \quad w_{\mathbf{n}'}^0(\alpha) = \frac{w_{\mathbf{n}}^0(\alpha)}{D(\alpha)},$$

and we deduce from Lemma 4.5 that

$$w_{\mathbf{n}}^0(\alpha) \langle n \rangle^s m_n \left| \frac{u_n z_{\mathbf{n}}(u, I(u) - \xi)}{|u_n|^2 - \xi_n} \right| \lesssim_{r,s} \eta^{4-1/5} N_{\alpha}^{-4s} \|u\|_{h^s}.$$

The desired inequality (4.11) follows when  $\varepsilon_*(s, r)$  is small enough.  $\square$

In the next lemma, we provide a fundamental example of a polynomial that operates at scale  $\alpha$ . Recall that  $\mathbf{n}_- = \min\{|n| \mid n \in \mathbb{Z}^d, k_n + \ell_n \geq 1\}$  and  $\mathfrak{I}$  denotes the set of integrable monomials, as defined in (3.7).

NOTATION 4.7. — Given  $\alpha \geq 1$  we define the set of multi-indices with an unpaired frequency smaller than  $N_{\alpha}$ :

$$(4.12) \quad \Lambda_{\alpha} := \{\mathbf{n} \in \mathfrak{N} \setminus \mathfrak{I} \mid \mathbf{n}_- < N_{\alpha}\},$$

Then, for  $H \in X(\varepsilon)$  and  $\xi \in \mathcal{U}_s(\varepsilon)$ , we define  $\Pi_{\Lambda_{\alpha}} H(\xi) \in \mathcal{H}(\mathbb{Z}_M^d)$  by

$$(4.13) \quad \Pi_{\Lambda_{\alpha}} H(\xi; u) := \sum_{\mathbf{n} \in \Lambda_{\alpha}} (H[\xi])_{\mathbf{n}} z_{\mathbf{n}}(u, I(u) - \xi).$$

Moreover, for  $q \geq 1$ ,

$$\begin{aligned} \Pi_{\deg=2q} Q(\xi; u) &:= \sum_{\mathbf{n} \in \mathfrak{N} : \deg \mathbf{n} = 2q} (Q[\xi])_{\mathbf{n}} z_{\mathbf{n}}(u, I(u) - \xi), \\ \Pi_{\deg \leq 2q} Q(\xi; u) &:= \sum_{p=1}^q \Pi_{\deg=2p} Q(\xi). \end{aligned}$$

LEMMA 4.8. — For all  $\alpha \geq 0$ , if  $\mathbf{n} \in \mathfrak{N} \setminus \Lambda_{\alpha+1}$  with  $\deg(\mathbf{n}) \geq 6$ , we have

$$w_{\mathbf{n}}^0(\alpha) \leq w_{\mathbf{n}}^0(\alpha+1), \quad w_{\mathbf{n}}^1(\alpha) \leq w_{\mathbf{n}}^1(\alpha+1).$$

REMARK 4.9. — As a consequence, if  $Q \in X_{\leq \bar{r}^2}(\varepsilon)$  satisfies, for all  $\xi \in \mathcal{U}_s(\varepsilon)$

$$\Pi_{\Lambda_{\alpha}} Q(\xi) = 0, \quad \Pi_{\deg \leq 4} Q(\xi) = 0,$$

(as defined in (4.13)), then  $Q$  operates at scale  $\alpha$  in the sense that

$$\|Q\|_{Y_{\alpha}^{\sup}} \leq \|Q\|_{Y_{\alpha-1}^{\sup}} \leq \dots \leq \|Q\|_{Y_0^{\sup}}.$$

We deduce from Proposition 4.6 the vector field estimate: for all  $\xi \in \mathcal{U}_s(\varepsilon)$  and  $u \in \mathcal{V}_{\alpha,s}(20\varepsilon, \xi)$ ,

$$\|\nabla Q(\xi; u)\|_{h^s} \leq \|Q\|_{Y_0^{\sup}} N_{\alpha}^{-4s} \varepsilon^3 \|u\|_{h^s}.$$

*Proof.* — Recall from Definition 4.1 that the weight  $w_{\mathbf{n}}^0(\alpha)$  is made of two parts: a prefactor  $N_{\alpha}^{-6s} \eta^6$ , which is a gain, and the product with the coefficients  $C_n(\alpha)$  and  $D(\alpha)$ .

When comparing the prefactor from scale  $\alpha$  to scale  $\alpha+1$  we have a loss:

$$(N_{\alpha}/N_{\alpha+1})^{-6s}.$$

On the other hand, increasing the weights  $C_n(\alpha)$  to  $C_n(\alpha+1)$  gives a saving when  $|n| \geq N_{\alpha+1}$ :

$$C_n(\alpha) = \eta^{-1} N_{\alpha}^{s+\tau} = (N_{\alpha}/N_{\alpha+1})^{s+\tau} C_n(\alpha+1).$$

Note also that increasing the weight  $D(\alpha)$  to  $D(\alpha + 1)$  always gives a saving:

$$D(\alpha) = (N_\alpha/N_{\alpha+1})^{2s} D(\alpha + 1).$$

We complete the proof of Lemma 4.8 from the assumption that  $\deg(\mathbf{n}) \geq 6$ :

$$\begin{aligned} w_{\mathbf{n}}^0(\alpha) &= (N_\alpha/N_{\alpha+1})^{-6s} \prod_{|n| \leq M} (N_\alpha/N_{\alpha+1})^{2m_n s + (k_n + \ell_n)(s+\tau)} w_{\mathbf{n}}^0(\alpha + 1) \\ &\leq (N_\alpha/N_{\alpha+1})^{(\deg(\mathbf{n})-6)s} w_{\mathbf{n}}^0(\alpha + 1). \end{aligned}$$

We proceed similarly to prove that  $w_{\mathbf{n}}^1(\alpha) \leq w_{\mathbf{n}}^1(\alpha + 1)$ .  $\square$

To conclude this subsection, we state a lemma in the same spirit. It is helpful when dealing with monomials of large degree.

LEMMA 4.10 (Weights for monomials with large degree). — *For all  $\mathbf{n} \in \mathfrak{N}$ , if*

$$(4.14) \quad \tau \deg(\mathbf{n}) \geq 6s,$$

*then for all  $\alpha \geq 0$ ,*

$$w_{\mathbf{n}}^0(\alpha) \leq w_{\mathbf{n}}^0(\alpha + 1).$$

REMARK 4.11. — As a consequence of Proposition 4.6, if a polynomial  $R \in X_{\leq \tau^2}(\varepsilon)$  has order greater than  $6s\tau^{-1}$  then for all  $u \in \mathcal{V}_{\beta,s}(20\varepsilon, \xi)$ ,

$$\|\nabla R(u)\|_{h^s} \lesssim_r \|R\|_{Z^{\sup}} N_\beta^{-4s} \varepsilon^3 \|u\|_{h^s}.$$

*Proof.* — We proceed as in the proof of Lemma 4.8. In this case we leverage on the the factor  $N_\alpha^\tau$  in  $C_n(\alpha)$  to get the saving: for all  $n \in \mathbb{Z}_M^d$ ,

$$C_n(\alpha) = \eta^{-1} \min(|n|, N_\alpha)^s N_\alpha^\tau \leq (N_\alpha/N_{\alpha+1})^\tau C_n(\alpha + 1).$$

This yields

$$\begin{aligned} w_{\mathbf{n}}^0(\alpha) &= (N_{\alpha+1}/N_\alpha)^{6s-2s \deg(\mathbf{m}) - \tau(\deg(\mathbf{k}) + \deg(\mathbf{\ell}))} w_{\mathbf{n}}^0(\alpha + 1) \\ &\leq (N_{\alpha+1}/N_\alpha)^{6s-\tau \deg(\mathbf{n})} w_{\mathbf{n}}^0(\alpha + 1), \end{aligned}$$

where we used that  $N_{\alpha+1} \geq N_\alpha$  and  $s \geq \tau$ . Assuming (4.14) concludes.  $\square$

4.4. POISSON BRACKET AND POINT-WISE ESTIMATE ON THE COEFFICIENTS. — We will prove in Proposition 4.15 (resp. Proposition 4.16) fundamental bounds on the  $Y_\alpha^{\sup}$ -norm ( $Y_\alpha^{\text{lip}}$ -norm) of the Poisson bracket of two polynomials. As a warm-up we first state some explicit computations for the Poisson bracket of two re-centered polynomials.

LEMMA 4.12 (Poisson bracket between two monomials). — *For fixed multi-indices  $\mathbf{n} = (\mathbf{k}, \mathbf{\ell}, \mathbf{m})$  and  $\mathbf{n}' = (\mathbf{k}', \mathbf{\ell}', \mathbf{m}')$  one has*

$$\begin{aligned} (4.15) \quad &\{z_{\mathbf{n}}(u, I(u) - \xi), z_{\mathbf{n}'}(u, I(u) - \xi)\} \\ &= 2i \sum_{n \in \mathbb{Z}_M^d} (k'_n \ell_n - k_n \ell'_n) \frac{z_{\mathbf{n}}(u, I(u) - \xi) z_{\mathbf{n}'}(u, I(u) - \xi)}{|u_n|^2} \\ &\quad + (m_n(k'_n - \ell'_n) + m'_n(\ell_n - k_n)) \frac{z_{\mathbf{n}}(u, I(u) - \xi) z_{\mathbf{n}'}(u, I(u) - \xi)}{|u_n|^2 - \xi_n}. \end{aligned}$$

In particular, we observe from (4.15) that a Poisson bracket either kills a pair  $u_n, \overline{u_n}$  when  $k'_n \ell_n - k_n \ell'_n \neq 0$ , and a term  $|u_n|^2 - \xi_n$  when  $m_n(k'_n - \ell'_n) + m'_n(\ell_n - k_n) \neq 0$ .

*Proof.* — We compute explicitly the Poisson bracket using (1.20):

$$\begin{aligned} & \{u^{\mathbf{k}} \overline{u}^{\ell}(I(u) - \xi)^{\mathbf{m}}, u^{\mathbf{k}'} \overline{u}^{\ell'}(I(u) - \xi)^{\mathbf{m}'}\} \\ &= 2i \sum_{n \in \mathbb{Z}_M^d} \left( \prod_{j \in \mathbb{Z}_M^d \setminus \{n\}} u_j^{k_j + k'_j} \overline{u}_j^{\ell_j + \ell'_j} (|u_j|^2 - \xi_j)^{m_j + m'_j} \right) \\ & \quad \times \left[ (k'_n \ell_n - k_n \ell'_n) \overline{u}_n^{\ell_n + \ell'_n - 1} u_n^{k_n + k'_n - 1} (|u_n|^2 - \xi_n)^{m_n + m'_n} \right. \\ & \quad \left. + (m_n(k'_n - \ell'_n) + m'_n(\ell_n - k_n)) u_n^{k_n + k'_n} \overline{u}_n^{\ell_n + \ell'_n} (|u_n|^2 - \xi_n)^{m_n + m'_n - 1} \right]. \end{aligned}$$

Organizing the terms gives the identity (4.15).  $\square$

The next lemma follows from (4.15), and gives a useful expression of the coefficients of the re-centered polynomial generated by the Poisson bracket of two re-centered polynomials.

LEMMA 4.13. — *If  $Q, H \in X(\varepsilon)$ , with  $\min(\deg(Q), \deg(H)) =: 2q$ , then  $\{Q, H\} \in X(\varepsilon)$  and we have for all  $\mathbf{n}'' \in \mathfrak{N}$  and  $\xi \in \mathcal{U}_s(\varepsilon)$ ,*

$$\begin{aligned} (4.16) \quad & (\{Q, H\}[\xi])_{\mathbf{n}''} = 2i \prod_{|j| \leq M} \sum_{0 \leq b_j \leq a_j \leq q} \binom{a_j}{b_j} \xi_j^{a_j - b_j} \sum_{n \in \mathbb{Z}_M^d} \sum_{(\mathbf{n}, \mathbf{n}') \in \mathfrak{N}^2} (Q[\xi])_{\mathbf{n}} (H[\xi])_{\mathbf{n}'} \\ & \times \left( \mathbf{1}_{E_{\mathbf{n}'', \mathbf{a}, \mathbf{b}}^{(1)}(n)}(\mathbf{n}, \mathbf{n}') (k'_n \ell_n - k_n \ell'_n) + \mathbf{1}_{E_{\mathbf{n}'', \mathbf{a}, \mathbf{b}}^{(2)}(n)}(\mathbf{n}, \mathbf{n}') (m_n(k'_n - \ell'_n) + m'_n(\ell_n - k_n)) \right), \end{aligned}$$

where, given  $\mathbf{n}'' = (m''_j, k''_j, \ell''_j)_{|j| \leq M}$ ,  $\mathbf{a} = (a_j)_{|j| \leq M}$ ,  $\mathbf{b} = (b_j)_{|j| \leq M}$  with  $0 \leq b_j \leq a_j \leq q$ , and  $n \in \mathbb{Z}_M^d$ , the sets of multi-indices

$$E_{\mathbf{n}'', \mathbf{a}, \mathbf{b}}^{(1)}(n), \quad E_{\mathbf{n}'', \mathbf{a}, \mathbf{b}}^{(2)}(n) \subset \mathfrak{N}^2$$

are defined as follows:

(1)  $(\mathbf{n}, \mathbf{n}') \in E_{\mathbf{n}'', \mathbf{a}, \mathbf{b}}^{(1)}(n)$  if and only if  $k'_n \ell_n - k_n \ell'_n \neq 0$ , with

$$a_n = \min(k_n + k'_n - 1, \ell_n + \ell'_n - 1),$$

and

$$m''_n = m_n + m'_n + b_n,$$

$$k''_n = k_n + k'_n - 1 - a_n,$$

$$\ell''_n = \ell_n + \ell'_n - 1 - a_n.$$

For all  $j \in \mathbb{Z}_M^d \setminus \{n\}$ ,

$$a_j = \min(k_j + k'_j, \ell_j + \ell'_j),$$

and

$$m''_j = m_j + m'_j + b_j,$$

$$k''_j = k_j + k'_j - a_j,$$

$$\ell''_j = \ell_j + \ell'_j - a_j,$$

(2)  $(\mathbf{n}, \mathbf{n}') \in E_{\mathbf{n}'', \mathbf{a}, \mathbf{b}}^{(2)}(n)$  if and only if  $m_n(k'_n - \ell'_n) + m'_n(\ell_n - k_n) \neq 0$ , with

$$a_n = \min(k_n + k'_n, \ell_n + \ell'_n),$$

and

$$m''_n = m_n + m'_n - 1 + b_n,$$

$$k''_n = k_n + k'_n - a_n,$$

$$\ell''_n = \ell_n + \ell'_n - a_n.$$

For all  $j \in \mathbb{Z}_M^d \setminus \{n\}$ , we impose the same conditions as for the above definition of  $E_{\mathbf{n}'', \mathbf{a}, \mathbf{b}}^{(1)}(n)$ .

Let us briefly comment on formula (4.16).

– For fixed  $n \in \mathbb{Z}_M^d$  we collect in  $E_{\mathbf{n}'', \mathbf{a}, \mathbf{b}}^{(1)}(n)$  (resp.  $E_{\mathbf{n}'', \mathbf{a}, \mathbf{b}}^{(2)}(n)$ ) the contributions of the Poisson bracket between two monomials  $(\mathbf{n}, \mathbf{n}')$  where a pair  $u_n, \overline{u_n}$  goes away (resp. a term  $y_n$  goes away).

– The parameters  $\mathbf{a} = (a_j)_{|j| \leq M}$  and  $\mathbf{b} = (b_j)_{|j| \leq M}$  come from the centering procedure:  $a_j$  represents the multiplicity of the pairing  $|u_j|^2$ , which is less than  $\min(\deg(Q), \deg(H))$ , while  $b_j$  determines the number of terms  $y_j$ . More precisely, when we center  $|u_j|^{2a_j} = (y_j + \xi_j)^{a_j}$  we obtain

$$|u_j|^{2a_j} = \sum_{b_j=0}^{a_j} \binom{a_j}{b_j} \xi_j^{a_j-b_j} y_j^{b_j}.$$

– Observe that the non-pairing condition is preserved:

$$k_j \ell_j = 0 \text{ and } k_{j'} \ell_{j'} = 0 \ \forall j \implies k''_j \ell''_j = 0 \ \forall j.$$

LEMMA 4.14. — Given  $\mathbf{n}'' \in \mathfrak{N}_{\leq 2q}$ , we let

$$\mathfrak{C}_{\mathbf{n}''}(q) := \{(\mathbf{n}, \mathbf{n}') \in \mathfrak{N}_{\leq 2q}^2 \mid \{z_{\mathbf{n}}, z_{\mathbf{n}'}\} \longrightarrow z_{\mathbf{n}''}\}$$

be the set of multi-indices whose Poisson bracket contributes to  $z_{\mathbf{n}''}$ . We have

$$(4.17) \quad \#(\mathfrak{C}_{\mathbf{n}''}(8\bar{r}^2)) \leq \varepsilon^{-1/10^4}.$$

To prove this lemma we count the total number of terms generated by the Poisson bracket between two monomials. The obtained estimate (4.17) is very crude, but it is sufficient for our needs.

*Proof.* — Suppose that  $(\mathbf{n}, \mathbf{n}') \in \mathfrak{N}_{\leq 2q}$ . According to (4.15) we see that  $\{z_{\mathbf{n}}, z_{\mathbf{n}'}\}$  generates  $2\#(\mathbb{Z}_M^d)$  terms. Each of them gives at most  $\#(\mathbb{Z}_M^d)q^2$  new terms, up to some universal constant  $C$ . Summing over all the possible pairs  $(\mathbf{n}, \mathbf{n}')$  concludes:

$$\#(\mathfrak{C}_{\mathbf{n}''}) \leq C\#(\mathfrak{N}_{\leq 2q})^2\#(\mathbb{Z}_M^d)^2q^2 \leq C\#(\mathbb{Z}_M^d)^{4q+2}q^2.$$

The inequality (4.17) then follows from our choices of parameters.  $\square$

The fundamental bilinear estimate is presented in the following proposition. Given two Hamiltonian polynomials  $Q$  and  $H$ , this estimate controls the  $Y_{\alpha}^{\text{sup}}$ -norm of the Poisson bracket between  $H$  and  $Q$  by the product of the  $Y_{\alpha}^{\text{sup}}$ -norms of  $Q$  and  $H$ .

PROPOSITION 4.15. — For all  $Q, H \in X_{\leq \bar{r}^2}(\mathbb{Z}_M^d)$  and for all  $\xi \in \mathcal{U}_s(\varepsilon)$ ,

$$(4.18) \quad \|\{Q, H\}\|_{Y_\alpha^{\text{sup}}} \leq \eta^{4-1/4} N_\alpha^{-4s} \|Q\|_{Y_\alpha^{\text{sup}}} \|H\|_{Y_\alpha^{\text{sup}}}.$$

Moreover, for  $Q \in X_{\leq \bar{r}^2}(\varepsilon)$  and for all integrable  $Z \in X_{4, \text{Int}}(\varepsilon)$ , we have

$$(4.19) \quad \|\{Q, Z\}\|_{Y_\alpha^{\text{sup}}} \leq \eta^{2+1/5-2/10^4} N_\alpha^{-2s} \|Q\|_{Y_\alpha^{\text{sup}}} \|Z\|_{Z^{\text{sup}}}.$$

The saving in  $\eta$  and in  $N_\alpha^{-s}$  is remarkable and it will compensate for the loss from the small divisor. In the second case (4.19), this saving is a consequence of the smallness of the centered actions  $|u_n|^2 - \xi_n$ , which is encoded in the weight  $D(\alpha)$ .

*Proof.* — Let us first prove (4.18). We denote

$$P(u) = \{Q, H\} = \sum_{\mathbf{n}'' \in \mathfrak{N}_{\leq 2\bar{r}^2}} (P[\xi])_{\mathbf{n}''} z_{\mathbf{n}''}(u, I(u) - \xi).$$

By homogeneity we can assume that

$$\|Q\|_{Y_\alpha^{\text{sup}}} = \|H\|_{Y_\alpha^{\text{sup}}} = 1.$$

In particular, for all  $(\mathbf{n}, \mathbf{n}') \in \mathfrak{N}$  we have

$$|(Q[\xi])_{\mathbf{n}}(H[\xi])_{\mathbf{n}'}| \leq w_{\mathbf{n}}^0(\alpha) w_{\mathbf{n}'}^0(\alpha),$$

and our goal is to prove that for all  $\mathbf{n}'' \in \mathfrak{N}$  and  $\xi$ ,

$$|(P[\xi])_{\mathbf{n}''}| \leq \eta^{4-1/4} N_\alpha^{-4s} w_{\mathbf{n}''}^0(\alpha).$$

*Reduction to fixed monomials  $(\mathbf{n}, \mathbf{n}')$ .* — Formula (4.16) (from Lemma 4.13) expresses  $(P[\xi])_{\mathbf{n}''}$  in terms of the coefficients  $(Q[\xi])_{\mathbf{n}}$  and  $(H[\xi])_{\mathbf{n}'}$  of the re-centered polynomials associated with  $Q$  and  $H$ . We deduce from it that for all  $\xi \in \mathcal{U}_s(\varepsilon)$

$$\begin{aligned} |(P[\xi])_{\mathbf{n}'}| &\lesssim_r \#(\mathfrak{C}_{\mathbf{n}''}(c\bar{r}^2)) \sup_{\mathbf{a}, \mathbf{b}, \mathbf{n}, \mathbf{n}'} \left( \prod_{|j| \leq M} |\xi_j|^{a_j - b_j} \right) w_{\mathbf{n}}^0(\alpha) w_{\mathbf{n}'}^0(\alpha) \\ &\left[ \mathbf{1}_{E_{\mathbf{n}''}^{(1)}(\mathbf{n}, \mathbf{a}, \mathbf{b})}(\mathbf{n}, \mathbf{n}') |k'_n \ell_n - k_n \ell'_n| + \mathbf{1}_{E_{\mathbf{n}''}^{(1)}(\mathbf{n}, \mathbf{a}, \mathbf{b})}(\mathbf{n}, \mathbf{n}') |m_n(k'_n - \ell_{n'}) + m_{n'}(\ell_n - k_n)| \right]. \end{aligned}$$

According to the (crude) counting estimate (4.17), we obtain

$$(4.20) \quad |(P[\xi])_{\mathbf{n}'}| \lesssim_r \varepsilon^{-1/10^4} \sup_{\mathbf{a}, \mathbf{b}, \mathbf{n}, \mathbf{n}'} \left( \prod_{|j| \leq M} |\xi_j|^{a_j - b_j} \right) w_{\mathbf{n}}^0(\alpha) w_{\mathbf{n}'}^0(\alpha) \left[ \mathbf{1}_{E_{\mathbf{n}''}^{(1)}(\mathbf{n}, \mathbf{a}, \mathbf{b})}(\mathbf{n}, \mathbf{n}') |k'_n \ell_n - k_n \ell'_n| + \mathbf{1}_{E_{\mathbf{n}''}^{(1)}(\mathbf{n}, \mathbf{a}, \mathbf{b})}(\mathbf{n}, \mathbf{n}') |m_n(k'_n - \ell_{n'}) + m_{n'}(\ell_n - k_n)| \right].$$

Let us fix  $n \in \mathbb{Z}_M^d$ ,  $\mathbf{a} = (a_j)_{|j| \leq M}$ ,  $\mathbf{b} = (b_j)_{|j| \leq M}$  with  $0 \leq b_j \leq a_j \leq \bar{r}$ . Let  $(\mathbf{n}, \mathbf{n}')$  be a pair of indices for which the Poisson bracket  $\{z_{\mathbf{n}}, z_{\mathbf{n}'}\}$  of the associated monomials contribute to  $z_{\mathbf{n}''}$ . We organize the terms in two cases.

*Case 1: A pair  $(u_n, \overline{u_n})$  goes away.* — This corresponds to the case when  $(\mathbf{n}, \mathbf{n}') \in E_{\mathbf{n}'', \mathbf{a}, \mathbf{b}}^{(1)}(n)$ . In such a case, we have

$$(4.21) \quad w_{\mathbf{n}}^0(\alpha) w_{\mathbf{n}'}^0(\alpha) = N_{\alpha}^{-6s} \eta^6 C_n(\alpha)^2 \left( \prod_{|j| \leq M} C_j(\alpha)^{2(a_j - b_j)} \left( \frac{C_j(\alpha)^2}{D(\alpha)} \right)^{b_j} \right) w_{\mathbf{n}''}^0(\alpha).$$

— The killing of the pair  $u_n, \overline{u_n}$  is responsible for the loss  $C_n(\alpha)^2$ , whereas the prefactor  $N_{\alpha}^{-6s} \eta^6$  is a saving that comes from this prefactor in one of the two weights  $w_{\mathbf{n}}^0(\alpha), w_{\mathbf{n}'}^0(\alpha)$ .

— In addition, the term  $\prod_{|j| \leq M} D(\alpha)^{-b_j} (C_{\alpha}(n))^{2b_j}$  is a second saving—which we will not exploit. It comes from the centering of  $b_j$ -terms  $|u_n|^2$  that were changed into  $|u_n|^2 - \xi_n$ .

— Note also that  $(a_j - b_j)$  terms  $|u_j|^2$  are also changed into  $\xi_j$ , and the extra weights (namely  $\prod_{|j| \leq M} C_j(\alpha)^{2(a_j - b_j)}$ ) will be absorbed by  $\prod_{|j| \leq M} \xi_j$ . We deduce

$$\begin{aligned} (4.21) &\leq N_{\alpha}^{-6s} \eta^6 C_n(\alpha)^2 \left( \prod_{|j| \leq M} C_j(\alpha)^{2(a_j - b_j)} \right) w_{\mathbf{n}''}^0(\alpha) \\ &\leq N_{\alpha}^{-4s+2\tau} \eta^4 \left( \prod_{|j| \leq M} C_j(\alpha)^{2(a_j - b_j)} \right) w_{\mathbf{n}''}^0(\alpha). \end{aligned}$$

Hence, for these contributions to (4.20) we have

$$\begin{aligned} (4.20) &\lesssim_r \varepsilon^{-1/10^4} N_{\alpha}^{-4s+2\tau} \eta^4 \left( \prod_{|j| \leq M} (C_j(\alpha)^2 |\xi_j|)^{a_j - b_j} \right) w_{\mathbf{n}''}^0(\alpha) \\ &\lesssim_r \varepsilon^{-1/10^4} N_{\alpha}^{-4s+2\tau} \eta^4 w_{\mathbf{n}''}^0(\alpha). \end{aligned}$$

*Case 2: A term  $|u_n|^2 - \xi_n$  goes away.* — In this case  $(\mathbf{n}, \mathbf{n}') \in E_{\mathbf{n}'', \mathbf{a}, \mathbf{b}}^{(2)}(n)$  and we simply modify the analysis from Case 1. It holds

$$\begin{aligned} w_{\mathbf{n}}^0(\alpha) w_{\mathbf{n}'}^0(\alpha) &= N_{\alpha}^{-6s} \eta^6 D(\alpha) \left( \prod_{|j| \leq M} C_j(\alpha)^{2(a_j - b_j)} \left( \frac{C_j(\alpha)^2}{D(\alpha)} \right)^{b_j} \right) w_{\mathbf{n}''}^0(\alpha). \\ (4.22) \quad &\leq N_{\alpha}^{-4s} \eta^{4-1/5} \left( \prod_{|j| \leq M} C_j(\alpha)^{2(a_j - b_j)} \right) w_{\mathbf{n}''}^0(\alpha). \end{aligned}$$

Hence, the contribution of these terms to (4.20) is controlled by

$$\begin{aligned} (4.20) &\lesssim_r \varepsilon^{-1/10^4} \eta^{4-1/5} N_{\alpha}^{-4s} \left( \prod_{|j| \leq M} (C_j(\alpha)^2 |\xi_j|)^{a_j - b_j} \right) w_{\mathbf{n}''}^0(\alpha) \\ &\lesssim_r \varepsilon^{-1/10^4} \eta^{4-1/5} N_{\alpha}^{-4s} w_{\mathbf{n}''}^0(\alpha). \end{aligned}$$

In both case, by choosing  $\varepsilon_*(s, r)$  sufficiently small we conclude that

$$(4.20) \leq \eta^{4-1/4} N_{\alpha}^{-4s} w_{\mathbf{n}''}^0(\alpha).$$

which is conclusive. We now turn to the proof of (4.19). Once again we may assume that

$$\|Q\|_{Y_{\alpha}^{\sup}} = 1, \quad \|Z\|_{Z^{\sup}} = 1.$$

By expanding  $Z$

$$Z = \sum_{|j_1|, |j_2| \leq M} (Z[\xi])_{\mathbf{e}_m(j_1) + \mathbf{e}_m(j_2)} (|u_{j_1}|^2 - \xi_{j_1}) (|u_{j_2}|^2 - \xi_{j_2}),$$

we obtain

$$\begin{aligned} \{Q, Z\}(\xi; u) &= 2i \sum_{\substack{\mathbf{n} \in \mathfrak{N}_{\leq \bar{r}^2} \\ |j_1|, |j_2| \leq M}} (k_{j_1} - \ell_{j_1}) (Q[\xi])_{\mathbf{n}} (Z[\xi])_{\mathbf{e}_m(j_1) + \mathbf{e}_m(j_2)} (|u_{j_2}|^2 - \xi_{j_2}) z_{\mathbf{n}}(u, I(u) - \xi) \\ &=: \sum_{\mathbf{n}''} (P[\xi])_{\mathbf{n}'} z_{\mathbf{n}''}(u, I(u) - \xi), \end{aligned}$$

In particular, a term  $|u_{j_1}|^2 - \xi_{j_1}$  goes away while a term  $|u_{j_2}|^2 - \xi_{j_2}$  appears. This term provides a saving  $D(\alpha)$ . We also stress out that there is no new pairing. Organizing the terms gives

$$(4.23) \quad (P[\xi])_{\mathbf{n}''} = 2i \sum_{\substack{\mathbf{n} \in \mathfrak{N}_{\leq \bar{r}^2} \\ |j_1|, |j_2| \leq M}} \mathbf{1}_{E_{\mathbf{n}'', 0, 0}^{(2)}(j_1)}(\mathbf{n}, \mathbf{e}_m(j_1) + \mathbf{e}_m(j_2)) (k_{j_1} - \ell_{j_1}) \times (Q[\xi])_{\mathbf{n}} (Z[\xi])_{\mathbf{e}_m(j_1) + \mathbf{e}_m(j_2)}.$$

For  $|j_1|, |j_2| \leq M$  and  $\mathbf{n} \in \mathfrak{N}_{\leq \bar{r}^2}$ ,  $\mathbf{n}'' \in \mathfrak{N}$ , we have that

$$(4.24) \quad (\mathbf{n}, \mathbf{e}_m(j_1) + \mathbf{e}_m(j_2)) \in E_{\mathbf{n}'', 0, 0}^{(2)}(j_1) \implies \mathbf{n}'' = \mathbf{n} + \mathbf{e}_m(j_2),$$

and, in such a case, we have

$$\mathbf{w}_{\mathbf{n}}^0(\alpha) = \frac{1}{D(\alpha)} \mathbf{w}_{\mathbf{n}''}^0(\alpha) = N_{\alpha}^{-2s} \eta^{2+1/5} \mathbf{w}_{\mathbf{n}''}^0(\alpha).$$

We deduce from the explicit formula (4.23) that for all  $\xi \in \mathcal{U}_s(\varepsilon)$ ,

$$(4.25) \quad |(P[\xi])_{\mathbf{n}'}| \lesssim_r \varepsilon^{-1/10^4} \mathbf{w}_{\mathbf{n}}^0(\alpha) \leq N_{\alpha}^{-2s} \eta^{2+1/5-2/10^4} \mathbf{w}_{\mathbf{n}''}^0(\alpha).$$

This completes the proof of Proposition 4.15.  $\square$

We now establish another bilinear estimate in the same spirit, but this time we measure the dependence on  $\xi$  of the coefficients of re-centered polynomials. The output polynomial  $P$  depends on the coefficients through the coefficients of  $Q$  and  $H$  and the terms  $\xi$  coming from the centering of new actions generated by the Poisson bracket.

**PROPOSITION 4.16.** — *For all  $Q, H \in X_{\leq \bar{r}^2}(\varepsilon)$  we have*

$$(4.26) \quad \|\{Q, H\}\|_{Y_{\alpha}^{\text{lip}}} \leq \eta^{4-1/4} N_{\alpha}^{-4s} (\|Q\|_{Y_{\alpha}^{\text{sup}}} \|H\|_{Y_{\alpha}^{\text{sup}}} + \|Q\|_{Y_{\alpha}^{\text{lip}}} \|H\|_{Y_{\alpha}^{\text{sup}}} + \|Q\|_{Y_{\alpha}^{\text{sup}}} \|H\|_{Y_{\alpha}^{\text{lip}}}).$$

Moreover, for all integrable quartic term  $Z \in X_{4, \text{Int}}(\varepsilon)$ ,

$$(4.27) \quad \|\{Q, Z\}\|_{Y_{\alpha}^{\text{lip}}} \leq \eta^{2+1/5-2/10^4} N_{\alpha}^{-2s} (\|Q\|_{Y_{\alpha}^{\text{lip}}} \|Z\|_{Z^{\text{sup}}} + \eta^2 N_{\alpha}^{-2s} \|Q\|_{Y_{\alpha}^{\text{sup}}} \|Z\|_{Z^{\text{lip}}}).$$

Since the proof of this proposition is relatively long (although close to the previous one) we postpone it to Section A.1 of the appendix.

## 5. SMALL-DIVISOR ESTIMATE FOR THE MODULATED FREQUENCIES

Initially, the frequencies are

$$\omega_n = \lambda_n^2 = g(n, n).$$

At each step of the iteration scheme, the centering procedure modulates the frequencies with internal parameters  $\xi$  as introduced in the beginning of Section 3. These parameters should be viewed as the actions of the initial data  $\phi \in \Pi_M h^s$  in the final coordinates. Recall that  $\mathbf{n}_-$ , defined in (3.6), is the size of the smallest index in  $\text{supp } \mathbf{n}$  that is not fully paired.

We also recall that  $\bar{r} = 10^5 r$ .

**DEFINITION 5.1** (Non-resonant set of initial data). — Given a sequence of modulated frequencies  $\omega(\xi) = (\omega_n(\xi))_{n \in \mathbb{Z}_M^d}$  and a multi-index  $\mathbf{n} \in \mathfrak{N}_{\leq 2\bar{r}}$ , we denote the (modulated) resonance function by

$$\Omega_{\mathbf{n}}(\omega(\xi)) := (\mathbf{k} - \ell) \cdot \omega(\xi) = \sum_{n \in \mathbb{Z}_M^d} (k_n - \ell_n) \omega_n(\xi).$$

For  $\gamma \in (0, 1)$  and  $\mathcal{E} \subset \mathfrak{N}_{\leq 2\bar{r}}$  the *non-resonant set* of internal parameters restricted to  $\mathcal{E}$  is

$$(5.1) \quad \Xi_{\varepsilon, \gamma}(\omega, \mathcal{E}) := \left\{ \xi \in \mathbb{R}^{\mathbb{Z}_M^d} \mid \min_{\mathbf{n} \in \mathcal{E}} |\Omega_{\mathbf{n}}(\xi)| \mathbf{n}_-^{2s} > \gamma \varepsilon^2 \right\},$$

and the *non-resonant set* of functions is

$$(5.2) \quad \mathcal{U}_{\varepsilon, \gamma}(\omega, \mathcal{E}) := \left\{ \phi \in \Pi_M B_s(\varepsilon) \mid \xi(\phi) \in \Xi_{\varepsilon, \gamma}(\omega, \mathcal{E}) \right\}.$$

When  $\mathcal{E} = \mathfrak{N}_{\leq 2\bar{r}}$  we just write  $\Xi_{\varepsilon, \gamma}(\omega)$  and  $\mathcal{U}_{\varepsilon, \gamma}(\omega)$ .

**REMARK 5.2.** — Note that the non-resonant set  $\mathcal{U}_{\varepsilon, \gamma}(\omega)$  is an open subset of  $\Pi_M B_s(\varepsilon)$ .

By construction, if  $\xi \in \Xi_{\varepsilon, \gamma}(\omega)$  then for all  $\mathbf{n} \in \mathfrak{N}_{\leq 2\bar{r}}$  we have

$$|\Omega_{\mathbf{n}}(\xi)| > \gamma \varepsilon^2 \mathbf{n}_-^{-2s}.$$

**LEMMA 5.3.** — Let  $\mathcal{A} \subset \mathbb{Z}^d$  be a non-empty finite set, let  $N = \#\mathcal{A} \geq 1$  and  $\rho > 0$ . We have

$$\text{meas}(\Pi_{\mathcal{A}} B_s(\rho)) = \frac{\pi^N \rho^{2N}}{(N+1)!} \prod_{n \in \mathcal{A}} \langle n \rangle^{-2s},$$

where  $\text{meas}$  is the (canonical) Lebesgue measure on  $\mathbb{C}^{\mathcal{A}} \cong (\mathbb{R}^2)^{\mathcal{A}}$ .

*Proof.* — Set

$$\begin{aligned} F: \Pi_{\mathcal{A}} B_s(\rho) &\longrightarrow B_{\mathbb{C}^{\mathcal{A}}}(0, 1) \\ (\phi)_n &\longmapsto (x_n)_n = \rho^{-1} \langle n \rangle^{2s} \phi_n. \end{aligned}$$

A (linear) change of variables gives

$$\begin{aligned} \text{meas}(\Pi_{\mathcal{A}} B_s(\rho)) &= \rho^{2\sharp \mathcal{A}} \prod_{n \in \mathcal{A}} \langle n \rangle^{-2s} \text{meas}(F(\Pi_{\mathcal{A}} B_s(\rho))) \\ &= \rho^{2N} \prod_{n \in \mathcal{A}} \langle n \rangle^{-2s} \text{meas}(B_{\mathbb{C}^{\mathcal{A}}}(0, 1)). \end{aligned}$$

We conclude by using the explicit formula for the volume of the unit Euclidean ball of  $\mathbb{R}^d$ :

$$\text{Vol}(B_{\mathbb{R}^d}(0, 1)) = \frac{\pi^{d/2}}{\Gamma(d/2 + 1)}.$$

This completes the proof of Lemma 5.3.  $\square$

We are now ready to state the measure estimate of the non-resonant set of functions, for a given sequence of modulated frequencies  $\omega(\xi) := (\omega_n(\xi))_{n \in \mathbb{Z}_M^d}$ .

**PROPOSITION 5.4** (Measure estimate of the non-resonant set)

*Let  $\omega \in C^1(\mathcal{U}_s(\varepsilon), \mathbb{R}^{\mathbb{Z}_M^d})$  be a sequence of modulated frequencies such that for all  $k, n \in \mathbb{Z}_M^d$ ,*

$$(5.3) \quad |\partial_{\xi_k}(\omega_n(\xi) - \xi_n)| \leq \varepsilon^{1/2}.$$

*Then, provided that  $\varepsilon_*(s, r)$  is small enough, we have that for all  $\gamma \in (0, 1)$ ,*

$$\text{meas}(\mathcal{U}_{\varepsilon, \gamma}(\omega)) \geq \text{meas}(\Pi_M B_s(\varepsilon))(1 - \gamma \varepsilon^{-1/10^4}).$$

*Proof.* — Fix  $\mathbf{n} \in \mathfrak{N}_{\leq 2\bar{r}}$ ,  $\gamma \in (0, 1)$ ,  $\varepsilon \in (0, \varepsilon_*)$  and set

$$\mathcal{S}_{\mathbf{n}, \gamma, \varepsilon}(\omega) := \left\{ \phi \in \Pi_M B_s(\varepsilon) \mid |\Omega_{\mathbf{n}}(\xi(\phi))| \leq \gamma \varepsilon^2 \mathbf{n}_-^{-2s} \right\}.$$

Let us prove that there exists  $C > 0$  such that for all  $\mathbf{n}$  and  $\gamma$

$$(5.4) \quad \text{meas}(\mathcal{S}_{\mathbf{n}, \gamma, \varepsilon}(\omega)) \lesssim \gamma (\sharp \mathbb{Z}_M^d) \text{meas}(\Pi_M B_s(\varepsilon)).$$

We shall conclude by observing that

$$\Pi_M B_s(\varepsilon) \setminus \mathcal{U}_{\gamma, \varepsilon} = \left\{ \phi \in \Pi_M B_s(\varepsilon) \mid \min_{\mathbf{n} \in \mathfrak{N}_{\leq 2\bar{r}}} |\Omega_{\mathbf{n}}(\xi(\phi))| \mathbf{n}_-^{2s} \leq \gamma \varepsilon^2 \right\} = \bigcup_{\mathbf{n} \in \mathfrak{N}_{\leq 2\bar{r}}} \mathcal{S}_{\mathbf{n}, \gamma, \varepsilon}(\omega),$$

and the result follows from (5.4):

$$\begin{aligned} \text{meas}(\Pi_M B_s(\varepsilon) \setminus \mathcal{U}_{\gamma, \varepsilon}) &\lesssim \gamma \sharp(\mathfrak{N}_{\leq 2\bar{r}}) \sharp(\mathbb{Z}_M^d) \text{meas}(\Pi_M B_s(\varepsilon)) \\ &\leq \gamma \varepsilon^{-1/10^4} \text{meas}(\Pi_M B_s(\varepsilon)), \end{aligned}$$

where we used the bound (4.4). It remains to prove (5.4), for fixed  $\mathbf{n} \in \mathfrak{N}_{\leq 2\bar{r}}$ . Take  $n_0 \in \text{supp}(\mathbf{n})$  such that

$$|n_0| = \mathbf{n}_-, \quad |k_{n_0} - \ell_{n_0}| \geq 1.$$

To decide whether a function belongs to  $\mathcal{S}_{\mathbf{n}, \gamma, \varepsilon}(\omega)$  or not, it suffices to move the its  $n_0$ -th action  $\xi_{n_0}(\phi)$ . Hence, we will freeze the other variables and, abusing notation, we write

$$\omega(\xi) = \omega(\xi_{n_0}),$$

and see the frequencies  $\omega_n$  as functions

$$\xi_{n_0} \in I_{n_0} \longmapsto \omega_n(\xi_{n_0}) \in \mathbb{R}, \quad I_{n_0} = [-\langle n_0 \rangle^{-2s} 400\varepsilon^2, \langle n_0 \rangle^{-2s} 400\varepsilon^2] \subset \mathbb{R}.$$

Once the variables  $(\xi_n)_{n \neq n_0}$  are fixed, we deduce from the mean value theorem and from the assumption (5.3) that for all  $(\xi_{n_0}, \xi'_{n_0}) \in I_{n_0}^2$ ,

$$|\omega_{n_0}(\xi_{n_0}) - \omega_{n_0}(\xi'_{n_0}) - (\xi_{n_0} - \xi'_{n_0})| \leq \varepsilon^{1/2} |\xi_{n_0} - \xi'_{n_0}|,$$

and, for all  $n \neq n_0$ ,

$$|\omega_n(\xi_{n_0}) - \omega_n(\xi'_{n_0})| \leq \varepsilon^{1/2} |\xi_{n_0} - \xi'_{n_0}|.$$

We deduce from the triangle inequality that

$$\begin{aligned} |\Omega_{\mathbf{n}}(\xi_{n_0}) - \Omega_{\mathbf{n}}(\xi'_{n_0})| &\geq |k_{n_0} - \ell_{n_0}| |\omega_{n_0}(\xi_{n_0}) - \omega_{n_0}(\xi'_{n_0})| \\ &\quad - \sum_{n \in \mathbb{Z}_M^d \setminus \{n_0\}} |k_n - \ell_n| |\omega_n(\xi_{n_0}) - \omega_n(\xi'_{n_0})| \\ &\geq |\xi_{n_0} - \xi'_{n_0}| (|k_{n_0} - \ell_{n_0}| - 2\bar{\varepsilon} \varepsilon^{1/2}) \\ &\geq \frac{1}{2} |\xi_{n_0} - \xi'_{n_0}|. \end{aligned}$$

It follows that for all  $\phi, \phi' \in \Pi_M B_s(\varepsilon)$  such that  $\phi_n = \phi'_n$  when  $n \neq n_0$ ,

$$(5.5) \quad \phi, \phi' \in E_{\mathbf{n}, \gamma, \varepsilon}(\omega) \implies \frac{1}{2} |\xi_{n_0}(\phi) - \xi_{n_0}(\phi')| \leq |\Omega_{\mathbf{n}}(\xi(\phi)) - \Omega_{\mathbf{n}}(\xi(\phi'))| \leq 2\gamma \mathbf{n}_-^{-2s} \varepsilon^2.$$

We deduce from this that

$$(5.6) \quad \sup_{z \in \Pi_{\mathbb{Z}_M^d \setminus \{n_0\}} B_s(\varepsilon)} \int_{\mathbb{C}} \mathbf{1}_{\mathcal{S}_{\mathbf{n}, \gamma, \varepsilon}(\omega)}(z; \phi_{n_0}) d\phi_{n_0} \leq 4\pi\gamma \mathbf{n}_-^{-2s} \varepsilon^2.$$

Indeed, using polar coordinates for  $\phi_{n_0}$  and using that the condition to belong to  $\mathcal{S}_{\mathbf{n}, \gamma, \varepsilon}$  only depends on  $|\phi_{n_0}|^2$ , we obtain, for fixed  $z \in \Pi_{\mathbb{Z}_M^d \setminus \{n_0\}}$ , that

$$\begin{aligned} \int_{\mathbb{C}} \mathbf{1}_{\mathcal{S}_{\mathbf{n}, \gamma, \varepsilon}(\omega)}(z; |\phi_{n_0}|^2) d\phi_{n_0} &= \int_{(0, 2\pi)} d\theta \int_{(0, 1)} \mathbf{1}_{\mathcal{S}_{\mathbf{n}, \gamma, \varepsilon}(\omega)}(z; \rho^2) \rho d\rho \\ &= \pi \int_{(0, 1)} \mathbf{1}_{\mathcal{S}_{\mathbf{n}, \gamma, \varepsilon}(\omega)}(z; \rho^2) d\rho^2. \end{aligned}$$

Then, to prove (5.6) we just observe from (5.5) that if  $\rho, \rho' \in (0, 1)^2$  are such that

$$\mathbf{1}_{\mathcal{S}_{\mathbf{n}, \gamma, \varepsilon}(\omega)}(z; \rho^2) = \mathbf{1}_{\mathcal{S}_{\mathbf{n}, \gamma, \varepsilon}(\omega)}(z; (\rho')^2) = 1,$$

then

$$|\rho^2 - (\rho')^2| \leq 2\gamma \mathbf{n}_-^{-2s} \varepsilon^2.$$

To conclude the proof of (5.4), we freeze the variables  $z = (\phi_n)_{n \in \mathbb{Z}_M^d \setminus \{n_0\}}$  in the measure estimate by applying Fubini's theorem:

$$\begin{aligned} \text{meas}(\mathcal{S}_{\mathbf{n}, \gamma, \varepsilon}(\omega)) &= \int_{\Pi_M B_s(\varepsilon)} \mathbf{1}_{\mathcal{S}_{\mathbf{n}, \gamma, \varepsilon}(\omega)}(\phi) d\phi \\ &\leq \int_{\Pi_{\mathbb{Z}_M^d \setminus \{n_0\}} B_s(\varepsilon)} \left( \int_{\mathbb{C}} \mathbf{1}_{\mathcal{S}_{\mathbf{n}, \gamma, \varepsilon}(\omega)}(z; \phi_{n_0}) d\phi_{n_0} \right) dz. \end{aligned}$$

According to (5.6), we deduce that

$$\text{meas}(\mathcal{S}_{\mathbf{n}, \gamma, \varepsilon}(\omega)) \leq 4\pi\gamma \mathbf{n}_-^{-2s} \varepsilon^2 \text{meas}(\Pi_{\mathbb{Z}_M^d \setminus \{n_0\}} B_s(\varepsilon)).$$

Lemma 5.3 gives

$$\text{meas}(\Pi_{\mathbb{Z}_M^d \setminus \{n_0\}} B_s(\varepsilon)) = \frac{\#\mathbb{Z}_M^d}{\pi \varepsilon^2} \langle n_0 \rangle^{2s} \text{meas}(\Pi_M B_s(\varepsilon)),$$

and we obtain

$$\text{meas}(\mathcal{S}_{\mathbf{n}, \gamma, \varepsilon}(\omega)) \leq 2\gamma (\#\mathbb{Z}_M^d) \mathbf{n}^{-2s} \langle n_0 \rangle^{2s} \text{meas}(\Pi_M B_s(\varepsilon)) \lesssim \gamma (\#\mathbb{Z}_M^d) \text{meas}(\Pi_M B_s(\varepsilon)).$$

This completes the proof of Proposition 5.4.  $\square$

## 6. NORMAL FORM WITH INTERNAL PARAMETERS

We state in this section our main normal form theorem from which we will deduce Theorem 2.7 in Section 7. This normal form is based on an induction on the frequency scales  $N_\alpha$ , for  $\alpha \in \{0, \dots, \beta\}$ .

**6.1. PREPARATION.** — Let  $H_{\text{lo}} \in \mathcal{H}_{\leq 4r}(\mathbb{Z}_M^d)$  be a real polynomial of degree less than or equal to  $4r$  as in the assumptions of Theorem 2.7 where, without loss of generality,<sup>(15)</sup> we supposed that  $f'(0)$  is suitably chosen such that

$$(6.1) \quad H_{\text{lo}}(u) = \frac{1}{2} \sum_{n \in \mathbb{Z}_M^d} (\lambda_n^2 + \frac{1}{2}|u_n|^2) |u_n|^2 + \sum_{j=3}^{2r} P^{(2j)}(u).$$

For  $j \in \{3, \dots, 2r\}$ , the real homogeneous polynomial  $P^{(2j)} \in \mathcal{H}_{2j}(\mathbb{Z}_M^d)$  satisfies

$$(6.2) \quad \|P^{(2j)}\|_\infty \leq \varepsilon^{-4c_* \nu r}.$$

Moreover, it is  $\varepsilon^{c_* \nu}$ -resonant in the sense that

$$P^{(2j)}(u) = \sum_{\vec{n} \in \mathcal{N}_{2j}} P_{\vec{n}}^{(2j)} u_{\vec{n}}, \quad P_{\vec{n}}^{(2j)} \neq 0 \implies \left| \sum_{i=1}^{2j} (-1)^i \lambda_{n_i}^2 \right| \leq \varepsilon^{c_* \nu}.$$

We re-center  $H_{\text{lo}}$  around parameters  $\xi$ , and we introduce a decomposition into quadratic terms, quartic integrable terms and terms of order greater than or equal to 6.

**LEMMA 6.1.** — *Let  $H_{\text{lo}}$  be as in (6.1). There exist two polynomials  $Q \in X_{\leq 4r}(\varepsilon)$ ,  $Z_4 \in X_{4, \text{Int}}(\varepsilon)$ , and some modulated frequencies, such that the following decomposition holds for all  $\xi \in \mathcal{U}_s(\varepsilon)$ :*

$$(6.3) \quad H_{\text{lo}}(u) = \frac{1}{2} \sum_{n \in \mathbb{Z}_M^d} \omega_n(\xi) (|u_n|^2 - \xi_n) + Z_4(\xi; u) + Q(\xi; u),$$

where

$$\Pi_{\deg=2} Q = \Pi_{\deg=4} Q = 0.$$

We have the bounds

$$\|Q\|_{Z^{\text{sup}}} + \|Q\|_{Z^{\text{lip}}} + \|Z_4\|_{Z^{\text{sup}}} + \|Z_4\|_{Z^{\text{lip}}} \lesssim \varepsilon^{-4c_* \nu r} \leq \varepsilon^{-1/10^4}.$$

<sup>(15)</sup>Since we are in the small data regime the sign of  $f'(0)$  has indeed no role to play.

The modulated frequencies are

$$\omega_n(\xi) = \lambda_n^2 + \xi_n + h_n(\xi),$$

where the last term  $h_n(\xi)$  is a polynomial function of  $(\xi_\ell)_{\ell \in \mathbb{Z}_M^d}$ . For all  $\xi \in \mathcal{U}_s(\varepsilon)$  and  $|k|, |n| \leq M$ ,

$$|\partial_{\xi_k}(\omega_n(\xi) - \xi_n)| \lesssim \varepsilon^{2-1/10^4} \leq \varepsilon.$$

*Proof.* — We set

$$Z_4[\xi] := \Pi_{\deg=4} H_{\text{lo}}[\xi], \quad Q[\xi] := H_{\text{lo}}[\xi] - Z_4[\xi] - Z_2[\xi],$$

where

$$Z_2[\xi](u, y) = \frac{1}{2} \sum_{n \in \mathbb{Z}_M^d} \omega_n(\xi) y_n$$

and we define

$$Q(\xi; u) := Q[\xi](u, I(u) - \xi), \quad Z_4(\xi; u) := Z_4[\xi](u, I(u) - \xi).$$

Then, the decomposition (6.3) follows from the identity

$$H_{\text{lo}}(u) = H_{\text{lo}}[\xi](u, I(u) - \xi).$$

Since  $P^{(2j)}$  is  $\varepsilon^{-c_*\nu}$ -resonant, we have from Step 1 in paragraph 2.6 that the quartic terms are integrable, which proves that  $Z_4 \in X_{4, \text{Int}}(\varepsilon)$ .

The bounds on the polynomial norms and on the modulated frequencies follow from the estimates (6.2) on the coefficients of  $P^{(2j)}$  (before the re-centering around  $\xi$ ), and from the fact that  $\xi$  is a small parameter.  $\square$

**6.2. STATEMENT.** — In the previous section, we described the structure of the initial Hamiltonian  $H_{\text{lo}} \in X_{\leq 2r}(\varepsilon)$ , written in (6.3). We are now ready to formulate the main normal form theorem, which is based on an induction on frequency scales. The current iteration step is  $\alpha \in \{0, \dots, \beta\}$ . Before giving the results we recall that the formalism was described in Section 3 and in particular in notation 3.10. We also recall that the non-resonant set of parameters  $\Xi(\omega)$  was defined in Definition 5.1.

For  $\alpha \in \{0, \dots, \beta\}$  and  $c > 0$  a universal constant to be determined later, we introduce

$$(6.4) \quad \varepsilon_\alpha := 10\varepsilon - c\alpha r \varepsilon^{3/2},$$

Given the parameter  $\gamma(\alpha) = 4^\alpha \gamma$ , the modulated frequencies  $\omega^{(\alpha)}(\xi)$  defined iteratively in Theorem 6.2, and the non-resonant set defined in (5.2), we fix

$$(6.5) \quad \Xi_\alpha := \Xi_{\varepsilon, \gamma(\alpha)}(\omega^{(\alpha)}, \Lambda_{\alpha+1}).$$

We are now ready to state the main normal form theorem.

**THEOREM 6.2** (Normal form at scale  $\alpha$ ). — *For all  $\alpha \in \{1, \dots, \beta\}$  and  $\xi \in \mathcal{U}_s(\varepsilon)$ , there exists a symplectic transformation  $\tau_{\alpha, \xi} : h^s(\mathbb{Z}_M^d) \rightarrow h^s(\mathbb{Z}_M^d)$  such that*

$$H_{\text{lo}} \circ \tau_{\alpha, \xi}(u) = Z_2^{(\alpha)}(\xi; u) + Z_4^{(\alpha)}(\xi; u) + Q^{(\alpha)}(\xi; u) + R^{(\alpha)}(\xi; u),$$

where the different terms are organized as follows:

(1)  $Z_2^{(\alpha)}$  collects the quadratic terms: for all  $\xi \in \mathcal{U}_s(\varepsilon)$

$$Z_2^{(\alpha)}(\xi; u) = \frac{1}{2} \sum_{n \in \mathbb{Z}_M^d} \omega_n^{(\alpha)}(\xi) (|u_n|^2 - \xi_n),$$

where  $\omega_n^{(\alpha)} \in C^1(\mathcal{U}_s(\varepsilon), \mathbb{R})$  for all  $|n| \leq M$ , with, for all  $|k| \leq M$  and  $\xi \in \mathcal{U}_s(\varepsilon)$ ,

$$(6.6) \quad |\partial_{\xi_k}(\omega_n^{(\alpha)}(\xi) - \xi_n)| \leq (\alpha + 1)\varepsilon.$$

In addition, the modulated frequencies are stable in the sense that (when  $\alpha \geq 1$ )

$$(6.7) \quad |\omega_n^{(\alpha)}(\xi) - \omega_n^{(\alpha-1)}(\xi)| \leq N_{\alpha-1}^{-4s} \varepsilon^3.$$

(2)  $Z_4^{(\alpha)} \in X_{4, \text{Int}}(\varepsilon)$  collects quartic terms, and

$$(6.8) \quad \max(\|Z_4^{(\alpha)}\|_{Z^{\text{sup}}}, \|Z_4^{(\alpha)}\|_{Z^{\text{lip}}}) \lesssim_{\alpha} \varepsilon^{-1/10^4}.$$

(3)  $Q^{(\alpha)} \in X_{\leq 2\bar{r}}(\varepsilon)$  operates at scale  $\alpha$ :

$$(6.9) \quad \|Q^{(\alpha)}\|_{Y_{\alpha}^{\text{sup}}} \lesssim_{\alpha} \varepsilon^{-1/10^4}, \quad \|Q^{(\alpha)}\|_{Y_{\alpha}^{\text{lip}}} \lesssim_{\alpha} \varepsilon^{-1/15}.$$

It is at least of order 6, and  $(cr)^{c\alpha r} \varepsilon^{c*\nu}$ -resonant for some universal constant  $c > 0$ :

$$(6.10) \quad (Q^{(\alpha)}[\xi])_{\mathbf{n}} \neq 0 \implies \deg(\mathbf{n}) \geq 6, \quad \left| \sum_{n \in \mathbb{Z}_M^d} (k_n - \ell_n) \lambda_n^2 \right| \leq (cr)^{c\alpha r} \varepsilon^{c*\nu}.$$

(4)  $R^{(\alpha)}$  is a smooth map on  $\mathcal{U}_s(\varepsilon) \times h^s(\mathbb{Z}_M^d)$ , and for all  $\xi \in \Xi_{\alpha-1}$  and  $u \in \mathcal{V}_{\alpha,s}(\varepsilon_{\alpha}, \xi)$ ,

$$(6.11) \quad \|\nabla_u R^{(\alpha)}(\xi; u)\|_{h^s} \lesssim_{\alpha,r} \varepsilon^{2r+1} \|u\|_{h^s}.$$

(5) The symplectic transformation satisfies, for all  $\xi \in \mathcal{U}_s(\varepsilon)$  and  $u \in \mathcal{V}_{\alpha,s}(\varepsilon_{\alpha}, \xi)$ ,

$$(6.12) \quad \|\tau_{\alpha,\xi}(u) - u\|_{h^s} \lesssim_{\alpha,r} \varepsilon^{3/4} \|u\|_{h^s},$$

$$(6.13) \quad \|\mathrm{d}\tau_{\alpha,\xi}(u) - \mathrm{Id}\|_{h^s \rightarrow h^s} \lesssim_{\alpha,r} \varepsilon^{3/4}.$$

Moreover, for all  $\phi \in \Pi_M B_s(\varepsilon_{\alpha})$  and  $u \in \mathcal{V}_{\alpha,s}(\varepsilon_{\alpha}, \xi(\phi))$ ,

$$(6.14) \quad \|\mathrm{d}_{\phi} \tau_{\alpha,\xi(\phi)}(u)\|_{h^s \rightarrow h^s} \lesssim_{\alpha,r} \varepsilon^{1-2/5}.$$

Here are some comments on the above theorem.

*Modulated frequencies.* — In the iterative procedure of increasing the scales  $\alpha$ , the successive polynomials we generate have new integrable terms that modulate the frequencies when we re-center the actions around  $\xi$ . The condition (6.6) guarantees that the modulated frequencies satisfy the Lipschitz assumption (5.3) from Proposition 5.4 (which is essential to prove the small-divisor bounds for parameters in a set with an asymptotically full measure).

*Terms of type  $Q$ .* — These terms (once re-centered) are of order greater than 6. They are either integrable, or associated to multi-indices  $\mathbf{n} \in \Lambda_\alpha$  (the unpaired indices are larger than  $N_\alpha$ ) (even if this property is not explicitly stated in Theorem 6.2). These polynomials operate at scale  $\alpha$  (in the sense that their  $Y_\alpha^{\text{sup}}$ -norm is controlled by a small negative power of  $\varepsilon$ ). If  $N_\alpha$  is very large, we see from the vector field estimate of Proposition 4.6 that they generate stable dynamics.

*Non-resonance condition and terms of order four.* — The condition (6.10) ensures that every term of order less than or equal to 4 are integrable. Indeed, we proved in Section 2.6 Step 1 that  $(cr)^{c\alpha\tau} \varepsilon^{c_*\nu}$ -resonant terms of order 4 are integrable.

*Modulation parameters.* — The non-resonance condition for the modulation parameters  $\xi \in \Xi_{\alpha-1}$  ensures that  $R^{(\alpha)}$  is a remainder order  $r$  (this is encoded by (6.11)). Recall that this condition is dictated by the modulated frequencies  $\omega^{(\alpha-1)}$  at scale  $\alpha-1$ , and allows us to use a Birkhoff normal form to remove  $\Pi_{\Lambda_\alpha} Q^{(\alpha-1)}$  (up to a term of order  $r$ ), which do not commute with the actions with a frequency lower than  $N_\alpha$  (and which a priori do not operate at scale  $\alpha$ ). We will prove in Lemma 6.5 that the stability condition (6.7) and the definition (6.5) of  $\Xi_\alpha$  imply that

$$\Xi_\alpha \subset \Xi_{\alpha-1}.$$

This ensures that non-resonant parameters  $\xi$  at scale  $\alpha$  (in the sense of (6.5)) are also non-resonant on the previous scale  $\alpha-1$ .

*The symplectic transformations.* — At scale  $\alpha$ , the symplectic transformations (which are well-defined on the finite dimensional space  $\Pi_M h^s$ ) depend on the modulation parameter  $\xi \in \mathcal{U}_s(\varepsilon)$ . In addition,  $\tau_{\alpha,\xi}$  is close to the identity on the annulus  $\mathcal{V}_{\alpha,s}(\varepsilon_\alpha, \xi)$ .

REMARK 6.3. — We observe from (6.12) that

$$\tau_{\alpha,\xi}(\mathcal{V}_{\alpha,s}(\varepsilon_\alpha, \xi)) \subset \mathcal{V}_{0,s}(10\varepsilon, \xi),$$

and from (6.13) that for all  $u \in \mathcal{V}_{\alpha,s}(\varepsilon_\alpha, \xi)$ ,

$$\|\mathrm{d}\tau_{\alpha,\xi}(u)\|_{h^s \rightarrow h^s} + \|(\mathrm{d}\tau_{\alpha,\xi}(u))^{-1}\|_{h^s \rightarrow h^s} \lesssim 1.$$

In particular, if  $u := \tau_{\alpha,\xi}^{-1}(v) \in \mathcal{V}_{0,s}(10\varepsilon, \xi)$ , then, for fixed  $\xi$ ,

$$(6.15) \quad \|\mathrm{d}\tau_{\alpha,\xi}^{-1}(u)\|_{h^s \rightarrow h^s} = \|(\mathrm{d}\tau_{\alpha,\xi}(v))^{-1}\|_{h^s \rightarrow h^s} \lesssim 1.$$

The rest of this section is devoted to the proof of the normal form Theorem 6.2. It consists in an induction on the frequency scales, and we first give an outline of the proof in the paragraph below.

**6.3. PROOF: INDUCTION ON THE FREQUENCY SCALES.** — We fix a scale  $\alpha \in \{1, \dots, \beta-1\}$ , and we assume that Theorem 6.2 is proved up to scale  $\alpha$ . In this paragraph, we briefly introduce the strategy we use to get from the scale  $\alpha$  to the scale  $\alpha+1$ . For this purpose, we will remove non-integrable terms from  $Q^{(\alpha)}$  which are associated to multi-indices in  $\Lambda_{\alpha+1}$  (defined in (4.12))

$$\mathbf{n} \in \Lambda_{\alpha+1},$$

up to some new terms that are remainders if  $\xi \in \Xi_\alpha$  is a non-resonant parameter. Using Lemma 4.8 and Remark 4.9, this will imply that the new Hamiltonian  $Q^{(\alpha+1)}$  operates at scale  $\alpha+1$ .

The algorithm to get from step  $\alpha$  to step  $\alpha+1$  requires at most  $\kappa \leq 100r$  steps, which are detailed in Sections 6.6 and 6.7. At each step, we introduce a new Lie transform, whose properties are presented in Section 6.4. To control the terms generated by the Taylor expansion of the Hamiltonian in the new coordinates, we will use the main multilinear estimates for the Poisson bracket with respect to the  $Y_\alpha^{\text{sup}}$ -norm, already proved in Section 4.4.

If we reach the scale  $\alpha = \beta$ , where  $\beta$  is sufficiently large, then the vector field generated by  $Q^{(\beta)}$  is smoothing when it is restricted to the open set  $\mathcal{V}_{\alpha,s}(10\varepsilon, \xi)$ , and the saving of  $N_\beta^{-4s}$  in the vector field estimate is sufficient to control the actions over a very long time  $T \sim N_\beta^{2s} = \varepsilon^{-r}$ .

**6.4. THE LIE TRANSFORMS.** — We introduce and study the Lie transforms used to remove the undesirable terms from the Hamiltonian (namely the ones with a frequency smaller than  $N_{\alpha+1}$  that is not paired).

**6.4.1. Truncation argument.** — Let  $\alpha \in \{0, \dots, \beta-1\}$  and  $\Xi_\alpha$  be the admissible sets of parameters defined in (6.5), just before the statement of Theorem 6.2. We first introduce a cutoff function, which essentially ensures that the Lie transform is defined for all  $\xi$ , but is only effective for non-resonant parameters  $\xi \in \Xi_\alpha$ .

**DEFINITION 6.4 (Truncation).** — Fix  $\varphi \in C_c^\infty(\mathbb{R})$  valued in  $[0, 1]$  such that

$$\text{supp } \varphi \subset [-1, 1], \quad \varphi \equiv 1 \quad \text{on } [-\tfrac{1}{2}, \tfrac{1}{2}], \quad \varphi < 1 \quad \text{on } \mathbb{R} \setminus [-\tfrac{1}{2}, \tfrac{1}{2}].$$

Then, given  $\omega^{(\alpha)}$  and  $\gamma(\alpha) = 4^\alpha \gamma$  (constructed at scale  $\alpha$  in Theorem 6.2 by the induction assumption) we define the cutoff-function

$$\mathfrak{h}^{(\alpha)} : \xi \in \mathcal{U}_s(\varepsilon) \mapsto \prod_{\substack{\mathbf{n} \in \Lambda_{\alpha+1} \\ \deg(\mathbf{n}) \leq 2\bar{r}}} \left( 1 - \varphi(\gamma(\alpha)^{-1} \mathbf{n}_-^{2s} \varepsilon^{-2} \Omega_{\mathbf{n}}(\omega^{(\alpha)}(\xi))) \right),$$

where  $\Lambda_{\alpha+1}$  was defined in (4.12).

We now prove important properties of the cutoff function  $\mathfrak{h}^{(\alpha)}$ .

**LEMMA 6.5.** — Let  $\alpha \in \{1, \dots, \beta\}$ , and suppose that Theorem 6.2 is proved up to scale  $\alpha$ . Then,

$$(6.16) \quad \xi \in \Xi_\alpha \implies \mathfrak{h}^{(\alpha)}(\xi) = 1.$$

Moreover, for all  $\xi \in \mathcal{U}_s(\varepsilon)$

$$(6.17) \quad \xi \in \text{supp } \mathfrak{h}^{(\alpha)} \implies \forall \mathbf{n} \in \Lambda_{\alpha+1}, \quad |\Omega_{\mathbf{n}}(\omega^{(\alpha)}(\xi))| > \frac{1}{2} \gamma(\alpha) \mathbf{n}_-^{-2s} \varepsilon^2,$$

and

$$(6.18) \quad \xi \in \text{supp } \mathfrak{h}^{(\alpha)} \implies \xi \in \Xi_{\alpha-1}.$$

REMARK 6.6. — In particular, we can deduce from (6.16) and (6.17) that  $\Xi_{\alpha} \subset \Xi_{\alpha-1}$ .

*Proof.* — The proof of (6.16) follows from the construction: suppose that  $\xi \in \Xi_{\alpha}$ , in which case for all  $\mathbf{n} \in \Lambda_{\alpha+1}$

$$|\Omega_{\mathbf{n}}(\omega^{(\alpha)}(\xi))| > \gamma(\alpha) \mathbf{n}_-^{-2s} \varepsilon^2.$$

Hence, we have that for all  $\mathbf{n} \in \Lambda_{\alpha+1}$ ,

$$\varphi(\gamma(\alpha)^{-1} \mathbf{n}_-^{2s} \varepsilon^{-2} \Omega_{\mathbf{n}}(\omega^{(\alpha)}(\xi))) = 0,$$

and we deduce that  $\mathfrak{h}^{(\alpha)}(\xi) = 1$ .

We now show (6.17), and suppose that  $\xi$  is in the support of  $\mathfrak{h}^{(\alpha)}$ . Then, we have from the construction of  $\varphi$  and  $\mathfrak{h}$  that, for all  $\mathbf{n} \in \Lambda_{\alpha+1}$ ,

$$|\Omega_{\mathbf{n}}(\omega^{(\alpha)}(\xi))| > \frac{1}{2} \gamma(\alpha) \mathbf{n}_-^{-2s} \varepsilon^2.$$

This proves (6.17). Lastly, to show (6.18), we observe from the stability property for the modulated frequencies claimed in (6.7) that

$$\begin{aligned} |\Omega_{\mathbf{n}}(\omega^{(\alpha-1)}(\xi))| &\geq |\Omega_{\mathbf{n}}(\omega^{(\alpha)}(\xi))| - |\Omega_{\mathbf{n}}(\omega^{(\alpha-1)}(\xi)) - \Omega_{\mathbf{n}}(\omega^{(\alpha)}(\xi))| \\ &\geq \frac{1}{2} \gamma(\alpha) \mathbf{n}_-^{-2s} \varepsilon^2 - 2\bar{r} N_{\alpha-1}^{-4s} \varepsilon^3. \end{aligned}$$

When  $\mathbf{n} \in \Lambda_{\alpha+1}$  then  $\mathbf{n}_- \leq N_{\alpha+1}$  and we obtain that for all  $\mathbf{n} \in \Lambda_{\alpha+1}$ ,

$$|\Omega_{\mathbf{n}}(\omega^{(\alpha-1)}(\xi))| > \frac{1}{4} \gamma(\alpha) \mathbf{n}_-^{-2s} \varepsilon^2 = \gamma(\alpha-1) \mathbf{n}_-^{-2s} \varepsilon^2,$$

under the condition

$$(N_{\alpha+1}/N_{\alpha-1})^{2s} N_{\alpha-1}^{-2s} \varepsilon < \frac{1}{4} \gamma(\alpha),$$

which, under our choice of parameters in Section 4.1, is equivalent to  $\varepsilon^{1-1/100} N_{\alpha}^{-2s} < 4^{\alpha-1} \varepsilon^{1/30}$ . This shows that  $\xi \in \Xi_{\alpha-1}$ , which is precisely the claim (6.18).  $\square$

LEMMA 6.7 (Small divisor estimate). — *Let  $\alpha \in \{0, \dots, \beta\}$  such that Theorem 6.2 is proved up to scale  $\alpha$ . Suppose that  $(\tilde{w}_n(\xi))_{n \in \mathbb{Z}_M^d}$  satisfies, for all  $\xi \in \mathcal{U}_s(\varepsilon)$ , and  $k, n \in \mathbb{Z}_M^d$ ,*

$$(6.19) \quad |\omega_n^{(\alpha)}(\xi) - \tilde{w}_n(\xi)| \lesssim_r \varepsilon^3 N_{\alpha}^{-4s},$$

$$(6.20) \quad |\partial_{\xi_k}(\omega_n^{(\alpha)}(\xi) - \tilde{w}_n(\xi))| \lesssim_r \varepsilon N_{\alpha}^{-2s}.$$

Then for all  $\xi \in \mathcal{U}_s(\varepsilon)$  and  $\mathbf{n} \in \Lambda_{\alpha+1}$ ,

$$(6.21) \quad \frac{\mathfrak{h}^{(\alpha)}(\xi)}{|\Omega_{\mathbf{n}}(\tilde{w}(\xi))|} \leq 4\gamma(\alpha)^{-1} \varepsilon^{-2} \mathbf{n}_-^{2s},$$

and, for all  $k \in \mathbb{Z}_M^d$ ,

$$(6.22) \quad \left| \partial_{\xi_k} \left( \frac{\mathfrak{h}^{(\alpha)}(\xi)}{\Omega_{\mathbf{n}}(\tilde{\omega}(\xi))} \right) \right| \leq 2\varepsilon^{-1/10^4} (\gamma(\alpha)^{-1} \varepsilon^{-2} \mathbf{n}_-^{2s})^2,$$

*Proof.* — First, we proceed as in the proof of (6.18) to deduce from (6.17), together with the stability assumption (6.19), that

$$(6.23) \quad \xi \in \text{supp } \mathfrak{h}^{(\alpha)} \implies \forall \mathbf{n} \in \Lambda_{\alpha+1}, |\Omega_{\mathbf{n}}(\tilde{\omega}(\xi))| \geq \frac{1}{4} \gamma(\alpha) \mathbf{n}_-^{2s} \varepsilon^2.$$

from which we deduce (6.21). To show (6.22), we take  $n \in \mathbb{Z}_M^d$  and we first observe that, thanks to (6.6), for all  $\mathbf{n} \in \Lambda_{\alpha+1}$  and  $\xi \in \mathcal{U}_s(\varepsilon)$ ,

$$|\partial_{\xi_n} \Omega_{\mathbf{n}}(\omega^{(\alpha)}(\xi))| + |\partial_{\xi_n} \Omega_{\mathbf{n}}(\tilde{\omega}(\xi))| \lesssim r.$$

We deduce that for all  $\xi \in \mathcal{U}_s(\varepsilon)$ ,

$$\begin{aligned} |\partial_{\xi_n} \mathfrak{h}^{(\alpha)}(\xi)| &\leq \sum_{\mathbf{n} \in \Lambda_{\alpha+1}} \gamma(\alpha)^{-1} \mathbf{n}_-^{2s} \varepsilon^{-2} \left| \partial_{\xi_n} \Omega_{\mathbf{n}}(\omega^{(\alpha)}(\xi)) \varphi'(\gamma(\alpha)^{-1} \mathbf{n}_-^{2s} \varepsilon^{-2} \Omega_{\mathbf{n}}(\omega^{(\alpha)}(\xi))) \right| \\ &\leq \varepsilon^{-1/10^4} \gamma(\alpha)^{-1} \varepsilon^{-2} N_{\alpha+1}^{2s}. \end{aligned}$$

Hence, for  $n \in \mathbb{Z}_M^d$ ,  $\mathbf{n} \in \Lambda_{\alpha+1}$  and  $\xi \in \mathcal{U}_s(\varepsilon)$ ,

$$\begin{aligned} \left| \partial_{\xi_n} \left( \frac{\mathfrak{h}^{(\alpha)}(\xi)}{\Omega_{\mathbf{n}}(\tilde{\omega}(\xi))} \right) \right| &\leq \left| \frac{\partial_{\xi_n} \mathfrak{h}^{(\alpha)}(\xi)}{\Omega_{\mathbf{n}}(\tilde{\omega}(\xi))} \right| + \mathfrak{h}^{(\alpha)}(\xi) \frac{|\partial_{\xi_n} \Omega_{\mathbf{n}}(\tilde{\omega}(\xi))|}{\Omega_{\mathbf{n}}(\tilde{\omega}(\xi))^2} \\ &\leq 2\varepsilon^{-1/10^4} (\gamma(\alpha)^{-1} \varepsilon^{-2} N_{\alpha+1}^{2s})^2, \end{aligned}$$

which concludes the proof of Lemma 6.7.  $\square$

6.4.2. *Definition of the transformation.* — Let us introduce linear operator used to solve the cohomological equations.

DEFINITION 6.8. — Given some modulated frequencies  $(\tilde{\omega}_n(\xi))_{n \in \mathbb{Z}_M^d}$  satisfying (6.19) and (6.20). We define the linear operator

$$(6.24) \quad \mathcal{L}_{\alpha, \tilde{\omega}} = \begin{cases} X_{\leq 2\bar{r}}(\varepsilon) \longrightarrow X_{\leq 2\bar{r}}(\varepsilon) \\ H \longmapsto \frac{i}{2} \mathfrak{h}^{(\alpha)}(\xi) \sum_{\mathbf{n} \in \Lambda_{\alpha+1}} \frac{(H[\xi])_{\mathbf{n}}}{\Omega_{\mathbf{n}}(\tilde{\omega}(\xi))} z_{\mathbf{n}}(u, I(u) - \xi) \end{cases}$$

When the context is clear we may abuse notations and simply write

$$\mathcal{L} := \mathcal{L}_{\alpha, \tilde{\omega}}.$$

The purpose of  $\mathcal{L}$  is to solve a cohomological equation.

LEMMA 6.9. — For all  $Q \in X_{\leq 2\bar{r}}(\varepsilon)$ , and frequencies  $(\tilde{\omega}_n)_{n \in \mathbb{Z}_M^d}$ , if

$$Z_2(\tilde{\omega}, \xi; u) := \sum_{|n| \leq M} \tilde{\omega}_n(\xi) (|u_n|^2 - \xi_n),$$

then for all  $\xi \in \mathcal{U}_s(\varepsilon)$  the polynomial  $\mathcal{L}(Q)(\xi) \in \mathcal{H}_{\leq 2\overline{r}}(\mathbb{Z}_M^d)$  solves the cohomological equation

$$(6.25) \quad Q(\xi; u) + \{Z_2(\tilde{\omega}), \mathcal{L}(Q)\}(\xi; u) = (\text{Id} - \mathfrak{h}^{(\alpha)}(\xi)\Pi_{\Lambda_{\alpha+1}})Q(\xi; u).$$

In particular, when  $\xi \in \Xi_\alpha$  then

$$Q(\xi; u) + \{Z_2(\tilde{\omega}), \mathcal{L}(Q)\}(\xi; u) = (\text{Id} - \Pi_{\Lambda_{\alpha+1}})Q(\xi; u).$$

*Proof.* — The equality (6.25) follows directly from the Definition 6.8 of  $\mathcal{L}_{\alpha, \tilde{\omega}}$ :

$$(6.26) \quad \begin{aligned} \{Z_2(\tilde{\omega}), \mathcal{L}(Q)(u)\}(\xi; u) &= -\mathfrak{h}^{(\alpha)}(\xi) \sum_{\mathbf{n} \in \Lambda_{\alpha+1}} (Q[\xi])_{\mathbf{n}} z_{\mathbf{n}}(u, I(u) - \xi) \\ &= -\mathfrak{h}^{(\alpha)}(\xi)\Pi_{\Lambda_{\alpha+1}}Q(\xi; u). \end{aligned}$$

The second equality is a consequence of the construction of the truncation  $\mathfrak{h}^{(\alpha)}$ , which satisfies  $\mathfrak{h}^{(\alpha)}(\xi) = 1$  for all  $\xi \in \Xi_\alpha$ .  $\square$

LEMMA 6.10. — For all  $Q \in X_{\leq 2\overline{r}}(\varepsilon)$ , we have

$$\begin{aligned} \|\mathcal{L}_{\alpha, \tilde{\omega}}(Q)\|_{Y_\alpha^{\text{sup}}} &\lesssim \gamma(\alpha)^{-1} \varepsilon^{-2} N_{\alpha+1}^{2s} \|\Pi_{\Lambda_{\alpha+1}} Q\|_{Y_\alpha^{\text{sup}}}, \\ \|\mathcal{L}_{\alpha, \tilde{\omega}}(Q)\|_{Y_\alpha^{\text{lip}}} &\lesssim \gamma(\alpha)^{-1} \varepsilon^{-2} N_{\alpha+1}^{2s} (\varepsilon^{-1/30-3/100-1/10^4} \|\Pi_{\Lambda_{\alpha+1}} Q\|_{Y_\alpha^{\text{sup}}} + \|\Pi_{\Lambda_{\alpha+1}} Q\|_{Y_\alpha^{\text{lip}}}). \end{aligned}$$

*Proof.* — The proof of the  $Y_\alpha^{\text{sup}}$ -bound is a direct consequence of the small divisor estimates (6.21) from Lemma 6.7. For the second bound, we also use the Leibniz rule: recall that for  $\mathbf{n} \in \mathfrak{N}$  and  $\xi \in \mathcal{U}_s(\varepsilon)$ ,

$$(\mathcal{L}_{\alpha, \tilde{\omega}}(Q)[\xi])_{\mathbf{n}} = \mathbf{1}_{\Lambda_{\alpha+1}}(\mathbf{n}) \frac{\mathfrak{h}^{(\alpha)}(\xi)}{\Omega_{\mathbf{n}}(\tilde{\omega}(\xi))} (Q[\xi])_{\mathbf{n}},$$

and given  $k \in \mathbb{Z}_M^d$  we have from the Leibniz rule that for  $\mathbf{n} \in \Lambda_{\alpha+1}$  and  $\xi \in \mathcal{U}_s(\varepsilon)$ ,

$$\partial_{\xi_k} (\mathcal{L}_{\alpha, \tilde{\omega}}(Q)[\xi])_{\mathbf{n}} = \partial_{\xi_k} \left( \frac{\mathfrak{h}^{(\alpha)}(\xi)}{\Omega_{\mathbf{n}}(\tilde{\omega}(\xi))} \right) (Q[\xi])_{\mathbf{n}} + \frac{\mathfrak{h}^{(\alpha)}(\xi)}{\Omega_{\mathbf{n}}(\tilde{\omega}(\xi))} \partial_{\xi_k} (Q[\xi])_{\mathbf{n}}.$$

According to the small divisor bounds (6.21) and (6.22), we obtain

$$\begin{aligned} &|\partial_{\xi_k} (\mathcal{L}_{\alpha, \tilde{\omega}}(Q)[\xi])_{\mathbf{n}}| \\ &\leq 2\varepsilon^{-1/10^4} (\gamma(\alpha)^{-1} \varepsilon^{-2} \mathbf{n}_-^{2s})^2 |(Q[\xi])_{\mathbf{n}}| + 4\gamma(\alpha)^{-1} \varepsilon^{-2} \mathbf{n}_-^{2s} |\partial_{\xi_k} (Q[\xi])_{\mathbf{n}}| \\ &\leq 2\varepsilon^{-1/10^4} (\gamma(\alpha)^{-1} \varepsilon^{-2} \mathbf{n}_-^{2s})^2 w_{\mathbf{n}}^0(\alpha) \|Q\|_{Y_\alpha^{\text{sup}}} + 4\gamma(\alpha)^{-1} \varepsilon^{-2} \mathbf{n}_-^{2s} w_{\mathbf{n}}^1(\alpha) \|Q\|_{Y_\alpha^{\text{lip}}} \\ &\lesssim w_{\mathbf{n}}^1(\alpha) \left( \varepsilon^{-1/10^4} (\gamma(\alpha)^{-1} \varepsilon^{-2} \mathbf{n}_-^{2s})^2 N_\alpha^{-2s} \eta^2 \|Q\|_{Y_\alpha^{\text{sup}}} + \gamma(\alpha)^{-1} \varepsilon^{-2} \mathbf{n}_-^{2s} \|Q\|_{Y_\alpha^{\text{lip}}} \right). \end{aligned}$$

Therefore, we have from the the assumption  $\mathbf{n} \in \Lambda_{\alpha+1}$  and the relation (4.6) between the parameters that

$$\begin{aligned} &(w_{\mathbf{n}}^1(\alpha))^{-1} |\partial_{\xi_k} (\mathcal{L}_{\alpha, \tilde{\omega}}(Q)[\xi])_{\mathbf{n}}| \\ &\lesssim \gamma(\alpha)^{-1} \varepsilon^{-2} N_{\alpha+1}^{2s} (\varepsilon^{-1/30-3/100-1/10^4} \|Q\|_{Y_\alpha^{\text{sup}}} + \|Q\|_{Y_\alpha^{\text{lip}}}). \end{aligned}$$

This completes the proof of the second estimate, recalling the Definition 4.3 of the  $Y_\alpha^{\text{lip}}$ -norm.  $\square$

COROLLARY 6.11. — For all  $Q \in X_{\leq 2\bar{\tau}}(\varepsilon)$ ,  $H \in X_{\leq \bar{\tau}^2}(\varepsilon)$  and  $Z_4 \in X_{4,\text{Int}}(\varepsilon)$  we have

$$\begin{aligned} \|\{H, \mathcal{L}_{\alpha, \tilde{\omega}}(Q)\}\|_{Y_{\alpha}^{\text{sup}}} &\leq \varepsilon^{3/2} N_{\alpha}^{-2s} \|H\|_{Y_{\alpha}^{\text{sup}}} \|\Pi_{\Lambda_{\alpha+1}} Q\|_{Y_{\alpha}^{\text{sup}}}, \\ \|\{Z_4, \mathcal{L}_{\alpha, \tilde{\omega}}(Q)\}\|_{Y_{\alpha}^{\text{sup}}} &\leq \varepsilon^{1/15} \|Z_4\|_{Z^{\text{sup}}} \|\Pi_{\Lambda_{\alpha+1}} Q\|_{Y_{\alpha}^{\text{sup}}}. \end{aligned}$$

Moreover, if we suppose that

$$(6.27) \quad \|\Pi_{\Lambda_{\alpha+1}} Q\|_{Y_{\alpha}^{\text{sup}}} + \|H\|_{Y_{\alpha}^{\text{sup}}} + \|Z_4\|_{Z^{\text{sup}}} + \|Z_4\|_{Z^{\text{lip}}} \leq \varepsilon^{-1/10^3}$$

and

$$(6.28) \quad \|\Pi_{\Lambda_{\alpha+1}} Q\|_{Y_{\alpha}^{\text{lip}}} + \|H\|_{Y_{\alpha}^{\text{lip}}} \lesssim \varepsilon^{-1/15},$$

then

$$\|\{H, \mathcal{L}_{\alpha, \tilde{\omega}}(Q)\}\|_{Y_{\alpha}^{\text{lip}}} \leq \varepsilon N_{\alpha}^{-2s}, \quad \|\{Z_4, \mathcal{L}_{\alpha, \tilde{\omega}}(Q)\}\|_{Y_{\alpha}^{\text{lip}}} \leq \varepsilon^{1/50}.$$

*Proof.* — It suffices to combine the (bilinear) estimate on the Poisson bracket between two polynomials given by Proposition 4.15 with the operator bound from Lemma 6.10, recalling the relation 4.6 between the parameters:

$$\begin{aligned} \|\{H, \mathcal{L}_{\alpha, \tilde{\omega}}(Q)\}\|_{Y_{\alpha}^{\text{sup}}} &\leq \eta^{4-1/4} N_{\alpha}^{-4s} \|H\|_{Y_{\alpha}^{\text{sup}}} \|\mathcal{L}_{\alpha, \tilde{\omega}}(Q)\|_{Y_{\alpha}^{\text{sup}}} \\ &\lesssim \gamma(\alpha)^{-1} (\eta \varepsilon^{-1})^2 (N_{\alpha+1}/N_{\alpha})^{2s} \eta^{2-1/4} N_{\alpha}^{-2s} \|H\|_{Y_{\alpha}^{\text{sup}}} \|\Pi_{\Lambda_{\alpha+1}} Q\|_{Y_{\alpha}^{\text{sup}}} \\ &\leq \varepsilon^{3/2} N_{\alpha}^{-2s} \|H\|_{Y_{\alpha}^{\text{sup}}} \|\Pi_{\Lambda_{\alpha+1}} Q\|_{Y_{\alpha}^{\text{sup}}}. \end{aligned}$$

As for the Lipschitz estimates, we have from the assumptions and from Lemma 6.10 that:

$$\|\mathcal{L}_{\alpha, \tilde{\omega}}(Q)\|_{Y_{\alpha}^{\text{sup}}} \lesssim \gamma(\alpha)^{-1} \varepsilon^{-2} N_{\alpha+1}^{2s} \varepsilon^{-1/10^3}, \quad \|\mathcal{L}_{\alpha, \tilde{\omega}}(Q)\|_{Y_{\alpha}^{\text{lip}}} \lesssim \gamma(\alpha)^{-1} \varepsilon^{-2} N_{\alpha+1}^{2s} \varepsilon^{-1/15}.$$

Then, Proposition 4.16 gives

$$\begin{aligned} &\|\{H, \mathcal{L}_{\alpha, \tilde{\omega}}\}\|_{Y_{\alpha}^{\text{lip}}} \\ &\leq \eta^{4-1/4} N_{\alpha}^{-4s} \gamma(\alpha)^{-1} \varepsilon^{-2} N_{\alpha+1}^{2s} (\varepsilon^{-1/10^3} \|H\|_{Y_{\alpha}^{\text{sup}}} + \varepsilon^{-1/15} \|H\|_{Y_{\alpha}^{\text{sup}}} + \varepsilon^{-1/10^3} \|H\|_{Y_{\alpha}^{\text{lip}}}) \\ &\leq \varepsilon N_{\alpha}^{-2s}. \end{aligned}$$

Similarly,

$$\begin{aligned} &\|\{Z_4, \mathcal{L}_{\alpha, \tilde{\omega}}\}\|_{Y_{\alpha}^{\text{lip}}} \\ &\lesssim \eta^{2+1/5-2/10^4} N_{\alpha}^{-2s} \gamma(\alpha)^{-1} \varepsilon^{-2} N_{\alpha+1}^{2s} (\varepsilon^{-1/15} \|Z\|_{Z_4^{\text{sup}}} + \eta^2 N_{\alpha}^{-2s} \varepsilon^{-1/10^3} \|Z_4\|_{Z^{\text{lip}}}) \\ &\leq \varepsilon^{1/50}. \end{aligned}$$

This completes the proof of Corollary 6.11.  $\square$

NOTATION 6.12. — Given  $Q \in X_{\leq \bar{\tau}^2}$ , and  $t \in \mathbb{R}$ , we denote by  $\Phi_{\mathcal{L}(Q)}^t$  the Hamiltonian flow on  $h^s(\mathbb{Z}_M^d)$  generated by  $\mathcal{L}(Q)(\xi)$ . For  $\xi \in \mathcal{U}_s(\varepsilon)$  and  $u \in h^s(\mathbb{Z}_M^d)$  we denote by  $v(\xi; t)$  (or implicitly  $v(t)$ ) the solution of the finite-dimensional ODE

$$(6.29) \quad \begin{cases} i\partial_t v = \nabla(\mathcal{L}(Q))(\xi)(v), \\ v(0) = u. \end{cases}$$

REMARK 6.13. — Using the preservation of the  $\ell^2$ -mass, we note that the ODE (6.29) is globally well-posed, and  $\Phi_{\mathcal{L}(Q)}^t$  is well-defined on  $h^s(\mathbb{Z}_M^d)$  for all  $t \in \mathbb{R}$ . Moreover,

$$\Phi_{\mathcal{L}(Q)}^{-t} = (\Phi_{\mathcal{L}(Q)}^t)^{-1}.$$

6.5. ESTIMATES ON THE LIE TRANSFORMS. — We fix some frequencies  $(\tilde{\omega}_n)_{n \in \mathbb{Z}_M^d}$  satisfying (6.19) and (6.20), and we consider the linear operator  $\mathcal{L} = \mathcal{L}_{\alpha, \tilde{\omega}}$  as in Definition 6.8. We prove important properties on the flow  $(\Phi_{\mathcal{L}(Q)}^t)_{t \in [-1, 1]}$ , when  $Q \in X_{\leq 2\bar{r}}(\varepsilon)$  operates at scale  $\alpha$  (in the sense that its  $Y_\alpha^{\text{sup}}$ -norm is no too large).

LEMMA 6.14 (Stability). — *Let  $Q \in X_{\leq 2\bar{r}}(\varepsilon)$  be such that*

$$(6.30) \quad \|Q\|_{Y_\alpha^{\text{sup}}} \leq \varepsilon^{-1/10^3}.$$

*For all  $\xi \in \mathcal{U}_s(\varepsilon)$ , the transformation  $\Phi_{\mathcal{L}(Q)(\xi)}^t$  is close to the identity on the annulus  $\mathcal{V}_{\alpha, s}(10\varepsilon, \xi)$ : for all  $t \in [-1, 1]$  and  $u \in \mathcal{V}_{\alpha, s}(10\varepsilon, \xi)$ ,*

$$(6.31) \quad \|\Phi_{\mathcal{L}(Q)}^t(u) - u\|_{h^s} \leq \varepsilon^{3/2} \|u\|_{h^s} N_\alpha^{-2s}.$$

COROLLARY 6.15. — *We deduce from (6.31) the bound*

$$(6.32) \quad \sum_{n \in \mathbb{Z}_M^d} \langle n \rangle^{2s} |(\Phi_{\mathcal{L}(Q)}^t(u))_n|^2 - |u_n|^2 \lesssim \varepsilon^{3/2} \|u\|_{h^s}^2 N_\alpha^{-2s},$$

*which gives the stability of the annulus  $\mathcal{V}_{\alpha, s}(10\varepsilon, \xi)$  under the transformations  $\Phi_{\mathcal{L}(Q)(\xi)}^t$ .*

*Proof of Lemma 6.14.* — We denote

$$v(t) = \Phi_{\mathcal{L}_{\alpha, \tilde{\omega}}(Q)}^t(u),$$

solution to the ODE (6.29). We have from the Duhamel's integral formula that for all  $t$ ,

$$v(t) = u - i \int_0^t [\nabla \mathcal{L}(Q)(\xi)](v(t')) dt'.$$

Moreover, we deduce from the vector field estimate of Proposition 4.6 and from the operator bound for  $\mathcal{L}$  obtained in Lemma 6.10 that

$$\begin{aligned} \|[\nabla \mathcal{L}(Q)(\xi)](v)\|_{h^s} &\leq N_\alpha^{-4s} \varepsilon^{4-1/4} \|\mathcal{L}(Q)\|_{Y_\alpha^{\text{sup}}} \|v\|_{h^s} \\ &\leq 4\gamma(\alpha)^{-1} (N_{\alpha+1}/N_\alpha)^{2s} N_\alpha^{-2s} \varepsilon^{2-1/4} \|Q\|_{Y_\alpha^{\text{sup}}} \|v\|_{h^s} \\ &\leq 4\gamma(\alpha)^{-1} (N_{\alpha+1}/N_\alpha)^{2s} N_\alpha^{-2s} \varepsilon^{2-1/4-1/10^3} \|v\|_{h^s}. \end{aligned}$$

Under the relation (4.6) for the parameters we deduce that

$$\|[\nabla \mathcal{L}(Q)(\xi)](v)\|_{h^s} \leq N_\alpha^{-2s} \varepsilon^{3/2} \|v\|_{h^s}.$$

Therefore, for all  $t$ ,

$$\|v(t) - u\|_{h^s} \leq N_\alpha^{-2s} \varepsilon^{3/2} \left| \int_0^t \|v(t')\|_{h^s} dt' \right|.$$

We conclude from a bootstrap argument, using that  $\|v(t')\|_{h^s} \lesssim \varepsilon$  when (6.31) is satisfied.  $\square$

In the next lemma we estimate the differential of the vector field. Since the proof is quite long and does not require new ideas, we postpone it to Section A.2 of the appendix.

LEMMA 6.16. — For all  $Q \in X_{\leq 2\bar{r}}$ ,  $\xi \in \mathcal{U}_s(\varepsilon)$  and  $u \in \mathcal{V}_{\alpha,s}(20\varepsilon, \xi)$ , we have

$$\|\mathrm{d}\nabla\mathcal{L}(Q)(\xi)(u)\|_{h^s \rightarrow h^s} \leq \varepsilon^{1-1/8} \|Q\|_{Y_\alpha^{\mathrm{sup}}}.$$

We now state and prove two important corollaries.

COROLLARY 6.17. — Let  $Q \in X_{\leq 2\bar{r}}(\varepsilon)$  be such that

$$\|Q\|_{Y_\alpha^{\mathrm{sup}}} \leq \varepsilon^{-1/10^3}.$$

We have that for all  $\xi \in \mathcal{U}_s(\varepsilon)$ ,  $u \in \mathcal{V}_{\alpha,s}(10\varepsilon, \xi)$  and  $t \in [-1, 1]$ ,

$$\|\mathrm{d}\Phi_{\mathcal{L}(Q)}^t(u) - \mathrm{Id}\|_{h^s \rightarrow h^s} \leq \varepsilon^{3/4}.$$

*Proof.* — We use the equation, the previous lemmas 6.14, 6.16, and a bootstrap argument. Let  $\varphi \in h^s$ , and  $u \in \mathcal{V}_{\alpha,s}(10\varepsilon, \xi)$ . Set

$$v(t) := \Phi_{\mathcal{L}(Q)}^t(u), \quad \tilde{\varphi}(t) := \mathrm{d}\Phi_{\mathcal{L}(Q)}^t(u)\varphi.$$

We have from the equation that

$$i\partial_t \tilde{\varphi}(t) = [\mathrm{d}\nabla\mathcal{L}(Q)(\xi)(v(t))] (\tilde{\varphi}(t)).$$

Since  $\Phi_{\mathcal{L}(Q)}^0 = \mathrm{Id}$ , we have

$$\tilde{\varphi}(0) = \varphi.$$

Hence, the Duhamel's integral formula gives

$$\|\tilde{\varphi}(t) - \varphi\|_{h^s} \leq \left| \int_0^t \left\| [\mathrm{d}_u \nabla\mathcal{L}(Q)(v(t'))](\tilde{\varphi}(t')) \right\|_{h^s} dt' \right|.$$

We have from Lemma 6.14 (and its consequence (6.32)) that for all  $t \in [-1, 1]$ ,

$$v(t) \in \mathcal{V}_{\alpha,s}(11\varepsilon, \xi).$$

According to Lemma 6.16,

$$\|\tilde{\varphi}(t) - \varphi\|_{h^s} \leq \varepsilon^{1-1/8} \left| \int_0^t \|Q\|_{Y_\alpha^{\mathrm{sup}}} \|\tilde{\varphi}(t')\|_{h^s} dt' \right|,$$

and, under the assumption that  $\|Q\|_{Y_\alpha^{\mathrm{sup}}} \leq \varepsilon^{-1/10^3}$  we can conclude from a bootstrap argument. More precisely, we use that if

$$\sup_{|t'| \leq |t|} \|\tilde{\varphi}(t') - \varphi\|_{h^s} \leq \varepsilon^{3/4} \|\varphi\|_{h^s},$$

then

$$\sup_{|t'| \leq |t|} \|\tilde{\varphi}(t')\|_{h^s} \leq 2\|\varphi\|_{h^s}.$$

This concludes the proof of Corollary 6.17.  $\square$

The second corollary is more technical, but in the same spirit. It is crucial to estimate the measure of the non-resonant initial data in the original coordinates (i.e., before the successive Lie transformations).

COROLLARY 6.18. — Let  $Q \in X_{\leq 2\bar{r}}(\varepsilon)$  be such that for all  $\xi \in \mathcal{U}_s(\varepsilon)$

$$\|Q\|_{Y_\alpha^{\text{sup}}} \leq \varepsilon^{-1/10^3}, \quad \|Q\|_{Y_\alpha^{\text{lip}}} \leq \varepsilon^{-1/15},$$

Then for all  $t \in [-1, 1]$ ,  $\phi \in \Pi_M B_s(10\varepsilon)$  and  $u \in \mathcal{V}_{\alpha,s}(10\varepsilon, \xi(\phi))$  then

$$(6.33) \quad \|\mathrm{d}_\phi \Phi_{\mathcal{L}(Q)(\xi(\phi))}^t(u)\|_{h^s \rightarrow h^s} \lesssim \varepsilon^{1-2/5}.$$

*Proof.* — Let  $\varphi \in h^s(\mathbb{Z}_M^d)$ ,  $\phi \in \Pi_M B_s(10\varepsilon)$  and  $u \in \mathcal{V}_{\alpha,s}(10\varepsilon, \xi(\phi))$ . For  $|t| \leq 1$ , set

$$v(t) = \Phi_{\mathcal{L}(Q)(\xi(\phi))}^t(u), \quad \tilde{\varphi}(t) := [\mathrm{d}_\phi v(t, \phi)](\varphi).$$

In particular,  $v$  is solution in  $C([-1, 1], h^s(\mathbb{Z}_M^d))$  to the Cauchy problem

$$\begin{cases} i\partial_t v(t) = \nabla[\mathcal{L}(Q)(\xi(\phi))](v(t)), \\ v(0) = u. \end{cases}$$

We deduce from the stability Lemma 6.14 (and more precisely from (6.32)) that  $v(t) \in \mathcal{V}_{\alpha,s}(20\varepsilon, \xi)$  for all  $t$ . In particular

$$\|v(t)\|_{h^s} \leq 2\|u\|_{h^s} \leq 20\varepsilon.$$

To control  $\tilde{\varphi}(t)$  in  $h^s(\mathbb{Z}_M^d)$  we use the equation and a bootstrap: differentiating (6.29) in  $\phi$  and applying the Leibniz rule and the chain rule gives

$$(6.34) \quad \begin{cases} i\partial_t \tilde{\varphi}(t) = \nabla H^{(1)}(\xi(\phi), \varphi)(v) + \nabla H^{(2)}(\xi(\phi), \varphi)(v) \\ \quad + [\mathrm{d}\nabla[\mathcal{L}(Q)(\xi(\phi))](v)](\tilde{\varphi}(t)), \\ \tilde{\varphi}(0) = 0, \end{cases}$$

where we collected in  $H^{(1)}$  (resp.  $H^{(2)}$ ) the contributions when  $\mathrm{d}_\phi$  falls on the re-centered variable  $z_{\mathbf{n}}(u, I(u) - \xi(\phi))$  (resp. on the coefficient of  $\mathcal{L}(Q)(\xi(\phi))$ ), namely

$$\begin{aligned} H^{(1)}(\xi(\phi), \varphi; v) &= - \sum_{\mathbf{n} \in \Lambda_{\alpha+1}} \sum_{|k| \leq M} \varphi_k \bar{\phi}_k \frac{\mathfrak{h}^{(\alpha)}(\xi(\phi))}{\Omega_{\mathbf{n}}(\tilde{\omega}(\xi(\phi)))} Q[\xi(\phi)]_{\mathbf{n}} m_k z_{\mathbf{n}-\mathbf{e}_m(k)}(v, I(v) - \xi), \\ H^{(2)}(\xi(\phi), \varphi; v) &= \sum_{\mathbf{n} \in \mathfrak{N}_{\leq 2\bar{r}}} [H^{(2)}(\xi, \varphi)]_{\mathbf{n}} z_{\mathbf{n}}(v, I(v) - \xi), \end{aligned}$$

with, for fixed  $\mathbf{n} \in \mathfrak{N}_{\leq 2\bar{r}}$ ,

$$\begin{aligned} &[H^{(2)}(\xi(\phi), \varphi)]_{\mathbf{n}} \\ &= \sum_{|k| \leq M} \varphi_k \bar{\phi}_k \partial_{\xi_k} \left( \frac{\mathfrak{h}^{(\alpha)}(\xi(\phi))}{\Omega_{\mathbf{n}}(\tilde{\omega}(\xi(\phi)))} \right) (Q[\xi(\phi)])_{\mathbf{n}} + \varphi_k \bar{\phi}_k \frac{\mathfrak{h}^{(\alpha)}(\xi(\phi))}{\Omega_{\mathbf{n}}(\tilde{\omega}(\xi(\phi)))} \partial_{\xi_k} (Q[\xi(\phi)])_{\mathbf{n}}. \end{aligned}$$

First, we easily control the last term on the right-hand side of (6.34) from Lemma 6.16: for all  $t$ ,

$$\left\| [\mathrm{d}\nabla \mathcal{L}(Q)(v(t))](\tilde{\varphi}(t)) \right\|_{h^s} \leq \varepsilon^{1-1/8} \|Q\|_{Y_\alpha^{\text{sup}}} \|\tilde{\varphi}(t)\|_{h^s} \leq \varepsilon^{1-2/5} \|\tilde{\varphi}(t)\|_{h^s}.$$

Then, we deduce from the vector field estimate of Proposition 4.6 that, for  $i \in \{1, 2\}$ ,

$$(6.35) \quad \begin{aligned} \|\nabla H^{(i)}(\xi(\phi), \varphi)(v)\|_{h^s} &\leq \|H^{(i)}(\varphi)\|_{Y_\alpha^{\text{sup}}} N_\alpha^{-4s} \varepsilon^3 \|v(t)\|_{h^s} \\ &\lesssim \|H^{(i)}(\varphi)\|_{Y_\alpha^{\text{sup}}} N_\alpha^{-4s} \varepsilon^4. \end{aligned}$$

Let us first control the second contribution ( $i = 2$ ), when the derivative falls on the coefficients. According to the small divisor bounds (6.21) and (6.22), we obtain that for all  $\mathbf{n} \in \mathfrak{N}$ ,

$$\begin{aligned} & |[H^{(2)}(\xi(\phi), \varphi)]_{\mathbf{n}}| \\ & \leq \|\varphi\|_{h^s} \|\phi\|_{h^s} \gamma(\alpha)^{-1} \varepsilon^{-2} N_{\alpha+1}^{2s} \\ & \quad \times \left( 2\varepsilon^{-1/10^4} \gamma(\alpha)^{-1} \varepsilon^{-2} N_{\alpha+1}^{2s} |(Q[\xi(\phi)])_{\mathbf{n}}| + \max_{|k| \leq M} |\partial_{\xi_k}(Q[\xi])_{\mathbf{n}}| \right) \\ & \lesssim w_{\mathbf{n}}^0(\alpha) \|\varphi\|_{h^s} \|\phi\|_{h^s} \gamma(\alpha)^{-1} (\varepsilon^{-2} N_{\alpha+1}^{2s})^2 \left( \gamma(\alpha)^{-1} \varepsilon^{-1/10^4} \|Q\|_{Y_{\alpha}^{\sup}} + \|Q\|_{Y_{\alpha}^{\text{lip}}} \right) \\ & \lesssim_{\alpha} w_{\mathbf{n}}^0(\alpha) \|\varphi\|_{h^s} \varepsilon^{1-1/30} (\varepsilon^{-2} N_{\alpha+1}^{2s})^2 (\varepsilon^{-1/30-1/10^4-1/10^3} + \varepsilon^{-1/15}). \end{aligned}$$

We deduce from the Definition 4.3 of the norms that

$$\|H^{(2)}(\phi)\|_{Y_{\alpha}^{\sup}} \lesssim_{\alpha} \|\varphi\|_{h^s} \varepsilon^{1-1/30-1/15} (\varepsilon^{-2} N_{\alpha+1}^{2s})^2,$$

and therefore

$$\begin{aligned} \|\nabla H^{(2)}(\xi(\phi), \varphi)(v)\|_{h^s} & \lesssim_{\alpha} \|\varphi\|_{h^s} (N_{\alpha+1}/N_{\alpha})^{4s} \varepsilon^{1-1/30-1/15} \\ & \lesssim_{\alpha} \|\varphi\|_{h^s} \varepsilon^{1-1/30-1/15-1/50} \leq \varepsilon^{1-2/5} \|\varphi\|_{h^s}. \end{aligned}$$

Similarly, we prove that

$$\begin{aligned} \|H^{(1)}(\xi(\phi), \varphi)\|_{Y_{\alpha}^{\sup}} & \leq \|\varphi\|_{h^s} \|\phi\|_{h^s} (\gamma(\alpha)^{-1} \varepsilon^{-2} N_{\alpha+1}^{2s}) D(\alpha) \|Q\|_{Y_{\alpha}^{\sup}} \\ & \lesssim_{\alpha} \|\varphi\|_{h^s} \varepsilon^{1-1/30-1/5-1/10^3} (\varepsilon^{-2} N_{\alpha+1}^{2s})^2, \end{aligned}$$

and therefore

$$\begin{aligned} \|\nabla H^{(1)}(\xi(\phi), \varphi)(v)\|_{h^s} & \lesssim (N_{\alpha+1}/N_{\alpha})^{4s} \varepsilon^{1-1/30-1/5-1/10^3} \|\varphi\|_{h^s} \\ & \lesssim_{\alpha} \varepsilon^{1-1/30-1/5-1/10^3-1/50} \|\varphi\|_{h^s} \leq \varepsilon^{1-2/5} \|\varphi\|_{h^s}. \end{aligned}$$

Hence, we obtain from Duhamel's integral formula that, for all  $t \in [-1, 1]$ ,

$$\|\tilde{\varphi}(t)\|_{h^s} \leq 2\varepsilon^{1-2/5} \|\varphi\|_{h^s} + \varepsilon^{1-2/5} \left| \int_0^t \|\tilde{\varphi}(t')\|_{h^s} dt' \right|,$$

which in turn gives (6.33) from the Gronwall inequality.  $\square$

**6.6. THE ITERATION AT SCALE  $\alpha$ .** — In this paragraph,  $\alpha \in \{1, \dots, \beta - 1\}$  is fixed and we suppose that Theorem 6.2 is proved up to scale  $\alpha$ . We detail the main iteration to remove the monomials associated to the multi-indices  $\mathbf{n} \in \Lambda_{\alpha+1}$ , in order to eventually go to scale  $\alpha + 1$ .

We now suppose that Theorem 6.2 is proved up to the step  $\alpha \in \{0, \dots, \beta - 1\}$ : for all  $\xi \in \mathcal{U}_s(\varepsilon)$ , there exists  $\tau_{\alpha, \xi}$  such that

$$(6.36) \quad H_{\text{lo}} \circ \tau_{\alpha, \xi}(u) = Z_2^{(\alpha)}(\xi; u) + Z_4^{(\alpha)}(\xi; u) + Q^{(\alpha)}(\xi; u) + R^{(\alpha)}(\xi; u),$$

satisfying points (1)–(5) of Theorem 6.2. We define

$$\kappa := 40r,$$

and, for  $j \in \{0, \dots, \kappa\}$ ,

$$\varepsilon_{j, \alpha} := \varepsilon_{\alpha} - j\varepsilon^{3/2} = 10\varepsilon - (c\alpha r + j)\varepsilon^{3/2}.$$

The strategy to pass from the normal form at scale  $\alpha$  to the scale  $\alpha + 1$  essentially consists in gaining *homogeneity* in  $\varepsilon$  by applying successively the Lie transformation introduced and analyzed in Section 6.4. After  $\kappa$  iterations (the number  $\kappa$  is comparable to a multiple of  $r$ ) all the terms that are not under the normal form at scale  $\alpha + 1$  (namely, which contain unpaired indices smaller than  $N_{\alpha+1}$ ) eventually come with a factor  $\varepsilon^{2r}$ , provided  $\xi$  is a non-resonant parameter, and are therefore acceptable remainders. The next proposition states this general second induction scheme (on the homogeneity), performed between two steps of the main induction scheme (on the frequency scales  $N_\alpha$ ).

**PROPOSITION 6.19** (Iteration at step  $\alpha$ ). — *For all  $\xi \in \mathcal{U}_s(\varepsilon)$  and  $j \in \{0, \dots, \kappa\}$ , there exists a symplectic transformation  $\tau_{j,\alpha,\xi}$  on  $h^s(\mathbb{Z}_M^d)$  such that*

$$H_{\text{lo}} \circ \tau_{j,\alpha,\xi}(u) = Z_2^{(j,\alpha)}(\xi; u) + Z_4^{(j,\alpha)}(\xi; u) + Q^{(j,\alpha)}(\xi; u) + R^{(j,\alpha)}(\xi; u),$$

where for all  $j \in \{1, \dots, \kappa\}$ ,

$$(6.37) \quad \|\mathbf{1}_{\Xi_\alpha}(\xi) \Pi_{\Lambda_{\alpha+1}} Q^{(j,\alpha)}\|_{Y_\alpha^{\text{sup}}} \leq \varepsilon^{j/30-1/10^4}.$$

Moreover, the new Hamiltonian and  $\tau_{j,\alpha,\xi}$  satisfy Theorem 6.2 at scale  $\alpha$ :

(1) The quadratic term is

$$Z_2^{(j,\alpha)}(\xi; u) = \frac{1}{2} \sum_{n \in \mathbb{Z}_M^d} \omega_n^{(j,\alpha)}(\xi) (|u_n|^2 - \xi_n),$$

where for all  $\xi \in \mathcal{U}_s(\varepsilon)$  and  $(k, n) \in (\mathbb{Z}_M^d)^2$ ,

$$(6.38) \quad |\omega_n^{(j,\alpha)}(\xi) - \omega_n^{(j-1,\alpha)}(\xi)| \leq \varepsilon^{3+1/10} N_\alpha^{-4s},$$

$$(6.39) \quad |\partial_{\xi_k}(\omega_n^{(j,\alpha)}(\xi) - \omega_n^{(j-1,\alpha)}(\xi))| \leq \varepsilon^{1+1/10} N_\alpha^{-2s}.$$

(2) The quartic integrable terms collected in  $Z_4^{(j,\alpha)} \in X_{4,\text{Int}}(\varepsilon)$  satisfy

$$(6.40) \quad \max(\|Z_4^{(j,\alpha)}\|_{Z^{\text{sup}}}, \|Z_4^{(j,\alpha)}\|_{Z^{\text{lip}}}) \lesssim_\alpha \varepsilon^{-1/10^4}$$

(3)  $Q^{(j,\alpha)} \in X_{\leq 2\bar{r}}(\varepsilon)$  operates at scale  $\alpha$ :

$$(6.41) \quad \|Q^{(j,\alpha)}\|_{Y_\alpha^{\text{sup}}} \lesssim_\alpha \varepsilon^{-1/10^4}, \quad \|Q^{(j,\alpha)}\|_{Y_\alpha^{\text{lip}}} \lesssim_\alpha \varepsilon^{-1/15},$$

and for all  $\xi \in \mathcal{U}_s(\varepsilon)$ ,

$$(6.42) \quad (Q^{(j,\alpha)}[\xi])_{\mathbf{n}} \neq 0 \implies \deg(\mathbf{n}) \geq 6, \quad \left| \sum_{n \in \mathbb{Z}_M^d} (k_n - \ell_n) \lambda_n^2 \right| \leq (cr)^{c\alpha r + j} \varepsilon^{-c_* \nu}.$$

(4) If  $\xi \in \Xi_{\alpha-1}$  then for all  $u \in \mathcal{V}_{\alpha,s}(\varepsilon_{j,\alpha}, \xi)$ ,

$$(6.43) \quad \|\nabla R^{(j,\alpha)}(\xi; u)\|_{h^s} \lesssim_{r,\alpha} \varepsilon^{2r+1} \|u\|_{h^s}.$$

(5) The symplectic transformation satisfies, for all  $\xi \in \mathcal{U}_s(\varepsilon)$  and  $u \in \mathcal{V}_{\alpha,s}(\varepsilon_{j,\alpha}, \xi)$ ,

$$(6.44) \quad \|\tau_{j,\alpha,\xi}(u) - u\|_{h^s} \lesssim_{r,\alpha} \varepsilon^{3/4} \|u\|_{h^s},$$

$$(6.45) \quad \|\text{d}_u \tau_{j,\alpha,\xi}(u) - \text{Id}\|_{h^s \rightarrow h^s} \lesssim_{r,\alpha} \varepsilon^{3/4},$$

Moreover, for all  $\phi \in \Pi_M B_s(\varepsilon_{j,\alpha})$  and  $u \in \mathcal{V}(\varepsilon_{j,\alpha}, \xi)$ ,

$$(6.46) \quad \|\text{d}_\phi \tau_{j,\alpha,\xi}(\phi)(u)\|_{h^s \rightarrow h^s} \lesssim_{r,\alpha} \varepsilon^{1-2/5}.$$

Let us make a few comments.

– The main idea is to remove from  $Q^{(\alpha)}$  the monomials associated to  $\mathbf{n} \in \Lambda_{\alpha+1}$ , when  $\xi$  is in the non-resonant set  $\Xi_\alpha$ . Specifically, after  $\kappa = 40r$  steps we are able to replace  $\Pi_{\Lambda_{\alpha+1}} Q^{(\alpha)}$  by a term which is a remainder when  $\xi \in \Xi_\alpha$ , as claimed in (6.37). Recalling Definition 4.3 of the norms, the condition (6.37) can be formulated this way: for all  $\mathbf{n} \in \mathfrak{N}$ ,

$$(6.47) \quad \mathbf{n} \in \Lambda_{\alpha+1} \quad \text{and} \quad \xi \in \Xi_\alpha \implies |(Q^{(j,\alpha)}[\xi])_{\mathbf{n}}| \leq \varepsilon^{j/30-1/10^4} w_{\mathbf{n}}^0(\alpha).$$

– According to Lemma 4.6, after at most  $\kappa$  steps, the bound (6.37) implies that  $\Pi_{\Lambda_{\alpha+1}} Q^{(j,\alpha)}$  is a remainder in the sense of (6.11). On the other hand, the polynomial  $(\text{Id} - \Pi_{\Lambda_{\alpha+1}}) Q^{(j,\alpha)}$  operates at scale  $\alpha + 1$  according to Lemma 4.8 and Remark 4.9.

– At each iteration  $j$  we generate additional terms of type  $Z_2$ ,  $Z_4$  and  $Q$ , and we need to propagate the estimates for their norms.

*Proof.* — We proceed by induction on  $j$ , and start with the Hamiltonian from Theorem 6.2 at scale  $\alpha$ , as written in (6.36).

*Initialization.* — The inputs are  $\tau_{0,\alpha,\xi} := \tau_{\alpha,\xi}$  and

$$\omega^{(0,\alpha)} := \omega^{(\alpha)}, \quad Z_4^{(0,\alpha)} := Z_4^{(\alpha)}, \quad Q^{(0,\alpha)} := Q^{(\alpha)}, \quad R^{(0,\alpha)} = R^{(\alpha)}.$$

At step  $j = 0$  we ask nothing more than the estimates claimed in Theorem 6.2 at scale  $\alpha$ .

*Iteration  $j \rightarrow j + 1$ .* — To pass from step  $j$  to step  $j + 1$  we need to remove from  $Q^{(j,\alpha)}$  the monomials associated to multi-indices in  $\Lambda_{\alpha+1}$ , up to some new terms  $\varepsilon^{1/30}$ -smaller. To achieve this we consider the Lie transformation introduced and studied in Section 6.4.

*The cohomological equation.* — We note from the recurrence assumptions (6.38) and (6.39) that the modulated frequencies  $(\omega_n^{(j,\alpha)})$  satisfy the stability estimate (6.19) and (6.20), and we let, for  $\xi \in \mathcal{U}_s(\varepsilon)$ ,

$$\tilde{\omega}_n(\xi) := \omega_n^{(j,\alpha)}(\xi).$$

We introduce the corresponding auxiliary Hamiltonian

$$(6.48) \quad \chi(\xi) := \mathcal{L}_{\alpha,\tilde{\omega}}(Q^{(j,\alpha)})(\xi) =: \mathcal{L}(Q^{(j,\alpha)})(\xi).$$

Lemma 6.7 provides the small divisor estimates (6.21) and (6.22) for the multiplication operator  $\mathcal{L}$ , defined in Definition 6.8. In addition, Lemma 6.9 gives that for all  $\xi \in \mathcal{U}_s(\varepsilon)$ ,

$$(6.49) \quad Q^{(j,\alpha)}(\xi; u) + \{Z_2^{(j,\alpha)}(\tilde{\omega}), \chi\}(\xi; u) = (\text{Id} - \mathfrak{h}^{(\alpha)}(\xi) \Pi_{\Lambda_{\alpha+1}}) Q^{(j,\alpha)}(\xi; u).$$

Hence, when  $\xi$  is non-resonant (i.e.,  $\xi \in \Xi_\alpha$ , and therefore  $\mathfrak{h}^{(\alpha)}(\xi) = 1$ ) the Hamiltonian  $\chi$  is tuned to solve the cohomological equation, see (6.25). In this way, the

monomials associated to multi-indices in  $\Lambda_{\alpha+1}$  can be filtered and removed. Once again we stress out that

$$\mathcal{L}(Q^{(j,\alpha)}) = \mathcal{L}(\Pi_{\Lambda_{\alpha+1}} Q^{(j,\alpha)}).$$

*The new variables.* — The Lie transformation was defined in Definition 6.8: for  $t \in [-1, 1]$ , let

$$\Phi_{j,\alpha,\xi}^t := \Phi_{\chi}^t,$$

where the auxiliary Hamiltonian  $\chi = \chi(j, \alpha, \xi)$  was defined above (6.48). According to the recurrence assumption (more precisely the bound (6.41)), the  $Y_{\alpha}^{\sup}$ -norm of  $Q^{(j,\alpha)}$  is bounded by  $\varepsilon^{-1/10^3}$ . Applying Lemma 6.14, we obtain that for all  $\xi \in \mathcal{U}_s(\varepsilon)$ ,  $u \in \mathcal{V}_{\alpha,s}(10\varepsilon, \xi)$  and  $t \in [-1, 1]$

$$(6.50) \quad \|\Phi_{j,\alpha,\xi}^t(u) - u\|_{h^s} \leq \varepsilon^{3/2} \|u(t)\|_{h^s} N_{\alpha}^{-2s}.$$

In particular, recalling that  $\varepsilon_{j,\alpha} := 10\varepsilon - (c\alpha r + j)\varepsilon^{3/2}$ , we have that for all  $\xi \in \mathcal{U}_s(\varepsilon)$ ,

$$(6.51) \quad \Phi_{j,\alpha,\xi}^t \left( \mathcal{V}_{\alpha,s}(\varepsilon_{j+1,\alpha}, \xi) \right) \subset \mathcal{V}_{\alpha,s}(\varepsilon_{j,\alpha}, \xi).$$

Corollary 6.17 gives that for all  $\xi \in \mathcal{U}_s(\varepsilon)$  and  $u \in \mathcal{V}_{\alpha,s}(10\varepsilon, \xi)$ ,

$$(6.52) \quad \|\mathrm{d}_u \Phi_{j,\alpha,\xi}^t(u) - \mathrm{Id}\|_{h^s \rightarrow h^s} \leq \varepsilon^{3/4},$$

and Corollary 6.18 gives that for all  $\phi \in \Pi_M B_s(10\varepsilon)$  and  $u \in \mathcal{V}_{\alpha,s}(10\varepsilon, \xi(\phi))$ ,

$$(6.53) \quad \|\mathrm{d}_{\phi} \Phi_{j,\alpha,\xi(\phi)}^t(u)\|_{h^s \rightarrow h^s} \lesssim \varepsilon^{1-2/5}.$$

We now define the new transformation, at scale  $j+1$ , by

$$\tau_{j+1,\alpha,\xi} := \tau_{j,\alpha,\xi} \circ \Phi_{j,\alpha,\xi}^1,$$

and we will deduce from the above bounds and from the recurrence assumption the bounds (6.44), (6.45) and (6.46) at step  $j+1$ . According to (6.51) and the recurrence assumption, we have that

$$(6.54) \quad \tau_{j+1,\alpha,\xi}(\mathcal{V}_{\alpha,s}(\varepsilon_{j+1,\alpha}, \xi)) \subset \tau_{j,\alpha,\xi}(\mathcal{V}_{\alpha,s}(\varepsilon_{j,\alpha}, \xi)) \subset \mathcal{V}_{\alpha,s}(10\varepsilon, \xi).$$

The bound (6.44) at scale  $j+1$  follows from the triangle inequality:

$$\|\tau_{j+1,\alpha,\xi}(u) - u\|_{h^s} \leq \|\tau_{j,\alpha,\xi}(\Phi_{j,\alpha,\xi}^1(u)) - \Phi_{j,\alpha,\xi}^1(u)\|_{h^s} + \|\Phi_{j,\alpha,\xi}^1(u) - u\|_{h^s}.$$

When  $u \in \mathcal{V}_{\alpha,s}(\varepsilon_{\alpha,j+1}, \xi)$ , the property (6.51) allows to apply the estimate (6.44) at step  $j$  and we obtain from the bound (6.50) that

$$\|\tau_{j+1,\alpha,\xi}(u) - u\|_{h^s} \lesssim_{r,\alpha,j} \varepsilon^{3/4} \|\Phi_{j,\alpha,\xi}^1(u)\|_{h^s} + \varepsilon^{3/2} \|u\|_{h^s} N_{\alpha}^{-2s} \lesssim_{r,\alpha,j} \varepsilon^{3/4} \|u\|_{h^s}.$$

This shows (6.44). The bound (6.45) is proved similarly, but using also the chain rule:

$$\mathrm{d} \tau_{j+1,\alpha,\xi}(u) - \mathrm{Id} = (\mathrm{d} \tau_{j,\alpha,\xi}(\Phi_{j,\alpha,\xi}^1(u)) - \mathrm{Id}) \circ \mathrm{d} \Phi_{j,\alpha,\xi}^1(u) + \mathrm{d} \Phi_{j,\alpha,\xi}^1(u) - \mathrm{Id}.$$

Thanks to (6.54), we can apply the recurrence bound (6.45) at step  $j$  together with (6.52) to obtain

$$\begin{aligned} & \|d\tau_{j+1,\alpha,\xi}(u) - \text{Id}\|_{h^s \rightarrow h^s} \\ & \leq \|d\tau_{j,\alpha,\xi}(\Phi_{j,\alpha,\xi}^1(u)) - \text{Id}\|_{h^s \rightarrow h^s} \|d\Phi_{j,\alpha,\xi}^1(u)\|_{h^s \rightarrow h^s} + \|d\Phi_{j,\alpha,\xi}^1(u) - \text{Id}\|_{h^s \rightarrow h^s} \\ & \lesssim_{r,\alpha,j} \varepsilon^{3/4} \|d_u \Phi_{j,\alpha,\xi}^1(u)\|_{h^s \rightarrow h^s} + \varepsilon^{3/4} \lesssim_{r,\alpha,j} \varepsilon^{3/4}, \end{aligned}$$

which proves (6.45) at step  $j+1$ . The proof of (6.46) goes along the same lines and we do not give the details.

*The new expansion.* — We do the Taylor expansion at  $t=0$  of the Hamiltonian in the new coordinates

$$H_{\text{lo}}^{(j+1,\alpha)}(\xi; u) := H_{\text{lo}}(\xi; \tau_{j,\alpha,\xi} \circ \Phi_{j,\alpha,\xi}^1(u)).$$

Recall the standard notation

$$\text{ad}_\chi^0(H) = H, \quad \text{ad}_\chi^1(H) = \{H, \chi\}, \quad \text{ad}_\chi^\ell(H) = \{\text{ad}_\chi^{\ell-1}(H), \chi\}, \quad \ell \geq 1.$$

First, we expand

$$\begin{aligned} (6.55) \quad Z_2^{(j,\alpha)}(\xi; \Phi_{j,\alpha,\xi}^1(u)) &= Z_2^{(j,\alpha)}(\xi; u) + \{Z_2^{(j,\alpha)}, \chi\}(\xi; u) + \sum_{\ell=2}^{\kappa-1} \frac{1}{\ell!} \text{ad}_\chi^\ell(Z_2^{(j,\alpha)})(\xi; u) \\ &\quad + \int_0^1 \frac{(1-t)^{\kappa-1}}{(\kappa-1)!} \text{ad}_\chi^\kappa(Z_2^{(j,\alpha)})(\xi; \Phi_{j,\alpha,\xi}^t(u)) dt, \end{aligned}$$

$$\begin{aligned} (6.56) \quad Z_4^{(j,\alpha)}(\xi; \Phi_{j,\alpha,\xi}^1(u)) &= Z_4^{(j,\alpha)}(\xi; u) + \sum_{\ell=1}^{\kappa-1} \frac{1}{\ell!} \text{ad}_\chi^\ell(Z_4^{(j,\alpha)})(\xi; u) \\ &\quad + \int_0^1 \frac{(1-t)^{\kappa-1}}{(\kappa-1)!} \text{ad}_\chi^\kappa(Z_4^{(j,\alpha)})(\xi; \Phi_{j,\alpha,\xi}^t(u)) dt, \end{aligned}$$

$$\begin{aligned} (6.57) \quad Q^{(j,\alpha)}(\xi; \Phi_{j,\alpha,\xi}^1(u)) &= Q^{(j,\alpha)} + \sum_{\ell=1}^{\kappa-1} \frac{1}{\ell!} \text{ad}_\chi^\ell(Q^{(j,\alpha)})(\xi; u) \\ &\quad + \int_0^1 \frac{(1-t)^{\kappa-1}}{(\kappa-1)!} \text{ad}_\chi^\kappa(Q^{(j,\alpha)})(\xi; \Phi_{j,\alpha,\xi}^t(u)) dt. \end{aligned}$$

We collect the integral remainders in

$$I^{(j,\alpha)}(\xi; u) := \int_0^1 \frac{(1-t)^\kappa}{\kappa!} \text{ad}_\chi^\kappa(Z_2^{(j,\alpha)} + Z_4^{(j,\alpha)} + Q^{(j,\alpha)})(\xi; \Phi_{j,\alpha,\xi}^t(u)) dt,$$

and the higher order polynomial terms in

$$P^{(j,\alpha)}(\xi; u) := \{Z_4^{(j,\alpha)} + Q^{(j,\alpha)}, \chi\}(\xi; u) + \sum_{\ell=2}^{\kappa-1} \frac{1}{\ell!} \text{ad}_\chi^\ell(Z_2^{(j,\alpha)} + Z_4^{(j,\alpha)} + Q^{(j,\alpha)})(\xi; u).$$

Note that  $P^{(j,\alpha)} \in X(\varepsilon)$  and, for all  $\xi \in \mathcal{U}_s(\varepsilon)$ ,  $\chi(\xi)$  has degree less than  $2\bar{r}$ , we have for all  $\xi \in \mathcal{U}_s(\varepsilon)$ ,

$$(6.58) \quad \deg(P^{(j,\alpha)}(\xi)) \leq 2\bar{r} + \kappa(2\bar{r} - 2) \leq 2\bar{r}^2.$$

We obtain the expansion

$$(6.59) \quad H_{\text{lo}}^{(j+1,\alpha)}(\xi; u) = Z_2^{(j,\alpha)}(\xi; u) + Z_4^{(j,\alpha)}(\xi; u) + (\text{Id} - \mathfrak{h}^{(\alpha)}(\xi)\Pi_{\Lambda_{\alpha+1}})Q^{(j,\alpha)}(\xi; u) \\ + P^{(j,\alpha)}(\xi; u) + I^{(j,\alpha)}(\xi; u) + R^{(j,\alpha)}(\xi; \Phi_{j,\alpha,\xi}^1(u)).$$

We recall that when  $\xi \in \Xi_\alpha$  then  $\mathfrak{h}^{(\alpha)}(\xi) = 1$ , which removes  $\Pi_{\Lambda_{\alpha+1}}Q^{(j,\alpha)}$  from  $H_{\text{lo}}^{(j+1,\alpha)}$ .

Let us now use the multilinear estimates obtained in Proposition 4.15 and the recurrence assumption to control the norms of the terms  $P^{(j,\alpha)}$  and  $I^{(j,\alpha)}$  generated in the expansions. We start with (6.55). When  $\ell = 1$ , recall that

$$(6.60) \quad \text{ad}_\chi(Z_2^{(j,\alpha)})(\xi; u) = -\mathfrak{h}^{(\alpha)}(\xi) \sum_{\mathbf{n} \in \Lambda_{\alpha+1}} (Q^{(j,\alpha)}[\xi])_{\mathbf{n}} z_{\mathbf{n}}(u, I(u) - \xi).$$

In particular,

$$\|\text{ad}_\chi(Z_2^{(j,\alpha)})\|_{Y_\alpha^{\text{sup}}} \leq \|\Pi_{\Lambda_{\alpha+1}}Q^{(j,\alpha)}\|_{Y_\alpha^{\text{sup}}}.$$

According to the recurrence assumptions (6.41) and (6.37)

$$(6.61) \quad \|\Pi_{\Lambda_{\alpha+1}}Q^{(j,\alpha)}\|_{Y_\alpha^{\text{sup}}} \lesssim \varepsilon^{-1/10^4}, \quad \|\mathbf{1}_{\Xi_\alpha}(\xi)\Pi_{\Lambda_{\alpha+1}}Q^{(j,\alpha)}\|_{Y_\alpha^{\text{sup}}} \leq \varepsilon^{j/30-1/10^4}.$$

We deduce from Corollary 6.11 that for all  $\ell \in \{2, \dots, \kappa\}$ ,

$$(6.62) \quad \|\text{ad}_\chi^\ell(Z_2^{(j,\alpha)})\|_{Y_\alpha^{\text{sup}}} \leq \varepsilon^{3/2} \|\text{ad}_\chi^{\ell-1}(Z_2^{(j,\alpha)})\|_{Y_\alpha^{\text{sup}}} \|\Pi_{\Lambda_{\alpha+1}}Q^{(j,\alpha)}\|_{Y_\alpha^{\text{sup}}} \\ \leq \varepsilon^{3/2(\ell-1)} \|\text{ad}_\chi(Z_2^{(j,\alpha)})\|_{Y_\alpha^{\text{sup}}} \|\Pi_{\Lambda_{\alpha+1}}Q^{(j,\alpha)}\|_{Y_\alpha^{\text{sup}}}^{\ell-1} \\ \lesssim_\ell \varepsilon^{\ell(3/2-1/10^4)-3/2-1/10^4},$$

and, for all  $\ell \in \{2, \dots, \kappa\}$ ,

$$(6.63) \quad \|\mathbf{1}_{\Xi_\alpha}(\xi)\text{ad}_\chi^\ell(Z_2^{(j,\alpha)})\|_{Y_\alpha^{\text{sup}}} \lesssim_\ell \varepsilon^{\ell(3/2-1/10^4+j/30)-3/2-1/10^4}.$$

Similarly, we obtain from Corollary 6.11 that for  $\ell \in \{1, \dots, \kappa\}$ ,

$$(6.64) \quad \|\text{ad}_\chi^\ell(Z_4^{(j,\alpha)})\|_{Y_\alpha^{\text{sup}}} \leq \varepsilon^{1/15+(3/2)(\ell-1)} \|Z_4^{(j,\alpha)}\|_{Z^{\text{sup}}} \|\Pi_{\Lambda_{\alpha+1}}Q^{(j,\alpha)}\|_{Y_\alpha^{\text{sup}}}^\ell \\ \lesssim_r \varepsilon^{\ell(3/2-1/10^4)+1/15-3/2-1/10^4},$$

and for all  $\ell \in \{2, \dots, \kappa\}$ ,

$$(6.65) \quad \|\mathbf{1}_{\Xi_\alpha}(\xi)\text{ad}_\chi^\ell(Z_4^{(j,\alpha)})\|_{Y_\alpha^{\text{sup}}} \lesssim \varepsilon^{\ell(3/2-1/10^4+j/30)+1/15-3/2-1/10^4}.$$

Finally, for all  $\ell \in \{1, \dots, \kappa\}$ ,

$$(6.66) \quad \|\text{ad}_\chi^\ell(Q^{(j,\alpha)})\|_{Y_\alpha^{\text{sup}}} \leq \varepsilon^{3/2\ell} \|\Pi_{\Lambda_{\alpha+1}}Q^{(j,\alpha)}(\xi)\|_{Y_\alpha^{\text{sup}}}^\ell \|Q^{(j,\alpha)}(\xi)\|_{Y_\alpha^{\text{sup}}} \\ \lesssim_r \varepsilon^{\ell(3/2-1/10^4)-1/10^4},$$

and, for all  $\ell \in \{1, \dots, \kappa\}$ ,

$$(6.67) \quad \|\mathbf{1}_{\Xi_\alpha}(\xi)\text{ad}_\chi^\ell(Q^{(j,\alpha)})\|_{Y_\alpha^{\text{sup}}} \leq \varepsilon^{\ell(3/2-1/10^4+j/30)-1/10^4}.$$

Collecting these contributions we deduce that

$$(6.68) \quad \|P^{(j,\alpha)}\|_{Y_\alpha^{\text{sup}}} \lesssim_r \varepsilon^{1/15-2/10^4} \leq \varepsilon^{1/20},$$

and

$$(6.69) \quad \|\mathbf{1}_{\Xi_\alpha}(\xi)P^{(j,\alpha)}\|_{Y_\alpha^{\text{sup}}} \leq \varepsilon^{1/20+j/30}.$$

Corollary 6.11 provides the analog estimates for the  $Y_\alpha^{\text{lip}}$  norms of the different terms (but this time we do not need the extra gain when  $\xi \in \Xi_\alpha$ ). We deduce from the Leibniz rule that

$$\|\text{ad}_\chi(Z_2^{(j,\alpha)})\|_{Y_\alpha^{\text{lip}}} \leq \varepsilon^{-1/15}.$$

Indeed, for  $|k| \leq M$  and  $\mathbf{n} \in \Lambda_{\alpha+1}$ ,

$$|\mathfrak{h}^{(\alpha)}(\xi)\partial_{\xi_k}(Q^{(j,\alpha)}[\xi])_{\mathbf{n}}| \leq \|Q^{(j,\alpha)}\|_{Y_\alpha^{\text{lip}}} w_{\mathbf{n}}^1(\alpha),$$

and, on the other hand, when  $\partial_{\xi_k}$  falls on the truncation function  $\mathfrak{h}^{(\alpha)}$  then we use the bound (6.4.1) (from the proof of the small divisor bound) to obtain

$$\begin{aligned} |\partial_{\xi_k} \mathfrak{h}^{(\alpha)}(\xi)Q_{\mathbf{n}}^{(j,\alpha)}(\xi)| &\leq \varepsilon^{-1/10^4} \gamma(\alpha)^{-1} \varepsilon^{-2} N_{\alpha+1}^{2s} \|Q^{(j,\alpha)}\|_{Y_\alpha^{\text{sup}}} w_{\mathbf{n}}^0(\alpha) \\ &\lesssim \varepsilon^{-2/10^4} \gamma(\alpha)^{-1} (\varepsilon^{-1}\eta)^2 (N_{\alpha+1}/N_\alpha)^{2s} w_{\mathbf{n}}^1(\alpha) \\ &\lesssim_\alpha \varepsilon^{-2/10^4-1/30-1/50-1/100} w_{\mathbf{n}}^1(\alpha) \leq \varepsilon^{-1/15} w_{\mathbf{n}}^1(\alpha). \end{aligned}$$

In particular  $\text{ad}_\chi(Z_2^{(j,\alpha)})$  verifies the assumptions (6.27) and (6.28) of Corollary 6.11, and we deduce by recurrence that for all  $\ell \geq 2$ ,

$$(6.70) \quad \|\text{ad}_\chi^\ell(Z_2^{(j,\alpha)})\|_{Y_\alpha^{\text{lip}}} = \|\text{ad}_\chi^{\ell-1}(\text{ad}_\chi(Z_2^{(j,\alpha)}))\|_{Y_\alpha^{\text{lip}}} \leq \varepsilon N_\alpha^{-2s}.$$

Similarly, using Corollary 6.11 we show that

$$\|\text{ad}_\chi(Z_4^{(j,\alpha)})\|_{Y_\alpha^{\text{lip}}} \leq \varepsilon^{1/50},$$

and therefore, for all  $\ell \in \{2, \dots, \kappa\}$ ,

$$(6.71) \quad \|\text{ad}_\chi^\ell(Z_4^{(j,\alpha)})\|_{Y_\alpha^{\text{lip}}} \leq \varepsilon N_\alpha^{-2s},$$

Finally, we deduce from Corollary 6.11 that for all  $\ell \geq 1$ ,

$$(6.72) \quad \|\text{ad}_\chi^\ell(Q^{(j,\alpha)})\|_{Y_\alpha^{\text{lip}}} \leq \varepsilon N_\alpha^{-2s}.$$

In particular, we proved that

$$(6.73) \quad \|P^{(j,\alpha)}\|_{Y_\alpha^{\text{lip}}} \lesssim \varepsilon^{1/50}.$$

Let us now explain how to organize the new terms generated from the expansions (6.55), (6.56) and (6.57), which are collected in  $P^{(j,\alpha)}$ , and the new remainder terms.

*New terms of type Q.* — We let

$$Q^{(j+1,\alpha)}(u) := (\text{Id} - \mathfrak{h}^{(\alpha)}\Pi_{\Lambda_{\alpha+1}})Q^{(j,\alpha)}(u) + \Pi_{6 \leq \deg \leq \bar{r}} P^{(j,\alpha)}(u).$$

Since  $\mathfrak{h}^{(\alpha)}(\xi) = 1$  for all  $\xi \in \Xi_\alpha$ , we have that for all  $\xi \in \Xi_\alpha$ ,

$$\Pi_{\Lambda_{\alpha+1}} Q^{(j+1,\alpha)}(\xi; u) = \Pi_{\Lambda_{\alpha+1}} \Pi_{6 \leq \deg \leq \bar{r}} P^{(j,\alpha)}(\xi; u),$$

which in turn implies from the bound (6.69) that

$$\|\mathbf{1}_{\Xi_\alpha}(\xi)\Pi_{\Lambda_{\alpha+1}} Q^{(j+1,\alpha)}(\xi)\|_{Y_\alpha^{\text{sup}}} \leq \varepsilon^{1/20+j/30} \leq \varepsilon^{(j+1)/30}.$$

This proves (6.37) at step  $j + 1$ . We now show that  $Q^{(j+1,\alpha)}$  still satisfies (6.41). On the one hand, since  $\mathfrak{h}^{(\alpha)}$  is valued in  $[0, 1]$  we have

$$\|(\text{Id} - \mathfrak{h}^{(\alpha)}(\xi)\Pi_{\Lambda_{\alpha+1}})Q^{(j,\alpha)}\|_{Y_{\alpha}^{\text{sup}}} \leq \|Q^{(j,\alpha)}\|_{Y_{\alpha}^{\text{sup}}},$$

and we conclude from the bound (6.68) that

$$\|Q^{(j+1,\alpha)}\|_{Y_{\alpha}^{\text{sup}}} \leq \|Q^{(j,\alpha)}\|_{Y_{\alpha}^{\text{sup}}} + \|P^{(j,\alpha)}\|_{Y_{\alpha}^{\text{sup}}} \leq \|Q^{(j,\alpha)}\|_{Y_{\alpha}^{\text{sup}}} + \varepsilon^{1/20} \lesssim_{\alpha} \varepsilon^{-1/10^4}.$$

On the other hand, we have from the chain rule and bound that

$$\|(\text{Id} - \mathfrak{h}^{(\alpha)}(\xi)\Pi_{\Lambda_{\alpha+1}})Q^{(j,\alpha)}\|_{Y_{\alpha}^{\text{lip}}} \leq \|Q^{(j,\alpha)}\|_{Y_{\alpha}^{\text{lip}}} + \varepsilon^{-1/15}.$$

We deduce from the bound (6.73) that

$$\begin{aligned} \|Q^{(j+1,\alpha)}\|_{Y_{\alpha}^{\text{lip}}} &\leq \|Q^{(j,\alpha)}\|_{Y_{\alpha}^{\text{lip}}} + \varepsilon^{-1/15} + \|P^{(j,\alpha)}\|_{Y_{\alpha}^{\text{lip}}} \\ &\leq \|Q^{(j,\alpha)}\|_{Y_{\alpha}^{\text{lip}}} + \varepsilon^{-1/15} + \varepsilon^{1/50} \lesssim_{\alpha} \varepsilon^{-1/15}, \end{aligned}$$

which gives the bound (6.41). Finally, to prove (6.42), we observe that the resonance function of the unmodulated frequencies is additive under the Poisson bracket of two monomials. Since there are at most  $\kappa$  Poisson brackets in the new terms, we obtain (6.42).

*New modulated frequencies.* — The polynomial  $P^{(j,\alpha)}$  contains terms of degree 2, corresponding to  $\mathbf{n}'' = \mathbf{e}_m(n)$  for  $n \in \mathbb{Z}_M^d$ :

$$\Pi_{\deg=2}P^{(j,\alpha)}(\xi; u) := \sum_{n \in \mathbb{Z}_M^d} (P^{(j,\alpha)}[\xi])_{\mathbf{e}_m(n)} (|u_n|^2 - \xi_n).$$

These terms modulate the frequencies. For  $n \in \mathbb{Z}_M^d$ , we set

$$\omega_n^{(j+1,\alpha)}(\xi) := \omega_n^{(j,\alpha)}(\xi) + (P^{(j,\alpha)}[\xi])_{\mathbf{e}_m(n)}(j, \xi).$$

By Definition 4.3 and from the bound (6.68), of the  $Y_{\alpha}^{\text{sup}}$ -norm we have, for all  $n \in \mathbb{Z}_M^d$  and  $\xi \in \mathcal{U}_s(\varepsilon)$

$$|(P^{(j,\alpha)}[\xi])_{\mathbf{e}_m(n)}| \leq w_{\mathbf{e}_m(n)}^0(\alpha) \|P^{(j,\alpha)}\|_{Y_{\alpha}^{\text{sup}}} \leq \varepsilon^{1/20} N_{\alpha}^{-6s} \eta^6 D(\alpha) \leq \varepsilon^{3+1/10} N_{\alpha}^{-4s}.$$

Similarly we deduce from the bound (6.73) that for  $(k, n) \in (\mathbb{Z}_M^d)^2$  and  $\xi \in \mathcal{U}_s(\varepsilon)$ ,

$$|\partial_{\xi_k}(P^{(j,\alpha)}[\xi])_{\mathbf{e}_m(n)}| \leq w_{\mathbf{e}_m(n)}^1(\alpha) \|P^{(j,\alpha)}\|_{Y_{\alpha}^{\text{lip}}} \lesssim \varepsilon^{1/50} \eta^4 N_{\alpha}^{-4s} D(\alpha) \leq \varepsilon^{1+1/10} N_{\alpha}^{-2s}.$$

This prove the stability estimates (6.38) and (6.39) for the frequencies.

*New integrable quartic terms.* — The polynomial  $P^{(j,\alpha)}$  also contains terms of degree 4 corresponding to monomials  $\mathbf{e}_m(n_1) + \mathbf{e}_m(n_2)$  for  $n_1, n_2 \in \mathbb{Z}_M^d$ . To estimate these contributions, and especially their  $Y_{\alpha}^{\text{lip}}$ -norms, we need to be slightly more precise. Actually, the terms

$$\text{ad}_X^{\ell}(Z_2^{(j,\alpha)}), \quad \text{ad}_X^{\ell}(Z_4^{(j,\alpha)})$$

generate new integrable terms only when  $\ell \geq 2$ . Therefore, we let

$$(6.74) \quad Z_4^{(j+1,\alpha)} := Z_4^{(j,\alpha)} + \Pi_{\deg=4} \tilde{P}^{(j,\alpha)},$$

where

$$\tilde{P}^{(j,\alpha)} := P^{(j,\alpha)} - \{Z_4^{(j,\alpha)}, \chi\} = \text{ad}_\chi(Q^{(j,\alpha)}) + \sum_{\ell=2}^{\kappa-1} \frac{1}{\ell!} \text{ad}_\chi^\ell(Z_2^{(j,\alpha)} + Z_4^{(j,\alpha)} + Q^{(j,\alpha)}).$$

The reason for introducing  $\tilde{P}^{(j,\alpha)}$  is that the estimates (6.68) and (6.73) are slightly better for  $\tilde{P}(j, \xi)$  than they are for  $P(j, \xi)$  (we gain an important factor  $\varepsilon$  for the  $Y_\alpha^{\text{lip}}$ -norm): indeed, we actually proved that

$$\|\tilde{P}(j, \xi)\|_{Y_\alpha^{\text{sup}}} \leq \varepsilon^{3/2-2/10^4}, \quad \|\tilde{P}(j, \xi)\|_{Y_\alpha^{\text{lip}}} \leq \varepsilon N_\alpha^{-2s}.$$

We write

$$\begin{aligned} \Pi_{\text{deg}=4} \tilde{P}^{(j,\alpha)}(\xi; u) &:= 2 \sum_{n_1, n_2 \in \mathbb{Z}_M^d} (\tilde{P}^{(j,\alpha)}[\xi])_{\mathbf{e}_m(n_1) + \mathbf{e}_m(n_2)} (|u_{n_1}|^2 - \xi_{n_1}) (|u_{n_2}|^2 - \xi_{n_2}) \\ &\quad - \sum_{n \in \mathbb{Z}_M^d} (\tilde{P}^{(j,\alpha)}[\xi])_{\mathbf{e}_m(n) + \mathbf{e}_m(n)} (|u_n|^2 - \xi_n)^2. \end{aligned}$$

For all  $\xi \in \mathcal{U}_s(\varepsilon)$  and  $(n_1, n_2) \in (\mathbb{Z}_M^d)^2$ , we have from the definition of the  $Y_\alpha^{\text{sup}}$ -norm that

$$\begin{aligned} |(\tilde{P}^{(j,\alpha)}[\xi])_{\mathbf{e}_m(n_1) + \mathbf{e}_m(n_2)}| &\leq w_{\mathbf{e}_m(n_1) + \mathbf{e}_m(n_2)}^0(\alpha) \|\tilde{P}^{(j,\alpha)}\|_{Y_\alpha^{\text{sup}}} \\ &\leq \varepsilon^{3/2-2/10^4} N_\alpha^{-6s} \eta^6 \mathbf{D}(\alpha)^2 \leq \varepsilon^{3/2-2/5-2/10^4} N_\alpha^{-2s} \eta^2. \end{aligned}$$

Similarly, for all  $k \in \mathbb{Z}_M^d$ ,

$$\begin{aligned} |\partial_{\xi_k}(\tilde{P}^{(j,\alpha)}[\xi])_{\mathbf{e}_m(n_1) + \mathbf{e}_m(n_2)}| &\leq w_{\mathbf{e}_m(n_1) + \mathbf{e}_m(n_2)}^1(\alpha) \|\tilde{P}^{(j,\alpha)}\|_{Y_\alpha^{\text{lip}}} \\ &\leq \varepsilon N_\alpha^{-2s} N_\alpha^{-4s} \eta^4 \mathbf{D}(\alpha)^2 \leq \varepsilon^{1-2/5} N_\alpha^{-2s}. \end{aligned}$$

This shows that for all  $\xi \in \mathcal{U}_s(\varepsilon)$ ,

$$\|\Pi_{\text{deg}=4} \tilde{P}^{(j,\alpha)}\|_{Y_\alpha^{\text{sup}}} \leq \varepsilon^{3/2} N_\alpha^{-2s}, \quad \|\Pi_{\text{deg}=4} \tilde{P}^{(j,\alpha)}\|_{Y_\alpha^{\text{lip}}} \leq \varepsilon^{1-2/5} N_\alpha^{-2s},$$

and therefore we have from the expression (6.74) of  $Z_4^{(j+1,\alpha)}$  that under the recurrence assumption (6.40)

$$\|Z_4^{(j+1,\alpha)}\|_{Z^{\text{sup}}} \leq \|Z_4^{(j,\alpha)}\|_{Z^{\text{sup}}} + \varepsilon^{3/2} N_\alpha^{-2s} \lesssim_\alpha \varepsilon^{-1/10^4},$$

and

$$\|Z_4^{(j+1,\alpha)}\|_{Z^{\text{lip}}} \leq \|Z_4^{(j,\alpha)}\|_{Z^{\text{lip}}} + \varepsilon^{1-2/5} N_\alpha^{-2s} \lesssim_\alpha \varepsilon^{-1/10^4}.$$

*New remainder terms.* — Four types of terms contribute to the remainder: the former remainder in the new coordinates, the terms with degree larger than  $\bar{r}$ , the integral remainders in the Taylor expansions and the constants. We define

$$\begin{aligned} R^{(j+1,\alpha)}(\xi; u) &:= R^{(j,\alpha)}(\xi; \Phi_{j,\alpha,\xi}^1(u)) + \Pi_{\text{deg} > \bar{r}} P^{(j,\alpha)}(\xi; u) \\ &\quad + I^{(j,\alpha)}(\xi; u) + \Pi_{\text{deg}=0} P^{(j,\alpha)}(\xi; u), \end{aligned}$$

and we justify that these terms are actually some remainders in the sense of the vector field estimate (6.43), when  $\xi \in \Xi_\alpha$  and  $u \in \mathcal{V}_{\alpha,s}(\varepsilon_{j+1,\alpha}, \xi)$ , with

$$\varepsilon_{j,\alpha} = 10\varepsilon - (\alpha\kappa + j)\varepsilon^{3/2}.$$

Fix  $\xi \in \Xi_\alpha$  and  $u \in \mathcal{V}_{\alpha,s}(\varepsilon_{j+1,\alpha}, \xi)$ . First, we note that, by definition,

$$\nabla \Pi_{\deg=0} P^{(j,\alpha)}(\xi; u) = 0,$$

and so that this term can be ignored. Then we deduce from the estimate on the symplectic transformation (6.51) that for all  $t \in [-1, 1]$ ,

$$v(t) := \Phi_{j,\alpha,\xi}^t(u) \in \mathcal{V}_{\alpha,s}(\varepsilon_{j,\alpha}, \xi).$$

Using that  $\Phi_{j,\alpha,\xi}^1$  is symplectic, we obtain

$$\begin{aligned} \nabla R^{(j,\alpha)}(\xi; \Phi_{j,\alpha,\xi}^1(u)) &= (d\Phi_{j,\alpha,\xi}^1(u))^* \nabla R^{(j,\alpha)}(\xi; v(1)) \\ &= -i(d\Phi_{j,\alpha,\xi}^1(u))^{-1} i \nabla R^{(j,\alpha)}(\xi; v(1)) \\ &= -id\Phi_{j,\alpha,\xi}^{-1}(v) i \nabla R^{(j,\alpha)}(\xi; v(1)). \end{aligned}$$

We deduce from the estimate on the symplectic transformation (6.52) and from the vector field estimate (6.43) at step  $j$  on  $\mathcal{V}_{\alpha,s}(\varepsilon_{j,\alpha}, \xi)$  that for all  $\xi \in \mathcal{U}_s(\varepsilon)$ ,

$$\begin{aligned} \|\nabla R^{(j,\alpha)}(\xi; \Phi_{j,\alpha,\xi}^1(u))\|_{h^s} &\leq \|d\Phi_{j,\alpha,\xi}^{-1}(v(1))\|_{h^s \rightarrow h^s} \|\nabla R^{(j,\alpha)}(\xi; v(1))\|_{h^s} \\ &\lesssim_{r,\alpha} (1 + \varepsilon^{1/2}) \varepsilon^{2r+1}. \end{aligned}$$

Similarly, given  $H \in X_{\leq 2\bar{r}}(\varepsilon)$  we have that for all  $t \in [0, 1]$ ,

$$\nabla \text{ad}_\chi^\ell(H)(\xi; \Phi_{j,\alpha,\xi}^t(u)) = -id\Phi_{j,\alpha,\xi}^{-t}(v) i \nabla \text{ad}_\chi^\ell(H)(\xi; v(t)).$$

The vector field estimate then follows from Proposition 4.6 and from the bounds (6.62), (6.64) and (6.66): for  $\xi \in \mathcal{U}_s(\varepsilon)$  and  $u \in \mathcal{V}_{\alpha,s}(\varepsilon_{j+1,\alpha}, \xi)$ ,

$$\begin{aligned} \|\nabla I^{(j,\alpha)}(\xi; u)\|_{h^s} &\lesssim \sup_{t \in [0,1]} \|d\Phi_{j,\alpha,\xi}^{-t}(v(t))\|_{h^s \rightarrow h^s} \|\nabla \text{ad}_\chi^\kappa(Z_2^{(j,\alpha)} + Z_4^{(j,\alpha)} + Q^{(j,\alpha)})(\xi; v(t))\|_{h^s} \\ &\lesssim \|\text{ad}_\chi^\kappa(Z_2^{(j,\alpha)} + Z_4^{(j,\alpha)} + Q^{(j,\alpha)})\|_{Y_\alpha^{\text{sup}}} N_\alpha^{-4s} \varepsilon^3 \sup_{t \in [0,1]} \|v(t)\|_{h^s} \leq \varepsilon^{3+\kappa/20} \|u\|_{h^s}, \end{aligned}$$

which is conclusive provided  $\kappa \geq 40r$ . As for the terms with large degree, we exploit Lemma 4.10 and the assumption  $\bar{r}\tau \geq 6s$ . We get

$$\|\Pi_{\deg > \bar{r}} P^{(j,\alpha)}\|_{Y_\beta^{\text{sup}}} \leq \|\Pi_{\deg > \bar{r}} P^{(j,\alpha)}\|_{Y_\alpha^{\text{sup}}} \leq \varepsilon^{1/20},$$

where we used the bound obtained in (6.68). On the other, since  $\deg(P^{(j,\alpha)}) \leq \bar{r}^2$ , we deduce from the vector field estimate at scale  $\beta$  of Proposition 4.6 that for all  $u \in \mathcal{V}_{\alpha,s}(10\varepsilon, \xi)$ ,

$$\|\nabla \Pi_{\deg > \bar{r}} P^{(j,\alpha)}(\xi; u)\|_{h^s} \leq \|\Pi_{\deg > \bar{r}} P^{(j,\alpha)}\|_{Y_\beta^{\text{sup}}} N_\beta^{-4s} \varepsilon^3 \|u\|_{h^s} \leq \varepsilon^{(1/20)+3} N_\beta^{-4s} \|u\|_{h^s},$$

and we conclude that under our choice of parameter (4.1) (at scale  $\beta$ ,  $N_\beta^{-2s} = \varepsilon^r$ )

$$\|\nabla \Pi_{\deg > \bar{r}} P^{(j,\alpha)}(\xi; u)\|_{h^s} \leq \varepsilon^{2r+3} \|u\|_{h^s}.$$

This concludes the proof of (6.43) at step  $j + 1$  and completes the proof of Proposition 6.19.  $\square$

6.7. FROM STEP  $\alpha$  TO STEP  $\alpha + 1$ : — We now have all the ingredients to deduce Theorem 6.2 at scale  $\alpha + 1$  from Theorem 6.2 at scale  $\alpha$ .

*Proof of Theorem 6.2.* — The initialization  $\alpha = 0$  was done in Section 6.1, with  $R^{(0)} = \Pi_{\deg=0}Q$ . We fix  $\alpha \in \{0, \dots, \beta - 1\}$  and we suppose that Theorem 6.2 is proved up to scale  $\alpha$ . The inputs are the transformation  $\tau_\alpha$  and the Hamiltonian functions

$$Z_2^{(\alpha)}, \quad Z_4^{(\alpha)}, \quad Q^{(\alpha)}, \quad R^{(\alpha)},$$

where  $Z_2^{(\alpha)}$  has modulated frequencies  $\omega^{(\alpha)}$ . We constructed the new objects at scale  $\alpha + 1$  in Proposition 6.19: for  $\varepsilon \in (0, 10^{-2}\varepsilon_*)$  and  $\xi \in \mathcal{U}_s(\varepsilon)$ , we set

$$\tau_{\alpha+1,\xi} := \tau_{\kappa,\alpha,\xi}, \quad Z_2^{(\alpha+1)} := Z_2^{(\kappa,\alpha)}, \quad Z_4^{(\alpha+1)} := Z_4^{(\kappa,\alpha)}.$$

We also set

$$Q^{(\alpha+1)} := (\text{Id} - \Pi_{\Lambda_{\alpha+1}})Q^{(\kappa,\alpha)}, \quad R^{(\alpha+1)} := \Pi_{\Lambda_{\alpha+1}}Q^{(\kappa,\alpha)} + R^{(\kappa,\alpha)}.$$

By construction

$$H_{\text{lo}} \circ \tau_{\alpha+1,\xi}(\xi; u) = Z_2^{(\alpha+1)}(\xi; u) + Z_4^{(\alpha+1)}(\xi; u) + Q^{(\alpha+1)}(\xi; u) + R^{(\alpha+1)}(\xi; u).$$

Then, we make sure that the new objects satisfy the properties claimed in Theorem 6.2 at scale  $\alpha + 1$ . First, note that

$$\varepsilon_{\kappa,\alpha} := 10\varepsilon - (\alpha r + \kappa)\varepsilon^{3/2} \geq 10\varepsilon - cr(\alpha + 1)\varepsilon^{3/2} = \varepsilon_{\alpha+1},$$

provided that  $c$  is large enough to ensure that  $\kappa \leq cr$  ( $c \geq 40$  is acceptable). Hence,  $\mathcal{V}_{\alpha+1,s}(\varepsilon_{\alpha+1}, \xi) \subset \mathcal{V}_{\alpha,s}(\varepsilon_{\kappa,\alpha}, \xi)$ .

*The new variables.* — We deduce from the above observation and from the estimates on the transformation (6.44), (6.45) and (6.46) proved at step  $\kappa$ , that the estimates (6.12), (6.13) and (6.14) (claimed in the normal form Theorem 6.2) hold at scale  $\alpha + 1$ .

*Modulated frequencies at scale  $\alpha + 1$ .* — We deduce the stability property from the bound (6.38): for all  $n \in \mathbb{Z}_M^d$ ,

$$|\omega_n^{(\alpha+1)}(\xi) - \omega_n^{(\alpha)}(\xi)| \leq \sum_{j=0}^{\kappa-1} |\omega_n^{(j+1,\alpha)}(\xi) - \omega_n^{(j,\alpha)}(\xi)| \leq \kappa N_\alpha^{-4s} \varepsilon^{3+1/10} \leq N_\alpha^{-4s} \varepsilon^3.$$

This proves (6.7) at scale  $\alpha + 1$ . Similarly, we deduce from (6.39) and (6.6) at scale  $\alpha$  that for all  $k \in \mathbb{Z}_M^d$ ,

$$\begin{aligned} |\partial_{\xi_k}(\omega_n^{(\alpha+1)}(\xi) - \xi_n)| &\leq |\partial_{\xi_k}(\omega_n^{(\alpha)}(\xi) - \xi_n)| + \sum_{j=0}^{\kappa-1} |\partial_{\xi_k}(\omega_n^{(j+1,\alpha)}(\xi) - \omega_n^{(j,\alpha)}(\xi))| \\ &\leq \alpha\varepsilon + \kappa\varepsilon^{1+1/10} N_\alpha^{-2s} \leq (\alpha + 1)\varepsilon. \end{aligned}$$

This proves (6.6) at scale  $\alpha + 1$ .

*Quartic integrable terms at scale  $\alpha + 1$ .* — According to (6.40),

$$\begin{aligned} \max(\|Z_4^{(\alpha+1)}\|_{Z^{\text{sup}}}, \|Z_4^{(\alpha+1)}\|_{Z^{\text{lip}}}) &= \max(\|Z_4^{(\kappa, \alpha)}\|_{Z^{\text{sup}}}, \|Z_4^{(\kappa, \alpha)}\|_{Z^{\text{lip}}}) \\ &\lesssim_{\alpha} \varepsilon^{-1/10^4} + \alpha + \kappa \varepsilon^{1/2} \lesssim_{\alpha} \varepsilon^{-1/10^4}, \end{aligned}$$

and this proves (6.8) at scale  $\alpha + 1$ .

*Terms of type  $Q$  at scale  $\alpha + 1$ .* — We now verify that  $Q^{(\alpha+1)}$  operates at scale  $\alpha + 1$ : we have from (6.41) that

$$\|Q^{(\kappa, \alpha)}\|_{Y_{\alpha}^{\text{sup}}} \lesssim_{\alpha} \varepsilon^{-1/10^4}, \quad \|Q^{(\kappa, \alpha)}\|_{Y_{\alpha}^{\text{lip}}} \lesssim \varepsilon^{-1/20}.$$

According to Lemma 4.8,

$$\|Q^{(\alpha+1)}\|_{Y_{\alpha+1}^{\text{sup}}} = \|(\text{Id} - \Pi_{\Lambda_{\alpha+1}})Q^{(\kappa, \alpha)}\|_{Y_{\alpha+1}^{\text{sup}}} \leq \|(\text{Id} - \Pi_{\Lambda_{\alpha+1}})Q^{(\kappa, \alpha)}\|_{Y_{\alpha}^{\text{sup}}} \lesssim_{\alpha} \varepsilon^{-1/10^4},$$

and, similarly,

$$\|Q^{(\alpha+1)}\|_{Y_{\alpha+1}^{\text{lip}}} = \|(\text{Id} - \Pi_{\Lambda_{\alpha+1}})Q^{(\kappa, \alpha)}\|_{Y_{\alpha+1}^{\text{lip}}} \leq \|(\text{Id} - \Pi_{\Lambda_{\alpha+1}})Q^{(\kappa, \alpha)}\|_{Y_{\alpha}^{\text{lip}}} \lesssim_{\alpha} \varepsilon^{-1/20}.$$

This proves (6.9) at scale  $\alpha + 1$ .

*Remainder term at scale  $\alpha + 1$ .* — As for the remainder term, when  $\xi \in \Xi_{\alpha}$  then  $\xi \in \Xi_{\alpha-1}$  and we deduce from (6.43) at step  $\kappa$  that for all  $u \in \mathcal{V}_{\alpha+1, s}(\varepsilon_{\alpha+1}, \xi)$ ,

$$\|\nabla R^{(\kappa, \alpha)}(\xi; u)\|_{h^s} \lesssim_{\alpha, r} \varepsilon^{2r+1} \|u\|_{h^s}.$$

Moreover, we have from (6.37) that

$$\|\Pi_{\Lambda_{\alpha+1}} Q^{(\kappa, \alpha)}\|_{Y_{\alpha}^{\text{sup}}} \leq \varepsilon^{\kappa/30-1/10^4} \leq \varepsilon^{2r},$$

provided, say, that  $\kappa \geq 100r$ . We conclude from Proposition 4.6 that for all  $u \in \mathcal{V}_{\alpha+1, s}(\varepsilon_{\alpha+1}, \xi)$ ,

$$\|\nabla \Pi_{\Lambda_{\alpha+1}} Q^{(\kappa, \alpha)}(\xi; u)\|_{h^s} \leq \varepsilon^{2r+1} \|u\|_{h^s}.$$

This proves (6.11) at scale  $\alpha + 1$  and completes the proof of Theorem 6.2.  $\square$

## 7. PROOF OF THE MAIN THEOREM

In this section we apply the normal form Theorem 6.2 at scale  $\beta$  together with the measure estimates from Section 5 to prove Theorem 2.7. Then, the finite dimensional reduction strategy exposed in Section 2 completes the proof of the main Theorem 1.3.

Let us start with the construction of the non-resonant set of initial data  $\Theta_{\varepsilon}^b$  and its measure estimate. We finally handle the dynamics of the low-frequency actions in Section 7.2.

7.1. THE GOOD MODULATION DATA SET, AND ITS MEASURE ESTIMATE. — Let  $\tau_\xi := \tau_{\beta,\xi}$  be the parameter-dependent symplectic transformations obtained in Theorem 6.2 after  $\beta$  steps. We set

$$(7.1) \quad \begin{aligned} \Psi: \Pi_M B_s(9\varepsilon) &\longrightarrow h^s(\mathbb{Z}_M^d) \\ \phi &\longmapsto \tau_{\beta,\xi(\phi)}(\phi). \end{aligned}$$

Let us state some properties of  $\Psi$  that we essentially deduce from the properties of  $\tau_{\xi,\beta}$  claimed in Theorem 6.2.

LEMMA 7.1. — *The function  $\Psi$  is a  $C^1$ -diffeomorphism from  $\Pi_M B_s(9\varepsilon)$  into its image which is close to the identity: for all  $\phi \in \Pi_M B_s(9\varepsilon)$ ,*

$$(7.2) \quad \|\Psi(\phi) - \phi\|_{h^s} \lesssim_r \varepsilon^{3/4} \|\phi\|_{h^s},$$

$$(7.3) \quad \|\mathrm{d}\Psi(\phi) - \mathrm{Id}\|_{h^s \rightarrow h^s} \leq \varepsilon^{1/2}.$$

*Proof.* — Observe that  $9\varepsilon \leq \varepsilon_\beta$  and, by construction, that  $\phi \in \mathcal{V}_{\beta,s}(\xi(\phi), \varepsilon_\beta)$ . Hence, the bound (7.2) corresponds to the bound (6.12). Applying the chain rule, we obtain that for all  $\phi \in \Pi_M B_s(9\varepsilon)$ ,

$$\mathrm{d}\Psi(\phi) = \mathrm{d}_\phi \tau_{\beta,\xi(\phi)}(u) \big|_{u=\phi} + \mathrm{d}\tau_{\beta,\xi(\phi)}(\phi).$$

Hence, (7.3) is a consequence of (6.14) at scale  $\beta$ , with  $u = \phi \in \mathcal{V}_s(\varepsilon_\beta, \xi(\phi))$ , and of (6.13).

It follows from (7.3) is that  $\Psi$  is a local diffeomorphism. According to the mean value theorem applied on the convex set  $\Pi_M B_s(9\varepsilon)$ ,  $\Psi - \mathrm{Id}$  is  $\varepsilon^{1/2}$ -Lipschitz, which in turn implies that  $\Psi$  is an injective function on  $\Pi_M B_s(9\varepsilon)$ . This proves that  $\Psi$  is a  $C^1$ -diffeomorphism on  $\Pi_M B_s(9\varepsilon)$ .  $\square$

Note that another consequence of (7.2) is that

$$\Psi(\Pi_M B_s(9\varepsilon)) \subset \Pi_M B_s(10\varepsilon).$$

The next lemma is useful to estimate the measure of sets transported by the function  $\Psi$ .

LEMMA 7.2. — *For all  $\phi \in \Pi_M B_s(9\varepsilon)$ ,*

$$(7.4) \quad \left| |\det(\mathrm{d}\Psi(\phi))| - 1 \right| \leq \varepsilon^{1/20}.$$

*Proof.* — Seeing  $\Pi_M B_s$  as an  $\mathbb{R}$ -vector space of dimension  $d_M := 2\sharp(\mathbb{Z}_M^d)$ , we deduce from (7.3) that

$$(1 - \varepsilon^{1/2})^{d_M} \leq |\det(\mathrm{d}\Psi(\phi))| \leq (1 + \varepsilon^{1/2})^{d_M}.$$

It remains to prove that

$$(7.5) \quad |(1 \pm \varepsilon^{1/2})^{d_M} - 1| \leq \varepsilon^{1/20}.$$

We have

$$(7.6) \quad |(1 \pm \varepsilon^{1/2})^{d_M} - 1| = |\exp(d_M \log(1 \pm \varepsilon^{1/2})) - 1| \leq 2d_M |\log(1 \pm \varepsilon^{1/2})|.$$

Recall that

$$d_M \leq 2\varepsilon^{-1/10^4}.$$

Using that for all  $x \in [0, 1)$ ,

$$|\log(1 - x)| \leq \frac{x}{1 - x}, \quad \log(1 + x) \leq x,$$

we obtain

$$d_M |\log(1 \pm \varepsilon^{1/2})| \leq 2\varepsilon^{1/2-1/10^4} \leq \frac{1}{2}\varepsilon^{1/20}.$$

We conclude from (7.6) that

$$|(1 \pm \varepsilon^{1/2})^{d_M} - 1| \leq \varepsilon^{1/20},$$

which gives (7.5) and completes the proof of (7.4).  $\square$

**DEFINITION 7.3** (Non-resonant initial data). — Let  $\omega^{(\beta-1)}$  be the modulated frequencies at step  $\beta-1$  of the normal form iteration scheme. The set of non-resonant internal parameters is  $\Xi_{\beta-1}$  as defined in (6.5) and the set of non-resonant data in final coordinates is

$$(7.7) \quad U_\beta = U_{\varepsilon, \gamma(\beta)}(\omega^{(\beta-1)}, \Lambda_\beta) = \{\phi \in \Pi_M B_s(\varepsilon) \mid \xi(\phi) \in \Xi_{\beta-1}\}.$$

We first deduce from Proposition 5.4 an estimate on the density of  $U_\beta$  in  $\Pi_M B_s(\varepsilon)$ .

**LEMMA 7.4.** — *For all  $\varepsilon \in (0, \varepsilon_*)$  we have*

$$(7.8) \quad \text{meas}(U_\beta) \geq (1 - \varepsilon^{1/38}) \text{meas}(\Pi_M B_s(\varepsilon)).$$

*Proof.* — For all  $\xi \in \mathcal{U}_s(\varepsilon)$ , the modulated frequencies  $(\omega_n^{(\beta-1)}(\xi))_n$  satisfy (6.6) at scale  $\beta-1$ . Hence, they satisfy the assumption of Proposition 5.4, and we obtain that the associated non-resonant set of data  $U_\beta$  has large density in  $\Pi_M B_s(\varepsilon)$ :

$$\text{meas}(\Pi_M B_s(\varepsilon) \setminus U_\beta) \leq \gamma(\beta) \varepsilon^{-1/10^4} \text{meas}(\Pi_M B_s(\varepsilon)).$$

Under our choice of parameters (in particular  $\gamma(\alpha) = 4^\alpha \gamma = 4^\alpha \varepsilon^{1/30}$ ) we deduce that

$$\text{meas}(\Pi_M B_s(\varepsilon) \setminus U_\beta) \leq \varepsilon^{1/38} \text{meas}(\Pi_M B_s(\varepsilon)).$$

This gives (7.8) and concludes the proof of the lemma.  $\square$

We are now ready to estimate from below the density of the set of non-resonant data in the coordinates obtained after the preliminary transformation  $\tau_0$  of Section 2.

**PROPOSITION 7.5.** — *For all  $\varepsilon \in (0, \varepsilon_*)$ , let*

$$(7.9) \quad \Theta_\varepsilon^b := \Psi(U_\beta) \cap \Pi_M B_s(\varepsilon).$$

*Then,  $\Theta_\varepsilon^b$  is an open subset of  $\Pi_M B_s(\varepsilon)$ , with*

$$(7.10) \quad \text{meas}(\Theta_\varepsilon^b) \geq (1 - \varepsilon^{1/39}) \text{meas}(\Pi_M B_s(\varepsilon)).$$

*Proof.* — Since  $U_\beta$  is also an open subset of  $\Pi_M B_s(\varepsilon)$  and  $\Psi$  is bi-Lipschitz,  $\Theta_\varepsilon^b$  is an open subset of  $\Pi_M B_s(\varepsilon)$ .

As for the measure estimate, we estimated the measure of  $U_\beta$  in Lemma 7.4 and it essentially remains to estimate its image by  $\Psi$ , intersected with  $\Pi_M B(\varepsilon)$ . Note that, according to (7.2), we have

$$\Psi(\Pi_M B_s(\varepsilon)) \subset \Pi_M B_s(\varepsilon + C(r)\varepsilon^{7/4}) \subset \Pi_M B_s(\varepsilon + \varepsilon^{3/2}),$$

provided  $\varepsilon_*(r)$  is small enough, and therefore, defining

$$\mathcal{A}_\varepsilon = \{u \in \Pi_M h^s \mid \varepsilon \leq \|u\|_{h^s} \leq \varepsilon + \varepsilon^{3/2}\}.$$

we obtain that

$$\Psi(U_\beta) = (\Psi(U_\beta) \cap \Pi_M B_s(\varepsilon)) \sqcup (\Psi(U_\beta) \cap \mathcal{A}_\varepsilon) = \Theta_\varepsilon^b \sqcup (\Psi(U_\beta) \cap \mathcal{A}_\varepsilon),$$

where the second equality follows from the definition of the set  $\Theta_\varepsilon^b$ . We deduce that

$$(7.11) \quad \text{meas}(\Theta_\varepsilon^b) \geq \text{meas}(\Psi(U_\beta)) - \text{meas}(\mathcal{A}_\varepsilon).$$

Then, we do a change of variable

$$\begin{aligned} \text{meas}(\Psi(U_\beta)) &= \int_{\Psi(U_\beta)} du = \int_{U_\beta} |\det(d\Psi(\phi))| d\phi \\ &\geq \text{meas}(U_\beta) \min_{\phi \in \Pi_M B_s(\varepsilon)} |\det(d\Psi(\phi))|, \end{aligned}$$

and we deduce from (7.4) and (7.8) that

$$\text{meas}(\Psi(U_\beta)) \geq (1 - \varepsilon^{1/38})(1 - \varepsilon^{1/20}) \text{meas}(\Pi_M B_s(\varepsilon)).$$

On the other hand, we deduce from the bound (7.5) that, by homogeneity,

$$\frac{\text{meas}(\mathcal{A}_\varepsilon)}{\text{meas}(\Pi_M(\varepsilon))} = (1 + \varepsilon^{1/2})^{d_M} - 1 \leq \varepsilon^{1/20}.$$

We conclude from (7.11) that

$$\text{meas}(\Theta_\varepsilon^b) \geq (1 - \tfrac{1}{2}\varepsilon^{1/39} - \varepsilon^{1/20}) \text{meas}(\Pi_M B_s(\varepsilon)) \geq (1 - \varepsilon^{1/39}) \text{meas}(\Pi_M B_s(\varepsilon)).$$

This completes the proof of Proposition 7.5.  $\square$

**7.2. DYNAMICS.** — It remains to prove the dynamical part of Theorem 2.7, which is precisely the stability of the low frequency actions for initial data in the non-resonant set  $\Theta_\varepsilon^b$  defined in (7.9).

*Proof of Theorem 2.7.* — Let  $u(0) \in \Pi_M B_s(\varepsilon)$ , with  $\|u(0)\|_{h^s} \leq \rho \leq \varepsilon$ . We suppose that  $0 < T \leq \varepsilon^{-r}$  and  $u \in C^1([-T, T], \Pi_M h^s)$  is such that

$$i\partial_t u = \nabla H_{10}(u) + f(t),$$

with  $u(t=0) = u(0)$  and, for all  $|t| \leq T$ ,

$$\|f(t)\|_{h^s} \leq \varepsilon^{3r} \rho.$$

The Hamiltonian  $H_{10}$  is given by (2.3). We constructed in Theorem 6.2 for each  $\xi \in \mathcal{U}_s(\varepsilon)$  a symplectic transformation  $\tau_{\beta, \xi}$ , and we defined the bi-Lipschitz function  $\Psi$  in (7.1). The set of non-resonant parameters  $\Theta_\varepsilon^b \subset \Pi_M B_s(\varepsilon)$  was introduced in (7.9).

We proved in Proposition 7.5 the density estimate (2.4) of  $\Theta_\varepsilon^b$  in  $\Pi_M B_s(\varepsilon)$ , and we now suppose that  $u(0) \in \Theta_\varepsilon^b$ . By construction of the set  $\Theta_\varepsilon^b$  (see (7.9)) one can find  $\phi \in U_\beta$  such that

$$u(0) = \Psi(\phi) = \tau_{\beta, \xi(\phi)}(\phi).$$

Then, we set for  $|t| \leq T$

$$\xi := \xi(\phi) = (|\phi_n|^2)_{|n| \leq M}, \quad v(t) := \tau_{\beta, \xi}^{-1}(u(t)).$$

In particular,  $v$  is solution in  $C([-T, T], \Pi_M h^s)$  to the Cauchy problem

$$(7.12) \quad \begin{cases} i\partial_t v = \nabla H^{(\beta)}(\xi; v(t)) - i(d\tau_{\beta, \xi}(v(t)))^{-1}(if(t)), \\ v(0) = \phi, \end{cases}$$

where

$$H^{(\beta)}(\xi; u) := H_{\text{lo}}^{(\beta)}(\xi; \tau_{\beta, \xi}(v)) = Z_2^{(\beta)}(\xi; u) + Z_4^{(\beta)}(\xi; u) + Q^{(\beta)}(\xi; u) + R^{(\beta)}(\xi; u),$$

as in Theorem 6.2. Note that we used

$$d\tau_{\beta, \xi}^{-1}(u(t)) = (d\tau_{\beta, \xi}(v(t)))^{-1}.$$

We now perform a bootstrap argument to show that for all  $|t| \leq T \leq \varepsilon^{-r}$ ,

$$(7.13) \quad \sum_{|n| \leq M} \langle n \rangle^{2s} ||v_n(t)|^2 - \xi_n| \leq N_\beta^{-2s} \varepsilon^{1/2} \rho^2.$$

This in particular implies that

$$(7.14) \quad v(t) \in \mathcal{V}_{\beta, s}(2\varepsilon, \xi), \quad \|v(t)\|_{h^s} \leq 2\rho,$$

and, according to the bound (6.12) on the symplectic transformation  $\tau_{\beta, \xi}$ , we deduce that

$$\begin{aligned} & \sum_{|n| \leq M} \langle n \rangle^{2s} ||u_n(t)|^2 - |u_n(0)|^2| \\ & \leq \sum_{|n| \leq M} \langle n \rangle^{2s} ||u_n(t)|^2 - |v_n(t)|^2| + \sum_{|n| \leq M} \langle n \rangle^{2s} ||v_n(t)|^2 - |v_n(0)|^2| \\ & \quad + \sum_{|n| \leq M} \langle n \rangle^{2s} ||v_n(0)|^2 - |u_n(0)|^2| \\ & \lesssim_r \varepsilon^{3/2} (\|v(t)\|_{h^s}^2 + \|\phi\|_{h^s}^2) + N_\beta^{-2s} \varepsilon^{1/2} \rho^2 \leq \frac{\varepsilon}{2} \rho^2, \end{aligned}$$

uniformly in  $|t| \leq T$ . This is precisely the desired bound (2.5) for the dynamics of  $u(t)$ .

It remains to prove (7.13). At time  $t = 0$ , since  $v(0) = \phi$  and  $\xi_n = |\phi_n|^2$  for all  $|n| \leq M$  the left-and-side of (7.13) vanishes. For  $|t| \leq T$ , we deduce from equation (7.12) a priori bounds for  $v$  under assumption (7.13), from which (7.13) follows by a standard bootstrap argument.

Since, by assumption,  $\phi \in U_{\beta-1}$  then  $\xi(\phi) \in \Xi_{\beta-1}$  and, assuming that  $v(t) \in \mathcal{V}_{\beta, s}(2\varepsilon, \xi)$ , we obtain

$$\|\nabla R^{(\beta)}(\xi, v(t))\|_{h^s} \lesssim_r \varepsilon^{2r+1} \|v(t)\|_{h^s}.$$

where we used the bound (6.11) for the remainder  $R$ . In addition, when  $v(t) \in \mathcal{V}_{\beta,s}(2\varepsilon, \xi)$ , we obtain from Proposition 4.6 and the bound (6.9) at scale  $\beta$  that

$$(7.15) \quad \|\nabla Q^{(\beta)}(\xi; v(t))\|_{h^s} \leq \|Q^{(\beta)}\|_{Y_\beta^{\text{sup}}} N_\beta^{-4s} \varepsilon^3 \|v(t)\|_{h^s} \lesssim N_\beta^{-4s} \varepsilon^{3-1/10^4} \rho.$$

Using the equation (7.12) we obtain that for all  $|n| \leq M$ ,

$$\begin{aligned} \frac{d}{dt} |v_n(t)|^2 &= 2 \operatorname{Im} \left( \partial_{\overline{v_n}} (Z_2^{(\beta)} + Z_4^{(\beta)} + Q^{(\beta)} + R^{(\beta)}) (\xi; v(t)) \overline{v_n(t)} \right) \\ &\quad + 2 \operatorname{Im} \left( i \left( (d\tau_{\beta,\xi}(v(t)))^{-1} (if(t)) \right)_n \overline{v_n(t)} \right). \end{aligned}$$

Using that  $Z_2^{(\beta)}, Z_4^{(\beta)} \in X_{\text{Int}}(\varepsilon)$  is integrable (and, consequently, does not contribute to the dynamics of the actions), we deduce that

$$\frac{d}{dt} |v_n(t)|^2 = 2 \operatorname{Im} \left( \partial_{\overline{v_n}} (Q^{(\beta)} + R^{(\beta)}) (\xi; v(t)) \overline{v_n} \right) + 2 \operatorname{Im} \left( i \left( (d\tau_{\beta,\xi}(v(t)))^{-1} (if(t)) \right)_n \overline{v_n} \right).$$

We deduce from the above an energy estimate that

$$\begin{aligned} \sum_{|n| \leq M} \langle n \rangle^{2s} \left| \frac{d}{dt} |v_n(t)|^2 \right| &\leq \|v(t)\|_{h^s} \\ &\quad \times \left( \|\nabla Q^{(\beta)}(\xi; v(t))\|_{h^s} + \|\nabla R^{(\beta)}(\xi; v(t))\|_{h^s} + \|(d\tau_{\beta,\xi}(v(t)))^{-1}\|_{h^s \rightarrow h^s} \|f(t)\|_{h^s} \right). \end{aligned}$$

Under the bootstrap assumption (7.13) and its consequences (7.14), we have therefore

$$\begin{aligned} \sum_{|n| \leq M} \langle n \rangle^{2s} \left| \frac{d}{dt} |v_n(t)|^2 \right| &\lesssim_r \rho^2 (N_\beta^{-4s} \varepsilon^{3-1/10^4} + \varepsilon^{2r+1} + \varepsilon^{3r}) \\ &\lesssim_r \varepsilon^r \rho^2 (N_\beta^{-4s} \varepsilon^{-r} \varepsilon^{3-1/10^4} + \varepsilon^{r+1} + \varepsilon^{2r}), \end{aligned}$$

where we used the bound (6.15) on the symplectic transformation and the assumption on  $f(t)$ . Under the conditions of subsection (4.1) on the parameters (in particular the choice (4.1) of  $N_\beta$  such that  $N_\beta^{-2s} = \varepsilon^r$ ), we deduce that

$$\sum_{|n| \leq M} \langle n \rangle^{2s} \left| \frac{d}{dt} |v_n(t)|^2 \right| \lesssim_r N_\beta^{-2s} \varepsilon^{r+1} \rho^2.$$

Hence, integrating in time (with  $|t| \leq T \leq \varepsilon^{-r}$ ) gives

$$\sum_{|n| \leq M} \langle n \rangle^{2s} | |v_n(t)|^2 - \xi_n | \lesssim_r N_\beta^{-2s} \varepsilon \rho^2.$$

This is a stronger bound than the bootstrap assumption (7.13), and this completes the proof of Theorem 2.7.  $\square$

## APPENDIX

## A.1. PROOF OF PROPOSITION 4.16

*Proof of Proposition 4.16.* — The proof follows the same lines as the proof of Proposition 4.15, with some adjustments. We first prove (4.26): given  $Q$  and  $H$  we set

$$P(\xi; u) := \{Q, H\}(\xi; u) = \sum_{\mathbf{n}'' \in \mathfrak{N}_{\leq 2\tau^2}} (P[\xi])_{\mathbf{n}''} z_{\mathbf{n}''}(u, I(u) - \xi).$$

As in the proof of Proposition 4.15 we first reduce to fixed monomials  $(\mathbf{n}, \mathbf{n}')$ : for all  $\mathbf{n}''$  and  $k \in \mathbb{Z}_M^d$  we have

$$(A.1) \quad |\partial_{\xi_k}(P[\xi])_{\mathbf{n}''}| \leq \varepsilon^{-1/10^4} \sup_{\mathbf{a}, \mathbf{b}, \mathbf{n}, \mathbf{n}'} \left| \partial_{\xi_k} \left( \prod_{|j| \leq M} \xi_j^{a_j - b_j} (Q[\xi])_{\mathbf{n}} (H[\xi])_{\mathbf{n}'} \right) \right| \\ \left[ \mathbf{1}_{E_{\mathbf{n}'', \mathbf{a}, \mathbf{b}}^{(1)}(\mathbf{n}, \mathbf{n}')} |k'_n \ell_n - k_n \ell'_n| + \mathbf{1}_{E_{\mathbf{n}'', \mathbf{a}, \mathbf{b}}^{(2)}(\mathbf{n}, \mathbf{n}')} |m_n(k_{n'} - \ell'_n) + m_{n'}(k_n - \ell_n)| \right]$$

We separate three main cases, depending on where the derivative  $\partial_{\xi_k}$  falls.

*Case 1:  $\partial_{\xi_k}$  falls on a term  $\xi_k$  coming from the re-centered actions.* — In this case, we have

$$(A.2) \quad \left| \partial_{\xi_k} \left( \prod_{|j| \leq M} \xi_j^{a_j - b_j} \right) (Q[\xi])_{\mathbf{n}} (H[\xi])_{\mathbf{n}'} \right| \\ \leq \|Q\|_{Y_\alpha^{\text{sup}}} \|H\|_{Y_\alpha^{\text{sup}}} \left( \mathbf{1}_{a_k - b_k \geq 1} (a_k - b_k) \xi_k^{a_k - b_k - 1} \prod_{j \neq k} \xi_j^{a_j - b_j} \right) w_{\mathbf{n}}^0(\alpha) w_{\mathbf{n}'}^0(\alpha).$$

Hence, when  $a_k - b_k \geq 1$ , a term  $\xi_k$  is removed, and this costs a factor  $C_k(\alpha)^2$ . Following the proof of Proposition 4.15 we deduce that the contribution of (A.2) to (A.1) is bounded by

$$(A.2) \leq \|Q\|_{Y_\alpha^{\text{sup}}} \|H\|_{Y_\alpha^{\text{sup}}} \eta^{4-1/5} N_\alpha^{-4s} w_{\mathbf{n}''}^0(\alpha) C_k(\alpha)^2 (a_k - b_k) \\ \lesssim_r \|Q\|_{Y_\alpha^{\text{sup}}} \|H\|_{Y_\alpha^{\text{sup}}} \eta^{2-1/5} N_\alpha^{-2(s-\tau)} w_{\mathbf{n}''}^0(\alpha) \\ = \|Q\|_{Y_\alpha^{\text{sup}}} \|H\|_{Y_\alpha^{\text{sup}}} \eta^{4-1/5} N_\alpha^{-4s+2\tau} w_{\mathbf{n}''}^1(\alpha) \\ \leq \|Q\|_{Y_\alpha^{\text{sup}}} \|H\|_{Y_\alpha^{\text{sup}}} \eta^{4-1/4} N_\alpha^{-4s} w_{\mathbf{n}''}^1(\alpha),$$

where we used Definition 4.1 for the weights and condition (4.2) on the parameters.

*Case 2:  $\partial_{\xi_k}$  falls on either  $(Q[\xi])_{\mathbf{n}}$  or  $(H[\xi])_{\mathbf{n}'}$ .* — Since the two situations are symmetric, we suppose that it falls on  $(Q[\xi])_{\mathbf{n}}$ . In Case 2 we have to control the following quantity:

$$(A.3) \quad \varepsilon^{-1/10^4} \|Q\|_{Y_\alpha^{\text{lip}}} \|H\|_{Y_\alpha^{\text{sup}}} \sup_{\mathbf{a}, \mathbf{b}, \mathbf{n}, \mathbf{n}'} \left( \prod_{|j| \leq M} \xi_j^{a_j - b_j} \right) w_{\mathbf{n}}^1(\alpha) w_{\mathbf{n}'}^0(\alpha) \\ \left[ \mathbf{1}_{E_{\mathbf{n}'', \mathbf{a}, \mathbf{b}}^{(1)}(\mathbf{n}, \mathbf{n}')} |k'_n \ell_n - k_n \ell'_n| + \mathbf{1}_{E_{\mathbf{n}'', \mathbf{a}, \mathbf{b}}^{(1)}(\mathbf{n}, \mathbf{n}')} |m_n(k'_n - \ell'_n) + m_{n'}(\ell_n - k_n)| \right].$$

In comparison to (4.20),  $w_{\mathbf{n}}^0(\alpha)$  is substituted with  $w_{\mathbf{n}}^1(\alpha)$  in (A.3). Note that the weight  $w_{\mathbf{n}''}^1(\alpha)$  has the same pre-factor as  $w_{\mathbf{n}}^1(\alpha)$ , so it suffices to prove that  $w_{\mathbf{n}'}^0(\alpha)$

absorbs the losses. Reproducing the analysis performed in Cases 1 and 2 in the proof of Proposition 4.15 gives

$$|(A.3)| \leq \|Q\|_{Y_\alpha^{\text{lip}}} \|H\|_{Y_\alpha^{\text{sup}}} \eta^{4-1/4} N_\alpha^{-4s} w_{\mathbf{n}''}^1(\alpha),$$

which is conclusive.

Let us now turn to the proof of (4.27). Given an integrable quartic term  $Z \in X_4(\varepsilon)$ , the Poisson bracket  $\{Q, Z\} =: P$  is explicitly written in (4.23). Once again we stress out that  $P$  has no new integrable terms (hence there is no Case 1 as above). Differentiating the coefficient  $(P[\xi])_{\mathbf{n}''}$  with respect to  $\xi_k$  gives

$$(A.4) \quad \partial_{\xi_k}(P[\xi])_{\mathbf{n}''} = 2i \sum_{|j_1|, |j_2| \leq M} \sum_{\mathbf{n} \in \mathfrak{N}_{\leq \overline{r}^2}} \mathbf{1}_{E_{\mathbf{n}'', 0, 0}(j_1)}(\mathbf{n}, \mathbf{e}_m(j_1) + \mathbf{e}_m(j_2))(k_{j_1} - \ell_{j_1}) \\ \left( \partial_{\xi_k}(Q[\xi])_{\mathbf{n}}(Z[\xi])_{\mathbf{e}_m(j_1) + \mathbf{e}_m(j_2)} + (Q[\xi])_{\mathbf{n}} \partial_{\xi_k}(Z[\xi])_{\mathbf{e}_m(j_1) + \mathbf{e}_m(j_2)} \right).$$

As detailed in the proof of Proposition 4.15, we have the relation (4.24) and therefore

$$w_{\mathbf{n}}^1(\alpha) = \frac{1}{D(\alpha)} w_{\mathbf{n}''}^1(\alpha) = N_\alpha^{-2s} \eta^{2+1/5} w_{\mathbf{n}''}^1(\alpha).$$

We deduce that when  $\partial_{\xi_k}$  falls on  $(Q[\xi])_{\mathbf{n}}$  the contribution is bounded by

$$\lesssim_r \|Q\|_{Y_\alpha^{\text{lip}}} \|Z\|_{Z^{\text{lip}}} \varepsilon^{-1/10^4} w_{\mathbf{n}}^1(\alpha) \leq \|Q\|_{Y_\alpha^{\text{lip}}} \|Z\|_{Z^{\text{sup}}} N_\alpha^{-2s} \eta^{2+1/5} w_{\mathbf{n}''}^1(\alpha).$$

On the other hand, when  $\partial_{\xi_k}$  falls on  $(Z[\xi])_{\mathbf{e}_m(j_1) + \mathbf{e}_m(j_2)}$  we reproduce the proof of (4.25) to deduce that this contribution is bounded by

$$\|Q\|_{Y_\alpha^{\text{sup}}} \|Z\|_{Z^{\text{lip}}} N_\alpha^{-2s} \eta^{2+1/5-2/10^4} w_{\mathbf{n}''}^0(\alpha) \\ \leq \|Q\|_{Y_\alpha^{\text{sup}}} \|Z\|_{Z^{\text{lip}}} N_\alpha^{-4s} \eta^{4+1/5-2/10^4} w_{\mathbf{n}''}^1(\alpha).$$

This concludes the proof of Proposition 4.16.  $\square$

## A.2. PROOF OF LEMMA 6.16

*Proof of Lemma 6.16.* — We consider  $Q \in X_{\leq 2\overline{r}}(\varepsilon)$ . By homogeneity we can assume that

$$\|Q\|_{Y_\alpha^{\text{sup}}} = 1.$$

Given  $\xi \in \mathcal{U}_s(\varepsilon)$ ,  $u \in \mathcal{V}_{\alpha, \xi}(20\varepsilon, \xi)$ , and a normalized  $w \in h^s(\mathbb{Z}_M^d)$  with  $\|w\|_{h^s} = 1$ , we have

$$[d\nabla \mathcal{L}(Q)(\xi)(u)](w) = \sum_{\mathbf{n} \in \Lambda_{\alpha+1}} \frac{\mathfrak{h}^{(\alpha)}(\xi)}{\Omega_{\mathbf{n}}(\tilde{\omega}(\xi))} (Q[\xi])_{\mathbf{n}} [d\nabla z_{\mathbf{n}}(u, I(u) - \xi)](w).$$

The small-divisor estimate (6.21) together with the definition of the  $Y_\alpha^{\text{sup}}$ -norm and the counting estimate (4.4) give

$$(A.5) \quad \|[d\nabla \mathcal{L}(Q)(\xi)(u)](w)\|_{h^s} \\ \lesssim \varepsilon^{-1/10^4} \gamma(\alpha)^{-1} \varepsilon^{-2} N_{\alpha+1}^{2s} \max_{\substack{\mathbf{n} \in \Lambda_{\alpha+1} \\ \deg(\mathbf{n}) \leq 2\overline{r}}} w_{\mathbf{n}}^0(\alpha) \|[d\nabla z_{\mathbf{n}}(u, I(u) - \xi)](w)\|_{h^s}.$$

We will prove that

$$(A.6) \quad \max_{\substack{\mathbf{n} \in \Lambda_{\alpha+1} \\ \deg(\mathbf{n}) \leq 2\bar{r}}} w_{\mathbf{n}}^0(\alpha) \| [d\nabla z_{\mathbf{n}}(u, I(u) - \xi)](w) \|_{h^s} \lesssim_r \varepsilon^{-1/10^4} \eta^{4-2/5} N_{\alpha}^{-2s},$$

and conclude from (A.5) and the relation (4.6) between the parameters:

$$|(A.5)| \lesssim_r \varepsilon^{-2/10^4} \gamma(\alpha)^{-1} (N_{\alpha+1}/N_{\alpha})^{2s} (\eta \varepsilon^{-1})^2 \eta^{2-2/5} \leq \varepsilon^{1-1/8}.$$

Let us show (A.6). Using that the Hamiltonian is real, we have

$$(A.7) \quad \begin{aligned} \| [d\nabla z_{\mathbf{n}}(u, I(u) - \xi)](w) \|_{h^s} &\lesssim \left( \sum_{\mathbf{n}' \in \mathbb{Z}_M^d} \langle \mathbf{n}' \rangle^{2s} \left| \sum_{\mathbf{n} \in \mathbb{Z}_M^d} \partial_{u_n} \partial_{\overline{u_{n'}}} z_{\mathbf{n}}(u, I(u) - \xi) w_n \right|^2 \right)^{1/2} \\ &\quad + \left( \sum_{\mathbf{n}' \in \mathbb{Z}_M^d} \langle \mathbf{n}' \rangle^{2s} \left| \sum_{\mathbf{n} \in \mathbb{Z}_M^d} \partial_{u_n} \partial_{\overline{u_{n'}}} z_{\mathbf{n}}(u, I(u) - \xi) \overline{w_n} \right|^2 \right)^{1/2} \\ &\lesssim \varepsilon^{-1/10^4} \max_{|\mathbf{n}|, |\mathbf{n}'| \leq M} \langle \mathbf{n}' \rangle^s |\partial_{u_n} \partial_{\overline{u_{n'}}} z_{\mathbf{n}}(u, I(u) - \xi)| |w_n|. \end{aligned}$$

For  $\mathbf{n} \in \Lambda_{\alpha+1}$  and  $(\mathbf{n}, \mathbf{n}') \in (\mathbb{Z}_M^d)^2$  we have

$$(A.8) \quad \partial_{u_n} \partial_{\overline{u_{n'}}} z_{\mathbf{n}}(u, I(u) - \xi) = k_n \ell_{n'} \frac{z_{\mathbf{n}}(u, I(u) - \xi)}{u_n \overline{u_{n'}}$$

$$(A.9) \quad + m_n \ell_{n'} \frac{z_{\mathbf{n}}(u, I(u) - \xi)}{(|u_n|^2 - \xi_n) \overline{u_{n'}}} \overline{u_n}$$

$$(A.10) \quad + k_n m_{n'} \frac{z_{\mathbf{n}}(u, I(u) - \xi)}{u_n (|u_{n'}|^2 - \xi_{n'})} u_{n'}$$

$$(A.11) \quad + m_n m_{n'} \frac{z_{\mathbf{n}}(u, I(u) - \xi)}{(|u_n|^2 - \xi_n)(|u_{n'}|^2 - \xi_{n'})} \overline{u_n} u_{n'}.$$

In the subsequent case by case analysis, we estimate each contribution separately.

We deduce from the zero momentum condition that when  $\ell_{n'} \geq 1$ , there exists  $j \in \mathbb{Z}_M^d \setminus \{\mathbf{n}\}$  such that

$$(A.12) \quad \max(k_j, \ell_j) \geq 1, \quad |j| \geq \frac{|\mathbf{n}'|}{2\bar{r}}.$$

Without loss of generality, we suppose that  $k_j \geq 1$ .

*Case 1.* — In (A.8), when  $k_n \ell_{n'} \geq 1$  a pair  $(u_n, \overline{u_{n'}})$  goes away. We separate two cases.

– *Case 1(a).* If  $|\mathbf{n}| \gtrsim |\mathbf{n}'|$ , then

$$\begin{aligned} \langle \mathbf{n}' \rangle^s k_n \ell_{n'} w_{\mathbf{n}}^0(\alpha) \left| \frac{z_{\mathbf{n}}(u, I(u) - \xi)}{u_n \overline{u_{n'}}} w_{n'} \right| &\leq \langle \mathbf{n} \rangle^s |w_n| w_{\mathbf{n}}^0(\alpha) \ell_n k_{n'} \left| \frac{z_{\mathbf{n}}(u, I(u) - \xi)}{u_n \overline{u_{n'}}} \right| \\ &\lesssim_r C_n(\alpha) C_{n'}(\alpha) w_{\mathbf{n}'}^0(\alpha) |z_{\mathbf{n}'}(u, I(u) - \xi)|, \end{aligned}$$

where

$$\mathbf{n}' = \mathbf{n} - \mathbf{e}_{\mathbf{t}}(\mathbf{n}) - \mathbf{e}_{\mathbf{t}}(\mathbf{n}') \in \tilde{\mathfrak{N}}, \quad \text{and} \quad w_{\mathbf{n}'}^0(\alpha) = C_n(\alpha)^{-1} C_{n'}(\alpha)^{-1} w_{\mathbf{n}}^0(\alpha).$$

We deduce from the weighted estimate (4.9) for the monomials that, in Case 1(a),

$$(A.6) \lesssim_r N_{\alpha}^{2(s+\tau)} \eta^{-2} N_{\alpha}^{-6s} \eta^6 \lesssim_r \eta^4 N_{\alpha}^{4s-2\tau}.$$

– *Case 1(b)*. In this case, however, we have  $|n| \ll |n'|$ , and the factor  $w_n$  cannot absorb the derivative  $\langle n' \rangle^s$ . Nevertheless, the frequency  $j$  given by (A.12) is different from  $n$  and it is therefore still available. There holds

$$\begin{aligned} \langle n' \rangle^s w_{\mathbf{n}}^0(\alpha) k_n \ell_{n'} \left| \frac{z_{\mathbf{n}}(u, I(u) - \xi)}{u_n \overline{u_{n'}}} w_{n'} \right| \\ \lesssim_r \langle j \rangle^s |u_j| |w_{n'}| w_{\mathbf{n}}^0(\alpha) k_n \ell_{n'} \left| \frac{z_{\mathbf{n}}(u, I(u) - \xi)}{u_n \overline{u_{n'}} u_j} \right| \\ \lesssim_r \varepsilon C_j(\alpha) C_n(\alpha) C_{n'}(\alpha) w_{\mathbf{n}'}^0(\alpha) |z_{\mathbf{n}'}(u, I(u) - \xi)|, \end{aligned}$$

where

$$\mathbf{n}' = \mathbf{n} - \mathbf{e}_{\mathbf{t}}(n) - \mathbf{e}_{\mathbf{l}}(n') - \mathbf{e}_{\mathbf{t}}(j) \in \tilde{\mathfrak{N}}, \quad w_{\mathbf{n}'}^0(\alpha) = (C_n(\alpha) C_{n'}(\alpha, \varepsilon) C_j(\alpha))^{-1} w_{\mathbf{n}}^0(\alpha)$$

We deduce from the weighted estimate (4.9) for monomials that, in Case 1(b),

$$(A.6) \lesssim_r \varepsilon N_{\alpha}^{3(s+\tau)} \eta^{-3} N_{\alpha}^{-6s} \eta^6 \lesssim_r \eta^4 N_{\alpha}^{-3(s-\tau)}.$$

*Case 2.* — In (A.9), when  $m_n \ell'_n \geq 1$  a pair  $(|u_n|^2 - \xi_n, \overline{u_{n'}})$  goes away and a term  $\overline{u_n}$  appears.

– *Case 2(a)*. If  $|n| \gtrsim |n'|$ , then

$$\begin{aligned} \langle n' \rangle^s w_{\mathbf{n}}^0(\alpha) m_n \ell_{n'} \left| \frac{z_{\mathbf{n}}(u, I(u) - \xi)}{(|u_n|^2 - \xi_n) \overline{u_{n'}}} \overline{u_n} w_{n'} \right| &\lesssim \langle n \rangle^s |w_n u_n| w_{\mathbf{n}}^0(\alpha) m_n \ell_{n'} \left| \frac{z_{\mathbf{n}}(u, I(u) - \xi)}{(|u_n|^2 - \xi_n) \overline{u_{n'}}} \right| \\ &\lesssim_r \varepsilon C_{n'}(\alpha) D(\alpha) w_{\mathbf{n}'}^0(\alpha) |z_{\mathbf{n}'}|, \end{aligned}$$

where

$$\mathbf{n}' = \mathbf{n} - \mathbf{e}_{\mathbf{m}}(n) - \mathbf{e}_{\mathbf{l}}(n') \in \tilde{\mathfrak{N}}, \quad w_{\mathbf{n}'}^0(\alpha) = (C_{n'}(\alpha) D(\alpha))^{-1} w_{\mathbf{n}}^0(\alpha).$$

We deduce from the weighted estimate for monomials that, in Case 2(a),

$$(A.6) \lesssim_r \varepsilon N_{\alpha}^{3s+\tau} \eta^{-3-1/5} N_{\alpha}^{-6s} \eta^6 \lesssim_r \eta^{4-1/5} N_{\alpha}^{-(3s-\tau)}.$$

– *Case 2(b)*. If  $|n| \ll |n'|$ , then we can find  $j \in \mathbb{Z}_M^d \setminus \{n, n'\}$  such that (A.12) holds, and deduce that

$$\begin{aligned} \langle n' \rangle^s w_{\mathbf{n}}^0(\alpha) m_n \ell_{n'} \left| \frac{z_{\mathbf{n}}(u, I(u) - \xi)}{(|u_n|^2 - \xi_n) \overline{u_{n'}}} \overline{u_n} w_{n'} \right| \\ \lesssim_r \langle j \rangle^s |u_j w_n u_n| w_{\mathbf{n}}^0(\alpha) m_n \ell_{n'} \left| \frac{z_{\mathbf{n}}(u, I(u) - \xi)}{(|u_n|^2 - \xi_n) \overline{u_{n'}} u_j} \right| \\ \lesssim_r \varepsilon^2 C_j(\alpha) C_{n'}(\alpha) D(\alpha) w_{\mathbf{n}'}^0(\alpha) |z_{\mathbf{n}'}(u, I(u) - \xi)|, \end{aligned}$$

where

$$\mathbf{n}' = \mathbf{n} - \mathbf{e}_{\mathbf{m}}(n) - \mathbf{e}_{\mathbf{l}}(n') - \mathbf{e}_{\mathbf{t}}(j), \quad w_{\mathbf{n}'}^0(\alpha) = (C_{n'}(\alpha) C_j(\alpha) D(\alpha))^{-1} w_{\mathbf{n}}^0(\alpha).$$

We deduce from the weighted bound (4.9) for the monomials that, in Case 2b),

$$(A.6) \lesssim_r \varepsilon^2 N_{\alpha}^{-4s-2\tau} \eta^{-4-1/5} N_{\alpha}^{6s} \eta^6 \lesssim_r \eta^{4-1/5} N_{\alpha}^{-2(s-\tau)}.$$

*Case 3.* — In (A.10), when  $k_n m_{n'} \geq 1$  a pair  $(u_n, |u_{n'}|^2 - \xi_{n'})$  goes away, and a term  $u_{n'}$  appears. It holds

$$\begin{aligned} \langle n' \rangle^s w_{\mathbf{n}}^0(\alpha) k_n m_{n'} & \left| \frac{z_{\mathbf{n}}(u, I(u) - \xi)}{u_n(|u_{n'}|^2 - \xi_{n'})} u_{n'} w_n \right| \\ & \leq \langle n' \rangle^s |u_{n'} w_n| w_{\mathbf{n}}^0(\alpha) k_n m_{n'} \left| \frac{z_{\mathbf{n}}(u, I(u) - \xi)}{u_n(|u_{n'}|^2 - \xi_{n'})} \right| \\ & \lesssim_r \varepsilon D(\alpha) C_n(\alpha) w_{\mathbf{n}'}^0(\alpha) |z_{\mathbf{n}'}(u, I(u) - \xi)|, \end{aligned}$$

with

$$\mathbf{n}' = \mathbf{n} - \mathbf{e}_t(n) - \mathbf{e}_m(n'), \quad w_{\mathbf{n}'}^0(\alpha) = (D(\alpha) C_n(\alpha))^{-1} w_{\mathbf{n}}^0(\alpha).$$

We deduce from the weighted estimate (4.9) for monomials that

$$(A.6) \lesssim_r \varepsilon N_{\alpha}^{3s+\tau} \eta^{-3-1/5} N_{\alpha}^{-6s} \eta^6 \lesssim_r \eta^{4-1/5} N_{\alpha}^{-3s+\tau}.$$

*Case 4.* — The last case (A.11) is a priori the worst since a pair of centered actions goes away, and induces a loss  $D(\alpha)^2$ . Nevertheless, a pair  $(\overline{u_n}, u_{n'})$  appears and will absorb the derivative in  $n$ . There is also a term  $w_{n'}$ , which we will not exploit. It holds

$$\begin{aligned} \langle n' \rangle^s w_{\mathbf{n}}^0(\alpha) m_n m_{n'} & \left| \frac{z_{\mathbf{n}}(u, I(u) - \xi)}{(|u_n|^2 - \xi_n)(|u_{n'}|^2 - \xi_{n'})} \overline{u_n} u_{n'} w_{n'} \right| \\ & \leq \langle n' \rangle^s |u_n| |u_{n'}| |w_{n'}| w_{\mathbf{n}}^0(\alpha) m_n m_{n'} \left| \frac{z_{\mathbf{n}}(u, I(u) - \xi)}{(|u_n|^2 - \xi_n)(|u_{n'}|^2 - \xi_{n'})} \right| \\ & \lesssim_r \varepsilon^2 D(\alpha)^2 w_{\mathbf{n}'}^0(\alpha) |z_{\mathbf{n}'}(u, I(u) - \xi)|, \end{aligned}$$

with

$$\mathbf{n}' = \mathbf{n} - \mathbf{e}_m(n) - \mathbf{e}_m(n'), \quad w_{\mathbf{n}'}^0(\alpha) = D(\alpha)^{-2} w_{\mathbf{n}}^0(\alpha).$$

We deduce from the weighted estimate for the monomials (4.9) that this contribution is bounded by

$$(A.6) \lesssim_r \varepsilon^2 N_{\alpha}^{4s} \eta^{-4-2/5} N_{\alpha}^{-6s} \eta^6 \lesssim_r \eta^{4-2/5} N_{\alpha}^{-2s}.$$

This proves that all contributions to (A.7) are acceptable to obtain (A.6), which completes the proof of Lemma 6.16.  $\square$

## REFERENCES

- [BH22] P. BALDI & E. HAUS — “Longer lifespan for many solutions of the Kirchhoff equation”, *SIAM J. Math. Anal.* **54** (2022), no. 1, p. 306–342.
- [Bam99] D. BAMBUSI — “Nekhoroshev theorem for small amplitude solutions in nonlinear Schrödinger equations”, *Math. Z.* **230** (1999), no. 2, p. 345–387.
- [BDGS07] D. BAMBUSI, J.-M. DELORT, B. GRÉBERT & J. SZEFTTEL — “Almost global existence for Hamiltonian semilinear Klein-Gordon equations with small Cauchy data on Zoll manifolds”, *Comm. Pure Appl. Math.* **60** (2007), no. 11, p. 1665–1690.
- [BFM24] D. BAMBUSI, R. FEOLA & R. MONTALTO — “Almost global existence for some Hamiltonian PDEs with small Cauchy data on general tori”, *Comm. Math. Phys.* **405** (2024), no. 1, article no. 15 (50 pages).
- [BG24] D. BAMBUSI & P. GÉRARD — “A Nekhoroshev theorem for some perturbations of the Benjamin-Ono equation with initial data close to finite gap tori”, *Math. Z.* **307** (2024), no. 3, article no. 54 (20 pages).
- [BG06] D. BAMBUSI & B. GRÉBERT — “Birkhoff normal form for partial differential equations with tame modulus”, *Duke Math. J.* **135** (2006), no. 3, p. 507–567.

- [BCGW24] J. BERNIER, N. CAMPS, B. GRÉBERT & Z. WANG – “Exponential stability of solutions to the Schrödinger–Poisson equation”, *Discrete Contin. Dynam. Systems. Ser. A* **44** (2024), no. 11, p. 3398–3442.
- [BFG20a] J. BERNIER, E. FAOU & B. GRÉBERT – “Long time behavior of the solutions of NLW on the  $d$ -dimensional torus”, *Forum Math. Sigma* **8** (2020), article no. e12 (26 pages).
- [BFG20b] ———, “Rational normal forms and stability of small solutions to nonlinear Schrödinger equations”, *Ann. PDE* **6** (2020), no. 2, article no. 14 (65 pages).
- [BG21] J. BERNIER & B. GRÉBERT – “Long time dynamics for generalized Korteweg-de Vries and Benjamin-Ono equations”, *Arch. Rational Mech. Anal.* **241** (2021), no. 3, p. 1139–1241.
- [BG25] ———, “Almost global existence for some nonlinear Schrödinger equations on  $\mathbb{T}^d$  in low regularity”, *Ann. Inst. Fourier (Grenoble)* (2025), Online first.
- [BGR23] J. BERNIER, B. GRÉBERT & T. ROBERT – “Dynamics of quintic nonlinear Schrödinger equations in  $H^{2/5+}(\mathbb{T})$ ”, 2023, To appear in Transactions of the AMS, [arXiv:2305.05236](#).
- [BM19] M. BERTI & A. MASPERO – “Long time dynamics of Schrödinger and wave equations on flat tori”, *J. Differential Equations* **267** (2019), no. 2, p. 1167–1200.
- [BD18] M. BERTI & J.-M. DELORT – *Almost global solutions of capillary-gravity water waves equations on the circle*, Lect. Notes of the Unione Matematica Italiana, vol. 24, Springer, Cham, 2018.
- [BFP23] M. BERTI, R. FEOLA & F. PUSATERI – “Birkhoff normal form and long time existence for periodic gravity water waves”, *Comm. Pure Appl. Math.* **76** (2023), no. 7, p. 1416–1494.
- [BMM24] M. BERTI, A. MASPERO & F. MURGANTE – “Hamiltonian Birkhoff normal form for gravity-capillary water waves with constant vorticity: almost global existence”, *Ann. PDE* **10** (2024), no. 2, article no. 22 (150 pages).
- [BMP20] L. BIASCO, J. E. MASSETTI & M. PROCESI – “An abstract Birkhoff normal form theorem and exponential type stability of the 1d NLS”, *Comm. Math. Phys.* **375** (2020), no. 3, p. 2089–2153.
- [BMP21] ———, “Almost periodic invariant tori for the NLS on the circle”, *Ann. Inst. H. Poincaré C Anal. Non Linéaire* **38** (2021), no. 3, p. 711–758.
- [Bou98] J. BOURGAIN – “Quasi-periodic solutions of Hamiltonian perturbations of 2D linear Schrödinger equations”, *Ann. of Math. (2)* **148** (1998), no. 2, p. 363–439.
- [Bou00a] ———, “On diffusion in high-dimensional Hamiltonian systems and PDE”, *J. Anal. Math.* **80** (2000), p. 1–35.
- [Bou00b] ———, “Problems in Hamiltonian PDE’s”, *Geom. Funct. Anal.* (2000), p. 32–56, GAFA 2000 (Tel Aviv, 1999).
- [Bou05] ———, “On invariant tori of full dimension for 1D periodic NLS”, *J. Functional Analysis* **229** (2005), no. 1, p. 62–94.
- [BK05] J. BOURGAIN & V. KALOSHIN – “On diffusion in high-dimensional Hamiltonian systems”, *J. Functional Analysis* **229** (2005), no. 1, p. 1–61.
- [Cam24] N. CAMPS – “Stability results for the nonlinear Schrödinger equation on Diophantine tori”, 2024, Journées EDP, Aussois, France, [HAL:hal-05038184](#).
- [CS24] N. CAMPS & G. STAFFILANI – “Modified scattering for the cubic Schrödinger equation on Diophantine waveguides”, 2024, [arXiv:2404.16817](#).
- [CKS<sup>+</sup>10] J. COLLIANDER, M. KEEL, G. STAFFILANI, H. TAKAOKA & T. TAO – “Transfer of energy to high frequencies in the cubic defocusing nonlinear Schrödinger equation”, *Invent. Math.* **181** (2010), no. 1, p. 39–113.
- [CG22] C. COLLOT & P. GERMAIN – “Derivation of the homogeneous kinetic wave equation: longer time scales”, 2022, [arXiv:2007.03508](#).
- [Del12] J.-M. DELORT – *A quasi-linear Birkhoff normal forms method. Application to the quasi-linear Klein-Gordon equation on  $\mathbb{S}^1$* , Astérisque, no. 341, Société Mathématique de France, Paris, 2012.
- [Den19] Y. DENG – “On growth of Sobolev norms for energy critical NLS on irrational tori: small energy case”, *Comm. Pure Appl. Math.* **72** (2019), no. 4, p. 801–834.
- [DG19] Y. DENG & P. GERMAIN – “Growth of solutions to NLS on irrational tori”, *Internat. Math. Res. Notices* (2019), no. 9, p. 2919–2950.

- [DGG17] Y. DENG, P. GERMAIN & L. GUTH – “Strichartz estimates for the Schrödinger equation on irrational tori”, *J. Functional Analysis* **273** (2017), no. 9, p. 2846–2869.
- [DGGRM22] Y. DENG, P. GERMAIN, L. GUTH & S. L. RYDIN MYERSON – “Strichartz estimates for the Schrödinger equation on non-rectangular two-dimensional tori”, *Amer. J. Math.* **144** (2022), no. 3, p. 701–745.
- [FGL13] E. FAOU, L. GAUCKLER & C. LUBICH – “Sobolev stability of plane wave solutions to the cubic nonlinear Schrödinger equation on a torus”, *Comm. Partial Differential Equations* **38** (2013), no. 7, p. 1123–1140.
- [FI21] R. FEOLA & F. IANDOLI – “Long time existence for fully nonlinear NLS with small Cauchy data on the circle”, *Ann. Scuola Norm. Sup. Pisa Cl. Sci. (5)* **22** (2021), no. 1, p. 109–182.
- [GG22] F. GIULIANI & M. GUARDIA – “Sobolev norms explosion for the cubic NLS on irrational tori”, *Nonlinear Anal.* **220** (2022), article no. 112865 (25 pages).
- [GIP09] B. GRÉBERT, R. IMEKRAZ & E. PATUREL – “Normal forms for semilinear quantum harmonic oscillators”, *Comm. Math. Phys.* **291** (2009), no. 3, p. 763–798.
- [GK15] M. GUARDIA & V. KALOSHIN – “Growth of Sobolev norms in the cubic defocusing nonlinear Schrödinger equation”, *J. Eur. Math. Soc. (JEMS)* **17** (2015), no. 1, p. 71–149, Erratum: *Ibid.* **19** (2017), no. 2, p. 601–602.
- [HPTV15] Z. HANI, B. PAUSADER, N. TZVETKOV & N. VISCIGLIA – “Modified scattering for the cubic Schrödinger equation on product spaces and applications”, *Forum Math. Pi* **3** (2015), article no. e4 (63 pages).
- [KP96] S. KUKSIN & J. PÖSCHEL – “Invariant Cantor manifolds of quasi-periodic oscillations for a nonlinear Schrödinger equation”, *Ann. of Math. (2)* **143** (1996), no. 1, p. 149–179.
- [LX24] J. LIU & D. XIANG – “Exact global control of small divisors in rational normal form”, *Nonlinearity* **37** (2024), no. 7, article no. 075020 (48 pages).
- [Pös02] J. PÖSCHEL – “On the construction of almost periodic solutions for a nonlinear Schrödinger equation”, *Ergodic Theory Dynam. Systems* **22** (2002), no. 5, p. 1537–1549.
- [SW20] G. STAFFILANI & B. WILSON – “Stability of the cubic nonlinear Schrödinger equation on an irrational torus”, *SIAM J. Math. Anal.* **52** (2020), no. 2, p. 1318–1342.
- [YZ14] X. YUAN & J. ZHANG – “Long time stability of Hamiltonian partial differential equations”, *SIAM J. Math. Anal.* **46** (2014), no. 5, p. 3176–3222.

Manuscript received 4th April 2024

accepted 13th May 2025

JOACKIM BERNIER, Nantes Université, CNRS, Laboratoire de Mathématiques Jean Leray, LMJL,  
F-44000 Nantes, France

E-mail : [joackim.bernier@univ-nantes.fr](mailto:joackim.bernier@univ-nantes.fr)

Url : <https://www.math.sciences.univ-nantes.fr/~bernier-j>

NICOLAS CAMPS, Univ Rennes, CNRS, IRMAR - UMR 6625,  
F-35000 Rennes, France

E-mail : [nicolas.camps@univ-rennes.fr](mailto:nicolas.camps@univ-rennes.fr)

Url : <https://perso.pages.math.cnrs.fr/users/nicolas.camps>