LONG-TIME INSTABILITY AND UNBOUNDED SOBOLEV ORBITS FOR SOME PERIODIC NONLINEAR SCHRÖDINGER EQUATIONS

ZAHER HANI

ABSTRACT. We study the energy cascade problematic for some nonlinear Schrödinger equations on $T^2$ in terms of the growth of Sobolev norms. We define the notion of long-time strong instability and establish its connection to the existence of unbounded Sobolev orbits. This connection is then explored for a family of cubic Schrödinger nonlinearities that are equal or closely related to the standard polynomial one $|u|^2u$. Most notably, we prove the existence of unbounded Sobolev orbits for a family of Hamiltonian cubic nonlinearities that includes the resonant cubic NLS equation (a.k.a. the first Birkhoff normal form).

1. Introduction

This manuscript is concerned with the dynamical behavior of solutions of the space-periodic nonlinear Schrödinger equation:

\[
\begin{aligned}
-i\partial_t u + \Delta u &= \mathcal{N}(u) \\
 u(0, x) &= u_0(x), \quad x \in T^2
\end{aligned}
\]  

(1.1)

where $\mathcal{N}(u)$ stands for a cubic Hamiltonian nonlinearity (i.e. $\mathcal{N}(u) = \frac{\partial}{\partial u} G(u, \overline{u})$) that is either equal or closely related to the polynomial cubic one given by $\mathcal{N}_\infty(u) = |u|^2u$. We will work in spatial dimension $d = 2$, but all the results extend directly to higher dimensions by considering solutions that depend only on the first two spatial coordinates. The initial data will be assumed to belong to some Sobolev space $H^s(T^2)$ with $s > 0$.

We are particularly interested in the broad question of whether there exists global solutions to (1.1) that exhibit oscillations at smaller and smaller spatial scales as time becomes larger and larger. In frequency space, this corresponds to solutions whose conserved energy moves from low-frequency concentration zones to high-frequency ones, in what is commonly referred to as the forward or direct cascade. One way to capture this energy dynamic in frequency space is to look at the behavior of Sobolev norms given by:

\[
\|u(t)\|_{H^s(T^2)} = \left( \sum_{n \in \mathbb{Z}^2} (n)^{2s} |\hat{u}(t, n)|^2 \right)^{1/2}
\]

whose growth in time, for $s > 1$, signals that a larger contribution to $\|u(t)\|_{H^s}$ is coming from higher and higher frequencies. The main question in this perspective is to find global solutions $u(t)$ of (1.1) for which:

\[
\sup_{t \in \mathbb{R}} \|u(t)\|_{H^s(T^2)} = +\infty? \tag{1.2}
\]

The author is partly supported by a Simons Postdoctoral Fellowship.
A different perspective to energy cascade is provided by the physical theory of weak turbulence, which can be, loosely speaking, defined as the out-of-equilibrium statistics of random nonlinear waves \(^{[14, 31]}\). This theory was invigorated in the 60s of the past century in plasma physics, where the subject adopted its name, and in the study of water waves. In these contexts, energy cascades were used to explain particle and energy transfer across tokomak plasma or how momentum is transferred from wind to the ocean. In contrast to the somewhat deterministic nature of question (1.2) above, weak turbulence theory addresses the cascade issue by looking at wave spectra (the average energy or mass density as a function of the wavelength) rather than individual wave trajectories. These spectra turn out to satisfy (after appropriate averaging and limiting operations) certain kinetic equations, for which Zakharov was able to find particular solutions (now called Kolmogorov-Zakharov spectra\(^1\)) that correspond to constant energy flux through scales \(^{[30]}\). Of course, a constant energy flux through scales is often not consistent with the conservation of energy of Hamiltonian systems like (1.1), a fact which necessitates the presence of feeding source and a dissipative sink of energy located at large and small scales respectively.

While these spectra offer a valid evidence (from a physics point of view) of some form of cascade phenomenon happening at the medium (unexcited) frequencies, many of the manipulations involved are still far from rigorous. In this paper, we will adopt the first perspective mentioned above of searching for unbounded Sobolev orbits as in (1.2). This seems to ask for a somewhat stronger form of cascade in comparison to the second perspective, which might explain why a satisfactory answer to (1.2) is still open for most equations of interest, including cubic NLS. For the polynomial cubic nonlinearity \(N_{\infty}(u) = |u|^2 u\), the conservation laws of mass given by:

\[
M[u(t)] := \int_{\mathbb{T}^2} |u(t, x)|^2 \, dx = M[u(0)]
\]  

and energy given by

\[
E[u(t)] := \int_{\mathbb{T}^2} \frac{1}{2} |\nabla u(t, x)|^2 + \frac{1}{4} |u(t, x)|^4 \, dx = E[u(0)]
\]

give global \textit{a priori} bounds on the \(L^2\) and \(H^1\) norms of the solution. So the answer to the question in (1.2) is negative for \(s = 0, 1\). We should note that the one-dimensional analogue of the cubic NLS equation is completely integrable \(^{[32]}\) and higher conservation laws give global a priori bounds for \(H^s(\mathbb{T}^2)\) at least when \(s \in \mathbb{N}\). However, it is expected that this behavior is particular to the one-dimensional case and that, for \(s > 0, s \neq 1\), solutions exhibiting unbounded orbits in \(H^s(\mathbb{T}^2)\) actually exist and are abundant. Proving this seems to be quite a difficult task that was considered by Bourgain in \(^{[4]}\) as one of the next century problems in Hamiltonian PDE.

It is worth mentioning that polynomial-in-time upper bounds on the growth of \(\|u(t)\|_{H^s}\) have been proved by several authors starting with pioneering works of Bourgain \(^{[3]}\) and Staffilani \(^{[28]}\) (see also \(^{[9, 10, 6]}\) and related work in \(^{[8, 33, 27, 12]}\)). But this polynomial growth is expected to be far from optimal. In fact, Bourgain conjectured in \(^{[3]}\) that the growth should be sub-polynomial in the sense that:

\[
\|u(t)\|_{H^s} \leq C_\epsilon t^\epsilon \quad \text{for any } \epsilon > 0,
\]

but this remains open as well.

\(^1\)Because they resemble Kolmogorov spectra in the theory of hydrodynamic turbulence.
We end this background overview with a remark concerning the effect of the growth of high Sobolev norms on the frequency dynamics of solutions. The conservation of mass and energy imply that any increase in high Sobolev norms is not only accompanied by a migration of energy to high frequencies, but also a migration of mass to low frequencies. This phenomenon is known in the physics literature as the inverse cascade of mass, and it balances the forward cascade of energy to keep the system conservative.

1.1. Long-time strong instability. Despite attracting a lot of attention, the existence of unbounded orbits as in (1.2) remains unproved for the polynomial cubic nonlinearity. However, some results exhibiting some form of cascade that is strictly weaker than (1.2) have been obtained in the past years (see for example [24, 11, 7, 19] and [16, 17, 25, 26, 18] for work on other Hamiltonian PDE). Of particular importance for this article is the result of Colliander, Keel, Staffilani, Takaoka, and Tao in [11] (see also [19] for a refinement) who proved that for any given parameters $s \in (0, \infty) \setminus \{1\}$, $K \gg 1$, and $0 < \delta \ll 1$, there exists a global solution $u(t,x)$ of (2.1) and a time $T > 0$ such that

$$\|u(0)\|_{H^s} \leq \delta \quad \text{and} \quad \|u(T)\|_{H^s} \geq K.$$  \hspace{1cm} (CKSTT)

The above result is equivalent to saying that the orbit of any $H^s$-neighborhood of the origin under the nonlinear flow of the (polynomial) cubic NLS is not uniformly bounded in $H^s$. This is still considerably weaker than (1.2) that requires a single orbit to be unbounded.

It will be useful to interpret the result (CKSTT) above as the long-time strong instability of the flow near 0, a concept that can be defined in great generality as follows:

**Definition 1.2** (Long-time strong instability). Let $X$ be a Banach space and suppose that $S(t)$ is a continuous dynamical system on $X$. We say that the flow exhibits long-time strong instability near $\phi \in X$ if for every $0 < \delta < 1$ and $K \geq 1$, there exists $\phi^* \in B_X(\phi, \delta)$ such that

$$\sup_{t > 0} \|S(t)\phi^*\|_X \geq K.$$ 

While the long-time strong instability near some $\phi$ is much weaker than the existence of an unbounded orbit in $X$ satisfying $\sup_t \|S(t)\phi\|_X = \infty$, it turns out that a sufficient extension of it is enough to guarantee the abundance of unbounded orbits. We state this observation in the proposition below also in the generality of continuous dynamical systems on Banach spaces. For this, we will need an additional assumption on the flow, which is well-posedness: We say that the flow $S(t)$ is well-posed if for every $T > 0$, the map

$$\phi \mapsto S(t)\phi$$

is continuous from $X \to \in C_t([-T, T] \times X)$.

**Proposition 1.3.** (From long-time strong instability to generic unbounded orbits) Let $X$ be a Banach space and $S(t)$ be a well-posed continuous dynamical system on $X$.

---

2This is the exact analogue of the forward and backward cascade of enstrophy and energy (respectively) discovered by Kraichnan [23] in the theory of hydrodynamic turbulence.

3The theorem is stated in [11] for $s > 1$. However, minor modifications to the proof allow one to extend that range to $s > 0, s \neq 1$. See [22] for details.

4This means that for every $x_0 \in X$, $S(t)x_0 \in C_t([-T, T]; X)$ for any $T > 0$, and satisfies $S(0)x_0 = x_0$ and $S(t_1 + t_2)x_0 = S(t_1)S(t_2)x_0$ for any $t_1, t_2 \in \mathbb{R}$.
(i) If the flow \( S(t) \) exhibits long-time strong instability near a dense set of initial data \( \phi_0 \in X \), then the set of unbounded orbits is generic in the sense of Baire category; in particular it is nonempty. More precisely, the set
\[
\{ \phi_0 \in X : \sup_{t>0} \| S(t)\phi_0 \| < \infty \}
\]
is meager (countable union of nowhere dense sets).

(ii) Let \( F \) be a closed subset of \( X \) and \( D \subset F \) be dense in \( F \). Suppose that for any \( \phi \in D, \delta > 0 \), \( K \geq 1 \), there exists \( \phi^* \in F \) such that \( \| \phi - \phi^* \|_X \leq \delta \) and \( \sup_{t>0} \| S(t)\phi^* \|_X \geq K \). Then the set \( \{ \psi \in F : \sup_{t \in \mathbb{R}} \| S(t)\psi \|_X = \infty \} \) is co-meager in \( F \) (in particular non-empty).

The above proposition can somehow be understood as a “nonlinear uniform boundedness principle” and its proof is a straightforward application of the Baire Category theorem. Indeed, let
\[
\mathcal{B} = \{ u_0 \in F : \sup_{t>0} \| S(t)u_0 \|_X < \infty \}
\]
and
\[
\mathcal{B}_N = \{ u_0 \in F : \sup_{t>0} \| S(t)u_0 \|_X \leq N \}.
\]
Obviously, \( \mathcal{B} = \bigcup_{N \in \mathbb{Z}, N \geq 1} \mathcal{B}_N \). \( \mathcal{B}_N \) is closed in \( X \) courtesy of the well-posedness of the flow \( S(t) \). The long-time strong instability of the flow near every \( u_0 \in D \) and the fact that \( D \) is dense in \( F \) implies that \( \mathcal{B}_N \) has no interior as a subset of \( F \) (because for any \( u_0 \in D \) and any \( \delta > 0 \), one can find \( \psi_0 \in B_X(0, \delta) \cap F \) whose orbit exits the ball \( B_X(0, 2N) \)). As a result, \( \mathcal{B} \subset F \) is a countable union of closed nowhere dense sets. Since \( F \) is a complete metric space, \( F \setminus \mathcal{B} \neq \emptyset \) and the set of unbounded orbits is co-meager in \( F \).

In practice, proving the genericity of unbounded orbits by establishing long-time strong instability near a dense set of \( X \), as in part (i) of the above proposition, can be quite a demanding task. Nonetheless, the existence of at least one unbounded orbit can be seen to trivially imply the validity of part (ii) of the proposition for some closed subset of \( X \), which makes the implication statement in that part an equivalence. In what follows, we will apply the above proposition for the NLS flow in \( H^s \) and prove long-time strong instability near some carefully chosen families of \( \phi \in X \). This will be sufficient to invoke part (ii) of Proposition \(^{15}\) in order to conclude the existence of unbounded orbits for certain cubic NLS nonlinearities that are arbitrarily close (but not quite equal) to the polynomial one.

1.4. Implications for NLS. Proposition \(^{13}\) suggests a program to proving the existence of unbounded orbits of \( H^s \) by proving analogues of the \( \text{CKSTT} \) instability result in \( H^s \) near sufficiently large families of initial data in \( H^s \). Indeed, with the Banach space \( X \) taken to be \( H^s(\mathbb{T}^2) \), the NLS flow defines a well-posed continuous dynamical system whenever equation \( \text{H}^s \) is globally well-posed in \( H^s(\mathbb{T}^2) \).

For the polynomial cubic nonlinearity \( |u|^2 u \), we take a rather timid step in the direction suggested by Proposition \(^{13}\) and prove long-time strong instability of the flow in \( H^s(\mathbb{T}^2) \) near plane wave initial data given by \( u_0^* = Ae^{inx} (A \in \mathbb{C}, n \in \mathbb{Z}^2) \) in the range \( 0 < s < 1 \):\(^5\)

\(^5\)I.e. the intersection of countably many dense open sets.

\(^6\)For instance, this is the case for the polynomial cubic nonlinearity whenever \( s > 2/3 \) \(^5\) \(^6\).
More precisely, our second result proves the existence of unbounded orbits for the following nonlinearity that are “almost polynomial”. By this we mean that they are obtained from the polynomial cubic nonlinearity, a positive answer to question (1.2) can be given for some cubic nonlinearities while we still cannot push the program suggested by Proposition 1.3 to its end for the polynomial cubic equation (2.1). Solutions to (2.1) with initial data $Ae^{inx}$ are explicit and are given by $Ae^{i(n.x+|n|^2t+|A|^2t)}$. While the proof of this theorem relies heavily on the constructions in [11], it offers a stark example of some of the difficulties encountered as one tries to prove long-time strong instability for (2.1) near non-zero initial data. In fact, the major issue in proving Theorem 1.5 is not only to find a tractable mechanism for norm growth (which will be a modification of that in [11]), but also to find ways to limit or incorporate interactions between two different parts of the solution: one is the “base solution” that has $O(1)$ size in $H^s$ and is a manifestation of the unperturbed plane wave solution $Ae^{i(n.x+|n|^2t+|A|^2t)}$; and the other is the “turbulent” part that has $O(\delta)$ size in $L^2$, but whose Sobolev norm has to grow to become larger than $K$ after time $T$.

While we believe that the restriction in Theorem 1.5 to the range $0 < s < 1$ is probably technical, we should mention that for large enough $s > 1$, plane wave solutions were shown in [15] to exhibit polynomial-time Nekhorashev-type stability in the following sense: For any $N \in \mathbb{N}$, there exists $\epsilon_0$ such if $\epsilon < \epsilon_0$, most solutions that are initially $\epsilon$-close to a plane wave $Ae^{inx}$ in $H^s$ (for $s$ large enough) will remain as such for a period of time that is $O(\epsilon^{-N})$. Of course, this does not prohibit growth of the Sobolev norm in infinite time, but it suggests the need for a much more elaborate growth mechanism that beats this long stability time interval. Indeed, it might be worth mentioning that the same interactions that require us to restrict the instability Theorem 1.5 to the range $0 < s < 1$ are the ones that allow the authors in [15] to prove their Nekhorashev-type stability result for large enough $s$.

While we still cannot push the program suggested by Proposition 1.3 to its end for the polynomial cubic nonlinearity, a positive answer to question (1.2) can be given for some cubic nonlinearities that are “almost polynomial”. By this we mean that they are obtained from the polynomial cubic nonlinearity $|u|^2u$ by eliminating highly non-resonant terms beyond any arbitrary threshold.

More precisely, our second result proves the existence of unbounded orbits for the following family of Hamiltonian nonlinearities:

$$\mathcal{N}_R(u) = \sum_{n \in \mathbb{Z}^2} \left( \sum_{\Gamma_R(n)} \hat{u}(n_1)\bar{u}(n_2)\hat{u}(n_3) \right) e^{2\pi in.x}, \quad R \in \{0\} \cup \mathbb{N}. \quad (1.6)$$

where $\Gamma_R(n) = \{(n_1, n_2, n_3) \in (\mathbb{Z}^2)^3 : n_1 - n_2 + n_3 = n \text{ and } \omega_4 := |n_1|^2 - |n_2|^2 + |n_3|^2 - |n|^2 \in [-R, R] \}.$

Note that the polynomial cubic nonlinearity $\mathcal{N}_\infty(u) = |u|^2u$ corresponds to taking $R = \infty$ in the above definition. Another well-known nonlinearity that is included in this family is the resonant

---

7I.e. they remain in a $2\epsilon$-neighborhood of a plane wave $A(t)e^{inx}$.

8Non-resonant interactions have a far weaker effect on the dynamical behavior than resonant ones (at least for early time scales), which makes the result in Theorem 1.5 further evidence of the existence of unbounded orbits for the polynomial cubic equation (2.1).
cubic NLS system (a.k.a. the first Birkhoff normal form) that corresponds to $R = 0$. We remark that for any $R \in \{0\} \cup \mathbb{N} \cup \{\infty\}$, the nonlinearity $\mathcal{N}_R(u)$ is Hamiltonian and conserves mass. In fact, in the two extreme cases $R = 0$ and $R = \infty$, the Hamiltonian is positive definite and controls the $\dot{H}^1$ norm of the solution (cf. (1.4) and (4.4)). The same is true for any $R \in \mathbb{N}$ provided that the mass of the solution is smaller than some positive threshold (cf. Section 4). This will allow us to conclude global well-posedness of the flow (cf. Section 4) in a way that allows to apply the paradigm of Proposition 1.3 and prove our main result:

**Theorem 1.6.** For any $R \in \{0\} \cup \mathbb{N}$ and any $s > 1$, there exists global solutions to (1.1) (with $\mathcal{N}(u) = \mathcal{N}_R(u)$) whose orbit is unbounded in $H^s(\mathbb{T}^2)$ as in (1.2). Moreover, initial data corresponding to unbounded orbits are residual in a closed subset of $H^s(\mathbb{T}^2)$.

A few remarks are in order: First, the same conclusion holds in the range of regularities $0 < s < 1$ under the conditional assumption that the system is globally well-posed for sufficiently small data (cf. Section 4). Second, the closed set $\mathcal{F} \subset H^s(\mathbb{T}^2)$ mentioned in Theorem 1.6 that contains the residual set of unbounded orbits, is defined in terms of a geometric and combinatorial condition imposed on the Fourier support. This condition allows one to prove long-time strong instability near a dense subset and conclude as in part (ii) of Proposition 1.3 that bounded orbits are a meager subset of $\mathcal{F}$. Finally, we should mention that Bourgain constructed in [2] a nonlinearity $\mathcal{N}(u)$ that grows linearly in $u$ for which he proved the existence of an unbounded orbit. The novelty in Theorem 1.6 comes from 1) proving the abundance of unbounded orbits for a family of Hamiltonian cubic nonlinearities $\mathcal{N}_R(u)$ that approximates the polynomial cubic one ($R = \infty$) and includes the well-known resonant cubic system ($R = 0$); and 2) illustrating how the program suggested by Proposition 1.3 can be applied to conclude existence and genericness (possibly in some metric subspace) of unbounded orbits.

The paper is organized as follows: in Section 2, we set up our problem in Fourier space and present some basic reductions and notions that will be used in the rest of the paper. In Section 3, we prove Theorem 1.5 and finally Theorem 1.6 is proved in Section 4. Throughout this paper, we write $A \lesssim B$ to signify that there is a constant $C > 0$ such that $A \leq CB$. We also write $A \sim B$ when $A \lesssim B \lesssim A$. If the constant $C$ involved has some explicit dependency, we emphasize it by a subscript. Thus $A \lesssim_u B$ means that $A \leq C(u)B$ for some constant $C(u)$ depending on $u$. In some instances, we use the notation $A \ll B$ to signify that the implicit constant $C$ is large. Finally, we omit all factors of $2\pi$ arising in the definition of the Fourier transform on $\mathbb{T}^2$ as they play no role in our analysis.

**Acknowledgements.** The author would like to thank Terry Tao for introducing him to this line of research as a Ph.D. student, and for all the invaluable discussions and suggestions. The author is supported by a Simons Postdoctoral Fellowship.

## 2. Resonant and finite reductions

We start by considering the equation:

$$\begin{cases} -i\partial_t u + \Delta u = |u|^2 u \\ u(0, x) = u_0(x) \end{cases} \quad (2.1)$$
In this section, we perform some reductions that will be used in the course of the paper. First, we will cast (2.1) as an infinite system of ODE which we call \( (F_{NLS}) \) in \( \{a_n(t)\}_{n \in \mathbb{Z}^2} \) where \( a_n(t) \) are closely related to the Fourier coefficients of the solution \( u(t,x) \). Afterwards we isolate the resonant reduction or the first Birkhoff normal form of this system \( (RF_{NLS}) \), which is obtained from \( (F_{NLS}) \) by deleting all non-resonant interactions. We then explain how to reduce further this resonant system into a finite “Toy system” of ODE as was done in [11] by choosing initial data supported on special subsets of \( \mathbb{Z}^2 \). These reductions, borrowed from [11] and summarized here for completeness, form an important ingredient for the proof of Theorems 1.5 and 1.6.

2.1. **NLS as an infinite system of ODE** \( F_{NLS} \). Equation (2.1) enjoys the following “guage freedom”:

\[
    v(t,x) = e^{iGt}u(t,x)
\]

where \( v \) satisfies the equation:

\[
    -i\partial_t v + \Delta v = (G + |v|^2)v
\]

Using the ansatz:

\[
    v(t,x) = \sum_n a_n(t)e^{i(n.x + |n|^2t)}
\]

we get after a direct calculation that the coefficients \( a_n(t) \) satisfy the following system of ODE:

\[
    -i\partial_t a_n(t) = (G + 2M)a_n - |a_n|^2a_n + \sum_{n_1 - n_2 + n_3 = n, n_1, n_3 \neq n} a_{n_1}a_{n_2}a_{n_3}e^{i\omega_4t} \tag{2.2}
\]

where \( M \) stands for the mass:

\[
    M = \sum_n |a_n|^2
\]

which is conserved in time. Choosing \( G = -2M \) we end up with an infinite system of ODE which we denote by \( (F_{NLS}) \) given by:

\[
    -i\partial_t a_n = -a_n|a_n|^2 + \sum_{n_1, n_2, n_3 \in \Gamma(n)} a_{n_1}a_{n_2}a_{n_3}e^{i\omega_4t} \tag{2.3}
\]

where \( \Gamma(n) = \{(n_1, n_2, n_3) \in (\mathbb{Z}^2)^3 : n_1 - n_2 + n_3 = n \text{ and } n_1, n_3 \neq n\} \).

We remark that at any time \( t \), one can recover \( u \) from \( v \) and vice-versa and the two solutions have the same Sobolev norms. It is easy to see that this system of ODE is locally well-posed in \( l^1(\mathbb{Z}^2) \) (see [11]).

2.2. **Resonant reduction** \( RF_{NLS} \). The first interaction term in (2.3) is a self interaction term, whereas the second term is a sum taken over all parallelograms (which may collapse to segments) with \( n \) as one of its vertices (and \( n_2 \) being the opposite vertex). Analyzing the resonance factor \( \omega_4 \) we get that the parallelogram is a rectangle exactly when \( \omega_4 = 0 \). In fact,

\[
    w_4 = |n_1|^2 - |n_2|^2 + |n_3|^2 - |n|^2 = |n_1 - n|^2 - |n_2 - n|^2 + |n_3 - n|^2 = -2(n_1 - n)(n_3 - n).
\]
The interactions for which $\omega_4 = 0$ are called resonant interaction, whereas those for which $\omega_4 \neq 0$ are non-resonant. Resonant interactions will play a particularly important role in the analysis, a fact which motivates us to define the set of all resonant non-self interactions $\Gamma_{\text{res}}(n) \subset \Gamma(n)$ as:

$$\Gamma_{\text{res}}(n) = \{(n_1, n_2, n_3) \in \Gamma(n) : \omega_4 = |n_1|^2 - |n_2|^2 + |n_3|^2 - |n|^2 = 0\}. \quad (2.4)$$

The resonant reduction ($\mathcal{RFNLS}$) of ($\mathcal{FNLS}$) is obtained by deleting all non-resonant interactions and retaining the resonant ones only. More precisely, ($\mathcal{RFNLS}$) is the system:

$$-i\partial_t r_n = -r_n|r_n|^2 + \sum_{n_1, n_2, n_3 \in \Gamma_{\text{res}}(n)} r_{n_1}r_{n_2}r_{n_3} \quad (2.5)$$

This system can be also obtained as the first Birkhoff normal form of ($2.3$). It will be important to notice that all non-self interactions (the second term on the RHS of ($2.5$)) of the resonant system ($\mathcal{RFNLS}$) occur in rectangles: each $n \in \mathbb{Z}^2$ interacts with other frequencies by forming a rectangle with $n$ as one of its vertices. This will be the key to collapsing this system to a finite system of ODE by choosing the initial data to be supported on special subsets of $\mathbb{Z}^2$ that are somewhat “closed under completing rectangles”. This is the content of the first two properties below.

2.3. Nuclear families and the finite reduction [11]. Suppose that for some subset $\Lambda \subset \mathbb{Z}^2$ we have:

- Property $I_\Lambda$ (Initial data): The initial data $r_n(0)$ is supported on $\Lambda$, i.e. $r_n(0) = 0$ unless $n \in \Lambda$.
- Property $II_\Lambda$ (Closure): If $n_1, n_2, n_3 \in \Lambda$ are three vertices of a rectangle, then the forth vertex of that rectangle is also in $\Lambda$. This implies that if $(n_1, n_2, n_3) \in \Gamma_{\text{res}}(n) \cap \Lambda^3$, then $n \in \Lambda$.

These two properties imply that the support of $r_n(t)$ remains restricted to $\Lambda$ for all later times. In fact, if $B(t) := \sum_{n \notin \Lambda} |r_n(t)|^2$, then it is easy to see that $|B'(t)| \leq C|B(t)|$ and the claim follows by applying Gronwall’s inequality (see [11] for details or Lemma 4.3 for a similar argument).

This observation allows to reduce the infinite system of ODE ($\mathcal{RFNLS}$) into a finite system of ODE by restricting the initial data to a finite subset $\Lambda$ that satisfies the closure property above. As was done in [11], we will further reduce this finite system into a smaller one by imposing some extra structure on the set $\Lambda$. This structure comes from the following hierarchy:

We will assume that $\Lambda$ splits into a disjoint union $\Lambda = \Lambda_1 \cup \ldots \cup \Lambda_P$ for some integer $P > 1$ and call $\Lambda_j$ the $j$–th generation of $\Lambda$. We define a nuclear family in $\Lambda$ as a rectangle $(n_1, n_2, n_3, n_4)$ where the frequencies $n_1, n_3$ (known as the “parents”) belong to the $j$–th generation $\Lambda_j$ while the “children” $n_2, n_4$ belong to $\Lambda_{j+1}$. Of course, if $(n_1, n_2, n_3, n_4)$ is a nuclear family then so are the trivial permutations $(n_3, n_2, n_1, n_4), (n_1, n_4, n_3, n_2)$, and $(n_3, n_4, n_1, n_2)$. In addition to the properties $I_\Lambda$ and $II_\Lambda$ above, we require:
• Property III\(\Lambda\) (Existence and uniqueness of spouse and children): For each \(1 \leq j < P\) and every \(n_1 \in \Lambda_j\), there exists a unique spouse \(n_3 \in \Lambda_j\) and unique (up to trivial permutations) children \(n_2, n_4 \in \Lambda_{j+1}\) such that \((n_1, n_2, n_3, n_4)\) is a nuclear family in \(\Lambda\).

• Property IV\(\Lambda\) (Existence and uniqueness of parents and siblings): For each \(1 \leq j < P\) and every \(n_2 \in \Lambda_{j+1}\) there exists a unique sibling \(n_4 \in \Lambda_{j+1}\) and unique (up to permutation) parents \(n_1, n_3 \in \Lambda_j\) such that \((n_1, n_2, n_3, n_4)\) is a nuclear family in \(\Lambda\).

• Property V\(\Lambda\) (Non-degeneracy): A sibling of any frequency \(n\) is never equal to its spouse.

• Property VI\(\Lambda\) (Faithfulness): Apart from nuclear families \(\Lambda\) contains no other rectangles. In fact, by the closure property II\(\Lambda\), this also means that it contains no right angled triangles other than those coming from vertices of nuclear families.

Assuming that such a set exists, the system \((RFNLS)\) can be written as:

\[
-i\partial_tr_n = -r_n|r_n|^2 + 2r_{n,\text{child}}r_{n,\text{child}} + 2r_{n,\text{parent}}r_{n,\text{parent}} - 2r_{n,\text{sibling}}
\]

where we used the fact that the only rectangles that contain \(n\) as a vertex are those coming from nuclear families where \(n\) is either a parent (second term on the RHS) or a child (third term on RHS). The constant factor 2 in front of these terms comes from the trivial permutations of the nuclear families.

The final property imposed on the initial data allows to collapse the \((RFNLS)\) system further.

• Property VII\(\Lambda\) (Intragenerational equality): The function \(n \mapsto r_n(0)\) is constant on each generation \(\Lambda_j\). In other words, \(r_{n_1}(0) = r_{n_2}(0)\) whenever \(n_1, n_2 \in \Lambda_j\).

One can verify (either by a Gronwall or a bootstrap argument) that Property VII\(\Lambda\) is preserved by the flow of \((RFNLS)\) (cf. [11]) in the sense that intragenerational equality remains true for all time. By setting \(r_n(t) = b_j(t)\) whenever \(n \in \Lambda_j\), we arrive at the following system for \(b_j\):

\[
-i\partial_tb_j = -b_j|b_j|^2 + 2b_{j+1}^2 + 2b_{j-1}^2
\]

with the convention\(\footnote{Actually, one can easily verify that if one regards \((2.6)\) as a system of ODE in \(b_j\) with \(j \in \mathbb{Z}\), then the property \(b_j(0) = 0\) for \(j \notin \{1, \ldots, P\}\) is propagated by the flow.}\) that \(b_0(t) = b_{P+1}(t) = 0\).

With these reductions at hand, the result \((\text{CKSTT})\) was proved in [11] in three main steps:

(i) Construction of the frequency set \(\Lambda\) satisfying Properties I\(\Lambda\)-VII\(\Lambda\) above, along with the norm explosion property:

\[
\sum_{n \in \Lambda_{P-2}} |n|^{2s} < K^2
\]

(ii) Construction of a solution of resonant system \((2.6)\) that is initially concentrated on a low frequency generation (namely \(\Lambda_3\) for technical reasons) and after some time \(T\) concentrated on a high-frequency generation \(\Lambda_{N-2}\). This amounts to proving the following property of the Toy system \((2.6)\):
The above two steps allow to prove the needed Sobolev norm growth for the resonant system (RFNLS) in (2.5). To transfer this growth to (FNLS), one needs:

(iii) **Stability lemma:** In this step, one proves the faithfulness of the approximation of $u(t)$ by $\tilde{u}(t)$ over a time interval that is strictly longer than that needed for the Sobolev norm of (2.5) to grow.

3. **LONG-TIME INSTABILITY OF PLANE WAVES**

In this section, we prove Theorem 1.5. We start by reducing to the case when $v = 0$ by resorting to the Galilean symmetry:

$$u(t, x) \leftrightarrow e^{i(v.x + |v|^2t)}u(t, x + 2vt)$$

(3.1)

Indeed, if $\tilde{u}$ solves (2.1), $\|\tilde{u}(0) - A\|_{H^s(T^2)} \leq \tilde{\delta}$, and $\|\tilde{u}(T)\|_{H^s(T^2)} \geq \tilde{K}$, then $u(t) = e^{i(v.x + |v|^2t)}\tilde{u}(t, x + 2vt)$ also solves (2.1) and satisfies $\|u(0) - Ae^{iv.x}\|_{H^s(T^2)} = \|e^{iv.x}(\tilde{u}(0) - A)\|_{H^s(T^2)} \lesssim v \|\tilde{u}(0) - A\|_{H^s(T^2)} \lesssim v \tilde{\delta}$, while $\|u(T)\|_{H^s(T^2)} = \|e^{iv.x}\tilde{u}(T)\|_{H^s(T^2)} \gtrsim v \tilde{K}$. The result then follows by choosing $\tilde{K}$ and $\tilde{\delta}$ appropriately in terms of $K, \delta, \text{and } v$. Also, we may assume without any loss of generality (thanks to phase translation invariance) that $A \in \mathbb{R}$ and positive.

The initial data will be taken to be $u(0) = A + w(0)$ where $w(0)$ is supported on a frequency set $\Lambda$ similar to that in the previous section. As before, the norm explosion will happen as the energy cascades between the different generations of $\Lambda$. For this to work, we need to limit the interactions between the large mode at 0 and the modes in $\Lambda$, so that this cascade process is not disturbed or reversed. Some of these interactions can be eliminated by imposing an additional “geometric” condition on $\Lambda$ (Property VIIIa below). Others cannot be easily eliminated and they are the reason why we restrict our result to the regime $0 < s < 1$. More precisely, we are referring to non-resonant interactions that happen when a mode at $n \in \Lambda$ interacts non-resonantly with the zero mode to excite the mode at $-n$ which in turn interacts with the zero mode to excite the mode at $n$. The net self-excitation of the mode at $n$ due to this interaction chain is actually resonant (but at a second order) and leads to the restriction on $s$.

In contrast to the argument in [11], it will be important to work on the dilated lattice $D = N\mathbb{Z}^2$ generated by $(N, 0)$ and $(0, N)$. It is easy to verify (via a Gronwall argument for example) that if the initial data is supported in frequency space on $D$, then the emanating solution will also be supported on $D$ for all time. This ultimately follows because the interactions of the (FNLS) system happen in parallelograms and $D$ trivially satisfies the parallelogram closure property: if three vertices of a parallelogram are in $D$, then so is the forth.
Consequently, we will consider the \((FNLS)\) system posed on \(D\):
\[
- i \partial_t c_n(t) = -|c_n|^2 c_n + \sum_{\Gamma(n)} c_{n_1} \overline{c_{n_2}} c_{n_3} e^{i \omega_4 t}
\]  
(3.2)
where \(\Gamma(n) = \{(n_1, n_2, n_3) \in D^3 : n_1 - n_2 + n_3 = n \text{ and } n_1, n_3 \neq n\}\) and \(\omega_4 = |n_1|^2 - |n_2|^2 + |n_3|^2 - |n|^2\). Notice that if \(\omega_4 \neq 0\), then we directly get that \(|\omega_4| \geq N^2\).

Since we are considering solutions with a relatively large zero-frequency mode, we have to isolate the terms of this sum involving this mode. First, we isolate the \(c_0(t)\) equation:
\[
- i \partial_t c_0(t) = -|c_0|^2 c_0 + c_0(t) \Gamma(0) \sum_{n_1 \neq 0} c_{n_1}(t) e^{2|n_1|^2 t} + \sum_{\Gamma(0)} c_{n_1} \overline{c_{n_2}} c_{n_3} e^{i \omega_4 t}
\]  
(3.3)
where \(\Gamma(0) = \{(n_1, n_2, n_3) \in D^3 : n_1 - n_2 + n_3 = 0 \text{ and } n_1, n_2, n_3 \neq 0\}\).

Another calculation shows that the equation of \(c_n\) for \(n \neq 0\) will have the following more elaborate form.
\[
- i \partial_t c_n(t) = -|c_n|^2 c_n + c_0(t) \Gamma(0) \sum_{n_1 \neq 0} c_{n_1}(t) e^{2i|n_1|^2 t} + 2 \left( \sum_{n_2 \neq 0, n} c_{n_2} c_{n+n_2} e^{2i n_2 t} \right) c_0(t) + \sum_{\Gamma(n)} c_{n_1} \overline{c_{n_2}} c_{n_3} e^{i \omega_4 t}
\]  
(3.4)
where \(\Gamma(n) = \{(n_1, n_2, n_3) \in \Gamma(n) : n_1, n_2, n_3 \neq 0\}\).

To construct a norm-explosive solution to (3.4), we will first construct an energy-cascading solution of the resonant system \((RFNLS)\) in a similar way to what was done in [11]. This solution to the resonant system will be blind to the large mode at 0 and will be supported on a finite set \(\Lambda\) located at high frequencies. The main difficulty is to show that on the set \(\Lambda\), solutions to (3.4) can still be approximated by solutions to the resonant system in spite of the additional interaction terms with the large mode at 0.

Denote by \((\tilde{c}_n(t))\) the solution to the resonant system:
\[
- i \partial_t \tilde{c}_n(t) = -|\tilde{c}_n(t)|^2 \tilde{c}_n(t) + \sum_{\Gamma_{res}(n)} \tilde{c}_{n_1}(t) \overline{\tilde{c}_{n_2}(t)} \tilde{c}_{n_3}(t)
\]  
(3.5)
with suitably chosen initial conditions. \((\tilde{c}_n(0))\) will be finitely supported on a resonant set \(\Lambda \subset D \setminus 0\) that is closed under completing rectangles in the sense that it satisfies Property II\(\Lambda\) from the previous section. Recall that this implies that \((\tilde{c}_n(t))\) remains supported on \(\Lambda\) for all later times. We will also impose the same conditions as before on the set \(\Lambda\), namely Properties II\(\Lambda\)-VI\(\Lambda\) as well as the norm explosion condition [27].

To be able to prove the closeness of \((\tilde{c}_n)\) and \((c_n)\) for \(n \in \Lambda\), we will impose a new condition that is used to limit the interaction of the frequencies in \(\Lambda\) with the 0-frequency mode \(c_0\).
• Property VIII\(\Lambda\) (No pair of modes from \(\Lambda\) forms a rectangle with 0): \(\Lambda\) contains at most one vertex of any rectangle having a vertex at 0, i.e. for any \(n_1, n_2 \in \Lambda\), neither one of the two parallelograms determined by the three vertices 0, \(n_1, n_2\) is a rectangle.

This property says that no resonant interactions can happen between two frequencies in \(\Lambda\) and the zero mode. Such interactions with the relatively large zero mode can leak energy out of \(\Lambda\) or inject energy into it, in a way that could destroy the delicate Arnold diffusion responsible for the energy cascade across the different generations of \(\Lambda\).

The existence of such a \(\Lambda\) is guaranteed by the following proposition:

**Proposition 3.1.** Let \(0 < s, s \neq 1\) and \(N \in \mathbb{N}\). Given parameters \(\delta \ll 1\) and \(K \gg 1\), there exists \(P = P(\delta, K)\) and a frequency set \(\Lambda = \Lambda_1 \cup \ldots \cup \Lambda_P \subseteq \mathbb{D} = N\mathbb{Z}^2\) satisfying Property II\(\Lambda\)-Property VIII\(\Lambda\) as well as the norm explosion condition (2.7). In particular, the cardinality of \(P\) can be chosen uniformly in \(N\).

**Proof.** This will be a special case of Proposition 4.7 in the next section (and the remark following its proof).

The growth of the Sobolev norm of \((\tilde{c}_n)\) will follow by choosing the initial data as we recalled in the previous section and using the diffusion result in Proposition 2.4. The key point is then to transfer this growth property of \((\tilde{c}_n)\) into a statement about the Sobolev norm growth of \((c_n)\). This will be a consequence of the following lemma, which constitutes the bulk of the proof:

**Lemma 3.2.** Fix \(0 < \sigma < 1\) and suppose that \(1 \ll B, B^{1+\sigma} \ll N, \text{ and } T \ll B^2 \log B\) (implicit constants are allowed to depend on \(\sigma\)). Suppose that \(\tilde{c}(t) = (\tilde{c}_n(t))\) is a solution to the resonant (3.5) satisfying:

- \(c_n(0) = 0\) unless \(n \in \Lambda\) for some finite set \(\Lambda\) satisfying properties I\(\Lambda\)-VIII\(\Lambda\) above.
- \(\|\tilde{c}\|_{L^1(\mathbb{Z}^2)} \leq B^{-1}\) for all \(0 \leq t \leq T\).

If \(c(t)\) is the solution to (\(\mathcal{F}\text{NLS}\)) with initial data \(c(0) = A + \tilde{c}(0)\) (i.e. \(c_0(0) = A\) and \(c_n(0) = \tilde{c}_n(0)\) for \(n \neq 0\)), then

\[
\|c(t) - \tilde{c}(t)\|_{L^1(\Lambda)} \lesssim A \frac{B^{1+\sigma}}{N^2}
\]

for all \(0 \leq t \leq T\).

3.3. **Proof of Theorem 1.5 assuming Lemma 3.2.** Suppose that \(0 < s < 1, \delta \ll 1, \text{ and } K \gg 1\) are given. Choose \(0 < \sigma < 1\) such that \(s(1 + \sigma) < 1\) and let \(P = P(\delta, K)\) and \(\Lambda = \Lambda_1 \cup \ldots \cup \Lambda_P\) be given by Proposition 3.1. Also let \(b(t) = (b_1(t), \ldots, b_P(t))\) be the solution to the "Toy system" (2.6) over a time interval \([0, T_0]\) given by Proposition 2.4 for some \(\epsilon = \epsilon(\delta, K)\) to be specified later. Note that the "Toy system" (2.6) satisfies the following scaling symmetry:

\[
b^\lambda(t) = \frac{1}{\lambda} b\left(\frac{t}{\lambda^2}\right)
\]

also solves (2.6) on the interval \([0, \lambda^2 T_0]\). Let \(\tilde{c}(t) = (\tilde{c}_n(t))\) be the solution of the resonant system obtained from \(b^\lambda(t)\) (i.e. \(\tilde{c}_n(t) = b^\lambda_n(t)\) if \(n \in \Lambda_j\) for \(1 \leq j \leq P\) and \(\tilde{c}_n(t) = 0\) otherwise).
Also, let \( c(t) = (c_n(t)) \) be the solution to (FNLs) with initial data \( c(0) = A + \tilde{c}(0) \) (i.e. \( c_0(0) = A, c_n(0) = \tilde{c}_n(0) \) for \( n \neq 0 \)). Since \( \|b(t)\|_{l^\infty} \sim 1 \), we have that \( \|\tilde{c}(t)\|_{l^\infty} \sim \lambda^{-1} \) for all \( 0 \leq t \leq \lambda^2 T_0 \).

It will be convenient to adopt the following norm:

\[
\|(a_n)\|_{H^s(\mathbb{Z}^2)} := \left( \sum_{n \in \mathbb{Z}^2} \langle n \rangle^{2s} |a_n|^2 \right)^{1/2} \sim \left\| \sum_{n \in \mathbb{Z}^2} a_n e^{inx} \right\|_{H^s(T^2)}. \tag{3.8}
\]

We will often abuse notation and refer to it as the Sobolev norm of \( (a_n) \).

The idea is to verify that 1) the Sobolev norm of \( \tilde{c}(t) \) grows from \( \delta \) to \( K \) in the time interval \([0, \lambda^2 T_0]\) and 2) the distance in \( H^s \) between \( c|_{\Lambda}(t) \) (i.e. restricting to the frequencies of \( c \) in \( \Lambda \)) and \( \tilde{c}(t) \) remains \( \leq 1 \) over this time interval. The latter is done using the stability lemma 3.2.

In the notation of Lemma 3.2, \( B \) will be taken to be \( C(P) \lambda \) (since \( \Lambda \) has \( C(P) \) elements and \( \tilde{c}(t) \lesssim \lambda^{-1} \)) and \( \sigma \) chosen as mentioned above so that \( s(1 + \sigma) < 1 \). The condition that \( 1 \ll B^{1+\sigma} \ll N \) will be guaranteed because \( \lambda \) will be chosen below to be \( \sim \delta_{\delta,K,P} N^s \) (and \( N^{s(1+\sigma)} \ll N \) ). Also, with \( \lambda \) large enough, one also guarantees that the stability interval \( B^2 \log B \) from the stability lemma 3.2 is longer than \( \lambda^2 T_0 \sim \delta_{\delta,K} B^2 \), which is the time needed for the Sobolev norm of \( (c_n(t)) \) to grow from \( \delta \) to \( K \).

We start by verifying that the Sobolev norm of \( \tilde{c}(t) = (\tilde{c}_n(t)) \) grows from \( \delta \) to \( K \) in the time interval \([0, \lambda^2 T_0]\). We will do this by first making sure that \( \|\tilde{c}(0)\|_{H^s(\mathbb{Z}^2)} \sim \delta \) (this will specify \( \lambda \) in terms of \( N \)) and then verifying that the ratio of Sobolev norms

\[
\frac{\|\tilde{c}(\lambda^2 T_0)\|_{H^s(\mathbb{Z}^2)}}{\|\tilde{c}(0)\|_{H^s(\mathbb{Z}^2)}} \gtrsim \frac{K}{\delta}. \tag{3.9}
\]

Notice that since

\[
\|\tilde{c}(0)\|_{H^s(\mathbb{Z}^2)} \sim \left( \sum_{n \in \Lambda} |n|^{2s} |\tilde{c}_n(0)|^2 \right)^{1/2} \sim_{P,\delta,K} N^s \lambda^{-1},
\]

one can make sure that \( \|\tilde{c}(0)\|_{H^s(\mathbb{Z}^2)} \sim \delta \) by choosing \( \lambda \sim_{P,\delta,K} N^s \). We remark that this implies also that \( \|c(0) - A\|_{H^s(\mathbb{Z}^2)} \sim \delta \). Moving to (3.9): Let \( Q \) be the ratio of the Sobolev norm of \( \tilde{c} \) at times \( \lambda^2 T_0 \) to that at time 0,

\[
Q \sim \frac{\sum_{n \in \Lambda} |n|^{2s} |\tilde{c}_n(\lambda^2 T_0)|^2}{\sum_{n \in \Lambda} |n|^{2s} |\tilde{c}_n(0)|^2} = \frac{\sum_{j=1}^P \sum_{n \in \Lambda_j} |n|^{2s} |b_j^n(\lambda^2 T_0)|^2}{\sum_{j=1}^{P'} \sum_{n \in \Lambda_j} |n|^{2s} |b_j^n(0)|^2}
\]

where we have used the fact that \( \tilde{c}(t) \) is supported on \( \Lambda \) and that it comes from the diffusion solution \( b(t) \). Adopting the notation \( S_j = \sum_{n \in \Lambda_j} |n|^{2s} \) and using the properties of \( b(t) \) given by proposition 2.3 we get that:
\[ Q \gtrsim_{\varepsilon} S_4 + S_5 + \ldots + S_P \]
\[ \gtrsim (1 - \varepsilon) \frac{S}{SP_{-2}} + O_{\delta,K}(\varepsilon) \]
\[ \gtrsim K^2 \]

by choosing \( \varepsilon \) small enough depending on \( P, \delta, K \).

Finally, we verify that \( \| c(\lambda^2 T_0)\|_{h^*(\mathbb{Z}^2)} \gtrsim K \) by showing that \( \| c|_{\Lambda}(\lambda^2 T_0)\|_{l^1(\Lambda)} \gtrsim K \). This follows from (3.7) as follows:

\[
\left( \sum_{n \in \Lambda} |n|^{2s} |c_n(\lambda^2 T_0) - \tilde{c}_n(\lambda^2 T_0)|^2 \right)^{1/2} \leq C(P) s^\|c(\lambda^2 T_0) - \tilde{c}(\lambda^2 T_0)|_{l^1(\Lambda)} \]
\[
\leq C(P, A) N^s \lambda^{1+\sigma} N^2 \leq C(P, A, \delta) \frac{N^{s(2+\sigma)}}{N^2} \leq 1 \text{ if } N \text{ is large enough,} \]

where we have used the fact that \( B = C(P) \lambda \) in the second inequality, that \( \lambda = C(P, \delta) N^s \) in the third, and that \( 2 > s(2+\sigma) \) in the last inequality.

\[ \square \]

3.4. **Proof of stability lemma** 3.2. Denote by \( D^* = D \setminus \{0\} \). We would like to show that \( (c_n(t)) \) remains close to \( (\tilde{c}_n(t)) \) in \( l^1(D^*) \) for a time interval \([0,T]\) under the assumption that \( \sum_{n \neq 0} |\tilde{c}_n(t)| \lesssim B^{-1} \) for some \( B \gg 1 \) and \( T \ll B^2 \log B \).

Denoting \( w_n = c_n - \tilde{c}_n \) for \( n \neq 0 \) we get the following difference equation:

\[
-i \partial_t w_n(t) = - |c_n(t)|^2 c_n(t) - |\tilde{c}_n(t)|^2 \tilde{c}_n(t) + \sum_{\Gamma_{res}(n) \cap \Lambda^3} \left( c_n c_{n_2} c_{n_3}(t) - \tilde{c}_n \tilde{c}_{n_2} \tilde{c}_{n_3}(t) \right) + c_0(t)^2 \tilde{c}_n(t) e^{-2i|n|^2 t} + 2 \left( \sum_{n_2 \neq 0, -n} c_{n_2} c_{n+n_2} e^{2i n_2 t} \right) c_0(t) + \left( \sum_{n_1 \neq 0, n} c_{n_1} c_{n-n_1}(t) e^{2i n_1 (n_1-n) t} \right) \tilde{c}_0(t) + \sum_{\tilde{\Gamma}(n) \setminus (\Gamma_{res}(n) \cap \Lambda^3)} c_{n_1} c_{n_2} c_{n_3} e^{i \omega_4 t} \]

Defining:
Lemma 3.5.\[A_n(t) = |c_n(t)|^2c_n(t) - |\tilde{c}_n(t)|^2\tilde{c}_n(t), \quad B_n(t) = \sum_{\Gamma_{res}(n) \cap \Lambda^3} \left( c_n \overline{c}_{n_2}(t) c_{n_3}(t) - \tilde{c}_n \overline{c}_{n_2}(t) \tilde{c}_{n_3}(t) \right)\]

\[C_n(t) = c_0(t)^2 \overline{c}_n(t)e^{-2|n|^2t}, \quad D_n(t) = \left( \sum_{n_2 \neq 0, -n} c_{n_2} \overline{c}_{n_2+n_2}e^{2i|n|^2t} \right) c_0(t)\]

\[E_n(t) = \left( \sum_{n_1 \neq 0, n} c_{n_1}(t) c_{n_1-n_1}(t) e^{2i|n|^2t} \right) c_0(t), \quad F_n(t) = \sum_{\Gamma(n) \setminus (\Gamma_{res}(n) \cap \Lambda^3)} c_{n_1} \overline{c}_{n_2} c_{n_3} e^{i\omega t}\]

we get that the equation satisfied by \((w_n(t))\) can be written in integral form as:

\[w_n(t) = i \int_0^t (A_n(t) + B_n(t) + C_n(t) + D_n(t) + E_n(t) + F_n(t)) \, dt \quad (3.11)\]

All the sums below are over \(D^* := D \setminus 0\) unless otherwise indicated and implicit constant are allowed to depend on \(A \sim 1\).

The proof follows a bootstrap argument: We will first adopt the following bootstrap hypotheses:

\[\sum_{n \neq 0} |w_n(t)| \lesssim B^{-1} \quad (3.12)\]

for all \(0 \leq t \leq T\) (and hence \(\sum_{n \in \Lambda} |c_n(t)| \lesssim B^{-1}\) over the same time interval). The goal will then be to prove that \(\sum_{n \neq 0} |w_n(t)|\) actually satisfies the stronger bound

\[\sum_{n \neq 0} |w_n(t)| \lesssim_A B^{1+\sigma} N^2 \ll B^{-1-\sigma} \quad (3.13)\]

on \([0, T]\) if \(T \ll B^2 \log B\).

Before we start estimating the contribution of the terms \(A_n, \ldots, F_n\), we will single out a useful identity that we will use repeatedly:

**Lemma 3.5.** For any \(0 \leq t \leq T\),

\[|A^2 - |c_0(t)|^2| \lesssim B^{-2} \quad (3.14)\]

In particular, \(|c_0(t)| \sim A \sim 1\) for all \(0 \leq t \leq T\).

**Proof.** This follows directly from the bootstrap hypothesis and the conservation of mass \(M[u(t)] = \sum_{n \in \mathcal{D}} |c_n(t)|^2\) which gives:

\[A^2 + \sum_{n \neq 0} |c_n(0)|^2 = M[u(0)] = M[u(t)] = |c_0(t)|^2 + \sum_{n \neq 0} |c_n(t)|^2\]
3.5.1. **Contribution of** $A_n$: Recall that

$$A_n(t) = |c_n(t)|^2 c_n(t) - |\bar{c}_n(t)|^2 \bar{c}_n(t) = |c_n(t)|^2 w_n(t) + c_n(t) w_n(t) \bar{c}_n(t) + |\bar{c}_n(t)|^2 w_n(t)$$

and hence:

$$\left| \sum_{n \neq 0} \int_0^t A_n(s) ds \right| \lesssim \sum_{n \neq 0} B^{-2} \int_0^t |w_n(s)| ds \quad (3.15)$$

3.5.2. **Contribution of** $B_n$: Recall that

$$B_n(t) = \sum_{\gamma \in \Lambda^3} \left( c_{n_1} \overline{c_{n_2}}(t) c_{n_3}(t) - \bar{c}_{n_1} \bar{c}_{n_2}(t) \bar{c}_{n_3}(t) \right)$$

Hence,

$$B_n(t) = \sum_{\gamma \in \Lambda^3} \left( w_{n_1} \bar{c}_{n_2}(t) c_{n_3}(t) + c_{n_1} w_{n_2}(t) c_{n_3}(t) + \bar{c}_{n_1} \bar{c}_{n_2}(t) w_{n_3}(t) \right)$$

and we also get:

$$\left| \sum_{n \neq 0} \int_0^t B_n(s) ds \right| \lesssim \sum_{n \neq 0} B^{-2} \int_0^t |w_n(s)| ds \quad (3.16)$$

3.5.3. **Contribution of** $C_n$: Recall that

$$C_n(t) = c_0(t) \overline{c_{n_{\perp}}(t)} e^{-2i|n|^2 t}$$

which gives upon integration by parts:

$$\int_0^t C_n(s) ds = \left. -\frac{1}{2i|n|^2} c_0(s) \overline{c_{n_{\perp}}(s)} e^{-2i|n|^2 s} \right|_0^t$$

$$\quad + \frac{1}{2i|n|^2} \int_0^t \left( 2c_0(s) \partial_t \overline{c_{n_{\perp}}(s)} + c_0(s) \partial_t c_{n_{\perp}}(s) \right) e^{-2i|n|^2 s} ds \quad (3.17)$$

Note that since $n \in D \setminus 0$, we have $|n| \geq N$ and hence the boundary terms (3.17) have the acceptable contribution of $O(B^{-1}N^{-2})$. Denote by:

$$I_n(t) = \int_0^t 2c_0(s) \partial_t \overline{c_{n_{\perp}}(s)} e^{-2i|n|^2 s} ds, \quad \text{and} \quad II_n(t) = \int_0^t c_0(s) \partial_t c_{n_{\perp}}(s) e^{-2i|n|^2 s} ds.$$

We start by estimating $I_n(t)$. From (3.18), we have

$$\partial_t c_0(t) = -i|c_0|^2 c_0(t) + O(AB^{-2}) = -iA^2 c_0(t) + i(A^2 - |c_0(t)|^2) c_0(t) + O(B^{-2})$$

$$= -iA^2 c_0(t) + O(AB^{-2}) \quad (3.19)$$
where we used (3.14) in the last equality. This leads to:

\[ |I_n(t)| \sim \int_0^t c_0(s) \partial_t c_n(s)e^{-2|n|^2 s} ds \]

\[ \lesssim |\int_0^t c_0(s)^2 e^{-2|n|^2 s} ds| + B^{-2} \int_0^t |c_n(s)| ds \]

\[ = |\int_0^t C_n(s) ds| + B^{-2} \int_0^t |c_n(s)| ds, \]

which gives that:

\[ \sum_{n \in D^n} \frac{1}{|n|^2} |I_n(t)| \lesssim \frac{1}{N^2} \sum_{n \in D^n} |\int_0^t C_n(s) ds| + B^{-2} \sum_{n \in D^n} \int_0^t |c_n(s)| ds \]  

(3.20)

To estimate \( II_n(t) \), we first use the expression of \( \partial_t c_n \) from (3.3) to get:

\[ II_n(t) = \int_0^t c_0(s)^2 \partial_t c_n(s)e^{-2|n|^2 s} ds \]  

(3.21)

\[ = i \int_0^t c_0(s)^2 |c_n(s)|^2 e^{-2|n|^2 s} ds \]  

(3.22)

\[ - i \int_0^t |c_0(s)|^4 c_n(s) ds \]  

(3.23)

\[ - 2i \int_0^t |c_0(s)|^2 c_0(s) \left( \sum_{n_2 \neq 0, n} c_{n_2} e^{-2in_2^2s} \right) e^{-2|n|^2 s} ds \]  

(3.24)

\[ - i \int_0^t c_0(s)^3 \left( \sum_{n_1 \neq 0, -n} c_{n_1} c_{-n_1} e^{-2in_1^2(n_1+n)s} \right) e^{-2|n|^2 s} ds \]  

(3.25)

\[ - i \int_0^t c_0(s)^2 \sum_{\Gamma(-n)} c_{n_2} c_{n_3} e^{-i\omega_4 s} e^{-2|n|^2 s} ds \]  

(3.26)

\[ = \Phi_1 + \Phi_2 + \Phi_3 + \Phi_4 + \Phi_5 \]  

(3.27)

where \( \Phi_1, \ldots, \Phi_5 \) are defined by (3.22), …, (3.26) respectively.

\[ \sum_{n \neq 0} |\Phi_1| \text{ and } \sum_{n \neq 0} |\Phi_5| \] can be easily estimated by \( \lesssim A B^{-2} \sum_{n \in D^n} \int_0^t |c_n(s)| ds \). Also, \( \sum_{n \neq 0} \Phi_2 \lesssim A \sum_{n \in D^n} \int_0^t |c_n(s)| ds \) and

\[ \sum_{n \neq 0} |\Phi_3|, \sum_{n \neq 0} |\Phi_4| \lesssim B^{-1} \sum_{n \neq 0} \int_0^t |c_n(s)| ds. \]

As a result, we get:

\[ \sum_{n \neq 0} |II_n(t)| \lesssim \int_0^t \sum_{n \neq 0} |c_n(s)| ds \]  

(3.28)

Remark. The term (3.22) gives the worst contribution in all our estimates and it is the reason why the result is restricted to the range \( 0 < s < 1 \). It is the result of a chain of two non-resonant interactions of the zero mode with the modes at \( n \) and \( -n \) respectively (via the term \( c_0^2 c_n e^{-2|n|^2 t} \))
that shows up in the equation of \(-i \partial_t c_{-n}\), and the term \(\frac{2}{c_0} e^{-2|n|^2 t}\) in the equation of \(-i \partial_t c_n\). The net effect of this interaction chain is a *second-order resonant* self-excitation of the mode at \(n\).

We are now in a good position to sum the contribution of \(\int_0^t C_n(t)\). Using (3.17), (3.20), and (3.28), and absorbing the first term of (3.20) in the left-hand side we get:

\[
\sum_{n \neq 0} \left| \int_0^t C_n(t) dt \right| \lesssim \frac{1}{N^2} \int_0^t \sum_{n \neq 0} |c_n(s)| ds + \frac{B^{-1} + B^{-1} t}{N^2} \tag{3.29}
\]

### 3.5.4. Contribution of \(D_n\): \(D_n\) involves resonant and non-resonant interactions, which we will separate as follows:

\[
D_n(t) = \left( \sum_{n_2 \neq 0, -n} \overline{c_{n_2 + n_2}} e^{2i n_2 t} \right) c_0(t)
\]

\[
= \left( \sum_{n_2 \neq 0, -n} \overline{c_{n_2 + n_2}} \right) c_0(t) + \left( \sum_{n_2 \neq 0, -n} \overline{c_{n_2 + n_2}} e^{2i n_2 t} \right) c_0(t)
\]

\[
= D_n(t) + \tilde{D}_n(t)
\]

We start by estimating \(\tilde{D}_n(t)\). For this we write first:

\[
c_0(t) = \frac{1}{A^2} A^2 c_0(t) = \frac{1}{A^2} |c_0(t)|^2 c_0(t) + (A^2 - |c_0(t)|^2) c_0(t)
\]

\[
= \frac{1}{A^2} (i \partial_t c_0(t)) + O(AB^{-2}) + (A^2 - |c_0(t)|^2) c_0(t)
\]

\[
= \frac{1}{A^2} (i \partial_t c_0(t)) + O(AB^{-2})
\]

where we have used (3.3) and (3.11). As a result, we get:

\[
\int_0^t \tilde{D}_n(s) ds = \sum_{n_2 \neq 0, -n} \int_0^t c_{n_2}(a) c_{n+n_2}(a) \left( \frac{1}{A^2} (i \partial_t c_0(s)) + O(AB^{-2}) \right) ds
\]

\[
= \frac{i}{A^2} \sum_{n_2 \neq 0, -n} \left( \overline{c_{n_2}(s)} c_{n+n_2}(s) c_0(s) \right) \Bigg|_0^t
\]

\[
- \int_0^t \left( \partial_t \overline{c_{n_2}(s)} c_{n+n_2}(s) c_0(s) + \overline{c_{n_2}(s)} \partial_t c_{n+n_2}(s) c_0(s) \right) ds
\]

\[
+ O \left( B^{-2} \sum_{n_2 \neq 0, -n} \int_0^t |c_{n_2}(s)||c_{n+n_2}(s)| ds \right)
\]

The contribution of the boundary terms above can be readily estimated in \(l^1(D^*)\) by \(\lesssim B^{-2}\). For the first integral, we again use equation (3.4) for \(\partial_t c_{n_2}\) and \(\partial_t c_{n-n_2}\) to get (we do the calculation for \(\partial_t c_{n_2}\), that of \(\partial_t c_{n-n_2}\) is similar):
Using (3.3) and (3.4) to estimate the case when \( n + n_2 \notin \Lambda \): if \( n + n_2 \notin \Lambda \), then \( c_{n+n_2} = w_{n+n_2} \) and we easily estimate:

\[
\left| \sum_{n_2 \neq 0} \int_0^t \partial_s c_{n+n_2}(s)c_{n+n_2}(s)c_0(s)ds \right| \lesssim B^{-2} \int_0^t \sum_{n \neq 0} |w_n(s)|ds \tag{3.31}
\]

where we integrated by parts in \( \Xi_1 \) (to gain a factor of \( N^{-2} \ll B^{-2} \)) and used direct estimates for \( \Xi_2, \ldots, \Xi_5 \).

The case when \( n + n_2 \in \Lambda \) (\( \Rightarrow n_2 \notin \Lambda \)) is slightly more tedious to bound and we need to consider the different contributions of \( \Xi_1, \ldots, \Xi_5 \). We start by estimating \( \Xi_1 \): Integrating by parts and using (3.34) and (3.31) to estimate \( \partial_s c_0 = O(1) \) and \( \sum_{n \in D} |\partial_s c_n| \lesssim B^{-1} \), we arrive at the estimate:

\[
\sum_{n, n_2 \neq 0} |\Xi_1| \lesssim \frac{B^{-1}}{N^2} \int_0^t \sum_{n \neq 0} |c_n(s)|ds + \frac{B^{-2}}{N^2} \lesssim \frac{B^{-1}}{N^2} \int_0^t \sum_{n \neq 0} |w_n(s)|ds + \frac{B^{-2}(1 + t)}{N^2}. \tag{3.32}
\]
Obviously, since \( n_2 \notin \Lambda \), we readily have the estimate:
\[
\sum_{n,n_2 \neq 0, n.n_2 = 0} |\Xi_2| \lesssim B^{-3} \int_0^t \sum_{n \neq 0} |w_n(s)| ds
\]  
(3.33)

As for \( \Xi_3 \), we split the sum over \( l \) into the resonant contribution when \( l.n_2 = 0 \) and the non-resonant contribution when \( l.n_2 \neq 0 \). For the resonant part, we use Property VIII\( \Lambda \) to conclude that either \( l \) or \( n_2 + l \notin \Lambda \) and hence we estimate the contribution of this part by:
\[
\lesssim B^{-2} \int_0^t \sum_{n \neq 0} |w_n(s)| ds
\]

As for the non-resonant part of \( \Xi_3 \) when \( n_2.l \neq 0 \), we integrate by parts to gain a factor of \( n_2.l \) in the denominator (recall that since \( n_2.l \in D, |n_2.l| \geq N^2 \)) and estimate the resulting contribution using (3.3) and (3.4) by:
\[
\sum_{n,n_2 \neq 0, n.n_2 = 0} |\Xi_3| \lesssim B^{-3} \int_0^t \sum_{n \neq 0} |w_n(s)| ds
\]
(3.34)

The estimate on \( \Xi_4 \) is similar to that on \( \Xi_3 \) and we won’t repeat the calculation. Finally, to estimate the contribution of \( \Xi_5 \), we notice that if one of \( l_1,l_2,l_3 \notin \Lambda \), we directly have the estimate:
\[
\sum_{n,n_2 \neq 0, n.n_2 = 0} |\Xi_5| \lesssim B^{-3} \int_0^t \sum_{n \neq 0} |w_n(s)| ds
\]

Otherwise, \( l_1,l_2,l_3 \in \Lambda \) and as a result of Property II\( \Lambda \) and the fact that \( n_2 \notin \Lambda \), we get that \( w_4 \neq 0 \) which implies that \( |\omega_4| > N^2 \). Integrating by parts and estimating as before we arrive at the estimate:
\[
\sum_{n,n_2 \neq 0, n.n_2 = 0} |\Xi_5| \lesssim B^{-3} \int_0^t \sum_{n \neq 0} |w_n(s)| ds + \frac{B^{-3}(1 + t)}{N^2}
\]  
(3.35)

Combining the estimates on \( \Xi_1 \ldots \Xi_5 \) in (3.31)–(3.35) we get the following estimate on \( \int_0^t \tilde{D}_n(t) dt \):
\[
\sum_{n \neq 0} |\int_0^t \tilde{D}_n(s) ds| \lesssim B^{-2} \int_0^t \sum_{n \neq 0} |w_n(s)| ds + \frac{B^{-2}(1 + t)}{N^2}
\]  
(3.36)

where we used the fact that \( B \lesssim N \).

To finish estimating the contribution of \( D_n \), we are left with its non-resonant part \( D_n(t) \) of \( D_n(t) \) given by:
\[
D_n(t) = \sum_{n_2.n \neq 0, n_2 \neq -n} c_{n_2} c_{n+n_2} e^{2in.n_2} c_0(t)
\]

Integrating by parts and estimating as before we get:
\[
\sum_{n \neq 0} \left| \int_0^t D_n(s) ds \right| \lesssim \frac{B^{-2}}{N^2} + \frac{B^{-1}}{N^2} \sum_{n \neq 0} \int_0^t |c_n(s)| ds
\]  
(3.37)

and consequently, we finally obtain:
\[
\sum_{n \neq 0} \left| \int_0^t D_n(s) ds \right| \lesssim B^{-2} \int_0^t \sum_{n \neq 0} |w_n(s)| ds + \frac{B^{-2}(1 + t)}{N^2}
\]  
(3.38)

3.5.5. **Contribution of \( E_n(t) \):** Estimating the contribution of \( E_n(t) \) is similar to that of \( D_n(t) \) and we won’t repeat the argument that gives:
\[
\sum_{n \neq 0} \left| \int_0^t E_n(s) ds \right| \lesssim B^{-2} \int_0^t \sum_{n \neq 0} |w_n(s)| ds + \frac{B^{-2}(1 + t)}{N^2}
\]  
(3.39)

3.5.6. **Contribution of \( F_n(t) \):** Recall that
\[
F_n(t) = \sum_{\Gamma(n) \cap \Lambda^3} c_{n_1} c_{n_2} c_{n_3} e^{i\omega t}
\]
Here the main concern are the terms for which \( n_1, n_2, n_3 \in \Lambda \). To deal with those we integrate by parts in the integral of \( F_n \) and use the fact that when \( w_4 \neq 0 \) then \( |w_4| \geq N^2 \). Noting also that \( \|\partial_t c_n\|_{L^2(D^*)} = O(B^{-1}) \) and using the bootstrap assumption we get:
\[
\sum_{n \neq 0} \left| \int_0^t \left( \sum_{\Gamma(n) \cap \Lambda^3 \cap \{\Gamma(v_{\gamma}(n) \cap \Lambda^3)\}} c_{n_1} c_{n_2} c_{n_3} e^{i\omega t}s \right) ds \right| \lesssim \frac{B^{-3}(1 + t)}{N^2} + \frac{B^{-2}}{N^2} \int_0^t \sum_{n \neq 0} |w_n(s)| ds
\]
As for the other terms of \( F_n(t) \), we will always have at least one of \( n_1, n_2, n_3 \) not in \( \Lambda \) and hence this sum can be estimated directly by \( B^{-2} \int_0^t |w_n(s)| ds \). Combining this with the above yields the estimate:
\[
\sum_{n \neq 0} \left| \int_0^t F_n(s) ds \right| \lesssim \frac{B^{-3}(1 + t)}{N^2} + B^{-2} \sum_{n \neq 0} \int_0^t |w_n(s)| ds
\]  
(3.40)

3.5.7. **End of proof:** Summing the contributions of \( A_n, \ldots, F_n \) in (3.15), (3.16), (3.29), (3.38), (3.39), (3.40), we get that:
\[
\sum_{n \neq 0} |w_n(s)| ds \lesssim B^{-2} \int_0^t \sum_{n \neq 0} |w_n(s)| ds + \frac{B^{-1}(1 + t)}{N^2}
\]
where we used the fact that \( B \lesssim N \). Gronwall’s inequality now gives:
\[
\sum_{n \neq 0} |w_n| \lesssim \left( \frac{B^{-1}(1 + t)}{N^2} \right) e^{CB^{-2}t}
\]  
(3.41)

and so for \( 0 \leq t \leq T \ll \sigma \sqrt{B^2 \log B} \), we have:
\[
\sum_{n \neq 0} |w_n| \lesssim \left( \frac{B^{-1}B^{2}(1 + t)}{N^2} \right) \ll \frac{B^{1+\sigma}}{N^2} \ll B^{-1-\sigma}
\]  
(3.42)
for $B$ large enough and $N \gg B^{1+\sigma}$.

\[\square\]

4. Unbounded orbits for almost-polynomial nonlinearities

In this section, we consider the equation:

\[-i\partial_t u + \Delta u = N_R(u)\]
\[u(0, x) = u_0(x)\]

(4.1)

with

\[N_R(u) = \sum_{n \in \mathbb{Z}^d} \left( \sum_{\Gamma_R(n)} \hat{u}(n_1) \overline{u(n_2)} \hat{u}(n_3) \right) e^{2\pi in.x}\]

(4.2)

and $\Gamma_R(n) = \{(n_1, n_2, n_3) \in (\mathbb{Z}^d)^3 : n_1 - n_2 + n_3 = n \}$ and $\omega_4 := |n_1|^2 - |n_2|^2 + |n_3|^2 - |n|^2 \in [-R, R]$.

This is obtained from the cubic nonlinearity (which corresponds to the case $R = \infty$) by deleting highly non-resonant interactions for which $|\omega_4| > R$.

We will prove that for any $R < \infty$, (4.1) admits unbounded orbits by following the program suggested by Proposition 1.3. Our results will be valid for all Sobolev spaces $H^s(\mathbb{T}^2)$ with $s > 0, s \neq 1$.

We start by making some remarks about the nonlinearity $N_R(u)$. We first note that for any $R$, one can easily verify that the flow defined by (4.1) conserves the mass:

\[M[u(t)] := \int_{\mathbb{T}^2} |u(t, x)|^2 dx.\]

Another conserved quantity is the Hamiltonian:

\[H_R(u) = \sum_n |n|^2 |\hat{u}(n)|^2 + \frac{1}{2} \sum_{n_1-n_2+n_3-n_4=0, |n_1|^2-|n_2|^2+|n_3|^2-|n_4|^2 \leq R} \hat{u}(n_1) \overline{u(n_2)} \hat{u}(n_3) \overline{u(n_4)}.\]

(4.3)

Notice that when $R = 0$, this is the Hamiltonian of the resonant system (or $(\mathcal{R}\mathcal{F}\mathcal{N}\mathcal{L}\mathcal{S})$) from the previous sections, and in this case $H_{R=0}(u)$ becomes coercive, in the sense that it controls the $H^1$ norm of the solution. Indeed, the second term in (4.3) can be written as:

\[\sum_{n \in \mathbb{Z}^2, k \in \mathbb{Z}} \left| \hat{u}(n_1) \hat{u}(n_3) \right|^{2} \geq 0\]

(4.4)

Remark. (Coercivity of $H_R$ and global existence) The local well-posedness theory of (4.1) is exactly the same as that of the polynomial cubic nonlinearity $|u|^3u$ in (2.1). This was proved in 22 (chapter 10.1) by adapting Bourgain’s proof for the polynomial cubic nonlinearity in 1. Indeed, just like the cubic nonlinearity $|u|^3u$, (4.1) is locally well-posed in $H^s(\mathbb{T}^2)$ for all $s > 0$ and the time of existence depends only on $\|u_0\|_{H^s(\mathbb{T}^2)}$ for $0 < s < 1$ and on $\|u_0\|_{H^s(\mathbb{T}^2)}$ for $s \geq 1$.

\[\text{In the range } 0 < s < 1, \text{ the result will be conditional on small-data global well-posedness, which can be proved in some cases at least when the mass } M \text{ is below a certain threshold (cf. following remark).} \]
When $R = 0$ or $\infty$, one can directly combine this local well-posedness result with the a priori control on the $H^1$ norm (given by the conservation and coercivity of the Hamiltonian $H_R$) to obtain global well-posedness of (4.1) for all $s \geq 1$. When $R \neq 0$, one can still extend the local well-posedness in $H^s(\mathbb{T}^2)$ for $s \geq 1$ into a global existence statement at least when the mass $M$ is smaller than some universal constant $C$. In fact, we have:

$$\frac{1}{2} \sum_{n_1-n_2+n_3-n_4=0} |\tilde{u}(n_1)|^2 - |\tilde{u}(n_2)|^2 + |\tilde{u}(n_3)|^2 - |\tilde{u}(n_4)|^2 \leq \frac{1}{2} \sum_{n_1-n_2+n_3-n_4=0} |\tilde{u}(n_1)|^2 - |\tilde{u}(n_2)|^2 + |\tilde{u}(n_3)|^2 - |\tilde{u}(n_4)|^2 \leq \frac{1}{2} C_{GN} \left( \sum_n |\tilde{u}(n)|^2 \right)^2 \left( \sum_n |n|^2 |\tilde{u}(n)|^2 \right)^2$$

where we applied the Gagliardo-Nirenberg inequality in the last step and denoted by $C_{GN}$ the optimal constant of this inequality. So if $M[u(0)] < \left( \frac{2}{C_{GN}} \right)^{1/2}$, the Hamiltonian $H_R(u)$ becomes coercive (uniformly in $R$), which gives (along with the conservation of mass) the needed a priori control on the $H^1$ norm. For concreteness, we summarize the contents of this remark in the following lemma, whose full proof can be found in [22] (chapter 10.1).

**Lemma 4.1.**

(i) For any $R \in 0 \cup \mathbb{N} \cup \{+\infty\}$, equation (4.1) is locally well-posed in $H^s(\mathbb{T}^2)$ for any $s > 0$. Moreover, the length of the local existence interval depends on $\|u_0\|_{H^s(\mathbb{T}^2)}$ when $0 < s < 1$ and only on $\|u_0\|_{H^1(\mathbb{T}^2)}$ when $s \geq 1$.

(ii) When $R = 0$ or $+\infty$, the equation is globally well-posed in $H^s(\mathbb{T}^2)$ for all $s \geq 1$.

(iii) There exists a universal constant $C^* := \left( \frac{2}{C_{GN}} \right)^{1/2}$, such that if $M[u_0] \leq C^*$, then the local solution in part i) exists globally in $H^s(\mathbb{T}^2)$ for any $s \geq 1$.

Finally, we end the remark by noting that minor modifications to the arguments in [20, 21] (where the case $R = \infty$ is considered) can extend the above global well-posedness statements in parts ii) and iii) above to the range $s > 2/3$. \[\square\]

Using the same reductions as in Section 2.1, we get that equation (4.1) is equivalent to the following system of ODE:

$$-i\partial_t a_n = -a_n|a_n|^2 + \sum_{n_1,n_2,n_3 \in \Gamma^R(n)} a_{n_1}a_{n_2}a_{n_3}e^{i\omega_4 t}$$

(4.5)

where $\Gamma^R(n) = \{(n_1,n_2,n_3) \in (\mathbb{Z}^2)^3 : n_1 - n_2 + n_3 = n \text{ and } n_1,n_3 \neq n : |\omega_4| \leq R\}$. Recall that at any time $t$, one can recover the solution $u$ of (4.1) from $v(t,x) = \sum_{n \in \mathbb{Z}^2} a_n e^{i(nx + |n|^2 t)}$, and vice-versa, and the two solutions have the same Sobolev norms.

**4.2. Proof of Theorem 1.6.** The non-self interactions that appear in system (4.5) happen in parallelograms with vertices $n, n_1, n_2, n_3$ satisfying $n_1, n_3 \neq n$ (which guarantees that the parallelogram does not degenerate to a point) and $|\omega_4| \leq R$. In what follows it will often be convenient to specify the arguments of the function $\omega_4$, so we precise:

$$\omega_4(n_1,n_2,n_3,n_4) := |n_1|^2 - |n_2|^2 + |n_3|^2 - |n_4|^2,$$

(4.6)

and remark that when $n_4 = n_1 - n_2 + n_3$, $\omega_4$ can be understood as a measure of an “eccentricity” of the parallelogram with vertices $n_1, n_2, n_3, n_4$ (in the sense that the smaller $|\omega_4|$ is, the closer is the parallelogram to being a rectangle).
Our initial data to (4.5) will be chosen to be supported on a set $S$ satisfying the following closure property:

**R–closure condition:** If $n_1, n_2, n_3 \in S$ are three vertices of a parallelogram with $n_4 := n_1 - n_2 + n_3$ being the forth vertex, and if $\omega_4 = |n_1|^2 - |n_2|^2 + |n_3|^2 - |n_4|^2 \in [-R, R]$, then $n_4 \in S$.

This condition is simply the generalization of the closure condition from Section 2.3, and it guarantees that the emanating solution $a_n(t)$ remains supported on $S$ for all later times. We include this in the following lemma:

**Lemma 4.3.** Suppose that the initial data $a_n(0) \in l^1(\mathbb{Z}^2)$ is supported in a set $S$ satisfying the R–closure condition above. If $a_n(t) \in L^\infty_{t,\text{loc}}(I; l^1(\mathbb{Z}^2))$ solves (4.5) on some time interval $I \ni 0$, then $a_n(t)$ is also supported on $S$, i.e. $a_n(t) = 0$ unless $n \in S$.

**Proof.** Let $B(t) = \sum_{n \notin S} |a_n|^2$, then

$$B'(t) = 2R \sum_{n \notin S} \left( |a_n|^4 + \sum_{n_1, n_2, n_3 \in \Gamma^R(n)} \overline{a_{n_1}}a_{n_2}a_{n_3}e^{i\omega_4 t} \right)$$

By the closure property of $S$ either $n_1, n_2,$ or $n_3 \notin S$ if $n \notin S$. This gives the bound:

$$B'(t) \leq CB(t)$$

for all $t$ in an arbitrary compact subinterval of $I$. Since $B(0) = 0$, we get the result by Gronwall’s inequality.

The plan is to apply Proposition 1.3 (ii) to conclude the existence of unbounded Sobolev orbits of (4.5). For this, we will prove long-time strong instability of the flow near all initial data whose support satisfies the $R$–closure condition. This is the content of the following main proposition:

**Proposition 4.4.** Let $s > 0, s \neq 1$. Suppose that $S$ is a finite subset of $\mathbb{Z}^2$ satisfying the $R$–closure condition and $\phi \in H^s(\mathbb{T}^2)$ is supported in Fourier space on $S$ (i.e. $\hat{\phi}(n) = 0$ unless $n \in S$). For any $\delta \ll 1, K \gg 1$, there exists a solution $(c_n(t))$ of (4.5) and a time $T > 0$ such that:

$$\| \sum_{n \in \mathbb{Z}^2} c_n(0)e^{in.x} - \phi \|_{H^s(\mathbb{T}^2)} \leq \delta \quad \text{and} \quad \| \sum_{n \in \mathbb{Z}^2} c_n(T)e^{in.x} \|_{H^s(\mathbb{T}^2)} \geq K.$$  \hspace{1cm} (4.7)

Moreover, $c_n(0)$ can be chosen to be supported on a finite set satisfying the $R$–closure condition.

---

\footnote{This condition is actually redundant as it is not hard to see that (4.5) is locally well-posed in $l^1(\mathbb{Z}^2)$.}
We delay the proof of this key proposition to the next section, but let us see how it gives a proof of Theorem 1.6. For \( s \geq 1 \), the flow map:

\[
X := \{ u_0 \in H^s : \| u_0 \|_{L^2} \leq C^s \} \rightarrow C^s_t H^s([0, T] \times \mathbb{T}^2)
\]

exists and gives a well-posed dynamical system on the closed set \( X \subset H^s(\mathbb{T}^2) \). Recall that this set is invariant under the flow thanks to the conservation of mass. As mentioned before, this restriction on the mass is included to guarantee global well-posedness of the flow and is not needed when \( R = 0 \).

Let \( D \subset X \) be the subset of all functions whose Fourier transform is supported on a finite set \( S \) satisfying the \( R \)-closure property. Define \( F := \overline{D} \) to be its closure in \( X \) (or equivalently in \( H^s \) since \( X \) is closed). The proof now proceeds exactly as that of Proposition 1.3 to conclude that, since \( F \) is a complete metric space, the set of unbounded orbits is non-empty and co-meager in \( F \).

Remark. Since Proposition 4.4 holds for all \( s > 0, s \neq 1 \), the above proof holds for \( 0 < s < 1 \) as long as we can prove that the flow map is globally well-posed on some open subset of \( H^s \) containing 0. As mentioned in the previous remark, this is the case for \( s > 2/3 \) by adapting the arguments in [21].

4.5. Proof of Proposition 4.4 The idea of the proof of Proposition 4.4 is as follows: Starting with an initial data \( b_n(0) = \hat{\phi}(n) \) supported on a finite set \( S \) satisfying the \( R \)-closure condition, the solution \( (b_n(t)) \) to (4.3) remains supported on the set \( S \) for all time. One another hand, one can construct a norm exploding solution \( (a_n(t)) \) to (4.3) that is supported on a disjoint set \( \Lambda \) also satisfying the \( R \)-closure condition, in addition to a combinatorial condition defined in terms of how it “interacts” with \( S \) (See Lemma 4.6). The set \( \Lambda \) will also have to satisfy a list of conditions similar to those in Section 4.5 and its construction is one of the main components of the proof. The careful choice of \( \Lambda \) will allow us to “paste” the two solutions \( (a_n(t)) \) and \( (b_n(t)) \) to obtain a solution \( (c_n(t)) \) satisfying (4.7). We start with specifying the above-mentioned geometric/combinatorial condition on two sets \( S_1 \) and \( S_2 \) of \( \mathbb{Z}^2 \) that allows one to paste solutions supported on \( S_1 \) and \( S_2 \).

Lemma 4.6 (Pasting lemma). Let \( S_1 \) and \( S_2 \) be two finite disjoint subsets of \( \mathbb{Z}^2 \) satisfying the \( R \)-closure condition. Suppose that \( S_1 \) and \( S_2 \) are not connected by any parallelogram with eccentricity \( \leq R \) in the following sense: if \( n_1, n_2 \in S_1, n_3 \in S_2 \), and \( n \) is the forth vertex of the parallelogram formed by \( n_1, n_2, \) and \( n_3 \), then:

\[
\begin{align*}
|\omega_4(n_1, n_2, n_3, n)| & \geq R & \text{if } n &= n_1 - n_2 + n_3, \\
|\omega_4(n_1, n_3, n_2, n)| & \geq R & \text{if } n &= n_1 - n_3 + n_2.
\end{align*}
\]

(4.8)

and vice-versa with \( S_1 \) and \( S_2 \) interchanged.

If \( (a_n(t)) \) and \( (b_n(t)) \) are solutions to (4.3) with \( (a_n(0)) \) and \( (b_n(0)) \) supported in \( S_1 \) and \( S_2 \) respectively, then \( (a_n(t) + b_n(t)) \) also solves (4.3).

Remark. The two conditions in (4.8) corresponds to the two ways to form a parallelogram out of the three points \( n_1, n_2, \) and \( n_3 \).
Proof. The proof is by direct verification. Since \( a_n(t) \) and \( b_n(t) \) satisfy finite dimensional systems of ODE and since (4.5) conserves the \( L^2 \) norms \( \sum_n |a_n(t)|^2 \) and \( \sum_n |b_n(t)|^2 \), the two solutions exist globally in time. Let \( c_n(t) = a_n(t) + b_n(t) \). Since \( S_1 \cup S_2 \) clearly satisfies the \( R \)-closure condition, we directly get that \( c_n(t) = 0 \) unless \( n \in S_1 \cup S_2 \). Say \( n \in S_1 \), then

\[
-i\partial_t c_n = -i\partial_t a_n = -a_n |a_n|^2 + \sum_{n_1, n_2, n_3 \in \Gamma^R(n)} a_{n_1} a_{n_2} a_{n_3} e^{i\omega_4 t}
\]

\[
= -c_n |c_n|^2 + \sum_{n_1, n_2, n_3 \in \Gamma^R(n)} c_{n_1} c_{n_2} c_{n_3} e^{i\omega_4 t}
\]

because the last summand would vanish unless \( n_1, n_2, n_3 \in S_1 \cup S_2 \) and in this case we must have that \( n_1, n_2, n_3 \in S_1 \) since \( n \in S_1 \). \( \square \)

As a result of the above lemma, the key point is to construct a solution \((a_n(t))\) supported on frequency set \( \Lambda \subset \mathbb{Z}^2 \) so that \( \Lambda \) and \( S \) are not connected by parallelograms with “eccentricity” \( |\omega_4| \leq R \) as defined in (4.8) of the pasting lemma above. Adopting the useful notation in (3.8), we will also make sure that:

\[
\|(a_n(0))\|_{h^\ast(\mathbb{Z}^2)} \leq \delta \tag{4.9}
\]

and at some later time \( T \):

\[
\|(a_n(t))\|_{h^\ast(\mathbb{Z}^2)} \geq K. \tag{4.10}
\]

By the pasting lemma, if \((b_n(t))\) solves (4.5) with initial data \((\hat{\phi}(n))\), then \((c_n = a_n(t) + b_n(t))\) solves (4.5) with initial data \((a_n(0) + \hat{\phi}(n))\). Clearly, this solution proves Theorem 4.4 since

\[
\|c_n(0) - \hat{\phi}(n)\|_{h^\ast(\mathbb{Z}^2)} = \|(a_n)\|_{h^\ast} \leq \delta
\]

and

\[
\|c_n(T)\|_{h^\ast(\mathbb{Z}^2)} \geq \|(a_n(T))\|_{h^\ast(\mathbb{Z}^2)} \geq K.
\]

4.6.1. Construction of \( \Lambda \). We start by compiling all the properties that we would like \( \Lambda \) to satisfy:

- **Property I\( \Lambda \)** (Initial data): The initial data \( a_n(0) \) is supported on \( \Lambda \), i.e. \( a_n(0) = 0 \) unless \( n \in \Lambda \).
- **Property II\( \Lambda \)** (\( R \)-disjointness with \( S \)): \( \Lambda \) and \( S \) are not connected by any parallelogram with eccentricity \( \leq R \) in the sense that if \( n_1, n_2 \in S, n_3 \in \Lambda \) and \( n \) is the forth vertex of the parallelogram formed by \( n_1, n_2, \) and \( n_3 \), then (4.8) holds; and vice-versa with \( \Lambda \) and \( S \) interchanged.

We would also like \( \Lambda \) to satisfy the \( R \)-closure property so we wouldn’t need to deal with interactions between \( \Lambda \) and the outside world.

- **Property III\( \Lambda \)** (Closure): If \( n_1, n_2, n_3 \in \Lambda \) are three vertices of a parallelogram for which:

\[
|\omega_4| = |n_1|^2 - |n_2|^2 + |n_3|^2 - |n|^2 \leq R,
\]

then the forth vertex \( n = n_1 - n_2 + n_3 \) is also in \( \Lambda \). Equivalently, if \( (n_1, n_2, n_3) \in \Gamma^R(n) \cap \Lambda^3 \), then \( n \in \Lambda \). Now we simplify the system (4.5) by stipulating that all parallelograms in \( \Lambda \) with \( |\omega_4| \leq R \) are actually rectangles:
• Property IVΛ (Rectangular structure): If \( n_1, n_2, n_3, n \in \Lambda \) are four vertices of a parallelogram with \( n_1 + n_3 = n + n_2 \) and \( |\omega_4| \leq R \), then this parallelogram is actually a rectangle with \( \omega_4 = 0 \).

This Property reduces system (4.11) for \( (a_n(t)) \) into the resonant system (2.5) that we recall here for convenience:

\[
- i \partial_t a_n = -a_n |a_n|^2 + \sum_{n_1, n_2, n_3 \in \Gamma_{res}(n)} a_{n_1} \overline{a_{n_2}} a_{n_3} \tag{4.11}
\]

Note that this is a finite system of ODE on \( \Lambda \) since \( a_n(t) = 0 \) unless \( n \in \Lambda \).

We now assume as in Section 2.3 that \( \Lambda \) splits into a disjoint union \( \Lambda = \Lambda_1 \cup \ldots \cup \Lambda_P \) for some integer \( P > 1 \), and we call \( \Lambda_j \) the \( j \)-th generation of \( \Lambda \). A nuclear family in \( \Lambda \) is defined as before as a rectangle \( (n_1, n_2, n_3, n_4) \) where the frequencies \( n_1, n_3 \) (called the “parents”) belong to the \( j \)-th generation \( \Lambda_j \) while the “children” \( n_2, n_4 \) belong to \( \Lambda_{j+1} \). Recall that if \( (n_1, n_2, n_3, n_4) \) is a nuclear family then so are the permutations \( (n_3, n_2, n_1, n_4), (n_1, n_4, n_3, n_2), \) and \( (n_3, n_4, n_1, n_2) \). In addition to the above properties we also impose the same rules of genealogy on the set \( \Lambda \) as in Section 2.3 namely:

• Property VΛ (Existence and uniqueness of spouse and children).
• Property VIΛ (Existence and uniqueness of parents and siblings).
• Property VIIΛ (Non-degeneracy).
• Property VIIIΛ (Faithfulness).

These are exactly Properties IIIΛ-VIΛ of Section 2.3. Note that by the closure property IVΛ, Property VIIIΛ implies that \( \Lambda \) contains no other parallelograms with \( |\omega_4| \leq R \) apart from nuclear families (which are genuine rectangles).

Assuming that such a set exists, the system (4.11) can now be written as:

\[
- i \partial_t a_n = -a_n |a_n|^2 + 2a_{n_{child} - 1} a_{n_{child} - 2} \overline{a_{n_{spouse}}} + 2a_{n_{parent} - 1} a_{n_{parent} - 2} \overline{a_{n_{sibling}}} \tag{4.12}
\]

where we used the fact that the only rectangles that contain \( n \) as a vertex are those coming from nuclear families where \( n \) is either a parent (second term on the RHS) or a child (third term on RHS). The constant factor 2 on the RHS comes from the trivial permutations of the nuclear families.

Finally, we reduce the resonant system furthermore by imposing intragenerational equality:

• Property IXΛ (Intragenerational equality): The function \( n \mapsto a_n(0) \) is constant on each generation \( \Lambda_j \). In other words, \( a_{n_1}(0) = a_{n_2}(0) \) if \( n_1, n_2 \in \Lambda_j \).

One can verify (either by a Gronwall or a bootstrap argument) that Property IXΛ is preserved by the flow of (4.11) in the sense that intragenerational equality remains true for all time if it is true initially. The reason for this is that every mode in generation \( \Lambda_j \) is excited in exactly the same way as other modes in generation \( \Lambda_j \) (cf. the RHS of (4.12)). Property IXΛ allows to collapse system (4.11), which is a system in \( (a_n(t)) \) with \( n \in \Lambda = \Lambda_1 \cup \ldots \cup \Lambda_P \), into a smaller “Toy system” of ODE in \( b_j(t) \) with \( 1 \leq j \leq N \) by setting \( a_n(t) = b_j(t) \) whenever \( n \in \Lambda_j \). The system satisfied by \( b_j \) is given by:
\[-i \partial_t b_j = -b_j |b_j|^2 + 2b_{j+1}^2 b_j + 2b_{j-1}^2 b_j\]  

with the convention that \(b_0(t) = b_{P+1}(t) = 0\).

The existence of the frequency set \(\Lambda\) is contained in the following proposition:

**Proposition 4.7.** Let \(0 < s, s \neq 1\). Given parameters \(\delta \ll 1\) and \(K \gg 1\) and a finite set \(S\) satisfying the \(R\)–closure condition, there exists a frequency set \(\Lambda = \Lambda_1 \cup \ldots \cup \Lambda_P \subset \mathbb{Z}^2\) for some \(P = P(\delta, K) \gg 1\) satisfying Property II\(\Lambda\)–Property VIII\(\Lambda\) as well as:

\[
\frac{\sum_{n \in \Lambda_{P-2}} |n|^2}{\sum_{n \in \Lambda_1} |n|^2} \geq \frac{K^2}{\delta^2}
\]

In addition, for any \(N \geq C(\delta, K, S)\) we can make sure that \(\Lambda\) consists of \(P2^{P-1}\) frequencies \(n\) satisfying \(N \leq |n| \leq C(\delta, K, S) N\).

**Proof.** Given \(\delta \ll 1, K \gg 1\), we first construct a set \(\Lambda^0 = \Lambda^0_1 \cup \ldots \cup \Lambda^0_P\) satisfying the \(0\)–closure condition and Property V\(\Lambda\)–Property VIII\(\Lambda\), as well as the norm explosion condition \((4.14)\). This is exactly the set \(\Lambda^0\) constructed\(^\text{12}\) in \textbf{11}.

Choose \(v_0 \in \mathbb{Z}^2\) such that for any distinct \(n_1, n_2 \in \Lambda^0\) and any distinct \(k_1, k_2 \in S\) we have:

\[v_0, (n_1 - n_2) \neq 0 \quad \text{and} \quad v_0, (k_1 - k_2) \neq 0\]

By replacing \(R\) by \([R] + 1\), we may assume without any loss of generality that \(R \in \mathbb{N}\). We will take \(\Lambda = N\Lambda^0 - v\) for any \(N \geq N_0(K, \delta, S, R)\), where \(N_0(K, \delta, S, R)\) and \(v\) will be chosen later such that \(|v| \ll N_0\). Notice that choosing \(N_0 > R\) guarantees that if \(\Lambda^0\) satisfies the \(0\)–closure condition then \(N\Lambda^0\) satisfies the \(R\)–closure condition. Denoting

\[w_4(n_1, n_2, n_3, n) = |n_1|^2 - |n_2|^2 - |n_3|^2 - |n|^2\]

and noticing that \(\omega_3(n_1, \ldots, n) = \omega_3(n_1 - w, n_2 - w, n_3 - w, n - w)\) for any \(w \in \mathbb{Z}^2\), we get that \(\Lambda = N\Lambda^0 - v\) satisfies Property III\(\Lambda\). Property IV\(\Lambda\) is also obviously satisfied if \(N_0 > R\) by a similar reasoning. Properties V\(\Lambda\)–VIII\(\Lambda\) are trivially inherited by \(\Lambda\) from \(\Lambda^0\).

Also, since \(N\Lambda^0\) satisfies \((4.14)\), then so does \(\Lambda\) because \(v \ll N \leq |n|\) for any \(n \in N\Lambda^0\). Similarly, the last statement in proposition \(4.7\) follows easily.

As a result, all that remains is to verify Property II\(\Lambda\). This will specify \(v\). We will choose \(v = lRv_0\), where \(l \in \mathbb{N}\) and \(L \leq l \leq 2L\) for some \(C(\delta, K, S) \leq L \ll \frac{N_0}{N_0^2}\). The exact choice of \(l \in [L, 2L]\) will follow from the following counting argument. There are four possible cases to consider:

\(^{12}\)Strictly speaking, the construction in \textbf{11} is stated for \(s > 1\), but an inspection of the proof there (namely pages 102-103) shows that the same holds for \(0 < s < 1\) by relabeling the indices of the generations \(\Lambda_j\) as \(j \rightarrow P - j\). See \textbf{11} or Proposition 8.2.1 of \textbf{22} for details.
Case 1: Suppose that \((n_1, n_2, k) \in \Lambda \times \Lambda \times S\) and \(n = n_1 - n_2 + k\) form a parallelogram. Then \(n_1 + v, n_2 + v, k\) and \(n\) also form a parallelogram. Defining \(\tilde{n}_i \in \Lambda^0\) such that \(n_i + v = N\tilde{n}_i\) \((i = 1, 2)\) we have:

\[
\omega_4(n_1, n_2, k, n) = |n_1|^2 - |n_2|^2 + |k|^2 - |n|^2
\]
\[
= |n_1 + v|^2 - |n_2 + v|^2 + |k|^2 - |n|^2 - 2v(n_1 - n_2)
\]
\[
= \omega_4(n_1 + v, n_2 + v, k, n) - 2vN(\tilde{n}_1 - \tilde{n}_2)
\]

Note that since \((n_1 + v, n_2 + v, k) \in NA^0 \times NA^0 \times S\), there are \(O_{\Lambda^0, S}(1)\) possible values for \(\omega_4(n_1 + v, n_2 + v, k, n)\) \((\text{the number of possible tuples in } \Lambda_0 \times \Lambda_0 \times S)\) and these values are independent of \(v\). Writing \(v = lN\tilde{v}_0\), we directly notice that for each such possible value of \(\omega_4(n_1 + v, n_2 + v, k, n)\), there is at most one choice of \(l \in \mathbb{Z}\) that would make \(|\omega_4| \leq R\) \((\text{since } v_0(\tilde{n}_1 - \tilde{n}_2) \neq 0)\).

Case 2: Now suppose that \((n_1, k, n_2) \in \Lambda \times S \times \Lambda\) form a parallelogram with the forth vertex being \(n = n_1 - k + n_2\). In this case, \(n_1 + v, k, n_2 + v\) and \(n + 2v\) also form a parallelogram and:

\[
\omega(n_1, k, n_2, n) = \omega(n_1 + v, k + v, n_2 + v, n + v) = |n_1 + v|^2 + |n_2 + v|^2 - |k + v|^2 - |n + v|^2
\]
\[
= |n_1 + v|^2 + |n_2 + v|^2 - |k|^2 - |n + 2v|^2 + 2v(n + v - k)
\]
\[
= \omega_4(n_1 + v, k, n_2 + v, k, n + 2v) + 2v(n + v - k)
\]

Notice that since \((n_1 + v, k, n_2 + v) \in NA_0 \times S \times NA_0\), there are \(O_{\Lambda_0, S}(1)\) possible values of \(\omega_4(n_1 + v, k, n_2 + v, k, n + 2v)\) and these values are independent of \(v\). Writing as before \(n_i + v = N\tilde{n}_i\) with \(\tilde{n}_i \in \Lambda_0\) for \(i = 1, 2\), we have:

\[
2v(n + v - k) = 2v((n_1 + v) + (n_2 + v) - v - 2k) = 2Rv_0 \cdot (N(\tilde{n}_1 + \tilde{n}_2) - Rv_0 - 2k)
\]
\[
= 2R(Nv_0 \cdot (\tilde{n}_1 + \tilde{n}_2) - Rv_0^2 - 2k) \in 2R\mathbb{Z}
\]

As a result, for each of the \(O_{\Lambda_0, S}(1)\) possible values of \(\omega_4(n_1 + v, k, n_2 + v, k, n_2v)\), there are at most two choices of \(l\) for which \(|\omega(n_1, k, n_2, n)| \leq R\).

Case 3: Now suppose that \((k_1, k_2, n_3) \in S \times S \times \Lambda\) and \(n = k_1 - k_2 + n_3\) form a parallelogram. We have that:

\[
\omega_4(k_1, k_2, n_3, n) = \omega_4(k_1, k_2, n_3 + v, n + v) - 2v(n_3 - n)
\]
\[
= \omega_4(k_1, k_2, n_3 + v, n + v) - 2v(k_1 - k_2)
\]
\[
= \omega_4(k_1, k_2, n_3 + v, n + v) - 2Rv_0(k_1 - k_2)
\]

As before, there are \(O_{\Lambda_0, S}(1)\) possible values for \(\omega_4(k_1, k_2, n_3 + v, n + v)\) and these values are independent of \(v\). For each such value there is at most one choice of \(l \in \mathbb{N}\) so that \(|\omega_4(k_1, k_2, n_3, n)| \leq R\) \((\text{since } v_0 \cdot (k_1 - k_2) \neq 0)\).

Case 4: Finally, suppose that \((k_1, n_2, k_3) \in S \times \Lambda \times S\) and \(n = k_1 - n_2 + k_3\) form a parallelogram. Then:

\[
\omega_4(k_1, n_2, k_3, n) = \omega_4(k_1, n_2 + v, k_3, n - v) + 2|v|^2 + 2(n_2 + n).v
\]
\[
= \omega_4(k_1, n_2 + v, k_3, n - v) + 2|v|^2 + 2(k_1 + k_3).v
\]
\[
= \omega_4(k_1, n_2 + v, k_3, n - v) + 2R(Rv_0^2 + 2(k_1 + k_3)v_0l)
\]
For each of the \( O_{\Lambda_0, S}(1) \) possible values of \( \omega_4(k_1, n_2 + v, k_3, n - v) \), there are at most two choices of \( l \) that can make \( |\omega_4(k_1, n_2, k_3, n)| \leq R \).

In conclusion, for each tuple in \( (S \times S \times \Lambda) \cup (\Lambda \times \Lambda \times S) \), there are at most two choices of \( l \in \mathbb{N} \) for which a parallelogram determined by this tuple satisfies \( |\omega_4| \leq R \). But there are only \( C(P, S) = C(\delta, K, S) \) such tuples, so by choosing \( L > C(\delta, K, S) \), the pigeonhole principle guarantees the existence of an \( L \leq l \leq 2L \) for which all such parallelograms joining \( \Lambda \) and \( S \) satisfy \( |\omega_4| \geq R \). This finishes the proof.

Remark. It is easy to notice that if \( S = \{0\} \) and if \( \Lambda^1 \) is a set satisfying the conditions of Proposition 4.7, then so is any dilation \( N \Lambda^1 \). In particular, one could construct \( \Lambda \) in the sublattice \( N \mathbb{Z}^2 \), which gives the proof of Proposition 3.1.

With Proposition 4.7 in hand, we are now in a position to finish the proof of Proposition 4.4.

Suppose that \( \delta \ll 1, K \gg 1 \) and that \( S \) is a finite set satisfying the \( R \)-closure condition. Recall that we only need to construct a solution \( (a_n(t)) \) to (4.5) that is supported in frequency space on a set \( \Lambda \) satisfying properties Property I_{\Lambda}–Property VIII_{\Lambda} \) and that satisfies (4.9) and (4.10).

Construct \( \Lambda \) as in Proposition 4.7 with \( P = P(\delta, K) \) and some \( N \gg C(\delta, K, S) \). Using Proposition 2.4 we construct a solution to (4.13) satisfying:

\[
|b_3(0)| \geq 1 - \epsilon \quad |b_j(0)| \leq \epsilon \quad j \neq 3 \\
|b_{P-2}(T)| \geq 1 - \epsilon \quad |b_j(T)| \leq \epsilon \quad j \neq P - 2
\]

where \( \epsilon = \epsilon(\delta, K) \) is to be specified later and \( T = T(\epsilon, P) \).

By construction, defining \( \tilde{a}_n(t) = b_j(t) \) for \( n \in \Lambda_j, j = 1, \ldots P \) gives a solution to (4.11), and hence to (4.5) thanks to Property IV_{\Lambda}. The requested solution \( (a_n(t)) \) will be a rescaling of \( (\tilde{a}_n) \) that preserves system (4.5), namely:

\[
a_n(t) = \lambda \tilde{a}_n(\frac{t}{\lambda^2}), \quad t \in [0, \lambda^2 T].
\]

We choose \( \lambda \) so that \( \|(a_n(0))\|_{h^s(\mathbb{Z}^2)} \sim \delta \). In fact, since

\[
\|(a_n(0))\|_{h^s(\mathbb{Z}^2)} \sim \frac{1}{\lambda} \left( \sum_{j=1}^{P} \sum_{n \in \Lambda_j} |n|^{2s} \right)^{1/2} \sim_{\delta, K, S} \lambda^{-1} N^s
\]

we only need to take \( \lambda \sim C(\delta, K, S) N^s \) for (4.9) to hold. The same calculation as that done in Section 3.3 gives that the ratio:

\[
Q = \frac{\|(a_n(\lambda^2 T))\|_{h^s(\mathbb{Z}^2)}}{\|(a_n(0))\|_{h^s(\mathbb{Z}^2)}} \geq \frac{K}{\delta}
\]

from which (4.10) follows.

References

[1] J. Bourgain, Fourier transform restriction phenomena for certain lattice subsets and applications to nonlinear evolution equations, Part I: Schrödinger Equations, GAFA 3 (1993), 107–156.
LONG-TIME INSTABILITY AND UNBOUNDED ORBITS

[2] J. Bourgain, Aspects of long time behaviour of solutions of nonlinear Hamiltonian evolution equations, Geom. Funct. Anal. 5 (1995), no. 2, 105–140.

[3] J. Bourgain, On the growth in time of higher Sobolev norms of smooth solutions of Hamiltonian PDE, Internat. Math. Res. Notices 1996, no. 6, 277–304.

[4] J. Bourgain, Problems in Hamiltonian PDE's, Geom. Funct. Anal. 2000, Special Volume, Part I, 32–56.

[5] J. Bourgain, A remark on normal forms and the I-method for periodic NLS, J. Anal. Math., 94 (2004), 125–157.

[6] J. Bourgain, Remarks on stability and diffusion in high-dimensional Hamiltonian systems and partial differential equations, Ergodic Theory Dynam. Systems 24 (2004), no. 5, 1331–1357.

[7] R. Carles, E. Faou, Energy cascades for NLS on $\mathbb{T}^d$, Discrete Contin. Dyn. Syst. 32 (2012), 2063–2077.

[8] F. Catoire, W.-M. Wang, Bounds on Sobolev norms for the defocusing nonlinear Schrödinger equation on general flat tori, Commun. Pure Appl. Anal. 9 (2010), no. 2, 33–56.

[9] J. E. Colliander, J.-M. Delort, C. E. Kenig, G. Staffilani, Bilinear estimates and applications to 2D NLS, Trans. Amer. Math. Soc. 353 (2001), no. 8, 3307–3325.

[10] J. Colliander, M. Keel, G. Staffilani, H. Takaoka, T. Tao, Polynomial growth and orbital instability bounds for $L^2$-subcritical NLS below the energy norm, Comm. Pure Appl. Anal. 2 (2003), 33–50.

[11] J. Colliander, M. Keel, G. Staffilani, H. Takaoka, T. Tao, Transfer of energy to high frequencies in the cubic defocusing nonlinear Schrödinger equation, Invent. Math. 181 (2010), no. 1, 39–113.

[12] J. Colliander, S. Kwon, T. Oh, A remark on normal forms and the "upside-down" I-method for periodic NLS: growth of higher Sobolev norms, preprint to appear in J. Anal. Math.

[13] D. De Silva, N. Pavlovic, G. Staffilani, N. Tzirakis, Global well-posedness for a periodic nonlinear Schrödinger equation in 1D and 2D, Discrete Contin. Dyn. Syst. 19 (2007) 37–65.

[14] J. Escher, A. 细野, A.C. Nour, A.C.A. Pusharev, V. E. Zakharov, Optical turbulence: weak turbulence, condensates and collapsing filaments in the nonlinear Schrödinger equation, Phys. D 57 (1992), no. 1-2, 96–160.

[15] E. Faou, L. Gauckler, C. Lubich, Sobolev stability of plane wave solutions to the cubic nonlinear Schrödinger equation on a torus, preprint arXiv:1109.2407.

[16] P. Gerard, S. Grellier, The cubic Szegő equation, Ann. Sci. Ec. Norm. Supér.(4), 43(5):761–810, 2010.

[17] P. Gerard, S. Grellier, Effective integrable dynamics for some nonlinear wave equation, preprint arXiv:1110.5719 (2011).

[18] M. Guardia, Growth of Sobolev norms in the cubic defocusing nonlinear Schrödinger equation with a convolution potential, preprint arXiv:1211.1267.

[19] M. Guardia, V. Kaloshin, Growth of Sobolev norms in the cubic defocusing nonlinear Schrödinger equation, preprint arXiv:1205.5188.

[20] Z. Hani, A bilinear oscillatory integral estimate and bilinear refinements to Strichartz estimates on closed manifolds, Analysis & PDE 5-2 (2012) 339–363.

[21] Z. Hani, Global well-posedness of the 2D–cubic nonlinear Schrödinger equation on compact manifolds without boundary, Comm. PDE. Volume 37, Issue 7, 1186–1236 (2012).

[22] Z. Hani, Global and dynamical aspects of nonlinear Schrödinger equations on compact manifolds, PhD Thesis, University of California, Los Angeles.

[23] R. H. Kraichman, Inertial ranges in two-dimensional turbulence, Phys. Fluids 10 (7) 1417–1423 (1997).

[24] S. B. Kuksin, Oscillations in space-periodic nonlinear Schrödinger equations, Geom. Funct. Anal. 7 (1997), no. 2, 338–363.

[25] O. Pocovnicu, Explicit formula for the solution of the Szegő equation on the real line and applications, Discrete Contin. Dyn. Syst., 31(3):607–649, 2011.

[26] O. Pocovnicu, First and second order approximations for a nonlinear wave equation, preprint arXiv:1111.6060 (2012).

[27] V. Sohinger, Bounds on the growth of high Sobolev norms of solutions to nonlinear Schrödinger Equations on $S^1$, Differential Integral Equations 24 (2011), no. 7-8, 653–715.

[28] G. Staffilani, On the growth of high Sobolev norms of solutions for KdV and Schrödinger equations, Duke Math. J. 86 (1997), no. 1, 109–142.

[29] T. Tao, Nonlinear dispersive equations: local and global analysis, Regional Conference Series in Mathematics, 106. American Mathematical Society, Providence, RI, 2006.
[30] V. E. Zakharov, Weak turbulence in media with decay spectrum. Zh. Priklad. Tech. Fix. 4, 35–39 (1965) [J. Appl. Mech. Tech. Phys. 4, 22–24 (1965)].

[31] V. E. Zakharov, V. S. L’vov, G. Falkovich, Kolmogorov Spectra of Turbulence I, Springer Series in Nonlinear Dynamics, Springer-Verlag, 1992.

[32] V. E. Zakharov, A. B. Shabat, Exact theory of two-dimensional self-focusing and one-dimensional self-modulation of waves in nonlinear media. Soviet Physics JETP 34 (1972), no. 1, 62–69.; translated from Ž. Éksper. Teoret. Fiz. 61 (1971), no. 1, 118–134.

[33] Zhong S., The growth in time of higher Sobolev norms of solutions to Schrödinger equations on compact Riemannian manifolds, J. Differential Equations 245 (2008), no. 2, 359–376.