J. Eur. Math. Soc. 25, 1497–1551 (2023) DOI 10.4171/JEMS/1200

Marcel Guardia · Zaher Hani · Emanuele Haus · Alberto Maspero · Michela Procesi

Strong nonlinear instability and growth of Sobolev norms near quasiperiodic finite gap tori for the 2D cubic NLS equation

Received October 22, 2019

Abstract. We consider the defocusing cubic nonlinear Schrödinger equation (NLS) on the two-dimensional torus. The equation admits a special family of elliptic invariant quasiperiodic tori called finite gap solutions. These are inherited from the integrable 1D model (cubic NLS on the circle) by considering solutions that depend only on one variable. We study the long-time stability of such invariant tori for the 2D NLS model and show that, under certain assumptions and over sufficiently long time scales, they exhibit a strong form of *transverse instability* in Sobolev spaces $H^s(\mathbb{T}^2)$ $(0 < s < 1)$. More precisely, we construct solutions of the 2D cubic NLS that start arbitrarily close to such invariant tori in the H^s topology and whose H^s norm can grow by any given factor. This work is partly motivated by the problem of infinite energy cascade for 2D NLS, and seems to be the first instance where (unstable) long-time nonlinear dynamics near (linearly stable) quasiperiodic tori is studied and constructed.

Keywords. Nonlinear Schrödinger equation, quasiperiodic, KAM, stability, growth of Sobolev norms

1. Introduction

A widely held principle in dynamical systems theory is that invariant quasiperiodic tori play an important role in understanding the complicated long-time behavior of Hamilto-

Marcel Guardia: Departament de Matemàtiques & IMTECH, Universitat Politècnica de Catalunya, Barcelona; and Centre de Recerca Matemàtica, Barcelona, Spain; marcel.guardia@upc.edu

Zaher Hani: Department of Mathematics, University of Michigan, 530 Church Street, Ann Arbor, MI 48109-1043, USA; zhani@umich.edu

Emanuele Haus: Dipartimento di Matematica e Fisica, Università di Roma Tre, Largo S.L. Murialdo, 00146 Roma, Italy; ehaus@mat.uniroma3.it

Alberto Maspero: International School for Advanced Studies (SISSA), Via Bonomea 265, 34136 Trieste, Italy; alberto.maspero@sissa.it

Michela Procesi: Dipartimento di Matematica e Fisica, Università di Roma Tre, Largo S.L. Murialdo, 00146 Roma, Italy; procesi@mat.uniroma3.it

Mathematics Subject Classification (2020): 35Q55, 37K45

nian ODE and PDE. In addition to being important in their own right, the hope is that such quasiperiodic tori can play an important role in understanding other, possibly more generic, dynamics of the system by acting as *islands* in whose vicinity orbits might spend long periods of time before moving to other such islands. The construction of such invariant sets for Hamiltonian PDE has witnessed an explosion of activity over the past thirty years after the success of extending KAM techniques to infinite dimensions. However, the dynamics near such tori is still poorly understood, and often restricted to the linear theory. The purpose of this work is to take a step in the direction of understanding and constructing nontrivial *nonlinear* dynamics in the vicinity of certain quasiperiodic solutions for the cubic defocusing NLS equation. In line with the above philosophy emphasizing the role of invariant quasiperiodic tori for other types of behavior, another aim is to push forward a program aimed at proving infinite Sobolev norm growth for the 2D cubic NLS equation, an outstanding open problem.

1.1. The dynamical system and its quasiperiodic objects

We start by describing the dynamical system and its quasiperiodic invariant objects at the center of our analysis. Consider the periodic cubic defocusing nonlinear Schrödinger equation (NLS),

$$\mathrm{i}\partial_t u + \Delta u = |u|^2 u \tag{2D-NLS}$$

where $(x, y) \in \mathbb{T}^2 = \mathbb{R}^2/(2\pi\mathbb{Z})^2$, $t \in \mathbb{R}$ and $u : \mathbb{R} \times \mathbb{T}^2 \to \mathbb{C}$. All the results in this paper extend trivially to higher dimensions $d \geq 3$ by considering solutions that only depend on two variables.[1] This is a Hamiltonian PDE with conserved quantities: (i) the Hamiltonian

$$H_0(u) = \int_{\mathbb{T}^2} \big(|\nabla u(x, y)|^2 + \tfrac{1}{2}|u(x, y)|^4\big)\mathrm{d}x\,\mathrm{d}y, \tag{1.1}$$

(ii) the mass

$$M(u) = \int_{\mathbb{T}^2} |u(x, y)|^2\,\mathrm{d}x\,\mathrm{d}y, \tag{1.2}$$

which is just the square of the L^2 norm of the solution, and (iii) the momentum

$$P(u) = \mathrm{i}\int_{\mathbb{T}^2} \overline{u(x, y)}\nabla u(x, y)\,\mathrm{d}x\,\mathrm{d}y. \tag{1.3}$$

We also remark that the equation is locally well-posed for data in $H^s(\mathbb{T}^d)$ for all $s > 0$ [6]. Thanks to the conservation of energy (and the subcritical nature of the local well-posedness result), one directly obtains global well-posedness in $H^s(\mathbb{T}^d)$ for $s \geq 1$. This global existence can be pushed further down to at least $s > 2/3$ using almost conservation inequalities (see for instance [15]). All the solutions constructed in this manuscript are infinitely smooth, and hence their global-in-time existence is guaranteed.

[1] We expect that the results also extend to the focusing sign of the nonlinearity ($-|u|^2 u$ on the R.H.S. of (2D-NLS)). The reason we restrict to the defocusing sign comes from the fact that the linear analysis around our quasiperiodic tori has only been established in full detail in [43] in this case.

Now, we describe the invariant objects around which we will study and construct our long-time nonlinear dynamics. Of course, such a task requires a very precise understanding of the linearized dynamics around such objects. For this reason, we take the simplest nontrivial family of invariant quasiperiodic tori admitted by (2D-NLS), namely those inherited from its completely integrable 1D counterpart

$$\mathrm{i}\partial_t q = -\partial_{xx} q + |q|^2 q, \quad x \in \mathbb{T}. \tag{1D-NLS}$$

This is a subsystem of (2D-NLS) if we consider solutions that depend only on the first spatial variable. It is well known that equation (1D-NLS) is integrable and its phase space is foliated by tori of finite or infinite dimension with periodic, quasiperiodic, or almost periodic dynamics. The quasiperiodic orbits are usually called *finite gap solutions*.

Such tori are Lyapunov stable (for all time!) as solutions of (1D-NLS) (as will be clear once we exhibit its integrable structure) and some of them are linearly stable as solutions of (2D-NLS), but we will be interested in their *long-time nonlinear stability* (or lack of it) as invariant objects for the 2D equation (2D-NLS). In fact, we shall show that they are *nonlinearly unstable* as solutions of (2D-NLS), and in a strong sense, in certain topologies and after very long time. Such instability is transversal in the sense that one drifts along the purely 2-dimensional directions: solutions which are initially very close to 1-dimensional become strongly 2-dimensional after some long time.[2]

1.2. Energy cascade, Sobolev norm growth, and Lyapunov instability

In addition to studying long-time dynamics close to invariant objects for NLS, another purpose of this work is to make progress on a fundamental problem in nonlinear wave theory, which is the transfer of energy between characteristically different scales for a nonlinear dispersive PDE. This is called the *energy cascade* phenomenon. It is a purely nonlinear phenomenon (energy is static in frequency space for the linear system), and will be the underlying mechanism behind the long-time instability of the finite gap tori mentioned above.

We shall exhibit solutions whose energy moves from very high frequencies towards low frequencies (*backward* or *inverse cascade*), as well as ones that exhibit a cascade in the opposite direction (*forward* or *direct cascade*). Such cascade phenomena have attracted a lot of attention in the past few years as they are central aspects of various theories of turbulence for nonlinear systems. For dispersive PDE, this goes by the name of *wave turbulence theory* which predicts the existence of solutions (and statistical states) of (2D-NLS) that exhibit a cascade of energy between very different length scales. In the mathematical community, Bourgain drew attention to such questions of energy cascade by first noting that it can be captured in a quantitative way by studying the behavior of the

[2] The tranversal instability phenomenon was already studied for solitary waves of the water waves equation [51] and the KP-I equation [52] by Rousset and Tzvetkov. However, their instability is a *linear effect*, in the sense that the linearized dynamics is unstable. In contrast, our result is a fundamentally nonlinear effect, as the linearized dynamics around some of the finite gap tori is stable.

Sobolev norms of the solution,

$$\|u\|_{H^s} = \Big(\sum_{n\in\mathbb{Z}^2} (1+|n|)^{2s} |\widehat{u}_n|^2 \Big)^{1/2}.$$

In his list of problems on Hamiltonian PDEs [9], Bourgain asked whether there exist solutions that exhibit a quantitative version of the forward energy cascade, namely solutions whose Sobolev H^s norms, with $s > 1$, are unbounded in time,

$$\sup_{t\geq 0} \|u(t)\|_{H^s} = +\infty, \quad s > 1. \tag{1.4}$$

We should point out here that such growth cannot happen for $s = 0$ or $s = 1$ due to the conservation laws of the equations. For other Sobolev indices, there exist polynomial upper bounds for the growth of Sobolev norms (see [7, 10, 11, 13, 14, 44, 53–56, 60]). Nevertheless, results proving actual growth of Sobolev norms are much more scarce. After seminal works by Bourgain himself [7] and Kuksin [37,39,40], the landmark result in [12] was of fundamental importance in the recent progress, including this work: It showed that for any $s > 1$, $\delta \ll 1$, $K \gg 1$, there exist solutions u of (2D-NLS) such that

$$\|u(0)\|_{H^s} \leq \delta \quad \text{and} \quad \|u(T)\|_{H^s} \geq K \tag{1.5}$$

for some $T > 0$. Even if not mentioned in that paper, the same techniques also lead to the same result for $s \in (0, 1)$. This paper induced a lot of activity in the area [26–28, 31–33] (see also [16, 22–24, 42, 45, 46] on results about growth of Sobolev norms with different techniques). Despite all that, Bourgain's question about solutions exhibiting (1.4) remains open on $\mathbb{T}^d$ (however, a positive answer has been given for the cylindrical domain $\mathbb{R} \times \mathbb{T}^d$ [32]).

The above-cited works revealed an intimate connection between Lyapunov instability and Sobolev norm growth. Indeed, the solution $u = 0$ of (2D-NLS) is an elliptic critical point and is linearly stable in all H^s. From this point of view, the result in [12] given in (1.5) can be interpreted as a strong form of Lyapunov instability (see item (6) in Section 1.4) in H^s, $s \neq 1$, of the elliptic critical point $u = 0$ (the first integrals (1.1) and (1.2) imply Lyapunov stability in the H^1 and L^2 topology). It turns out that this connection goes further, particularly in relation to the question of finding solutions exhibiting (1.4). As was observed in [31], one way to prove the existence of such solutions is to prove that, for sufficiently many $\phi \in H^s$, an instability similar to that in (1.5) holds, but with $\|u(0) - \phi\|_{H^s} \leq \delta$. In other words, proving long-time instability as in (1.5) but with solutions starting δ-close to ϕ, and for sufficiently many $\phi \in H^s$, implies the existence (and possible genericness) of unbounded orbits satisfying (1.4). Such a program (based on a Baire category argument) was applied successfully for the Szegő equation on $\mathbb{T}$ in [24].

Motivated by this, one is naturally led to studying this strong form of Lyapunov instability for more general invariant objects of (2D-NLS) (or other Hamiltonian PDEs), or equivalently to investigate whether one can achieve Sobolev norm explosion starting arbitrarily close to a given invariant object. The first work in this direction is by one of the present authors [31]. He considers the plane waves $u(t, x) = Ae^{\mathrm{i}(mx-\omega t)}$ with $\omega = m^2 + A^2$, periodic orbits of (2D-NLS), and proves that there are orbits which start

δ-close to them and undergo H^s Sobolev norm explosion, $0 < s < 1$. This implies that the plane waves are strongly Lyapunov unstable in these topologies. Stability results for plane waves in H^s, $s > 1$, on shorter time scales are provided in [20].

1.3. Statement of results

Roughly speaking, we will construct solutions to (2D-NLS) that start very close to finite gap tori in appropriate topologies, and exhibit either a backward cascade of energy from high to low frequencies, or a forward cascade from low to high frequencies. In the former case, the solutions that exhibit backward cascade start in an arbitrarily small vicinity of a finite gap torus in a Sobolev space $H^s(\mathbb{T}^2)$ with $0 < s < 1$, but grow to become larger than any pre-assigned $K \gg 1$ in the same H^s (higher Sobolev H^s norms with $s > 1$ decrease, but they are large for all times). In the latter case, the solutions that exhibit a forward cascade start in an arbitrarily small vicinity of a finite gap torus in $L^2(\mathbb{T}^2)$, but their Sobolev H^s norm (for all $s > 1$) exhibits growth by a large multiplicative factor $K \gg 1$ after a large time. We shall comment further on those results after we state the theorems precisely.

To do that, we need to introduce the Birkhoff coordinates for equation (1D-NLS). Grébert and Kappeler showed in [25] that there exists a globally defined map, called the *Birkhoff map*, such that for all $s \geq 0$,

$$\Phi : H^s(\mathbb{T}) \to h^s(\mathbb{Z}) \times h^s(\mathbb{Z}), \quad q \mapsto (z_m, \overline{z}_m)_{m\in\mathbb{Z}}, \tag{1.6}$$

such that equation (1D-NLS) is transformed in the new coordinates $(z_m, \overline{z}_m)_{m\in\mathbb{Z}} = \Phi(q)$ to

$$\mathrm{i}\dot{z}_m = \alpha_m(I) z_m \tag{1.7}$$

where $I = (I_m)_{m\in\mathbb{Z}}$ and $I_m = |z_m|^2$ are the actions, which are conserved in time (since $\alpha_m(I) \in \mathbb{R}$). Therefore in these coordinates, called *Birkhoff coordinates*, equation (1D-NLS) becomes a chain of nonlinear harmonic oscillators. Of course the solutions of (1.7) live on finite- and infinite-dimensional tori with periodic, quasiperiodic or almost periodic dynamics, depending on how many of the actions I_m (which are constant!) are nonzero and on the properties of rational dependence of the frequencies. Hence the (1D-NLS) equation admits a family of finite-dimensional integrable subsystems, denoted by $\mathcal{G}_\mathcal{S}$, where $\mathcal{S}$ runs through the nonempty, finite subsets of $\mathbb{Z}$; $\mathcal{G}_\mathcal{S}$ is contained in $\bigcap_{n\geq 0} H^n(\mathbb{T}, \mathbb{C})$ and its elements are called $\mathcal{S}$-*gap solutions*. In particular, $\mathcal{G}_\mathcal{S}$ is foliated by $\mathcal{T}_\mathcal{S}^I := \Phi^{-1}(\mathbb{T}_\mathcal{S}^I)$ where

$$\mathbb{T}_\mathcal{S}^I := \{z \in \ell^2 : |z_m|^2 = I_m \text{ for } m \in \mathcal{S},\ |z_m|^2 = 0 \text{ for } m \notin \mathcal{S}\}$$

is a torus of dimension $|\mathcal{S}|$ parametrized by the action variables $I = (I_j)_{j\in\mathcal{S}} \in \mathbb{R}^{\mathcal{S}}_{>0}$. This torus, as an invariant object of equation (1D-NLS), is stable for all times in the sense of Lyapunov.

We will abuse notation, and identify $H^s(\mathbb{T})$ with the closed subspace of $H^s(\mathbb{T}^2)$ of functions depending only on the x variable. Consequently, $\mathcal{T}_\mathcal{S}^I$ is a closed torus in $H^s(\mathbb{T}) \subset H^s(\mathbb{T}^2)$ which is invariant for the (2D-NLS) dynamics.

The main result of this paper will show the instability (in the sense of Lyapunov) of many of these invariant objects under the dynamics of (2D-NLS). Roughly speaking, we show that, under certain assumptions on the choices of modes $\mathcal{S}$ and actions I, these tori are unstable in the $H^s(\mathbb{T}^2)$ topology for $s \in (0,1)$. Even more, there exist orbits which start arbitrarily close to these tori and undergo an arbitrarily large H^s norm explosion. In order to state our result precisely, we introduce the definition of generic set $\mathcal{S}$:

Definition 1.1 ($\mathtt{L}$-genericity). Given $\mathtt{L} \in \mathbb{N}$, we say that $\mathcal{S} = \{\mathtt{m}_1, \dots, \mathtt{m}_\mathtt{d}\}$ is $\mathtt{L}$-*generic* if it satisfies the condition

$$\sum_{i=1}^{\mathtt{d}} \ell_i \mathtt{m}_i \neq 0 \quad \text{for all } \ell \in \mathbb{Z}^\mathtt{d} \text{ with } 0 < |\ell| := \sum_{i=1}^{\mathtt{d}} |\ell_i| \leq \mathtt{L}, \tag{1.8}$$

where $\mathtt{d}$ is the cardinality of $\mathcal{S}$.

Our main result is the following:

Theorem 1.2. *Fix a positive integer* $\mathtt{d} \geq 2$ *and a sufficiently large* $\mathtt{L} \in \mathbb{N}$. *Assume that* $\mathcal{S}_0$ *has cardinality* $\mathtt{d}$ *and is* $\mathtt{L}$*-generic. Then there exists* $\varepsilon_* > 0$ *such that for any* $\varepsilon \in (0, \varepsilon_*)$ *there exists a positive measure Cantor-like set* $\mathcal{I} \subset (\varepsilon/2, \varepsilon)^\mathtt{d}$ *such that the torus* $\mathcal{T}^I_{\mathcal{S}_0}$ *with* $I = (I_j)_{j \in \mathcal{S}_0} \in \mathcal{I}$ *has the following properties:*

(1) (Long-time instability of $\mathcal{S}_0$-gap solutions in $H^s(\mathbb{T}^2, \mathbb{C})$ for $0 < s < 1$) *For any* $s \in (0,1)$, $\delta > 0$ *small enough, and* $K > 1$ *sufficiently large, there exists a smooth solution* $u(t)$ *of* (2D-NLS), $u : \mathbb{R} \to \bigcap_{n \geq 0} H^n(\mathbb{T}^2, \mathbb{C})$, *and a time* $0 < T \leq e^{(K/\delta)^\beta}$ *such that*
$$\operatorname{dist}(u(0), \mathcal{T}^I_{\mathcal{S}_0})_{H^s(\mathbb{T}^2)} \leq \delta \quad \textit{and} \quad \|u(T)\|_{H^s(\mathbb{T}^2)} \geq K.$$
Here the exponent $\beta > 1$ *can be chosen independently of* K, δ. *In particular, the* $\mathcal{S}_0$*-gap solutions in* $\mathcal{T}^I_{\mathcal{S}_0}$ *are Lyapunov* (*orbitally*) *unstable.*

(2) (Long-time instability of $\mathcal{S}_0$-gap solutions in $H^s(\mathbb{T}^2, \mathbb{C})$ for $s > 1$) *For any* $s > 1$ *and any* $K > 1$ *sufficiently large, there exists a smooth solution* $u : \mathbb{R} \to \bigcap_{n \geq 0} H^n(\mathbb{T}^2, \mathbb{C})$ *of* (2D-NLS) *and a time* $0 < T \leq e^{K^\sigma}$ *such that*
$$\operatorname{dist}(u(0), \mathcal{T}^I_{\mathcal{S}_0})_{L^2(\mathbb{T}^2)} \leq K^{-\sigma'} \quad \textit{and} \quad \|u(T)\|_{H^s(\mathbb{T}^2)} \geq K \|u(0)\|_{H^s(\mathbb{T}^2)}.$$
Here $\sigma, \sigma' > 0$ *depend on* s, *but not on* K. *Note that* $\operatorname{dist}(u(0), \mathcal{T}^I_{\mathcal{S}_0})_{H^s(\mathbb{T}^2)}$ *might not be small.*

1.4. Comments and remarks on Theorem 1.2

(1) The relative measure of the set $\mathcal{I}$ of admissible actions can be taken as close to 1 as desired. Indeed, by taking smaller ε_*, the relative measure satisfies
$$|1 - \operatorname{meas}(\mathcal{I})| \leq C \varepsilon_*^\kappa$$
for some constant $C > 0$ and $0 < \kappa < 1$ independent of $\varepsilon_* > 0$. The genericity condition on the set $\mathcal{S}_0$ and the actions $(I_\mathtt{m})_{\mathtt{m} \in \mathcal{S}_0} \in \mathcal{I}$ ensure that the *linearized* dynamics around

the resulting torus $\mathcal{T}^I_{\mathcal{S}_0}$ is stable for the perturbations we need to induce the nonlinear instability. In fact, a subset of those tori is even linearly stable for much more general perturbations as we remark below.

(2) *Why does the finite gap solution need to be small?* To prove Theorem 1.2 we need to analyze the linearization of equation (2D-NLS) at the finite gap solution (see Section 4). Roughly speaking, this leads to a Schrödinger equation with a quasi-periodic potential. Luckily, such operators can be *reduced* to constant coefficients via a KAM scheme. This is known as *reducibility theory* which allows one to construct a change of variables that transforms the linearized operator into an essentially constant coefficient diagonal one. This KAM scheme was carried out in [43], and requires the quasi-periodic potential, given by the finite gap solution here, to be small for the KAM iteration to converge. That being said, we suspect a similar result to be true for nonsmall finite gap solutions.

(3) To put the complexity of this result in perspective, it is instructive to compare it with the stability result in [43]. In that paper, it is shown that a proper *subset* $\mathcal{I}' \subset \mathcal{I}$ of the tori considered in Theorem 1.2 are Lyapunov stable in H^s, $s > 1$, but for shorter time scales than those considered in this theorem. More precisely, all orbits that are initially δ-close to $\mathcal{T}^I_{\mathcal{S}_0}$ in H^s stay $C\delta$-close for some fixed $C > 0$ for time scales $t \sim \delta^{-2}$. The same stability result (with a completely identical proof) holds if we replace H^s by $\mathcal{F}\ell_1$ norm (functions whose Fourier series is in ℓ^1). In fact, by trivially modifying the proof, one could also prove stability on the δ^{-2} time scale in $\mathcal{F}\ell_1 \cap H^s$ for $0 < s < 1$. What this means is that the solutions in the first part of Theorem 1.2 remain within $C\delta$ of $\mathcal{T}^I_{\mathcal{S}_0}$ up to times $\sim \delta^{-2}$ but can diverge vigorously afterwards at much longer time scales.

It is also worth mentioning that the complementary subset $\mathcal{I} \setminus \mathcal{I}'$ has a positive measure subset where tori are linearly unstable since they possess a finite set of modes that exhibit hyperbolic behavior. In principle, hyperbolic directions are good for instability, but they are not useful for our purposes since they live at very low frequencies, and hence cannot be used (at least not by themselves alone) to produce a substantial growth of Sobolev norms. We avoid dealing with these linearly unstable directions by restricting our solution to an invariant subspace on which these modes are at rest.[3]

(4) The growth in part (1) of the theorem is the result of the so-called *inverse cascade of mass* from high frequencies towards smaller ones, whereas the growth for $s > 1$ in part (2) is the result of a *forward cascade of kinetic energy* from low to high frequencies. Both phenomena are predicted by the physical theory of *wave turbulence* but their rigorous justification is highly nontrivial from a mathematical viewpoint as we discussed earlier. For part (1), initially the mass of the perturbation is concentrated on the "high frequency set" Λ_{03} in Theorem 7.3, but becomes concentrated on the "low frequency set" $\Lambda_{0,\mathtt{g}-1}$ at time T. This leads to the inflation of the H^s norm for $0 < s < 1$ (cf. (7.5)) whereas the Sobolev norms for $s > 1$ actually contract. In contrast, in part (2), the initial kinetic energy

[3]We expect that such hyperbolic directions should imply a transverse instability result similar to the one obtained by Rousset and Tzvetkov [51,52] for solitary waves.

of the perturbation is concentrated on the set $\Lambda_{0,\mathtt{g}-1}$ and ends up being concentrated on the "high frequency set" Λ_{03}, which yields the growth of Sobolev norms for $s > 1$. It is here that the dependence of the solution on s starts to make a difference in the proof (cf. Sections 7 and 8).

(5) It is expected that a similar statement to the first part of Theorem 1.2 is also true for $s > 1$. This would be a stronger instability compared to that in the second part (for which the initial perturbation is small in L^2 but not in H^s). Nevertheless, this case cannot be tackled with the techniques considered in this paper. Indeed, one of the key points in the proof is to perform a (partial) Birkhoff normal form up to order 4 around the finite gap solution. The terms which lead to the instabilities in Theorem 1.2 are quasi-resonant instead of being completely resonant. Working in the H^s topology with $s \in (0, 1)$, such terms can be considered completely resonant with little error on the time scales where instability happens. However, this cannot be done for $s > 1$, for which one might be able to eliminate those terms by a higher order normal form ($s > 1$ gives a stronger topology and can thus handle worse small divisors). This would mean that one needs other resonant terms to achieve growth of Sobolev norms. The same difficulties were encountered in [31] to prove the instability of the plane waves of (2D-NLS).

(6) The first part of the result of Theorem 1.2 can be interpreted as a *strong* form of Lyapunov instability in H^s norm ($0 < s < 1$) of the tori $\mathcal{T}^I_{\mathcal{S}_0}$ where the $\mathcal{S}_0$-gap solutions are supported. Indeed, for an invariant subset X of the phase space, being Lyapunov stable means that for all $\epsilon > 0$ there exists $\delta > 0$ such that all solutions that are δ-close to X at time $t = 0$ stay ϵ-close to X for all times. Thus, Lyapunov instability of X means that there exists $K > 0$ such that for all $\delta > 0$ there exist a solution $u(t)$ and a time T such that $\mathrm{dist}(u(0), X) < \delta$ and $\mathrm{dist}(u(T), X) > K$. In the first part of Theorem 1.2, we prove that for $X = \mathcal{T}^I_{\mathcal{S}_0}$ such an instability property holds true for all $\delta > 0$ and **for all** $K > 0$ in H^s norm (with $s \in (0, 1)$). Thus, a stronger form of instability holds: one can start as close to $\mathcal{T}^I_{\mathcal{S}_0}$ as desired but still end up *as far as desired* from $\mathcal{T}^I_{\mathcal{S}_0}$ after some time $T = T(K, \delta) > 0$.

(7) For finite-dimensional Hamiltonian dynamical systems, proving Lyapunov instability for quasi-periodic Diophantine elliptic (or maximal-dimensional Lagrangian) tori is an extremely difficult task. Actually all the results obtained [30, 59] deal with C^r or C^∞ Hamiltonians, and not a single example of such instability is known for analytic Hamiltonian systems. In fact, there are no results of instabilities in the vicinity of nonresonant elliptic critical points or periodic orbits for analytic Hamiltonian systems (see [17, 18, 34] for results on the C^∞ topology). The present paper proves the existence of unstable Diophantine elliptic tori in an analytic infinite-dimensional Hamiltonian system. Obtaining such instabilities in infinite dimensions is, in some sense, easier: having infinite dimensions gives "more room" for instabilities.

(8) It is well known that many Hamiltonian PDEs possess quasiperiodic invariant tori [1,3–5,8,19,21,38,47–50,57,58]. Most of those tori are normally elliptic and thus linearly stable. It is widely expected that the behavior given by Theorem 1.2 also arises in the

neighborhoods of (many of) those tori. Nevertheless, it is not clear how to apply the techniques of the present paper to these settings.

1.5. Scheme of the proof

Let us explain the main steps to prove Theorem 1.2.

(1) *Analysis of the* 1*-dimensional cubic Schrödinger equation.* We express the 1-dimensional cubic NLS in terms of the Birkhoff coordinates. We need a quite precise knowledge of the Birkhoff map (see Theorem 3.1). In particular, we need that it "behaves well" in ℓ^1. This is done in [41] and summarized in Section 3. In Birkhoff coordinates, the finite gap solutions are supported in a finite set of variables. We use such coordinates to express the Hamiltonian (1.1) in a more convenient way.

(2) *Reducibility of the* 2*-dimensional cubic NLS around a finite gap solution.* We reduce the linearization of the vector field around the finite gap solutions to a constant coefficients diagonal vector field. This is done in [43] and explained in Section 4. In Theorem 4.3 we give conditions leading to full reducibility. In effect, this transforms the linearized operator around the finite gap into a constant coefficient diagonal (in Fourier space) operator, with eigenvalues $\{\Omega_{\vec{j}}\}_{\vec{j}\in\mathbb{Z}^2\setminus\mathcal{S}_0}$. We give the asymptotics of these eigenvalues in Theorem 4.4, which roughly speaking look like

$$\Omega_{\vec{j}} = |\vec{j}|^2 + O(J^{-2}) \tag{1.9}$$

for frequencies $\vec{j} = (m, n)$ satisfying $|m|, |n| \sim J$. This seemingly harmless $O(J^{-2})$ correction to the unperturbed Laplacian eigenvalues is sharp and will be responsible for the restriction to $s \in (0, 1)$ in the first part of Theorem 1.2 as we shall explain below.

(3) *Degree* 3 *Birkhoff normal form around the finite gap solution.* This is done in [43], but we shall need more precise information from this normal form that will be crucial for Steps (5) and (6) below. This is done in Section 5 (see Theorem 5.2).

(4) *Partial normal form of degree* 4. We remove all degree 4 monomials which are not (too close to) resonant. This is done in Section 6, and leaves us with a Hamiltonian with (close to) resonant degree 4 terms plus a higher degree part which will be treated as a remainder in our construction.

(5) We follow the paradigm set forth in [12, 28] to construct solutions to the truncated Hamiltonian consisting of the (close to) resonant degree 4 terms isolated above, and then to the full Hamiltonian by an approximation argument. This construction will be done at frequencies $\vec{j} = (m, n)$ such that $|m|, |n| \sim J$ with J very large, and for which the dynamics is effectively given by the following system of ODEs:

$$\begin{cases} i\dot{a}_{\vec{j}} = -|a_{\vec{j}}|^2 a_{\vec{j}} + \sum_{\mathcal{R}(\vec{j})} a_{\vec{j}_1}\overline{a_{\vec{j}_2}} a_{\vec{j}_3} e^{i\Gamma t}, \\ \mathcal{R}(\vec{j}) := \{(\vec{j}_1, \vec{j}_2, \vec{j}_3) \in \mathbb{Z}^2 \setminus \mathcal{S}_0 : \vec{j}_1, \vec{j}_3 \neq \vec{j},\ \vec{j}_1 - \vec{j}_2 + \vec{j}_3 = \vec{j}, \\ \qquad\qquad\qquad\qquad\qquad\qquad\qquad |\vec{j}_1|^2 - |\vec{j}_2|^2 + |\vec{j}_3|^2 = |\vec{j}|^2\}, \\ \Gamma := \Omega_{\vec{j}_1} - \Omega_{\vec{j}_2} + \Omega_{\vec{j}_3} - \Omega_{\vec{j}}. \end{cases}$$

We remark that the conditions of the set $\mathcal{R}(\vec{j})$ are essentially equivalent to saying that $(\vec{j}_1, \vec{j}_2, \vec{j}_3, \vec{j})$ form a rectangle in $\mathbb{Z}^2$. Also note that by the asymptotics of $\Omega_{\vec{j}}$ mentioned above in (1.9), one obtains $\Gamma = O(J^{-2})$ if all the frequencies involved are in $\mathcal{R}(\vec{j})$ and satisfy $|m|, |n| \sim J$. The idea now is to reduce this system to a finite-dimensional system called the "Toy Model", which is tractable enough for us to construct a solution that cascades energy. An obstruction to this plan is the presence of the oscillating factor $e^{\mathrm{i}\Gamma t}$ for which Γ is not zero (in contrast to [12]) but rather $O(J^{-2})$. The only way to proceed with this reduction is to approximate $e^{\mathrm{i}\Gamma t} \sim 1$ which is only possible provided $J^{-2}T \ll 1$. The solution coming from the Toy Model is supported on a finite number of modes $\vec{j} \in \mathbb{Z}^2 \setminus \mathcal{S}_0$ satisfying $|j| \sim J$, and the time it takes for the energy to diffuse across its modes is $T \sim O(\nu^{-2})$ where ν is the characteristic size of the modes in ℓ^1 norm. Requiring the solution to be initially close in H^s to the finite gap solution would necessitate that $\nu J^s \lesssim \delta$, which gives $T \gtrsim_\delta J^{-2s}$, and hence the condition $J^{-2}T \ll 1$ translates into $s < 1$. This explains the restriction to $s < 1$ in the first part of Theorem 1.2. If we only require our solutions to be close to the finite gap solution in L^2, then no such restriction on ν is needed, and hence there is no restriction on s beyond being $s > 0$ and $s \neq 1$, which is the second part of the theorem.

This analysis is done in Sections 7 and 8. In the former, we perform the reduction to the effective degree 4 Hamiltonian taking into account all the changes of variables performed in the previous sections; while in Section 8 we perform the above approximation argument allowing us to shadow the Toy Model solution mentioned above with a solution of (2D-NLS) exhibiting the needed norm growth, thus completing the proof of Theorem 1.2.

In Appendix B we give a list of notations and parameters used throughout the paper.

2. Notation and functional setting

2.1. *Notation*

For a complex number z, it is often convenient to use the notation

$$z^\sigma = \begin{cases} z & \text{if } \sigma = +1, \\ \bar{z} & \text{if } \sigma = -1. \end{cases}$$

For any subset $\Gamma \subset \mathbb{Z}^2$, we denote by $h^s(\Gamma)$ the set of sequences $(a_{\vec{j}})_{\vec{j} \in \Gamma}$ with norm

$$\|a\|_{h^s(\Gamma)} = \Big(\sum_{\vec{j} \in \Gamma} \langle \vec{j} \rangle^{2s} |a_{\vec{j}}|^2 \Big)^{1/2} < \infty.$$

Our phase space will be obtained by an appropriate linearization around the finite gap solution with $\mathtt{d}$ frequencies/actions. For a finite set $\mathcal{S}_0 \subset \mathbb{Z} \times \{0\}$ of $\mathtt{d}$ elements, we consider the phase space $\mathcal{X} = (\mathbb{C}^{\mathtt{d}} \times \mathbb{T}^{\mathtt{d}}) \times \ell^1(\mathbb{Z}^2 \setminus \mathcal{S}_0) \times \ell^1(\mathbb{Z}^2 \setminus \mathcal{S}_0)$. The first part $(\mathbb{C}^{\mathtt{d}} \times \mathbb{T}^{\mathtt{d}})$ corresponds to the finite gap sites in action-angle coordinates, whereas $\ell^1(\mathbb{Z}^2 \setminus \mathcal{S}_0) \times \ell^1(\mathbb{Z}^2 \setminus \mathcal{S}_0)$ corresponds to the remaining orthogonal sites in frequency space. We shall often denote the ℓ^1 norm by $\|\cdot\|_1$. We shall denote variables on $\mathcal{X}$ by

$$\mathcal{X} \ni (\mathcal{Y}, \theta, \mathbf{a}) : \quad \mathcal{Y} \in \mathbb{C}^{\mathrm{d}},\ \theta \in \mathbb{T}^{\mathrm{d}},\ \mathbf{a} = (a, \bar{a}) \in \ell^1(\mathbb{Z}^2 \setminus \mathcal{S}_0) \times \ell^1(\mathbb{Z}^2 \setminus \mathcal{S}_0).$$

We shall use multi-index notation to write monomials like $\mathcal{Y}^l$ and $\mathfrak{m}_{\alpha,\beta} = a^\alpha \bar{a}^\beta$ where $l \in \mathbb{N}^{\mathrm{d}}$ and $\alpha, \beta \in \mathbb{N}^{\mathbb{Z}^2 \setminus \mathcal{S}_0}$. Oftentimes, we will abuse notation, and simply write $\mathbf{a} \in \ell^1$ to mean $\mathbf{a} = (a, \bar{a}) \in \ell^1(\mathbb{Z}^2 \setminus \mathcal{S}_0) \times \ell^1(\mathbb{Z}^2 \setminus \mathcal{S}_0)$, and $\|\mathbf{a}\|_1 = \|a\|_{\ell^1(\mathbb{Z}^2 \setminus \mathcal{S}_0)}$.

Definition 2.1. For a monomial of the form $e^{\mathrm{i}\ell\cdot\theta}\mathcal{Y}^l \mathfrak{m}_{\alpha,\beta}$, we define its degree to be $2|l| + |\alpha| + |\beta| - 2$, where the modulus of a multi-index is given by its ℓ^1 norm.

2.2. *Regular Hamiltonians*

Given a Hamiltonian function $F(\mathcal{Y}, \theta, \mathbf{a})$ on the phase space $\mathcal{X}$, we associate to it the Hamiltonian vector field

$$X_F := \{-\partial_\theta F, \partial_{\mathcal{Y}} F, -\mathrm{i}\partial_{\bar{a}} F, \mathrm{i}\partial_a F\},$$

where we have used the standard complex notation to denote the Fréchet derivatives of F with respect to the variable $\mathbf{a} \in \ell^1$.

We will often need to complexify the variable $\theta \in \mathbb{T}^{\mathrm{d}}$ into the domain

$$\mathbb{T}^{\mathrm{d}}_\rho := \{\theta \in \mathbb{C}^{\mathrm{d}} : \mathrm{Re}(\theta) \in \mathbb{T}^{\mathrm{d}},\ |\mathrm{Im}(\theta)| \le \rho\}$$

and consider vector fields which are functions

$$\mathbb{C}^{\mathrm{d}} \times \mathbb{T}^{\mathrm{d}}_\rho \times \ell^1 \to \mathbb{C}^{\mathrm{d}} \times \mathbb{C}^{\mathrm{d}} \times \ell^1, \quad (\mathcal{Y}, \theta, \mathbf{a}) \mapsto (X^{(\mathcal{Y})}, X^{(\theta)}, X^{(a)}, X^{(\bar{a})}),$$

which are analytic in $\mathcal{Y}, \theta, \mathbf{a}$. Our vector fields will be defined on the domain

$$D(\rho, r) := \mathbb{T}^{\mathrm{d}}_\rho \times D(r) \quad \text{where} \quad D(r) := \{|\mathcal{Y}| \le r^2,\ \|\mathbf{a}\|_1 \le r\}. \tag{2.1}$$

For the vector fields, we use the norm

$$\| X \|_r := |X^{(\theta)}| + \frac{|X^{(\mathcal{Y})}|}{r^2} + \frac{\|X^{(a)}\|_1}{r} + \frac{\|X^{(\bar{a})}\|_1}{r}.$$

All Hamiltonians F considered in this article are analytic, real-valued and can be expanded in Taylor Fourier series which are well defined and pointwise absolutely convergent,

$$F(\mathcal{Y}, \theta, \mathbf{a}) = \sum_{\alpha,\beta \in \mathbb{N}^{\mathbb{Z}^2 \setminus \mathcal{S}_0},\, \ell \in \mathbb{Z}^{\mathrm{d}},\, l \in \mathbb{N}^{\mathrm{d}}} F_{\alpha,\beta,l,\ell}\, e^{\mathrm{i}\ell\cdot\theta}\mathcal{Y}^l \mathfrak{m}_{\alpha,\beta}. \tag{2.2}$$

Correspondingly we expand vector fields in Taylor Fourier series (again well defined and pointwise absolutely convergent)

$$X^{(v)}(\mathcal{Y}, \theta, \mathbf{a}) = \sum_{\alpha,\beta \in \mathbb{N}^{\mathbb{Z}^2 \setminus \mathcal{S}_0},\, \ell \in \mathbb{Z}^{\mathrm{d}},\, l \in \mathbb{N}^{\mathrm{d}}} X^{(v)}_{\alpha,\beta,l,\ell}\, e^{\mathrm{i}\ell\cdot\theta}\mathcal{Y}^l \mathfrak{m}_{\alpha,\beta},$$

where v denotes the components $\theta_i, \mathcal{Y}_i$ for $1 \le i \le \mathtt{d}$ or $a_{\vec{j}}, \bar{a}_{\vec{j}}$ for $\vec{j} \in \mathbb{Z}^2 \setminus \mathcal{S}_0$.

To a vector field we associate its majorant

$$\underline{X}_{\rho}^{(v)}[\mathcal{Y}, \mathbf{a}] := \sum_{\ell \in \mathbb{Z}^{\mathtt{d}},\, l \in \mathbb{N}^{\mathtt{d}},\, \alpha, \beta \in \mathbb{N}^{\mathbb{Z}^2}} |X^{(v)}_{\alpha,\beta,l,\ell}| e^{\rho|\ell|} \mathcal{Y}^l \mathfrak{m}_{\alpha,\beta}$$

and require that this is an analytic map on $D(r)$. Such a vector field is called *majorant analytic*. Since Hamiltonian functions are defined modulo constants, we give the following definition of the *norm* of F:

$$|F|_{\rho,r} := \sup_{(\mathcal{Y},\mathbf{a}) \in D(r)} \left\| \underline{(X_F)}_{\rho} \right\|_r .$$

Note that the norm $|\cdot|_{\rho,r}$ controls $|\cdot|_{\rho',r'}$ whenever $\rho' < \rho$, $r' < r$.

Finally, we will also consider Hamiltonians $F(\lambda; \theta, a, \bar{a}) \equiv F(\lambda)$ depending on an external parameter $\lambda \in \mathcal{O} \subset \mathbb{R}^{\mathtt{d}}$. For those, we define the *inhomogeneous Lipschitz* norm

$$|F|^{\mathcal{O}}_{\rho,r} := \sup_{\lambda \in \mathcal{O}} |F(\lambda)|_{\rho,r} + \sup_{\lambda_1 \neq \lambda_2 \in \mathcal{O}} \frac{|F(\lambda_1) - F(\lambda_2)|_{\rho,r}}{|\lambda_1 - \lambda_2|}.$$

2.3. Commutation rules

Given two Hamiltonians F and G, we define their Poisson bracket as $\{F, G\} := \mathrm{d}F(X_G)$; in coordinates,

$$\{F, G\} = -\partial_{\mathcal{Y}} F \cdot \partial_\theta G + \partial_\theta F \cdot \partial_{\mathcal{Y}} G + \mathrm{i}\Big(\sum_{\vec{\jmath} \in \mathbb{Z}^2 \setminus \mathcal{S}_0} \partial_{\bar{a}_{\vec{\jmath}}} F \partial_{a_{\vec{\jmath}}} G - \partial_{a_{\vec{\jmath}}} F \partial_{\bar{a}_{\vec{\jmath}}} G \Big).$$

Given $\alpha, \beta \in \mathbb{N}^{\mathbb{Z}^2 \setminus \mathcal{S}_0}$ we denote $\mathfrak{m}_{\alpha,\beta} := a^\alpha \bar{a}^\beta$. To the monomial $e^{\mathrm{i}\ell\cdot\theta} \mathcal{Y}^l \mathfrak{m}_{\alpha,\beta}$ with $\ell \in \mathbb{Z}^{\mathtt{d}}$, $l \in \mathbb{N}^{\mathtt{d}}$ we associate various numbers. We denote

$$\eta(\alpha, \beta) := \sum_{\vec{\jmath} \in \mathbb{Z}^2 \setminus \mathcal{S}_0} (\alpha_{\vec{\jmath}} - \beta_{\vec{\jmath}}), \quad \eta(\ell) := \sum_{i=1}^{\mathtt{d}} \ell_i. \tag{2.3}$$

We also associate to $e^{\mathrm{i}\ell\cdot\theta} \mathcal{Y}^l \mathfrak{m}_{\alpha,\beta}$ the quantities $\pi(\alpha, \beta) = (\pi_x, \pi_y)$ and $\pi(\ell)$ defined by

$$\pi(\alpha, \beta) = \begin{bmatrix} \pi_x(\alpha, \beta) \\ \pi_y(\alpha, \beta) \end{bmatrix} = \sum_{\vec{\jmath} = (m,n) \in \mathbb{Z}^2 \setminus \mathcal{S}_0} \begin{bmatrix} m \\ n \end{bmatrix} (\alpha_{\vec{\jmath}} - \beta_{\vec{\jmath}}), \quad \pi(\ell) = \sum_{i=1}^{\mathtt{d}} \mathtt{m}_i \ell_i. \tag{2.4}$$

The above quantities are associated with the mass $\mathcal{M}$ and momentum $\mathcal{P} = (\mathcal{P}_x, \mathcal{P}_y)$ functionals given by

$$\begin{aligned} \mathcal{M} &:= \sum_{i=1}^{\mathtt{d}} \mathcal{Y}_i + \sum_{\vec{\jmath} \in \mathbb{Z}^2 \setminus \mathcal{S}_0} |a_{\vec{\jmath}}|^2, \\ \mathcal{P}_x &:= \sum_{i=1}^{\mathtt{d}} \mathtt{m}_i \mathcal{Y}_i + \sum_{(m,n) \in \mathbb{Z}^2 \setminus \mathcal{S}_0} m |a_{(m,n)}|^2, \\ \mathcal{P}_y &:= \sum_{(m,n) \in \mathbb{Z}^2 \setminus \mathcal{S}_0} n |a_{(m,n)}|^2, \end{aligned} \tag{2.5}$$

via the following commutation rules: given a monomial $e^{\mathrm{i}\ell\cdot\theta} \mathcal{Y}^l \mathfrak{m}_{\alpha,\beta}$,

$$\{\mathcal{M}, e^{\mathrm{i}\ell\cdot\theta} y^l \mathfrak{m}_{\alpha,\beta}\} = \mathrm{i}(\eta(\alpha,\beta)+\eta(\ell))e^{\mathrm{i}\ell\cdot\theta} y^l \mathfrak{m}_{\alpha,\beta},$$
$$\{\mathcal{P}_x, e^{\mathrm{i}\ell\cdot\theta} y^l \mathfrak{m}_{\alpha,\beta}\} = \mathrm{i}(\pi_x(\alpha,\beta)+\pi(\ell))e^{\mathrm{i}\ell\cdot\theta} y^l \mathfrak{m}_{\alpha,\beta},$$
$$\{\mathcal{P}_y, e^{\mathrm{i}\ell\cdot\theta} y^l \mathfrak{m}_{\alpha,\beta}\} = \mathrm{i}\pi_y(\alpha,\beta)e^{\mathrm{i}\ell\cdot\theta} y^l \mathfrak{m}_{\alpha,\beta}.$$

Remark 2.2. An analytic Hamiltonian function $\mathcal{F}$ (expanded as in (2.2)) commutes with the mass $\mathcal{M}$ and the momentum $\mathcal{P}$ if and only if the following *selection rules* on its coefficients hold:

$$\{\mathcal{F},\mathcal{M}\}=0 \iff \mathcal{F}_{\alpha,\beta,l,\ell}\,(\eta(\alpha,\beta)+\eta(\ell))=0,$$
$$\{\mathcal{F},\mathcal{P}_x\}=0 \iff \mathcal{F}_{\alpha,\beta,l,\ell}\,(\pi_x(\alpha,\beta)+\pi(\ell))=0,$$
$$\{\mathcal{F},\mathcal{P}_y\}=0 \iff \mathcal{F}_{\alpha,\beta,l,\ell}\,(\pi_y(\alpha,\beta))=0,$$

where $\eta(\alpha,\beta), \eta(\ell)$ are defined in (2.3) and $\pi(\alpha,\beta), \pi(\ell)$ are defined in (2.4).

Definition 2.3. We will denote by $\mathcal{A}_{\rho,r}$ the set of all real-valued Hamiltonians of the form (2.2) with finite $|\cdot|_{\rho,r}$ norm and which Poisson commute with $\mathcal{M}$, $\mathcal{P}$. Given a compact set $\mathcal{O} \subset \mathbb{R}^{\mathrm{d}}$, we denote by $\mathcal{A}^{\mathcal{O}}_{\rho,r}$ the Banach space of Lipschitz maps $\mathcal{O} \to \mathcal{A}_{\rho,r}$ with the norm $|\cdot|^{\mathcal{O}}_{\rho,r}$.

From now on, all our Hamiltonians will belong to $\mathcal{A}_{\rho,r}$ for some $\rho, r > 0$.

3. Adapted variables and Hamiltonian formulation

3.1. Fourier expansion and phase shift

Let us start by expanding u in Fourier coefficients,

$$u(x,y,t) = \sum_{\vec{j}=(m,n)\in\mathbb{Z}^2} u_{\vec{j}}(t)e^{\mathrm{i}(mx+ny)}.$$

Then the Hamiltonian H_0 introduced in (1.1) can be written as

$$H_0(u) = \sum_{\vec{j}\in\mathbb{Z}^2} |\vec{j}|^2|u_{\vec{j}}|^2 + \frac{1}{2}\sum_{\substack{\vec{j}_i\in\mathbb{Z}^2\\ \vec{j}_1-\vec{j}_2+\vec{j}_3-\vec{j}_4=0}} u_{\vec{j}_1}\bar{u}_{\vec{j}_2}u_{\vec{j}_3}\bar{u}_{\vec{j}_4}$$
$$= \sum_{\vec{j}\in\mathbb{Z}^2} |\vec{j}|^2|u_{\vec{j}}|^2 - \frac{1}{2}\sum_{\vec{j}\in\mathbb{Z}^2}|u_{\vec{j}}|^4 + 2\overbrace{\Big(\sum_{\vec{j}\in\mathbb{Z}^2}|u_{\vec{j}}|^2\Big)^2}^{M(u)^2} + \frac{1}{2}\sideset{}{^\star}\sum_{\substack{\vec{j}_i\in\mathbb{Z}^2\\ \vec{j}_1-\vec{j}_2+\vec{j}_3-\vec{j}_4=0}} u_{\vec{j}_1}\bar{u}_{\vec{j}_2}u_{\vec{j}_3}\bar{u}_{\vec{j}_4}$$

where $\sum^\star$ means the sum over the quadruples $\vec{j}_i$ such that $\{\vec{j}_1,\vec{j}_3\}\neq\{\vec{j}_2,\vec{j}_4\}$.

Since the mass $M(u)$ in (1.2) is a constant of motion, we make a trivial phase shift and consider an equivalent Hamiltonian $H(u) = H_0(u) - M(u)^2$,

$$H(u) = \int_{\mathbb{T}^2} |\nabla u(x,y)|^2\,\mathrm{d}x\,\mathrm{d}y + \frac{1}{2}\int_{\mathbb{T}^2}|u(x,y)|^4\,\mathrm{d}x\,\mathrm{d}y - M(u)^2 \tag{3.1}$$

corresponding to the Hamilton equation

$$\mathrm{i}\partial_t u = -\Delta u + |u|^2 u - 2M(u)u, \quad (x, y) \in \mathbb{T}^2. \tag{3.2}$$

Clearly the solutions of (3.2) differ from the solutions of (2D-NLS) only by a phase shift.[4] Then

$$H(u) = \sum_{\vec{j}\in\mathbb{Z}^2} |\vec{j}|^2 |u_{\vec{j}}|^2 - \frac{1}{2}\sum_{\vec{j}\in\mathbb{Z}^2} |u_{\vec{j}}|^4 + \frac{1}{2}\sum^{\star}_{\substack{\vec{j}_i\in\mathbb{Z}^2\\ \vec{j}_1-\vec{j}_2+\vec{j}_3-\vec{j}_4=0}} u_{\vec{j}_1}\bar{u}_{\vec{j}_2}u_{\vec{j}_3}\bar{u}_{\vec{j}_4}. \tag{3.3}$$

3.2. The Birkhoff map for the 1D cubic NLS

We devote this section to gathering some properties of the Birkhoff map for the integrable 1D NLS equation. These will be used to write the Hamiltonian (3.3) in a more convenient way. The main reference for this section is [41].

We shall denote by $B^s(r)$ the ball of radius r and center 0 in the topology of $h^s \equiv h^s(\mathbb{Z})$.

Theorem 3.1. *There exist $r_* > 0$ and a symplectic, real analytic map Φ with $\mathrm{d}\Phi(0) = \mathbb{I}$ such that for all $s \geq 0$ one has the following:*

(i) $\Phi : B^s(r_*) \to h^s$. *More precisely, there exists a constant $C > 0$ such that for all $0 \leq r \leq r_*$,*

$$\sup_{\|q\|_{h^s}\leq r} \|(\Phi - \mathbb{I})(q)\|_{h^s} \leq C r^3.$$

The same estimate holds for $\Phi^{-1} - \mathbb{I}$ or with h^s replaced by ℓ^1.

(ii) *Moreover, if $q \in h^s$ for $s \geq 1$, then Φ introduces local Birkhoff coordinates for* (1D-NLS) *in h^s as follows: the integrals of motion of* (1D-NLS) *are real analytic functions of the actions $I_j = |z_j|^2$ where $(z_j)_{j\in\mathbb{Z}} = \Phi(q)$. In particular, the Hamiltonian $H_{\text{1D-NLS}}(q) \equiv \int_{\mathbb{T}} |\partial_x q(x)|^2\,\mathrm{d}x - M(q)^2 + \frac{1}{2}\int_{\mathbb{T}} |q(x)|^4\,\mathrm{d}x$, the mass $M(q) := \int_{\mathbb{T}} |q(x)|^2\,\mathrm{d}x$ and the momentum $P(q) := -\int_{\mathbb{T}} \bar{q}(x)\mathrm{i}\partial_x q(x)\,\mathrm{d}x$ have the form*

$$\begin{aligned}(H_{\text{1D-NLS}} \circ \Phi^{-1})(z) &\equiv h_{\text{1D-NLS}}((|z_m|^2)_{m\in\mathbb{Z}})\\ &= \sum_{m\in\mathbb{Z}} m^2|z_m|^2 - \frac{1}{2}\sum_{m\in\mathbb{Z}} |z_m|^4 + O(|z|^6), \qquad (3.4)\\ (M \circ \Phi^{-1})(z) &= \sum_{m\in\mathbb{Z}} |z_m|^2,\\ (P \circ \Phi^{-1})(z) &= \sum_{m\in\mathbb{Z}} m|z_m|^2.\end{aligned}$$

[4] To show the equivalence we consider any solution $u(x, t)$ of (3.2) and consider the invertible map

$$u \mapsto v = u e^{-2\mathrm{i}M(u)t} \quad \text{with inverse} \quad v \mapsto u = v e^{2\mathrm{i}M(v)t}.$$

Then a direct computation shows that v solves (2D-NLS).

(iii) *Define the* (1D-NLS) *action-to-frequency map* $I \mapsto \alpha^{\text{1D-NLS}}(I)$ *by* $\alpha_m^{\text{1D-NLS}}(I) := \frac{\partial h_{\text{1D-NLS}}}{\partial I_m}$ *for* $m \in \mathbb{Z}$. *Then one has the asymptotic expansion*

$$\alpha_m^{\text{1D-NLS}}(I) = m^2 - I_m + \frac{\varpi_m(I)}{\langle m \rangle} \tag{3.5}$$

where $\varpi_m(I)$ *is at least quadratic in* I.

Proof. Item (i) is the main content of [41], where it is proved that the Birkhoff map is majorant analytic between some Fourier–Lebesgue spaces. Item (ii) is proved in [25]. Item (iii) is [36, Theorem 1.3]. ∎

Remark 3.2. Theorem 3.1 implies that all solutions of (1D-NLS) have Sobolev norms uniformly bounded in time (as it happens for other integrable systems, like KdV and Toda lattice, see e.g. [2,35]). On the contrary, the Szegő equation is an integrable system which exhibits growth of Sobolev norms [24].

3.3. Adapted variables

The aim of this section is to write the Hamiltonian (3.1), the mass M (1.2) and the momentum P (1.3) in the local variables around the finite gap solution corresponding to

$$\begin{cases} |z_m|^2 = I_m, & m \in \mathcal{S}_0, \\ z_m = 0, & m \in \mathbb{Z} \setminus \mathcal{S}_0. \end{cases}$$

We start from the Hamiltonian in Fourier coordinates (3.3), and set

$$q_m := u_{(m,0)} \quad \text{if } m \in \mathbb{Z}, \qquad a_{\vec{\jmath}} = u_{\vec{\jmath}} \quad \text{if } \vec{\jmath} = (m,n) \in \mathbb{Z}^2,\ n \neq 0.$$

We rewrite the Hamiltonian accordingly in increasing degree in a, obtaining

$$\begin{aligned}
H(q,a) = {} & \sum_{m\in\mathbb{Z}} m^2 |q_m|^2 - \frac{1}{2}\sum_{m\in\mathbb{Z}} |q_m|^4 + \frac{1}{2} \sideset{}{^\star}\sum_{\substack{m_i\in\mathbb{Z} \\ m_1-m_2+m_3-m_4=0}} q_{m_1}\bar{q}_{m_2}q_{m_3}\bar{q}_{m_4} \\
& + \sum_{\vec{\jmath}\in\mathbb{Z}^2\setminus\mathbb{Z}} |\vec{\jmath}|^2 |a_{\vec{\jmath}}|^2 + 2 \sideset{}{^\star}\sum_{\substack{\vec{\jmath}_i=(m_i,n_i),\, i=3,4,\, n_i\neq 0 \\ m_1-m_2+m_3-m_4=0 \\ n_3-n_4=0}} q_{m_1}\bar{q}_{m_2}a_{\vec{\jmath}_3}\bar{a}_{\vec{\jmath}_4} \\
& + \operatorname{Re} \sum_{\substack{\vec{\jmath}_i=(m_i,n_i),\, i=2,4,\, n_i\neq 0 \\ m_1-m_2+m_3-m_4=0 \\ n_2+n_4=0}} \bar{q}_{m_1}a_{\vec{\jmath}_2}\bar{q}_{m_3}a_{\vec{\jmath}_4} + 2\operatorname{Re} \sum_{\substack{\vec{\jmath}_i=(m_i,n_i),\, i=2,3,4,\, n_i\neq 0 \\ m_1-m_2+m_3-m_4=0 \\ -n_2+n_3-n_4=0}} q_{m_1}\bar{a}_{\vec{\jmath}_2}a_{\vec{\jmath}_3}\bar{a}_{\vec{\jmath}_4} \\
& + \frac{1}{2} \sideset{}{^\star}\sum_{\substack{\vec{\jmath}_i=(m_i,n_i),\, i=1,2,3,4,\, n_i\neq 0 \\ \vec{\jmath}_1-\vec{\jmath}_2+\vec{\jmath}_3-\vec{\jmath}_4=0}} a_{\vec{\jmath}_1}\bar{a}_{\vec{\jmath}_2}a_{\vec{\jmath}_3}\bar{a}_{\vec{\jmath}_4} - \frac{1}{2}\sum_{\vec{\jmath}\in\mathbb{Z}^2\setminus\mathbb{Z}} |a_{\vec{\jmath}}|^4 \\
=: {} & H_{\text{1D-NLS}}(q) + H^{\text{II}}(q,a) + H^{\text{III}}(q,a) + H^{\text{IV}}(a).
\end{aligned}$$

Step 1. First we make the following change of coordinates, which amounts to introducing Birkhoff coordinates on the line $\mathbb{Z}\times\{0\}$. We set

$$\begin{aligned}&((z_m)_{m\in\mathbb{Z}},(a_{\vec{j}})_{\vec{j}\in\mathbb{Z}^2\setminus\mathbb{Z}})\mapsto((q_m)_{m\in\mathbb{Z}},(a_{\vec{j}})_{\vec{j}\in\mathbb{Z}^2\setminus\mathbb{Z}}),\\&(q_m)_{m\in\mathbb{Z}}=\Phi^{-1}((z_m)_{m\in\mathbb{Z}}),\quad a_{\vec{j}}=u_{\vec{j}},\quad \vec{j}\in\mathbb{Z}^2\setminus\mathbb{Z}.\end{aligned}$$

In those new coordinates, the Hamiltonian becomes

$$\begin{aligned}H(z,a)=\;&H_{\text{1D-NLS}}(\Phi^{-1}(z))+H^{\text{II}}(\Phi^{-1}(z),a)\\&+H^{\text{III}}(\Phi^{-1}(z),a)+H^{\text{IV}}(a),\end{aligned}$$

where

$$H_{\text{1D-NLS}}(\Phi^{-1}(z))=h_{\text{1D-NLS}}((|z_m|^2)_{m\in\mathbb{Z}}).$$

Step 2. Next, we go to action-angle coordinates only on the set

$$\mathcal{S}_0=\{\mathtt{m}_1,\dots,\mathtt{m}_d\}\subset\mathbb{Z}\times\{0\}$$

and rename z_m for $m\notin\mathcal{S}_0$ as $a_{(m,0)}$, as follows:

$$\begin{aligned}&(\mathcal{Y}_i,\theta_i,a_{\vec{j}})_{\substack{1\le i\le \mathtt{d}\\ \vec{j}\in\mathbb{Z}^2\setminus\mathcal{S}_0}}\mapsto(z_m,a_{\vec{j}})_{m\in\mathbb{Z},\,\vec{j}\in\mathbb{Z}^2\setminus\mathbb{Z}},\\
&\qquad z_{\mathtt{m}_i}=\sqrt{I_{\mathtt{m}_i}+\mathcal{Y}_i}\,e^{\mathrm{i}\theta_i},\qquad \mathtt{m}_i\in\mathcal{S}_0,\\
&\qquad z_m=a_{(m,0)},\qquad\qquad\qquad m\in\mathbb{Z}\setminus\mathcal{S}_0,\\
&\qquad a_{\vec{j}}=a_{\vec{j}},\qquad\qquad\qquad\quad \vec{j}\in\mathbb{Z}^2\setminus\mathbb{Z}.\end{aligned}$$

In those coordinates, the Hamiltonian becomes (using (3.4))

$$\mathcal{H}(\mathcal{Y},\theta,a)=\;h_{\text{1D-NLS}}\big(I_{\mathtt{m}_1}+\mathcal{Y}_1,\dots,I_{\mathtt{m}_\mathtt{d}}+\mathcal{Y}_\mathtt{d},(|a_{(m,0)}|^2)_{m\notin\mathcal{S}_0}\big) \tag{3.6}$$
$$+H^{\text{II}}\big(\Phi^{-1}(\sqrt{I_{\mathtt{m}_1}+\mathcal{Y}_1}\,e^{\mathrm{i}\theta_1},\dots,\sqrt{I_{\mathtt{m}_\mathtt{d}}+\mathcal{Y}_\mathtt{d}}\,e^{\mathrm{i}\theta_\mathtt{d}},(a_{(m,0)})_{m\notin\mathcal{S}_0}),(a_{(m,n)})_{n\neq0}\big) \tag{3.7}$$
$$+H^{\text{III}}\big(\Phi^{-1}(\sqrt{I_{\mathtt{m}_1}+\mathcal{Y}_1}\,e^{\mathrm{i}\theta_1},\dots,\sqrt{I_{\mathtt{m}_\mathtt{d}}+\mathcal{Y}_\mathtt{d}}\,e^{\mathrm{i}\theta_\mathtt{d}},(a_{(m,0)})_{m\notin\mathcal{S}_0}),(a_{(m,n)})_{n\neq0}\big) \tag{3.8}$$
$$+H^{\text{IV}}((a_{(m,n)})_{n\neq0}). \tag{3.9}$$

We first remark that $\mathcal{T}^I_{\mathcal{S}_0}$ is described in the $(\mathcal{Y},\theta,a)$ coordinates by $\mathcal{Y}=0$, $a=0$. Furthermore, it is proved in [43, Proposition 4.2] that a neighborhood of $(0,\theta,0)$ corresponds in the original variables to a neighborhood of the torus $\mathcal{T}^I_{\mathcal{S}_0}$; in particular

$$|\mathcal{Y}|\le r^2,\ \|a\|_{h^s(\mathbb{Z}^2\setminus\mathcal{S}_0)}\le r\implies \operatorname{dist}(u(\mathcal{Y},\theta,a),\mathcal{T}^I_{\mathcal{S}_0})_{H^s(\mathbb{T}^2)}\le cr \tag{3.10}$$

for some $c>0$ and any sufficiently small $r\ge0$.

Step 3. Now, we expand each line separately. By Taylor expanding around the finite gap torus corresponding to $(\mathcal{Y},\theta,a)=(0,\theta,0)$ we obtain, up to an additive constant,

$$
\begin{aligned}
& h_{\text{1D-NLS}}(I_{\mathtt{m}_1}+\mathcal{Y}_1,\dots,I_{\mathtt{m}_{\mathtt{d}}}+\mathcal{Y}_{\mathtt{d}},(|a_{(m,0)}|^2)_{m\notin\mathcal{S}_0}) \\
&\quad = \sum_{i=1}^{\mathtt{d}} \partial_{\mathtt{m}_i} h_{\text{1D-NLS}}(I_{\mathtt{m}_1},\dots,I_{\mathtt{m}_{\mathtt{d}}},0)\mathcal{Y}_i + \sum_{m\in\mathbb{Z}\setminus\mathcal{S}_0} \partial_m h_{\text{1D-NLS}}(I_{\mathtt{m}_1},\dots,I_{\mathtt{m}_{\mathtt{d}}},0)|a_{(m,0)}|^2 \\
&\qquad -\frac{1}{2}(|\mathcal{Y}|^2+\sum_{m\in\mathbb{Z}\setminus\mathcal{S}_0}|a_{(m,0)}|^4)+O\Big(|I|\Big\{\sum_{j=1}^{\mathtt{d}}\mathcal{Y}_j+\sum_{m\notin\mathcal{S}_0}|a_{(m,0)}|^2\Big\}^2\Big) \\
&\qquad + O\Big(\Big\{\sum_{j=1}^{\mathtt{d}}\mathcal{Y}_j+\sum_{m\notin\mathcal{S}_0}|a_{(m,0)}|^2\Big\}^3\Big),
\end{aligned}
$$

where we have used formula (3.4) in order to deduce that $\frac{\partial^2 h_{\text{1D-NLS}}}{\partial I_m \partial I_n}(0)=-\delta_n^m$ where δ_n^m is the Kronecker delta.

The following lemma follows easily from Theorem 3.1 (particularly formulae (3.4) and (3.5)):

Lemma 3.3 (Frequencies around the finite gap torus). *Denote*

$$
\partial_{I_{\mathtt{m}_j}} h_{\text{1D-NLS}}(I_{\mathtt{m}_1},\dots,I_{\mathtt{m}_{\mathtt{d}}},0) \equiv \alpha_{\mathtt{m}_j}^{\text{1D-NLS}}(I_{\mathtt{m}_1},\dots,I_{\mathtt{m}_{\mathtt{d}}},0) = \mathtt{m}_j^2 - \widetilde{\lambda}_j(I_{\mathtt{m}_1},\dots,I_{\mathtt{m}_{\mathtt{d}}}).
$$

Then:

(1) *The map $(I_{\mathtt{m}_1},\dots,I_{\mathtt{m}_{\mathtt{d}}}) \mapsto \widetilde{\lambda}(I_{\mathtt{m}_1},\dots,I_{\mathtt{m}_{\mathtt{d}}}) = (\widetilde{\lambda}_i(I_{\mathtt{m}_1},\dots,I_{\mathtt{m}_{\mathtt{d}}}))_{1\le i\le \mathtt{d}}$ is a diffeomorphism from a small neighborhood of 0 in $\mathbb{R}^{\mathtt{d}}$ to a small neighborhood of 0 in $\mathbb{R}^{\mathtt{d}}$. Indeed, $\widetilde{\lambda}$ = Identity + (quadratic in I). More precisely, there exists $\varepsilon_{1d}>0$ such that if $0<\varepsilon<\varepsilon_{1d}$ and*

$$
\widetilde{\lambda}(I_{\mathtt{m}_1},\dots,I_{\mathtt{m}_{\mathtt{d}}}) = \varepsilon\lambda, \quad \lambda\in(1/2,1)^{\mathtt{d}},
$$

then $(I_{\mathtt{m}_1},\dots,I_{\mathtt{m}_{\mathtt{d}}}) = \varepsilon\lambda + O(\varepsilon^2)$. From now on, and to simplify notation, we will use the vector λ as a parameter as opposed to $(I_{\mathtt{m}_1},\dots,I_{\mathtt{m}_{\mathtt{d}}})$, and we write

$$
\omega_i(\lambda) = \mathtt{m}_i^2 - \varepsilon\lambda_i, \quad 1\le i\le \mathtt{d},
$$

for the frequencies at the tangential sites in $\mathcal{S}_0$.

(2) *For $m\in\mathbb{Z}\setminus\mathcal{S}_0$, denoting $\Omega_m(\lambda) := \partial_{I_m} h_{\text{1D-NLS}}(I_{\mathtt{m}_1}(\lambda),\dots,I_{\mathtt{m}_{\mathtt{d}}}(\lambda),0)$, we have*

$$
\Omega_m(\lambda) := m^2 + \frac{\varpi_m(I(\lambda))}{\langle m\rangle} \quad \text{with} \quad \sup_{\lambda\in(1/2,1)^{\mathtt{d}}}\ \sup_{m\in\mathbb{Z}} |\varpi_m(I(\lambda))| \le C\varepsilon^2.
$$

With this in mind, line (3.6) becomes

$$
\begin{aligned}
& h_{\text{1D-NLS}}\big(I_{\mathtt{m}_1}+\mathcal{Y}_1,\dots,I_{\mathtt{m}_{\mathtt{d}}}+\mathcal{Y}_{\mathtt{d}},(|a_{(m,0)}|^2)_{m\notin\mathcal{S}_0}\big) \\
&\quad = \omega(\lambda)\cdot\mathcal{Y} + \sum_{m\in\mathbb{Z}\setminus\mathcal{S}_0}\Omega_m(\lambda)|a_{(m,0)}|^2 - \frac{1}{2}\Big(|\mathcal{Y}|^2+\sum_{m\in\mathbb{Z}\setminus\mathcal{S}_0}|a_{(m,0)}|^4\Big) \\
&\qquad + O\Big(|I|\Big\{\sum_{j=1}^{\mathtt{d}}\mathcal{Y}_j+\sum_{m\notin\mathcal{S}_0}|a_{(m,0)}|^2\Big\}^2\Big) + O\Big(\Big\{\sum_{j=1}^{\mathtt{d}}\mathcal{Y}_j+\sum_{m\notin\mathcal{S}_0}|a_{(m,0)}|^2\Big\}^3\Big).
\end{aligned}
$$

We now analyze (3.7). This is given by

$$(3.7) = \sum_{\vec{j}\in\mathbb{Z}^2\setminus\mathbb{Z}} |\vec{j}|^2 |a_{\vec{j}}|^2 + 2 \sum^{\star}_{\substack{\vec{j}_i=(m_i,n_i),\, i=3,4,\, n_i\neq 0\\ m_1-m_2+m_3-m_4=0\\ n_3-n_4=0}} q_{m_1}\bar{q}_{m_2} a_{\vec{j}_3}\bar{a}_{\vec{j}_4}$$
$$+ \operatorname{Re} \sum_{\substack{\vec{j}_i=(m_i,n_i),\, i=2,4,\, n_i\neq 0\\ m_1-m_2+m_3-m_4=0\\ n_2+n_4=0}} \bar{q}_{m_1} a_{\vec{j}_2} \bar{q}_{m_3} a_{\vec{j}_4}$$

where we now think of q_m as a function of $\mathcal{Y}, \theta, a$. By Taylor expanding it at $\mathcal{Y} = 0$ and $a = 0$, we get

$$\begin{aligned} q_m &= q_m(\lambda; \mathcal{Y}, \theta, (a_{(m_1,0)})_{m_1\in\mathbb{Z}\setminus\mathcal{S}_0}) \\ &= \overbrace{q_m(\lambda; 0, \theta, 0)}^{=:q_m^{\rm fg}(\lambda;\theta)} + \sum_{i=1}^{\mathtt{d}} \frac{\partial q_m}{\partial \mathcal{Y}_i}(\lambda; 0, \theta, 0)\mathcal{Y}_i \\ &\quad + \sum_{m_1\in\mathbb{Z}\setminus\mathcal{S}_0} \left(\frac{\partial q_m}{\partial a_{(m_1,0)}}(\lambda; 0, \theta, 0) a_{(m_1,0)} + \frac{\partial q_m}{\partial \bar{a}_{(m_1,0)}}(\lambda; 0, \theta, 0)\bar{a}_{(m_1,0)} \right) \\ &\quad + \sum_{\substack{m_1,m_2\in\mathbb{Z}\setminus\mathcal{S}_0\\ \sigma_1,\sigma_2=\pm 1}} Q^{\sigma_1\sigma_2}_{m,m_1m_2}(\lambda;\theta) a^{\sigma_1}_{(m_1,0)} a^{\sigma_2}_{(m_2,0)} + \mathcal{O}(\mathcal{Y}^2, \mathcal{Y}a, a^3), \end{aligned} \tag{3.11}$$

where we have denoted by $(q_m^{\rm fg}(\lambda;\theta))_{m\in\mathbb{Z}}$ the finite gap torus (which corresponds to $\mathcal{Y} = 0$, $\mathbf{a} = 0$), $\mathcal{O}(\mathcal{Y}^2, \mathcal{Y}a, a^3)$ are terms that invoke $\mathcal{Y}^2$, $\mathcal{Y}a$ or a^3, and

$$Q^{\sigma_1\sigma_2}_{m,m_1m_2}(\lambda;\theta) = \frac{1}{2} \frac{\partial^2 q_m}{\partial a^{\sigma_1}_{m_1} \partial a^{\sigma_2}_{m_2}}(\lambda; 0, \theta, 0).$$

Therefore, we obtain

$$\begin{aligned} (3.7) &= \sum_{\vec{j}\in\mathbb{Z}^2\setminus\mathbb{Z}} |\vec{j}|^2 |a_{\vec{j}}|^2 + 2 \sum^{\star}_{\substack{\vec{j}_i=(m_i,n_i),\, i=3,4,\, n_i\neq 0\\ m_1-m_2+m_3-m_4=0\\ n_3-n_4=0}} q^{\rm fg}_{m_1}(\lambda;\theta)\bar{q}^{\rm fg}_{m_2}(\lambda;\theta) a_{\vec{j}_3}\bar{a}_{\vec{j}_4} \\ &\quad + \operatorname{Re} \sum_{\substack{\vec{j}_i=(m_i,n_i),\, i=2,4,\, n_i\neq 0\\ m_1-m_2+m_3-m_4=0\\ n_2+n_4=0}} \bar{q}^{\rm fg}_{m_1}(\lambda;\theta) a_{\vec{j}_2} \bar{q}^{\rm fg}_{m_3}(\lambda;\theta) a_{\vec{j}_4} \\ &\quad + \Bigg\{ 2 \sum^{\star}_{\substack{\vec{j}_i=(m_i,n_i),\, i=3,4,\, n_i\neq 0\\ m_1-m_2+m_3-m_4=0\\ n_3-n_4=0}} \sum_{m_2'\in\mathbb{Z}\setminus\mathcal{S}_0} \frac{\partial \bar{q}_{m_2}}{\partial \bar{a}_{(m_2',0)}}(\lambda; 0, \theta, 0) q^{\rm fg}_{m_1}(\lambda;\theta)\bar{a}_{(m_2',0)} a_{\vec{j}_3}\bar{a}_{\vec{j}_4} \\ &\quad + \text{similar cubic terms in } (a, \bar{a}) \Bigg\} \\ &\quad + (3.7)^{(2)} + (3.7)^{(\geq 3)} \end{aligned}$$

where (3.7)$^{(2)}$ are degree 2 terms (cf. Definition 2.1), and (3.7)$^{(\geq 3)}$ are those of degree ≥ 3. More precisely,

$$(3.7)^{(2)} = 2 \sum^{\star}_{\substack{\vec{j}_i=(m_i,n_i),\, i=3,4,\, n_i\neq 0\\ m_1-m_2+m_3-m_4=0\\ n_3-n_4=0\\ 1\leq i\leq \mathtt{d}}} q^{\mathrm{fg}}_{m_1}(\lambda;\theta)\frac{\partial \bar{q}_{m_2}}{\partial \mathcal{Y}_i}(\lambda;0,\theta,0)\mathcal{Y}_i a_{\vec{j}_3}\bar{a}_{\vec{j}_4} + \text{similar terms}$$
$$+ \sum^{\star}_{\substack{\vec{j}_i=(m_i,n_i),\, i=3,4,\, n_i\neq 0\\ m_1-m_2+m_3-m_4=0\\ n_3-n_4=0\\ \sigma_1,\sigma_2=\pm 1,\, m_1',m_2'\in\mathbb{Z}\setminus\mathcal{S}_0}} L^{\sigma_1,\sigma_2}_{m_1,m_2,m_1',m_2'}(\lambda;\theta)a^{\sigma_1}_{(m_1',0)}a^{\sigma_2}_{(m_2',0)}a_{\vec{j}_3}\bar{a}_{\vec{j}_4} + \text{similar terms}, \tag{3.12}$$

for some uniformly bounded coefficients $L^{\sigma_1,\sigma_2}_{m_1,m_2,m_1',m_2'}$.

Next, we move on to (3.8), for which we have, using (3.11),

$$(3.8) = \underbrace{2\,\mathrm{Re} \sum_{\substack{\vec{j}_i=(m_i,n_i),\, i=2,3,4,\, n_i\neq 0\\ m_1-m_2+m_3-m_4=0\\ -n_2+n_3-n_4=0}} q^{\mathrm{fg}}_{m_1}(\lambda;\theta)\bar{a}_{\vec{j}_2}a_{\vec{j}_3}\bar{a}_{\vec{j}_4} + 2\,\mathrm{Re} \sum_{\substack{\vec{j}_i=(m_i,n_i),\, i=2,3,4,\, n_i\neq 0\\ m_1-m_2+m_3-m_4=0\\ -n_2+n_3-n_4=0}} \frac{\partial q_{m_1}}{\partial a_{(m_1',0)}}(\lambda;0,\theta,0)a_{(m_1',0)}\bar{a}_{\vec{j}_2}a_{\vec{j}_3}\bar{a}_{\vec{j}_4} + \text{similar terms}}_{(3.8)^{(2)}}$$
$$+ (3.8)^{(\geq 3)}, \tag{3.13}$$

where (3.8)$^{(2)}$ are terms of degree 2 and (3.8)$^{(\geq 3)}$ are terms of degree ≥ 3.

In conclusion, we obtain

$$\mathcal{H}(\lambda;\mathcal{Y},\theta,\mathbf{a}) = \mathcal{N} + \mathcal{H}^{(0)}(\lambda;\theta,\mathbf{a}) + \mathcal{H}^{(1)}(\lambda;\theta,\mathbf{a}) + \mathcal{H}^{(2)}(\lambda;\mathcal{Y},\theta,\mathbf{a})$$
$$+ \mathcal{H}^{(\geq 3)}(\lambda;\mathcal{Y},\theta,\mathbf{a}), \tag{3.14}$$

where

$$\mathcal{N} = \sum_{i=1}^{\mathtt{d}} \omega_{\mathtt{m}_i}(\lambda)\mathcal{Y}_i + \sum_{m\notin\mathcal{S}_0}\Omega_m(\lambda)|a_{(m,0)}|^2 + \sum_{\substack{\vec{j}=(m,n)\in\mathbb{Z}^2\\ n\neq 0}} |\vec{j}|^2|a_{\vec{j}}|^2, \tag{3.15}$$

$$\mathcal{H}^{(0)}(\lambda;\theta,\mathbf{a}) = 2 \sum^{\star}_{\substack{\vec{j}_i=(m_i,n_i),\, i=3,4,\, n_i\neq 0\\ m_1-m_2+m_3-m_4=0\\ n_3-n_4=0}} q^{\mathrm{fg}}_{m_1}(\lambda;\theta)\bar{q}^{\mathrm{fg}}_{m_2}(\lambda;\theta)a_{\vec{j}_3}\bar{a}_{\vec{j}_4}$$
$$+ \mathrm{Re} \sum_{\substack{\vec{j}_i=(m_i,n_i),\, i=2,4,\, n_i\neq 0\\ m_1-m_2+m_3-m_4=0\\ n_2+n_4=0}} \bar{q}^{\mathrm{fg}}_{m_1}(\lambda;\theta)a_{\vec{j}_2}\bar{q}^{\mathrm{fg}}_{m_3}(\lambda;\theta)a_{\vec{j}_4}, \tag{3.16}$$

$$
\begin{aligned}
&\mathcal{H}^{(1)}(\lambda;\theta,\mathbf{a})\\
&\quad = 2\,\mathrm{Re} \sum_{\substack{\vec{j}_i=(m_i,n_i),\, i=2,3,4,\, n_i\neq 0\\ m_1-m_2+m_3-m_4=0\\ -n_2+n_3-n_4=0}} q^{\mathrm{fg}}_{m_1}(\lambda;\theta)\bar{a}_{\vec{j}_2} a_{\vec{j}_3}\bar{a}_{\vec{j}_4}\\
&\qquad + 2 \sum^{\star}_{\substack{\vec{j}_i=(m_i,n_i),\, i=3,4,\, n_i\neq 0\\ m_1-m_2+m_3-m_4=0\\ n_3-n_4=0}} \sum_{m_2'\in\mathbb{Z}\setminus\mathcal{S}_0} \frac{\partial \bar{q}_{m_2}}{\partial \bar{a}_{(m_2',0)}}(\lambda;0,\theta,0) q^{\mathrm{fg}}_{m_1}(\lambda;\theta)\bar{a}_{(m_2',0)} a_{\vec{j}_3}\bar{a}_{\vec{j}_4}\\
&\qquad + \text{similar cubic terms in } (a,\bar{a}),
\end{aligned}
\tag{3.17}
$$

$$
\begin{aligned}
\mathcal{H}^{(2)}(\lambda;\theta,\mathbf{a}) = {}& H^{\mathrm{IV}}((a_{(m,n)})_{n\neq 0}) - \frac{1}{2}\Big(|\mathcal{Y}|^2 + \sum_{m\in\mathbb{Z}\setminus\mathcal{S}_0} |a_{(m,0)}|^4\Big)\\
&+ O\Big(\varepsilon\Big\{\sum_{j=1}^{\mathtt{d}} \mathcal{Y}_j + \sum_{m\notin\mathcal{S}_0} |a_{(m,0)}|^2\Big\}^2\Big) + (3.7)^{(2)} + (3.8)^{(2)},
\end{aligned}
\tag{3.18}
$$

where $(3.7)^{(2)}$ and $(3.8)^{(2)}$ were defined in (3.12) and (3.13) respectively. Finally, $\mathcal{H}^{(\geq 3)}$ collects all remainder terms of degree ≥ 3.

For short we write $\mathcal{N}$ as $\mathcal{N} = \omega(\lambda)\cdot\mathcal{Y} + \mathcal{D}$ where $\mathcal{D}$ is the diagonal operator

$$
\mathcal{D} := \sum_{\vec{j}=(m,n)\in\mathbb{Z}^2\setminus\mathcal{S}_0} \Omega^{(0)}_{\vec{j}} |a_{\vec{j}}|^2
$$

and the normal frequencies $\Omega^{(0)}_{\vec{j}}$ are defined by

$$
\Omega^{(0)}_{\vec{j}} := \begin{cases} |\vec{j}|^2 & \text{if } \vec{j} = (m,n) \text{ with } n\neq 0,\\ \Omega_m(\lambda) & \text{if } \vec{j} = (m,0),\ m\notin\mathcal{S}_0. \end{cases}
\tag{3.19}
$$

Proceeding as in [43], one can prove the following result:

Lemma 3.4. *Fix $\rho > 0$. There exists $\varepsilon_* > 0$ and for any $0 \leq \varepsilon \leq \varepsilon_*$ there exist $r_* \leq \sqrt{\varepsilon}/4$ and $C > 0$ such that $\mathcal{H}^{(0)}$, $\mathcal{H}^{(1)}$, $\mathcal{H}^{(2)}$ and $\mathcal{H}^{(\geq 3)}$ belong to $\mathcal{A}^{\mathcal{O}}_{\rho,r_*}$ and for $0 < r \leq r_*$,*

$$
|\mathcal{H}^{(0)}|^{\mathcal{O}}_{\rho,r} \leq C\varepsilon, \quad |\mathcal{H}^{(1)}|^{\mathcal{O}}_{\rho,r} \leq C\sqrt{\varepsilon}\, r, \quad |\mathcal{H}^{(2)}|^{\mathcal{O}}_{\rho,r} \leq C r^2, \quad |\mathcal{H}^{(\geq 3)}|^{\mathcal{O}}_{\rho,r} \leq C\frac{r^3}{\sqrt{\varepsilon}}.
\tag{3.20}
$$

4. Reducibility of the quadratic part

In this section, we review the reducibility of the quadratic part $\mathcal{N} + \mathcal{H}^{(0)}$ (see (3.15) and (3.16)) of the Hamiltonian, which is the main part of [43]. This will be a symplectic linear change of coordinates that transforms the quadratic part into an effectively diagonal, time independent expression.

4.1. Restriction to an invariant sublattice $\mathbb{Z}_N^2$

For $N \in \mathbb{N}$, we define the sublattice $\mathbb{Z}_N^2 := \mathbb{Z} \times N\mathbb{Z}$ and remark that it is invariant for the flow in the sense that the subspace

$$E_N := \{a_{\vec{j}} = \bar{a}_{\vec{j}} = 0 \text{ for } \vec{j} \notin \mathbb{Z}_N^2\}$$

is invariant for the original NLS dynamics and that of the Hamiltonian (3.14). From now on, we restrict our system to this invariant sublattice, with

$$N > \max_{1 \le i \le \mathtt{d}} |\mathtt{m}_i|. \tag{4.1}$$

The reason for this restriction is that it simplifies (actually eliminates the need for) some genericity requirements that are needed for [43] as well as some of the normal form calculations that we will perform later.

It will also be important to introduce the following two subsets of $\mathbb{Z}_N^2$:

$$\mathscr{S} := \{(\mathtt{m}, n) : \mathtt{m} \in \mathcal{S}_0,\ n \in N\mathbb{Z},\ n \neq 0\}, \quad \mathfrak{Z} = \mathbb{Z}_N^2 \setminus (\mathscr{S} \cup \mathcal{S}_0). \tag{4.2}$$

4.2. Admissible monomials and reducibility

The reducibility of the quadratic part of the Hamiltonian will introduce a change of variables that modifies the expression of the mass $\mathcal{M}$ and momentum $\mathcal{P}$ as follows. Let us set

$$\begin{aligned} \widetilde{\mathcal{M}} &:= \sum_{i=1}^{\mathtt{d}} \mathcal{Y}_i + \sum_{(m,n)\in\mathfrak{Z}} |a_j|^2, \\ \widetilde{\mathcal{P}}_x &:= \sum_{i=1}^{\mathtt{d}} \mathtt{m}_i \mathcal{Y}_i + \sum_{(m,n)\in\mathfrak{Z}} m|a_{(m,n)}|^2, \\ \widetilde{\mathcal{P}}_y &:= \sum_{(m,n)\in\mathbb{Z}_N^2} n|a_{(m,n)}|^2. \end{aligned} \tag{4.3}$$

These will be the expressions for the mass and momentum after the change of variables introduced in the following two theorems. Notice the absence of the terms $\sum_{1\le i\le \mathtt{d},\, n\in N\mathbb{Z}} |a_{(\mathtt{m}_i,n)}|^2$ and $\sum_{1\le i\le \mathtt{d},\, n\in N\mathbb{Z}} \mathtt{m}_i |a_{(\mathtt{m}_i,n)}|^2$ from the expressions of $\widetilde{\mathcal{M}}$ and $\widetilde{\mathcal{P}}_x$ above. These terms are absorbed in the new definition of the $\mathcal{Y}$ and $\mathbf{a}$ variables.

Definition 4.1 (Admissible monomials). Given $\mathbf{j} = (\vec{j}_1, \dots, \vec{j}_p) \in (\mathbb{Z}_N^2 \setminus \mathcal{S}_0)^p$, $\ell \in \mathbb{Z}^{\mathtt{d}}$, $l \in \mathbb{N}^{\mathtt{d}}$, and $\sigma = (\sigma_1, \dots, \sigma_p) \in \{-1, 1\}^p$, we say that $(\mathbf{j}, \ell, \sigma)$ is *admissible*, and denote $(\mathbf{j}, \ell, \sigma) \in \mathfrak{A}_p$, if the monomial $\mathfrak{m} = e^{\mathrm{i}\theta\cdot\ell} \mathcal{Y}^l a_{\vec{j}_1}^{\sigma_1} \dots a_{\vec{j}_p}^{\sigma_p}$ Poisson commutes with $\widetilde{\mathcal{M}}, \widetilde{\mathcal{P}}_x, \widetilde{\mathcal{P}}_y$. We call a monomial $e^{\mathrm{i}\theta\cdot\ell} \mathcal{Y}^l a_{\vec{j}_1}^{\sigma_1} \dots a_{\vec{j}_p}^{\sigma_p}$ admissible if $(\mathbf{j}, \ell, \sigma)$ is admissible.

Definition 4.2. We define the *resonant set at degree 0* by

$$\mathfrak{R}_2 := \{(\vec{j}_1, \vec{j}_2, \ell, \sigma_1, \sigma_2)\} \in \mathfrak{A}_2 : \ell = 0,\ \sigma_1 = -\sigma_2,\ \vec{j}_1 = \vec{j}_2\}. \tag{4.4}$$

Theorem 4.3. *Fix $\varepsilon_0 > 0$ sufficiently small. There exist positive $\rho_0, \gamma_0, \tau_0, r_0, \mathtt{L}_0$ (with $\mathtt{L}_0$ depending only on $\mathtt{d}$) such that the following holds true uniformly for all $0 < \varepsilon \le \varepsilon_0$: For an $\mathtt{L}_0$-generic choice of the set $\mathcal{S}_0$ (in the sense of Definition 1.1), there exist a compact* domain $\mathcal{O}_0 \subseteq (1/2, 1)^{\mathtt{d}}$, *satisfying $|(1/2,1)^{\mathtt{d}} \setminus \mathcal{O}_0| \le \varepsilon_0$, and Lipschitz (in λ) functions $\{\Omega_{\vec{j}}\}_{\vec{j}\in\mathbb{Z}_N^2\setminus\mathcal{S}_0}$ defined on $\mathcal{O}_0$ (described more precisely in Theorem 4.4 below) such that:*

(1) *The set*

$$\mathcal{C}^{(0)} := \{\lambda \in \mathcal{O}_0 : |\omega\cdot\ell + \sigma_1\Omega_{\vec{j}_1}(\lambda,\varepsilon) + \sigma_2\Omega_{\vec{j}_2}(\lambda,\varepsilon)| \ge \gamma_0\varepsilon/\langle\ell\rangle^{\tau_0}, \quad \forall(\vec{j},\ell,\sigma)\in\mathfrak{A}_2\setminus\mathfrak{R}_2\} \tag{4.5}$$

has positive measure. In fact $|\mathcal{O}_0\setminus\mathcal{C}^{(0)}| \lesssim \varepsilon_0^{\kappa_0}$ for some $\kappa_0 > 0$ independent of ε_0.

(2) *For each $\lambda\in\mathcal{C}^{(0)}$ and all $r\in[0,r_0]$, $\rho\in[\frac{\rho_0}{64},\rho_0]$, there exists an invertible symplectic change of variables $\mathcal{L}^{(0)}$ that is well defined and majorant analytic from $D(\rho/8,\zeta_0 r)$ to $D(\rho,r)$ (here $\zeta_0>0$ is a constant depending only on ρ_0, $\max|\mathtt{m}_k|^2$) and such that if $\mathbf{a}\in h^1(\mathbb{Z}_N^2\setminus\mathcal{S}_0)$, then*

$$(\mathcal{N}+\mathcal{H}^{(0)})\circ\mathcal{L}^{(0)}(\mathcal{Y},\theta,\mathbf{a}) = \omega\cdot\mathcal{Y} + \sum_{\vec{j}\in\mathbb{Z}_N^2\setminus\mathcal{S}_0}\Omega_{\vec{j}}\,|a_{\vec{j}}|^2.$$

(3) *The mass $\mathcal{M}$ and the momentum $\mathcal{P}$ (defined in (2.5)) in the new coordinates are given by*

$$\mathcal{M}\circ\mathcal{L}^{(0)} = \widetilde{\mathcal{M}}, \quad \mathcal{P}\circ\mathcal{L}^{(0)} = \widetilde{\mathcal{P}}, \tag{4.6}$$

where $\widetilde{\mathcal{M}}$ and $\widetilde{\mathcal{P}}$ are defined in (4.3).

(4) *The map $\mathcal{L}^{(0)}$ maps h^1 to itself and has the form*

$$\mathcal{L}^{(0)}: \quad \mathbf{a}\mapsto L(\lambda;\theta,\varepsilon)\mathbf{a}, \quad \mathcal{Y}\mapsto\mathcal{Y}+(\mathbf{a},Q(\lambda;\theta,\varepsilon)\mathbf{a}), \quad \theta\mapsto\theta.$$

The same holds for the inverse map $(\mathcal{L}^{(0)})^{-1}$.

(5) *The linear maps $L(\lambda;\theta,\varepsilon)$ and $Q(\lambda;\theta,\varepsilon)$ are block diagonal in the y Fourier modes, in the sense that $L = \mathrm{diag}_{n\in N\mathbb{N}}(L_n)$ with each L_n acting on the sequence $\{a_{(m,n)}, a_{(m,-n)}\}_{m\in\mathbb{Z}}$ (and similarly for Q). Moreover, $L_0 = \mathrm{Id}$ and L_n is of the form* $\mathrm{Id} + S_n$ *where S_n is a smoothing operator in the following sense: with the smoothing norm $\lceil\cdot\rfloor_{\rho,-1}$ defined in (4.7) below,*

$$\sup_{n\neq 0}\lceil S_n\circ P_{\{|m|\ge(\mathtt{m}_{\mathtt{d}}+1)\}}\rfloor_{\rho,-1} \lesssim \varepsilon,$$

where $P_{\{|m|\ge K\}}$ is the orthogonal projection of a sequence $(c_m)_{m\in\mathbb{Z}}$ onto the modes $|m|\ge K$.

The above smoothing norm is defined as follows: Let $S(\lambda;\theta,\varepsilon)$ be an operator acting on sequences $(c_k)_{k\in\mathbb{Z}}$ through its matrix elements $S(\lambda;\theta,\varepsilon)_{m,k}$. Let us denote by $S(\lambda;\ell,\varepsilon)_{m,k}$ the θ-Fourier coefficients of $S(\lambda;\theta,\varepsilon)_{m,k}$. For $\rho,\nu>0$ we define

$$\lceil S(\lambda;\theta,\varepsilon)\rfloor_{\rho,\nu} := \sup_{\|c\|_{\ell^1}\le 1}\Big\|\Big(\sum_{\substack{k\in\mathbb{Z}\\ \ell\in\mathbb{Z}^{\mathtt{d}}}} e^{\rho|\ell|}|S_{m,k}(\lambda;\ell,\varepsilon)|\langle k\rangle^{-\nu}c_k\Big)\Big\|_{\ell^1}. \tag{4.7}$$

This definition is equivalent to the more general norm used in [43, Definition 3.9]. Roughly speaking, the boundedness of this norm means that, in terms of its action on sequences, S maps $\langle k\rangle^{\nu}\ell^1 \to \ell^1$. As observed in [43, Remark 3.10], thanks to the conservation of momentum this also means that S maps $\ell^1 \to \langle k\rangle^{-\nu}\ell^1$.

Proof of Theorem 4.3. The result follows from [43], by applying first the change of variables in Theorem 5.1 and then the one in Theorem 7.1 to the quadratic part of the NLS Hamiltonian (hence ignoring the terms $\widetilde{\mathcal{H}}^{(1)}$, $\widetilde{\mathcal{H}}^{(\geq 2)}$ in [43, (5.2)] and the terms $\mathcal{K}^{(1)}, \mathcal{K}^{(\geq 2)}$ in [43, (7.3)]). Note that in [43] the results are proved in h^s norm with $s > 1$, for instance in (4.7) the ℓ^1 norm is substituted with the h^s norm. However, the proofs only rely on momentum conservation and on the fact that h^s is an algebra with respect to convolution, which holds true also for ℓ^1. Hence the proof of our case is identical and we do not repeat it.

We are able to describe quite precisely the asymptotics of the frequencies $\Omega_{\vec{j}}$ of Theorem 4.3.

Theorem 4.4. *For any $0 < \varepsilon \leq \varepsilon_0$ and $\lambda \in \mathcal{C}^{(0)}$, the frequencies $\Omega_{\vec{j}} \equiv \Omega_{\vec{j}}(\lambda, \varepsilon)$, $\vec{j} = (m, n) \in \mathbb{Z}^2_N \setminus \mathcal{S}_0$, introduced in Theorem 4.3 have the following asymptotics:*

$$\Omega_{\vec{j}}(\lambda,\varepsilon) = \begin{cases} \widetilde{\Omega}_{\vec{j}}(\lambda,\varepsilon) + \dfrac{\varpi_m(\lambda,\varepsilon)}{\langle m\rangle}, & n = 0, \\ \widetilde{\Omega}_{\vec{j}}(\lambda,\varepsilon) + \dfrac{\Theta_m(\lambda,\varepsilon)}{\langle m\rangle^2} + \dfrac{\Theta_{m,n}(\lambda,\varepsilon)}{\langle m\rangle^2 + \langle n\rangle^2}, & n \neq 0, \end{cases} \tag{4.8}$$

where

$$\widetilde{\Omega}_{\vec{j}}(\lambda,\varepsilon) := \begin{cases} m^2, & \vec{j} = (m,0),\ m \notin \mathcal{S}_0, \\ m^2 + n^2, & \vec{j} = (m,n) \in \mathfrak{Z},\ n \neq 0, \\ \varepsilon\mu_i(\lambda) + n^2, & \vec{j} = (\mathtt{m}_i, n) \in \mathscr{S},\ n \neq 0, \end{cases}$$

where $\mathfrak{Z}$ and $\mathscr{S}$ are the sets defined in (4.2).

Here the $\{\mu_i(\lambda)\}_{1\leq i\leq \mathtt{d}}$ are the roots of the polynomial

$$P(t,\lambda) := \prod_{i=1}^{\mathtt{d}} (t+\lambda_i) - 2\sum_{i=1}^{\mathtt{d}} \lambda_i \prod_{k\neq i}(t+\lambda_k),$$

which is irreducible over $\mathbb{Q}(\lambda)[t]$.

Finally, $\mu_i(\lambda)$, $\{\varpi_m(\lambda,\varepsilon)\}_{m\in\mathbb{Z}\setminus\mathcal{S}_0}$, $\{\Theta_m(\lambda,\varepsilon)\}_{m\in\mathbb{Z}}$ and $\{\Theta_{m,n}(\lambda,\varepsilon)\}_{(m,n)\in\mathbb{Z}^2_N\setminus\mathcal{S}_0}$ fulfill

$$\sum_{1\leq i\leq \mathtt{d}} |\mu_i(\cdot)|^{\mathcal{O}_0} + \sup_{\varepsilon\leq\varepsilon_0} \frac{1}{\varepsilon^2}\Big(\sup_{m\in\mathbb{Z}\setminus\mathcal{S}_0} |\varpi_m(\cdot,\varepsilon)|^{\mathcal{O}_0} + \sup_{m\in\mathbb{Z}} |\Theta_m(\cdot,\varepsilon)|^{\mathcal{O}_0} + \sup_{\substack{(m,n)\in\mathbb{Z}^2_N \\ n\neq 0}} |\Theta_{m,n}(\cdot,\varepsilon)|^{\mathcal{O}_0}\Big) \leq \mathtt{M}_0 \tag{4.9}$$

for some $\mathtt{M}_0$ independent of ε.

Theorem 4.4 follows from [43, Theorem 2.10 and Corollary 7.5)], together with the observation that the set $\mathscr{C}$ defined in [43, Definition 2.3] satisfies $\mathscr{C} \cap \mathbb{Z}^2_N = \emptyset$ if $N > \max_i |\mathtt{m}_i|$.

We conclude this section with a series of remarks.

Remark 4.5. Notice that the $\{\mu_i(\lambda)\}_{1\le i\le \mathtt{d}}$ depend on the number $\mathtt{d}$ of tangential sites but *not on* $\{\mathtt{m}_i\}_{1\le i\le \mathtt{d}}$.

Remark 4.6. The asymptotic expansion (4.8) of the normal frequencies does not contain any constant term. The reason is that we canceled such a term when we subtracted the quantity $M(u)^2$ from the Hamiltonian at the very beginning (see the footnote in Section 3.1). Of course if we had not removed $M(u)^2$, we would have had a constant correction to the frequencies, equal to $\|q(\omega t,\cdot)\|^2_{L^2}$. Since $q(\omega t, x)$ is a solution of (2D-NLS), it enjoys mass conservation, and thus $\|q(\omega t,\cdot)\|^2_{L^2} = \|q(0,\cdot)\|^2_{L^2}$ is independent of time.

Remark 4.7. In the new variables, the *selection rules* of Remark 2.2 become (with $\mathcal{H}$ expanded as in (2.2))

$$\begin{aligned}
\{\mathcal{H}, \widetilde{\mathcal{M}}\} = 0 &\iff \mathcal{H}_{\alpha,\beta,\ell}\,(\widetilde{\eta}(\alpha,\beta) + \eta(\ell)) = 0,\\
\{\mathcal{H}, \widetilde{\mathcal{P}}_x\} = 0 &\iff \mathcal{H}_{\alpha,\beta,\ell}\,(\widetilde{\pi}_x(\alpha,\beta) + \pi(\ell)) = 0,\\
\{\mathcal{H}, \widetilde{\mathcal{P}}_y\} = 0 &\iff \mathcal{H}_{\alpha,\beta,\ell}\,(\pi_y(\alpha,\beta)) = 0,
\end{aligned}$$

where $\eta(\ell)$ is defined in (2.3), $\pi_y(\alpha,\beta), \pi(\ell)$ in (2.4), while

$$\widetilde{\eta}(\alpha,\beta) := \sum_{\vec{\jmath}\in\mathfrak{Z}} (\alpha_{\vec{\jmath}} - \beta_{\vec{\jmath}}), \quad \widetilde{\pi}_x(\alpha,\beta) := \sum_{\vec{\jmath}=(m,n)\in\mathfrak{Z}} m(\alpha_{\vec{\jmath}} - \beta_{\vec{\jmath}}).$$

5. Elimination of cubic terms

If we apply the change $\mathcal{L}^{(0)}$ obtained in Theorem 4.3 to the Hamiltonian (3.14), we obtain

$$\begin{aligned}
\mathcal{K}(\lambda; \mathcal{Y}, \theta, \mathbf{a}) &:= \mathcal{H} \circ \mathcal{L}^{(0)}(\lambda; \mathcal{Y}, \theta, \mathbf{a})\\
&= \omega\cdot\mathcal{Y} + \sum_{\vec{\jmath}\in\mathbb{Z}^2_N\setminus\mathcal{S}_0} \Omega_{\vec{\jmath}}|a_{\vec{\jmath}}|^2 + \mathcal{K}^{(1)} + \mathcal{K}^{(2)} + \mathcal{K}^{(\ge 3)},\\
\mathcal{K}^{(j)} = \mathcal{H}^{(j)}\circ\mathcal{L}^{(0)} \quad (j = 1,2),& \quad \mathcal{K}^{(\ge 3)} = \mathcal{H}^{(\ge 3)}\circ\mathcal{L}^{(0)}.
\end{aligned} \tag{5.1}$$

As a direct consequence of Lemma 3.4 and Theorem 4.3, estimates (3.20) hold also for $\mathcal{K}^{(j)}$, $j = 1, 2$, and $\mathcal{K}^{(\ge 3)}$.

We now perform one step of Birkhoff normal form change of variables which cancels out $\mathcal{K}^{(1)}$ completely. In order to define such a change of variables we need to impose third order Melnikov conditions, which hold true on a subset of the set $\mathcal{C}^{(0)}$ of Theorem 4.3.

Lemma 5.1. *Fix* $0 < \varepsilon_1 < \varepsilon_0$ *sufficiently small and* $\tau_1 > \tau_0$ *sufficiently large. There exist constants* $\gamma_1 > 0$ *and* $\mathtt{L}_1 > \mathtt{L}_0$ *(with* $\mathtt{L}_1$ *depending only on* $\mathtt{d}$*) such that for all* $0 < \varepsilon \le \varepsilon_1$

and for an $\mathtt{L}_1$*-generic choice of the set* $\mathcal{S}_0$ (*in the sense of Definition* 1.1), *the set*

$$\mathcal{C}^{(1)} := \{\lambda \in \mathcal{C}^{(0)} : |\omega \cdot \ell + \sigma_1 \Omega_{\vec{J}_1}(\lambda, \varepsilon) + \sigma_2 \Omega_{\vec{J}_2}(\lambda, \varepsilon) + \sigma_3 \Omega_{\vec{J}_3}(\lambda, \varepsilon)| \geq \gamma_1 \varepsilon / \langle \ell \rangle^{\tau_1}, \\ \forall (\vec{J}, \ell, \sigma) \in \mathfrak{A}_3\},$$

where $\mathfrak{A}_3$ *is introduced in Definition* 4.1, *has positive measure. More precisely, we have* $|\mathcal{C}^{(0)} \setminus \mathcal{C}^{(1)}| \lesssim \varepsilon_1^{\kappa_1}$ *for some constant* $\kappa_1 > 0$ *independent of* ε_1.

This lemma is proven in [43, Appendix C].

The main result of this section is the following theorem.

Theorem 5.2. *Assume the hypotheses and the notation of Lemma* 5.1. *Consider the constants* $\mathtt{L}_1$, γ_1, τ_1 *given by Lemma* 5.1, *the associated set* $\mathcal{C}^{(1)}$, *and the constants* ε_0, ρ_0 *and* r_0 *given in Theorem* 4.3. *There exist* $0 < \varepsilon_1 \leq \varepsilon_0$, $0 < \rho_1 \leq \rho_0/64$, *and* $0 < r_1 \leq r_0$ *such that the following holds true for all* $0 < \varepsilon \leq \varepsilon_1$. *For each* $\lambda \in \mathcal{C}^{(1)}$ *and all* $0 < r \leq r_1$ *and* $0 < \rho \leq \rho_1$, *there exists a symplectic change of variables* $\mathcal{L}^{(1)}$ *that is well defined and majorant analytic from* $D(\rho/2, r/2)$ *to* $D(\rho, r)$ *and such that applied to the Hamiltonian* $\mathcal{K}$ *in* (5.1) *it leads to*

$$\mathcal{Q} := \mathcal{K} \circ \mathcal{L}^{(1)}(\lambda; \mathcal{Y}, \theta, \mathbf{a}) = \omega \cdot \mathcal{Y} + \sum_{\vec{J} \in \mathbb{Z}_N^2 \setminus \mathcal{S}_0} \Omega_{\vec{J}}(\lambda, \varepsilon)|a_{\vec{J}}|^2 + \mathcal{Q}^{(2)} + \mathcal{Q}^{(\geq 3)}, \quad (5.2)$$

where:

(i) *The map* $\mathcal{L}^{(1)}$ *is the time-*1 *flow of a cubic Hamiltonian* $\chi^{(1)}$ *such that* $|\chi^{(1)}|^{\mathcal{C}^{(1)}}_{\rho/2, r/2} \lesssim r/\sqrt{\varepsilon}$.

(ii) $\mathcal{Q}^{(2)}$ *is of degree* 2 (*in the sense of Definition* 2.1), *it is given by*

$$\mathcal{Q}^{(2)} = \mathcal{K}^{(2)} + \tfrac{1}{2}\{\mathcal{K}^{(1)}, \chi^{(1)}\}, \quad (5.3)$$

and satisfies $|\mathcal{Q}^{(2)}|_{\rho/2, r/2} \lesssim r^2$.

(iii) $\mathcal{Q}^{(\geq 3)}$ *is of degree at least* 3 *and satisfies*

$$|\mathcal{Q}^{(\geq 3)}|^{\mathcal{C}^{(1)}}_{\rho/2, r/2} \lesssim r^3/\sqrt{\varepsilon}. \quad (5.4)$$

(iv) $\mathcal{L}^{(1)}$ *satisfies* $\widetilde{\mathcal{M}} \circ \mathcal{L}^{(1)} = \widetilde{\mathcal{M}}$ *and* $\widetilde{\mathcal{P}} \circ \mathcal{L}^{(1)} = \widetilde{\mathcal{P}}$.

(v) $\mathcal{L}^{(1)}$ *maps* $D(\rho/2, r/2) \cap h^1 \to D(\rho, r) \cap h^1$, *and if we denote* $(\widetilde{\mathcal{Y}}, \widetilde{\theta}, \widetilde{\mathbf{a}}) = \mathcal{L}^{(1)}(\mathcal{Y}, \theta, \mathbf{a})$, *then*

$$\|\widetilde{\mathbf{a}} - \mathbf{a}\|_{\ell^1} \lesssim \|\mathbf{a}\|^2_{\ell^1}. \quad (5.5)$$

To prove this theorem, we use the following lemma, which is proved in [43].

Lemma 5.3. *For every* $\rho, r > 0$ *the following holds true:*

(i) *Let* $h, f \in \mathcal{A}^{\mathcal{O}}_{\rho,r}$. *For any* $0 < \rho' < \rho$ *and* $0 < r' < r$, *one has*

$$|\{f, g\}|^{\mathcal{O}}_{\rho', r'} \leq \upsilon^{-1} C |f|^{\mathcal{O}}_{\rho,r} |g|^{\mathcal{O}}_{\rho,r}.$$

where $\upsilon := \min(1 - r'/r, \rho - \rho')$. *If* $\upsilon^{-1}|f|^{\mathcal{O}}_{\rho,r} < \zeta$ *is sufficiently small then the* (*time*-1 *flow of the*) *Hamiltonian vector field* X_f *defines a close-to-identity canonical change of variables* $\mathcal{T}_f$ *such that*

$$|h \circ \mathcal{T}_f|^{\mathcal{O}}_{\rho',r'} \le (1 + C\zeta)|h|^{\mathcal{O}}_{\rho,r} \quad \text{for all } 0 < \rho' < \rho,\ 0 < r' < r.$$

(ii) *Let* $f, g \in \mathcal{A}^{\mathcal{O}}_{\rho,r}$ *be of minimal degree respectively* $\mathtt{d}_f$ *and* $\mathtt{d}_g$ (*see Definition* 2.1) *and define the function*

$$\mathcal{T}_{\mathtt{i}}(f; h) = \sum_{l=\mathtt{i}}^{\infty} \frac{(\mathrm{ad} f)^l}{l!} h, \quad \mathrm{ad}(f)h := \{h, f\}. \tag{5.6}$$

Then $\mathcal{T}_{\mathtt{i}}(f; g)$ *is of minimal degree* $\mathtt{d}_f\,\mathtt{i} + \mathtt{d}_g$ *and*

$$|\mathcal{T}_{\mathtt{i}}(f; h)|^{\mathcal{O}}_{\rho',r'} \le C(\rho)\upsilon^{-\mathtt{i}}(|f|^{\mathcal{O}}_{\rho,r})^{\mathtt{i}}|g|^{\mathcal{O}}_{\rho,r}, \quad \forall 0 < \rho' < \rho,\ 0 < r' < r.$$

Proof of Theorem 5.2. We look for $\mathcal{L}^{(1)}$ as the time-1 flow of a Hamiltonian $\chi^{(1)}$. With

$$\widehat{\mathcal{N}} := \omega \cdot \mathcal{Y} + \sum_{\vec{\jmath} \in \mathbb{Z}^2_N \setminus \mathcal{S}_0} \Omega_{\vec{\jmath}}(\lambda, \varepsilon)|a_{\vec{\jmath}}|^2 \quad \text{and} \quad \mathcal{T}_j(\chi^{(1)}; \cdot) = \sum_{k \ge j} \frac{\mathrm{ad}(\chi^{(1)})^{k-1}[\{\cdot, \chi^{(1)}\}]}{k!},$$

we have

$$\mathcal{K} \circ \mathcal{L}^{(1)} = \widehat{\mathcal{N}} + \{\widehat{\mathcal{N}}, \chi^{(1)}\} + \mathcal{K}^{(1)} \tag{5.7}$$

$$+ \mathcal{T}_2(\chi^{(1)}; \widehat{\mathcal{N}}) + \{\mathcal{K}^{(1)}, \chi^{(1)}\} + \mathcal{T}_2(\chi^{(1)}; \mathcal{K}^{(1)}) \tag{5.8}$$

$$+ \mathcal{K}^{(2)} + \mathcal{T}_1(\chi^{(1)}; \mathcal{K}^{(2)}) + \mathcal{K}^{(\ge 3)} \circ \mathcal{L}^{(1)} \tag{5.9}$$

We choose $\chi^{(1)}$ to solve the homological equation $\{\widehat{\mathcal{N}}, \chi^{(1)}\} + \mathcal{K}^{(1)} = 0$. Thus we set

$$\mathcal{K}^{(1)} = \sum_{\ell, \mathbf{j}, \vec{\sigma} \in \mathfrak{A}_3} K^{\vec{\sigma}}_{\ell,\mathbf{j}}(\lambda, \varepsilon)\, e^{\mathrm{i}\theta\cdot\ell} a^{\sigma_1}_{\vec{\jmath}_1} a^{\sigma_2}_{\vec{\jmath}_2} a^{\sigma_3}_{\vec{\jmath}_3}, \quad \chi^{(1)} = \sum_{\ell, \mathbf{j}, \vec{\sigma} \in \mathfrak{A}_3} \chi^{\vec{\sigma}}_{\ell,\mathbf{j}}(\lambda, \varepsilon)\, e^{\mathrm{i}\theta\cdot\ell} a^{\sigma_1}_{\vec{\jmath}_1} a^{\sigma_2}_{\vec{\jmath}_2} a^{\sigma_3}_{\vec{\jmath}_3}$$

with

$$\chi^{\vec{\sigma}}_{\ell,\mathbf{j}}(\lambda, \varepsilon) := \frac{\mathrm{i} K^{\vec{\sigma}}_{\ell,\mathbf{j}}(\lambda, \varepsilon)}{\omega \cdot \ell + \sigma_1 \Omega_{\vec{\jmath}_1}(\lambda, \varepsilon) + \sigma_2 \Omega_{\vec{\jmath}_2}(\lambda, \varepsilon) + \sigma_3 \Omega_{\vec{\jmath}_3}(\lambda, \varepsilon)}.$$

Since $\lambda \in \mathcal{C}^{(1)}$, we have

$$|\chi^{(1)}|^{\mathcal{C}^{(1)}}_{\rho/2,r} \lesssim r/\sqrt{\varepsilon},$$

since the terms q^{fg}_m appearing in $\mathcal{H}^{(1)}$ (and hence $\mathcal{K}^{(1)}$) are $O(\sqrt{\varepsilon})$. We turn to the terms of line (5.8). First we use the homological equation $\{\widehat{\mathcal{N}}, \chi^{(1)}\} + \mathcal{K}^{(1)} = 0$ to get

$$\mathcal{T}_2(\chi^{(1)}; \widehat{\mathcal{N}}) = \sum_{k \ge 2} \frac{\mathrm{ad}(\chi^{(1)})^{k-1}[\{\widehat{\mathcal{N}}, \chi^{(1)}\}]}{k!} = -\frac{1}{2}\{\mathcal{K}^{(1)}, \chi^{(1)}\} - \sum_{k \ge 2} \frac{\mathrm{ad}(\chi^{(1)})^k[\mathcal{K}^{(1)}]}{(k+1)!}.$$

Therefore, we define $\mathcal{Q}^{(2)}$ as in (5.3) and

$$\mathcal{Q}^{(\geq 3)} = \mathcal{T}_2(\chi^{(1)}; \mathcal{K}^{(1)}) + \mathcal{T}_1(\chi^{(1)}; \mathcal{K}^{(2)}) + \mathcal{K}^{(\geq 3)} \circ \mathcal{L}^{(1)} - \sum_{k\geq 2} \frac{\mathrm{ad}(\chi^{(1)})^k[\mathcal{K}^{(1)}]}{(k+1)!}.$$

By Lemma 5.3, $\mathcal{Q}^{(\geq 3)}$ has degree at least 3 and fulfills the quantitative estimate (5.4). To prove (iv), we use the fact that $\{\widetilde{\mathcal{M}}, \chi^{(1)}\} = \{\widetilde{\mathcal{P}}, \chi^{(1)}\} = 0$ since $\mathcal{K}^{(1)}$ commutes with $\widetilde{\mathcal{M}}$ and $\widetilde{\mathcal{P}}$, hence its monomials fulfill the selection rules of Remark 4.7. By the explicit formula for $\chi^{(1)}$ above, the same selection rules hold for $\chi^{(1)}$, and consequently $\mathcal{L}^{(1)}$ preserves $\widetilde{\mathcal{M}}$ and $\widetilde{\mathcal{P}}$.

It remains to prove the mapping properties of the operator $\mathcal{L}^{(1)}$. First we show that it maps $D(\rho/2, r/2) \to D(\rho, r)$. Let $(\widetilde{\mathcal{Y}}, \widetilde{\theta}, \widetilde{\mathbf{a}}) = \mathcal{L}^{(1)}(\mathcal{Y}, \theta, \mathbf{a})$. Then $(\widetilde{\mathcal{Y}}, \widetilde{\theta}, \widetilde{\mathbf{a}}) = (\widetilde{\mathcal{Y}}(s), \widetilde{\theta}(s), \widetilde{\mathbf{a}}(s))|_{s=1}$ where $(\widetilde{\mathcal{Y}}(s), \widetilde{\theta}(s), \widetilde{\mathbf{a}}(s))$ is the Hamiltonian flow generated by $\chi^{(1)}$ at time $0 \leq s \leq 1$. Using the identity

$$(\widetilde{\mathcal{Y}}(t), \widetilde{\theta}(t), \widetilde{\mathbf{a}}(t)) = (\mathcal{Y}, \theta, \mathbf{a}) + \int_0^t X_{\chi^{(1)}}(\widetilde{\mathcal{Y}}(s), \widetilde{\theta}(s), \widetilde{\mathbf{a}}(s))\, \mathrm{d}s$$

where $X_{\chi^{(1)}}$ is the Hamiltonian vector field associated with $\chi^{(1)}$ above, and a standard continuity (bootstrap) argument, we conclude that $(\widetilde{\mathcal{Y}}, \widetilde{\theta}, \widetilde{\mathbf{a}}) \in D(\rho, r)$. Similarly, one also deduces estimate (5.5). Finally, to prove that $\mathcal{L}^{(1)}$ maps $D(\rho/2, r/2) \cap h^1 \to h^1$, we note that $\widehat{\mathcal{N}}$ is equivalent to the square of the h^1 norm, and

$$\widehat{\mathcal{N}} \circ \mathcal{L}^{(1)} = \widehat{\mathcal{N}} + \mathcal{T}_1(\chi^{(1)}; \widehat{\mathcal{N}}) = \widehat{\mathcal{N}} - \sum_{k\geq 0} \frac{\mathrm{ad}(\chi^{(1)})^k[\mathcal{K}^{(1)}]}{(k+1)!} = \widehat{\mathcal{N}} + O(\sqrt{\varepsilon}\, r^3),$$

and this completes the proof. ∎

6. Analysis of the quartic part of the Hamiltonian

At this stage, we are left with the Hamiltonian $\mathcal{Q}$ given in (5.2). The aim of this section is to eliminate nonresonant terms from $\mathcal{Q}^{(2)}$. First note that $\mathcal{Q}^{(2)}$ contains monomials of one of the following two forms:

$$e^{\mathrm{i}\theta\cdot\ell}\, a_{\vec{j}_1}^{\sigma_1} a_{\vec{j}_2}^{\sigma_2} a_{\vec{j}_3}^{\sigma_3} a_{\vec{j}_4}^{\sigma_4} \quad \text{or} \quad e^{\mathrm{i}\theta\cdot\ell}\, \mathcal{Y}^l a_{\vec{j}_1}^{\sigma_1} a_{\vec{j}_2}^{\sigma_2} \quad \text{with } |l| = 1.$$

In order to cancel out the terms quadratic in a by a Birkhoff normal form procedure, we only need the *second Melnikov conditions* imposed in (4.5). In order to cancel out the quartic tems in a we need the *fourth Melnikov conditions*, namely to control expressions of the form

$$\omega(\lambda)\cdot\ell + \sigma_1\Omega_{\vec{j}_1}(\lambda,\varepsilon) + \sigma_2\Omega_{\vec{j}_2}(\lambda,\varepsilon) + \sigma_3\Omega_{\vec{j}_3}(\lambda,\varepsilon) + \sigma_4\Omega_{\vec{j}_4}(\lambda,\varepsilon), \quad \sigma_i = \pm 1. \tag{6.1}$$

We start by defining the following set $\mathfrak{R}_4 \subset \mathfrak{A}_4$ (see Definition 4.1):

$$\begin{aligned}\mathfrak{R}_4 := \big\{(\mathbf{j}, \ell, \sigma)\colon & \ell = 0 \text{ and } \vec{j}_1, \vec{j}_2, \vec{j}_3, \vec{j}_4 \notin \mathscr{S} \text{ form a rectangle, or} \\ & \ell = 0,\ \vec{j}_1, \vec{j}_2 \notin \mathscr{S}, \vec{j}_3, \vec{j}_4 \in \mathscr{S} \text{ form a horizontal rectangle (even degenerate), or} \\ & \ell \neq 0,\ \vec{j}_1, \vec{j}_2, \vec{j}_3 \in \mathscr{S},\ \vec{j}_4 \notin \mathscr{S} \text{ and } |m_4| < M_0, \text{ where } M_0 \text{ is a universal constant, or} \\ & \ell = 0,\ \vec{j}_1, \vec{j}_2, \vec{j}_3, \vec{j}_4 \in \mathscr{S} \text{ form a horizontal trapezoid}\big\} \end{aligned} \tag{6.2}$$

where $\mathscr{S}$ is the set defined in (4.2). Here a trapezoid (or a rectangle) is said to be *horizontal* if two of its sides are parallel to the x-axis.

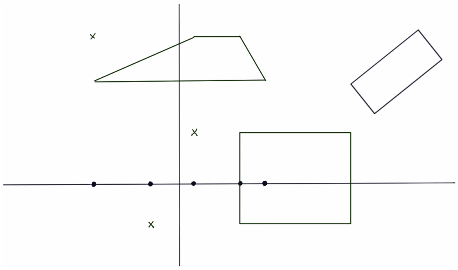

Fig. 1. The black dots are the points in $\mathcal{S}_0$. The two rectangles and the trapezoid correspond to cases 1, 2, 4 in $\mathfrak{R}_4$. In order to represent case 3, we have highlighted three points in $\mathcal{S}$. To each such triple we may associate at most one $\ell \neq 0$ and one $\vec{j}_4 \in \mathfrak{Z}$, which form a resonance of type 3.

Proposition 6.1. *Fix $0 < \varepsilon_2 < \varepsilon_1$ sufficiently small and $\tau_2 > \tau_1$ sufficiently large. There exist $\gamma_2 > 0$ and $\mathtt{L}_2 \geq \mathtt{L}_1$ (with $\mathtt{L}_2$ depending only on $\mathtt{d}$) such that for all $0 < \varepsilon \leq \varepsilon_2$ and for an $\mathtt{L}_2$-generic choice of the set $\mathcal{S}_0$ (in the sense of Definition 1.1), the set*

$$\mathcal{C}^{(2)} := \big\{\lambda \in \mathcal{C}^{(1)} : |\omega \cdot \ell + \sigma_1 \Omega_{\vec{j}_1}(\lambda, \varepsilon) + \sigma_2 \Omega_{\vec{j}_2}(\lambda, \varepsilon) + \sigma_3 \Omega_{\vec{j}_3}(\lambda, \varepsilon) + \sigma_4 \Omega_{\vec{j}_4}(\lambda, \varepsilon)| \\ \geq \gamma_2 \varepsilon / \langle \ell \rangle^{\tau_2},\ \forall (\vec{j}, \ell, \sigma) \in \mathfrak{A}_4 \setminus \mathfrak{R}_4\big\},$$

has positive measure and $|\mathcal{C}^{(1)} \setminus \mathcal{C}^{(2)}| \lesssim \varepsilon_2^{\kappa_2}$ for some $\kappa_2 > 0$ independent of ε_2.

The proof of the proposition, being quite technical, is postponed to Appendix A.

An immediate consequence, following the same strategy as for the proof of Theorem 5.2, is the following result. We define $\Pi_{\mathfrak{R}_4}$ as the projection of a function in $D(\rho, r)$ onto the sum of monomials with indices in $\mathfrak{R}_4$. Abusing notation, we define analogously $\Pi_{\mathfrak{R}_2}$ as the projection onto monomials $e^{\mathrm{i}\ell\cdot\theta} \mathcal{Y}^l a_{\vec{j}_1}^{\sigma_1} a_{\vec{j}_2}^{\sigma_2}$ with $|l| = 1$ and $(\vec{j}_1, \vec{j}_2, \ell, \sigma_1, \sigma_2) \in \mathfrak{R}_2$.

Theorem 6.2. *There exist $0 < r_2 \leq r_1$ and $0 < \rho_2 \leq \rho_1$ such that for all $0 < \varepsilon \leq \varepsilon_2$, all $\lambda \in \mathcal{C}^{(2)}$ and all $r \in [0, r_2]$, $\rho \in [\rho_2/2, \rho_2]$ there exists a symplectic change of variables $\mathcal{L}^{(2)}$ well defined and majorant analytic from $D(\rho/2, r/2)$ to $D(\rho, r)$ such that*

$$\mathcal{Q} \circ \mathcal{L}^{(2)}(\mathcal{Y}, \theta, \mathbf{a}) = \omega \cdot \mathcal{Y} + \sum_{\vec{j} \in \mathbb{Z}^2 \setminus \mathcal{S}_0} \Omega_{\vec{j}}(\lambda, \varepsilon) |a_{\vec{j}}|^2 + \mathcal{Q}^{(2)}_{\mathrm{Res}} + \widetilde{\mathcal{Q}}^{(\geq 3)}, \tag{6.3}$$

where

$$\mathcal{Q}^{(2)}_{\mathrm{Res}} = \Pi_{\mathfrak{R}_4}\mathcal{Q}^{(2)} + \Pi_{\mathfrak{R}_2}\mathcal{Q}^{(2)} \tag{6.4}$$

with $\mathfrak{R}_4$ *defined in* (6.2), $\mathfrak{R}_2$ *defined in* (4.4) *and*

$$|\mathcal{Q}^{(2)}_{\mathrm{Res}}|_{\rho/2,r/2} \lesssim r^2, \quad |\widetilde{\mathcal{Q}}^{(\geq 3)}|_{\rho/2,r/2} \lesssim r^3/\sqrt{\varepsilon}.$$

Moreover, $\mathcal{L}^{(2)}$ *maps* $D(\rho/2, r/2) \cap h^1 \to D(\rho, r) \cap h^1$, *and if we denote* $(\widetilde{\mathcal{Y}}, \widetilde{\theta}, \widetilde{\mathbf{a}}) = \mathcal{L}^{(2)}(\mathcal{Y}, \theta, \mathbf{a})$, *then*

$$\|\widetilde{\mathbf{a}} - \mathbf{a}\|_{\ell^1} \lesssim \|\mathbf{a}\|^3_{\ell^1}.$$

Proof. The proof is analogous to the one of Theorem 5.2, and we skip it. ∎

7. Construction of the toy model

Once we have performed (partial) Birkhoff normal form up to order 4, we can start applying the ideas developed in [12] to the Hamiltonian (6.3). Note that throughout this section $\varepsilon > 0$ is a fixed parameter. Namely, we do not use its smallness and we do not modify it.

We first apply the (time dependent) change of variables to rotating coordinates

$$a_{\vec{\jmath}} = \beta_{\vec{\jmath}} e^{\mathrm{i}\Omega_{\vec{\jmath}}(\lambda,\varepsilon)t}, \tag{7.1}$$

to the Hamiltonian (6.3), which leads to the corrected Hamiltonian

$$\mathcal{Q}_{\mathrm{rot}}(\mathcal{Y}, \theta, \beta, t) = \mathcal{Q} \circ \mathcal{L}^{(2)}(\mathcal{Y}, \theta, \{\beta_{\vec{\jmath}} e^{\mathrm{i}\Omega_{\vec{\jmath}}(\lambda,\varepsilon)t}\}_{\vec{\jmath}\in\mathbb{Z}^2_N\setminus\mathcal{S}_0}) - \sum_{\vec{\jmath}\in\mathbb{Z}^2_N\setminus\mathcal{S}_0} \Omega_{\vec{\jmath}}(\lambda,\varepsilon)|\beta_{\vec{\jmath}}|^2. \tag{7.2}$$

We split this Hamiltonian as a suitable first order truncation $\mathcal{G}$ plus two remainders,

$$\mathcal{Q}_{\mathrm{rot}}(\mathcal{Y}, \theta, \beta, t) = \mathcal{G}(\mathcal{Y}, \theta, \beta) + \mathcal{J}_1(\mathcal{Y}, \theta, \beta, t) + \mathcal{R}(\mathcal{Y}, \theta, \beta, t)$$

with

$$\begin{aligned}
\mathcal{G}(\mathcal{Y}, \theta, \beta) &= \omega \cdot \mathcal{Y} + \mathcal{Q}^{(2)}_{\mathrm{Res}}(\mathcal{Y}, \theta, \beta),\\
\mathcal{J}_1(\mathcal{Y}, \theta, \beta, t) &= \mathcal{Q}^{(2)}_{\mathrm{Res}}(\mathcal{Y}, \theta, \{\beta_{\vec{\jmath}} e^{\mathrm{i}\Omega_{\vec{\jmath}}(\lambda,\varepsilon)t}\}_{\vec{\jmath}\in\mathbb{Z}^2_N\setminus\mathcal{S}_0}) - \mathcal{Q}^{(2)}_{\mathrm{Res}}(\mathcal{Y}, \theta, \beta),\\
\mathcal{R}(\mathcal{Y}, \theta, \beta, t) &= \widetilde{\mathcal{Q}}^{(\geq 3)}(\mathcal{Y}, \theta, \{\beta_{\vec{\jmath}} e^{\mathrm{i}\Omega_{\vec{\jmath}}(\lambda,\varepsilon)t}\}_{\vec{\jmath}\in\mathbb{Z}^2_N\setminus\mathcal{S}_0}),
\end{aligned} \tag{7.3}$$

where $\mathcal{Q}^{(2)}_{\mathrm{Res}}$ and $\widetilde{\mathcal{Q}}^{(\geq 3)}$ are the Hamiltonians introduced in Theorem 6.2.

For the rest of this section we focus on the truncated Hamiltonian $\mathcal{G}$. Note that the remainder $\mathcal{J}_1$ is not smaller than $\mathcal{G}$. Nevertheless it will be smaller when evaluated on the particular solutions we consider. The term $\mathcal{R}$ is smaller than $\mathcal{G}$ for small data since it is the remainder of the normal form obtained in Theorem 6.2. Later in Section 8 we show that including the dismissed terms $\mathcal{J}_1$ and $\mathcal{R}$ barely alters the dynamics of the solutions of $\mathcal{G}$ that we analyze.

7.1. The finite set Λ

We now start constructing special dynamics for the Hamiltonian $\mathcal{G}$ with the aim of treating the contributions of $\mathcal{J}_1$ and $\mathcal{R}$ as remainder terms. Following [12], we do not study the full dynamics of $\mathcal{G}$ but we restrict the dynamics to invariant subspaces. Indeed, we shall construct a set $\Lambda \subset \mathfrak{Z} := (\mathbb{Z} \times N\mathbb{Z}) \setminus (\mathcal{S}_0 \cup \mathscr{S})$, for some large N, in such a way that it generates an invariant subspace (for the dynamics of $\mathcal{G}$) given by

$$U_\Lambda := \{\beta_{\vec{j}} = 0 : \vec{j} \notin \Lambda\}. \tag{7.4}$$

Thus, we consider the following definition.

Definition 7.1 (Completeness). We say that a set $\Lambda \subset \mathfrak{Z}$ is *complete* if U_Λ is invariant under the dynamics of $\mathcal{G}$.

Remark 7.2. It can be easily seen that if Λ is complete, then U_Λ is also invariant under the dynamics of $\mathcal{G} + \mathcal{J}_1$.

We construct a complete set $\Lambda \subset \mathfrak{Z}$ (see Definition 7.1) and we study the restriction on it of the dynamics of the Hamiltonian $\mathcal{G}$ in (7.3). Following [12], we impose several conditions on Λ to obtain dynamics as simple as possible.

The set Λ is constructed in two steps. First we construct a preliminary set $\Lambda_0 \subset \mathbb{Z}^2$ on which we impose numerous geometrical conditions. Later on we scale Λ_0 by a factor N to obtain $\Lambda \subset N\mathbb{Z} \times N\mathbb{Z} \subset \mathfrak{Z}$.

The set Λ_0 is "essentially" the one described in [12]. The crucial point in that paper is to choose carefully the modes so that each mode in Λ_0 only belongs to two rectangles with vertices in Λ_0. This allows one to simplify the dynamics considerably and makes it easier to analyze. Certainly, this requires imposing several conditions on Λ_0. We add some extra conditions to adapt the set Λ_0 to the particular setting of the present paper.

We split Λ_0 into $\mathfrak{g}$ disjoint generations, $\Lambda_0 = \Lambda_{01} \cup \cdots \cup \Lambda_{0\mathfrak{g}}$. We call a quadruplet $(\vec{j}_1, \vec{j}_2, \vec{j}_3, \vec{j}_4) \in \Lambda_0^4$ a *nuclear family* if $\vec{j}_1, \vec{j}_3 \in \Lambda_{0k}$, $\vec{j}_2, \vec{j}_4 \in \Lambda_{0,k+1}$, and the four vertices form a nondegenerate rectangle. Then, we require the following conditions.

- Property I_{Λ_0} (Closure): If $\vec{j}_1, \vec{j}_2, \vec{j}_3 \in \Lambda_0$ are three vertices of a rectangle, then the fourth vertex of that rectangle is also in Λ_0.
- Property II_{Λ_0} (Existence and uniqueness of spouse and children): For each $1 \le k < \mathfrak{g}$ and every $\vec{j}_1 \in \Lambda_{0k}$, there exists a unique spouse $\vec{j}_3 \in \Lambda_{0k}$ and unique (up to trivial permutations) children $\vec{j}_2, \vec{j}_4 \in \Lambda_{0,k+1}$ such that $(\vec{j}_1, \vec{j}_2, \vec{j}_3, \vec{j}_4)$ is a nuclear family in Λ_0.
- Property III_{Λ_0} (Existence and uniqueness of parents and siblings): For each $1 \le k < \mathfrak{g}$ and every $\vec{j}_2 \in \Lambda_{0,k+1}$ there exists a unique sibling $\vec{j}_4 \in \Lambda_{0,k+1}$ and unique (up to permutation) parents $\vec{j}_1, \vec{j}_3 \in \Lambda_{0k}$ such that $(\vec{j}_1, \vec{j}_2, \vec{j}_3, \vec{j}_4)$ is a nuclear family in Λ_0.
- Property IV_{Λ_0} (Nondegeneracy): A sibling of any frequency $\vec{j}$ is never equal to its spouse.

- Property V_{Λ_0} (Faithfulness): Apart from nuclear families, Λ_0 contains no other rectangles. In fact, by the closure property I_{Λ_0}, this also means that it contains no right angled triangles other than those coming from vertices of nuclear families.
- Property VI_{Λ_0}: There are no two elements in Λ_0 such that $\vec{j}_1 \pm \vec{j}_2 = 0$. There are no three elements in Λ_0 such that $\vec{j}_1 - \vec{j}_2 + \vec{j}_3 = 0$. If four points in Λ_0 satisfy $\vec{j}_1 - \vec{j}_2 + \vec{j}_3 - \vec{j}_4 = 0$ then either the relation is trivial or such points form a family.
- Property VII_{Λ_0}: There are no points in Λ_0 with one of the coordinates equal to zero, i.e.
$$\Lambda_0 \cap (\mathbb{Z} \times \{0\} \cup \{0\} \times \mathbb{Z}) = \emptyset.$$
- Property $VIII_{\Lambda_0}$: There are no two points in Λ_0 which form a right angle with 0.

Property I_{Λ_0} is just a rephrasing of the completeness condition introduced in Definition 7.1. Properties II_{Λ_0}, III_{Λ_0}, IV_{Λ_0}, V_{Λ_0} correspond to being a family tree as stated in [12].

Theorem 7.3. *Fix* $\mathtt{K} \gg 1$ *and* $s \in (0, 1)$*. Then there exist* $\mathfrak{g} \gg 1$*,* $A_0 \gg 1$*,* $\eta > 0$*, and a set* $\Lambda_0 \subset \mathbb{Z}^2$ *with*
$$\Lambda_0 = \Lambda_{01} \cup \cdots \cup \Lambda_{0\mathfrak{g}},$$
which satisfies conditions I_{Λ_0}*–*$VIII_{\Lambda_0}$ *and also*
$$\frac{\sum_{\vec{j} \in \Lambda_{0,\mathfrak{g}-1}} |\vec{j}|^{2s}}{\sum_{\vec{j} \in \Lambda_{03}} |\vec{j}|^{2s}} \geq \frac{1}{2} 2^{(1-s)(\mathfrak{g}-4)} \geq \mathtt{K}^2. \tag{7.5}$$
Moreover, for any $A \geq A_0$*, there exists a function* $f(\mathfrak{g})$ *satisfying*
$$e^{A^{\mathfrak{g}}} \leq f(\mathfrak{g}) \leq e^{2(1+\eta)A^{\mathfrak{g}}} \quad \text{for } \mathfrak{g} \text{ large enough}, \tag{7.6}$$
such that each generation Λ_{0k} *has* $2^{\mathfrak{g}-1}$ *disjoint frequencies* $\vec{j}$ *satisfying*
$$C^{-1} f(\mathfrak{g}) \leq |\vec{j}| \leq C 3^{\mathfrak{g}} f(\mathfrak{g}), \quad \vec{j} \in \Lambda_{0k}, \tag{7.7}$$
and
$$\frac{\sum_{\vec{j} \in \Lambda_{0k}} |\vec{j}|^{2s}}{\sum_{\vec{j} \in \Lambda_{0i}} |\vec{j}|^{2s}} \leq C e^{s\mathfrak{g}} \tag{7.8}$$
for any $1 \leq i < k \leq \mathfrak{g}$ *and some constant* $C > 0$ *independent of* $\mathfrak{g}$.

The construction of such sets was first done in [12] (see also [26–29]) where the authors construct sets Λ satisfying Properties I_Λ–V_Λ and estimate (7.8). The proof of Theorem 7.3 follows the same lines. Indeed, Properties VI_Λ–$VIII_\Lambda$ can be obtained through the same density argument. Finally, the estimate (7.7), even if not stated explicitly in [12], is an easy consequence of the proof in that paper (in [27–29] a slightly weaker estimate is used).

Remark 7.4. Note that $s \in (0,1)$ implies that we are constructing a backward cascade orbit (energy is transferred from high to low modes). This means that the modes in each generation of Λ_0 are just labeled in reverse order $\Lambda_{0j} \leftrightarrow \Lambda_{0,\mathfrak{g}-j+1}$ compared to the ones constructed in [12]. The second statement of Theorem 1.2 concerns $s > 1$ and therefore a forward cascade orbit (energy transferred from low to high modes). For this result, we need a set Λ_0 of the same kind as that of [12], which thus satisfies

$$\frac{\sum_{\vec{j}\in\Lambda_{0,\mathfrak{g}-1}} |\vec{j}|^{2s}}{\sum_{\vec{j}\in\Lambda_{03}} |\vec{j}|^{2s}} \geq \frac{1}{2} 2^{(s-1)(\mathfrak{g}-4)} \geq \mathrm{K}^2$$

instead of estimate (7.5).

We now scale Λ_0 by a factor N satisfying (4.1) and we denote $\Lambda := N\Lambda_0$. Note that the listed properties I_{Λ_0}–$\mathrm{VIII}_{\Lambda_0}$ are invariant under scaling. Thus, if they are satisfied by Λ_0, they are satisfied by Λ too.

Lemma 7.5. *There exists a set Λ satisfying all statements of Theorem 7.3 (with a different $f(\mathfrak{g})$ satisfying (7.6)) and also the following additional properties.*

(1) *If two points $\vec{j}_1, \vec{j}_2 \in \Lambda$ form a right angle with a point $(m, 0) \in \mathbb{Z} \times \{0\}$, then*

$$|m| \geq \sqrt{f(\mathfrak{g})}.$$

(2) *$\Lambda \subset N\mathbb{Z} \times N\mathbb{Z}$ with $N = f(\mathfrak{g})^{4/5}$.*

Proof. Consider any of the sets Λ obtained in Theorem 7.3. By Property $\mathrm{VIII}_{\Lambda_0}$ one has $m \neq 0$. Define $\vec{j}_3 = (m, 0)$. The condition for orthogonality is either

$$\text{(i) } (\vec{j}_1 - \vec{j}_2)\cdot(\vec{j}_3 - \vec{j}_2) = 0 \quad \text{or} \quad \text{(ii) } (\vec{j}_1 - \vec{j}_3)\cdot(\vec{j}_2 - \vec{j}_3) = 0.$$

Taking $\vec{j}_i = (m_i, n_i)$, $i = 1, 2$, condition (i) implies (after some computations) that

$$m = \frac{(n_1 - n_2)n_2 + (m_1 - m_2)m_2}{m_1 - m_2}.$$

Then since $|m_1 - m_2| \leq 2Cf(\mathfrak{g})3^{\mathfrak{g}}$ and the numerator is not zero, we have

$$|m| \geq \frac{1}{4Cf(\mathfrak{g})3^{\mathfrak{g}}} \geq \frac{1}{f(\mathfrak{g})^{3/2}}. \tag{7.9}$$

Now we consider condition (ii). One finds that m is a root of the quadratic equation

$$m^2 - (m_1 + m_2)m + (m_1 m_2 + n_1 n_2) = 0.$$

First we note that $m_1 m_2 + n_1 n_2 \neq 0$ by Property $\mathrm{VIII}_{\Lambda_0}$, since $m = 0$ cannot be a solution. Now consider the discriminant $\Delta = (m_1 + m_2)^2 - 4(m_1 m_2 + n_1 n_2)$. If $\Delta < 0$, then no right angle is possible. If $\Delta = 0$, then clearly $|m| \geq 1/2$, since once again $m = 0$ is not a solution. Finally, let $\Delta > 0$. Then

$$m = \frac{m_1 + m_2}{2}\left(1 \pm \sqrt{1 - \frac{4(m_1 m_2 + n_1 n_2)}{(m_1 + m_2)^2}}\right).$$

Denoting $\gamma := \frac{4(m_1 m_2 + n_1 n_2)}{(m_1+m_2)^2}$, the condition $\Delta > 0$ implies that $-\infty < \gamma < 1$. Splitting in two cases: $|\gamma| \leq 1$ and $\gamma < -1$ one can easily see that in either case m satisfies (7.9). Now it only remains to scale the set Λ by a factor $f(\mathfrak{g})^4$. Then, taking as new $f(\mathfrak{g})$, $\tilde{f}(\mathfrak{g}) := f(\mathfrak{g})^5$, the resulting set Λ satisfies all statements of Theorem 7.3 and also the statements of Lemma 7.5. ∎

7.2. *The truncated Hamiltonian on the finite set Λ and the toy model of* [12]

We use the properties of the set Λ given by Theorem 7.3 and Lemma 7.5 to compute the restriction of the Hamiltonian $\mathcal{G}$ in (7.3) to the invariant subset U_Λ (see (7.4)).

Lemma 7.6. *Consider the set $\Lambda \subset N\mathbb{Z} \times N\mathbb{Z}$ obtained in Theorem 7.3. Then the set*

$$\mathcal{M}_\Lambda = \{(\mathcal{Y}, \theta, \beta) : \mathcal{Y} = 0,\ \beta \in U_\Lambda\}$$

is invariant under the flow associated to the Hamiltonian $\mathcal{G}$. Moreover, $\mathcal{G}$ restricted to $\mathcal{M}_\Lambda$ can be written as

$$\mathcal{G}|_{\mathcal{M}_\Lambda}(\theta, \beta) = \mathcal{G}_0(\beta) + \mathcal{J}_2(\theta, \beta) \tag{7.10}$$

where

$$\mathcal{G}_0(\beta) = -\frac{1}{2}\sum_{\vec{j}\in\Lambda} |\beta_{\vec{j}}|^4 + \frac{1}{2} \sum^{\star}_{\substack{(\vec{j}_1,\vec{j}_2,\vec{j}_3,\vec{j}_4)\in\Lambda^4 \\ \vec{j}_i \text{ form a rectangle}}} \beta_{\vec{j}_1}\bar{\beta}_{\vec{j}_2}\beta_{\vec{j}_3}\bar{\beta}_{\vec{j}_4} \tag{7.11}$$

and the remainder $\mathcal{J}_2$ satisfies

$$|\mathcal{J}_2|_{\rho,r} \lesssim r^2 f(\mathfrak{g})^{-4/5}. \tag{7.12}$$

Proof. First we note that, since $\mathcal{Y} = 0$ on $\mathcal{M}_\Lambda$,

$$\mathcal{G}|_{\mathcal{M}_\Lambda} = \mathcal{Q}^{(2)}_{\text{Res}}|_{\mathcal{M}_\Lambda} = \Pi_{\mathfrak{R}_4}\mathcal{Q}^{(2)}|_{\mathcal{M}_\Lambda}$$

where $\mathcal{Q}^{(2)}_{\text{Res}}$ is the Hamiltonian defined in Theorem 6.2. We start by analyzing the Hamiltonian $\mathcal{Q}^{(2)}$ introduced in Theorem 5.2, which is defined as

$$\mathcal{Q}^{(2)} = \mathcal{K}^{(2)} + \tfrac{1}{2}\{\mathcal{K}^{(1)}, \chi^{(1)}\}.$$

We analyze each term. Here it plays a crucial role that $\Lambda \subset N\mathbb{Z} \times N\mathbb{Z}$ with $N = f(\mathfrak{g})^{4/5}$.

In order to estimate $\mathcal{K}^{(2)}$, defined in (5.1), we recall that Λ does not have any mode in the x-axis and therefore the original quartic Hamiltonian has not been modified by the Birkhoff map (1.6) (this is evident from the formula for $\mathcal{H}^{(2)}$ in (3.18)). Thus, it is enough to analyze how the quartic Hamiltonian has been modified by the linear change $\mathcal{L}^{(0)}$ analyzed in Theorems 4.3 and 4.4. Using the smoothing property of the change of coordinates $\mathcal{L}^{(0)}$ given in Theorem 4.3(5), one obtains

$$\Pi_{\mathfrak{R}_4}\mathcal{K}^{(2)}|_{\mathcal{M}_\Lambda} = -\frac{1}{2}\sum_{\vec{j}\in\Lambda}|a_{\vec{j}}|^4 + \frac{1}{2}\sum_{\text{Rectangles}\subset\Lambda} a_{\vec{j}_1}\bar{a}_{\vec{j}_2}a_{\vec{j}_3}\bar{a}_{\vec{j}_4} + O\left(\frac{r^2}{N}\right).$$

Now we deal with the term $\{\mathcal{K}^{(1)}, \chi^{(1)}\}$. Since we only need to analyze $\Pi_{\mathfrak{R}_4}\{\mathcal{K}^{(1)}, \chi^{(1)}\}|_{\mathcal{M}_\Lambda}$, we only need to consider monomials in $\mathcal{K}^{(1)}$ and in $\chi^{(1)}$ which have at least two indices in Λ. We represent this by setting

$$\chi^{(1)} = \chi^{(1)}_{\#\Lambda\leq 1} + \chi^{(1)}_{\#\Lambda\geq 2},$$

where $\#\Lambda \geq 2$ means that we restrict to those monomials which have at least two indices in Λ. We then have

$$\{\mathcal{K}^{(1)}, \chi^{(1)}\}|_{\mathcal{M}_\Lambda} = \{\mathcal{K}^{(1)}, \chi^{(1)}_{\#\Lambda\geq 2}\}|_{\mathcal{M}_\Lambda}.$$

We estimate the size of $\chi^{(1)}_{\#\Lambda\geq 2}$. As explained in the proof of Theorem 5.2, $\chi^{(1)}_{\#\Lambda\geq 2}$ has coefficients

$$\chi^{(1)}_{\ell,\mathbf{j},\vec{\sigma}} = \frac{\mathrm{i}\mathcal{K}^{(1)}_{\ell,\mathbf{j},\vec{\sigma}}}{\omega\cdot\ell + \sigma_1\Omega_{\vec{j}_1}(\lambda,\varepsilon) + \sigma_2\Omega_{\vec{j}_2}(\lambda,\varepsilon) + \sigma_3\Omega_{\vec{j}_3}(\lambda,\varepsilon)} \tag{7.13}$$

with $\vec{j}_2, \vec{j}_3 \in \Lambda$.

We first estimate the tails (in ℓ) of $\chi^{(1)}$ and then we analyze the finite number of cases left. For the tails, it is enough to use Theorem 5.2 to deduce the following estimate for any $\rho \leq \rho_1/2$, where ρ_1 is the constant introduced in that theorem:

$$\Big|\sum_{|\ell|>\sqrt[4]{N}} \chi^{(1)}_{\ell,\mathbf{j},\vec{\sigma}} e^{\mathrm{i}\theta\cdot\ell} a^{\sigma_1}_{\vec{j}_1} a^{\sigma_2}_{\vec{j}_2} a^{\sigma_3}_{\vec{j}_3}\Big|^{\mathcal{C}^{(1)}}_{\rho,r} \lesssim e^{-(\rho_1-\rho)\sqrt[4]{N}} |\chi^{(1)}|^{\mathcal{C}^{(1)}}_{\rho_1,r} \leq r e^{-(\rho_1-\rho)\sqrt[4]{N}}.$$

We restrict our attention to monomials with $|\ell| \leq \sqrt[4]{N}$. We take $\vec{j}_2, \vec{j}_3 \in \Lambda$ and we consider different cases depending on $\vec{j}_1$ and the properties of the monomial. In each case we show that the denominator of (7.13) is larger than N.

Case 1: $\vec{j}_1 \notin \mathscr{S}$. The selection rules are (according to Remark 4.7)

$$\eta(\ell)+\sigma_1+\sigma_2+\sigma_3=0,\;\; \vec{\mathtt{m}}\cdot\ell+\sigma_1 m_1+\sigma_2 m_2+\sigma_3 m_3=0,\;\; \sigma_1 n_1+\sigma_2 n_2+\sigma_3 n_3=0,$$

and the leading term in the denominator of (7.13) is

$$\vec{\mathtt{m}}^2\cdot\ell + \sigma_1|\vec{j}_1|^2 + \sigma_2|\vec{j}_2|^2 + \sigma_3|\vec{j}_3|^2 \tag{7.14}$$

where $\vec{\mathtt{m}}^2 = (\mathtt{m}_1^2, \ldots, \mathtt{m}_\mathtt{d}^2)$. We consider the following subcases:

A1: $\sigma_3 = \sigma_1 = +1, \sigma_2 = -1$. In this case $\vec{j}_1 - \vec{j}_2 + \vec{j}_3 - \mathtt{v} = 0$, where $\mathtt{v} := (-\vec{\mathtt{m}}\cdot\ell, 0)$. We rewrite (7.14) as

$$\vec{\mathtt{m}}^2\cdot\ell + (\vec{\mathtt{m}}\cdot\ell)^2 - (\vec{\mathtt{m}}\cdot\ell)^2 + |\vec{j}_1|^2 - |\vec{j}_2|^2 + |\vec{j}_3|^2 = \vec{\mathtt{m}}^2\cdot\ell + (\vec{\mathtt{m}}\cdot\ell)^2 - 2(\mathtt{v}-\vec{j}_3, \vec{j}_3-\vec{j}_2).$$

Assume first $\vec{j}_2 \neq \vec{j}_3$. Since the set Λ satisfies Lemma 7.5 (1), and $|\vec{\mathtt{m}}\cdot\ell| \lesssim \sqrt[4]{N} \lesssim f(\mathfrak{g})^{1/5}$, we can ensure that $\vec{j}_2$ and $\vec{j}_3$ do not form a right angle with $\mathtt{v}$, thus

$$(\mathtt{v}-\vec{j}_2, \vec{j}_3-\vec{j}_2) \in \mathbb{Z}\setminus\{0\}.$$

Actually by Lemma 7.5 (ii), $\vec{\jmath}_3 - \vec{\jmath}_2 \in N\mathbb{Z}^2$ and hence, using also $|\ell| \leq \sqrt[4]{N}$,

$$|\vec{\mathtt{m}}^2 \cdot \ell + (\vec{\mathtt{m}} \cdot \ell)^2 - 2(\mathtt{v} - \vec{\jmath}_3, \vec{\jmath}_3 - \vec{\jmath}_2)| \geq 2N - N/8 > N.$$

Now there remains the case $\vec{\jmath}_2 = \vec{\jmath}_3$. Such monomials cannot exist in $\mathcal{H}^{(1)}$ in (3.17) since the monomials with two equal modes have been removed in (3.3) (it does not support degenerate rectangles). Naturally a degenerate rectangle may appear after we apply the change $\mathcal{L}^{(0)}$ introduced in Theorem 4.3. Nevertheless, the map $\mathcal{L}^{(0)}$ is identity plus smoothing (see statement (5) of that theorem), which leads to the needed N^{-1} factor.

B1: $\sigma_3 = \sigma_2 = +1, \sigma_1 = -1$. Now the selection rule reads $-\vec{\jmath}_1 + \vec{\jmath}_2 + \vec{\jmath}_3 - \mathtt{v} = 0$, with again $\mathtt{v} = (-\vec{\mathtt{m}} \cdot \ell, 0)$. We rewrite (7.14) as

$$\vec{\mathtt{m}}^2 \cdot \ell + (\vec{\mathtt{m}} \cdot \ell)^2 - (\vec{\mathtt{m}} \cdot \ell)^2 - |\vec{\jmath}_1|^2 + |\vec{\jmath}_2|^2 + |\vec{\jmath}_3|^2 = \vec{\mathtt{m}}^2 \cdot \ell + (\vec{\mathtt{m}} \cdot \ell)^2 - 2(\mathtt{v} - \vec{\jmath}_3, \mathtt{v} - \vec{\jmath}_2).$$

By Lemma 7.5 (1), $(\mathtt{v} - \vec{\jmath}_2, \mathtt{v} - \vec{\jmath}_3) \neq 0$. By Property VIII$_\Lambda$ and Lemma 7.5 (2), one has $|(\vec{\jmath}_2, \vec{\jmath}_3)| \geq N^2$ and estimate (7.7) implies $|\vec{\jmath}_2|, |\vec{\jmath}_3| \leq N^{3/2}$. Then

$$|(\mathtt{v} - \vec{\jmath}_2, \mathtt{v} - \vec{\jmath}_3)| \geq |(\vec{\jmath}_2, \vec{\jmath}_3)| - |(\mathtt{v}, \vec{\jmath}_2 + \vec{\jmath}_3)| - |\mathtt{v}|^2 \geq N^2/4$$

and one concludes as in **A1**.

C1: $\sigma_1 = \sigma_3 = \sigma_2 = +1$. The denominator (7.14) satisfies

$$\left|\vec{\mathtt{m}}^2 \cdot \ell + |\vec{\jmath}_1|^2 + |\vec{\jmath}_2|^2 + |\vec{\jmath}_3|^2\right| \geq 2N - |\vec{\mathtt{m}}^2 \cdot \ell| \geq 2N - N/8 \geq N.$$

This completes the proof of Case 1.

Case 2: $\vec{\jmath}_1 \in \mathcal{S}$. The selection rules are

$$\eta(\ell) + \sigma_2 + \sigma_3 = 0, \quad \vec{\mathtt{m}} \cdot \ell + \sigma_2 m_2 + \sigma_3 m_3 = 0, \quad \sigma_1 n_1 + \sigma_2 n_2 + \sigma_3 n_3 = 0,$$

and the leading term in the denominator is

$$\vec{\mathtt{m}}^2 \cdot \ell + \sigma_1 n_1^2 + \sigma_2 |\vec{\jmath}_2|^2 + \sigma_3 |\vec{\jmath}_3|^2, \tag{7.15}$$

where $\vec{\mathtt{m}}^2 = (\mathtt{m}_1^2, \ldots, \mathtt{m}_\mathtt{d}^2)$. We can reduce Case 2 to Case 1.

B2: $\sigma_2 = \sigma_3 = +1, \sigma_1 = -1$. We can assume that $\vec{\jmath}_1 = (\mathtt{m}_1, n_1)$. Define $\tilde{\ell} = \ell + \mathbf{e}_1$, and deduce from the selection rules and (7.15) that

$$\vec{\mathtt{m}} \cdot \tilde{\ell} - \mathtt{m}_1 + m_2 + m_3 = \vec{\mathtt{m}} \cdot \ell + m_2 + m_3 = 0.$$

Then the leading term in the denominator becomes

$$\vec{\mathtt{m}}^2 \cdot \tilde{\ell} - (\mathtt{m}_1^2 + n_1^2) + |\vec{\jmath}_2|^2 + |\vec{\jmath}_3|^2$$

and one proceeds as in case **B1** with $\tilde{\ell}$ in place of ℓ.

Cases **A2** and **C2**, defined analogously to **A1** and **C1** in Case 1, are completely equivalent.

In conclusion, we have proved that

$$\left|\chi^{(1)}_{\#\Lambda\geq 2}|_{\mathcal{M}_\Lambda}\right|^{\mathcal{C}(1)}_{\rho,r} \leq rN^{-1}. \tag{7.16}$$

Item (i) of Lemma 5.3, jointly with estimate (7.16), implies that, for $\rho' \in (0, \rho/2]$ and $r' \in (0, r/2]$

$$\left|\{\mathcal{K}^{(1)}, \chi^{(1)}_{\#\Lambda\geq 2}\}|_{\mathcal{M}_\Lambda}\right|^{\mathcal{C}(1)}_{\rho',r'} \lesssim r^2 N^{-1}.$$

This completes the proof of Lemma 7.6. ∎

The Hamiltonian $\mathcal{G}_0$ in (7.11) is the Hamiltonian that the I-team derived to construct their toy model. A posteriori we will check that the remainder $\mathcal{J}_2$ plays a minor role in our analysis.

The properties of Λ imply that the equation associated to $\mathcal{G}_0$ reads

$$\mathrm{i}\dot{\beta}_{\vec{j}} = -\beta_{\vec{j}}|\beta_{\vec{j}}|^2 + 2\beta_{\vec{j}_{\text{child}_1}}\beta_{\vec{j}_{\text{child}_2}}\overline{\beta_{\vec{j}_{\text{spouse}}}} + 2\beta_{\vec{j}_{\text{parent}_1}}\beta_{\vec{j}_{\text{parent}_2}}\overline{\beta_{\vec{j}_{\text{sibling}}}} \tag{7.17}$$

for each $\vec{j} \in \Lambda$. In the first and last generations, the parents and children are set to zero respectively. Moreover, the particular form of this equation implies the following corollary.

Corollary 7.7 ([12]). *Consider the subspace*

$$\widetilde{U}_\Lambda = \{\beta \in U_\Lambda : \beta_{\vec{j}_1} = \beta_{\vec{j}_2} \textit{ for all } \vec{j}_1, \vec{j}_2 \in \Lambda_k \textit{ for some } k\},$$

where all the members of a generation take the same value. Then $\widetilde{U}_\Lambda$ is invariant under the flow associated to the Hamiltonian $\mathcal{G}_0$. Therefore, equation (7.17) restricted to $\widetilde{U}_\Lambda$ becomes

$$\mathrm{i}\dot{b}_k = -b_k^2\overline{b}_k + 2\overline{b}_k(b_{k-1}^2 + b_{k+1}^2), \quad k = 1, \dots, \mathfrak{g}, \tag{7.18}$$

where

$$b_k = \beta_{\vec{j}} \quad \textit{for any } \vec{j} \in \Lambda_k. \tag{7.19}$$

The dimension of $\widetilde{U}_\Lambda$ is $2\mathfrak{g}$, where $\mathfrak{g}$ is the number of generations. In [12] and [28], the authors construct certain orbits of the *toy model* (7.18) which shift its mass from being localized at b_3 to being localized at $b_{\mathfrak{g}-1}$. These orbits will lead to orbits of the original equation (2D-NLS) undergoing growth of Sobolev norms.

Theorem 7.8 ([28]). *Fix a large $\gamma \gg 1$. Then for any large enough $\mathfrak{g}$ and $\mu = e^{-\gamma\mathfrak{g}}$, there exist an orbit of system (7.18) and $T_0 > 0$ such that*

$$|b_3(0)| > 1 - \mu, \qquad |b_{\mathfrak{g}-1}(T_0)| > 1 - \mu,$$
$$|b_i(0)| < \mu \quad \textit{for } i \neq 3, \qquad |b_i(T_0)| < \mu \quad \textit{for } i \neq \mathfrak{g} - 1.$$

Moreover, there exists a constant $C > 0$ independent of $\mathfrak{g}$ such that

$$0 < T_0 < C\mathfrak{g}\ln(1/\mu) = C\gamma\mathfrak{g}^2.$$

This theorem is proven in [12] without time estimates. The time estimates were obtained in [28].

8. The approximation argument

In Sections 4–6 we have applied several transformations and in Sections 6 and 7 we have removed certain small remainders. This has allowed us to derive a simple equation, called the toy model in [12]; then, in Section 7, we have analyzed some special orbits of this system. The last step of the proof of Theorem 1.2 is to show that when incorporating back the removed remainders ($\mathcal{J}_1$ and $\mathcal{R}$ in (7.3) and $\mathcal{J}_2$ in (7.10)) and undoing the changes of coordinates performed in Theorems 4.3 and 5.2, in Proposition 6.2 and in (7.1), the toy model orbit obtained in Theorem 7.8 leads to a solution of the original equation (2D-NLS) undergoing growth of Sobolev norms.

Now we analyze each remainder and each change of coordinates. From the orbit obtained in Theorem 7.8 and using (7.19) one can obtain an orbit of the Hamiltonian (7.11). Moreover, both (7.11) and (7.18) are invariant under the scaling

$$b^{\nu}(t) = \nu^{-1} b(\nu^{-2} t). \tag{8.1}$$

By Theorem 7.8, the solution $b(t)$ is thus defined on the interval $[0, T]$, where

$$T = \nu^2 T_0 \leq \nu^2 C \gamma \mathtt{g}^2, \tag{8.2}$$

where T_0 is the time obtained in Theorem 7.8.

Now we prove that one can construct a solution of the Hamiltonian (7.2) "close" to the orbit β^{ν} of the Hamiltonian (7.11) defined as

$$\beta^{\nu}_{\vec{j}}(t) = \begin{cases} \nu^{-1} b_k(\nu^{-2} t) & \text{for } \vec{j} \in \Lambda_k, \\ 0 & \text{for } \vec{j} \notin \Lambda, \end{cases} \tag{8.3}$$

where $b(t)$ is the orbit given by Theorem 7.8. Note that this implies incorporating the remainders in (7.3) and (7.10).

We take a large ν so that (8.3) is small. In the original coordinates this will correspond to solutions close to the finite gap solution. Taking $\mathcal{J} = \mathcal{J}_1 + \mathcal{J}_2$ (see (7.3) and (7.10)), the equations for β and $\mathcal{Y}$ associated to the Hamiltonian (7.2) can be written as

$$\begin{aligned} \mathrm{i}\dot{\beta} &= \partial_{\overline{\beta}} \mathcal{G}_0(\beta) + \partial_{\overline{\beta}} \mathcal{J}(\mathcal{Y}, \theta, \beta) + \partial_{\overline{\beta}} \mathcal{R}(\mathcal{Y}, \theta, \beta), \\ \dot{\mathcal{Y}} &= -\partial_{\theta} \mathcal{J}(\mathcal{Y}, \theta, \beta) - \partial_{\theta} \mathcal{R}(\mathcal{Y}, \theta, \beta). \end{aligned} \tag{8.4}$$

Now we estimate the closeness of the orbit of the toy model obtained in Theorem 7.8 and orbits of (7.2).

Theorem 8.1. *Fix* $0 < s_1 < s_2 < 1$. *Consider a solution*

$$(\mathcal{Y}, \theta, \beta) = (0, \theta_0, \beta^{\nu}(t))$$

of the Hamiltonian (7.11) *for any* $\theta_0 \in \mathbb{T}^{\mathtt{d}}$, *where* $\beta^{\nu}(t) = \{\beta^{\nu}_{\vec{j}}(t)\}_{\vec{j} \in \mathbb{Z}^2_N \setminus \mathcal{S}_0}$ *is the solution given by* (8.3). *Assume*

$$f(\mathtt{g})^{s_1} \leq \nu \leq f(\mathtt{g})^{s_2}. \tag{8.5}$$

Then there exists σ *(depending on* s_1, s_2 *but independent of* $\mathfrak{g}$ *and* γ*) such that any solution* $(\mathcal{Y}(t), \theta(t), \widetilde{\beta}(t))$ *of* (7.2) *with initial condition* $\widetilde{\beta}(0) = \widetilde{\beta}^0 \in \ell^1$, $\mathcal{Y}(0) = \mathcal{Y}^0 \in \mathbb{R}^{\mathrm{d}}$ *with* $\|\widetilde{\beta}^0 - \beta^\nu(0)\|_{\ell^1} \le \nu^{-1-4\sigma}$ *and* $|\mathcal{Y}^0| \le \nu^{-2-4\sigma}$ *and any* $\theta(0) = \theta_1 \in \mathbb{T}^{\mathrm{d}}$ *satisfies*

$$\|\widetilde{\beta}_{\vec{j}}(t) - \beta^\nu_{\vec{j}}(t)\|_{\ell^1} \le \nu^{-1-\sigma}, \quad |\mathcal{Y}(t)| \le \nu^{-2-\sigma},$$

for $0 < t < T$*, where* T *is the time defined in* (8.2).

The proof is deferred to Section 8.1. Note that the change to rotating coordinates in (7.1) does not alter the ℓ^1 norm and therefore a similar result can be stated for orbits of the Hamiltonian (6.3) (the modulus in (6.3) adding the rotating phase).

Proof of Theorem 1.2. We use Theorem 8.1 to obtain a solution of the Hamiltonian (3.14) undergoing growth of Sobolev norms. Then the same property will hold true for the corresponding solution of the Hamiltonian (3.3), by applying the inverse of the Birkhoff map Φ in Theorem 3.1, which leaves untouched the modes $a_{(m,n)}$ with $n \neq 0$. We consider the solution $(\mathcal{Y}^*(t), \theta^*(t), \mathbf{a}^*(t))$ of the former Hamiltonian with initial condition

$$\begin{aligned} \mathcal{Y}^* &= 0, \\ \theta^* &= \theta_0, \\ a^*_{\vec{j}} &= \begin{cases} \nu^{-1} b_k(0) & \text{for } \vec{j} \in \Lambda_k, \\ 0 & \text{for } \vec{j} \notin \Lambda_k, \end{cases} \end{aligned} \tag{8.6}$$

for an arbitrary choice of $\theta_0 \in \mathbb{T}^{\mathrm{d}}$. Since this initial condition has finite support, it follows by applying (the inverse of) the Birkhoff map Φ in Theorem 3.1 that the corresponding initial condition in the original coordinates $u_{\vec{j}}$ belongs to h^s for all $s \ge 0$. Then the well-posedness properties of equation (2D-NLS) imply that the solution $u(t,x)$ with this initial condition belongs to $\bigcap_{s \ge 0} H^s(\mathbb{T}^2, \mathbb{C})$ for all times. Also, note that the distance from $u(0)$ to the torus $\mathcal{T}^I_{\mathcal{S}_0}$ is measured by the amplitude of $\{a^*_{\vec{j}}\}$ (see (3.10)).

We need to prove that Theorem 8.1 applies to this solution with ν satisfying (8.5). To this end, we perform the changes of coordinates given in Theorems 4.3, 5.2 and 6.2, keeping track of the ℓ^1 norm.

For $\mathcal{L}^{(j)}$, $j = 1, 2$, Theorems 5.2 and 6.2 imply the following. For $(\mathcal{Y}, \theta, \mathbf{a}) \in D(\rho, r)$ define $\pi_{\mathbf{a}}(\mathcal{Y}, \theta, \mathbf{a}) := \mathbf{a}$. Then

$$\|\pi_{\mathbf{a}} \mathcal{L}^{(j)}(\mathcal{Y}, \theta, \mathbf{a}) - \mathbf{a}\|_{\ell^1} \lesssim \|\mathbf{a}\|^2_{\ell^1}. \tag{8.7}$$

This estimate is not true for the change of coordinates $\mathcal{L}^{(0)}$ given in Theorem 4.3. Nevertheless, this change is smoothing (see Theorem 4.3 (5)). This implies that if all $\vec{j} \in \operatorname{supp}\{\mathbf{a}\}$ satisfy $|\vec{j}| \ge J$ then

$$\|\pi_{\mathbf{a}} \mathcal{L}^{(0)}(\mathcal{Y}, \theta, \mathbf{a}) - \mathbf{a}\|_{\ell^1} \lesssim J^{-1} \|\mathbf{a}\|_{\ell^1}. \tag{8.8}$$

Thanks to Theorem 7.3 (more precisely (7.7)), we can apply this estimate to (8.6) with $J = Cf(\mathfrak{g})$. Using the fact that $\|\mathbf{a}^*\|_{\ell^1} \lesssim \nu^{-1} \mathfrak{g} 2^{\mathfrak{g}}$ and the condition on ν in (8.5) (which implies $f(\mathfrak{g})^{-1} \le \nu^{-1}$), one can check

$$\|\pi_{\mathbf{a}} \mathcal{L}^{(0)}(0, \theta^*, \mathbf{a}^*) - \mathbf{a}^*\|_{\ell^1} \lesssim \nu^{-1} \mathfrak{g} 2^{\mathfrak{g}} f(\mathfrak{g})^{-1} \lesssim \nu^{-3/2}.$$

Therefore,

$$\|\pi_{\mathbf{a}}(\mathcal{L}^{(2)} \circ \mathcal{L}^{(1)} \circ \mathcal{L}^{(0)}(0, \theta^*, \mathbf{a}^*)) - \mathbf{a}^*\|_{\ell^1} \lesssim \nu^{-3/2}.$$

We define $(\widetilde{\mathcal{Y}}^*, \widetilde{\theta}^*, \widetilde{\mathbf{a}}^*)$ to be the image of the point (8.6) under the composition of these three changes. We apply Theorem 8.1 to the solution of (7.2) with this initial condition. Note that Theorem 8.1 is stated in rotating coordinates (see (7.1)). Nevertheless, since this change is the identity on the initial conditions, one does not need to make any further modification. Moreover, the change (7.1) leaves invariant both the ℓ^1 and Sobolev norms. We show that the solution $(\widetilde{\mathcal{Y}}^*(t), \widetilde{\theta}^*(t), \widetilde{\mathbf{a}}^*(t))$ expressed in the original coordinates satisfies the desired growth of Sobolev norms.

Define

$$S_i = \sum_{\vec{j} \in \Lambda_i} |\vec{j}|^{2s} \quad \text{for } i = 1, \dots, \mathfrak{g}.$$

We first estimate the initial and final Sobolev norms of the solution $(\mathcal{Y}^*(t), \theta^*(t), \mathbf{a}^*(t))$ in terms of the constants S_i; namely, we prove

$$\frac{1}{2}\nu^{-2} S_3 \leq \|\mathbf{a}^*(0)\|_{h^s}^2 \leq 2\nu^{-2} S_3 \quad \text{and} \quad \|\mathbf{a}^*(T)\|_{h^s}^2 \geq \frac{\nu^{-2}}{4} S_{\mathfrak{g}-1}. \tag{8.9}$$

The initial condition (8.6) for the orbit under consideration has support Λ (recall that $\mathcal{Y} = 0$). Therefore,

$$\|\mathbf{a}^*(0)\|_{h^s}^2 = \sum_{i=1}^{\mathfrak{g}} \sum_{\vec{j} \in \Lambda_i} |\vec{j}|^{2s} \nu^{-2} |b_i(0)|^2.$$

Then, taking into account Theorem 7.8,

$$\left|\|\mathbf{a}^*(0)\|_{h^s}^2 - \nu^{-2} S_3\right| \leq 3\nu^{-2}\mu S_3 + \nu^{-2}\mu^2 \sum_{i \neq 3} S_i \leq \nu^{-2} S_3 \left(3\mu + \mu^2 \sum_{i \neq 3} \frac{S_i}{S_3}\right).$$

From Theorem 7.3 we know that $S_i/S_3 \lesssim e^{s\mathfrak{g}}$ for $i \neq 3$. Therefore, to bound these terms we use the definition of μ from Theorem 7.8. Taking $\gamma > 1/2$ and $\mathfrak{g}$ large enough, we obtain the first estimate in (8.9).

To obtain the second estimate in (8.9) (final Sobolev norm), note that

$$\|\mathbf{a}^*(T)\|_{h^s}^2 \geq \sum_{\vec{j} \in \Lambda_{\mathfrak{g}-1}} |\vec{j}|^{2s} |a_{\vec{j}}^*(T)|^2 \geq S_{\mathfrak{g}-1} \inf_{\vec{j} \in \Lambda_{\mathfrak{g}-1}} |a_{\vec{j}}^*(T)|^2. \tag{8.10}$$

Thus, it is enough to obtain a lower bound for $|a_{\vec{j}}^*(T)|$ for $\vec{j} \in \Lambda_{\mathfrak{g}-1}$. To do so, we need to express $\mathbf{a}^*$ in normal form coordinates and use Theorem 8.1. We split $|a_{\vec{j}}^*(T)|$ as follows. Let $(\widetilde{\mathcal{Y}}^*(t), \widetilde{\theta}^*(t), \widetilde{\mathbf{a}}^*(t))$ be the image of the orbit with initial condition (8.6) under the changes of variables in Theorems 4.3 and 5.2, in Proposition 6.2 and in (7.1). Then

$$|a_{\vec{j}}^*(T)| \geq |\beta_{\vec{j}}^{\nu}(T)| - |\widetilde{a}_{\vec{j}}^*(T) - \beta_{\vec{j}}^{\nu}(T) e^{\mathrm{i}\Omega_{\vec{j}}(\lambda,\varepsilon)T}| - |\widetilde{a}_{\vec{j}}^*(T) - a_{\vec{j}}^*(T)|.$$

The first term, by Theorem 7.8, satisfies $|\beta^{\nu}_{\vec{j}}(T)| \geq \nu^{-1}/2$. For the second one, using Theorem 8.1, we have

$$|\widetilde{a}^*_{\vec{j}}(T) - \beta^{\nu}_{\vec{j}}(T) e^{\mathrm{i}\Omega_{\vec{j}}(\lambda,\varepsilon)T}| \leq \nu^{-1-\sigma}.$$

Finally, taking into account the estimates (8.7) and (8.8), one can bound the third term as

$$|\widetilde{a}^*_{\vec{j}}(T) - a^*_{\vec{j}}(T)| \leq \|\widetilde{\mathbf{a}}^*(T) - \mathbf{a}^*(T)\|_{\ell^1} \lesssim \|\mathbf{a}^*(T)\|^2_{\ell^1} + \frac{\|\mathbf{a}^*(T)\|_{\ell^1}}{|\vec{j}|}.$$

Now, by Theorems 8.1 and 7.3 (more precisely the fact that $|\vec{j}| \gtrsim f(\mathfrak{g})$ for $\vec{j} \in \Lambda$),

$$|\widetilde{a}^*_{\vec{j}}(T) - a^*_{\vec{j}}(T)| \leq \nu^{-1-\sigma}$$

(taking a smaller σ if necessary). Thus, by (8.10), we obtain the second estimate in (8.9).

The last step to prove Theorem 1.2 is to choose suitable ν and $\mathfrak{g}$ in terms of the parameters $K \gg 1$ and $0 < \delta \ll 1$.

To measure the growth of Sobolev norms, note that (8.9) implies

$$\frac{\|\mathbf{a}^*(T)\|^2_{h^s}}{\|\mathbf{a}^*(0)\|^2_{h^s}} \geq \frac{S_{\mathfrak{g}-1}}{8 S_3} \geq \frac{1}{16} 2^{(1-s)(\mathfrak{g}-4)}.$$

Thus, taking $\mathfrak{g} \sim \ln(K/\delta)$, one obtains the growth of Sobolev norm

$$\frac{\|\mathbf{a}^*(T)\|^2_{h^s}}{\|\mathbf{a}^*(0)\|^2_{h^s}} \gtrsim \frac{K^2}{\delta^2}.$$

To control the initial Sobolev norm, we need $\|\mathbf{a}^*(0)\|^2_{h^s} \sim \nu^{-2} S_3 \sim \delta^2$. Note that this estimate and the ones just obtained imply $\|\mathbf{a}^*(T)\|^2_{h^s} \gtrsim K^2$. To estimate $\|\mathbf{a}^*(0)\|^2_{h^s}$, it is enough to choose a suitable ν (as a function of $\mathfrak{g}$). To this end, note that by Theorem 7.3,

$$C^{-1} 2^{\mathfrak{g}} f(\mathfrak{g})^{2s} \leq S_3 \leq C 2^{\mathfrak{g}} 3^{\mathfrak{g}} f(\mathfrak{g})^{2s}.$$

Thus, one can take $\nu^2 \sim \delta^{-2} S_3$, which satisfies

$$f(\mathfrak{g})^s \lesssim \nu \lesssim \delta^{-1} 2^{\mathfrak{g}/2} 3^{\mathfrak{g}/2} f(\mathfrak{g})^s.$$

Then, choosing the parameters s_1, s_2 introduced in Theorem 8.1 such that $0 < s_1 < s < s_2 < 1$, one sees that, taking $\mathfrak{g}$ large enough (recall that we have chosen $\mathfrak{g} \sim \ln(K/\delta)$, which we can make arbitrarily big by enlarging K if necessary), the chosen ν belongs to the range admitted for ν in Theorem 8.1. This gives

$$C^{-1}\delta^2 \leq \|\mathbf{a}^*(0)\|^2_{h^s} \leq C\delta^2$$

for some $C > 0$ independent of δ. Note that in the reasoning above, to obtain small initial Sobolev h^s norm it has been crucial that $s \in (0, 1)$.

Remark 8.2. In case we require only the ℓ^2 norm of $\mathbf{a}^*(0)$ to be small, we can drop the condition $s < 1$. Indeed, $\|\mathbf{a}^*(0)\|_{\ell^2} \lesssim \nu^{-1} 2^{\mathfrak{g}} \mathfrak{g}$, which can be made arbitrarily small by simply taking $\mathfrak{g}$ large enough (and ν as in (8.5)).

The time estimates can be easily deduced from (8.2), (8.5), (7.6) and Theorem 7.8, which concludes the proof of Theorem 1.2 (1).

For the proof of the second statement of Theorem 1.2 it is enough to point out that the condition $s < 1$ has only been used in imposing that the initial Sobolev norm is small. The estimate for the ℓ^2 norm can be obtained as explained in Remark 8.2. ∎

8.1. *Proof of Theorem* 8.1

To prove Theorem 8.1, we define

$$\xi = \beta - \beta^\nu(t).$$

We use the equations in (8.4) to deduce an equation for ξ. It can be written as

$$\mathrm{i}\dot{\xi} = \mathcal{Z}_0(t) + \mathcal{Z}_1(t)\xi + \mathcal{Z}_1'(t)\overline{\xi} + \mathcal{Z}_1''(t)\mathcal{Y} + \mathcal{Z}_2(\xi, \mathcal{Y}, t), \tag{8.11}$$

where

$$\begin{aligned}
\mathcal{Z}_0(t) &= \partial_{\overline{\beta}}\mathcal{J}(0, \theta, \beta^\nu) + \partial_{\overline{\beta}}\mathcal{R}(0, \theta, \beta^\nu),\\
\mathcal{Z}_1(t) &= \partial_{\beta\overline{\beta}}\mathcal{G}_0(\beta^\nu) + \partial_{\beta\overline{\beta}}\mathcal{J}(0, \theta, \beta^\nu),\\
\mathcal{Z}_1'(t) &= \partial_{\overline{\beta}\overline{\beta}}\mathcal{G}_0(\beta^\nu) + \partial_{\overline{\beta}\overline{\beta}}\mathcal{J}(0, \theta, \beta^\nu),\\
\mathcal{Z}_1''(t) &= \partial_{\mathcal{Y}\overline{\beta}}\mathcal{G}_0(\beta^\nu) + \partial_{\mathcal{Y}\overline{\beta}}\mathcal{J}(0, \theta, \beta^\nu),\\
\mathcal{Z}_2(t) &= \partial_{\overline{\beta}}\mathcal{G}_0(\beta^\nu + \xi) - \partial_{\overline{\beta}}\mathcal{G}_0(\beta^\nu) - \partial_{\beta\overline{\beta}}\mathcal{G}_0(\beta^\nu)\xi\\
&\quad - \partial_{\overline{\beta}\overline{\beta}}\mathcal{G}_0(\beta^\nu)\overline{\xi} + \partial_{\overline{\beta}}\mathcal{J}(\mathcal{Y}, \theta, \beta^\nu + \xi) - \partial_{\overline{\beta}}\mathcal{J}(0, \theta, \beta^\nu)\\
&\quad - \partial_{\beta\overline{\beta}}\mathcal{J}(0, \theta, \beta^\nu)\xi - \partial_{\overline{\beta}\overline{\beta}}\mathcal{J}(0, \theta, \beta^\nu)\overline{\xi} - \partial_{\mathcal{Y}\overline{\beta}}\mathcal{J}(0, \theta, \beta^\nu)\mathcal{Y}\\
&\quad + \partial_{\overline{\beta}}\mathcal{R}(\mathcal{Y}, \theta, \beta^\nu + \xi) - \partial_{\overline{\beta}}\mathcal{R}(0, \theta, \beta^\nu).
\end{aligned} \tag{8.12}$$

We analyze the equations for ξ in (8.11) and $\mathcal{Y}$ in (8.4).

Lemma 8.3. *Assume that* $(\beta^\nu, \mathcal{Y}), (\beta^\nu + \xi, \mathcal{Y}) \in D(r_2)$ (*see* (2.1)) *where* r_2 *has been given by Theorem* 6.2. *Then*

$$\begin{aligned}
\frac{d}{dt}\|\xi\|_{\ell^1} \leq{}& C\nu^{-4}\mathfrak{g}^4 2^{4\mathfrak{g}} + C\nu^{-3}\mathfrak{g}^3 2^{3\mathfrak{g}}(f(\mathfrak{g})^{-4/5} + tf(\mathfrak{g})^{-2}) + C\nu^{-2}\mathfrak{g}^2 2^{2\mathfrak{g}}\|\xi\|_{\ell^1}\\
&+ C\nu^{-1}\mathfrak{g}2^{\mathfrak{g}}|\mathcal{Y}| + C\nu^{-1}\mathfrak{g}2^{\mathfrak{g}}\|\xi\|_{\ell^1}^2 + C\|\xi\|_{\ell^1}|\mathcal{Y}| + C|\mathcal{Y}|^2
\end{aligned}$$

for some constant $C > 0$ *independent of* ν.

Proof. We compute estimates for each term in (8.12). For $\mathcal{Z}_0$, we use the fact that the definition of $\mathcal{R}$ in (7.3) and Theorem 6.2 imply $\|\partial_{\overline{\beta}}\mathcal{R}(0, \theta, \beta^\nu)\|_{\ell^1} \leq \mathcal{O}(\|\beta^\nu\|_{\ell^1}^4)$. Thus, it only remains to use the results in Theorems 7.8 (using (8.1)) and 7.3, to obtain

$$\|\partial_{\overline{\beta}}\mathcal{R}(0, \theta, \beta^\nu)\|_{\ell^1} \leq C\nu^{-4}\mathfrak{g}^4 2^{4\mathfrak{g}}.$$

To bound $\partial_{\overline{\beta}}\mathcal{J}(0,\theta,\beta^\nu)$, the other term in $\mathcal{Z}_0$, recall that $\mathcal{J} = \mathcal{J}_1 + \mathcal{J}_2$ (see (7.3) and (7.10)). Then we split into two terms $\partial_{\overline{\beta}}\mathcal{J}(0,\theta,\beta^\nu) = \partial_{\overline{\beta}}\mathcal{J}_1(0,\theta,\beta^\nu) + \partial_{\overline{\beta}}\mathcal{J}_2(\theta,\beta^\nu)$ as

$$\begin{aligned}\partial_{\overline{\beta}}\mathcal{J}_1(0,\theta,\beta^\nu) &= \partial_{\overline{\beta}}\{\mathcal{G}(0,\theta,(\beta^\nu_{\vec{j}}e^{\mathrm{i}\Omega_{\vec{j}}(\lambda,\varepsilon)t})_{\vec{j}\in\mathbb{Z}^2_N\setminus\mathcal{S}_0}) - \mathcal{G}(0,\theta,\beta^\nu)\}\\ &= \partial_{\overline{\beta}}\{\mathcal{Q}^{(2)}_{\mathrm{Res}}(0,\theta,(\beta^\nu_{\vec{j}}e^{\mathrm{i}\Omega_{\vec{j}}(\lambda,\varepsilon)t})_{\vec{j}\in\mathbb{Z}^2_N\setminus\mathcal{S}_0}) - \mathcal{Q}^{(2)}_{\mathrm{Res}}(0,\theta,\beta^\nu)\}, \qquad (8.13)\\ \partial_{\overline{\beta}}\mathcal{J}_2(\theta,\beta^\nu) &= \partial_{\overline{\beta}}\{\mathcal{G}(0,\theta,(\beta^\nu_{\vec{j}}e^{\mathrm{i}\Omega_{\vec{j}}(\lambda,\varepsilon)t})_{\vec{j}\in\mathbb{Z}^2_N\setminus\mathcal{S}_0}) - \mathcal{G}_0((\beta^\nu_{\vec{j}}e^{\mathrm{i}\Omega_{\vec{j}}(\lambda,\varepsilon)t})_{\vec{j}\in\mathbb{Z}^2_N\setminus\mathcal{S}_0})\},\end{aligned} \tag{8.14}$$

To bound (8.13), recall that $\mathcal{Q}^{(2)}_{\mathrm{Res}}$ defined in (6.4) is the sum of two terms. Since $\Pi_{\mathfrak{R}_2}\mathcal{Q}^{(2)}$ is action preserving, the only terms contributing to (8.13) are the ones coming from $\Pi_{\mathfrak{R}_4}\mathcal{Q}^{(2)}$. Since β^ν is supported on Λ, it follows from (6.2) that

$$\partial_{\overline{\beta}}\mathcal{J}_1(0,\theta,\beta^\nu) = \Bigg(\sum_{\substack{\vec{j}_1,\vec{j}_2,\vec{j}_3\in\Lambda\\ |\vec{j}_1|^2-|\vec{j}_2|^2-|\vec{j}_3|^2-|\vec{j}|^2=0}} (e^{\mathrm{i}t(\Omega_{\vec{j}_1}-\Omega_{\vec{j}_2}+\Omega_{\vec{j}_3}-\Omega_{\vec{j}})}-1)\,\mathcal{J}_{\vec{j}_1\vec{j}_2\vec{j}_3\vec{j}}\beta^\nu_{\vec{j}_1}\overline{\beta^\nu_{\vec{j}_2}}\beta^\nu_{\vec{j}_3}\Bigg)_{\vec{j}\in\Lambda}. \tag{8.15}$$

In order to bound the oscillating factor, we use the formula for the eigenvalues given in Theorem 4.4 to find that, for $\vec{j}_1,\vec{j}_2,\vec{j}_3,\vec{j}\in\Lambda$, one has

$$|e^{\mathrm{i}t(\Omega_{\vec{j}_1}-\Omega_{\vec{j}_2}+\Omega_{\vec{j}_3}-\Omega_{\vec{j}})}-1| \lesssim |t|\,|\Omega_{\vec{j}_1}-\Omega_{\vec{j}_2}+\Omega_{\vec{j}_3}-\Omega_{\vec{j}}| \lesssim |t|/f(\mathfrak{g})^2.$$

Hence, for $t\in[0,T]$, using the estimate for $\mathcal{Q}^{(2)}_{\mathrm{Res}}$ given by Theorem 6.2,

$$\|\partial_{\overline{\beta}}\mathcal{J}_1(0,\theta,\beta^\nu)\|_{\ell^1} \le Ctf(\mathfrak{g})^{-2}\|\beta^\nu\|^3_{\ell^1} \le Ct\nu^{-3}\mathfrak{g}^3 2^{3\mathfrak{g}}f(\mathfrak{g})^{-2}.$$

To bound (8.14), it is enough to use (7.12) and (7.6) to obtain

$$\|\partial_{\overline{\beta}}\mathcal{J}_2(\theta,\beta^\nu)\|_{\ell^1} \le Cf(\mathfrak{g})^{-4/5}\|\beta^\nu\|^3_{\ell^1} \le C\nu^{-3}\mathfrak{g}^3 2^{3\mathfrak{g}}f(\mathfrak{g})^{-4/5}.$$

For the linear terms, one can easily see that

$$\|\mathcal{Z}_1(t)\xi\|_{\ell^1} \le C\|\beta^\nu\|^2_{\ell^1}\|\xi\|_{\ell^1} \le C\nu^{-2}\mathfrak{g}^2 2^{2\mathfrak{g}}\|\xi\|_{\ell^1},$$

and the same for $\|\mathcal{Z}_1'(t)\overline{\xi}\|_{\ell^1}$. Analogously,

$$\|\mathcal{Z}_1''(t)\mathcal{Y}\|_{\ell^1} \le C\|\beta^\nu\|_{\ell^1}|\mathcal{Y}| \le C\nu^{-1}\mathfrak{g}2^{\mathfrak{g}}|\mathcal{Y}|.$$

Finally, it is enough to use the definition of $\mathcal{Z}_2$, the definition of $\mathcal{R}$ in (7.3) and Theorem 6.2 to show

$$\begin{aligned}\|\mathcal{Z}_2\| &\le \|\beta^\nu\|_{\ell^1}\|\xi\|^2_{\ell^1} + \|\beta^\nu\|^2_{\ell^1}|\mathcal{Y}| + \|\xi\|_{\ell^1}|\mathcal{Y}| + \|\beta^\nu\|^3_{\ell^1}\|\xi\|_{\ell^1} + |\mathcal{Y}|^2\\ &\le C\nu^{-1}\mathfrak{g}2^{\mathfrak{g}}|\mathcal{Y}|\,\|\xi\|^2_{\ell^1} + C\nu^{-2}\mathfrak{g}^2 2^{2\mathfrak{g}}\|\xi\|_{\ell^1}|\mathcal{Y}| + C\nu^{-3}\mathfrak{g}^3 2^{3\mathfrak{g}}\|\xi\|_{\ell^1} + |\mathcal{Y}|^2. \qquad \blacksquare\end{aligned}$$

Lemma 8.4. *Assume that* $(\beta^\nu, \mathcal{Y}), (\beta^\nu + \xi, \mathcal{Y}) \in D(r_2)$ *(see* (2.1)*) where* r_2 *has been given by Theorem* 6.2. *Then*

$$\frac{d}{dt}|\mathcal{Y}| \le C\nu^{-5}\mathtt{g}^5 2^{5\mathtt{g}} + C\nu^{-3}\mathtt{g}^3 2^{3\mathtt{g}}\|\xi\|_{\ell^1}^2 + C\nu^{-1}\mathtt{g}2^{\mathtt{g}}\|\xi\|_{\ell^1}^3 + C\|\xi\|_{\ell^1}|\mathcal{Y}| + C\nu^{-3}\mathtt{g}^3 2^{3\mathtt{g}}|\mathcal{Y}|^2$$

for some constant $C > 0$ *independent of* ν.

Proof. Proceeding as for $\dot{\xi}$, we write the equation for $\dot{\mathcal{Y}}$ as

$$\dot{\mathcal{Y}} = \mathcal{X}_0(t) + \mathcal{X}_1(t)\xi + \mathcal{X}_1'(t)\overline{\xi} + \mathcal{X}_1''(t)\mathcal{Y} + \mathcal{X}_2(\xi, \mathcal{Y}, t), \tag{8.16}$$

with

$$\begin{aligned}
\mathcal{X}_0(t) &= -\partial_\theta \mathcal{J}(0,\theta,\beta^\nu) - \partial_\theta \mathcal{R}(0,\theta,\beta^\nu),\\
\mathcal{X}_1(t) &= \partial_{\beta\theta}\mathcal{J}(0,\theta,\beta^\nu),\\
\mathcal{X}_1'(t) &= \partial_{\overline{\beta}\theta}\mathcal{J}(0,\theta,\beta^\nu),\\
\mathcal{X}_1''(t) &= \partial_{\mathcal{Y}\theta}\mathcal{J}(0,\theta,\beta^\nu),\\
\mathcal{X}_2(t) &= -\partial_\theta \mathcal{J}(\mathcal{Y},\theta,\beta^\nu+\xi) + \partial_\theta\mathcal{J}(0,\theta,\beta^\nu) - \partial_\theta\mathcal{R}(\mathcal{Y},\theta,\beta^\nu+\xi) + \partial_\theta\mathcal{R}(0,\theta,\beta^\nu).
\end{aligned}$$

We claim that $\mathcal{X}_1(t)$ and $\mathcal{X}_1'(t)$ are identically zero. Then, proceeding as in the proof of Lemma 8.3, one can bound each term and complete the proof of Lemma 8.4.

To explain the absence of linear terms, consider first $\partial_{\beta\theta}\mathcal{J}(0,\theta,\beta^\nu)$. It contains two types of monomials: those coming from $\mathfrak{K}_2$ (see (4.4)), which however do not depend on θ, and those coming from $\mathfrak{K}_4$ (see (6.2)). But also those last monomials do not depend on θ once they are restricted to the set Λ (indeed, the only monomials of $\mathfrak{K}_4$ which are θ-dependent are those of the third line of (6.2), which are supported outside Λ). Therefore $\partial_{\beta\theta}\mathcal{J}(0,\theta,\beta^\nu) \equiv 0$ (and so $\partial_{\overline{\beta}\theta}\mathcal{J}(0,\theta,\beta^\nu)$ and $\partial_{\mathcal{Y}\theta}\mathcal{J}(0,\theta,\beta^\nu)$). ■

We define

$$M = \|\xi\|_{\ell^1} + \nu|\mathcal{Y}|.$$

As a conclusion of these two lemmas, we can deduce that

$$\dot{M} \le C\left(\nu^{-4}\mathtt{g}^4 2^{4\mathtt{g}} + \nu^{-3}\mathtt{g}^3 2^{3\mathtt{g}}(f(\mathtt{g})^{-4/5} + tf(\mathtt{g})^{-2})\right) + C\nu^{-2}\mathtt{g}^2 2^{2\mathtt{g}}M + \nu^{-1}\mathtt{g}2^{\mathtt{g}}M^2.$$

Now we apply a bootstrap argument. Assume that for some $T^* > 0$ and $0 < t < T^*$ we have

$$M(t) \le C\nu^{-1-\sigma/2}$$

(the constant σ will be determined later).

Recall that, by assumption, for $t = 0$ we know that the above is already satisfied since

$$M(0) \le \nu^{-1-4\sigma}. \tag{8.17}$$

A posteriori we will show that the time T in (8.2) satisfies $0 < T < T^*$ and therefore the bootstrap assumption holds. Note that, taking $\mathtt{g}$ large enough (and recalling (7.6) and (8.5)), the bootstrap estimate implies that $(\beta^\nu, \mathcal{Y})$ and $(\beta^\nu + \xi, \mathcal{Y})$ satisfy the assumption of Lemmas 8.3 and 8.4. With the boostrap assumption, we have

$$\dot{M} \leq C\big(\nu^{-4}\mathfrak{g}^4 2^{4\mathfrak{g}} + \nu^{-3}\mathfrak{g}^3 2^{3\mathfrak{g}}(f(\mathfrak{g})^{-4/5} + tf(\mathfrak{g})^{-2})\big) + C\nu^{-2}\mathfrak{g}^2 2^{2\mathfrak{g}} M.$$

Applying the Gronwall inequality one obtains

$$M \leq C\big(M(0) + \nu^{-4}\mathfrak{g}^4 2^{4\mathfrak{g}} t + \nu^{-3}\mathfrak{g}^3 2^{3\mathfrak{g}}(tf(\mathfrak{g})^{-4/5} + t^2 f(\mathfrak{g})^{-2})\big)e^{\nu^{-2}\mathfrak{g}^2 2^{2\mathfrak{g}} t}.$$

Thus, using (8.2), the estimates for T_0 in Theorem 7.8, (8.17) and (8.5) (note that it implies $f(\mathfrak{g})^{-1} \leq \nu^{-1}$), we get

$$\begin{aligned} M &\leq C\big(M(0) + \nu^{-2}\mathfrak{g}^6 2^{4\mathfrak{g}} + \nu^{-1}\mathfrak{g}^5 2^{3\mathfrak{g}} f(\mathfrak{g})^{-4/5} + \nu\mathfrak{g}^7 2^{3\mathfrak{g}} f(\mathfrak{g})^{-2}\big)e^{C\mathfrak{g}^4 2^{2\mathfrak{g}}} \\ &\leq C(\nu^{-1-4\sigma} + \nu^{-2}\mathfrak{g}^6 2^{4\mathfrak{g}} + \nu^{-9/5}\mathfrak{g}^5 2^{3\mathfrak{g}} + \nu^{-3}\mathfrak{g}^7 2^{3\mathfrak{g}})e^{C\mathfrak{g}^4 2^{2\mathfrak{g}}}. \end{aligned}$$

Now, taking A large enough (see Theorem 7.3), there exists $\sigma > 0$ such that for $t \in [0, T]$, provided $\mathfrak{g}$ is sufficiently large,

$$M(t) \leq \nu^{-1-\sigma}.$$

This implies that $T \leq T^*$. That is, the bootstrap assumption was valid. This completes the proof.

Appendix A. Proof of Proposition 6.1

We split the proof into several steps. We first perform an algebraic analysis of the nonresonant monomials.

A.1. Analysis of monomials of the form $e^{\mathrm{i}\theta\cdot\ell}\, a_{\vec{j}_1}^{\sigma_1}\, a_{\vec{j}_2}^{\sigma_2}\, a_{\vec{j}_3}^{\sigma_3}\, a_{\vec{j}_4}^{\sigma_4}$

We analyze the small divisors (6.1) related to these monomials. Taking advantage of the asymptotics of the eigenvalues given in Theorem 4.4, we consider a “good” first order approximation of the small divisor given by

$$\omega(\lambda)\cdot\ell + \sigma_1\widetilde{\Omega}_{\vec{j}_1}(\lambda,\varepsilon) + \sigma_2\widetilde{\Omega}_{\vec{j}_2}(\lambda,\varepsilon) + \sigma_3\widetilde{\Omega}_{\vec{j}_3}(\lambda,\varepsilon) + \sigma_4\widetilde{\Omega}_{\vec{j}_4}(\lambda,\varepsilon). \tag{A.1}$$

Note that this is an affine function in ε and therefore it can be written as

$$(\text{A.1}) \equiv \mathtt{K}^{\sigma}_{\mathbf{j},\ell} + \varepsilon\mathtt{F}^{\sigma}_{\mathbf{j},\ell}(\lambda).$$

We say that a monomial is *Birkhoff nonresonant* if, for any $\varepsilon > 0$, this expression is not 0 as a function of λ.

Lemma A.1. *Assume that the* $\mathtt{m}_k$*'s do not solve any of the linear equations defined in* (A.5) *below* (*this determines* $\mathtt{L}_2$ *in the statement of Theorem* 6.1). *Consider a monomial of the form* $e^{\mathrm{i}\theta\cdot\ell}\, a_{j_1}^{\sigma_1}\, a_{j_2}^{\sigma_2}\, a_{j_3}^{\sigma_3}\, a_{j_4}^{\sigma_4}$ *with* $(\mathbf{j},\ell,\sigma) \in \mathfrak{A}_4$. *If* $(\mathbf{j},\ell,\sigma) \notin \mathfrak{R}_4$, *then it is Birkhoff nonresonant.*

Proof. We write the functions $\mathtt{K}^{\sigma}_{\mathbf{j},\ell}$ and $\mathtt{F}^{\sigma}_{\mathbf{j},\ell}(\lambda)$ explicitly as

$$\mathtt{K}^{\sigma}_{\mathbf{j},\ell} := \omega^{(0)}\cdot\ell+\sigma_1\widetilde{\Omega}_{\vec{\jmath}_1}(\lambda,0)+\sigma_2\widetilde{\Omega}_{\vec{\jmath}_2}(\lambda,0)+\sigma_3\widetilde{\Omega}_{\vec{\jmath}_3}(\lambda,0)+\sigma_4\widetilde{\Omega}_{\vec{\jmath}_3}(\lambda,0), \tag{A.2}$$

$$\begin{aligned}\mathtt{F}^{\sigma}_{\mathbf{j},\ell}(\lambda) &:= \partial_\varepsilon\big(\omega(\lambda)\cdot\ell+\sigma_1\widetilde{\Omega}_{\vec{\jmath}_1}(\lambda,\varepsilon)+\sigma_2\widetilde{\Omega}_{\vec{\jmath}_2}(\lambda,\varepsilon)+\sigma_3\widetilde{\Omega}_{\vec{\jmath}_3}(\lambda,\varepsilon)+\sigma_4\widetilde{\Omega}_{\vec{\jmath}_4}(\lambda,\varepsilon)\big)\big|_{\varepsilon=0}\\ &= -\lambda\cdot\ell+\sigma_1\vartheta_{\vec{\jmath}_1}(\lambda)+\sigma_2\vartheta_{\vec{\jmath}_2}(\lambda)+\sigma_3\vartheta_{\vec{\jmath}_3}(\lambda)+\sigma_4\vartheta_{\vec{\jmath}_4}(\lambda).\end{aligned} \tag{A.3}$$

As in [43], $\mathtt{K}^{\sigma}_{\mathbf{j},\ell}$ is an integer while the functions $\vartheta_{\vec{\jmath}}(\lambda)$ belong to the finite list of functions $\vartheta_{\vec{\jmath}}(\lambda) \in \{0, \{\mu_i(\lambda)\}_{1\le i\le \mathtt{d}}\}$ defined in Theorem 4.4. Clearly to prove that the resonance (A.1) $= 0$ does not hold identically, it is enough to ensure that

$$\mathtt{K}^{\sigma}_{\mathbf{j},\ell} = 0 \quad \text{and} \quad \mathtt{F}^{\sigma}_{\mathbf{j},\ell}(\lambda) \equiv 0 \tag{A.4}$$

cannot occur for $(\mathbf{j},\ell,\sigma) \in \mathfrak{A}_4 \setminus \mathfrak{R}_4$. We study all the possible combinations; each time we assume that (A.4) holds and we deduce a contradiction.

(1): $\vec{\jmath}_i \in \mathfrak{Z}$ for any $1 \le i \le 4$. If $\ell \neq 0$, then $\mathtt{F}^{\sigma}_{\mathbf{j},\ell}(\lambda) = -\lambda\cdot\ell$ is not identically 0. Now take $\ell = 0$. By conservation of $\widetilde{\mathcal{P}}_x$, $\widetilde{\mathcal{P}}_y$ we find that $\sum_{i=1}^4 \sigma_i \vec{\jmath}_i = 0$ and $\mathtt{K}^{\sigma}_{\mathbf{j},\ell} = 0$ implies $\sum_{i=1}^4 \sigma_i |\vec{\jmath}_i|^2 = 0$. Then, using mass conservation (see Remark 4.7), since $\ell = 0$, one has $\sum_{i=1}^4 \sigma_i = 0$ and therefore the $\vec{\jmath}_i$'s form a rectangle (and thus $(\mathbf{j},0,\sigma)$ belongs to $\mathfrak{R}_4$).

(2): $\vec{\jmath}_1, \vec{\jmath}_2, \vec{\jmath}_3 \in \mathfrak{Z}$, $\vec{\jmath}_4 \in \mathcal{S}$. Then $\mathtt{F}^{\sigma}_{\mathbf{j},\ell}(\lambda) = -\lambda\cdot\ell + \sigma_3\,\mu_i(\lambda)$ for some $1 \le i \le \mathtt{d}$. If $\mathtt{F}^{\sigma}_{\mathbf{j},\ell}(\lambda) \equiv 0$ then $\mu_i(\lambda) = \sigma_3\lambda\cdot\ell$ is a root in $\mathbb{Z}[\lambda]$ of the polynomial $P(t,\lambda)$ defined in Theorem 4.4; however, $P(t,\lambda)$ is irreducible over $\mathbb{Q}(\lambda)[t]$, yielding a contradiction.

(3): $\vec{\jmath}_1, \vec{\jmath}_2 \in \mathfrak{Z}$, $\vec{\jmath}_3, \vec{\jmath}_4 \in \mathcal{S}$. We can assume $\vec{\jmath}_3 = (\mathtt{m}_i, n_3)$, $\vec{\jmath}_4 = (\mathtt{m}_k, n_4)$ for some $1 \le i, k \le \mathtt{d}$. Then

$$\mathtt{F}^{\sigma}_{\mathbf{j},\ell}(\lambda) = -\lambda\cdot\ell + \sigma_3\,\mu_i(\lambda) + \sigma_4\mu_k(\lambda).$$

Case $\ell \neq 0$. Then $\mathtt{F}^{\sigma}_{\mathbf{j},\ell}(\lambda) \equiv 0$ iff $\mu_i(\lambda) \equiv -\sigma_3\sigma_4\mu_k(\lambda) + \sigma_3\lambda\cdot\ell$. This means that $\mu_k(\lambda)$ is a common root of $P(t,\lambda)$ and $P(-\sigma_3\sigma_4 t + \sigma_3\lambda\cdot\ell, \lambda)$. However this last polynomial is irreducible as well, being the translation of an irreducible polynomial. Hence the two polynomials must be equal (or opposite). A direct computation shows that this does not happen (see [43, Lemma 6.1] for details).

Case $\ell = 0$. Then $\mathtt{F}^{\sigma}_{\mathbf{j},\ell}(\lambda) \equiv 0$ iff $\mu_i(\lambda) \equiv -\sigma_3\sigma_4\mu_k(\lambda)$.

- If $i \neq k$ and $\sigma_3\sigma_4 = -1$, then $P(t,\lambda)$ would have a root with multiplicity 2. But $P(t,\lambda)$, being an irreducible polynomial, has no multiple roots.
- If $i \neq k$ and $\sigma_3\sigma_4 = 1$, then $P(t,\lambda)$ and $P(-t,\lambda)$ would have $\mu_k(\lambda)$ as a common root. However, $P(-t,\lambda)$ is irreducible on $\mathbb{Z}[\lambda]$ as well, and two irreducible polynomials sharing a common root must coincide (up to sign), i.e. $P(t,\lambda) \equiv \pm P(-t,\lambda)$. A direct computation using the explicit expression of $P(t,\lambda)$ shows that this is not true.
- If $i = k$ and $\sigma_3\sigma_4 = 1$ then $\mu_i(\lambda) \equiv 0$ would be a root of $P(t,\lambda)$. But $P(t,\lambda)$ is irreducible over $\mathbb{Z}[\lambda]$, so it cannot have 0 as a root.

- If $i = k$ and $\sigma_3\sigma_4 = -1$ (we can assume $\sigma_3 = 1$, $\sigma_4 = -1$), by mass conservation one has $\sigma_1 + \sigma_2 = 0$ and by conservation of $\widetilde{\mathcal{P}}_x$ one has $\sigma_1 m_1 + \sigma_2 m_2 = 0$, thus $m_1 = m_2$. Then by conservation of $\widetilde{\mathcal{P}}_y$ we get $n_1 - n_2 + n_3 - n_4 = 0$, which together with $0 = \mathtt{K}^{\sigma}_{\mathbf{j},\ell} = n_1^2 - n_2^2 + n_3^2 - n_4^2$ gives $\{n_1, n_3\} = \{n_2, n_4\}$. One verifies easily that in that case the sites $\vec{j}_r$'s form a horizontal rectangle (which could even be degenerate), and therefore they belong to $\mathfrak{R}_4$.

(4): $\vec{j}_1, \vec{j}_2, \vec{j}_3 \in \mathscr{S}$, $\vec{j}_4 \in \mathfrak{Z}$. We can assume that $\vec{j}_1 = (\mathtt{m}_{i_1}, n_1)$, $\vec{j}_2 = (\mathtt{m}_{i_2}, n_2)$, $\vec{j}_3 = (\mathtt{m}_{i_3}, n_3)$ for some $1 \le i_1, i_2, i_3 \le \mathtt{d}$ and $n_1, n_2, n_3 \neq 0$. Then

$$\mathtt{F}^{\sigma}_{\mathbf{j},\ell}(\lambda) = -\lambda \cdot \ell + \sigma_1 \mu_{i_1}(\lambda) + \sigma_2 \mu_{i_2}(\lambda) + \sigma_3 \mu_{i_3}(\lambda) \, .$$

By conservation of mass $\eta(\ell) + \sigma_4 = 0$, hence $\ell \neq 0$. Assume $\mathtt{F}^{\sigma}_{\mathbf{j},\ell}(\lambda) \equiv 0$. This can only happen for (at most) a unique choice of $\ell^{(\mathbf{i},\sigma)} \in \mathbb{Z}^{\mathtt{d}}$, $\mathbf{i} := (i_1, i_2, i_3)$. By conservation of $\widetilde{\mathcal{P}}_x$ we have

$$\sum_k \mathtt{m}_k \ell_k^{(\mathbf{i},\sigma)} + \sigma_4 m_4 = 0.$$

These two conditions fix $m_4 \equiv m_4^{(\mathbf{i},\sigma)}$ uniquely. In particular, if m_4 is sufficiently large, we have a contradiction.

(5): $\vec{j}_r \in \mathscr{S}$ for all $1 \le r \le 4$. Then

$$\mathtt{F}^{\sigma}_{\mathbf{j},\ell}(\lambda) = -\lambda \cdot \ell + \sigma_1 \mu_{i_1}(\lambda) + \sigma_2 \mu_{i_2}(\lambda) + \sigma_3 \mu_{i_3}(\lambda) + \sigma_4 \mu_{i_4}(\lambda).$$

If $\ell \neq 0$, the condition $\mathtt{F}^{\sigma}_{\mathbf{j},\ell}(\lambda) \equiv 0$ fixes $\ell^{(\mathbf{i},\sigma)} \in \mathbb{Z}^{\mathtt{d}}$ uniquely, $\mathbf{i} := (i_1, i_2, i_3, i_4)$. By conservation of $\widetilde{\mathcal{P}}_x$ we have the condition

$$\sum_k \mathtt{m}_k \ell_k^{(\mathbf{i},\sigma)} = 0 \tag{A.5}$$

defining a hyperplane, which can be excluded by suitably choosing the tangential sites $\mathtt{m}_k$ (recall that the functions $\mu_i(\lambda)$ are independent of this choice, see Remark 4.5).

If $\ell = 0$, we have $\sum_r \sigma_r n_r = \sum_r \sigma_r n_r^2 = 0$. Then $\{n_1, n_3\} = \{n_2, n_4\}$. One verifies easily that in that case the sites $\vec{j}_r$ form a horizontal trapezoid (which could even be degenerate). ∎

A.2. Analysis of monomials of the form $e^{\mathrm{i}\theta\cdot\ell} \mathcal{Y}^l a_{\vec{j}_1}^{\sigma_1} a_{\vec{j}_2}^{\sigma_2}$

In this case, since the factor $\mathcal{Y}^l$ does not affect the Poisson brackets, admissible monomials (in the sense of Definition 4.1) are nonresonant provided they do not belong to the set $\mathfrak{R}_2$ introduced in Definition 4.2.

Lemma A.2. *Any monomial of the form $e^{\mathrm{i}\theta\cdot\ell} a_{\vec{j}_1}^{\sigma_1} a_{\vec{j}_2}^{\sigma_2} \mathcal{Y}_i$ with $(\mathbf{j}, \ell, \sigma) \notin \mathfrak{R}_2$ admissible in the sense of Definition 4.1 is Birkhoff nonresonant.*

Proof. We skip the proof since it is analogous to that of [43, Lemma 6.1]. ∎

A.3. Quantitative measure estimate

We are now in a position to prove our quantitative nonresonance estimate. Recall that, by Theorem 4.3, the frequencies $\Omega_{\vec{j}}(\lambda,\varepsilon)$ of the Hamiltonian (5.1) have the form (4.8). Expanding $\Omega_{\vec{j}}(\lambda,\varepsilon)$ in Taylor series in powers of ε we get

$$\begin{aligned}\omega(\lambda)\cdot\ell+\sigma_1\Omega_{\vec{j}_1}(\lambda,\varepsilon)+\sigma_2\Omega_{\vec{j}_2}(\lambda,\varepsilon)+\sigma_3\Omega_{\vec{j}_3}(\lambda,\varepsilon)+\sigma_4\Omega_{\vec{j}_4}(\lambda,\varepsilon)\\ =\mathtt{K}^{\boldsymbol{\sigma}}_{\mathbf{j},\ell}+\varepsilon\mathtt{F}^{\boldsymbol{\sigma}}_{\mathbf{j},\ell}(\lambda)+\varepsilon^2\mathtt{G}^{\boldsymbol{\sigma}}_{\mathbf{j},\ell}(\lambda,\varepsilon),\end{aligned}\tag{A.6}$$

where $\mathtt{K}^{\boldsymbol{\sigma}}_{\mathbf{j},\ell}$ is defined in (A.2) and $\mathtt{F}^{\boldsymbol{\sigma}}_{\mathbf{j},\ell}(\lambda)$ is defined in (A.3). We wish to prove that the set of $\lambda\in\mathcal{C}^{(2)}_{\varepsilon}$ such that

$$\begin{aligned}|\omega(\lambda)\cdot\ell+\sigma_1\Omega_{\vec{j}_1}(\lambda,\varepsilon)+\sigma_2\Omega_{\vec{j}_2}(\lambda,\varepsilon)+\sigma_3\Omega_{\vec{j}_3}(\lambda,\varepsilon)+\sigma_4\Omega_{\vec{j}_4}(\lambda,\varepsilon)|\geq\varepsilon\gamma_2/\langle\ell\rangle^{\tau_2},\\ \forall(\mathbf{j},\ell,\boldsymbol{\sigma})\in\mathfrak{A}_4\setminus\mathfrak{R}_4,\end{aligned}\tag{A.7}$$

has positive measure for γ_2 and ε small enough and τ_2 large enough. We deal with the cases $|\ell|\leq 4\mathtt{M}_0$ and $|\ell|>4\mathtt{M}_0$ separately.

A.3.1. Case $|\ell|\leq 4\mathtt{M}_0$. We start with the following lemma.

Lemma A.3. *There exists $\mathtt{k}\in\mathbb{N}$ such that for any $\gamma_c>0$ sufficiently small, there exists a compact domain $\mathcal{C}_c\subset\mathcal{O}_0$ with $|\mathcal{O}_0\setminus\mathcal{C}_c|\sim\gamma_c^{1/\mathtt{k}}$ and*

$$\min\{|\mathtt{F}^{\boldsymbol{\sigma}}_{\mathbf{j},\ell}(\lambda)|:\lambda\in\mathcal{C}_c,\ (\ell,\mathbf{j},\boldsymbol{\sigma})\in\mathfrak{A}_4\setminus\mathfrak{R}_4,\ |\ell|\leq 4\mathtt{M}_0,\ \mathtt{K}^{\boldsymbol{\sigma}}_{\mathbf{j},\ell}=0\}\geq\gamma_c>0.$$

Proof. See [43, Lemma 6.4]. The estimate on the measure follows from classical results on sublevels of analytic functions. ■

We can now prove the following result.

Proposition A.4. *There exist $\varepsilon_c>0$ and a set $\mathcal{C}_c\subset\mathcal{O}_0$ such that for any $\varepsilon\leq\varepsilon_c$ and any $\lambda\in\mathcal{C}_c$ one has*

$$\Big|\omega(\lambda)\cdot\ell+\sum_{l=1}^{4}\sigma_l\Omega_{\vec{j}_l}(\lambda,\varepsilon)\Big|\geq\frac{\gamma_c\varepsilon}{2},\quad\forall(\mathbf{j},\ell,\boldsymbol{\sigma})\in\mathfrak{A}_4\setminus\mathfrak{R}_4,\ |\ell|\leq 4\mathtt{M}_0.\tag{A.8}$$

Moreover, $|\mathcal{O}_0\setminus\mathcal{C}_c|\leq\alpha\varepsilon_c^{\kappa}$ where α,κ do not depend on ε_c.

Proof. By the very definition of $\mathtt{M}_0$ in (4.9) and the estimates on the eigenvalues given in Theorem 4.4, one has $\sup_{\lambda\in\mathcal{O}_0}|\mathtt{F}^{\boldsymbol{\sigma}}_{\mathbf{j},\ell}(\lambda)|\leq 8\mathtt{M}_0$ and $\sup_{\lambda\in\mathcal{O}_0}|\mathtt{G}^{\boldsymbol{\sigma}}_{\mathbf{j},\ell}(\lambda)|\leq 4\mathtt{M}_0$. Assume first that $\mathtt{K}^{\boldsymbol{\sigma}}_{\mathbf{j},\ell}\in\mathbb{Z}\setminus\{0\}$. Then if ε_c is sufficiently small and for $\varepsilon<\varepsilon_c$ one has

$$|(\mathrm{A.6})|\geq|\mathtt{K}^{\boldsymbol{\sigma}}_{\mathbf{j},\ell}|-\varepsilon 8\mathtt{M}_0-\varepsilon^2 4\mathtt{M}_0\geq 1/2.$$

Hence, for such ℓ's, (A.8) is trivially fulfilled for all $\lambda\in\mathcal{O}_0$. If instead $\mathtt{K}^{\boldsymbol{\sigma}}_{\mathbf{j},\ell}=0$, we use Lemma A.3 with $\gamma_c=10\mathtt{M}_0\varepsilon_c$ to obtain a set $\mathcal{C}_c\subset\mathcal{O}_0$ such that for any $\lambda\in\mathcal{C}_c$ and any $(\mathbf{j},\ell,\boldsymbol{\sigma})\in\mathfrak{A}_4\setminus\mathfrak{R}_4$ with $|\ell|\leq 4\mathtt{M}_0$,

$$|(\mathrm{A.6})|\geq\varepsilon\gamma_c-\varepsilon^2 4\mathtt{M}_0\geq\varepsilon\gamma_c/2.$$

■

A.3.2. Case $|\ell| > 4\mathtt{M}_0$. In this case we prove the following result.

Proposition A.5. *Fix $\varepsilon_\star > 0$ sufficiently small and $\tau_\star > 0$ sufficiently large. For any $\varepsilon < \varepsilon_\star$, there exists a set $\mathcal{C}_\star \subset \mathcal{O}_0$ such that $|\mathcal{O}_0 \setminus \mathcal{C}_\star| \lesssim \varepsilon_\star^\kappa$ (with α, κ independent of $\varepsilon_\star$), and for any $\lambda \in \mathcal{C}_\star$ and $|\ell| > 4\mathtt{M}_0$ one has*

$$\left|\omega(\lambda)\cdot\ell + \sum_{l=1}^{4} \sigma_l \Omega_{\vec{\jmath}_l}(\lambda,\varepsilon)\right| \geq \gamma_\star \frac{\varepsilon}{\langle \ell\rangle^{\tau_\star}} \tag{A.9}$$

for some constant $\gamma_\star$ depending on $\varepsilon_\star$.

To prove the proposition, first define, for $1 \leq i \leq \mathtt{d}$ and $0 \leq k \leq \mathtt{d}$, the functions

$$\widehat{\mathtt{F}}_{i,k}(\lambda) = \begin{cases} \varepsilon\mu_i(\lambda) & \text{if } k = 0, \\ \varepsilon\mu_{i,k}^+(\lambda) & \text{if } 1 \leq i < k \leq \mathtt{d}, \\ \varepsilon\mu_{i,k}^-(\lambda) & \text{if } 1 \leq k < i \leq \mathtt{d}, \\ 0 & \text{if } 1 \leq i = k \leq \mathtt{d}. \end{cases}$$

The right hand side of (A.6) is always of the form

$$\begin{aligned} &\omega(\lambda)\cdot\ell + K + \eta_1\widehat{\mathtt{F}}_{i_1,k_1}(\lambda) + \eta_2\widehat{\mathtt{F}}_{i_2,k_2}(\lambda) + \eta_3\widehat{\mathtt{F}}_{i_3,k_3}(\lambda) + \eta_4\widehat{\mathtt{F}}_{i_4,k_4}(\lambda) \\ &\quad + \eta_{11}\frac{\Theta_{m_1}(\lambda,\varepsilon)}{\langle m_1\rangle^2} + \eta_{12}\frac{\Theta_{m_2}(\lambda,\varepsilon)}{\langle m_2\rangle^2} + \eta_{13}\frac{\Theta_{m_3}(\lambda,\varepsilon)}{\langle m_3\rangle^2} + \eta_{14}\frac{\Theta_{m_4}(\lambda,\varepsilon)}{\langle m_4\rangle^2} \\ &\quad + \eta_{21}\frac{\Theta_{m_1,n_1}(\lambda,\varepsilon)}{\langle m_1\rangle^2+\langle n_1\rangle^2} + \eta_{22}\frac{\Theta_{m_2,n_2}(\lambda,\varepsilon)}{\langle m_2\rangle^2+\langle n_2\rangle^2} + \eta_{23}\frac{\Theta_{m_3,n_3}(\lambda,\varepsilon)}{\langle m_3\rangle^2+\langle n_3\rangle^2} + \eta_{24}\frac{\Theta_{m_4,n_4}(\lambda,\varepsilon)}{\langle m_4\rangle^2+\langle n_4\rangle^2} \\ &\quad + \eta_{31}\frac{\varpi_{m_1}(\lambda,\varepsilon)}{\langle m_1\rangle} + \eta_{32}\frac{\varpi_{m_2}(\lambda,\varepsilon)}{\langle m_2\rangle} + \eta_{33}\frac{\varpi_{m_3}(\lambda,\varepsilon)}{\langle m_3\rangle} + \eta_{34}\frac{\varpi_{m_4}(\lambda,\varepsilon)}{\langle m_4\rangle} \end{aligned} \tag{A.10}$$

for a particular choice of $K \in \mathbb{Z}, m_i \in \mathbb{Z}, n_i \in N\mathbb{Z} \setminus \{0\}$ and $\eta_r, \eta_{jj'} \in \{-1,0,1\}$. Therefore it is enough to show (A.9) where the left hand side is replaced by (A.10).

Proof of Proposition A.5. If the integer K is large, namely $|K| \geq 4|\ell| \max_{1\leq i\leq \mathtt{d}} \mathtt{m}_i^2$, then the quantity on the left hand side of (A.9) is far from zero. More precisely,

$$\begin{aligned} |(\text{A.10})| &\geq |K| - |\omega(\lambda)|\,|\ell| - \sum_{r=1}^{4} |\widehat{\mathtt{F}}_{i_r,k_r}|^{\mathcal{O}_1} - \sum_{r=1}^{4}\frac{|\Theta_{m_r}(\cdot,\varepsilon)|^{\mathcal{O}_1}}{\langle m_r\rangle^2} \\ &\quad - \sum_{r=1}^{4}\frac{|\Theta_{m_r,n_r}(\cdot,\varepsilon)|^{\mathcal{O}_1}}{\langle m_r\rangle^2+\langle n_r\rangle^2} - \sum_{r=1}^{4}\frac{|\varpi_{m_r}(\cdot,\varepsilon)|^{\mathcal{O}_1}}{\langle m_r\rangle} \\ &\geq 4\max_{1\leq i\leq \mathtt{d}} \mathtt{m}_i^2|\ell| - \max_{1\leq i\leq \mathtt{d}} \mathtt{m}_i^2|\ell| - \varepsilon|\ell| - 4\varepsilon\mathtt{M}_0 - 4\varepsilon^2\mathtt{M}_0 \geq \mathtt{M}_0. \end{aligned}$$

So from now on we restrict ourselves to the case $|K| \leq 4|\ell| \max_{1\leq i\leq \mathtt{d}} \mathtt{m}_i^2$. We will repeatedly use the following result, which is an easy variant of [47, Lemma 5].

Lemma A.6. *Fix $K \in \mathbb{Z}$, $m_i \in \mathbb{Z}, n_i \in \mathbb{Z} \setminus \{0\}$, $\eta_j, \eta_{jj'} \in \{-1, 0, 1\}$. For any $\alpha > 0$ one has*

$$\text{meas}(\{\lambda \in \mathcal{O}_0 : |(\text{A.10})| < \varepsilon\alpha\}) < 16\alpha|\ell|^{-1}.$$

The proof relies on the fact that all the functions appearing in (A.10) are Lipschitz in λ; for full details see e.g. [43, Lemma C.2].

Now, let us fix

$$\gamma_\star = \frac{\varepsilon_\star \mathtt{M}_0}{100}. \tag{A.11}$$

We construct the set $\mathcal{C}_\star$ by induction on the number n defined by

$$n := |\eta_{1,1}| + \cdots + |\eta_{3,4}| \leq 12,$$

which is nothing but the number of nonzero coefficients in (A.10) . For every $0 \leq n \leq 12$ we construct (i) a positive increasing sequence τ_n and (ii) a sequence of nested sets $\mathcal{C}^n = \mathcal{C}^n(\gamma_\star, \tau_n)$ such that:

(1) There exists $C > 0$, independent of ε and $\gamma_\star$, such that

$$\text{meas}(\mathcal{O}_0 \setminus \mathcal{C}^0) \leq C\gamma_\star, \quad \text{meas}(\mathcal{C}^n \setminus \mathcal{C}^{n+1}) \leq C\gamma_\star. \tag{A.12}$$

(2) For $\lambda \in \mathcal{C}^n$ and $|\ell| \geq 4\mathtt{M}_0$ one has

$$|(\text{A.10})| \geq \varepsilon\gamma_\star/\langle\ell\rangle^{\tau_n}. \tag{A.13}$$

Then the proposition follows by taking $\mathcal{C}_\star := \mathcal{C}^{12}$, $\tau_\star = \tau_{12}$, so that $|\mathcal{O}_0 \setminus \mathcal{C}_\star| \leq 13C\gamma_\star \sim \gamma_\star$ provided $\gamma_\star$ is small enough.

Case $n = 0$. Define

$$G^0_{K,\mathbf{i},\mathbf{k},\eta,\ell}(\gamma_\star, \tau_0) := \{\lambda \in \mathcal{O}_0 : |(\text{A.10})| \leq \varepsilon\gamma_\star/\langle\ell\rangle^{\tau_0} \text{ and } \eta_{jj'} = 0 \; \forall j, j'\},$$

where $K \in \mathbb{Z}$ with $|K| \leq 4\max_{1\leq i\leq \mathtt{d}} \mathtt{m}_i^2|\ell|$, $\mathbf{i} = (i_1, i_2, i_3, i_4) \in \{1, \ldots, \mathtt{d}\}^4$, $\mathbf{k} = (k_1, k_2, k_3, k_4) \in \{0, \ldots, \mathtt{d}\}^4$, $\ell \in \mathbb{Z}^{\mathtt{d}}$ with $|\ell| \geq 4\mathtt{M}_0$, $\eta = (\eta_1, \eta_2, \eta_3, \eta_4) \in \{-1, 0, 1\}^4$. By Lemma A.6 with $\alpha = \gamma_\star\langle\ell\rangle^{-\tau_0}$ we have

$$\text{meas}(G^0_{K,\mathbf{i},\mathbf{k},\eta,\ell}(\gamma_\star, \tau_0)) \leq 16\gamma_\star/\,\langle\ell\rangle^{\tau_0+1}.$$

Taking the union over all the possible values of $K, \mathbf{i}, \mathbf{k}, \eta, \ell$ one gets

$$\text{meas}\Big(\bigcup_{\substack{|\ell|\geq 4\mathtt{M}_0,\, \mathbf{i},\mathbf{k},\eta \\ |K|\leq 4\max_i \mathtt{m}_i^2|\ell|}} G^0_{K,\mathbf{i},\mathbf{k},\eta,\ell}(\gamma_\star, \tau_0)\Big) \leq C(\mathtt{d})\,\gamma_\star \sum_{|\ell|\geq 4\mathtt{M}_0} \frac{1}{\langle\ell\rangle^{\tau_0}} \leq C\gamma_\star,$$

which is finite provided $\tau_0 \geq \mathtt{d} + 1$. Letting

$$\mathcal{C}^0 := \mathcal{O}_0 \setminus \bigcup_{\substack{|\ell|\geq 4\mathtt{M}_0,\, \mathbf{i},\mathbf{k},\eta \\ |K|\leq 4\max_i \mathtt{m}_i^2|\ell|}} G^0_{K,\mathbf{i},\mathbf{k},\eta,\ell}(\gamma_\star, \tau_0)$$

one clearly has meas$(\mathcal{O}_0 \setminus \mathcal{C}^0) \le C\gamma_\star$, and for $\lambda \in \mathcal{C}^0$,

$$\Big|\omega(\lambda)\cdot\ell + K + \sum_{r=1}^{4} \eta_j \widehat{\mathtt{F}}_{i_r,k_r}(\lambda)\Big| \ge \varepsilon\gamma_\star/\langle\ell\rangle^{\tau_0}$$

for any admissible choice of $\ell, K, \mathbf{i}, \mathbf{k}, \eta$. This proves the inductive step for $n = 0$.

Case $n \rightsquigarrow n+1$**.** Assume that (A.13) holds for any possible choice of $\eta_{11}, \dots, \eta_{34}$ such that $|\eta_{11}| + \dots + |\eta_{34}| \le n \le 11$ for some $(\tau_j)_{j=1}^n$. We now prove step $n+1$. Let us fix $\tau_{n+1} \ge \mathtt{d} + 1 + 6\tau_n$. We shall show that for each $|\ell| \ge 4\mathtt{M}_0$, the set

$$G_\ell^{n+1} := \{\lambda \in \mathcal{C}^n \colon |(\text{A.10})| \le \varepsilon\gamma_\star/\langle\ell\rangle^{\tau_{n+1}},\ |\eta_{11}| + \dots + |\eta_{34}| = n+1\} \tag{A.14}$$

has measure $\le C(\mathtt{d})\gamma_\star/\langle\ell\rangle^{\mathtt{d}+1}$. Thus defining

$$\mathcal{C}^{n+1} := \mathcal{C}^n \setminus \bigcup_{|\ell|\ge 4\mathtt{M}_0} G_\ell^{n+1}(\gamma_\star, \tau_{n+1}),$$

we obtain the claimed estimates (A.12) and (A.13). To estimate the measure of (A.14) we consider three cases.

Case 1: Assume that

$$\exists m_i, \quad |m_i| \ge \langle\ell\rangle^{\tau_n}$$

(of course we also assume that one of the coefficients $\eta_{1i}, \eta_{2i}, \eta_{3i}$ is not null); we can assume it is m_4. Then we treat all the terms in (A.10) which contain m_4 as perturbations, and we estimate all the other terms using the inductive assumption. Here are the details: First we have

$$\left|\frac{\Theta_{m_4}(\lambda,\varepsilon)}{\langle m_4\rangle^2}\right| + \left|\frac{\Theta_{m_4,n_4}(\lambda,\varepsilon)}{\langle m_4\rangle^2 + \langle n_4\rangle^2}\right| + \left|\frac{\varpi_{m_4}(\lambda,\varepsilon)}{\langle m_4\rangle}\right| \le \frac{\mathtt{M}_0\,\varepsilon^2}{\langle\ell\rangle^{\tau_n}}.$$

By the inductive assumption (A.13) and (A.11), for any $\lambda \in \mathcal{C}^n$ one has

$$\begin{aligned}
|(\text{A.10})| \ge{}& \Big|\omega(\lambda)\cdot\ell + K + \sum_{j=1}^{4}\eta_i \widehat{\mathtt{F}}_{i_j,k_j}(\lambda) + \sum_{j=1}^{3}\eta_{1j}\frac{\Theta_{m_r}(\lambda,\varepsilon)}{\langle m_j\rangle^2} \\
&+ \sum_{j=1}^{3}\eta_{2j}\frac{\Theta_{m_j,n_j}(\lambda,\varepsilon)}{\langle m_j\rangle^2 + \langle n_j\rangle^2} + \sum_{j=1}^{3}\eta_{3j}\frac{\varpi_{m_j}(\lambda,\varepsilon)}{\langle m_j\rangle}\Big| - \frac{\mathtt{M}_0\varepsilon^2}{\langle\ell\rangle^{\tau_n}} \\
\ge{}& \frac{\varepsilon\gamma_\star}{\langle\ell\rangle^{\tau_n}} - \frac{\mathtt{M}_0\varepsilon^2}{\langle\ell\rangle^{\tau_n}} \ge \frac{\varepsilon\gamma_\star}{2\langle\ell\rangle^{\tau_n}} \ge \frac{\varepsilon\gamma_\star}{\langle\ell\rangle^{\tau_{n+1}}}
\end{aligned}$$

provided $\tau_{n+1} \ge \tau_n + 1$. Therefore, in this case, there are no λ's contributing to the set (A.14).

Case 2: Assume that

$$\exists n_i, \quad |n_i|^2 \ge \langle\ell\rangle^{\tau_n}$$

(and again we also assume that one of the coefficients η_{2i} is not null); we can assume it is n_4. Similarly to the previous case, we treat the term in (A.10) which contains n_4 as a perturbation, and we estimate all the other terms using the inductive assumption. More precisely, we have

$$\left|\frac{\Theta_{m_4,n_4}(\lambda,\varepsilon)}{\langle m_4\rangle^2+\langle n_4\rangle^2}\right| \le \frac{\mathtt{M}_0\varepsilon^2}{\langle \ell\rangle^{\tau_n}},$$

so by the inductive assumption (A.13) and (A.11),

$$\begin{aligned}
|(\mathrm{A}.10)| \ge & \left|\omega(\lambda)\cdot\ell + K + \sum_{j=1}^{4}\eta_i \widehat{\mathtt{F}}_{i_j,k_j}(\lambda) + \sum_{j=1}^{4}\eta_{1j}\frac{\Theta_{m_j}(\lambda,\varepsilon)}{\langle m_j\rangle^2}\right. \\
& \left. + \sum_{j=1}^{3}\eta_{2j}\frac{\Theta_{m_j,n_j}(\lambda,\varepsilon)}{\langle m_j\rangle^2+\langle n_j j\rangle^2} + \sum_{j=1}^{4}\eta_{3j}\frac{\varpi_{m_j}(\lambda,\varepsilon)}{\langle m_j\rangle}\right| - \frac{\mathtt{M}_0\,\varepsilon^2}{\langle\ell\rangle^{\tau_n}} \\
\ge & \frac{\varepsilon\gamma_\star}{2\langle\ell\rangle^{\tau_n}} \ge \frac{\varepsilon\gamma_\star}{\langle\ell\rangle^{\tau_{n+1}}}
\end{aligned}$$

provided $\tau_{n+1}\ge\tau_n+1$. Also in this case, there are no λ's contributing to the set (A.14).

Case 3: We have

$$|m_i|, |n_i|^2 \le \langle\ell\rangle^{\tau_n}$$

for all the m_i, n_i that appear in (A.10) with nonzero coefficients. Furthermore, recall that we are considering the case $|K|\le 4\max_i \mathtt{m}_i^2|\ell|$. Thus we are left with a finite number of cases and we can impose a finite number of Melnikov conditions. So define

$$G^{n+1}_{K,\mathbf{i},\mathbf{k},\eta,\ell,\mathbf{m},\mathbf{n}}(\gamma_\star,\tau_{n+1}) := \{\lambda\in\mathcal{C}^n : |(\mathrm{A}.10)|\le\varepsilon\gamma_\star/\langle\ell\rangle^{\tau_{n+1}},\ |\eta_{11}|+\cdots+|\eta_{34}|=n+1\}.$$

By Lemma A.6 with $\alpha=\gamma/\langle\ell\rangle^{\tau_{n+1}}$ we have

$$\mathrm{meas}(G^{n+1}_{K,\mathbf{i},\mathbf{k},\eta,\ell,\mathbf{m},\mathbf{n}}(\gamma_\star,\tau_{n+1})) \le 16\gamma_\star/\langle\ell\rangle^{\tau_{n+1}+1}, \tag{A.15}$$

and taking the union over the possible values of $K, \mathbf{i}, \mathbf{k}, \eta, \mathbf{m}, \mathbf{n}$ one gets

$$G^{n+1}_\ell \equiv \bigcup_{\mathbf{i},\mathbf{k},\eta}\ \bigcup_{\substack{|m_i|,|n_i|^2\le\langle\ell\rangle^{\tau_n}\\ |K|\le 4\max_i \mathtt{m}_i^2|\ell|}} G_{K,\mathbf{i},\mathbf{k},\eta,\ell,\mathbf{m},\mathbf{n}}(\gamma_\star,\tau_{n+1}).$$

Estimate (A.15) gives immediately

$$\mathrm{meas}(G^{n+1}_\ell) \le C(\mathtt{d})\gamma_\star\frac{\langle\ell\rangle^{1+6\tau_n}}{\langle\ell\rangle^{\tau_{n+1}+1}} \le \frac{C(\mathtt{d})\gamma_\star}{\langle\ell\rangle^{\mathtt{d}+1}},$$

as claimed. ∎

We can finally prove Proposition 6.1.

Proof of Proposition 6.1. Fix $\gamma_c=\gamma_\star=:\gamma_2$ sufficiently small, and put $\varepsilon_2:=\min(\varepsilon_c,\varepsilon_\star)$, $\tau_2:=\tau_\star$ and $\mathcal{C}^{(2)}:=\mathcal{C}_c\cap\mathcal{C}_\star$. Propositions A.4 and A.5 guarantee that for any $\lambda\in\mathcal{C}^{(2)}$, estimate (A.7) is fulfilled. Finally, one has $|\mathcal{C}^{(1)}\setminus\mathcal{C}^{(2)}|\lesssim\gamma_2^{1/\mathtt{k}}+\gamma_2\sim\gamma_2^{1/\mathtt{k}}$. ∎

Appendix B. List of notations

We give a list of notations and parameters. We also specify the relations between the parameters needed to prove the first statement of Theorem 1.2.

- $\mathtt{d} \in \mathbb{N}$ – Dimension of the torus $\mathcal{T}^I_{\mathcal{S}_0}$
- $\mathcal{S}_0 \subset \mathbb{Z} \times \{0\}$ – Set of (Birkhoff) modes where the torus $\mathcal{T}^I_{\mathcal{S}_0}$ is supported. It has cardinality $\mathtt{d}$.
- $I = (I_{\mathtt{m}_1}, \dots, I_{\mathtt{m}_\mathtt{d}}) \in \mathbb{R}^{\mathtt{d}}_{>0}$ – Actions which define the torus $\mathcal{T}^I_{\mathcal{S}_0}$.
- $\varepsilon \in \mathbb{R}$ – Size of the actions I.
- $s \in (0,1)$ – The index for the Sobolev H^s norm in the first statement of Theorem 1.2.
- $\delta \ll 1$ – It measures the initial distance from the torus $\mathcal{T}^I_{\mathcal{S}_0}$ (in the Sobolev H^s norm) in the first statement of Theorem 1.2.
- $K \gg 1$ – It measures the final Sobolev norm in the first statement of Theorem 1.2.
- $r \in \mathbb{R}$ – Size of the neighborhood of 0 where the several steps of Birkhoff normal form are performed (see (2.1)).
- $\rho \in \mathbb{R}$ – Width of the analyticity domain in the angles θ (see (2.1)).
- $\lambda \in (1/2, 1)^{\mathtt{d}}$ – Parameter used to modulate the actions I (see Lemma 3.3).
- $\omega(\lambda) \in \mathbb{R}^{\mathtt{d}}$ – Tangential frequencies of the torus (see Lemma 3.3).
- $N \in \mathbb{N}$ – It is introduced in (4.1) and defines the set lattice $\mathbb{Z} \times N\mathbb{Z}$ where equation (2D-NLS) is restricted. It will be chosen depending on $\mathfrak{g}$ (see below).
- $\Omega_{\vec{J}}(\lambda, \varepsilon)$ – Normal frequencies of the torus $\mathcal{T}^I_{\mathcal{S}_0}$ (see Theorem 4.3).
- $\Lambda \subset \mathbb{Z} \times N\mathbb{Z}$ – Set where the solution undergoing growth of Sobolev norms is essentially supported (see Theorem 7.3).
- $\mathfrak{g} \in \mathbb{N}$ – Number of generations of the set Λ. In Section 8 it is $\mathfrak{g} \sim \ln(K/\delta)$.
- $f(\mathfrak{g})$ – It gives the size of the modes in Λ (see (7.6) and (7.7)).
- $\mu \in \mathbb{R}$ – It measures the errors in the toy model orbit (see Theorem 7.8). It satisfies $\mu = e^{-\gamma \mathfrak{g}}$ for some $\gamma \gg 1$.
- $T_0 \in \mathbb{R}$ – Transition time for the toy model orbit (see Theorem 7.8). It satisfies $T_0 \sim \mathfrak{g}^2$.
- $\nu \in \mathbb{R}$ – Scaling applied to the toy model solution. In the case $s \in (0,1)$ (first part of the statement of Theorem 1.2) it satisfies $f(\mathfrak{g})^{s_1} \le \nu \le f(\mathfrak{g})^{s_2}$ for some $0 < s_1 < s < s_2 < 1$.
- $T \in \mathbb{R}$ – Transition time for the orbit in the first statement of Theorem 1.2 (see also (8.2)). It satisfies $T \sim \nu^2 \mathfrak{g}^2 \lesssim e^{(K/\delta)^\beta}$ for some $\beta > 1$.

Acknowledgements. The authors would like to thank the referees for their careful reading of the manuscript and their excellent suggestions that helped improve the exposition significantly.

Funding. This project has received funding from the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation programme (grant agreement No 757802) and under FP7-IDEAS (grant agreement No 306414). M.G. has also been partly supported by the Spanish MINECO-FEDER Grant PGC2018-098676-B-100 (AEI/FEDER/UE) and by the Catalan Institution for Research and Advanced Studies via an ICREA Academia Prize 2019. Z.H. was partly supported by a Sloan Fellowship, NSF grants DMS-1600561 and DMS-1654692, and a Simons Collaboration Grant. A.M. was partly supported by Progetto di Ricerca GNAMPA - INdAM 2018 "Moti stabili ed instabili in equazioni di tipo Schrödinger". M.P. and E.H. were partially supported by PRIN 2015 "Variational methods in analysis, geometry and physics".

References

[1] Baldi, P., Berti, M., Haus, E., Montalto, R.: Time quasi-periodic gravity water waves in finite depth. Invent. Math. **214**, 739–911 (2018) Zbl 1445.76017 MR 3867631

[2] Bambusi, D., Maspero, A.: Birkhoff coordinates for the Toda lattice in the limit of infinitely many particles with an application to FPU. J. Funct. Anal. **270**, 1818–1887 (2016) Zbl 1335.37047 MR 3452718

[3] Berti, M., Biasco, L.: Branching of Cantor manifolds of elliptic tori and applications to PDEs. Comm. Math. Phys. **305**, 741–796 (2011) Zbl 1230.37092 MR 2819413

[4] Berti, M., Bolle, P.: Quasi-periodic solutions with Sobolev regularity of NLS on $\mathbb{T}^d$ with a multiplicative potential. J. Eur. Math. Soc. **15**, 229–286 (2013) Zbl 1260.35196 MR 2998835

[5] Berti, M., Corsi, L., Procesi, M.: An abstract Nash-Moser theorem and quasi-periodic solutions for NLW and NLS on compact Lie groups and homogeneous manifolds. Comm. Math. Phys. **334**, 1413–1454 (2015) Zbl 1312.35157 MR 3312439

[6] Bourgain, J.: Fourier transform restriction phenomena for certain lattice subsets and applications to nonlinear evolution equations. I. Schrödinger equations. Geom. Funct. Anal. **3**, 107–156 (1993) Zbl 0787.35097 MR 1209299

[7] Bourgain, J.: On the growth in time of higher Sobolev norms of smooth solutions of Hamiltonian PDE. Int. Math. Res. Notices **1996**, 277–304 Zbl 0934.35166 MR 1386079

[8] Bourgain, J.: Quasi-periodic solutions of Hamiltonian perturbations of 2D linear Schrödinger equations. Ann. of Math. (2) **148**, 363–439 (1998) Zbl 0928.35161 MR 1668547

[9] Bourgain, J.: Problems in Hamiltonian PDE's. Geom. Funct. Anal., Special Volume, Part I, 32–56 (2000) Zbl 1050.35016 MR 1826248

[10] Bourgain, J.: Remarks on stability and diffusion in high-dimensional Hamiltonian systems and partial differential equations. Ergodic Theory Dynam. Systems **24**, 1331–1357 (2004) Zbl 1087.37056 MR 2104588

[11] Catoire, F., Wang, W.-M.: Bounds on Sobolev norms for the defocusing nonlinear Schrödinger equation on general flat tori. Comm. Pure Appl. Anal. **9**, 483–491 (2010) Zbl 1189.35301 MR 2600446

[12] Colliander, J., Keel, M., Staffilani, G., Takaoka, H., Tao, T.: Transfer of energy to high frequencies in the cubic defocusing nonlinear Schrödinger equation. Invent. Math. **181**, 39–113 (2010) Zbl 1197.35265 MR 2651381

[13] Colliander, J., Kwon, S., Oh, T.: A remark on normal forms and the "upside-down" I-method for periodic NLS: growth of higher Sobolev norms. J. Anal. Math. **118**, 55–82 (2012) Zbl 1310.35213 MR 2993022

[14] Colliander, J. E., Delort, J.-M., Kenig, C. E., Staffilani, G.: Bilinear estimates and applications to 2D NLS. Trans. Amer. Math. Soc. **353**, 3307–3325 (2001) Zbl 0970.35142 MR 1828607

[15] De Silva, D., Pavlović, N., Staffilani, G., Tzirakis, N.: Global well-posedness for a periodic nonlinear Schrödinger equation in 1D and 2D. Discrete Contin. Dynam. Systems **19**, 37–65 (2007) Zbl 1293.35291 MR 2318273

[16] Delort, J.-M.: Growth of Sobolev norms of solutions of linear Schrödinger equations on some compact manifolds. Int. Math. Res. Notices **2010**, 2305–2328 Zbl 1229.35040 MR 2652223

[17] Douady, R.: Stabilité ou instabilité des points fixes elliptiques. Ann. Sci. École Norm. Sup. (4) **21**, 1–46 (1988) Zbl 0656.58020 MR 944100

[18] Douady, R., Le Calvez, P.: Exemple de point fixe elliptique non topologiquement stable en dimension 4. C. R. Acad. Sci. Paris Sér. I Math. **296**, 895–898 (1983) Zbl 0535.58015 MR 715330

[19] Eliasson, L. H., Kuksin, S. B.: KAM for the nonlinear Schrödinger equation. Ann. of Math. (2) **172**, 371–435 (2010) Zbl 1201.35177 MR 2680422

[20] Faou, E., Gauckler, L., Lubich, C.: Plane wave stability of the split-step Fourier method for the nonlinear Schrödinger equation. Forum Math. Sigma **2**, art. e5, 45 pp. (2014) Zbl 1297.65127 MR 3264246

[21] Geng, J., Xu, X., You, J.: An infinite dimensional KAM theorem and its application to the two dimensional cubic Schrödinger equation. Adv. Math. **226**, 5361–5402 (2011) Zbl 1213.37104 MR 2775905

[22] Gérard, P., Grellier, S.: The cubic Szegő equation. Ann. Sci. École Norm. Sup. (4) **43**, 761–810 (2010) Zbl 1228.35225 MR 2721876

[23] Gérard, P., Grellier, S.: Effective integrable dynamics for a certain nonlinear wave equation. Anal. PDE **5**, 1139–1155 (2012) Zbl 1268.35013 MR 3022852

[24] Gérard, P., Grellier, S.: An explicit formula for the cubic Szegő equation. Trans. Amer. Math. Soc. **367**, 2979–2995 (2015) Zbl 1318.37024 MR 3301889

[25] Grébert, B., Kappeler, T.: The Defocusing NLS Equation and Its Normal Form. EMS Ser. Lect. Math., Eur. Math. Soc., Zürich (2014) Zbl 1298.35002 MR 3203027

[26] Guardia, M.: Growth of Sobolev norms in the cubic nonlinear Schrödinger equation with a convolution potential. Comm. Math. Phys. **329**, 405–434 (2014) Zbl 1296.35169 MR 3207008

[27] Guardia, M., Haus, E., Procesi, M.: Growth of Sobolev norms for the analytic NLS on $\mathbb{T}^2$. Adv. Math. **301**, 615–692 (2016) Zbl 1353.35260 MR 3539385

[28] Guardia, M., Kaloshin, V.: Growth of Sobolev norms in the cubic defocusing nonlinear Schrödinger equation. J. Eur. Math. Soc. **17**, 71–149 (2015) Zbl 1311.35284 MR 3312404

[29] Guardia, M., Kaloshin, V.: Erratum to "Growth of Sobolev norms in the cubic defocusing nonlinear Schrödinger equation" J. Eur. Math. Soc. **19**, 601–602 (2017) Zbl 1433.35364 MR 3605027

[30] Guardia, M., Kaloshin, V.: Orbits of nearly integrable systems accumulating to KAM tori. arXiv:1412.7088 (2014)

[31] Hani, Z.: Long-time instability and unbounded Sobolev orbits for some periodic nonlinear Schrödinger equations. Arch. Ration. Mech. Anal. **211**, 929–964 (2014) Zbl 1293.35298 MR 3158811

[32] Hani, Z., Pausader, B., Tzvetkov, N., Visciglia, N.: Modified scattering for the cubic Schrödinger equation on product spaces and applications. Forum Math. Pi **3**, art. e4, 63 pp. (2015) Zbl 1326.35348 MR 3406826

[33] Haus, E., Procesi, M.: Growth of Sobolev norms for the quintic NLS on T^2. Anal. PDE **8**, 883–922 (2015) Zbl 1322.35126 MR 3366006

[34] Kaloshin, V., Mather, J. N., Valdinoci, E.: Instability of resonant totally elliptic points of symplectic maps in dimension 4. Astérisque **297**, 79–116 (2004) Zbl 1156.37313 MR 2135676

[35] Kappeler, T., Maspero, A., Molnar, J., Topalov, P.: On the convexity of the KdV Hamiltonian. Comm. Math. Phys. **346**, 191–236 (2016) Zbl 1353.35251 MR 3528420

[36] Kappeler, T., Schaad, B., Topalov, P.: Scattering-like phenomena of the periodic defocusing NLS equation. Math. Res. Lett. **24**, 803–826 (2017) Zbl 1390.35330 MR 3696604

[37] Kuksin, S. B.: Growth and oscillations of solutions of nonlinear Schrödinger equation. Comm. Math. Phys. **178**, 265–280 (1996) Zbl 0862.35112 MR 1389904

[38] Kuksin, S., Pöschel, J.: Invariant Cantor manifolds of quasi-periodic oscillations for a nonlinear Schrödinger equation. Ann. of Math. (2) **143**, 149–179 (1996) Zbl 0847.35130 MR 1370761

[39] Kuksin, S. B.: On turbulence in nonlinear Schrödinger equations. Geom. Funct. Anal. **7**, 783–822 (1997) Zbl 0912.35143 MR 1465602

[40] Kuksin, S. B.: Oscillations in space-periodic nonlinear Schrödinger equations. Geom. Funct. Anal. **7**, 338–363 (1997) Zbl 0874.35113 MR 1445390

[41] Maspero, A.: Tame majorant analyticity for the Birkhoff map of the defocusing nonlinear Schrödinger equation on the circle. Nonlinearity **31**, 1981–2030 (2018) Zbl 1391.35356 MR 3816662

[42] Maspero, A.: Lower bounds on the growth of Sobolev norms in some linear time dependent Schrödinger equations. Math. Res. Lett. **26**, 1197–1215 (2019) Zbl 1435.35139 MR 4028118

[43] Maspero, A., Procesi, M.: Long time stability of small finite gap solutions of the cubic nonlinear Schrödinger equation on $\mathbb{T}^2$. J. Differential Equations **265**, 3212–3309 (2018) Zbl 1428.35529 MR 3812225

[44] Planchon, F., Tzvetkov, N., Visciglia, N.: On the growth of Sobolev norms for NLS on 2- and 3-dimensional manifolds. Anal. PDE **10**, 1123–1147 (2017) Zbl 1371.35276 MR 3668586

[45] Pocovnicu, O.: Explicit formula for the solution of the Szegö equation on the real line and applications. Discrete Contin. Dynam. Systems **31**, 607–649 (2011) Zbl 1235.35263 MR 2825631

[46] Pocovnicu, O.: First and second order approximations for a nonlinear wave equation. J. Dynam. Differential Equations **25**, 305–333 (2013) Zbl 1270.65060 MR 3054639

[47] Pöschel, J.: A KAM-theorem for some nonlinear partial differential equations. Ann. Scuola Norm. Sup. Pisa Cl. Sci. (4) **23**, 119–148 (1996) Zbl 0870.34060 MR 1401420

[48] Procesi, C., Procesi, M.: A KAM algorithm for the resonant non-linear Schrödinger equation. Adv. Math. **272**, 399–470 (2015) Zbl 1312.37047 MR 3303238

[49] Procesi, M., Procesi, C.: A normal form for the Schrödinger equation with analytic nonlinearities. Comm. Math. Phys. **312**, 501–557 (2012) Zbl 1277.35318 MR 2917174

[50] Procesi, M., Xu, X.: Quasi-Töplitz functions in KAM theorem. SIAM J. Math. Anal. **45**, 2148–2181 (2013) Zbl 1304.37056 MR 3072759

[51] Rousset, F., Tzvetkov, N.: Transverse instability of the line solitary water-waves. Invent. Math. **184**, 257–388 (2011) Zbl 1225.35024 MR 2793858

[52] Rousset, F., Tzvetkov, N.: Stability and instability of the KdV solitary wave under the KP-I flow. Comm. Math. Phys. **313**, 155–173 (2012) Zbl 1252.35052 MR 2928221

[53] Sohinger, V.: Bounds on the growth of high Sobolev norms of solutions to nonlinear Schrödinger equations on $\mathbb{R}$. Indiana Univ. Math. J. **60**, 1487–1516 (2011) Zbl 1254.35212 MR 2996998 .

[54] Sohinger, V.: Bounds on the growth of high Sobolev norms of solutions to nonlinear Schrödinger equations on S^1. Differential Integral Equations **24**, 653–718 (2011) Zbl 1249.35310 MR 2830703

[55] Sohinger, V.: Bounds on the growth of high Sobolev norms of solutions to 2D Hartree equations. Discrete Contin. Dynam. Systems **32**, 3733–3771 (2012) Zbl 1247.35116 MR 2945837

[56] Staffilani, G.: Quadratic forms for a 2-D semilinear Schrödinger equation. Duke Math. J. **86**, 79–107 (1997) Zbl 0872.35104 MR 1427846

[57] Wang, W.-M.: Energy supercritical nonlinear Schrödinger equations: quasiperiodic solutions. Duke Math. J. **165**, 1129–1192 (2016) Zbl 1342.35358 MR 3486416

[58] Wayne, C. E.: Periodic and quasi-periodic solutions of nonlinear wave equations via KAM theory. Comm. Math. Phys. **127**, 479–528 (1990) Zbl 0708.35087 MR 1040892

[59] Zhang, J., Cheng, C.-Q.: Asymptotic trajectories of KAM torus. arXiv:1312.2102 (2013)

[60] Zhong, S.: The growth in time of higher Sobolev norms of solutions to Schrödinger equations on compact Riemannian manifolds. J. Differential Equations **245**, 359–376 (2008) Zbl 1143.58012 MR 2428002