THE WEAKLY NONLINEAR LARGE BOX LIMIT OF THE 2D CUBIC NONLINEAR SCHRÖDINGER EQUATION

ERWAN FAOU, PIERRE GERMAIN, AND ZAHER HANI

ABSTRACT. We consider the cubic nonlinear Schrödinger (NLS) equation set on a two dimensional box of size L with periodic boundary conditions. By taking the large box limit $L \to \infty$ in the weakly nonlinear regime (characterized by smallness in the critical space), we derive a new equation set on $\mathbb{R}^2$ that approximates the dynamics of the frequency modes. This nonlinear equation turns out to be Hamiltonian and enjoys interesting symmetries, such as its invariance under Fourier transform, as well as several families of explicit solutions. A large part of this work is devoted to a rigorous approximation result that allows to project the long-time dynamics of the limit equation into that of the cubic NLS equation on a box of finite size.

CONTENTS

1. Introduction

The long-time behavior of nonlinear dispersive equations on compact domains is a subject that is remarkably rich yet very poorly understood. This complexity arises since dispersion does not translate into decay, neither at the linear nor at the nonlinear level, in contrast to the case of non-compact domains like $\mathbb{R}^d$. As a consequence, one loses all asymptotic stability results around equilibrium solutions (like the zero solution); such results are often the starting point in the study of asymptotic behavior in that setting. In fact, these equilibria are not expected to be stable in the long run, and completely out-of-equilibrium dynamics is anticipated.

The purpose of the present work is to uncover new, and structured or coherent, aspects of this out-of-equilibrium behavior. Such coherent dynamics are best seen in frequency space and will be revealed in the *large-volume limit* and *weakly nonlinear* regime. The latter is determined by how the size of the data compares to that of the box, a factor that plays a key role in determining the limiting system as we argue below.

1.1. Presentation of the equation.

Our model equation will be the 2D cubic nonlinear Schrödinger equation given by

$$\begin{cases} -i\partial_t v + \Delta v &= \pm|v|^2 v \\ v(0,x) &= v_0(x) \qquad x \in \mathbb{T}_L^2 = [0,L] \times [0,L]. \end{cases} \tag{NLS}$$

where $v = v(t,x)$ is a complex-valued field, $t \in \mathbb{R}$ represents time, and the spatial domain is taken to be the box $\mathbb{T}_L^2$ of size L with periodic boundary conditions[1]. Recall that this equation conserves mass and the Hamiltonian given respectively by:

$$M[v(t)] \overset{def}{=} \int_{\mathbb{T}_L^2} |v(t,x)|^2 \, dx \quad \text{and} \quad H[v(t)] \overset{def}{=} \frac{1}{2} \int_{\mathbb{T}_L^2} |\nabla v(t,x)|^2 \, dx \pm \frac{1}{4} \int_{\mathbb{T}_L^2} |v(t,x)|^4 \, dx. \tag{1.1}$$

---

[1]Since the equation respects the parity of the initial data, the periodic boundary condition include both Dirichlet and Neumann conditions corresponding respectively to odd and even solutions.

The sign of the non-linearity (+ for defocusing equation and − for focusing equation) will not play a central part in the analysis; this is partly due to the fact that we focus on the long-time behavior of small initial data, i.e. ones that are close to the zero equilibrium $v = 0$ in the scale-invariant topology (namely $L^2(\mathbb{T}_L^2)$).

The size of the data is measured in comparison to the size L of the box, and their relationship will be a defining factor for various regimes of long-time behavior. Consequently, it will be both convenient and more transparent to adopt the ansatz $v(t, x) = \epsilon u(t, x)$, where $\epsilon > 0$. The field u satisfies

$$\begin{cases} -i\partial_t u + \Delta u & = \pm\epsilon^2 |u|^2 u \\ u(0, x) & = u_0(x). \end{cases} \tag{NLS$_\epsilon$}$$

When working with (NLS$_\epsilon$), it makes most sense to consider that $\|u_0\|_{L^2} \sim 1$, since the ϵ parameter already accounts for the size of the data.

1.2. Cubic NLS on $\mathbb{T}^2$ vs. $\mathbb{R}^2$.
Before starting our analysis, let us review known results on the large-time behavior of solutions to (NLS) on $\mathbb{R}^2$ and $\mathbb{T}^2$ (of course, it would be equivalent, but less convenient, to discuss (NLS$_\epsilon$)) .

The equation is mass-critical in the sense that the scaling symmetry given by $v \mapsto v_\lambda \overset{def}{=} \lambda v(\lambda^2 t, \lambda x)$ leaves both the equation and the mass invariant[2]. This leads to global well-posedness in $H^s(\mathbb{R}^2)$ for all $s \geq 0$ and small mass ([13] or see [57] for a textbook treatment). This restriction on the mass can be removed completely in the defocusing case and relaxed in the focusing case up to the mass of the ground state [41, 19, 20]. More importantly, the asymptotic behavior near the zero equilibrium is completely understood in this Euclidean setting: any small mass solution scatters to a linear solution as $t \to \pm\infty$. This means that if $v_0 \in H^s(\mathbb{R}^2)$ for some $s \geq 0$, there exists $\varphi_{\pm\infty} \in H^s(\mathbb{R}^2)$ such that

$$\|v(t) - e^{-it\Delta_{\mathbb{R}^2}}\varphi_{\pm\infty}\|_{H^s(\mathbb{R}^2)} \to 0 \qquad \text{as } t \to \pm\infty.$$

Moreover, one has the following asymptotic stability statement which is a direct consequence of the small data theory: If $v_0 \in H^s(\mathbb{R}^2)$ for some $s \geq 0$ is sufficiently small, then

$$\|v(t) - e^{-it\Delta}v_0\|_{H^s(\mathbb{R}^2)} = O(\|v_0\|_{H^s}^3) \qquad \text{for all } t \in \mathbb{R}. \tag{1.2}$$

On $\mathbb{T}^2$, local well-posedness was proved by Bourgain [5] for data in H^s, with $s > 0$. This leads when combined with (1.1) to global well-posedness for data with finite energy (and small mass in the focusing case). We remark that since we will be only interested in the dynamics of *small* solutions of (NLS) with finite energy, or equivalently solutions of (NLS$_\epsilon$) with $\epsilon \ll 1$, all solutions exist globally and conserve mass and energy. It is also worth pointing out that local well-posedness in the scale-invariant norm L^2 remains an open question on $\mathbb{T}^2$ due to the failure of the scale-invariant $L_{t,x}^4$ Strichartz estimate (see (3.1)). Rather surprisingly, although we will only be interested in sufficiently regular solutions that exist globally by elementary arguments, this endpoint Strichartz estimate and the sharp constants associated to it will play a central role in our analysis (see Section 3).

When it comes to long-time behavior of (NLS) on $\mathbb{T}^2$, we observe a number of phenomena rather than the single-scenario situation of the Euclidean setting. Indeed, the expectation is that (NLS) is a *non-integrable* infinite dimensional Hamiltonian system, and as a consequence

---

[2]On the torus, scaling is not a perfectly honest symmetry: v_λ is a solution of (NLS) on $\mathbb{T}_{\lambda^{-1}}^2$ rather than $\mathbb{T}^2$. Nevertheless, this symmetry keeps the mass $M[u]$ invariant, which extends the crucial notion of criticality to the compact setting.

it should sustain, just as in the finite dimensional case, orbits exhibiting strikingly different asymptotic behaviors. Known results include KAM tori and quasi-periodic solutions [24, 54, 7, 21] and heteroclinic and unbounded Sobolev orbits (see [16, 29, 28] for partial results in this direction). In particular, no long-time stability result anywhere close to (1.2) can hold for general data. For instance, one of the byproducts of the constructions in this paper is a whole regime of initial data violating (1.2) at time scales of roughly $\|v_0\|_{L^2}^{-2}$.

1.3. **The weakly nonlinear large-box limit.** Equation (NLS$_\epsilon$) has a particularly transparent form in Fourier space: Denoting by $\mathbb{Z}_L^2 \overset{def}{=} L^{-1}\mathbb{Z}^2$ and expanding

$$v(t,x) = L^{-2} \sum_{K \in \mathbb{Z}_L^2} a_K(t) e^{2\pi i K \cdot x}, \qquad \text{where} \qquad a_K(t) = \int_{\mathbb{T}_L^2} v(t,x) e^{-2\pi i K \cdot x}\, dx$$

one obtains that $a_K(t)$ solves the infinite dimensional system of ODE:

$$-i\partial_t a_K(t) - 4\pi^2 |K|^2 a_K(t) = \pm \frac{\epsilon^2}{L^4} \sum_{(K_1,K_2,K_3) \in \mathcal{S}(K)} a_{K_1}(t) \overline{a_{K_2}(t)} a_{K_3}(t)$$

where $\mathcal{S}(K) = \{(K_1, K_2, K_3) \in \mathbb{Z}_L^2 : K_1 - K_2 + K_3 = K\}$. Obviously, the linear flow evolves much faster than the non-linear one, which makes it convenient to pass to the so-called "interaction representation picture" by defining $\widetilde{a}_K(t) = e^{-4\pi^2 i |K|^2 t} a_K(t)$. The equation satisfied by $\widetilde{a}_K(t)$ is:

$$-i\partial_t \widetilde{a}_K(t) = \pm \frac{\epsilon^2}{L^4} \sum_{\mathcal{S}(K)} \widetilde{a}_{K_1}(t) \overline{\widetilde{a}_{K_2}(t)} \widetilde{a}_{K_3}(t) e^{4\pi^2 i \Omega t} \tag{1.3}$$

where $\Omega \overset{def}{=} |K_1|^2 - |K_2|^2 + |K_3|^2 - |K|^2$. This equation describes how the frequency mode a_K is excited by other modes through the nonlinear interactions included in $\mathcal{S}(K)$.

Our aim is now to consider the weakly nonlinear ($\epsilon \to 0$) large-box ($L \to \infty$) limit, while keeping a balance between ϵ and L. We elaborate on this below.

- The weakly nonlinear limit $\epsilon \to 0$ enables to restrict the above sum $\sum_{\mathcal{S}(K)}$ to *resonant interactions*, which correspond to those $(K_1, K_2, K_3) \in \mathcal{S}(K)$ such that $\Omega = 0$. This can be achieved by a normal form transformation. Restricting nonlinear interactions to resonant ones, one obtains the related *resonant system*

$$-i\partial_t r_K(t) = \pm \frac{\epsilon^2}{L^4} \sum_{\mathcal{R}(K)} r_{K_1}(t) \overline{r_{K_2}(t)} r_{K_3}(t) \tag{1.4}$$

  where $\mathcal{R}(K) = \{(K_1, K_2, K_3) \in \mathcal{S}(K) : \Omega \overset{def}{=} |K_1|^2 - |K_2|^2 + |K_3|^2 - |K|^2 = 0\}$. This system is also the first Birkhoff normal form approximation of (1.3) and will play a central role in our analysis. Its fundamental importance in describing the long-time behavior of (NLS) has been realized and exploited in previous works ([16, 12, 29, 30]).

- The large box limit $L \to \infty$ is the thermodynamic or infinite volume limit: we are not concerned too much about the behavior of each single Fourier mode, but rather by "macroscopic" interactions and variables. This perspective can be compared to that taken in statistical mechanics to study many-particle systems. Passing to the large-box limit is also one of the main ingredients of *kinetic weak turbulence theory*, which seeks a statistical description of the out-of-equilibrium frequency dynamics of

equations like (NLS_ϵ). Taking the large-box limit, one obtains from (1.4) a simpler system now set on $\mathbb{R}^2$. A very pedestrian explanation is the following: as L goes to infinity, the lattice $\mathbb{Z}_L^2$ becomes more and more refined, allowing (in some sense which will be explained) to consider the right-hand side of (1.4) as a Riemann sum, approaching a certain integral as $L \to \infty$.

- <u>The relative sizes of ϵ and L:</u> as measured by the quantity $\frac{(-\log \epsilon)}{\log L}$ are crucial in determining which limiting system best describes the dynamics in the corresponding range of ϵ and L. The main focus of this work is the regime where $\epsilon L \ll 1$. As we argue below, this is exactly the regime where the resonant dynamics dictated by (1.4) dominate non-resonant interactions in (1.3).

Taking the large-box limit of (1.4) in a rigorous fashion entails several difficulties due to the non-linear lattice relation "$\Omega = 0$" defining the resonant manifold $\mathcal{R}(K)$. This is the reason why tools from analytic and geometric number theory will play a crucial role in the analysis, first at the level of deriving the limit equation and then at the level of proving the needed uniform estimates on resonant sums like that in (1.4). Without going into the details of the derivation for now, the offspring of this analysis is a nonlinear integro-differential equation, now set on $\mathbb{R}^2$, given by

$$
\begin{aligned}
-i\partial_t g(\xi, t) &= \mathcal{T}(g, g, g)(\xi, t); \qquad \xi \in \mathbb{R}^2 \\
\mathcal{T}(g, g, g)(\xi, t) &= \int_{-1}^{1} \int_{\mathbb{R}^2} g(\xi + \lambda z, t)\overline{g}(\xi + \lambda z + z^\perp)g(\xi + z^\perp)\, d\lambda\, dz.
\end{aligned}
\tag{CR}
$$

where we used $z^\perp$ to denote the rotation of $z \in \mathbb{R}^2$ by the angle of $+\pi/2$. The name (CR) stands for *continuous resonant* as it corresponds to a continuous limit of (1.4). Morally speaking, one should think of the frequency modes a_K of (1.3) as being approximated by the trace[3] of a *rescaled version* of $g(t, \xi)$ on the lattice $\mathbb{Z}_L^2$. To the best of our knowledge, equation (CR) is new. Upon analyzing it, one soon realizes that it is somewhat special in the sense that it enjoys rather unusual properties, symmetries, and even explicit solutions that we postpone elaborating on till the next section.

How does the new equation inform on the long-time behavior of (NLS_ϵ) on a box of finite size? In other words, can we prove that the nonlinear dynamics of (CR) is a reflection or a "subset" of the dynamics of (NLS_ϵ)? The positive answer to these questions is a major part of the analysis and is one of the novelties of this work. It is contained in Theorems 2.4 and 2.6, whose precise statements we delay until we have set up all our notations and parameters in the next section. Roughly speaking, Theorem 2.4 allows to project the dynamics of (CR) onto that of (NLS_ϵ) posed on a box of size L, whereas Theorem 2.6 shows how the dynamics of (CR) embed into that of (NLS) on the torus $\mathbb{T}^2$.

The weakly nonlinear, large-box regime is the usual set up for the theory of weak turbulence, also called wave turbulence. We present this theory briefly in the next section to draw the similarities and differences with what is done here. Equation (CR) can be viewed as a deterministic version of the famous Kolmogorov-Zakharov (KZ)[4] equation of weak turbulence theory. The (KZ) equation is widely used in the physics and applied sciences (oceanography

---

[3]It is customary in the physics literature not to distinguish between the notation for $a_K(t)$ and its "large-box limit" $g(t, K)$ and denote both by $a_K(t)$ with the understanding that $K \in \mathbb{R}^2$ after the large-box limit is taken.

[4]Also called sometime *wave-wave kinetic equation.*

and atmospheric sciences) to understand several aspects of the frequency dynamics of equations like (NLS$_\epsilon$). Despite many similarities, the most notable difference between this work and the vast physics literature on (kinetic) wave turbulence is that our derivation does not involve any randomization of the data. This has the effect that the limiting equation (CR) is time-reversible (even Hamiltonian), as opposed to (KZ) (cf. Section 2.5). It should be noted that the proper and rigorous justification of the Kolmogorov-Zakharov equation is a very important, and difficult, open question both from the physical and mathematical viewpoint. We hope that this work is a step forward in this direction.

1.4. **Organization of the paper.** The paper is organized as follows. We end this introduction by listing the notation used in the rest of the manuscript. In the next section, we sketch briefly the derivation of (CR), state the main approximation theorems 2.4 and 2.6 mentioned above, and summarize the properties of (CR). We also include in that section some heuristic interpretations of (CR) and elaborate on its relation to weak turbulence theory. In section 3 we prove the crucial trilinear estimates on resonant sums like that in (1.4), which are of interest in their own right as they are intimately related to some sharp Strichartz estimates on $\mathbb{T}^2$. In Section 4, we prove that the right-hand-side of (1.4) converges in the limit of large L to the right-hand-side of (CR) up to important rescaling factors and obtain quantitative estimates on the errors incurred. In Section 5 we give the proof of the approximation theorems. Afterwards, we embark on analyzing equation (CR) in Sections 6-8, starting from its Hamiltonian properties, symmetries and invariances, explicit solutions, global well-posedness results, as well as variational properties and characterizations of some of the explicit solutions.

1.5. **Notations.** The following notations are used. Recall that we denote $\mathbb{Z}_L^2 = L^{-1}\mathbb{Z}^2$ and for any subset Λ of $\mathbb{Z}_L^2$, $\Lambda^* = \Lambda \setminus \{0\}$.

1.5.1. *Fourier transform.* <u>On $\mathbb{R}^2$:</u> The Fourier transform of a function f on $\mathbb{R}^2$ is denoted $\mathcal{F}f$ or $\widehat{f}$, and given by

$$\mathcal{F}f(\xi) = \widehat{f}(\xi) = \frac{1}{2\pi}\int_{\mathbb{R}^2} e^{-ix\cdot\xi} f(x)\, dx \quad \text{so that} \quad f(x) = \frac{1}{2\pi}\int_{\mathbb{R}^2} e^{ix\cdot\xi}\widehat{f}(\xi)\, d\xi.$$

The map $f \mapsto \widehat{f}$ is an isometry on $L^2(\mathbb{R}^d)$.

<u>On $\mathbb{T}_L^2$:</u> The Fourier transform of a function u on $\mathbb{T}_L^2$ is the sequence $a_K = \widehat{u}(K)$ with $K \in \mathbb{Z}_L^2 = L^{-1}\mathbb{Z}^2$ given by

$$a_K = \int_{\mathbb{T}_L^2} u(x)e^{-2\pi iK.x}\, dx \quad \text{so that} \quad u(x) = \frac{1}{L^2}\sum_{K\in\mathbb{Z}_L^2} a_K(t)e^{2\pi iK.x}.$$

The map $u \mapsto \{a_K\}$ is an isometry from $L^2(\mathbb{T}_L^2)$ to ℓ_L^2 defined below.

1.5.2. *Function spaces.* <u>On $\mathbb{Z}_L^2$:</u> We will consider sequences $(a_K)_{K\in\mathbb{Z}_L^2}$ and use the normalized counting measure $L^{-2}\sum_{K\in\mathbb{Z}_L^2}\delta_K$ to define $\ell^p(\mathbb{Z}_L^2)$ norms for $p < \infty$. Therefore we set for $1 \le p < +\infty$,

$$\|a_K\|_{\ell_L^p} \overset{def}{=} \left(L^{-2}\sum_{K\in\mathbb{Z}_L^2} |a_K|^p\right)^{1/p} \quad \text{and} \quad \|a_K\|_{\ell_L^\infty} \overset{def}{=} \sup_{K\in\mathbb{Z}_L^2} |a_K|, \quad \text{for} \quad p = +\infty.$$

The weighted $\ell_L^p(\mathbb{Z}_L^2)$ spaces will be denoted $\ell_L^{p,\sigma}$ with norms

$$\|a_K\|_{\ell_L^{p,\sigma}} \overset{def}{=} \|\langle K \rangle^\sigma a_K\|_{\ell_L^p,}$$

Of particular importance in the analysis will be the space $X_L^\sigma \overset{def}{=} \ell_L^{\infty,\sigma}$ with $\sigma \in \mathbb{R}$, the space associated with the norm

$$\|a_K\|_{X_L^\sigma} \overset{def}{=} \|\langle K \rangle^\sigma a_K\|_{\ell_L^\infty}. \tag{1.5}$$

Notice that with our normalization of the $\ell_L^{2,s}(\mathbb{Z}_L^2)$ spaces we have that

$$\|a_K\|_{\ell_L^{2,s}(\mathbb{Z}_L^2)} \lesssim \|a_K\|_{X_L^\sigma(\mathbb{Z}_L^2)} \quad \text{whenever } \sigma > s+1.$$

On $\mathbb{R}^2$ and $\mathbb{T}_L^2$: We use the standard definitions and notations for Lebesgue, weighted Lebesgue (in their homogeneous and inhomogeneous versions), and Sobolev spaces on $\mathbb{R}^2$ and $\mathbb{T}_L^2$. For example, with M standing for either $\mathbb{R}^2$ or $\mathbb{T}_L^2$ we denote the L^p, homogeneous weighted L^p, inhomogeneous weighted L^p, and Sobolev norms respectively by

$$\|g\|_{L^p(M)} \overset{def}{=} \left(\int_M |g(x)|^p \, dx \right)^{1/p}$$

$$\|g\|_{\dot{L}^{p,\sigma}(M)} \overset{def}{=} \||x|^\sigma g(x)\|_{L^p(M)}$$

$$\|g\|_{L^{p,\sigma}(M)} \overset{def}{=} \|\langle x \rangle^\sigma g(x)\|_{L^p(M)}$$

$$\|g\|_{H^m(M)} \overset{def}{=} \|\langle \nabla \rangle^m g(x)\|_{L^2(M)},$$

for $p \in [1,\infty]$ (with the usual modification for $p = \infty$), and $m, \sigma \in \mathbb{R}$. Notice that the Sobolev norm of a function $\phi \in H^s(\mathbb{T}_L^2)$ is equivalent to $\|\widehat{\phi}\|_{\ell_L^{2,s}}$. The inner product on $L^2(M)$ is denoted by

$$\langle f, g \rangle \overset{def}{=} \int_M f(x)\overline{g(x)} \, dx.$$

Finally, because of their particular importance in our analysis, we will denote by $\dot{X}^\sigma = \dot{L}^{\infty,\sigma}(\mathbb{R}^2)$ and $X^\sigma := L^{\infty,\sigma}(\mathbb{R}^2)$ with norm

$$\|g\|_{X^\sigma} \overset{def}{=} \|\langle x \rangle^\sigma g\|_{L^\infty(\mathbb{R}^2)}. \tag{1.6}$$

Note that for a given function in $g \in X^\sigma(\mathbb{R}^2)$, we can associate the trace or projection sequence $\{g_K\}_{K \in \mathbb{Z}_L^2} \in X_L^\sigma(\mathbb{Z}_L^2)$ by the formula $g_K = g(K)$.

1.5.3. *Miscellaneous.*

- The counter-clockwise rotation of center 0 and angle $\frac{\pi}{2}$ on $\mathbb{R}^2$ will be denoted by the superscript $\cdot^\perp$.
- The "Japanese brackets" stand for $\langle x \rangle \overset{def}{=} \sqrt{1 + |x|^2}$.
- The normalized Gaussian with L^2 mass one reads $G(x) \overset{def}{=} \frac{1}{\sqrt{\pi}} e^{-\frac{|x|^2}{2}}$.
- $A \lesssim B$ if $A \leq CB$ for some implicit, universal constant C. $A \lesssim_\delta B$ means that the implicit constant C depends on δ.
- $A \sim B$ if $A \lesssim B$ and $B \lesssim A$.
- For any $m \in \mathbb{R}_+$, we denote by $\mathcal{B}(m)$ the box $[-\frac{m}{2}, \frac{m}{2}) \times [-\frac{m}{2}, \frac{m}{2})$ and by $B(x, m)$ the Euclidean ball with center x and radius m.

2. Statement of the results

In this section, we present a very formal sketch of the derivation of (CR) from (1.3). This is done rigorously in the following sections. We then state the main results in this work starting with the main approximation theorems that allow to compare the dynamics of (CR) and that of (NLS) along with the crucial estimates involved in proving them. We believe the latter to be of interest in their own right. Afterwards, we state the properties of the (CR) equation that are proved in Sections 6-8. Finally we present a heuristic interpretation of (CR) and elaborate on how it fits in the weak turbulence picture.

2.1. **Formal derivation of (CR).** Recall that we denoted by $v(t, x)$ the solution of (NLS), by $u(t, x) \stackrel{def}{=} \epsilon^{-1} v(t, x)$ the solution of (NLS$_\epsilon$), and by $a_K(t) \stackrel{def}{=} \hat{u}(t, K)$ for $K \in \mathbb{Z}_L^2$. We then went to the interaction representation picture and defined $\tilde{a}_K(t) = a_K(t) e^{-4\pi^2 i |K|^2 t}$ which satisfies the system (1.3)

$$-i\partial_t \tilde{a}_K(t) = \frac{\epsilon^2}{L^4} \sum_{\mathcal{S}(K)} \tilde{a}_{K_1}(t) \overline{\tilde{a}_{K_2}(t)} \tilde{a}_{K_3}(t) e^{4\pi^2 i \Omega t}$$

with $\Omega \stackrel{def}{=} |K_1|^2 - |K_2|^2 + |K_3|^2 - |K|^2$ and $\mathcal{S}(K) = \{(K_1, K_2, K_3) : K_1 - K_2 + K_3 = K\}$. Note that $a_K(t)$ can be directly reconstructed from $\tilde{a}_K(t)$ and vice-versa. Splitting the interactions on the right-hand-side of the above equation into resonant and non-resonant gives

$$-i\partial_t \tilde{a}_K(t) = \underbrace{\frac{\epsilon^2}{L^4} \sum_{\mathcal{R}(K)} \tilde{a}_{K_1}(t) \overline{\tilde{a}_{K_2}(t)} \tilde{a}_{K_3}(t)}_{\text{Resonant interactions}} + \underbrace{\frac{\epsilon^2}{L^4} \sum_{\mathcal{S}(K) \backslash \mathcal{R}(K)} \tilde{a}_{K_1}(t) \overline{\tilde{a}_{K_2}(t)} \tilde{a}_{K_3}(t) e^{4\pi^2 i \Omega t}}_{\text{Non-resonant interactions}}$$

where $\mathcal{R}(K) = \{(K_1, K_2, K_3) \in \mathcal{S}(K) : \Omega = 0\}$. It is well-known that if ϵ is small enough resonant interactions have considerably more importance on the dynamics than non-resonant ones [16, 12, 29]. This is often made precise by resorting to a *normal form transformation*. A normal form transformation is an invertible change of variable $\tilde{a}_K \rightarrow c_K = \mathcal{U}(\tilde{a}_K)$ that is sufficiently close to the identity operator and which transforms the equation for $\tilde{a}_K$ into the following equation for $c_K(t)$:

$$-i\partial_t c_K(t) = \frac{\epsilon^2}{L^4} \sum_{\mathcal{R}(K)} c_{K_1}(t) \overline{c_{K_2}(t)} c_{K_3}(t) + \epsilon^4 (\text{quintic term}).$$

Several normal form transformations give the above result, we mention Poincaré-Dulac and Birkhoff normal forms. We will use the somewhat simpler Poincaré-Dulac normal form, but leave the details to Section 5 when we prove Theorem 2.4. If ϵ is small enough (depending on L and the control we have on relevant norms of $\{c_K\}$), the resonant sum above dominates the remainder quintic term and drives the dynamics. It is here that the weak non-linearity assumption comes into the picture and dictates the range of $\frac{(-\log \epsilon)}{\log L}$ under consideration.

The upshot of the above discussion is that we can pretend, at least formally, that $\tilde{a}_K(t)$ satisfies the resonant system (1.4), namely

$$-i\partial_t \tilde{a}_K(t) \text{ “} = \text{ ” } \frac{\epsilon^2}{L^4} \sum_{\mathcal{R}(K)} \tilde{a}_{K_1}(t) \overline{\tilde{a}_{K_2}(t)} \tilde{a}_{K_3}(t) \tag{2.1}$$

At this step, one can interpret " $=$ " as an equality sign with some lower order terms added to the right-hand-side, but things will get even more formal below. The next step is to take the large-box limit of the above equation. This is equivalent to changing the above discrete sums in (2.1) to integrals via a suitable equidistribution analysis. Recall that if $u(z)$ is a sufficiently regular and decaying function on $\mathbb{R}^2$ then

$$L^{-2} \sum_{K \in \mathbb{Z}_L^2} u(K) \to \int_{\mathbb{R}^2} u(z)\,dz, \quad \text{when} \quad L \to +\infty.$$

Of course the sum in (2.1) is much more complicated than a standard Riemann sum, so our first task is to re-parametrize it in a way that allows us to pass to the large-box limit. For this, we change variables in $\mathcal{R}(K)$ and write $K_i = K + N_i$ for $i = 1, 2, 3$. The defining relations of $\mathcal{R}(K)$ now translate to $N_1 + N_3 = N_2$ and $|N_1|^2 + |N_3|^2 = |N_2|^2$. As a result, we have by Pythagoras's theorem that

$$\mathcal{R}(K) = \{(K + N_1, K + N_2 + N_3, K + N_3), \quad N_1 \cdot N_3 = 0, \quad (N_1, N_3) \in \mathbb{Z}_L^2\}. \tag{2.2}$$

The degenerate sums when either N_1 or N_3 is zero disappear in the large-box limit and can be treated as error terms because they account for much fewer lattice points than the non-degenerate case when $N_1, N_3 \neq 0$. This will be very clear when we do the estimates in Section 4, but for now it can be seen from Figure 1 which shows the plots of rectangles in $\mathcal{R}(K)$ with $K = (0,0)$ on lattices $\mathbb{Z}_L^2$ with different size of L. Notice how taking L larger and larger makes the density of rectangles tend to a continuous one.
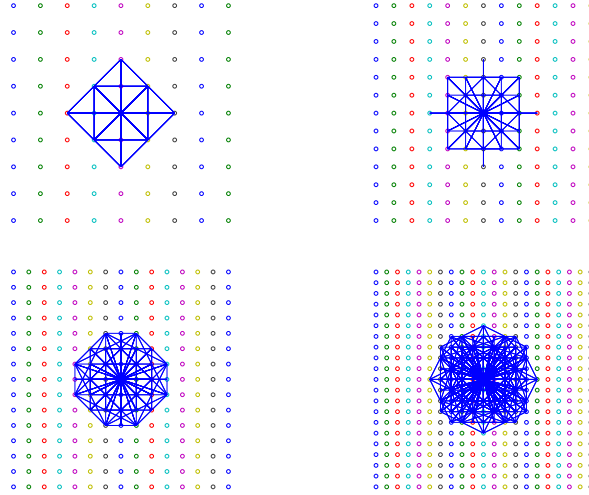
FIGURE 1. Representations of the resonant sets $\mathcal{R}(0,0)$ when L increases

As a result, we can again pretend that $\widetilde{a}_K(t)$ satisfies the following system:

$$-i\partial_t \widetilde{a}_K(t) \text{ " } = \text{ " } \frac{\epsilon^2}{L^4} \sum_{\substack{N_1, N_3 \in \mathbb{Z}_L^{2*} \\ N_1 \perp N_3}} \widetilde{a}_{K+N_1}(t)\overline{\widetilde{a}_{K+N_1+N_3}(t)}\widetilde{a}_{K+N_3}(t). \tag{2.3}$$

The above sum is still not very amenable to taking a large-box limit, so we further parametrize it as follows. Write $N_1 = \frac{\alpha}{L}(p, q)$ where $\alpha \in \mathbb{N}$ and $p, q \in \mathbb{Z}$ are relatively prime in the sense

that $g.c.d.(|p|,|q|) = 1$[5]. Since $N_3 \perp N_1$, then it has to be of the form $\frac{\beta}{L}(-q,p) = \frac{\beta}{L}(p,q)^\perp$ for some $\beta \in \mathbb{Z}^*$. Lattice points in $\mathbb{Z}_L^2$ of the form $\frac{(p,q)}{L}$ are special and are called *visible lattice points*.

Definition 2.1. *For a given number $L > 0$, we say that a lattice point $z \in (\mathbb{Z}_L^2)^*$ is* visible *(from 0) if the segment $[0,z]$ does not intersect the lattice $\mathbb{Z}_L^2$ apart from the endpoints 0 and z. It is easy to see that a lattice point $z \in {\mathbb{Z}_L^2}^*$ is visible if and only if $z = \frac{1}{L}(p,q)$ where (p,q) are co-prime.*

As a result, we can again morally pretend that $\widetilde{a}_K(t)$ satisfies

$$-i\partial_t \widetilde{a}_K(t) \text{``} = \text{''} \frac{\epsilon^2}{L^4} \sum_{\alpha \in \mathbb{N}, \beta \in \mathbb{Z}^*} \sum_{\substack{J \in {\mathbb{Z}_L^2}^* \\ J \, visible}} \widetilde{a}(K + \alpha J, t)\overline{\widetilde{a}(K + \alpha J + \beta J^\perp, t)}\widetilde{a}(K + \beta J^\perp, t) \quad (2.4)$$

where $J^\perp$ is the rotation of J by $+\frac{\pi}{2}$. We are now in a better position to take the large-box limit. The only remaining obstruction is that the sum over J is not a standard equidistribution sum over all lattice points in $\mathbb{Z}_L^2$ but only over visible ones. It turns out that visible lattice points have density $\zeta(2)^{-1} = \frac{6}{\pi^2}$ in $\mathbb{Z}_L^2$ where $\zeta(2)$ is the Riemann zeta function evaluated at 2. In other words, if one picks a lattice point at random then with probability $\zeta(2)^{-1}$ this lattice point is visible. The translation of this statement to equidistribution theory is the fact that if $u(z)$ is a sufficiently regular and decaying function then

$$L^{-2} \sum_{\substack{J \in \mathbb{Z}_L^2 \\ J \, visible}} u(z)\, dz \to \frac{1}{\zeta(2)} \int_{\mathbb{R}^2} u(z) dz.$$

The proof of the claim on the density of visible lattice points relies on the Möbius inversion formula (see for example [31]). We briefly repeat it in Section 4.2 as it serves as a prototype for the argument needed to replace the more complicated trilinear sum (over J visible) in (2.4) by an integral over $\mathbb{R}^2$ and obtain good estimates on the discrepancy error (see Lemma 4.1).

Now that we know how to replace sums over visible lattice points by integrals, we are ready to pass to the large-box limit, at least formally. Pretending that $\{\widetilde{a}(K,t)\}_{K \in \mathbb{Z}_L^2}$ is the trace of a sufficiently smooth and decaying function $\{\widetilde{a}(\xi,t)\}_{\xi \in \mathbb{R}^2}$ and replacing the sums in (2.4) by integrals, one obtains after a somewhat tedious calculation (performed in Section 4 with the needed error estimates) that $\widetilde{a}(K,t)$ satisfies the equation:

$$-i\partial_t \widetilde{a}(K,t) \text{``` } = \text{ '''} \frac{4\epsilon^2 \log L}{\zeta(2)L^2} \int_{-1}^{1} \int_{\mathbb{R}^2} \widetilde{a}(K+z)\overline{\widetilde{a}(K+z+\lambda z^\perp)}\widetilde{a}(K + \lambda z^\perp)\, dz\, d\lambda, \quad (2.5)$$

where we used the notation ``` `` = '' ''' to emphasize the heuristic nature of this last analysis.

Reparameterizing time gives the equation

$$-i\partial_t \widetilde{a}(K,t) \text{``` } = \text{ '''} \int_{-1}^{1} \int_{\mathbb{R}^2} \widetilde{a}(K+z)\overline{\widetilde{a}(K+z+\lambda z^\perp)}\widetilde{a}(K + \lambda z^\perp)\, dz\, d\lambda,$$

which is exactly (CR)! Making sense of the above formal argument is not a straightforward task and several technical difficulties have to be overcome before any sort of rigorous statement can be made in this direction. We elaborate on some of them below.

---

[5]Here we use the convention that if p (resp. q) is zero, then p and q are relatively prime if and only if $q = \pm 1$ (resp. $p = \pm 1$).

2.2. **Uniform bounds and approximation theorems.** The ultimate aim here is to understand and prove in what sense and to what extent do the long-time dynamics of (CR) inform on that of (1.3). Some answers are provided by Theorem 2.4, which pertains to (NLS_ϵ) on the box of size L, and Theorem 2.6 which pertains to the (NLS) on the torus $\mathbb{T}^2$ (or $L = 1$).

2.2.1. *Uniform bounds on resonant sums.* A crucial part in the proof of the above-mentioned theorems is obtaining *sharp* bounds on resonant sums like that appearing on the right-hand-side of (2.1). More concretely, let us define the normalized trilinear operator

$$\mathcal{T}_L(e, f, g)(K) \overset{def}{=} \frac{\zeta(2)}{2L^2 \log L} \sum_{\mathcal{R}(K)} e_{K_1} f_{K_2} g_{K_3}, \tag{2.6}$$

acting on sequences $e = \{e_K\}$, $f = \{f_K\}$ and $g = \{g_K\}$ on $\mathbb{Z}_L^2$. The normalization of $\mathcal{T}_L$ was chosen merely for notational convenience so that the following sharp bounds appear uniform in L. Estimates on such operators are intimately connected to periodic Strichartz estimates and are in fact equivalent to the $L_x^2 \to L_{t,x}^4$ estimate if one is bent on working in L^2-based spaces (see Section 3). Notice that since we are essentially passing to the $L \to \infty$ limit, sharp estimates are indespensible and even a logarithmic loss in L cannot be tolerated. We will see that estimates with very small loss (L^ϵ for any ϵ) can be somewhat easily proved from the existing literature on periodic Strichartz estimates (see [5, 8] and Section 3.3), but they turn out to be almost completely useless by themselves to prove the approximation results. Unfortunately, the sharp version of the Strichartz estimate we need is open and is widely considered to be a very difficult problem at the intersection of harmonic analysis and number theory [6, 42]. Our way around this dilemma is to work with norms that are not L^2-based, namely spaces $X_L^\sigma(\mathbb{Z}_L^2)$ defined in (1.5), and prove the *sharp* Stichartz-type estimate in these norms, which turns out to be more tractable than the L^2-based Strichartz estimate. The argument here relies on quite delicate estimates on lattices and the end result is the following theorem proved in Section 3

Theorem 2.2. *Let $\sigma > 2$ and $\{a_K\}, \{b_K\}$, and $\{c_K\}$ be three sequences in X_L^σ. The following estimate holds uniformly in L:*

$$\|\mathcal{T}_L(a_K, b_K, c_K)\|_{X_L^\sigma} \lesssim \|a_K\|_{X_L^\sigma} \|b_K\|_{X_L^\sigma} \|c_K\|_{X_L^\sigma}. \tag{2.7}$$

Moreover, the dependance in L in this estimate is sharp.

The true meaning of the last statement will become clear below where we show that the operator $\mathcal{T}_L$ is not only uniformly bounded in X_L^σ, but also converges as $L \to \infty$ to a *non-zero* entity when it acts on sequences that are traces of sufficiently smooth and decaying functions. It is worth mentioning that as a corollary to Theorem 2.2, we obtain what seems to be the first $L_{t,x}^4$ Strichartz-type estimate on $\mathbb{T}^2$ with the sharp L^2-critical scaling (See Corollary 3.2).

2.2.2. *Discrete to continuous convergence.* Let us first define the trilinear operator $\mathcal{T}$ acting on complex valued functions $f, g, h : \mathbb{R}^2 \to \mathbb{C}$ as follows

$$\text{if } \xi \in \mathbb{R}^2, \qquad \mathcal{T}(f, g, h)(\xi) \overset{def}{=} \int_{-1}^1 \int_{\mathbb{R}^2} f(\xi + x)\overline{g(\xi + x + \lambda x^\perp)}h(\xi + \lambda x^\perp)\, dx\, d\lambda. \tag{2.8}$$

The crucial observation is that, as $L \to \infty$, $\mathcal{T}_L$ converges to $\mathcal{T}$ in a sense made precise by the following theorem:

Theorem 2.3 (Discrete to continuous approximation in X^σ)**.** *Let $\sigma > 2$ and suppose that $g : \mathbb{R}^2 \to \mathbb{C}$ satisfies the following bounds:*

$$\|g\|_{X^{\sigma+1}(\mathbb{R}^2)} + \|\nabla g\|_{X^{\sigma+1}(\mathbb{R}^2)} \leq B. \tag{2.9}$$

The following estimate holds for $L > 1$

$$\|\mathcal{T}(g, g, g)(K) - \mathcal{T}_L(g, g, g)(K)\|_{X_L^\sigma(\mathbb{Z}_L^2)} \lesssim \frac{B^3}{\log L}. \tag{2.10}$$

This theorem is proved in Section 4. As it stands, estimate (2.10) is sharp in terms of the decay of $(\log L)^{-1}$, however it might be possible to isolate the terms that decay as such, move them to the left-hand-side and obtain better convergence rates. For the purposes of this paper, (2.10) is more than enough.

2.2.3. *Approximation theorems.* We are finally ready to state our approximation theorems. We would like to start with a solution $g(t, \xi)$ of (CR) defined on *an arbitrarily long* time interval $[0, M]$ and construct a solution of (NLS$_\epsilon$) whose dynamics in Fourier space is approximated by that of $g(t, \xi)$. The fact that M is chosen arbitrarily allows to carry "any" nonlinear dynamic of the limit equation (CR) and project it into that of (1.3). However, our heuristic calculation in Subsection 2.1 (precisely equation (2.5)) indicates that the two equations do not evolve with the same time scales. In fact, since the time-scale of (CR) is normalized to 1, the "clock" of (1.3) is much slower and its time-scale can be read from (2.5) to be $T_* = \frac{\zeta(2)L^2}{4\epsilon^2 \log L}$. As a result, we should actually expect to approximate the evolution of (1.3) with that of the rescaled function $g(\frac{t}{T^*}, \xi)$ for $0 \leq t \leq MT^*$ where $g(t, \xi)$ solves (CR) on $[0, M]$. This is the content of the following theorem.

Theorem 2.4. *Fix $\sigma > 2$ and $0 < \gamma < 1$. Suppose that $g(t, \xi)$ is a solution of (CR) over a time interval $[0, M]$ with initial data $g_0 = g(t = 0)$ such that $g_0, \nabla g_0 \in X^{\sigma+1}(\mathbb{R}^2)$. By Theorem 2.8 below, (CR) is globally well-posed for such initial data. Denote*

$$B \stackrel{def}{=} \sup_{t \in [0,M]} \|g(t)\|_{X^{\sigma+1}(\mathbb{R}^2)} + \|\nabla g(t)\|_{X^{\sigma+1}(\mathbb{R}^2)}.$$

There exists $c = c(\gamma)$ such that if $L > 1$ and $\epsilon < cL^{-1-\gamma}B^{-1}$, the solution $\widetilde{a}_K(t)$ of (1.3) with initial data $\widetilde{a}_K(0) \stackrel{def}{=} g_0(K)$ satisfies the following estimate:

$$\|\widetilde{a}_K(t) - g(\frac{t}{T^*}, K)\|_{X_L^\sigma(\mathbb{Z}_L^2)} \lesssim \frac{B}{(\log L)^{1-\gamma}} \qquad \text{with } T^* \stackrel{def}{=} \frac{\zeta(2)L^2}{2\epsilon^2 \log L}, \tag{2.11}$$

for all $0 \leq t \leq T^* \min\left(M, \frac{\gamma \log \log L}{cB^2}\right)$.

In particular, if L is chosen so that $L \geq \exp \exp(c\gamma^{-1}MB^2)$, then (2.11) holds for all $t \in [0, MT^*]$ (the full domain of definition of $g(\frac{t}{T^*})$).

Note that by the approximation interval in (2.11) corresponds for the limiting equation (CR) to the interval $\left[0, \min\left(M, \frac{\gamma \log \log L}{cB^2}\right)\right]$, which means that if L is large enough, one can project the full dynamics of $g(t)$ over the interval $[0, M]$ into that of (NLS$_\epsilon$).

Remark 2.5. *How optimal is the condition on L and ϵ in the statement of the theorem? This question leads to distinguishing several natural regimes in the weakly nonlinear, large-box limit for (NLS):*

- If $\epsilon^2 L \to \infty$, it is natural to expect that the nonlinear dynamics will be effectively described (upon appropriate rescaling) by (NLS) on $\mathbb{R}^2$. Indeed, for data of size one, localized around $B(0,1)$ in space and frequency, the nonlinear time-scale[6] is of order ϵ^{-2} while the wave travels at a speed $O(1)$. Thus, it takes a time L before the wave can tell the difference between $\mathbb{R}^2$ and $\mathbb{T}_L^2$. Therefore, if $\frac{1}{\epsilon^2} \ll L$, the effective limiting equation is (NLS) on $\mathbb{R}^2$.
- The regime described in the above theorem is essentially $\epsilon L \to 0$. It seems optimal for the type of result we prove: indeed, the nonlinear time scale is $\frac{1}{\epsilon^2}$, and the necessary time to average over non-resonant interactions is L^2 (which also corresponds to the period of the linear Schrödinger equation). If $L^2 \ll \frac{1}{\epsilon^2}$, non-resonant interactions are averaged out before the nonlinear time-scale is reached, and only the resonant ones contribute to the dynamics. This numerology will become even more transparent once we perform the normal form analysis in the proof of Theorem 2.4.
- There remains the regime $\frac{1}{L} < \epsilon < \frac{1}{\sqrt{L}}$. What should be expected there is not very clear, but it seems likely that not only exactly resonant interactions should play a significant role.

It is easy by a scaling argument to translate the above theorem into a result for (NLS) on $\mathbb{T}^2$ in the weakly non-linear, high frequency regime. Here, one can interpret (CR) as an equation for the frequency envelopes of (NLS). Starting with solutions of (CR), we will be constructing solutions of (NLS) having size ~ 1 in some Sobolev space $H^s(\mathbb{T}^2)$ with $s > 1$, and describing the dynamics in that space as well. The parameter ϵ which is comparable to size of the initial data in $L^2(\mathbb{T}^2)$ will be used to normalize the data under consideration in $H^s(\mathbb{T}^2)$.

Theorem 2.6. *Fix $s > 1$ and $0 < \gamma < 1$. Suppose that $g(t, \xi)$ is a solution of (CR) over a time interval $[0, M]$ with initial data $g_0(\xi) = g(0, \xi)$ such that $g_0, \nabla g_0 \in X^\sigma(\mathbb{R}^2)$ with $\sigma > s+2$. Recall that (CR) is globally well-posed for such initial data by Theorem 2.8. Denote by*

$$B \overset{def}{=} \sup_{t \in [0,M]} \|g(t)\|_{X^\sigma(\mathbb{R}^2)} + \|\nabla g\|_{X^\sigma(\mathbb{R}^2)}.$$

Let $N > 1 + CB^{\frac{2}{s-1}}$ and $v(t)$ be the solution to the cubic NLS equation

$$-i\partial_t v + \Delta v = \mu |v|^2 v, \quad v(0, x) = v_0(x), \quad \mu = \pm 1, \quad x \in \mathbb{T}^2, \tag{2.12}$$

where the Fourier transform of the initial data is given by

$$\widehat{v_0}(k) = \frac{1}{N^{s+1}} g_0(\frac{k}{N}) \quad \forall k \in \mathbb{Z}^2, \tag{2.13}$$

(consequently $\|v(0)\|_{H^s(\mathbb{T}^2)} \sim \|\langle \xi \rangle^s g_0\|_{L^2(\mathbb{R}^2)} \sim 1$ uniformly in N). Then the following estimate holds:

$$\|\widehat{v}(k, t) - e^{4\pi^2 i |k|^2 t} N^{-1-s} \mu g(\frac{t}{T_N}, \frac{k}{N})\|_{\ell^{2,s}(2\pi \mathbb{Z}^2)} \lesssim_s \frac{B}{(\log N)^{1-\gamma}}, \qquad with \ T_N = \frac{\zeta(2) N^{2s}}{2 \log N} \tag{2.14}$$

for all $0 \le t \le T_N \min\left(M, \frac{\gamma \log \log N}{cB^2}\right)$.

---

[6]Prior to this time-scale, the solution behaves linearly, and thus remains almost constant in the interaction-representation picture (i.e. constant linear profile).

In particular, if $N \geq \exp\exp(c\gamma^{-1}MB^2)$, then (2.14) holds for all $t \in [0, MT_N]$ (the full domain of definition of $g(\frac{t}{T_N})$).

The condition on $s > 1$ is the translation of the condition $\epsilon < L^{-1-\gamma}$ in Theorem 2.4. Notice that as we increase N, v_0 carries more and more refined information about g_0 and the approximation becomes better and better. The Sobolev norm of $v(t)$ is uniformly comparable to $\|\langle\xi\rangle^s g(\frac{t}{T_N}, \xi)\|_{L^2(\mathbb{R}^2)}$ for all $0 \leq t \leq T_N M$. As in Theorem 2.4, the approximation interval is long enough to carry all the nonlinear behavior of $g(t)$ on the interval $[0, M]$ of interest. This shows that we can find solutions to the (NLS) equation (2.12) in $H^s(\mathbb{T}^2)$ over very large times (both in the focusing and defocusing case) with rather interesting nonlinear behaviors inherited from (CR). Note, however, that these solutions are small in $H^r(\mathbb{T}^2)$ for $r < s$ (Their L^2 norm is $\sim N^{-s} \ll 1$ which guarantees global existence), and large in $H^r(\mathbb{T}^2)$ for $r > s$. The only "non-trivial" dynamics described in the limit of large N (i.e. uniformly in N) happens in $H^s(\mathbb{T}^2)$.

A striking application of this corollary happens when the data is taken to be any of the explicit solutions of (CR) that will be discussed below. The simplest of which is when we take $g_0 = G(x) = \frac{1}{\sqrt{\pi}}e^{-\frac{1}{2}|x|^2}$. We will see that the solution $g(t, x)$ is given explicitly by $g(t, x) = e^{i\omega_0 t}G(x)$ for some *non-zero* constant ω_0. In this case, B is uniformly bounded by a universal constant independent of M. The previous corollary can be interpreted as a stability result saying that the solution of the (NLS) equation (2.12) starting with the initial value (2.13) (which is close to a gaussian with small variance in the x-space) remains close in H^s over long times to the function

$$v_N(t, x) = \sum_{k \in \mathbb{Z}^2} e^{2\pi i k \cdot x} e^{i(4\pi^2 |k|^2 t + C_N t)} \frac{1}{N^{s+1}} G(\frac{k}{N}), \qquad (2.15)$$

where the phase shift $C_N = \omega_0 T_N^{-1}$ is very small (of order N^{-2s}), but its effect is felt for times $t \geq T_N$. In other words, the function $v(t, x)$ behaves like a solution that is almost periodic in time, and is of size one in H^s. Note that in particular at times $t \in (2\pi)^{-1}\mathbb{Z}$ the solution (2.15) focuses again to a Gaussian (up to unimodular factor), while for the other times, the solution is not concentrated anymore in the x-space. This is observed numerically as well. The nonlinear effect is only exhibited via the modulation factor $e^{iC_N t}$ with $C_N = \omega_0 T_N^{-1}$. Since $\omega_0 \neq 0$, this gives an explicit example of a dynamic of the cubic NLS equation on $\mathbb{T}^2$ that violates (1.2) at the tractable time scales.

Let us conclude by mentioning that this result can be compared with the long-time Sobolev stability of plane wave solutions[7] of (NLS) in [22], but note that the construction is radically different: Plane waves in the defocusing case are of order 1 in H^s for *all* H^s, $s \geq 0$, and their stability comes from the creation of generically non-resonant frequencies around the orbit of the plane wave. The same game could be played with the other explicit solutions of g in Section 6.

2.3. **Analysis of equation (CR).** It turns out that equation (CR) has a surprisingly rich structure in terms of its Hamiltonian nature, symmetries, explicit solutions, and global well-posedness properties. We describe them in this section.

---

[7]These are the solutions of the form $Ae^{i(2\pi n \cdot x + 4\pi^2 |n|^2 t \pm |A|^2 t)}$.

2.3.1. *Hamiltonian structure.* Equation (CR) is Hamiltonian: it derives from the Hamiltonian functional

$$\mathcal{H}(f) = \frac{1}{4} \int_{-1}^{1} \int_{\mathbb{R}^2} \int_{\mathbb{R}^2} f(x+z) f(\lambda x^\perp + z) \overline{f(x + \lambda x^\perp + z)} \, \overline{f(z)} \, dx \, dz \, d\lambda \qquad (2.16)$$

with the symplectic form on $L^2(\mathbb{R}^2)$ given by

$$\omega(f,g) = \mathfrak{Im} \int_{\mathbb{R}^2} f(x) \overline{g(x)} \, dx.$$

A small computation performed in Section 6 reveals that the Hamiltonian functional can also be written

$$\mathcal{H}(f) = \frac{\pi}{2} \int_{\mathbb{R}} \int_{\mathbb{R}^2} |e^{it\Delta_{\mathbb{R}^2}} \widehat{f}|^4 \, dx \, dt. \qquad (2.17)$$

Thus $\mathcal{H}(f)$ is simply a multiple of the $L^4_{t,x}$ Strichartz norm on $\mathbb{R}^2$! While $\mathcal{H}(f)$ can be easily seen, from (2.16) and Cauchy-Schwartz, to be bounded on $L^2(\mathbb{R}^2)$ in the sense that

$$\mathcal{H}(f) \lesssim \|f\|_{L^2(\mathbb{R}^2)}^4 \qquad (2.18)$$

this estimate is nothing but the $L^4_{t,x}$-Strichartz estimate on $\mathbb{R}^2$ ([56, 61]). The Hamiltonian structure of (CR) is examined in detail in Section 6, let us mention here a few striking facts. In all what follows suppose that $g(t)$ is a solution of (CR).

- Conserved quantities. The coercive conserved quantities are $\mathcal{H}(g)$, $\int |g|^2$, $\int |\nabla g|^2$ and $\int |xg|^2$.
- Group Symmetries. The equation (CR) enjoys a large group of symmetries, including in particular $g(x) \mapsto g(x + x_0)$, $g(x) \mapsto g(x)e^{ix\xi_0}$, $g \mapsto e^{i\lambda x^2} g$, $g \mapsto e^{i\lambda\Delta} g$ for any $x_0, \xi_0 \in \mathbb{R}^2$ and $\lambda \in \mathbb{R}$.
- Invariance under the Fourier transform. The Fourier transform $\widehat{g}(t) = \mathcal{F}_{\mathbb{R}^2} g(t)$ is also a solution of (CR)!
- Invariance of the Eigenspaces of the harmonic oscillator. Recall that the harmonic oscillator $-\Delta + |x|^2$ admits a discrete orthonormal basis of $L^2(\mathbb{R}^2)$ formed of eigenvectors. Each eigenspace $\{E_k\}_{k\in\mathbb{N}}$ is finite dimensional of dimension k. The flow of (CR) leaves each of the eigenspaces of the harmonic oscillator operator $-\Delta + |x|^2$ invariant, i.e. if $g_0 \in E_k$ for some k, then $g(t) \in E_k$ for all $t \in \mathbb{R}$.
- Gaussians are stationary solutions. By the previous point, the ground state of $-\Delta + |x|^2$ is invariant. It is given by $\mathbb{R}e^{-\frac{x^2}{2}}$. Thus, a stationary solution of (CR) is given by $e^{-\frac{x^2}{2} + i\omega_0 t}$ for a constant ω_0.
- Other stationary solutions. We mention here $\frac{e^{iCt}}{|x|}$, for a suitable constant C. This solution corresponds exactly to the Raleigh-Jeans solution of the Kolmogorov-Zakharov equation (cf. Section 2.5). Other explicit solutions are given in Section 6.

The rich structure of symmetries of (CR) is somewhat reminiscent of the Szegö equation studied by Gérard and Grellier [26, 25, 27] on $\mathbb{T}$ and Pocovnicu [52, 53] on $\mathbb{R}$. It turns out that the Szegö equation is completely integrable; however the question of whether one can show that (CR) is completely integrable or not is very interesting and open with remarkable consequences on the long-time behavior of (NLS).

2.3.2. *Analytic properties.* The first step is to describe boundedness properties of $\mathcal{T}$. Recall that

$$\dot{X}^1(\mathbb{R}^2) \stackrel{def}{=} \{f(x) \mid |x|f(x) \in L^\infty(\mathbb{R}^2)\} \tag{2.19}$$

Proposition 2.7. *The trilinear operator $\mathcal{T}$ is bounded from $(L^2)^3$ to L^2, and from $(\dot{X}^1)^3$ to $\dot{X}^1$.*

This statement easily gives obvious local well-posedness results in L^2 and $\dot{X}^1$; combining it with the conservation laws of (CR) as well as its invariance with respect to the Fourier transform, one obtains the following global well-posedness theorem.

Theorem 2.8.
- *The equation (CR) is locally well-posed in $\dot{X}^1(\mathbb{R}^2)$, and globally well-posed in $L^2(\mathbb{R}^2)$.*
- *(CR) is globally well-posed in $H^s(\mathbb{R}^2)$ and $L^{2,s}(\mathbb{R}^2)$ for any $s \geq 0$.*
- *(CR) is globally well-posed in $X^\sigma(\mathbb{R}^2)$ for any $\sigma > 2$. Moreover, if $\nabla g(0) \in X^\sigma$, then $g, \nabla g \in C_{\mathrm{loc}}(\mathbb{R}; X^\sigma)$.*

Further well-posedness results are mentioned in Theorem 7.5. Next, we observe that the trilinear operator $\mathcal{T}$ has a remarkable weak boundedness property, reminiscent of the div-curl lemma [48, 60]:

Theorem 2.9. *Assume that $\{f^n\}_{n\in\mathbb{N}}$, $\{g^n\}_{n\in\mathbb{N}}$ and $\{h^n\}_{n\in\mathbb{N}}$ are sequences converging weakly in L^2 to f, g and h respectively. Then the sequence $\{\mathcal{T}(f^n, g^n, h^n)\}_{n\in\mathbb{N}}$ converges weakly in L^2 to $\mathcal{T}(f, g, h)$.*

This implies a weak compactness result for solutions of (CR).

Theorem 2.10. *Given a sequence $\{f^n\}_{n\in\mathbb{N}}$ of solutions of (CR), uniformly bounded in the space $L^\infty([0,T], L^2)$ for some $T > 0$, there exists a solution f of (CR) towards which a subsequence of $\{f^n\}_{n\in\mathbb{N}}$ converges in the weak-* topology of $L^\infty([0,T], L^2)$.*

Theorems 2.7–2.10 are proved, along with other results, in section 7.

2.3.3. *Variational properties.* The last of the properties of (CR) that we mention here relates to the variational characterization of the Gaussian family of explicit solutions in relation to the Hamiltonian function $\mathcal{H}(f)$. The theorem below is not new given the realization that $\mathcal{H}(f)$ can be equivalently written as (2.17) and the study of extremizers of Strichartz estimates in [23, 34]. Still, we chose to give a full proof in order to keep the paper as self-contained as possible. It relies on the heat flow monotonicity following ideas in [11, 4, 2].

Theorem 2.11. *For a fixed mass $\|f\|_{L^2} = M$, (2.18) is only saturated by Gaussians with mass M, up to the finite group of symmetries of $\mathcal{H}$. Furthermore, the optimal implicit constant in the (2.18) is $\frac{\pi}{8}$.*

Using in addition the profile decomposition for the L^4 Strichartz norm [47] gives compactness of maximizing sequences and the L^2−orbital stability of the Gaussian family.

This last result and Theorem 2.11 are proved in section 8.

2.4. **Physical-space interpretation.** Here we try to give a heuristic explanation why the limiting equation (CR) is natural; this could play a role in the dynamics beyond the regime $\epsilon < L^{-1}$ mentioned above.

Let us first approach the equation in Fourier space. The previous arguments can then be recast as follows. For L finite, the (discrete) resonant system derives from the Hamiltonian $\mathcal{H}_L(f) = \frac{1}{4}\langle \mathcal{T}_L(f, f, f), f \rangle$. We saw that $\mathcal{T}_L \to \mathcal{T}$ as $L \to \infty$, therefore $\mathcal{H}_L$ converges to $\mathcal{H} = \frac{1}{4}\langle \mathcal{T}(f, f, f), f \rangle$, and the equation deriving from $\mathcal{H}$ is (CR).

It is perhaps physically more illuminating to examine the convergence to (CR) in physical space. There are then two key elements to be taken into account.

(1) The flow of the linear Schrödinger equation is periodic on $\mathbb{T}_L^2$ with period $T = L^2$ (the situation is actually more complicated: the period is actually shorter depending on the considered frequency, and this is ultimately responsible for the $\zeta(2)$ factor appearing in our derivation. We will however gloss over this additional difficulty).

(2) Let S be the scattering operator for (NLS) on $\mathbb{R}^2$ (This is the operator that maps the profile at $t = -\infty$ to the profile at $t = +\infty$), and denote by $\widetilde{S} = \mathcal{F}S\mathcal{F}^{-1}$ its Fourier space incarnation. Under appropriate assumptions, and for a real constant C, $\widetilde{S}$ can be expanded around 0 as

$$\widetilde{S}f = f + Ci\mathcal{T}(f, f, f) + \text{higher order terms in } f \qquad (2.20)$$

as can be seen by a small computation from Duhamel's formula.

Let $v(t)$ be a solution of (NLS) (with $\epsilon = 1$) and let $f(t, K) := e^{-it|K|^2}\widehat{v}(t, K)$ be the Fourier transform of its linear profile. We can now explain why (CR) comes out as a natural limiting system governing the dynamics of f. Suppose that at the initial time $t = 0$, the data $f(0, K)$ is small and localized (so that all (NLS) solutions considered are global). Since L is very large, the corresponding solution $v(t)$ of (NLS) on $\mathbb{T}_L^2$ will initially evolve like a solution of (NLS) on $\mathbb{R}^2$: it will scatter, and be well-approximated by the linear flow; thus for t sufficiently large the linear profile will keep a nearly constant value given in Fourier space by $f\left(\frac{T}{2}\right)$... until the period T when the linear flow will focus v again. The nonlinear interaction will then come into play, and since L is very large it will essentially amount to applying the scattering operator $\widetilde{S}$. Thus, essentially, $f\left(\frac{3T}{2}\right) \simeq \widetilde{S}f\left(\frac{T}{2}\right)$. The same reasoning gives for any integer n that $f\left(\frac{T}{2} + nT\right) \simeq \widetilde{S}f\left(\frac{T}{2} + (n-1)T\right)$. Using in addition the smallness of f and (2.20), this yields

$$f\left(\frac{T}{2} + nT\right) - f\left(\frac{T}{2} + (n-1)T\right) \simeq Ci\mathcal{T}(f, f, f)\left(\frac{T}{2} + (n-1)T\right),$$

which, rescaling time, can be approximated by (CR).

2.5. **Relation to weak turbulence.** Weak turbulence, also known as wave turbulence, seeks a statistical description[8] of the out-of-equilibrium dynamics of equations ranging from (NLS_ϵ) to water waves or Vlasov-Maxwell. The fundamental equations of weak turbulence theory, sometimes called the Kolmogorov-Zakharov equations, appeared early on in the past century [51, 32, 33]; however the theory was most invigorated after the work of Zakharov in the 60s [62, 63], who sought to parallel Kolmogorov's theory of hydrodynamic turbulence.

The paradigm of kinetic weak turbulence can be briefly summarized as follows. We refer the reader to the classical works of Zakharov, L'vov and Falkovich [64] and Nazarenko [49], as well as to Spohn [55] for more precise information. Roughly speaking, one starts with initial data $a_K(0) = \sqrt{J_K(\omega)}\phi_K(\omega)$ where $J_K(\omega) \geq 0$ and $\phi_K(\omega) \in \mathbb{S}^1$ are random variables such

---

[8]More precisely, this refers to kinetic wave turbulence, as opposed to more recent theories like *discrete* or *mesoscopic* wave turbulence.

that $\mathbb{E}\phi_k\overline{\phi_L} = \delta_{J=K}$: this is the *random phase approximation*. Then, one tries to write down an equation for the *wave spectrum* $n_K(t) = \mathbb{E}|a_K|^2 = \mathbb{E}J_K$. The equation of $n_K(t)$ is obtained in three steps: 1) Performing a sequence of statistical and time averages[9], 2) taking the large-box limit $L \to \infty$ thus making $n_K(t)$ a field on $\mathbb{R}^2$, and 3) taking the weak non-linearity limit $\epsilon \to 0$ which has the effect of restricting to the resonant manifold in the obtained equation. The latter is called the Kolmogorov-Zakharov equation and, for (NLS_ϵ), it often appears in the physics literature in the following form [64, 49]:

$$\partial_t n_K(t) = \epsilon^4 \iiint n(K_1)n(K_2)n(K_3)n(K)\left(\frac{1}{n(K_1)} - \frac{1}{n(K_2)} + \frac{1}{n(K_3)} - \frac{1}{n(K)}\right) \quad \text{(KZ)}$$
$$\delta(K_1 - K_2 + K_3 - K)\,\delta(|K_1|^2 - |K_2|^2 + |K_3|^2 - |K|^2)dK_1\,dK_2\,dK_3.$$

The derivation of this equation is very formal in the physics literature and a rigorous justification seems to be a difficult open question of great importance both in physics and mathematics (see however Lukkarinen and Spohn [44] for first steps in this direction). The importance of the (KZ) equation is mostly due to the existence of four explicit power-type stationary solutions which are a manifestation of the equilibrium and out-of-equilibrium frequency dynamics of (NLS_ϵ). Two of the those solutions can be easily seen by inspecting (KZ): namely the *thermodynamic equilibrium solution* $n_K(t) = 1$ and the *Raleigh-Jeans* solution $n_K(t) = |K|^{-2}$. The other two solutions are the interesting ones from the point of view of out-of-equilibrium dynamics and correspond to the constant flux of energy through scales (*backward and forward cascade* spectra). Nonetheless, such constant-flux solutions can only be sustained by adding external sources and sinks of energy to the system (say at very high and low frequencies).

There are strong similarities between weak turbulence theory and the ideas followed in the present paper: indeed, we follow the steps 2) and 3) in the above paradigm. Formal similarities are more obvious if we write (CR) in the form

$$-i\partial_t f_K(t) = \iiint f(K_1)\overline{f(K_2)}f(K_3)\delta(K_1 - K_2 + K_3 - K)\,\delta(|K_1|^2 - |K_2|^2 + |K_3|^2 - |K|^2)dK_1\,dK_2\,dK_3$$

(in the above, we think of K, K_1, K_2, K_3 as continuous variables). Also notice that the exact solution $e^{iCt}|K|^{-1}$ of (CR) plays the role of the Raleigh-Jeans solution $n(K) = |K|^{-2}$ of (KZ).

The central difference is that, while we do perform the large-box and small non-linearity limit, no randomness, or averaging over statistical ensembles, is used in the present paper. As a result, the equation (CR) is Hamiltonian, time reversible, and exhibits no trend to equilibrium. This should be contrasted with (KZ), which is not Hamiltonian, and for which a version of the H-theorem applies. Another important difference is in terms of time-scales. It is obvious that the time scale of (KZ) is $O(\epsilon^{-4})$ (called the *kinetic time-scale*, as opposed to the *dynamic time-scale* ϵ^{-2}); whereas the time scale in our case (see (2.5)) was roughly $T^* = \frac{L^2}{\epsilon^2 \log L}$ which is $\ll \epsilon^{-4}$ in the regime we are in. Clearly, an equation like (NLS_ϵ) undergoes a lot of nonlinear dynamics prior to the kinetic time scale ϵ^{-4}. (KZ) seems too course to sense those intermediate time-scales, which (CR) seems to capture.

---

[9]This step is particularly far from being rigorous due to the assumption of "preservation of chaos" i.e. preservation of the random phase approximation which is usually made in the physics literature.

Finally, we remark that it has been recognized in the wave turbulence community that some aspects of the turbulent dynamics are not captured by (KZ) especially starting with the highly-influential works of Majda, Mclaughlin, and Tabak [46, 10]. Several theories were formulated to address these and other concerns [65, 39, 38, 50] and new regimes of wave turbulence emerged like *mesoscopic* and *discrete* wave turbulence [39, 38, 50, 66, 45] in order to account for some finite-size effects that are not captured by the kinetic description mentioned above. We refer the reader to Nazarenko's modern textbook treatment in [49] for more details and references. A particularly relevant theory for our work stands out, namely that of *discrete wave turbulence* (see for example [36, 17, 59, 37, 18, 38, 39, 40]) in which one retains the Hamiltonian nature of the equations and emphasizes the importance of the *"discrete"* exactly resonant interactions on the dynamics, the same perspective and regime we adopt here. It is conceivable that the (CR) equation provides a good model for this framework.

3. Proof of Theorem 2.2: uniform boundedness of the $\mathcal{T}_L$

The purpose of this section is to prove Theorem 2.2 which is of interest in its own right. It gives a uniform bound on the operators $\mathcal{T}_L$ on the spaces $X^\sigma(\mathbb{Z}_L^2)$.

3.1. Sharp Strichartz and multilinear estimates.
Estimates on resonant sums such as (2.7) are directly related to the endpoint[10] Strichartz estimates on $\mathbb{T}^2$ given by

$$\|e^{it\Delta_{\mathbb{T}^2}} P_N \phi\|_{L_{t,x}^4([0,1]\times\mathbb{T}^2)} \leq C(N)\|\phi\|_{L^2(\mathbb{T}^2)}, \tag{3.1}$$

where P_N denotes frequency projection onto frequencies $\lesssim N$. To see the relation, one simply notices that:

$$\|e^{it\Delta_{\mathbb{T}^2}} P_N \phi\|_{L_{t,x}^4([0,(2\pi)^{-1}]\times\mathbb{T}^2)}^4 = \int_0^{(2\pi)^{-1}} \int_{\mathbb{T}^2} \left| \sum_{n\in\mathbb{Z}^2} \widehat{\phi}(n) e^{i(2\pi n\cdot x + 4\pi^2 |n|^2 t)} \right|^4 dxdt$$

$$= \sum_{(k_1,k_2,k_3,k_4)} \int_0^{(2\pi)^{-1}} \int_{\mathbb{T}^2} e^{i2\pi(k_1-k_2+k_3-k_4)\cdot x + 4\pi^2 i(|k_1|^2-|k_2|^2+|k_3|^2-|k|^2)t} \widehat{\phi}(k_1)\overline{\widehat{\phi}(k_2)}\widehat{\phi}(k_3)\overline{\widehat{\phi}(k)} \, dx \, dt$$

$$= \sum_{k\in\mathbb{Z}^2} \sum_{\mathcal{R}(k)} \widehat{\phi}(k_1)\overline{\widehat{\phi}(k_2)}\widehat{\phi}(k_3)\overline{\widehat{\phi}(k)} = \left\langle \sum_{\mathcal{R}(k)} \widehat{\phi}(k_1)\overline{\widehat{\phi}(k_2)}\widehat{\phi}(k_3), \widehat{\phi}(k) \right\rangle_{\ell^2\times\ell^2}.$$

Bourgain initiated the systematic study of such estimates and proved in [5] that the optimal $C(N)$ is $\lesssim_\epsilon N^\epsilon$ for any positive ϵ. More precisely, using the divisor bound on integral domains and a counting argument, he shows that

$$C(N) \lesssim \exp\left(\frac{\log N}{\log\log N}\right) \lesssim_\epsilon N^\epsilon. \tag{3.2}$$

A lower bound on $C(N)$ is implicit in Bourgain's work [5, 6] (but can be found explicitly in [42]), and is given by

$$C(N) \gtrsim (\log N)^{1/4}.$$

This is in sharp contrast with the analogous situation when $\mathbb{T}^2$ is replaced by $\mathbb{R}^2$ in which case $C(N)$ is bounded uniformly in N. The analogue of Theorem 2.2 with X_L^σ replaced by $\ell_L^{2,\sigma}$ is

---

[10]We call this estimate endpoint because $p = 4$ is the smallest $L_{t,x}^p$ exponent in (3.1) where one might hope for an upper bound on $C(N)$ in (3.1) independent of N. This is indeed the case on $\mathbb{R}^2$ where $p = 4$ is also the minimum exponent where any estimate of the form (3.1) can hold due to Knapp's counter-example.

effectively equivalent (see proof of Corollary 3.2 below) to proving that $(\log N)^{1/4}$ is also an upper bound (up to a constant factor) for the optimal $C(N)$. Indeed, based on Theorem 2.2 and Corollary 3.2 below, we conjecture the following

Conjecture 3.1. *For all N sufficiently large and $\phi \in L^2(\mathbb{T}^2)$, it holds that:*

$$\|e^{it\Delta_{\mathbb{T}^2}} P_N \phi\|_{L^4_{t,x}([0,1]\times\mathbb{T}^2)} \lesssim (\log N)^{1/4} \|\widehat{\phi}\|_{\ell^2(\mathbb{Z}^2)}. \tag{3.3}$$

While this particular estimate remains open and is widely expected to be quite difficult, it is equivalent to (2.7) with X^σ replaced by the L^2-based weighted spaces $\ell_L^{2,s}$ (the factor of L^2 in the definition (2.6) of the operator $\mathcal{T}_L$ is acquired once one rescales the estimate to the torus $\mathbb{T}_L^2$ and the factor of $\log L$ is the effect of the constant $C(N)$ raised to the forth power). In short, the normalization factor in the definition (2.6) of the operator $\mathcal{T}_L$ in the estimate (2.7) corresponds exactly to the sharp constant of this Strichartz estimate.

Nonetheless, Theorem 2.2 implies the following alternative to (3.3):

Corollary 3.2. *For N sufficiently large and ϕ a function on $\mathbb{T}^2$, the following estimate holds:*

$$\|e^{it\Delta_{\mathbb{T}^2}} P_N \phi\|_{L^4_{t,x}([0,1]\times\mathbb{T}^2)} \lesssim (\log N)^{1/4} N^{1/2} \|\widehat{\phi}\|_{\ell^\infty}^{3/4} \|\widehat{\phi}\|_{\ell^1}^{1/4} \tag{3.4}$$

Moreover, this estimate is sharp in its dependence on N.

The interest of (3.4) as opposed to other variants of (3.3) (like the one obtained from Bourgain's bound (3.2) for $C(N)$) is that it is scales exactly like the conjectured optimal estimate (3.3)[11]. A quick way to see this is to notice that (3.4) would follow from (3.3) using the following scale-invariant estimates $\|\widehat{P_N\phi}\|_{\ell^2} \lesssim N^{1/2}\|\widehat{\phi}\|_{\ell^4} \lesssim N^{1/2}\|\widehat{\phi}\|_{\ell^1}^{1/4}\|\widehat{\phi}\|_{\ell^\infty}^{3/4}$. This scale-invariance turns out to be crucial for our purposes. In fact, the analysis in Section 5 vitally requires us to know the sharp dependence on L in Theorem 2.2 to extent that any less refined analysis that slightly misses the right dependence (namely $L^2 \log L$ appearing in our normalization of $\mathcal{T}_L$ in (2.6)), such as that given in Proposition 3.9 (which yields the dependence $L^{2+\epsilon}$ for any ϵ), is completely useless.

Proof of Corollary 3.2. Without loss of generality, we assume that $\widehat{\phi}$ is supported in the ball $B(0, 2N)$. As we noticed before,

$$\|e^{it\Delta_{\mathbb{T}^2}} P_N \phi\|^4_{L^4_{t,x}([0,(2\pi)^{-1}]\times\mathbb{T}^2)} = \int_0^{(2\pi)^{-1}} \int_{\mathbb{T}^2} \left| \sum_{n\in\mathbb{Z}^2} \widehat{\phi}(n) e^{i2\pi(n\cdot x + 2\pi|n|^2 t)} \right|^4 dx\, dt$$

$$= \sum_{k\in\mathbb{Z}^2} \sum_{\mathcal{R}(k)} \widehat{\phi}(k_1)\overline{\widehat{\phi}(k_2)}\widehat{\phi}(k_3)\overline{\widehat{\phi}(k)} = \sum_{K\in\mathbb{Z}_N^2} \sum_{\mathcal{R}(K)} a_{K_1}\overline{a_{K_2}}a_{K_3}\overline{a_K}$$

where $a_K = \widehat{\phi}(NK)$ for any $K \in \mathbb{Z}_N^2$. Corollary 3.2 now follows by applying Theorem 2.2 to the sequence $\{a_K\}$ and noting that since it is supported in $B(0,2)$ we have that $\|a_K\|_{X_N^\sigma} \sim \|\widehat{\phi}\|_{\ell^\infty}$ and $\sum_K \langle K\rangle^{-\sigma} a_K \sim \sum_K a_K = \|\widehat{\phi}\|_{\ell^1}$. The claim of sharpness follows either from the same example as that in [42] or by simply noticing it as a consequence to Theorem 2.3 proved in the next section. $\square$

We now proceed with the proof of Theorem 2.2. The sharpness of the estimate follows by the same argument just given above.

---

[11]In the sense that it translates into an estimate on $\mathbb{T}_L^2$ of the same form with ℓ^p replaced by ℓ_L^p and N replaced by LN.

3.2. **Proof of Theorem 2.2.** Without loss of generality, we assume that $\|a_K\|_{X_L^\sigma} = \|b_K\|_{X_L^\sigma} = \|c_K\|_{X_L^\sigma} = 1$. We first parameterize $\mathcal{R}(K)$ as in (2.2) and write

$$\sum_{\mathcal{R}(K)} a_{K_1} b_{K_2} c_{K_3} = \sum_{\substack{N_1, N_3 \in \mathbb{Z}_L^2 \\ N_1 \cdot N_3 = 0}} a_{K+N_1} b_{K+N_1+N_3} c_{K+N_3}$$

$$= a_K \sum_{N \in \mathbb{Z}_L^2} b_N c_N + c_K \sum_{N \in \mathbb{Z}_L^2} a_N b_N - a_K b_K c_K + \sum_{\substack{N_1, N_3 \in (\mathbb{Z}_L^2)^* \\ N_1 \cdot N_3 = 0}} a_{K+N_1} b_{K+N_1+N_3} c_{K+N_3}.$$

The estimate on the first three terms above is a triviality, so we only focus on proving that

$$\left\| \sum_{\substack{N_1, N_3 \in (\mathbb{Z}_L^2)^* \\ N_1 \cdot N_3 = 0}} a_{K+N_1} b_{K+N_1+N_3} c_{K+N_3} \right\|_{X_L^\sigma} \lesssim L^2 \log L.$$

Now using the parametrization (2.4) of the set $\{N_1, N_3 \in \mathbb{Z}_L^{2\,*} : N_1 \cdot N_3 = 0\}$, we write the above sum as

$$\sum_{\substack{N_1, N_3 \in (\mathbb{Z}_L^2)^* \\ N_1 \cdot N_3 = 0}} a_{K+N_1} b_{K+N_1+N_3} c_{K+N_3} = \sum_{\substack{\alpha \geq 1 \\ \beta \in \mathbb{Z}^*}} \sum_{J \in \mathbb{Z}_L^2 \text{ visible}} a_{K+\alpha J} b_{K+\alpha J+\beta J^\perp} c_{K+\beta J^\perp}.$$

Next we split the sum in α and β into four pieces $\{1 \leq \alpha \leq \beta\}, \{1 \leq \alpha \leq -\beta\}, \{1 \leq \beta \leq \alpha\}$ and $\{1 \leq -\beta \leq \alpha\}$. Using the symmetries in the above sum (like $J \leftrightarrow -J$, $J \leftrightarrow J^\perp$), it is sufficient to prove the estimate for the first piece, namely:

$$\left\| \sum_{1 \leq \alpha \leq \beta} \sum_{J \in \mathbb{Z}_L^2 \text{ visible}} a_{K+\alpha J} b_{K+\alpha J+\beta J^\perp} c_{K+\beta J^\perp} \right\|_{X_L^\sigma} \lesssim L^2 \log L.$$

This is equivalent to showing that for any $\sigma > 2$ and any $K \in \mathbb{Z}_L^2$

$$\langle K \rangle^\sigma \sum_{1 \leq \alpha \leq \beta} \sum_{J \in \mathbb{Z}_L^2 \text{ visible}} \langle K + \alpha J \rangle^{-\sigma} \langle K + \alpha J + \beta J^\perp \rangle^{-\sigma} \langle K + \beta J^\perp \rangle^{-\sigma} \lesssim L^2 \log L. \qquad (3.5)$$

Recall the notation $\mathcal{B}(m) = [-\frac{m}{2}, \frac{m}{2})^2$. If $\gamma \in \mathbb{N}$, then any $N \in \mathbb{Z}_L^2$ can be written as $r_\gamma + \gamma Q_\gamma$ for some $r_\gamma \in \mathcal{B}(\gamma/L) \cap \mathbb{Z}_L^2$ and $Q_\gamma \in \mathbb{Z}_L^2$. This allows to write for any α, β of the sum (3.5)

$$K = r_\alpha + \alpha Q_\alpha; \qquad K = r_\beta + \beta Q_\beta^\perp. \qquad (3.6)$$

Note that we have $|r_\alpha| \leq \frac{\alpha}{\sqrt{2}L}$ and $|r_\beta| \leq \frac{\beta}{\sqrt{2}L}$. The proof of (3.5) turns out to be particularly sensitive to two values of J that we treat independently, namely $J = -Q_\alpha$ and $J = -Q_\beta$. So we denote by $S \overset{def}{=} \{-Q_\alpha, -Q_\beta\}$, and we split the sum in J into two pieces

$$\sum_{J \in \mathbb{Z}^2 \text{ visible}} = \sum_{J \notin S \text{ visible}} + \sum_{J \in S \text{ visible}}.$$

and treat each case independently. We start with the first case.

Case 1: Contribution of the $J \notin S$ Here we prove that

$$\langle K \rangle^\sigma \sum_{1 \leq \alpha \leq \beta} \sum_{J \in \mathbb{Z}_L^2 \setminus S \, visible} \langle K + \alpha J \rangle^{-\sigma} \langle K + \alpha J + \beta J^\perp \rangle^{-\sigma} \langle K + \beta J^\perp \rangle^{-\sigma} \lesssim L^2 \log L. \qquad (3.7)$$

We further split the analysis into two cases: $|K| \leq 100$ and $|K| \geq 100$.

Subcase 1(a): $|K| \leq 100$. Here we only need to show that

$$\sum_{1 \leq \alpha \leq \beta} \sum_{\substack{J \in \mathbb{Z}_L^2 \setminus S \\ visible}} \langle K + \alpha J \rangle^{-\sigma} \langle K + \alpha J + \beta J^\perp \rangle^{-\sigma} \langle K + \beta J^\perp \rangle^{-\sigma} \lesssim L^2 \log L. \qquad (3.8)$$

This follows from the estimate

$$\sum_{1 \leq \alpha \leq \beta} \sum_{\substack{J \in \mathbb{Z}_L^2 \setminus S \\ visible}} \langle K + \alpha J + \beta J^\perp \rangle^{-\sigma} \langle K + \beta J^\perp \rangle^{-\sigma} \lesssim L^2 \log L,$$

which is proved in Lemma 3.5. To prove Lemma 3.5, we need the following elementary inequality, which we isolate in a lemma since we will use it often:

Lemma 3.3. *Let* $0 \leq \alpha \leq \beta$ *be non-negative integers with* β *positive, and* $\sigma > 2$. *Let* $r \in \mathbb{R}^2$ *be such that* $|r| \leq \frac{3}{4} \frac{\sqrt{\alpha^2 + \beta^2}}{L}$. *Then the following estimate holds:*

$$\sum_{J \in (\mathbb{Z}_L^2)^*} \langle r + \alpha J + \beta J^\perp \rangle^{-\sigma} \lesssim \min(\frac{L^2}{\beta^2}, \frac{L^\sigma}{\beta^\sigma}) = \begin{cases} \frac{L^2}{\beta^2} & \text{if } \beta \leq L, \\ \frac{L^\sigma}{\beta^\sigma} & \text{if } \beta \geq L. \end{cases} \qquad (3.9)$$

Proof. In the sum (3.9), J runs in $(\mathbb{Z}_L^2)^*$ and can be written j/L for some $j \in \mathbb{Z}^2$. Now we have $|r + \alpha J + \beta J^\perp| \geq \sqrt{\alpha^2 + \beta^2}|J| - |r| \geq \frac{1}{4L}\sqrt{\alpha^2 + \beta^2}|j| \geq \frac{\beta}{4L}|j|$. Hence the previous sum can always be estimated by

$$\sum_{j \in (\mathbb{Z}^2)^*} \left\langle \frac{\beta|j|}{4L} \right\rangle^{-\sigma} \lesssim \begin{cases} \frac{L^\sigma}{\beta^\sigma}\big(\sum_{j \in (\mathbb{Z}^2)^*} \frac{1}{|j|^\sigma}\big) \lesssim \frac{L^\sigma}{\beta^\sigma} & \text{if } \beta \geq L, \\ \frac{L^2}{\beta^2} + \frac{L^\sigma}{\beta^\sigma}\big(\sum_{\substack{j \in (\mathbb{Z}^2)^* \\ |j| \geq L\beta^{-1}}} \frac{1}{|j|^\sigma}\big) \lesssim \frac{L^2}{\beta^2} & \text{if } \beta \leq L, \end{cases}$$

as $\sigma > 2$. $\qquad \square$

Remark 3.4. *Note that when* $\beta \leq L$, *the previous estimate by* L^2/β^2 *is still valid when the sum* (3.9) *is taken over all the point of lattice, including the origin* $J = (0,0)$.

Now (3.8) follows from the following lemma.

Lemma 3.5. *For* $L \geq 5$ *and any* $K \in \mathbb{Z}_L^2$, *there holds*

$$\sum_{1 \leq \alpha \leq \beta} \sum_{\substack{J \in \mathbb{Z}_L^2 \setminus S \\ visible}} \langle K + \alpha J + \beta J^\perp \rangle^{-\sigma} \langle K + \beta J^\perp \rangle^{-\sigma} \lesssim L^2 \log L. \qquad (3.10)$$

Proof of Lemma 3.5. Since $J \neq -Q_\beta$, then by Lemma 3.3 with $\alpha = 0$, we have that

$$\sum_{J \in \mathbb{Z}_L^2 \setminus S \, visible} \langle K + \alpha J + \beta J^\perp \rangle^{-\sigma} \langle K + \beta J^\perp \rangle^{-\sigma} \lesssim \min(\frac{L^2}{\beta^2}, \frac{L^\sigma}{\beta^\sigma})$$

from which we deduce that the sum in (3.10) can be estimated up to a constant by

$$\sum_{\substack{1\leq\alpha\leq\beta\\\beta\leq L}}\frac{L^2}{\beta^2}+\sum_{\substack{1\leq\alpha\leq\beta\\\beta\geq L}}\frac{L^\sigma}{\beta^\sigma}\leq L^2\sum_{1\leq\beta\leq L}\frac{1}{\beta}+L\sum_{\beta\geq L}\left(\frac{L}{\beta}\right)^{\sigma-1}\lesssim L^2\log L+L^2,$$

from which we easily deduce the result.

Remark 3.6. *Notice from the previous equation that if we restrict the sum in (3.8) to $\beta\geq L$, then we can estimate it by L^2 rather than $L^2\log L$.*

Subcase 1(b): $|K|\geq 100$. We start by noting that we can restrict the sum in J to the case $|K+\alpha J|\leq\frac{|K|}{100}$. In fact, if $|K+\alpha J|\geq\frac{|K|}{100}$, then estimate (3.5) follows directly from Lemma 3.5 above.

The restriction $|K+\alpha J|\leq\frac{|K|}{100}$ has several implications that we enumerate below:

(1) First we notice that $|\alpha J|\geq\frac{99}{100}|K|$ and consequently $|\beta J^\perp|\geq\frac{99}{100}|K|$ since $\beta\geq\alpha$.
(2) From the orthogonality relation

$$|K|^2+|K+\alpha J+\beta J^\perp|^2=|K+\alpha J|^2+|K+\beta J^\perp|^2$$

we conclude that $|K+\beta J^\perp|\geq\frac{99}{100}|K|$. Consequently, we also have that

$$|K+\alpha J+\beta J^\perp|\geq|\beta J^\perp|-|K+\alpha J|\geq\frac{98}{100}|K|.$$

(3) Since $|K+\alpha J|\leq\frac{1}{100}|K|$, then $\alpha J\in B(-K,\frac{|K|}{100})$ and hence $\beta J^\perp\in B(-\frac{\beta}{\alpha}K^\perp,\frac{\beta}{100\alpha}|K|)$. So $K+\beta J^\perp\in B(K-\frac{\beta}{\alpha}K^\perp,\frac{\beta}{100\alpha}|K|)$. But $|K-\frac{\beta}{\alpha}K^\perp|\geq\frac{\beta}{\alpha}|K|$, so we finally deduce that in fact we have:

$$|K+\beta J^\perp|\geq\frac{99}{100}\frac{\beta}{\alpha}|K|.$$

As a consequence of this analysis, (3.5) follows once we show that

$$\sum_{1\leq\alpha\leq\beta}\sum_{\substack{J\in\mathbb{Z}_L^2\backslash S\\J\,visible}}\langle K+\alpha J\rangle^{-\sigma}\langle K+\beta J^\perp\rangle^{-\sigma}\lesssim L^2\log L. \tag{3.11}$$

For this we use the third point above to estimate as follows:

$$\sum_{1\leq\alpha\leq\beta}\sum_{\substack{J\in\mathbb{Z}_L^2\backslash S\\J\,visible}}\langle K+\alpha J\rangle^{-\sigma}\langle K+\beta J^\perp\rangle^{-\sigma}$$

$$\lesssim\sum_{1\leq\alpha\leq\beta}\left(\sum_{\substack{J\in\mathbb{Z}_L^2\backslash S\\J\,visible}}\langle K+\alpha J\rangle^{-2\sigma}\right)^{1/2}\left(\sum_{\substack{J\in\mathbb{Z}_L^2\backslash S\\|K+\beta J^\perp|\geq\frac{99}{100}\frac{\beta}{\alpha}|K|}}\langle K+\beta J^\perp\rangle^{-2\sigma}\right)^{1/2}$$

$$\lesssim\sum_{1\leq\alpha\leq\beta}\frac{L}{\alpha}\left(\sum_{\substack{J\in\mathbb{Z}_L^2\backslash S\\|K+\beta J^\perp|\geq\frac{99}{100}\frac{\beta}{\alpha}|K|}}\langle K+\beta J^\perp\rangle^{-2\sigma}\right)^{1/2}, \tag{3.12}$$

by Lemma 3.3 and the fact that $J \neq -Q_\alpha$. The estimate on the term in parentheses above should be handled with some extra care. Recalling that $K = r_\beta + \beta Q_\beta^\perp$ with $|r_\beta| \leq \frac{\beta}{\sqrt{2}L}$ and $J \neq -Q_\beta$, we have

$$\sum_{\substack{J \in \mathbb{Z}_L^2 \setminus S \\ |K+\beta J^\perp| \geq \frac{99}{100}\frac{\beta}{\alpha}|K|}} \langle K + \beta J^\perp \rangle^{-2\sigma} \leq \sum_{\substack{J \in (\mathbb{Z}_L^2)^* \\ |r_\beta+\beta J^\perp| \geq \frac{99}{100}\frac{\beta}{\alpha}|K|}} \langle r_\beta + \beta J^\perp \rangle^{-2\sigma} \lesssim \sum_{\substack{J \in (\mathbb{Z}_L^2)^* \\ |\beta J^\perp| \geq \frac{99}{200}\frac{\beta}{\alpha}|K|}} |\beta J|^{-2\sigma}$$

since $|\beta J| = |\beta J^\perp| \geq \frac{1}{2}|r + \beta J^\perp|$. As a consequence, we get that:

$$\sum_{\substack{J \in \mathbb{Z}_L^2 \setminus S \\ |K+\beta J^\perp| \geq \frac{99}{100}\frac{\beta}{\alpha}|K|}} \langle K + \beta J^\perp \rangle^{-2\sigma} \lesssim \beta^{-2\sigma} \sum_{\substack{J \in \mathbb{Z}_L^2 \setminus \{0\} \\ |J^\perp| \geq \frac{99}{200\alpha}|K|}} |J|^{-2\sigma}$$

$$\lesssim \begin{cases} L^{2\sigma}\beta^{-2\sigma} & \text{if } \alpha \geq 100L|K|, \\ L^2\beta^{-2\sigma}(\frac{|K|}{\alpha})^{-2\sigma+2} & \text{if } \alpha \leq 100L|K|. \end{cases}$$

Going back to (3.12), we get that

$$\sum_{\substack{1 \leq \alpha \leq \beta}} \sum_{\substack{J \in \mathbb{Z}_L^2 \setminus S \\ J \text{ visible}}} \langle K + \alpha J \rangle^{-\sigma} \langle K + \beta J^\perp \rangle^{-\sigma} \lesssim \sum_{\substack{1 \leq \alpha \leq \beta \\ \alpha \leq 100L|K|}} \frac{L^2 \alpha^{\sigma-2}}{\beta^\sigma |K|^{\sigma-1}} + \sum_{100L|K| \leq \alpha \leq \beta} \frac{L^{\sigma+1}}{\alpha\beta^\sigma}$$

$$\lesssim \sum_{1 \leq \alpha \leq 100L|K|} \frac{L^2}{\alpha |K|^{\sigma-1}} + \sum_{100L|K| \leq \alpha} \frac{L^{\sigma+1}}{\alpha^\sigma} \lesssim \frac{L^2 \log(L|K|)}{|K|^{\sigma-1}} \lesssim L^2 \log L.$$

which finishes the proof of this case.

Remark 3.7. *Repeating this last estimate with the additional restriction that $\beta \geq L$ we get the improved bound*

$$\sum_{\substack{1 \leq \alpha \leq \beta \\ \beta \geq L}} \sum_{\substack{J \in \mathbb{Z}_L^2 \setminus S \\ J \text{ visible}}} \langle K + \alpha J \rangle^{-\sigma} \langle K + \beta J^\perp \rangle^{-\sigma} \lesssim \sum_{\substack{1 \leq \alpha \leq \beta \\ \alpha \leq 100L|K| \, ; \, \beta \geq L}} \frac{L^2 \alpha^{\sigma-2}}{\beta^\sigma |K|^{\sigma-1}} + \sum_{100L|K| \leq \alpha \leq \beta} \frac{L^{\sigma+1}}{\alpha\beta^\sigma}$$

$$\lesssim \sum_{1 \leq \alpha \leq 100L|K|} \frac{L^2 \alpha^{\sigma-2}}{\max(\alpha, L)^{\sigma-1}|K|^{\sigma-1}} + \sum_{100L|K| \leq \alpha} \frac{L^{\sigma+1}}{\alpha^\sigma} \lesssim L^2 \frac{\log|K|}{|K|^{\sigma-1}} \lesssim L^2.$$

This, along with Remark 3.6, shows that if we restrict the β sum in (3.7) to $\beta \geq L$, we also have the better estimate

$$\langle K \rangle^\sigma \sum_{\substack{1 \leq \alpha \leq \beta \\ \beta \geq L}} \sum_{\substack{J \in \mathbb{Z}_L^2 \setminus S \\ \text{visible}}} \langle K + \alpha J \rangle^{-\sigma} \langle K + \alpha J + \beta J^\perp \rangle^{-\sigma} \langle K + \beta J^\perp \rangle^{-\sigma} \lesssim L^2. \tag{3.13}$$

This will be useful for us later.

Case 2: Contribution of $J \in S$. We now deal with the case when J is equal to one of the forbidden values in $S \overset{def}{=} \{-Q_\alpha, -Q_\beta\}$. Recall that since J is visible, we only consider the cases when $Q_\alpha, Q_\beta \neq 0$. We start with the simpler case $J = -Q_\alpha \neq 0$.

Subcase 2(a): $J = -Q_\alpha$. We will show that for any $K \in \mathbb{Z}_L^2$, it holds that

$$\langle K \rangle^\sigma \sum_{1 \leq \alpha \leq \beta} \langle K - \alpha Q_\alpha \rangle^{-\sigma} \langle K - \alpha Q_\alpha - \beta Q_\alpha^\perp \rangle^{-\sigma} \langle K - \beta Q_\alpha^\perp \rangle^{-\sigma} \lesssim L^2. \tag{3.14}$$

We start with some elementary observations.

(1) Recall that $K = r_\alpha + \alpha Q_\alpha$. Since $Q_\alpha \neq 0$, the α has to be less than $2L|K|$, since otherwise $r_\alpha = K$ and $Q_\alpha = 0$.

(2) We have that $K + \alpha J = r_\alpha$, $K + \beta J^\perp = r_\alpha + \alpha Q_\alpha - \beta Q_\alpha^\perp$, and $K + \alpha J + \beta J^\perp = r_\alpha - \beta Q_\alpha^\perp$.

(3) Notice that since $\alpha \leq \beta$, then $|K + \beta J^\perp| \geq \beta |Q_\alpha| - \frac{\alpha}{\sqrt{2}L} \gtrsim \beta |Q_\alpha|$. Similarly, $|K + \alpha J + \beta J^\perp| \gtrsim \beta |Q_\alpha|$.

(4) Finally, notice that since $Q_\alpha \neq 0$ and $K = r_\alpha + \alpha Q_\alpha$ with $|r_\alpha| \leq \frac{\alpha}{\sqrt{2}L} \leq \frac{1}{\sqrt{2}}\alpha |Q_\alpha|$, then $|Q_\alpha| \sim \frac{|K|}{\alpha}$.

As a result, $|K + \beta J^\perp|, |K + \alpha J + \beta J^\perp| \gtrsim \frac{\beta}{\alpha}|K|$ and estimate (3.5) would be satisfied for $J = -Q_\alpha$ once we prove that

$$\sum_{\substack{1 \leq \alpha \leq \beta \\ \alpha \leq 2L|K|}} \langle r_\alpha \rangle^{-\sigma} \langle \frac{\beta}{\alpha}|K| \rangle^{-2\sigma} \lesssim L^2 \langle K \rangle^{-\sigma}. \tag{3.15}$$

Now we have

$$\sum_{\substack{1 \leq \alpha \leq \beta \\ \alpha \leq 2L|K|}} \langle r_\alpha \rangle^{-\sigma} \langle \frac{\beta}{\alpha}|K| \rangle^{-2\sigma} \lesssim \sum_{1 \leq \alpha \leq \beta} \langle \frac{\beta}{\alpha}|K| \rangle^{-2\sigma}$$

$$\lesssim \left[\sum_{\substack{1 \leq \alpha \leq \beta \leq \alpha|K|^{-1} \\ \alpha \leq 2L|K|}} 1 \right] + |K|^{-2\sigma} \left[\sum_{\substack{1 \leq \alpha \leq \beta \\ \beta \geq \alpha|K|^{-1}, \alpha \leq 2L|K|}} \frac{\alpha^{2\sigma}}{\beta^{2\sigma}} \right],$$

where the first sum is zero if $|K| > 1$. A direct computation now gives the result.

Subcase 2(b): $J = -Q_\beta$. Here the analysis is more delicate and relies heavily on the fact that if $J = -Q_\beta$, then Q_β is a visible lattice point. We will show that

$$\langle K \rangle^\sigma \sum_{1 \leq \alpha \leq \beta} \langle K - \alpha Q_\beta \rangle^{-\sigma} \langle K - \alpha Q_\beta - \beta Q_\beta^\perp \rangle^{-\sigma} \langle K - \beta Q_\beta^\perp \rangle^{-\sigma} \lesssim L^2 \tag{3.16}$$

As before, we start with some useful observations:

(1) Recall that $K = r_\beta + \beta Q_\beta^\perp$. Since $Q_\beta \neq 0$, then $\beta \leq 2L|K|$. Also, since $|r_\beta| \leq \frac{1}{\sqrt{2}}\frac{\beta}{L} \leq \frac{1}{\sqrt{2}}\beta|Q_\beta|$, then $|Q_\beta| \sim \frac{|K|}{\beta}$.

(2) Since $Q_\beta^\perp$ is visible, then $\beta = g.c.d(LK - Lr_\beta)$ (For $m = (m_1, m_2) \in (\mathbb{Z}^2)^*$, we use the notation $g.c.d(m)$ to denote the greatest common divisor of the $|m_1|$ and $|m_2|$).

(3) With $J = -Q_\beta$, we have that $K + \beta J^\perp = r_\beta$, $K + \alpha J = r_\beta - \alpha Q_\beta + \beta Q_\beta^\perp$, $K + \alpha J + \beta J^\perp = r_\beta - \alpha Q_\beta$.

(4) Due to the orthogonality relation,

$$|K|^2 + |r_\beta - \alpha Q_\beta|^2 = |r_\beta|^2 + |r_\beta - \alpha Q_\beta + \beta Q_\beta^\perp|^2,$$

we have that either $|r_\beta| \geq |K|/\sqrt{2}$ or $|r_\beta - \alpha Q_\beta + \beta Q_\beta^\perp| \geq |K|/\sqrt{2}$. But $|r_\beta - \alpha Q_\beta + \beta Q_\beta^\perp| \geq (\sqrt{2} - 1)|r_\beta|$ (since $|r_\beta| \leq \frac{\beta}{\sqrt{2}L} \leq \frac{1}{\sqrt{2}}\beta|Q_\beta^\perp| \leq \frac{1}{\sqrt{2}}| - \alpha Q_\beta + \beta Q_\beta^\perp|$). Therefore, we always have that

$$|r_\beta - \alpha Q_\beta + \beta Q_\beta^\perp| \geq \frac{|K|}{100}.$$

As a consequence of Observation 4, we will be done once we show that

$$\sum_{1 \le \alpha \le \beta \le 2L|K|} \langle r_\beta \rangle^{-\sigma} \langle r_\beta - \alpha Q_\beta \rangle^{-\sigma} \lesssim L^2 \tag{3.17}$$

(5) Next we reduce (3.17) to the case when $\beta \ge 200L$. Indeed,

$$\sum_{1 \le \alpha \le \beta \le 200L} \langle r_\beta \rangle^{-\sigma} \langle r_\beta - \alpha Q_\beta \rangle^{-\sigma} \le \sum_{1 \le \alpha \le \beta \le 200L} 1 \lesssim L^2.$$

Since the complementary range of β is non-empty only when $|K| \ge 100$ (from observation (1) above), we are now reduced to proving that for any $|K| \ge 100$, we have

$$\sum_{\substack{1 \le \alpha \le \beta \\ 200L \le \beta \le 2L|K|}} \langle r_\beta \rangle^{-\sigma} \langle r_\beta - \alpha Q_\beta \rangle^{-\sigma} \lesssim L^2 \tag{3.18}$$

(6) Finally, we can also assume that $|r_\beta| \le |K|/100$ and $|r_\beta - \alpha Q_\beta| \le |K|/100$, since otherwise, one can trivially estimate the left-hand-side of (3.18) by

$$|K|^{-\sigma} \sum_{\substack{1 \le \alpha \le \beta \\ 200L \le \beta \le 2L|K|}} 1 \lesssim L^2 |K|^{2-\sigma} \lesssim L^2.$$

We start by estimating the sum in α in (3.18) by

$$\sum_{1 \le \alpha \le \beta} \langle r_\beta - \alpha Q_\beta \rangle^{-\sigma} \lesssim \min\left(\beta, \sum_{\alpha \ge 0} \langle \alpha Q_\beta \rangle^{-\sigma} \right) \lesssim \begin{cases} \min(\beta, |Q_\beta|^{-1}), & \text{if } |Q_\beta| \le 1 \\ 1 & \text{if } |Q_\beta| \ge 1. \end{cases}$$

As a result, we have that

$$\sum_{\substack{1 \le \alpha \le \beta \\ 200L \le \beta \le 2L|K|}} \langle r_\beta \rangle^{-\sigma} \langle r_\beta - \alpha Q_\beta \rangle^{-\sigma} \lesssim \sum_{\substack{200L \le \beta \le 2L|K| \\ |Q_\beta| \le 1}} \langle r_\beta \rangle^{-\sigma} \min(\beta, |Q_\beta|^{-1}) + \sum_{\substack{200L \le \beta \le 2L|K| \\ |Q_\beta| \ge 1}} \langle r_\beta \rangle^{-\sigma}$$

$$\tag{3.19}$$

The second sum above is easy to estimate: as β ranges over $[200L, 2L|K|]$, r_β varies in the ball $B(0, \frac{|K|}{100})$ (see Observation 6 above). Moreover, for each $r \in B(0, \frac{|K|}{100})$, there is at most one β for which $r = r_\beta$ (in fact, $\beta = g.c.d(LK - Lr)$ by Observation 2. As a result, the second sum on the right-hand-side of (3.19) can be estimated by:

$$\sum_{r \in B(0, \frac{|K|}{100})} \langle r \rangle^{-\sigma} \lesssim L^2.$$

Therefore, we will be done once we show that

$$\sum_{\substack{200L \le \beta \le 2L|K| \\ |Q_\beta| \le 1}} \langle r_\beta \rangle^{-\sigma} \min(\beta, |Q_\beta|^{-1}) \lesssim L^2. \tag{3.20}$$

Recall from Observation 1 above, that $|Q_\beta| \sim \frac{|K|}{\beta}$, which means, since $|K| \ge 100$, that $\min(\beta, |Q_\beta|^{-1}) = |Q_\beta|^{-1} \sim \frac{\beta}{|K|}$. Again, we change variables here from β to r_β due to the

one-to-one correspondence between β and r ($\beta = g.c.d(LK - Lr_\beta)$ since Q_β is visible). Consequently the left-hand side of (3.20) is estimated by

$$|K|^{-1} \sum_{\substack{100L \leq \beta \leq 2L|K| \\ |Q_\beta| \leq 1}} \langle r_\beta \rangle^{-\sigma} \beta \lesssim |K|^{-1} \sum_{\substack{|r| \leq \frac{|K|}{100} \\ L \leq g.c.d(LK-Lr) \leq 2L|K|}} \langle r \rangle^{-\sigma} g.c.d(LK - Lr).$$

Let $k = LK \in \mathbb{Z}^2$, and $m = Lr \in \mathbb{Z}^2$, then the above sum is equal to

$$\frac{L}{|k|} \sum_{\substack{|m| \leq \frac{|k|}{100} \\ L \leq g.c.d(k-m) \leq 2|k|}} \langle \frac{m}{L} \rangle^{-\sigma} g.c.d(k - m) = \frac{L}{|k|} \sum_{\substack{n \in \mathbb{Z}^2, |k-n| \leq \frac{|k|}{100} \\ L \leq g.c.d(n) \leq 2|k|}} \langle \frac{k-n}{L} \rangle^{-\sigma} g.c.d(n)$$

$$\lesssim \frac{L}{|k|} \sum_{j=L}^{2|k|} j \left(\sum_{\substack{n \in \mathbb{Z}^2, |k-n| \leq \frac{|k|}{100} \\ g.c.d(n)=j}} \langle \frac{k-n}{L} \rangle^{-\sigma} \right)$$

We claim that the term in parentheses is bounded by $\frac{L^2}{j^2}$. Assuming this claim, we get that

$$\frac{L}{|k|} \sum_{\substack{|m| \leq \frac{|k|}{100} \\ L \leq g.c.d(k-m) \leq 2|k|}} \langle \frac{m}{L} \rangle^{-\sigma} g.c.d(k-m) \lesssim \frac{L}{|k|} \sum_{j=L}^{2|k|} j \left(\frac{L^2}{j^2} \right) \lesssim L^2 \frac{\log(|k|/L)}{|k|/L} = L^2 \frac{\log|K|}{|K|} \lesssim L^2$$

since $|K| \geq 1$. To prove the above-mentioned claim we notice that if $g.c.d(n) = j$, then n is an element of the lattice $j\mathbb{Z}^2$, and hence

$$\sum_{\substack{n \in B(k, \frac{|k|}{100}) \\ g.c.d(n)=j}} \langle \frac{k-n}{L} \rangle^{-\sigma} \sim \sum_{l=0}^{\log_2(\frac{|k|}{L})} 2^{-\sigma l} \sum_{\substack{n \in B(k, 2^l L) \\ g.c.d(n)=j}} 1 \lesssim \sum_{l=0}^{\log_2(\frac{k}{L})} 2^{-\sigma l} \frac{2^{2l} L^2}{j^2} \lesssim \frac{L^2}{j^2},$$

since $\sigma > 2$. This finishes the proof of Theorem 2.2. $\qquad \square$

Remark 3.8. *Combining Remark 3.7 with estimates (3.14) and (3.16), we obtain that restricting the β sum to $\beta \geq L$, one obtains the better estimate*

$$\left\| \sum_{\substack{1 \leq \alpha \leq \beta \\ \beta \geq L}} \sum_{\substack{J \in \mathbb{Z}_L^2 \\ visible}} a_{K+\alpha J} \, b_{K+\alpha J+\beta J^\perp} \, c_{K+\beta J^\perp} \right\|_{X_L^\sigma} \lesssim L^2 \|a_K\|_{X_L^\sigma} \|b_K\|_{X_L^\sigma} \|c_K\|_{X_L^\sigma}. \qquad (3.21)$$

This will be useful in the following section.

3.3. **Estimates on non-resonant level sets.** In order to effectively bound non-resonant interactions in Section 5, we will need estimates on trilinear sums taken over non-resonant tuples $(K_1, K_2, K_3) \in \mathcal{S}(K)$ that belong to some fixed non-resonant level set of the function $\Omega = |K_1|^2 - |K_2|^2 + |K_3|^2 - |K|^2$, i.e. tuples (K_1, K_2, K_3) belonging to sets of the form

$$\mathcal{R}_\mu(K) \overset{def}{=} \{(K_1, K_2, K_3) \in (\mathbb{Z}_L^2)^3 : K_1 - K_2 + K_3 = K, \text{ and } |K_1|^2 - |K_2|^2 + |K_3|^2 - |K|^2 = \mu\}, \tag{3.22}$$

for $\mu \in \mathbb{R}$. In contrast to the resonant sums (corresponding to $\mu = 0$) treated above, finding the sharpest dependence on L in the estimates on these sums is not as crucial. The following proposition is sufficient for our purposes.

Proposition 3.9. *Let $L \geq 1$ and suppose that $|\mu| \leq L^{10}$, then for any $\epsilon > 0$ and any $\sigma > 2$, the following estimate holds*

$$\left\| \sum_{\mathcal{R}_\mu(K)} a_{K_1} b_{K_2} c_{K_3} \right\|_{X_L^\sigma} \lesssim_{\sigma, \epsilon} L^{2+\epsilon} \|a_K\|_{X_L^\sigma} \|b_K\|_{X_L^\sigma} \|c_K\|_{X_L^\sigma}. \tag{3.23}$$

The proof of this lemma depends on lattice counting arguments similar to those used by Bourgain to prove periodic Strichartz estimates [5, 8]. In particular, we rely on the following lemma, whose proof is implicit in the above cited works.

Lemma 3.10. *Let $C(R)$ denote the circle in $\mathbb{R}^2$ centered at $a \in \mathbb{Z}^2$ of radius R. For any $\epsilon > 0$, the number of integer lattice points on the circle and inside a box of size L is bounded by $C_\epsilon L^\epsilon$ independent of a and R.*

Proof. We split into two cases:

Case 1: $R \leq L^{10}$. In this case, we use the upper bound on the number of lattice points on the circle $C(R)$. If z is such a lattice point, then $(m, n) = z - a \in \mathbb{Z}^2$ and $(m, n) \in \mathbb{Z}^2$ satisfies $R^2 = m^2 + n^2 = (m + in)(m - in)$. Therefore $m + in$ is a divisor of R^2 in the integral domain $\mathbb{Z} + i\mathbb{Z}$, which is bounded by $C_\epsilon R^{\frac{\epsilon}{100}}$ for any $\epsilon > 0$ (see for instance [31, 58]), which is sufficient in this case.

Case 2: $R \geq L^{10}$. Here we will use an idea that goes back to Jarnick [35] relying on the fact that the area of any triangle in $\mathbb{R}^2$ with vertices in the lattice $\mathbb{Z}^2$ is lower bounded by $1/2$ (being $1/2$ of the cross product of two adjacent edges). We will show that the if $L \ll R^{1/3}$, then there cannot be more than two lattice points on the circle $C(R)$ inside a box of size L. To see this, suppose that A, B, C are three such lattice points with $\theta \overset{def}{=} \widehat{AC} < \pi$ and B being on the arc between A and C. Then

$$\frac{1}{2} \leq \text{Area of triangle } ABC \leq \text{Area of the outer cap between segment and the arc described by } AC.$$

As θ is the angle of the sector described by AC, a simple calculation shows that the area of this cap is $\frac{R^2}{2}(\theta - \sin\theta) \lesssim R^2\theta^3$ if θ is small enough. But $\theta \lesssim LR^{-1}$, and hence we get that

$$\frac{1}{2} \lesssim L^3 R^{-1}.$$

This gives the needed contradiction if $R \ll L^{1/3}$. $\square$

Proof of Proposition 3.9. Due to the symmetry between K_1 and K_3, we may assume without any loss of generality that $|K_1| \geq |K_3|$. We split the analysis into two cases:

Case 1: $|K_1| \geq \frac{|K|}{100}$ or $|K| \leq 100$. In this case, (3.23) reduces to showing that

$$\sup_{K \in \mathbb{Z}_L^2} \sum_{\mathcal{R}_\mu(K)} \langle K_2 \rangle^{-\sigma} \langle K_3 \rangle^{-\sigma} \lesssim_\epsilon L^{2+\epsilon}. \qquad (3.24)$$

Now we change variable and define $N_i = K_i - K$ $(i = 1, 2, 3)$ so that the new variables satisfy the relations:

$$N_1 + N_3 = N_2 \qquad 2(N_3 - N_2) \cdot N_3 = |N_1|^2 + |N_3|^2 - |N_2|^2 = \mu.$$

As a result, we have that

$$\sum_{\mathcal{R}_\mu(K)} \langle K_2 \rangle^{-\sigma} \langle K_3 \rangle^{-\sigma} = \sum_{N_2 \in \mathbb{Z}_L^2} \langle K + N_2 \rangle^{-\sigma} \sum_{l=0}^{\infty} 2^{-\sigma l} \sum_{\substack{N_3 \in \mathbb{Z}_L^2 \cap B(-K, 2^l) \\ 2(N_3 - N_2) \cdot N_3 = \mu}} 1 \qquad (3.25)$$

Let $n_3 = LN_3, n_2 = LN_2, m = \frac{L^2 \mu}{2} \in \mathbb{Z}$ (otherwise there is nothing to prove). Then we have that

$$(n_3 - n_2) \cdot n_3 = m \Leftrightarrow |2n_3 - n_2|^2 = 4m + |n_2|^2,$$

which means that for a fixed n_2, $2n_3$ belongs to the circle centered at n_2 of radius $\sqrt{4m + |n_2|^2}$. Now in the last sum in (3.25), for a given $l \in \mathbb{N}$, we have that $n_3 = LN_3$ belongs to balls of size $L2^l$. Using Lemma 3.10, we have that for any n_2, $l \geq 0$, and $\epsilon > 0$,

$$\sum_{\substack{N_3 \in \mathbb{Z}_L^2 \cap B(-K, 2^l) \\ 2(N_3 - N_2) \cdot N_3 = \mu}} 1 \quad \lesssim_\epsilon (2^l L)^\epsilon.$$

Consequently, we have that

$$\sum_{\mathcal{R}_\mu(K)} \langle K_2 \rangle^{-\sigma} \langle K_3 \rangle^{-\sigma} \lesssim_\epsilon L^\epsilon \sum_{N_2 \in \mathbb{Z}_L^2} \langle K + N_2 \rangle^{-\sigma} \sum_{l=0}^{\infty} 2^{-\sigma l} 2^{\epsilon \ell} \lesssim L^{2+\epsilon},$$

which is (3.24).

Case 2: $|K_1| < \frac{|K|}{100}$ and $|K| > 100$. This case is non-empty only when $\mu < 0$ due to the relation

$$|K_1|^2 + |K_3|^2 = |K|^2 + |K_2|^2 + \mu. \qquad (3.26)$$

From the relation $K_2 = K_1 - K_3 - K$ and recalling that $|K_3| \leq |K_1|$, we deduce that $|K_2| \geq \frac{|K|}{50}$, and as a result, we are reduced to showing that

$$\sup_{K \in \mathbb{Z}_L^2} \sum_{\mathcal{R}_\mu(K)} \langle K_1 \rangle^{-\sigma} \langle K_3 \rangle^{-\sigma} \lesssim_\epsilon L^{2+\epsilon} \qquad (3.27)$$

Now we change variable and define $N_i = K_i - K$ $(i = 1, 2, 3)$ so that the new variables satisfy the relations:

$$N_1 + N_3 = N_2 \qquad -2N_1 \cdot N_3 = |N_1|^2 + |N_3|^2 - |N_2|^2 = \mu.$$

Note that since $|K_3| \leq |K_1| \leq \frac{|K|}{100}$, we have that $|N_1| \sim |K| \sim |N_3|$. Now we write, $N_1 = \alpha J$ where $\alpha \in \mathbb{N}$ and J is visible and denote by $m \stackrel{def}{=} L^2 \frac{\mu}{2} \in \mathbb{Z}$ (otherwise there is nothing to prove). But if $2N_1 \cdot N_3 = -\mu$, we must have that α divides m.

The fact that N_3 satisfies the relation $J \cdot N_3 = \frac{m}{L^2\alpha}$ shows that $K + N_3$ ranges over a one-dimensional lattice of spacing $|J|$: Indeed, if we denote $n_3 = LN_3$ and $j = LJ$, then the equation $j \cdot n_3 = \frac{m}{2}$ has a general solution of the form $n_3 = n_3^{(0)} + \beta j^\perp$ with $\beta \in \mathbb{Z}$ and $n_3^{(0)}$ being any particular solution of the equation $j \cdot n_3^{(0)} = \frac{m}{2}$ (see [31] for the elementary argument). Consequently, for any fixed α, J, we have that

$$\sum_{N_3 \cdot J = \frac{m}{L^2\alpha}} \langle K + N_3 \rangle^{-\sigma} \lesssim \sum_{\beta \in \mathbb{Z}} \langle \beta J^\perp \rangle^{-\sigma} \lesssim \max(1, |J|^{-1}).$$

As a result, if we write as in the previous section $K = r_\alpha + \alpha Q_\alpha$, with $r_\alpha \in \mathcal{B}(\alpha/L) \cap \mathbb{Z}_L^2$ and $Q_\alpha \in \mathbb{Z}_L^2$ (see (3.6)), then we can estimate

$$
\begin{aligned}
\sum_{\mathcal{R}_\mu(K)} \langle K_1 \rangle^{-\sigma} \langle K_3 \rangle^{-\sigma} &\leq \sum_{\alpha | m} \sum_{\substack{J \in \mathbb{Z}_L^2 \\ J\, visible}} \langle K + \alpha J \rangle^{-\sigma} \sum_{N_3 \cdot J = \frac{m}{L^2\alpha}} \langle K + N_3 \rangle^{-\sigma} \\
&\lesssim \sum_{\alpha | m} \sum_{\substack{J \in \mathbb{Z}_L^2 \setminus \{-Q_\alpha\} \\ J\, visible}} \langle r_\alpha + \alpha(J + Q_\alpha) \rangle^{-\sigma} \sum_{N_3 \cdot J = \frac{m}{L^2\alpha}} \langle K + N_3 \rangle^{-\sigma} \\
&\quad + \sum_{\alpha | m} \langle r_\alpha \rangle^{-\sigma} \sum_{N_3 \cdot Q_\alpha = \frac{-m}{L^2\alpha}} \langle r_\alpha + \alpha Q_\alpha + N_3 \rangle^{-\sigma} \\
&\lesssim \sum_{\alpha | m} \sum_{\substack{J \in \mathbb{Z}_L^2 \setminus \{-Q_\alpha\} \\ J\, visible}} \langle r_\alpha + \alpha(J + Q_\alpha) \rangle^{-\sigma} \max(1, |J|^{-1}) \\
&\quad + \sum_{\alpha | m} \max(1, |Q_\alpha|^{-1}).
\end{aligned}
$$

For the first sum above we recall that $|J| = \frac{|N_1|}{\alpha} \sim \frac{|K|}{\alpha}$ and use Lemma 3.3, as for the second sum we simply estimate $|Q_\alpha|^{-1} \leq L$ and use that the bound on the number of divisors [31, 58] of m given by $C_\epsilon |m|^{\epsilon/100} \lesssim_\epsilon L^\epsilon$, as $|m| = L^2 \frac{|\mu|}{2} \lesssim L^{12}$ by assumption on μ. As a result, we get that

$$\sum_{\mathcal{R}_\mu(K)} \langle K_1 \rangle^{-\sigma} \langle K_3 \rangle^{-\sigma} \lesssim \left[\sum_{\alpha | m} \frac{L^2}{\alpha^2} \max\left(\frac{\alpha}{|K|}, 1\right) \right] + L^{1+\epsilon} \lesssim L^{2+\epsilon},$$

which is (3.27). This finishes the proof of the proposition. $\square$

4. Proof of Theorem 2.3: convergence of $\mathcal{T}_L$ to $\mathcal{T}$

The aim of this section is to prove Theorem 2.3. We first give the proof by assuming the co-prime equidistribution Lemma 4.1. In a second part, we prove this lemma.

4.1. Proof of Theorem 2.3.
We will use the following expression for $\mathcal{T}_L$ (see (2.6) and (2.2)):

$$\mathcal{T}_L(g, g, g)(K) = \frac{1}{\mathfrak{G}} \sum_{K_1 \cdot K_2 = 0} g(K + K_1)\bar{g}(K + K_1 + K_2)g(K + K_2) \quad \text{with} \quad \mathfrak{G} \overset{def}{=} \frac{2L^2 \log L}{\zeta(2)}.$$

Without loss of generality we assume that $B = 1$ in the statement of Theorem 2.3.

Step 1: decomposition of $\mathcal{T}_L$. Start by writing

$$\frac{1}{\mathfrak{G}} \sum_{K_1 \cdot K_2 = 0} g(K + K_1) \bar{g}(K + K_1 + K_2) g(K + K_2)$$

$$= \frac{1}{\mathfrak{G}} \sum_{\substack{K_1 \cdot K_2 = 0 \\ K_1, K_2 \neq 0}} g(K + K_1) \bar{g}(K + K_1 + K_2) g(K + K_2) + \frac{2}{\mathfrak{G}} \left(\sum_{K'} |g(K')|^2 \right) g(K) - \frac{1}{\mathfrak{G}} |g(K)|^2 g(K)$$

$$= I_1 + I_2 + I_3.$$

The terms I_2 and I_3 can be easily bounded by $\frac{L^2}{\mathfrak{G}} \|g\|_{\ell_L^2}^2 \|g\|_{\ell_L^\infty}$ and $\mathfrak{G}^{-1} \|g\|_{\ell_L^\infty}^3$ respectively. As for the main term I_1, we parameterize the sum over $\{k_1, k_2 \in \mathbb{Z}^2 \setminus \{0\} : k_1 \cdot k_2 = 0\}$ as $k_1 = \alpha(p, q), k_2 = \beta(-q, p)$ with $\alpha, \beta, p, q \in \mathbb{Z}$ such that $\alpha \geq 1, \beta \neq 0$ and $g.c.d(|p|, |q|) = 1$ (see Definition 2.1). We then write it as:

$$I_1 = \frac{1}{\mathfrak{G}} \sum_{\substack{\alpha \geq 1, \, g.c.d(p,q)=1 \\ \beta \in \mathbb{Z}^*}} \sum g(K + \frac{\alpha}{L}(p,q)) \bar{g}(K + \frac{\alpha}{L}(p,q) + \frac{\beta}{L}(-q,p)) g(K + \frac{\beta}{L}(-q,p))$$

$$= \frac{1}{\mathfrak{G}} \sum_{\substack{\alpha \geq 1, \, g.c.d(p,q)=1 \\ \beta \geq 1}} \sum g(K + \frac{\alpha}{L}(p,q)) \bar{g}(K + \frac{\alpha}{L}(p,q) + \frac{\beta}{L}(-q,p)) g(K + \frac{\beta}{L}(-q,p)) \quad (4.1)$$

$$+ \frac{1}{\mathfrak{G}} \sum_{\substack{\alpha \geq 1, \, g.c.d(p,q)=1 \\ \beta \geq 1}} \sum g(K - \frac{\alpha}{L}(p,q)) \bar{g}(K - \frac{\alpha}{L}(p,q) + \frac{\beta}{L}(-q,p)) g(K + \frac{\beta}{L}(-q,p)), \quad (4.2)$$

where we used the symmetry $(p, q) \to -(p, q)$. Now we write:

$$(4.1) = \frac{1}{\mathfrak{G}} \sum_{1 \leq \alpha \leq \beta} \sum_{g.c.d(p,q)=1} + \frac{1}{\mathfrak{G}} \sum_{1 \leq \beta \leq \alpha} \sum_{g.c.d(p,q)=1} - \frac{1}{\mathfrak{G}} \sum_{1 \leq \alpha = \beta} \sum_{(p,q)=1}$$

$$= \frac{1}{\mathfrak{G}} \sum_{1 \leq \alpha \leq \beta} \sum_{g.c.d(p,q)=1} g(K + \frac{\alpha}{L}(p,q)) \bar{g}(K + \frac{\alpha}{L}(p,q) + \frac{\beta}{L}(-q,p)) g(K + \frac{\beta}{L}(-q,p))$$

$$+ \sum_{1 \leq \beta \leq \alpha} \sum_{g.c.d(p',q')=1} g(K + \frac{\alpha}{L}(-q',p')) \bar{g}(K + \frac{\alpha}{L}(-q',p') - \frac{\beta}{L}(p',q')) g(K - \frac{\beta}{L}(p',q'))$$

$$- \frac{1}{\mathfrak{G}} \sum_{K' \in (\mathbb{Z}_L^2)^*} g(K + K') \bar{g}(K + K' + K'^\perp) g(K + K'^\perp)$$

$$= \frac{1}{\mathfrak{G}} \sum_{1 \leq \alpha \leq \beta} \sum_{g.c.d(p,q)=1} g(K + \frac{\alpha}{L}(p,q)) \bar{g}(K + \frac{\alpha}{L}(p,q) + \frac{\beta}{L}(-q,p)) g(K + \frac{\beta}{L}(-q,p))$$

$$+ \sum_{1 \leq \alpha \leq \beta} \sum_{g.c.d(p,q)=1} g(K - \frac{\alpha}{L}(p,q)) \bar{g}(K - \frac{\alpha}{L}(p,q) + \frac{\beta}{L}(-q',p')) g(K + \frac{\beta}{L}(-q,p))$$

$$- \frac{1}{\mathfrak{G}} \sum_{K' \in (\mathbb{Z}_L^2)^*} g(K + K') \bar{g}(K + K' + K'^\perp) g(K + K'^\perp)$$

(recall that the superscript $\perp$ the rotation in $\mathbb{R}^2$ by $\pi/2$ in the counter-clockwise direction). Arguing similarly

$$(4.2) = \frac{1}{\mathfrak{G}} \sum_{1 \leq \alpha \leq \beta} \sum_{(p,q)=1} (\cdots) + \frac{1}{\mathfrak{G}} \sum_{1 \leq \beta \leq \alpha} \sum_{(p,q)=1} (\cdots) - \frac{1}{\mathfrak{G}} \sum_{1 \leq \alpha = \beta} \sum_{(p,q)=1} (\cdots)$$

$$= \frac{1}{\mathfrak{G}} \sum_{1 \leq \alpha \leq \beta} \sum_{(p,q)=1} g(K - \frac{\alpha}{L}(p,q))\bar{g}(K - \frac{\alpha}{L}(p,q) + \frac{\beta}{L}(-q,p))g(K + \frac{\beta}{L}(-q,p))$$

$$+ \sum_{1 \leq \alpha \leq \beta} \sum_{(p,q)=1} g(K + \frac{\alpha}{L}(p,q))\bar{g}(K + \frac{\alpha}{L}(p,q) + \frac{\beta}{L}(-q,p))g(K + \frac{\beta}{L}(-q,p))$$

$$- \frac{1}{\mathfrak{G}} \sum_{K' \in (\mathbb{Z}_L^2)^*} g(K + K')\bar{g}(K + K' + K'^\perp)g(K + K'^\perp).$$

As a result,

$$I_1 = \frac{2}{\mathfrak{G}} \sum_{1 \leq \alpha \leq \beta} \sum_{(p,q)=1} g(K + \frac{\alpha}{L}(p,q))\bar{g}(K + \frac{\alpha}{L}(p,q) + \frac{\beta}{L}(-q,p))g(K + \frac{\beta}{L}(-q,p))$$

$$+ \frac{2}{\mathfrak{G}} \sum_{1 \leq \alpha \leq \beta} \sum_{(p,q)=1} g(K - \frac{\alpha}{L}(p,q))\bar{g}(K - \frac{\alpha}{L}(p,q) + \frac{\beta}{L}(-q,p))g(K + \frac{\beta}{L}(-q,p)) \qquad (4.3)$$

$$- \frac{2}{\mathfrak{G}} \sum_{K' \in (\mathbb{Z}_L^2)^*} g(K + K')\bar{g}(K + K' + K'^\perp)g(K + K'^\perp)$$

$$= II_1 + II_2 + II_3. \qquad (4.4)$$

Notice that the sum in II_3 is a sum over squares with vertex at K. Using the relation

$$|K|^2 + |K + K' + K'^\perp|^2 = |K + K'|^2 + |K + K'^\perp|^2$$

one notices that either $|K + K'|$ or $|K + K'^\perp|$ is larger than $|K|/2$. This allows to estimate $II_3(K)$ as follows:

$$II_3(K) \lesssim \mathfrak{G}^{-1} \langle K \rangle^{-\sigma} \sum_{K' \in (\mathbb{Z}_L^2)^*} \langle K + K' + K'^\perp \rangle^{-\sigma} \lesssim \langle K \rangle^{-\sigma} \frac{L^2}{\mathfrak{G}}.$$

which means that II_3 has an acceptable contribution in X^σ, namely

$$\|II_3(K)\|_{X^\sigma} \lesssim \frac{L^2}{\mathfrak{G}}.$$

We will perform the analysis only on II_1 since that for II_2 is exactly the same except for the signature of α.

Step 2: restricting the range of β. Before performing the analysis on II_1, recall that we call points z of $\mathbb{Z}^2$ *visible* if $z = (p,q)$ with $g.c.d(|p|,|q|) = 1$ (see Definition 2.1). Consequently, we can write II_1 as:

$$II_1 \overset{def}{=} \frac{2}{\mathfrak{G}} \sum_{1 \leq \alpha \leq \beta} \sum_{J \in \mathbb{Z}^2 \ visible} g(K + \frac{\alpha}{L}J)\bar{g}(K + \frac{\alpha}{L}J + \frac{\beta}{L}J^\perp)g(K + \frac{\beta}{L}J^\perp). \qquad (4.5)$$

We split II_1 into a sum over $\beta < L$ and another over $\beta \geq L$ to get:

$$II_1 = III_1 + III_2 \qquad (4.6)$$

where

$$III_1 = \frac{2}{\mathfrak{G}} \sum_{\substack{1 \leq \alpha \leq \beta \\ \beta < L}} \sum_{J \in \mathbb{Z}^2 \ visible} g(K + \frac{\alpha}{L}J)\bar{g}(K + \frac{\alpha}{L}J + \frac{\beta}{L}J^\perp)g(K + \frac{\beta}{L}J^\perp) \tag{4.7}$$

and

$$III_2 = \frac{2}{\mathfrak{G}} \sum_{\substack{1 \leq \alpha \leq \beta \\ \beta \geq L}} \sum_{J \in \mathbb{Z}^2 \ visible} g(K + \frac{\alpha}{L}J)\bar{g}(K + \frac{\alpha}{L}J + \frac{\beta}{L}J^\perp)g(K + \frac{\beta}{L}J^\perp). \tag{4.8}$$

Remark 3.8 says exactly that

$$\|III_2\|_{X^\sigma} \lesssim \frac{L^2}{\mathfrak{G}}, \tag{4.9}$$

which means that the contribution of III_2 is an acceptable error.

Step 3: discrete to continuous limit in J. We move on to analyzing III_1. For this, we define $N \stackrel{def}{=} L/\beta$ and write the sum:

$$\sum_{\substack{J \in \mathbb{Z}^2 \\ visible}} g(K + \frac{\alpha}{L}J)\bar{g}(K + \frac{\alpha}{L}J + \frac{\beta}{L}J^\perp)g(K + \frac{\beta}{L}J^\perp)$$

$$= \sum_{\substack{J' \in \mathbb{Z}_N^2 \\ visible}} g(K + \frac{\alpha}{\beta}J')\bar{g}(K + \frac{\alpha}{\beta}J' + J'^\perp)g(K + J'^\perp).$$

The estimate on this sum is contained in the following lemma.

Lemma 4.1 (Co-prime equidistribution lemma). *Let $\sigma > 2$. For all $B > 0$, if $g : \mathbb{R}^2 \to \mathbb{C}$ satisfies the following bounds:*

$$\|g\|_{X^{\sigma+1}(\mathbb{R}^2)} + \|\nabla g\|_{X^{\sigma+1}(\mathbb{R}^2)} \leq B,$$

then for all $N > 1$, we have

$$\mathcal{E}_N \stackrel{def}{=} \left\| \frac{1}{N^2} \sum_{\substack{J \in \mathbb{Z}_N^2 \\ J \ visible}} g(K + \frac{\alpha}{\beta}J)\bar{g}(K + \frac{\alpha}{\beta}J + J^\perp)g(K + J^\perp) \right.$$

$$\left. - \frac{1}{\zeta(2)} \int_{\mathbb{R}^2} g(K + \frac{\alpha}{\beta}z)\bar{g}(K + \frac{\alpha}{\beta}z + z^\perp)g(K + z^\perp)dz \right\|_{X^\sigma(\mathbb{R}^2)} \lesssim B^3 \frac{1 + \log N}{N}. \tag{4.10}$$

We leave the proof of this lemma for later in order not to distract the reader. This allows us to write:

$$III_1 = \frac{2L^2}{\zeta(2)\mathfrak{G}} \sum_{1 \leq \alpha \leq \beta \leq L} \frac{1}{\beta^2} \int_{\mathbb{R}^2} g(K + \frac{\alpha}{\beta}x)\bar{g}(K + \frac{\alpha}{\beta}x + x^\perp)g(K + x^\perp)\, dx + \frac{2L^2}{\mathfrak{G}} \sum_{1 \leq \alpha \leq \beta \leq L} \frac{1}{\beta^2}\mathcal{E}_{L/\beta}$$

$$= IV_1 + IV_2$$

with

$$\|IV_2\|_{X^\sigma} \lesssim \frac{L}{\mathfrak{G}} \sum_{1 \leq \beta \leq L} (1 + \log(L/\beta)) \lesssim \frac{L^2}{\mathfrak{G}}.$$

Step 4: discrete to continuous limit in α. In order to analyze IV_1, we view the α sum as a Riemann sum, which we will replace by the corresponding integral. For each $1 \leq \beta \leq L$, denote:

$$\mathcal{D}_\beta \stackrel{def}{=} \frac{1}{\beta} \sum_{1 \leq \alpha \leq \beta} \int_{\mathbb{R}^2} g(K + \frac{\alpha}{\beta}x)\bar{g}(K + \frac{\alpha}{\beta}x + x^\perp)g(K + x^\perp)\, dx$$

$$- \int_0^1 \int_{\mathbb{R}^2} g(K + \lambda x)\bar{g}(K + \lambda x + x^\perp)g(K + x^\perp)\, dx\, d\lambda$$

$$= \sum_{1 \leq \alpha \leq \beta} \int_{\frac{\alpha-1}{\beta}}^{\frac{\alpha}{\beta}} \int_{\mathbb{R}^2} \left(g(K + \frac{\alpha}{\beta}x) - g(K + \lambda x) \right) \bar{g}(K + \frac{\alpha}{\beta}x + x^\perp)g(K + x^\perp)dx\, d\lambda$$

$$+ \sum_{1 \leq \alpha \leq \beta} \int_{\frac{\alpha-1}{\beta}}^{\frac{\alpha}{\beta}} \int_{\mathbb{R}^2} g(K + \lambda x) \left(\bar{g}(K + \frac{\alpha}{\beta}x + x^\perp) - \bar{g}(K + \lambda x + x^\perp) \right) g(K + x^\perp)dx\, d\lambda.$$

We only estimate the first term as the second is similar. It can be written

$$\sum_{1 \leq \alpha \leq \beta} \int_{\frac{\alpha-1}{\beta}}^{\frac{\alpha}{\beta}} \int_\lambda^{\alpha/\beta} \int_{\mathbb{R}^2} x \cdot \nabla g(K + tx)\bar{g}(K + \frac{\alpha}{\beta}x + x^\perp)g(K + x^\perp)dx\, dt\, d\lambda.$$

We will show that for each $1 \leq \alpha \leq \beta$ and each $t \in [0, 1]$,

$$\left\| \int_{\mathbb{R}^2} x \cdot \nabla g(K + tx)\bar{g}(K + \frac{\alpha}{\beta}x + x^\perp)g(K + x^\perp)\, dx \right\|_{X^\sigma} \lesssim 1. \tag{4.11}$$

To see this we split into two cases: Either $|x| \leq 100|K|$, in which case we use the orthogonality relation

$$|K|^2 + |K + tx + x^\perp|^2 = |K + tx|^2 + |K + x^\perp|^2$$

to obtain that either $|K + x^\perp| \geq |K|/2$ or $|K + tx| \geq |K|/2$. In both cases, we have since $|g(x)| + |\nabla g(x)| \lesssim \langle x \rangle^{-\sigma-1}$:

$$\left| \int_{|x| \leq 100|K|} x \cdot \nabla g(K + tx)\bar{g}(K + \frac{\alpha}{\beta}x + x^\perp)g(K + x^\perp)\, dx \right|$$

$$\lesssim \frac{|K|}{\langle K \rangle^{\sigma+1}} \int_{\mathbb{R}^2} \langle K + \frac{\alpha}{\beta}x + x^\perp \rangle^{-\sigma}\, dx \lesssim \langle K \rangle^{-\sigma},$$

which is (4.11). In the case when $|x| \geq 100|K|$, then $|K + x^\perp| \geq |K|$ and $|K + \frac{\alpha}{\beta}x + x^\perp| \geq \frac{1}{2}|\frac{\alpha}{\beta}x + x^\perp| \geq \frac{1}{2}|x|$, and hence the estimate (4.11) follows directly.

With (4.11) in hand, Minkowski's inequality gives

$$\|\mathcal{D}_\beta\|_{X^\sigma} \lesssim \frac{1}{\beta}. \tag{4.12}$$

Consequently,

$$IV_1 = \frac{2L^2}{\zeta(2)\mathfrak{G}} \sum_{1 \le \beta \le L} \frac{1}{\beta} \int_0^1 \int_{\mathbb{R}^2} g(K + \lambda x)\bar{g}(K + \lambda x + x^\perp)g(K + x^\perp)\, dx\, d\lambda$$

$$+ \frac{2L^2}{\zeta(2)\mathfrak{G}} \sum_{1 \le \beta \le L} \frac{1}{\beta}\mathcal{D}_\beta(K)$$

and hence

$$IV_1 = \frac{2\log^D(L)L^2}{\zeta(2)\mathfrak{G}} \int_0^1 \int_{\mathbb{R}^2} g(K + \lambda x)\bar{g}(K + \lambda x + x^\perp)g(K + x^\perp)\, dx\, d\lambda + \frac{2L^2}{\zeta(2)\mathfrak{G}} \sum_{1 \le \beta \le L} \frac{1}{\beta}\mathcal{D}_\beta(K)$$

$$= V_1 + V_2,$$

where $\log^D n = \sum_{k=1}^n \frac{1}{k}$, and where V_2 can be bounded as explained above by

$$\|V_2\|_{X^\sigma} \lesssim \frac{L^2}{\mathfrak{G}}. \tag{4.13}$$

Finally, using that $|\log^D L - \log L| = O(1)$, the definition of $\mathfrak{G}$, and Theorem 2.2, we obtain

$$V_1 = \int_0^1 \int_{\mathbb{R}^2} g(K + \lambda x)\bar{g}(K + \lambda x + x^\perp)g(K + x^\perp)\, dx\, d\lambda + O_{X^\sigma}\left(\frac{1}{\log L}\right)$$

(with the notation $A = B + O_{X^\sigma}(C)$ if $\|A - B\|_{X^\sigma} \lesssim C$).

Step 5: conclusion. Combining all the above, we get that if $\mathfrak{G} = \frac{2\log^D(L)L^2}{\zeta(2)}$ then:

$$II_1 = \int_0^1 \int_{\mathbb{R}^2} g(K + \lambda z)\bar{g}(K + \lambda z + z^\perp)g(K + z^\perp)dz\, d\lambda + O_{X^\sigma}\left(\frac{1}{\log L}\right).$$

The same analysis for II_2 gives (recalling that the only difference with the above is the sign of α)

$$II_2 = \int_0^1 \int_{\mathbb{R}^2} g(K - \lambda z)\bar{g}(K - \lambda z + z^\perp)g(K + z^\perp)dz\, d\lambda + O_{X^\sigma}\left(\frac{1}{\log L}\right)$$

$$= \int_{-1}^0 \int_{\mathbb{R}^2} g(K + \lambda z)\bar{g}(K + \lambda z + z^\perp)g(K + z^\perp)dz\, d\lambda + O_{X^\sigma}\left(\frac{1}{\log L}\right),$$

and as a final result we get that:

$$\left\| \frac{1}{\mathfrak{G}} \sum_{K_1 \cdot K_3 = 0} g(K + K_1)\bar{g}(K + K_1 + K_2)g(K + K_2) - \mathcal{T}(g, g, g)(K) \right\|_{X^\sigma} \lesssim \frac{1}{\log L}, \tag{4.14}$$

which finishes the proof.

4.2. Proof of the co-prime equidistribution Lemma 4.1. We start by recalling some elementary number theory.

Step 0: number theoretic preliminaries. For any $N > 0$, define for any subset P of $\mathbb{R}^2$ the function:

$$g_P(N) = \#\{z \in (\mathbb{Z}_N^2)^* \cap P\}.$$

Note that for reasonably nice regions P, the area of P is equal to the quantity:

$$V_P = \lim_{N \to +\infty} \frac{g_P(N)}{N^2}.$$

Recall that a lattice point $z \in (\mathbb{Z}_N^2)^*$ is *visible* (from 0) if the segment $[0z]$ does not intersect the lattice $\mathbb{Z}_N^2$ apart from the endpoints 0 and z; equivalently, $z = N^{-1}(p, q)$ where (p, q) are co-prime (see Definition 2.1). This definition allows one to define the following counter-part of g_P:

$$f_P(N) \stackrel{def}{=} \#\{z \in \mathbb{Z}_N^2 \cap P : z \text{ is visible}\}.$$

It is a classical result in analytic number theory [31] that

$$\lim_{N \to +\infty} \frac{f_P(N)}{N^2} = \frac{V_P}{\zeta(2)}, \tag{4.15}$$

where $\zeta(2) = \frac{\pi^2}{6}$ is the Riemann zeta function at 2. The generalization of this result from characteristic functions to more general functions, along with the relevant discrepancy or error estimates, will be a byproduct of the analysis in this section, and we include it at the end in Corollary 4.5.

To prove Lemma 4.1, we need to estimate the discrepancy between the equidistribution sum in (4.10) and the integral in the strong norm $X^\sigma(\mathbb{Z}_L^2)$. The proof of (4.15), which is actually the prototype for our proof, is based on an inversion formula for the multiplicative Möbius function $\mu(m) : \mathbb{N} \to \mathbb{Z}$, whose definition we recall first:

$$\mu(k) = \begin{cases} 1 & \text{if } k = 1, \\ (-1)^\kappa & \text{if } k = p_1 p_2 \ldots p_\kappa \text{ and all the primes } p_1, \ldots, p_\kappa \text{ are different,} \\ 0 & \text{if } k \text{ is not square-free.} \end{cases} \tag{4.16}$$

Recall that the Möbius function is multiplicative (i.e. $\mu(ab) = \mu(a)\mu(b)$ whenever a and b are co-prime) and satisfies the following fundamental inversion formula proved by Möbius [31]:

$$g(n) = \sum_{d|n} f(d) \Rightarrow f(n) = \sum_{d|n} \mu(\frac{n}{d}) g(d),$$

from which follows:

Lemma 4.2 ([31])**.** *Suppose that $f : \mathbb{N} \to \mathbb{C}$ satisfies, for all $x \in \mathbb{N}$, $\sum_{m,n} |f(mnx)| = \sum_k d(k)|f(kx)| < \infty$, where $d(k)$ is the number of divisors of k. Then for any $x \in \mathbb{N}$, there holds*

$$g(x) = \sum_{m=1}^{\infty} f(mx) \Leftrightarrow f(x) = \sum_{n=1}^{\infty} \mu(n) g(nx). \tag{4.17}$$

The final property of $\mu(n)$ that will be needed in the proof is its connection to the Riemann zeta function defined for $\mathrm{Re}\, s > 1$ by

$$\zeta(s) = \sum_{n \in \mathbb{N}} \frac{1}{n^s}.$$

Applying Lemma 4.2 with $f(x) = x^{-s}$ and $g(x) = \zeta(s) x^{-s}$, one gets that

$$\frac{1}{\zeta(s)} = \sum_{n=1}^{\infty} \frac{\mu(n)}{n^s}.$$

Step 1: a new formulation for (4.10). Again we assume without any loss of generality that $B = 1$. Given $u : \mathbb{R}^2 \to \mathbb{C}$ and such that $\sup_{N>1} N^{-2} \sum_{z \in \mathbb{Z}_N^2} |u(z)| < \infty$, we define the following two sums:

$$S(u, N) \overset{def}{=} \sum_{\substack{z \in \mathbb{Z}_N^2 \\ visible}} u(z) \quad \text{and} \quad D(u, N) = \sum_{z \in (\mathbb{Z}_N^2)^*} u(z). \tag{4.18}$$

Note that if $u \in X^\kappa(\mathbb{R}^2)$ for some $\kappa > 2$, then for any $N > 0$, we have that

$$N^{-2} D(|u|, N) \lesssim 1 \tag{4.19}$$

We will first compare the discrepancy between $S(u, N)$ and $D(u, N)$.

Notice that every $z \in \mathbb{Z}_N^2$ is a visible point for exactly one lattice in the family $\{\mathbb{Z}_{\frac{N}{m}}^2\}_{m \in \mathbb{N}}$. As a result, we have

$$\sum_{z \in \mathbb{Z}_N^2} u(z) = \sum_{m=1}^{\infty} \sum_{\substack{z \in \mathbb{Z}_{\frac{N}{m}}^2 \\ visible}} u(z)$$

and hence for all $N > 0$,

$$D(u, N) = \sum_{m=1}^{\infty} S(u, \frac{N}{m}).$$

Consequently, applying Lemma 4.2 to the functions $f(y) = D(u, \frac{N}{y})$ and $g(y) = S(u, \frac{N}{y})$ at $x = 1$ we obtain for all $N > 0$,

$$S(u, N) = \sum_{n=1}^{\infty} \mu(n) D(u, \frac{N}{n}), \tag{4.20}$$

provided that $\sum_{m,n \in \mathbb{N}} |S(u, \frac{N}{mn})| < \infty$. But this follows from (4.19) as soon as $u \in X^\kappa(\mathbb{R}^2)$ with $\kappa > 2$, and the fact that $S(u, \sigma) \leq D(|u|, \sigma)$ for any $\sigma > 0$.

With these notations, the error term in the estimate (4.10) can be written as follows

$$\mathcal{E}_N(K) \overset{def}{=} \frac{1}{N^2} \sum_{z \in \mathbb{Z}_N^2 \ visible} u_K(z) - \frac{1}{\zeta(2)} \int_{\mathbb{R}^2} u_K(z) dz = \frac{1}{N^2} S(u_K, N) - \frac{1}{\zeta(2)} \int_{\mathbb{R}^2} u_K(z) dz,$$

where

$$u_K(z) = g(K + \frac{\alpha}{\beta} z) \bar{g}(K + \frac{\alpha}{\beta} z + z^\perp) g(K + z^\perp). \tag{4.21}$$

Now by (4.20)

$$\frac{1}{N^2} S(u_K, N) = \sum_{n=1}^{\infty} \frac{\mu(n)}{n^2} \left(\frac{n^2}{N^2} D(u_K, \frac{N}{n}) \right)$$

$$= \sum_{n=1}^{\infty} \frac{\mu(n)}{n^2} \int_{\mathbb{R}^2} u_K(z) \, dz + \sum_{n=1}^{\infty} \frac{\mu(n)}{n^2} \left(\frac{n^2}{N^2} D(u_K, \frac{N}{n}) - \int_{\mathbb{R}^2} u_K(z) \, dz \right)$$

$$= \frac{1}{\zeta(2)} \int_{\mathbb{R}^2} u_K(z) \, dz + \sum_{n=1}^{\infty} \frac{\mu(n)}{n^2} \left(\frac{n^2}{N^2} D(u_K, \frac{N}{n}) - \int_{\mathbb{R}^2} u_K(z) \, dz \right).$$

As a result, we have that

$$\rho^2 S(u_K, \rho) - \frac{1}{\zeta(2)} \int_{\mathbb{R}^2} u_K(z)\,dz = \sum_{n=1}^{\infty} \frac{\mu(n)}{n^2} \left(\frac{n^2}{N^2} D(u_K, \frac{N}{n}) - \int_{\mathbb{R}^2} u_K(z)\,dz \right)$$

$$\overset{def}{=} \mathfrak{D}_1 + \mathfrak{D}_2.$$

with

$$\mathfrak{D}_1 = \sum_{n=1}^{[N]} \frac{\mu(n)}{n^2} \left(\frac{n^2}{N^2} D(u_K, \frac{N}{n}) - \int_{\mathbb{R}^2} u_K(z)\,dz \right) \quad \text{and}$$

$$\mathfrak{D}_2 = \sum_{n=[N]+1}^{\infty} \frac{\mu(n)}{n^2} \left(\frac{n^2}{N^2} D(u_K, \frac{N}{n}) - \int_{\mathbb{R}^2} u_K(z)\,dz \right).$$

Step 2: estimates of $\mathfrak{D}_1$ and $\mathfrak{D}_2$. The estimate of $\mathfrak{D}_1$ and $\mathfrak{D}_2$ relies on Lemma 4.3 below, whose proof will be given in the next paragraph.

For any given $R > 0$ and $K \in \mathbb{R}^2$, there exists a unique $r \in \mathcal{B}(\frac{1}{R})$ and $Q \in \mathbb{Z}_R^2$ such that $K = r + Q$. Define

$$D^*(u_K, R) \overset{def}{=} \sum_{M \in (\mathbb{Z}_R^2)^* \setminus \{Q^\perp\}} u_K(M). \tag{4.22}$$

Note that $D^*(u_K, R) = D(u_K, R)$ if $Q = 0$. In particular, if

$$|K| \le \frac{1}{2R} \implies D^*(u_K, R) = D(u_K, R). \tag{4.23}$$

Lemma 4.3. *With the previous notations, we have*

(1) *If $R \le 1$, then*

$$\|R^{-2} D^*(u_K, R) - \int_{\mathbb{R}^2} u_K(z)\,dz\|_{X^\sigma(\mathbb{R}^2)} \lesssim 1. \tag{4.24}$$

(2) *If $R \ge 1$, then*

$$\|R^{-2} D(u_K, R) - \int_{\mathbb{R}^2} u_K(z)\,dz\|_{X^\sigma(\mathbb{R}^2)} \lesssim \frac{1}{R}. \tag{4.25}$$

To estimate $\mathfrak{D}_1$, we use the second part this Lemma (with $R = N/n \ge 1$) to get that

$$\|\mathfrak{D}_1\|_{X^\sigma(\mathbb{R}^2)} \lesssim \sum_{n=1}^{[N]} \frac{1}{n^2} \frac{n}{N} \lesssim \frac{1 + \log N}{N}$$

which is acceptable.

To estimate $\mathfrak{D}_2$, we must isolate from the sum defining $D(u_K, \frac{N}{n})$ the problematic point $Q^\perp$ (depending on K and $R = \frac{N}{n}$) and define $D^*(u_K, \frac{N}{n})$ as in (4.22). Using (4.23) with $R = N/n \le 1$, we can write

$$\mathfrak{D}_2 = \sum_{n=[N]+1}^{\infty} \frac{\mu(n)}{n^2} \left(\frac{n^2}{N^2} D^*(u_K, \frac{N}{n}) - \int_{\mathbb{R}^2} u_K(z)\,dz \right) + \frac{1}{N^2} \sum_{n=[N]+1}^{2[N|K|]+1} \mu(n) u_K(Q^\perp)$$

$$\overset{def}{=} \mathfrak{D}_2^{(1)} + \mathfrak{D}_2^{(2)},$$

The estimate on $\mathfrak{D}_2^{(1)}$ follows from the first part of Lemma 4.3 below, which gives that

$$\|\mathfrak{D}_2^{(1)}\|_{X^\sigma(\mathbb{R}^2)} \lesssim \sum_{n=[N]+1}^{\infty} \frac{1}{n^2} \lesssim \frac{1}{N}.$$

To estimate the contribution of $\mathfrak{D}_2^{(2)}$, we first note that

$$u_K(Q^\perp) = g(r + \frac{\alpha}{\beta}Q^\perp + Q)\overline{g}(r + \frac{\alpha}{\beta}Q^\perp)g(r)$$

and hence due to the orthogonality relation

$$|K|^2 + |r + \frac{\alpha}{\beta}Q^\perp|^2 = |r|^2 + |r + \frac{\alpha}{\beta}Q^\perp + Q|^2$$

we get that $|r + \frac{\alpha}{\beta}Q^\perp + Q| \geq \frac{|K|}{2}$ or $|r| \geq \frac{|K|}{2}$. Using the fact that $g \in X^{\sigma+1}$, we can estimate

$$|\mathfrak{D}_2^{(2)}| \lesssim \frac{1}{N^2} \sum_{n=[N]+1}^{2[N|K|]+1} \langle K \rangle^{-\sigma-1} \lesssim \frac{1}{N^2} \frac{N|K|}{\langle K \rangle^{\sigma+1}} \lesssim \frac{1}{N\langle K \rangle^\sigma},$$

and consequently we have that

$$\|\mathfrak{D}_2^{(2)}\|_{X^\sigma(\mathbb{R}^2)} \lesssim \frac{1}{N}$$

as needed. This finishes the proof of Lemma 4.1, modulo Lemma 4.3 that we now prove.

Step 3: Proof of Lemma 4.3. We start with part (1). In this case $R \leq 1$ and we will show that

$$\|R^{-2}D^*(u_K, R)\|_{X^\sigma(\mathbb{R}^2)} + \|\int_{\mathbb{R}^2} u_K(z)\|_{X^\sigma(\mathbb{R}^2)} \lesssim 1. \tag{4.26}$$

Recall that we have written $K = r + Q$ with $Q \in \mathbb{Z}_R^2$ and $r \in \mathcal{B}(\frac{1}{R})$, and defined $D^*(u_K, R)$ so that:

$$\frac{1}{R^2}D^*(u_K, R) = \frac{1}{R^2} \sum_{M \in (\mathbb{Z}_R^2)^* \backslash \{Q^\perp\}} g(K + \frac{\alpha}{\beta}M)\overline{g}(K + \frac{\alpha}{\beta}M + M^\perp)g(K + M^\perp). \tag{4.27}$$

The bound on (4.27) follows by splitting into two cases: either $|K + \frac{\alpha}{\beta}M| \geq \frac{|K|}{100}$ or $|K + \frac{\alpha}{\beta}M| \leq \frac{|K|}{100}$. The contribution of the first case ($|K + \frac{\alpha}{\beta}M| \geq \frac{|K|}{100}$) to the X^σ norm of (4.27) can be bounded by

$$\sup_{K \in \mathbb{R}^2} \frac{1}{R^2} \sum_{M \in \mathbb{Z}_R^2 \backslash \{Q^\perp\}} \langle K + \frac{\alpha}{\beta}M + M^\perp \rangle^{-\sigma} \langle K + M^\perp \rangle^{-\sigma}$$

$$\lesssim \sup_{K \in \mathbb{R}^2} \frac{1}{R^2} \sum_{M \in \mathbb{Z}_R^2 \backslash \{Q^\perp\}} \langle K + M^\perp \rangle^{-\sigma} \lesssim 1,$$

where we have used Lemma 3.3 (in the case $\alpha = 0$, $\beta = 1$ and $L \leq 1$) and the fact that $M \neq Q^\perp$ in the last step.

In the case when $|K + \frac{\alpha}{\beta}M| \leq \frac{|K|}{100}$, we can assume that $|K| \geq 100$ since otherwise the same argument as in the case $|K + \frac{\alpha}{\beta}M| > \frac{|K|}{100}$ applies. But $|K + \frac{\alpha}{\beta}M| \leq \frac{K}{100}$ implies that $|M| \geq |\frac{\alpha}{\beta}M| \geq \frac{99}{100}|K|$ (recall that $\alpha \leq \beta$ in (4.21)). This also implies that $|K + \frac{\alpha}{\beta}M + M^\perp| \geq$

$|M^\perp| - |K + \frac{\alpha}{\beta}M| \geq \frac{98}{100}|K|$ and consequently we have that the contribution of the second case to the X^σ norm of (4.27) can be bounded by

$$\sup_{K \in \mathbb{R}^2} \frac{1}{R^2} \sum_{M \in \mathbb{Z}_R^2 \setminus \{Q^\perp\}} \langle K + \frac{\alpha}{\beta}M \rangle^{-\sigma} \langle K + M^\perp \rangle^{-\sigma} \lesssim \sup_{K \in \mathbb{R}^2} \frac{1}{R^2} \sum_{M \in \mathbb{Z}_R^2 \setminus \{Q^\perp\}} \langle K + M^\perp \rangle^{-\sigma} \lesssim 1,$$

where we used again Lemma 3.3 and the fact that $M \neq Q^\perp$. In conclusion, we have that

$$\|(4.27)\|_{X^\sigma} \lesssim 1.$$

As a result,

$$\|R^{-2}D^*(u_K, R)\|_{X^\sigma} \lesssim 1 \tag{4.28}$$

The estimate on $\int_{\mathbb{R}^2} u_K(z)dz$ is similar to (even simpler than) that performed above for the discrete sum, and in view of (4.21), one gets the bound

$$\left\| \int_{\mathbb{R}^2} u_K(z)\, dz \right\|_{X^\sigma} \lesssim 1,$$

which finishes the proof of (4.24).

Now we move on to proving Part (2). Here $R \geq 1$. Let $\mathcal{B}_M(\frac{1}{R})$ denote the box $M + [0, \frac{1}{R})^2$. Then

$$\frac{1}{R^2}D(u_K, R) - \int_{\mathbb{R}^2} u_K(z)\, dz = \sum_{M \in \mathbb{Z}_R^2} \int_{\mathcal{B}_M(\frac{1}{R})} [g(K + \frac{\alpha}{\beta}M)\bar{g}(K + \frac{\alpha}{\beta}M + M^\perp)g(K + M^\perp)$$

$$- g(K + \frac{\alpha}{\beta}z)\bar{g}(K + \frac{\alpha}{\beta}z + z^\perp)g(K + z^\perp)]\, dz$$

$$= \sum_{M \in \mathbb{Z}_R^2} \int_{\mathcal{B}_M(\frac{1}{R})} \left[g(K + \frac{\alpha}{\beta}M) - g(K + \frac{\alpha}{\beta}z) \right] \bar{g}(K + \frac{\alpha}{\beta}M + M^\perp)g(K + M^\perp)dz \tag{4.29}$$

$$+ \sum_{M \in \mathbb{Z}_R^2} \int_{\mathcal{B}_M(\frac{1}{R})} g(K + \frac{\alpha}{\beta}z) \left[\bar{g}(K + \frac{\alpha}{\beta}M + M^\perp) - \bar{g}(K + \frac{\alpha}{\beta}z + z^\perp) \right] g(K + M^\perp)\, dz$$

$$\tag{4.30}$$

$$+ \sum_{M \in \mathbb{Z}_R^2} \int_{\mathcal{B}_M(\frac{1}{R})} g(K + \frac{\alpha}{\beta}z)\bar{g}(K + \frac{\alpha}{\beta}z + z^\perp) \left[g(K + M^\perp) - g(K + z^\perp) \right]\, dz. \tag{4.31}$$

We start by bounding (4.29). For this we write

$$|(4.29)| =$$

$$\left| \sum_{M \in \mathbb{Z}_R^2} \int_{\mathcal{B}_M(\frac{1}{R})} \int_0^1 \frac{\alpha}{\beta}(M - z).\nabla g(K + \frac{\alpha}{\beta}z + \frac{\alpha}{\beta}t(M - z))\bar{g}(K + \frac{\alpha}{\beta}M + M^\perp)g(K + M^\perp)dt\, dz \right|$$

$$\leq \frac{1}{R} \sum_{M \in \mathbb{Z}_R^2} \int_{\mathcal{B}_0(\frac{1}{R})} \int_0^1 |\nabla g|(K + \frac{\alpha}{\beta}M + \frac{\alpha}{\beta}z - \frac{\alpha}{\beta}tz)|g|(K + \frac{\alpha}{\beta}M + M^\perp)|g|(K + M^\perp)dt\, dz$$

$$= \frac{1}{R^3} \fint_{\mathcal{B}_0(\frac{1}{R})} \int_0^1 \sum_{M \in \mathbb{Z}_R^2} |\nabla g|(K + \frac{\alpha}{\beta}M + \frac{\alpha}{\beta}z - \frac{\alpha}{\beta}tz)|g|(K + \frac{\alpha}{\beta}M + M^\perp)|g|(K + M^\perp)dt\, dz,$$

where we denoted by $\fint_{\mathcal{B}_0(\frac{1}{R})} = R^2 \int_{\mathcal{B}_0(\frac{1}{R})}$. To prove that $\|(4.29)\| \lesssim \frac{1}{R}$, it suffices -given our bounds on g and ∇g in X^σ- to prove that for any $0 \leq t \leq 1$ and $z \in \mathcal{B}_0(\frac{1}{R})$

$$\frac{1}{R^2} \sum_{M \in \mathbb{Z}_R^2} \langle K + \frac{\alpha}{\beta}M + \frac{\alpha}{\beta}(1-t)z \rangle^{-\sigma} \langle K + \frac{\alpha}{\beta}M + M^\perp \rangle^{-\sigma} \langle K + M^\perp \rangle^{-\sigma} \lesssim \langle K \rangle^{-\sigma}. \quad (4.32)$$

But this follows by repeating the by-now-familiar argument: If $|K| \leq 200$ or if $|K + \frac{\alpha}{\beta}M| \geq \frac{|K|}{100}$, then since $(1-t)z \in \mathcal{B}_0(\frac{1}{R})$, $\langle K + \frac{\alpha}{\beta}M + \frac{\alpha}{\beta}z - \frac{\alpha}{\beta}tz \rangle \gtrsim \langle K \rangle$; and (4.32) follows from Lemma 3.3 (with $\beta = 1$ and $L \geq 1$) since

$$\frac{1}{R^2} \sum_{M \in \mathbb{Z}_R^2} \langle K + M^\perp \rangle^{-\sigma} \lesssim 1.$$

Otherwise, we have that $|K| \geq 200$ and $|K + \frac{\alpha}{\beta}M| \leq \frac{|K|}{100}$. In this case, (4.32) follows again from Lemma 3.3 since $|K + \frac{\alpha}{\beta}M + M^\perp| \gtrsim |K|$ (using the orthogonality relation).

The argument for (4.30) and (4.31) are similar upon using again Lemma (3.3) and the bounds on g and ∇g. This finishes the proof of Lemma 4.3.

Remark 4.4. *A direct consequence of the analysis performed in the previous section is the following co-prime equidistribution result, which we state for functions u satisfying $(|u|, |\nabla u|) \in X^\sigma(\mathbb{R}^2)$, but can be extended easily to much larger classes of functions.*

Corollary 4.5. *Suppose that $u : \mathbb{R}^2 \to \mathbb{C}$ satisfies $\|u\|_{X^\sigma} + \|\nabla u\|_{X^\sigma} \leq B$ for $\sigma > 2$. Then for any $L \geq 1$, the following discrepancy bound holds:*

$$\left| L^{-2} \sum_{\substack{J \in \mathbb{Z}_L^2 \\ visible}} u(J) - \frac{1}{\zeta(2)} \int_{\mathbb{R}^2} u(z)\,dz \right| \lesssim B \frac{1 + \log L}{L}. \quad (4.33)$$

5. Proof of Approximation Theorems: convergence of (NLS) to (CR)

The purpose of this section is to prove Theorems 2.4 and 2.6. We will start by adopting a notationally convenient rescaling. Recall that $\widetilde{a}_K(t) = e^{-4\pi^2 i |K|^2 t} a_K(t)$ satisfies the equation

$$-i\partial_t \widetilde{a}_K(t) = \frac{\epsilon^2}{L^4} \sum_{\mathcal{S}(K)} \widetilde{a}_{K_1}(t) \overline{\widetilde{a}_{K_2}}(t) \widetilde{a}_{K_3}(t) e^{4\pi^2 i \Omega t}.$$

where $\Omega(K, K_1, K_2, K_3) = -|K|^2 + |K_1|^2 - |K_2|^2 + |K_3|^2$. Set $C_0 = \dfrac{\zeta(2)}{2}$, and define the renormalized profile b_K as

$$b_K(t) = \widetilde{a}_K(\frac{C_0 L^2}{\log L}t) = a_K(\frac{C_0 L^2}{\log L}t) e^{-i\frac{4\pi^2 C_0 L^2}{\log L}|K|^2 t}. \quad (5.1)$$

It will be much more notationally convenient to work with $b_K(t)$ instead of $a_K(t)$. It satisfies

$$-i\partial_t b_K = \frac{C_0 \epsilon^2}{L^2 \log L} \sum_{\mathcal{S}(K)} b_{K_1}(t) \overline{b_{K_2}(t)} b_{K_3}(t) e^{i\frac{4\pi^2 C_0 L^2}{\log L}\Omega(K, K_1, K_2, K_3)t},$$

Recall the definition of the operator $\mathcal{T}_L$ as

$$\mathcal{T}_L(e,f,g)(K) \overset{def}{=} \frac{C_0}{L^2 \log L} \sum_{\mathcal{R}(K)} e_{K_1} f_{K_2} g_{K_3},$$

for sequences $e = \{e_K\}$, $e = \{f_K\}$ and $g = \{g_K\}$ on $\mathbb{Z}_L^2$. Hence the equation can be written

$$-i\partial_t b_K = \underbrace{\epsilon^2 \mathcal{T}_L(b,b,b)(K)}_{\text{resonant interactions}} + \underbrace{\frac{\epsilon^2 C_0}{L^2 \log L} \sum_{\mathcal{S}(K)\backslash\mathcal{R}(K)} b_{K_1}(t)\overline{b_{K_2}(t)}b_{K_3}(t)e^{i\frac{4\pi^2 C_0 L^2}{\log L}\Omega t}}_{\text{non-resonant interactions}}. \quad (5.2)$$

In sequel, the initial value $b_K(0)$ will be the projection of a function $g_0 : \mathbb{R}^2 \to \mathbb{C}$, that is $b_K(0) = g_0(K)$ for $K \in \mathbb{Z}_L^2$. With this rescaling, Theorem 2.4 follows once we show that for any $0 < \gamma < 1$, there exists c_γ such that if $\epsilon < c_\gamma L^{-1-\gamma}B^{-1}$, it holds that

$$\|b_K(t) - g(\epsilon^2 t, K)\|_{X_L^\sigma} \lesssim \frac{B}{(\log L)^{1-\gamma}} \quad (5.3)$$

for all $0 \le t \le \epsilon^{-2}\min(M, \gamma\frac{\log\log L}{c_\gamma B^2})$. If we require from the start that $L \ge \exp\exp\gamma^{-1}c_\gamma MB^2$, Theorem 2.4 follows as stated.

Before proving Theorem 2.4, we start with the following elementary lemma:

Lemma 5.1. *Let $\sigma > 2$. For any sequences $\{a_K\}, \{b_K\}$, and $\{c_K\}$ in X_L^σ, we have*

$$\left\| \sum_{K_1 - K_2 + K_3 = K} a_{K_1} b_{K_2} c_{K_3} \right\|_{X_L^\sigma} \lesssim L^4 \|a_K\|_{X_L^\sigma} \|b_K\|_{X_L^\sigma} \|c_K\|_{X_L^\sigma} \quad (5.4)$$

Proof. The proof is straightforward. Without loss of generality, we assume that the sequences $\{a_K\}, \{b_K\}$, and $\{c_K\}$ are non-negative. Then we have for $K \in \mathbb{Z}_L^2$,

$$\langle K\rangle^\sigma \sum_{K_1 - K_2 + K_3 = K} a_{K_1} b_{K_2} c_{K_3} \le \sum_{K_1 - K_2 + K_3 = K} \langle K_1\rangle^\sigma a_{K_1} b_{K_2} c_{K_3} + \sum_{K_1 - K_2 + K_3 = K} a_{K_1} \langle K_2\rangle^\sigma b_{K_2} c_{K_3}$$

$$+ \sum_{K_1 - K_2 + K_3 = K} a_{K_1} b_{K_2} \langle K_3\rangle^\sigma c_{K_3}.$$

We bound the contribution of the first terms; the other being similar.

$$\sum_{K_1 - K_2 + K_3 = K} \langle K_1\rangle^\sigma a_{K_1} b_{K_2} c_{K_2} \le \|a_K\|_{X_L^\sigma} \sum_{K_2, K_3} b_{K_2} c_{K_2} \le L^4 \|a_K\|_{X^\sigma} \|b_K\|_{\ell_L^1} \|c_K\|_{\ell_L^1}$$

$$\lesssim L^4 \|a_K\|_{X_L^\sigma} \|b_K\|_{X_L^\sigma} \|c_K\|_{X_L^\sigma}.$$

$\square$

Proof of Theorem 2.4. Let $g^\epsilon(t,\xi) \overset{def}{=} g(\epsilon^2 t, \xi)$. It satisfies the equation

$$-i\partial_t g^\epsilon(t,\xi) = \epsilon^2 \mathcal{T}(g^\epsilon, g^\epsilon, g^\epsilon)(t,\xi). \quad (5.5)$$

Note that $b_K(0) = g_0(K) \in X_L^{\sigma+1}$. From (5.2), the sequence $\{b_K\}$ satisfies the equation

$$-i\partial_t b_K = \epsilon^2 \mathcal{T}_L(g^\epsilon, g^\epsilon, g^\epsilon)(K) + \epsilon^2 \left[\mathcal{T}_L(b,b,b)(K) - \mathcal{T}_L(g^\epsilon, g^\epsilon, g^\epsilon)(K)\right]$$

$$+ \frac{\epsilon^2}{L^2 \log L} \sum_{\mathcal{S}(K)\backslash\mathcal{R}(K)} b_{K_1}\overline{b_{K_2}}b_{K_3}e^{i\frac{4\pi^2 C_0 L^2 \Omega}{\log L}t}.$$

Let $w_K(t) \overset{def}{=} b_K(t) - g(\epsilon^2 t, K)$ for $K \in \mathbb{Z}_L^2$ and $t \leq \varepsilon^{-2} M$. Then the equation satisfied by w_K is

$$
-i\partial_t w_K(t) = \epsilon^2 \left[T_L(b,b,b)(K) - T_L(g^\epsilon, g^\epsilon, g^\epsilon)(K) \right] + \epsilon^2 \left[\mathcal{T}_L(g^\epsilon, g^\epsilon, g^\epsilon)(K) - \mathcal{T}(g^\epsilon, g^\epsilon, g^\epsilon)(K) \right]
$$
$$
+ \frac{C_0 \epsilon^2}{L^2 \log L} \sum_{\mathcal{S}(K) \backslash \mathcal{R}(K)} b_{K_1} \overline{b_{K_2}} b_{K_3} e^{i \frac{4\pi^2 C_0 L^2 \Omega}{\log L} t}.
$$

Our first step is an application of a normal form transformation: For this we write

$$
\frac{C_0 \epsilon^2}{L^2 \log L} \sum_{\mathcal{S}(K) \backslash \mathcal{R}(K)} b_{K_1} \overline{b_{K_2}} b_{K_3} e^{i \frac{4\pi^2 C_0 L^2 \Omega}{\log L} t} = \partial_t \left(\frac{\epsilon^2}{L^2} \sum_{\mathcal{S}(K) \backslash \mathcal{R}(K)} \frac{1}{4\pi^2 i L^2 \Omega} b_{K_1} \overline{b_{K_2}} b_{K_3} e^{i \frac{4\pi^2 C_0 L^2 \Omega}{\log L} t} \right)
$$
$$
- \frac{\epsilon^2}{L^2} \sum_{\mathcal{S}(K) \backslash \mathcal{R}(K)} \frac{1}{4\pi^2 i L^2 \Omega} \partial_t \left(b_{K_1} \overline{b_{K_2}} b_{K_3} \right) e^{i \frac{4\pi^2 C_0 L^2 \Omega}{\log L} t}.
$$

As a result, if we define the change of coordinates

$$
e_K(t) \overset{def}{=} w_K(t) - \frac{\epsilon^2}{L^2} \sum_{\mathcal{S}(K) \backslash \mathcal{R}(K)} \frac{1}{4\pi^2 L^2 \Omega} b_{K_1}(t) \overline{b_{K_2}}(t) b_{K_3}(t) e^{i \frac{4\pi^2 C_0 L^2 \Omega}{\log L} t},
$$

then the equation satisfied by $e_K(t)$ is

$$
-i\partial_t e_K(t) = \epsilon^2 \left[\mathcal{T}_L(b,b,b)(K) - \mathcal{T}_L(g^\epsilon, g^\epsilon, g^\epsilon)(K) \right] + \epsilon^2 \left[\mathcal{T}_L(g^\epsilon, g^\epsilon, g^\epsilon)(K) - \mathcal{T}(g^\epsilon, g^\epsilon, g^\epsilon)(K) \right]
$$

$$
\tag{5.6}
$$

$$
- \frac{\epsilon^2}{L^2} \sum_{\mathcal{S}(K) \backslash \mathcal{R}(K)} \frac{1}{4\pi^2 i L^2 \Omega} \partial_t \left(b_{K_1} \overline{b_{K_2}} b_{K_3} \right) e^{i \frac{4\pi^2 C_0 L^2 \Omega}{\log L} t}, \tag{5.7}
$$

$$
e_K(0) = - \frac{\epsilon^2}{L^2} \sum_{\mathcal{S}(K) \backslash \mathcal{R}(K)} \frac{1}{4\pi^2 L^2 \Omega} b_{K_1}(0) \overline{b_{K_2}(0)} b_{K_3}(0) \tag{5.8}
$$

Of course, the advantage of working with e_K as opposed to b_K is the fact that (5.7) is quintic in b_K.

The proof follows a bootstrap argument. Recall that $\|g^\epsilon\|_{X^\sigma} \leq B$ for all $0 \leq t \leq \epsilon^{-2} M$. We will show that we can find a constant $c = c_\gamma$, such that if

$$
\epsilon < c_\gamma L^{-1-\gamma} B^{-1} \quad \text{and} \quad \sup_{t \in [0,T]} \|b_K(t)\|_{X_L^\sigma} \leq 2B
$$

for some $T \leq \epsilon^{-2} \min(M, cB^{-2} \log \log L)$, then (5.3) holds on the interval $[0, T]$. The result then follows on the full interval $[0, \epsilon^{-2} \min(M, cB^{-2} \log \log L)]$ via a continuity argument.

We start by proving that our normal form transformation from $w_K \to e_K$ is close to the identity. Indeed, recalling the definition of (3.22) of the set $\mathcal{R}_\mu$ for $\mu \in \mathbb{R}$, we can write

$$
\begin{aligned}
\|w_K(t) - e_K(t)\|_{X_L^\sigma} \lesssim &\frac{\epsilon^2}{L^2} \left\| \sum_{m \in \mathbb{Z} \backslash \{0\}} m^{-1} \sum_{\mathcal{R}_{mL^{-2}}(K)} b_{K_1} \overline{b_{K_2}} b_{K_3} e^{i \frac{4\pi^2 C_0 L^2 \Omega}{\log L} t} \right\|_{X_L^\sigma} \\
\leq &\frac{\epsilon^2}{L^2} \left\| \sum_{0 < |m| < L^{10}} m^{-1} \sum_{\mathcal{R}_{mL^{-2}}(K)} b_{K_1} \overline{b_{K_2}} b_{K_3} e^{i \frac{4\pi^2 C_0 L^2 \Omega}{\log L} t} \right\|_{X_L^\sigma} \\
&+ \frac{\epsilon^2}{L^2} \left\| \sum_{|m| \geq L^{10}} m^{-1} \sum_{\mathcal{R}_{mL^{-2}}(K)} b_{K_1} \overline{b_{K_2}} b_{K_3} e^{i \frac{4\pi^2 C_0 L^2 \Omega}{\log L} t} \right\|_{X_L^\sigma} \\
\leq &\frac{\epsilon^2}{L^2} \sum_{0 < |m| < L^{10}} |m|^{-1} \left\| \sum_{\mathcal{R}_{mL^{-2}}(K)} b_{K_1} \overline{b_{K_2}} b_{K_3} e^{i \frac{4\pi^2 C_0 L^2 \Omega}{\log L} t} \right\|_{X_L^\sigma} \\
&+ \frac{\epsilon^2}{L^2} L^{-10} \left\| \sum_{\mathcal{S}(K)} |b_{K_1}| |b_{K_2}| |b_{K_3}| \right\|_{X_L^\sigma} \\
\lesssim_\gamma &\epsilon^2 L^\gamma B^3
\end{aligned}
$$

for any $\gamma > 0$, where we have used the bootstrap assumption that $\|b_K\|_{X_L^\sigma} \leq 2B$, Proposition 3.9 to bound the first sum, and Lemma 5.1 to bound the second sum. As a result,

$$
\|w_K(t) - e_K(t)\|_{X^\sigma} \lesssim_\gamma \epsilon^2 L^\gamma B^3. \tag{5.9}
$$

We start by estimating the terms on the right-hand-side of (5.6). We start with the first term:

$$
\epsilon^2 \|\mathcal{T}_L(b, b, b)(K) - \mathcal{T}_L(g^\epsilon, g^\epsilon, g^\epsilon)(K)\|_{X^\sigma} \lesssim \epsilon^2 B^2 \|w_K\|_{X_L^\sigma} \lesssim \epsilon^2 B^2 \|e_K\|_{X_L^\sigma} + c_\gamma \epsilon^4 L^\gamma B^5.
$$

where we used Theorem 2.7 and the trilinearity of $\mathcal{T}_L$ in the first step. The second term on the right-hand-side of (5.6) can be estimated using Theorem 2.3, which gives

$$
\epsilon^2 \|T_L(g^\epsilon, g^\epsilon, g^\epsilon)(K) - \mathcal{T}(g^\epsilon, g^\epsilon, g^\epsilon)(K)\|_{X^\sigma} \lesssim \frac{\epsilon^2 B^3}{\log L}.
$$

Now notice that (5.7) can be written as:

$$
(5.7) = -\frac{\epsilon^2}{4\pi^2 L^2} \sum_{m \in \mathbb{Z} \backslash \{0\}} (im)^{-1} \sum_{\mathcal{R}_{mL^{-2}}(K)} (\partial_t b_{K_1}) \overline{b_{K_2}} b_{K_3} e^{4\pi^2 i \frac{L^2 \Omega}{\log L} t}
$$

$$
+ \text{ two similar terms.}
$$

As a result, we can estimate as before using Proposition 3.9 and Lemma 5.1

$$
\begin{aligned}
\|(5.7)\|_{X_L^\sigma} \lesssim &\frac{\epsilon^2}{L^2} \sum_{0 < |m| < L^{10}} |m|^{-1} \left\| \sum_{\mathcal{R}_{mL^{-2}}(K)} \partial_t b_{K_1} \overline{b_{K_2}} b_{K_3} e^{i \frac{4\pi^2 C_0 L^2 \Omega}{\log L} t} \right\|_{X_L^\sigma} + \frac{\epsilon^2}{L^2} L^{-10} \left\| \sum_{\mathcal{S}(K)} |\partial_t b_{K_1}| |b_{K_2}| |b_{K_3}| \right\|_{X_L^\sigma} \\
\lesssim_\gamma &\frac{\epsilon^2}{L^2} L^{2+\gamma} \|\partial_t b_K\|_{X_L^\sigma} \|b_K\|_{X_L^\sigma}^2.
\end{aligned}
$$

Using once again Lemma 5.1 gives

$$\|(5.7)\|_{X_L^\sigma} \lesssim_\gamma \epsilon^4 L^{2+\gamma} \|b_K\|_{X_L^\sigma}^5 \le c_\gamma \epsilon^4 L^{2+\gamma} B^5.$$

Collecting all the above estimates, we get that $\|e_K\|_{X_L^\sigma}$ satisfies the following differential inequality:

$$\|\partial_t e_K(t)\|_{X^\sigma(\mathbb{Z}_L^2)} \le c\epsilon^2 B^2 \|e_K\|_{X_L^\sigma} + c\frac{\epsilon^2 B^3}{\log L} + c_\gamma \epsilon^4 L^{2+\gamma} B^5 \quad \text{and} \quad \|e_K(0)\|_{X_L^\sigma} \le c_\gamma \epsilon^2 L^\gamma B^3,$$

for some universal constant c and a constant c_γ depending on γ. An application of Gronwall's lemma then implies that, if $0 \le t \le T \le \epsilon^{-2} M$, we have

$$\|e_K(t)\|_{X_L^\sigma} \le \|e_K(0)\|_{X_L^\sigma} e^{c\epsilon^2 B^2 t} + (e^{c\epsilon^2 B^2 t} - 1)\left(\frac{B}{\log L} + \frac{c_\gamma}{c}\epsilon^2 L^{2+\gamma} B^3\right).$$

As a consequence, if $\epsilon < c'_\gamma L^{-1-\gamma} B^{-1}$, then it holds that

$$\|e_K(t)\|_{X_L^\sigma} \le e^{c\epsilon^2 B^2 t} \frac{2B}{\log L}.$$

Using in addition $t \le T \le \gamma \frac{\log \log L}{c\epsilon^2 B^2}$, we get that

$$\|e_K(t)\|_{X_L^\sigma} \le \frac{2B}{(\log L)^{1-\gamma}}.$$

In conclusion, combining the above estimate with (5.9) we get

$$\|b_K - g^\epsilon\|_{X_L^\sigma} = \|w_K\|_{X_L^\sigma} \le \|e_K\|_{X_L^\sigma} + \|w_K - e_K\|_{X_L^\sigma} \le \frac{B}{L^{2+\gamma}} + \frac{2B}{(\log L)^{1-\gamma}}.$$

This completes the bootstrap argument, and the proof. $\square$

Finally, we give the simple proof of how to deduce Corollary 2.6 from the above theorem.

Proof of Corollary 2.6. Shrink γ if needed so that $\gamma \le \frac{s-1}{100}$. Define $(\widetilde{a}_K(t))_{K \in \mathbb{Z}_L^2}$ such that $e^{-4\pi^2 it|k|^2}\widehat{v}(t,k) = \frac{\epsilon}{N}\widetilde{a}_{\frac{k}{N}}(N^2 t)$ for $k \in \mathbb{Z}$. Then $\widetilde{a}_K(t)$ satisfies (1.3).

Choose $\epsilon = N^{-s} < c_\gamma N^{-1-\gamma} B^{-1}$ if $N \ge CB^{\frac{s-1}{2}}$ thanks to our choice γ, and apply Theorem 2.4 for $\widetilde{a}_K$ to get for $T^* = \frac{\zeta(2)N^2}{2\epsilon^2 \log N}$

$$\|\widetilde{a}_K(t) - g(\frac{t}{T^*}, K)\|_{X_L^\sigma(\mathbb{Z}_N^2)} \lesssim \frac{B}{(\log N)^{1-\gamma}}$$

for all $0 \le t \le T^* \min\left(M, \frac{\gamma \log \log N}{cB^2}\right)$. This gives that

$$\|e^{-4\pi^2 itN^2|K|^2}\widehat{v}(t, NK) - \frac{\epsilon}{N}g(\frac{N^2 t}{T^*}, K)\|_{X_N^\sigma(\mathbb{Z}_N^2)} \lesssim \frac{\epsilon B}{N(\log N)^{1-\gamma}}$$

for all $0 \le t \le N^{-2}T^* \min\left(M, \frac{\gamma \log \log N}{cB^2}\right)$. Letting $T_N \overset{def}{=} \frac{\zeta(2)N^2}{2\epsilon^2 \log N}$, we get that since $\sigma > s+1$ that

$$\|e^{-4\pi^2 itN^2|K|^2}\widehat{v}(t, NK) - \frac{\epsilon}{N}g(\frac{t}{T_N}, K)\|_{\ell_N^{2,s}(\mathbb{Z}_N^2)} \lesssim \frac{\epsilon B}{N(\log N)^{1-\gamma}},$$

for all $0 \le t \le T_N \min\left(M, \frac{\gamma \log \log N}{cB^2}\right)$. This translates on $\mathbb{Z}^2$ to the estimate

$$\|e^{-4\pi^2 it|k|^2}\widehat{v}(t,k) - \frac{\epsilon}{N}g(\frac{t}{T_N}, \frac{k}{N})\|_{\ell^{2,s}(\mathbb{Z}^2)} \lesssim \frac{(\epsilon N^s)B}{(\log N)^{1-\gamma}},$$

over the same time interval. Recalling that $\epsilon = N^{-s}$ gives (2.14).   □

6. Hamiltonian structure of (CR)

Recall that the (CR) equation is given by

$$-i\partial_t g(\xi, t) = \mathcal{T}(g, g, g)(\xi, t); \qquad \xi \in \mathbb{R}^2$$

$$\mathcal{T}(g, g, g)(\xi, t) = \int_{-1}^{1} \int_{\mathbb{R}^2} g(\xi + \lambda z, t)\overline{g}(\xi + \lambda z + z^\perp)g(\xi + z^\perp)\, dz\, d\lambda. \qquad \text{(CR)}$$

By changing variables (formally for now) one can notice that the trilinear operator $T(g, g, g)$ can be written in several different ways like

$$\mathcal{T}(g, g, g)(\xi, t) = \int_{-1}^{1} \int_{\mathbb{R}^2} g(\xi + z)\overline{g}(\xi + \lambda z^\perp + z)g(\xi + \lambda z^\perp, t)\, dz\, d\lambda$$

$$= \frac{1}{2} \int_{\mathbb{R}} \int_{\mathbb{R}^2} g(\xi + \lambda z, t)\overline{g}(\xi + \lambda z + z^\perp)g(\xi + z^\perp)\, dz\, d\lambda.$$

In the latter integral, λ runs over $\mathbb{R}$ instead of $[-1, 1]$, as can be seen by a simple change of integration variables $\lambda \to \lambda^{-1}$.

6.1. The Hamiltonian functional. The equation (CR) derives from the Hamiltonian functional

$$\mathcal{H}(f) = \frac{1}{4} \int_{-1}^{1} \int_{\mathbb{R}^2} \int_{\mathbb{R}^2} f(\xi + z)f(\xi + \lambda z^\perp)\overline{f(\xi + z + \lambda z^\perp)}\,\overline{f(\xi)}\, dz\, d\xi\, d\lambda \qquad (6.1)$$

given the symplectic form

$$\omega(f, g) = \mathfrak{Im} \int_{\mathbb{R}^2} f(x)\overline{g(x)}\, dx. \qquad (6.2)$$

Let us notice right away that the Hamiltonian can be equivalently written

$$\mathcal{H}(f) = \frac{1}{8} \int_{\mathbb{R}} \int_{\mathbb{R}^2} \int_{\mathbb{R}^2} f(\xi + z)f(\xi + \lambda z^\perp)\overline{f(\xi + z + \lambda z^\perp)}\,\overline{f(\xi)}\, dz\, d\xi\, d\lambda. \qquad (6.3)$$

A remarkable identity is the fact that

$$\mathcal{H}(f) = \frac{\pi}{2} \int_{\mathbb{R}_t \times \mathbb{R}_x^2} \left|e^{it\Delta_{\mathbb{R}^2}}\widehat{f}\right|\, dx\, dt = \frac{\pi}{2}\|e^{it\Delta_{\mathbb{R}^2}}\widehat{f}\|_{L_{t,x}^4(\mathbb{R}_t \times \mathbb{R}_x^2)}^4 \qquad (6.4)$$

which can be seen by expanding the right-hand-side as follows. Suppose first that $f \in \mathcal{S}(\mathbb{R}^2)$ (the case of general $f \in L^2$ follows by a limiting argument using (2.18) which we prove later). Then

$$\|e^{it\Delta}\widehat{f}\|_{L^4}^4 = \||e^{it\Delta}\widehat{f}|^2\|_{L^2}^2 = \frac{1}{4\pi^2}\|e^{-it|\xi|^2}f(\xi) * e^{it|\xi|^2}\overline{f(\xi)}\|_{L^2}^2$$

$$= \frac{1}{4\pi^2} \int_{\mathbb{R}^2} \int_{\mathbb{R}^2} \int_{\mathbb{R}^2} e^{it2\xi \cdot \alpha} f(\eta + \alpha)\overline{f(\eta + \alpha + \xi)}\,\overline{f(\eta)}f(\eta + \xi)\, d\xi\, d\eta\, d\alpha$$

$$= \frac{1}{4\pi^2} \int_{\mathbb{R}^2} \int_{\mathbb{R}^2} \int_{\mathbb{R}} \int_{\mathbb{R}} e^{i2t\mu|\alpha|^2} f(\eta + \alpha)\overline{f(\eta + \alpha + \mu\alpha + \lambda\alpha^\perp)}\,\overline{f(\eta)}f(\eta + \mu\alpha + \lambda\alpha^\perp)|\alpha|^2 d\lambda\, d\mu\, d\eta\, d\alpha,$$

where we changed in the last line coordinates by $\xi = \mu\alpha + \lambda\alpha^\perp$. Integrating in time and using the identity $\int_\mathbb{R} e^{ixy} dy = 2\pi\delta_{x=0}$ gives

$$\|e^{it\Delta}\widehat{f}\|_{L^4}^4 = \frac{1}{4\pi} \int \int \int f(\eta + \alpha)\overline{f(\eta + \alpha + \lambda\alpha^\perp)}\,\overline{f(\eta)}f(\eta + \lambda\alpha^\perp)d\lambda\,d\eta\,d\alpha = \frac{2}{\pi}\mathcal{H}(f).$$

In particular, $\mathcal{H}(f)$ is positive definite, i.e. $\mathcal{H}(f) \geq 0$ and $\mathcal{H}(f) = 0$ if and only if $f = 0$.

6.2. Symmetries.

Proposition 6.1. *The following symmetries leave the Hamiltonian $\mathcal{H}$ invariant:*
  (i) *Phase rotation: $f \mapsto e^{i\theta}f$.*
  (ii) *Translation: $f \mapsto f(\cdot + x_0)$ for any $x_0 \in \mathbb{R}^2$.*
  (iii) *Modulation: $f \mapsto e^{ix\cdot\xi_0}f$ for any $\xi_0 \in \mathbb{R}^2$.*
  (iv) *Quadratic modulation: $f \mapsto e^{i\tau|x|^2}f$ for any $\tau \in \mathbb{R}$.*
  (v) *Schrödinger group: $f \mapsto e^{i\tau\Delta}f$ for any $\tau \in \mathbb{R}$.*
  (vi) *Rotation: $f \mapsto f(R_\theta\,\cdot)$, for any θ, where R_θ is the rotation of angle θ.*
  (vii) *Scaling $f \mapsto \mu f(\mu\,\cdot)$ for any $\mu > 0$.*
  (viii) *Fourier transform $f \mapsto \widehat{f}$.*

Proof. The points (i), (ii), (iii), (iv), (vi), (vii) follow simply from the formula giving $\mathcal{H}$, in particular from the fact that the arguments of f in this formula satisfy the relations

$$\begin{cases} (\xi + z) + (\xi + \lambda z^\perp) = (\xi + z + \lambda z^\perp) + (\xi) \quad \text{and} \\ |\xi + z|^2 + |\xi + \lambda z^\perp|^2 = |\xi + z + \lambda z^\perp|^2 + |\xi|^2. \end{cases}$$

The point (v) is a consequence of (iv) and (viii). Thus we are left with proving (viii); this will be done in Lemma 6.5. □

Since the transformations (i) to (vi) are generated by Hamiltonian flows, and as a result they give, by Noether's theorem, conserved quantities for the equation (CR).

Corollary 6.2. *The following quantities are conserved by the flow of* (CR):
  (i) *Mass: $\int |f(x)|^2\,dx$.*
  (ii) *Momentum: $\int \xi \left|\widehat{f}(\xi)\right|^2 d\xi$.*
  (iii) *Position: $\int x|f(x)|^2\,dx$.*
  (iv) *First moment $\int |x|^2|f(x)|^2\,dx$.*
  (v) *Kinetic energy: $\int |\nabla f(x)|^2\,dx$.*
  (vi) *Angular momentum $i\int (x \times \nabla)f(x)\overline{f(x)}\,dx$.*
  (vii) *Hamiltonian $\mathcal{H}(f)$.*

The following is essentially a corollary of Theorem 6.1.

Corollary 6.3. *The symmetries mentioned in Proposition 6.1 (acting only in the x variable) leave the set of solutions of* (CR) *invariant. This is also the case for the transformation*

$$f \mapsto \lambda^\alpha \left(\lambda^{2\alpha - 2\gamma}t\,,\,\lambda^\gamma x\right) \qquad \text{where } \lambda, \alpha, \gamma > 0. \tag{6.5}$$

Remark 6.4 (scaling)**.** *It is natural to ask which is the critical space for* (CR)*. Since it enjoys a family of possible scalings, as noted above, the answer is not as clear as for, say,* (NLS)*. However, it seems natural to think of L^2 as a critical space for data. Indeed,*
  • *The space L^2 is critical for* (NLS)*, from which* (CR) *is derived.*

- *It is essentially the largest Sobolev space for which local well-posedness can be proved, see Section 7.*
- *Finally, the only scaling above which does not act on the time variable, namely $f \mapsto \lambda f(\lambda \cdot)$, leaves the norm of L^2 invariant.*

Lemma 6.5. *If $f, g, h \in L^2$,*

$$\mathcal{F}\left(\mathcal{T}(f, g, h)\right) = \mathcal{T}\left(\widehat{f}, \widehat{g}, \widehat{h}\right) \quad and \quad \mathcal{H}(f) = \mathcal{H}(\widehat{f}).$$

Remark 6.6. *Why this Fourier transform symmetry actually holds is somewhat mysterious. It is reminiscent of the invariance of the cubic nonlinear Schrödinger equation set on $\mathbb{R}^2$ by the pseudo-conformal transformation*

$$u(t, x) \mapsto u_c(t) \overset{def}{=} \frac{1}{t} e^{i \frac{|x|^2}{4t}} u\left(-\frac{1}{t}, \frac{x}{t}\right).$$

which can also be written at the level of profiles $f = e^{-it\Delta_{\mathbb{R}^2}} u$ and $f_c = e^{-it\Delta_{\mathbb{R}^2}} u_c$ as

$$f(t) = \overline{\mathcal{F} f_c}(t^{-1}).$$

Also notice that the well known fact that

$$as \ t \to \infty, \qquad \left\| e^{it\Delta} \psi(x) - \frac{1}{4\pi i t} e^{i \frac{|x|^2}{4t}} \widehat{\psi}\left(\frac{x}{2t}\right) \right\|_{L^2} \longrightarrow 0,$$

if $\psi \in \mathcal{S}(\mathbb{R}^2)$, combined with L^2 boundedness of $\mathcal{H}$, implies that points (iii) and (vii) of Proposition 6.1 are equivalent.

Proof. Again we assume that all functions are in Schwartz class $\mathcal{S}(\mathbb{R}^2)$. The case of functions in L^2 follows from a standard limiting argument using Lemma 7.1. The second equality is a consequence of the first and Plancherel's theorem:

$$\mathcal{H}(f) = \frac{1}{4} \left\langle \mathcal{T}(f, f, f), f \right\rangle = \frac{1}{4} \left\langle \mathcal{F}\mathcal{T}(f, f, f), \widehat{f} \right\rangle = \frac{1}{4} \left\langle \mathcal{T}\left(\widehat{f}, \widehat{f}, \widehat{f}\right), \widehat{f} \right\rangle = \mathcal{H}\left(\widehat{f}\right).$$

We now prove the first identity: By Fourier inversion and using the identity $\frac{1}{4\pi^2} \int e^{ix \cdot \xi} \, dx = \delta_{\xi=0}$ we compute

$$\mathcal{F}\mathcal{T}(f, g, h)(\xi) = \frac{1}{2\pi} \int_{\mathbb{R}^2} \int_{-1}^{1} \int_{\mathbb{R}^2} e^{-iz \cdot \xi} f(x + z) g(\lambda x^\perp + z) \overline{h(x + \lambda x^\perp + z)} \, dx \, d\lambda \, dz$$

$$= \frac{1}{16\pi^4} \int \int_{-1}^{1} \int \int \int \int e^{iz \cdot (-\xi + \alpha + \beta - \gamma)} e^{ix \cdot (\alpha - \lambda \beta^\perp + \lambda \gamma^\perp - \gamma)} \widehat{f}(\alpha) \widehat{g}(\beta) \overline{\widehat{h}(\gamma)} \, d\alpha \, d\beta \, d\gamma \, dx \, d\lambda \, dz$$

$$= \int_{-1}^{1} \int \widehat{f}\left(\xi + \frac{1}{\lambda}\xi^\perp - \frac{1}{\lambda}\beta^\perp\right) \widehat{g}(\beta) \overline{\widehat{h}\left(\frac{1}{\lambda}\xi^\perp + \beta - \frac{1}{\lambda}\beta^\perp\right)} \, d\beta \, \frac{d\lambda}{\lambda^2}.$$

Changing variables to $\beta' = \frac{1}{\lambda}(\xi - \beta)^\perp$ and then omitting the primes gives

$$\mathcal{F}\mathcal{T}(f, g, h)(\xi) = \int_{-1}^{1} \int \widehat{f}(\xi + \beta) \widehat{g}(\xi + \lambda \beta^\perp) \overline{\widehat{h}(\beta + \lambda \beta^\perp + \xi)} \, d\beta \, d\lambda,$$

which is the desired result. $\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad \square$

6.3. **Invariance of eigenspaces of the Harmonic oscillator.** Recall that the eigenvalues of the harmonic oscillator $-\Delta + |x|^2$ on $L^2(\mathbb{R}^2)$ are given by $2k$, $k \in \mathbb{N}$. The smallest eigenvalue, 2, has an eigenspace E_2 of dimension 1, generated by the Gaussian $e^{-|x|^2/2}$. The eigenspace E_{2k} has dimension k, and is generated by Hermite functions.

Proposition 6.7. *The eigenspaces E_k, $k \in 2\mathbb{N} + 2$, of the harmonic oscillator are invariant by the flow of* (CR).

Proof. Let us introduce some notations: Given a functional $\mathcal{F}$, denote its symplectic gradient with respect to the symplectic form ω (see (6.2)) by $\nabla_\omega \mathcal{F}$, thus $d\mathcal{F}(X) = \omega(X, \nabla_\omega \mathcal{F})$; denote further $\{\mathcal{F}, \mathcal{G}\}$ for the Poisson bracket $\omega(\nabla_\omega \mathcal{F}, \nabla_\omega \mathcal{G})$; finally $[X, Y]$ is the Lie bracket of the vector fields X and Y. Then the Hamiltonian $\mathcal{G}$ is invariant by the Hamiltonian flow induced by $\mathcal{F}$ if and only if $\{\mathcal{F}, \mathcal{G}\} = 0$. Furthermore, the Hamiltonian flows of $\mathcal{F}$ and $\mathcal{G}$ commute if and only if $[\nabla_\omega \mathcal{F}, \nabla_\omega \mathcal{G}]$. The formula $[\nabla_\omega \mathcal{F}, \nabla_\omega \mathcal{G}] = \nabla_\omega \{\mathcal{F}, \mathcal{G}\} = 0$ (see [1]) implies that: if $\mathcal{G}$ is invariant by the flow of $\mathcal{F}$, then the flows of $\mathcal{F}$ and $\mathcal{G}$ commute.

We saw that the transformations $e^{it\Delta}$ (the flow associated with the Hamiltonian $\mathcal{K}_1 \overset{def}{=} \int |\nabla f|^2$) and $e^{it|x|^2}$ (the flow of the Hamiltonian $\mathcal{K}_2 \overset{def}{=} \int |xf|^2$) both leave $\mathcal{H}$ invariant. If we denote by $\{\cdot, \cdot\}$ the Poisson bracket associated with the symplectic form (6.2) this implies that $\{\mathcal{H}, \mathcal{K}_1\} = 0$ and $\{\mathcal{H}, \mathcal{K}_2\} = 0$, and hence $\{\mathcal{H}, \mathcal{K}_1 + \mathcal{K}_2\} = 0$. Now, the flow associated with the Hamiltonian $\mathcal{K}_1 + \mathcal{K}_2$ is the transformation $e^{it(-\Delta + |x|^2)}$. By the preceding paragraph, this implies that the unitary flow of the harmonic oscillator and the flow of $\mathcal{H}$ commute.

Using in addition that phase rotations commute with the flow of $\mathcal{H}$ (which we denote here by $U(t)$), we get if $f \in E_k$: for any s, t

$$e^{is(-\Delta + |x|^2)} U(t) f = U(t) e^{is(-\Delta + |x|^2)} f = U(t) e^{isk} f = e^{isk} U(t) f.$$

The equality $e^{is(-\Delta + |x|^2)} U(t) f = e^{isk} U(t) f$ implies that $U(t) f \in E_k$. $\square$

6.4. **Stationary waves.** We are concerned here with solutions of the type $f = e^{i\omega t} \phi$; thus ϕ verifies

$$\omega \phi = \mathcal{T}(\phi, \phi, \phi).$$

First notice that the boundedness of $\mathcal{T}$ on L^2 (see Section 7 and (8.2)) implies, after taking the scalar product of the above with ϕ, that

$$0 \leq \omega \leq \frac{\pi}{8} \|\phi\|_2^2.$$

(this is in conformity the scaling of the equation of course). Many stationary solutions can be uncovered by examining the eigenspaces of the harmonic oscillator which, as we know from the previous subsection, are invariant by the flow of (CR). We give a few examples:

- The first eigenspace E_2 is generated by the Gaussian $e^{-\frac{|x|^2}{2}}$; its stability and the conservation of the L^2 norm imply that for some constant ω_0, $e^{i\omega_0 t} e^{-\frac{|x|^2}{2}}$ is an exact solution. In fact, this is equivalent to saying that

$$\mathcal{T}(e^{-\frac{|x|^2}{2}}, e^{-\frac{|x|^2}{2}}, e^{-\frac{|x|^2}{2}}) = \omega_0 e^{-\frac{|x|^2}{2}}, \tag{6.6}$$

  which can also easily be verified explicitly.

Letting the symmetries of the equation act on this Gaussian, we obtain the family of stationary waves: for $(\alpha, \beta, x_0, v_0, \gamma) \in \mathbb{R}_+ \times \mathbb{R} \times \mathbb{R}^2 \times \mathbb{R}^2 \times \mathbb{C}$,

$$\gamma e^{-(\frac{\alpha}{2}+i\beta)|x-x_0|^2+i\left(v_0\cdot x+\omega_0\frac{|\gamma|^2}{\alpha^2}t\right)}.$$

- The second eigenspace E_4 is generated by $x_1 e^{-\frac{|x|^2}{2}}$ and $x_2 e^{-\frac{|x|^2}{2}}$. Using the invariance of E_4, conservation of the L^2 norm, and of the oddness with respect to x_1, we obtain that $x_1 e^{-\frac{|x|^2}{2}+i\omega_1 t}$ is an exact solution. By invariance by rotation, we obtain the family of stationary solutions

$$(\cos\theta x_1 + \sin\theta x_2)e^{-\frac{|x|^2}{2}+i\omega_1 t},$$

  where ω_1 is a constant and θ arbitrary in $\mathbb{R}$ (more stationary solutions can then be obtained by applying the symmetries).
- The third eigenspace E_6 is generated by $x_1 x_2 e^{-\frac{|x|^2}{2}}$, $(2x_1^2-1)e^{-\frac{|x|^2}{2}}$ and $(2x_2^2-1)e^{-\frac{|x|^2}{2}}$. By conservation of the L^2 norm, invariance of E_6, and invariance of the set of radial functions, we find new exact solutions of the form

$$(2|x|^2 - 1)e^{-\frac{|x|^2}{2}+\omega_2 t}$$

  for a constant ω_2.

Orbital stability of Gaussians with respect to perturbations in $L^{2,1} \cap H^1$ can be easily established by relying on the conserved quantities $\int |f|^2$, $\int |xf|^2$, $\int |\nabla f|^2$. This is done in the following proposition.

Proposition 6.8. *Consider data f_0 such that $\|f_0 - e^{-\frac{|x|^2}{2}}\|_{H^1 \cap L^{2,1}} = \delta > 0$, and let $f(t)$ be the corresponding solution of* (CR). *Then*

$$\sup_t \mathrm{dist}_{H^1 \cap L^{2,1}}(f(t), \mathbb{S}^1 e^{-\frac{|x|^2}{2}}) \lesssim \sqrt{\delta},$$

where we denoted $\mathbb{S}^1 e^{-\frac{|x|^2}{2}}$ for $\{\gamma e^{-\frac{|x|^2}{2}}, \gamma \in \mathbb{C} \text{ and } |\gamma| = 1\}$.

Proof. Let us denote by P_k, with $k \in 2\mathbb{N}$, the projectors on the eigenspaces of $-\Delta + |x|^2$. Notice first that the assumption that $\|f_0 - e^{-\frac{|x|^2}{2}}\|_{H^1 \cap L^{2,1}} = \delta$, implies that

$$\left| \|f_0\|_{L^2}^2 - \|e^{-\frac{|x|^2}{2}}\|_{L^2}^2 \right| \lesssim \delta$$

and that

$$\left| \langle(-\Delta + |x|^2)f_0, f_0\rangle - \langle(-\Delta + |x|^2)e^{-\frac{|x|^2}{2}}, e^{-\frac{|x|^2}{2}}\rangle \right| \lesssim \delta.$$

Furthermore, these two inequalities remain true if one replaces f_0 by $f(t)$, which follows by using the conserved quantities of the equation. In terms of the spectral projectors P_k, this

means

$$\left| \sum_{k \in 2\mathbb{N}} \|P_k f(t)\|_{L^2}^2 - 4\pi^2 \right| \lesssim \delta$$

$$\left| \sum_{k \in 2\mathbb{N}} k^2 \|P_k f(t)\|_{L^2}^2 - 16\pi^2 \right| \lesssim \delta.$$

This implies that

$$\sum_{k \in 2\mathbb{N}+2} k^2 \|P_k f(t)\|_{L^2}^2 \lesssim \delta,$$

which easily implies the desired result. □

Finally, another interesting explicit solution is given by

$$e^{i\omega_3 t} \frac{1}{|x|} \qquad \text{for a constant } \omega_3.$$

This follows easily from the identity $\mathcal{T}(\frac{1}{|x|}, \frac{1}{|x|}, \frac{1}{|x|}) = \frac{\omega_3}{|x|}$, which will be established in Section 7. Once again, applying the symmetries of the equation (detailed in Section 6) gives new solutions. The action of these symmetries is easily described, except for the expression $e^{is\Delta} \frac{1}{|x|}$. We refer to the book of Cazenave [14], where it is shown that it is L^∞, smooth, and decaying like $\frac{1}{|x|}$ for $s > 0$.

As mentioned before the solution $g(t, \xi) = e^{i\omega_3 t} |\xi|^{-1}$ carries particular significance in terms of its relation to wave turbulence theory as it corresponds to the Raleigh-Jeans solution $n(\xi) = |\xi|^{-2}$ of the (KZ) equation mentioned in Section 2 (recall that $n(\xi, t)$ is related to the square of the Fourier modes).

7. Analytic properties

The analytic properties of solutions of (CR) reflect essentially those of $\mathcal{T}$. This operator does not seem to belong to a well-studied class of operators. If one fixes λ and integrates only over y in the definition (2.8), the operator under consideration belongs to the class studied by Brascamp and Lieb [9]; however, the integration over λ seems responsible for some important properties of $\mathcal{T}$, such as Theorem 7.8 below.

The feature of $\mathcal{T}$ which differs from the setting of Brascamp and Lieb is that the arguments of f, g, and h are nonlinear functions of λ and x. A nonlinear Brascamp-Lieb theory is not available, but a nonlinear analog of the related Loomis-Whitney inequality has been established by Bennett, Carbery and Wright [3]; we will make use of it.

7.1. Boundedness of $\mathcal{T}$.

Proposition 7.1. *For the following spaces X, the trilinear operator $\mathcal{T}$ is bounded from X^3 to X:*

  (i) L^2.
  (ii) $\dot{L}^{\infty,1}$.
  (iii) $\dot{L}^{p,1-\frac{2}{p}}$ *for $p \geq 2$.*
  (iv) $L^{2,\sigma}$ *for any $\sigma \geq 0$.*
  (v) H^σ *for any $\sigma \geq 0$.*
  (vi) $Y^{\sigma,p} \stackrel{def}{=} \{f(x) \,|\, e^{\sigma|x|^2} f(x) \in L^p\}$ *for any $\sigma \geq 0$, $p \geq 2$.*

Remark 7.2. *Due to (6.4), the statement (i) is equivalent to the classical L^4 Strichartz estimate [61, 56]; the proof we give is very simple, and seems not be found in the literature.*

Notice that the three first statements above are at the scaling of the equation; the three last spaces are subcritical, thus they will give for the equation the possibility to propagate smoothness, or localization.

Proof. To prove (i), we argue by duality and use the Cauchy-Schwarz inequality to obtain

$$\langle \mathcal{T}(f,g,h), F \rangle = \int_{-1}^{1} \int \int f(x+z) g(\lambda x^\perp + z) \overline{h(x + \lambda x^\perp + z)} \, \overline{F(z)} \, dx \, dz \, d\lambda$$

$$\lesssim \int_{-1}^{1} \left(\int \int |f(x+z)|^2 |F(z)|^2 \, dx \, dz \right)^{1/2}$$

$$\left(\int \int |h(x+z)|^2 |g(z)|^2 \, dx \, dz \right)^{1/2} d\lambda$$

$$\lesssim \|f\|_{L^2} \|g\|_{L^2} \|h\|_{L^2} \|F\|_{L^2}.$$

To prove (ii), it suffices to establish that

$$\mathcal{T} \left(\frac{1}{|x|}, \frac{1}{|x|}, \frac{1}{|x|} \right) \leq \frac{C}{|x|},$$

this will be done in Lemma 7.3 below. Indeed, equality holds!

The point (iii) can be obtained from (i) and (ii) and complex interpolation.

To prove (iv), proceed as for (i), using in addition the inequality

$$\text{for any } \lambda \in \mathbb{R}, \, \xi, z \text{ in } \mathbb{R}^2, \qquad \frac{\langle \xi \rangle^\sigma}{\langle \xi + z \rangle^\sigma \langle \xi + z + \lambda z^\perp \rangle^\sigma \langle \xi + \lambda z^\perp + z \rangle^\sigma} \lesssim 1.$$

which follows from the relation $\xi = (\xi + z) - (\xi + z + \lambda z^\perp) + (\xi + \lambda z^\perp)$.

The assertion (v) is a consequence of (iv) and the invariance of $\mathcal{T}$ under Fourier transform proved in Proposition 6.1.

Finally, (vi) follows from the interpolation between the two cases $p = 2$ and $p = \infty$. The endpoint $p = \infty$ follows from (6.6); as for the endpoint $p = 2$, it can be proved by combining (i) with the fact that for any λ, x, z,

$$\text{for any } \lambda \in \mathbb{R}, \, x, z \text{ in } \mathbb{R}^2, \qquad \frac{e^{\sigma |z|^2}}{e^{\sigma |x+z|^2} e^{\sigma |\lambda x^\perp + z|^2} e^{\sigma |x + \lambda x^\perp + z|^2}} \lesssim 1.$$

$$\square$$

We now prove the identity that was alluded to earlier.

Lemma 7.3. *For some constant C, we have*

$$\mathcal{T} \left(\frac{1}{|x|}, \frac{1}{|x|}, \frac{1}{|x|} \right) = \frac{C}{|x|}.$$

Proof. Denote e_1 for the point $(1,0)$ in $\mathbb{R}^2$. By scale invariance, it suffices to show that

$$\mathcal{T} \left(\frac{1}{|x|}, \frac{1}{|x|}, \frac{1}{|x|} \right)(e_1) = \int_{-1}^{1} \int_{\mathbb{R}^2} \frac{1}{|x + e_1|} \frac{1}{|\lambda x^\perp + e_1|} \frac{1}{|x + \lambda x^\perp + e_1|} \, dx \, d\lambda < \infty.$$

We prove this bound by splitting the integration domain into several regions.

Where $|x| \ll 1$. The contribution of this set is easily seen to be bounded.

Where $|x| >> 1$ and $|\lambda| >> \frac{1}{|x|}$. The contribution of this set is bounded by

$$C \int_{|x|>>1} \int_{\frac{1}{|x|} \ll |\lambda| \ll 1} \frac{1}{|x|} \frac{1}{|\lambda x|} \frac{1}{|x|} \, d\lambda \, dx \lesssim \int_{|x|>>1} \frac{1}{|x|^3} \log |x| \, dx < \infty.$$

Where $|x| >> 1$ and $|\lambda| \ll \frac{1}{|x|}$. This contributes at most

$$C \int_{|x|>>1} \int_{|\lambda| \ll \frac{1}{|x|}} \frac{1}{|x|^2} d\lambda \, dx \lesssim \int_{|x|>>1} \frac{1}{|x|^3} \, dx < \infty.$$

Where $|x| >> 1$ and $|\lambda| \sim \frac{1}{|x|}$. This gives at most

$$C \int_{|x|>>1} \int_{\lambda \sim \frac{1}{|x|}} \frac{1}{|\lambda x^\perp + e_1|} \frac{1}{|x|^2} d\lambda \, dx = \int_{|\lambda| \ll 1} \lambda^2 \int_{|x| \sim \frac{1}{\lambda}} \frac{1}{|\lambda x^\perp + e_1|} \, dx \, d\lambda.$$

Changing variable $y = \lambda x^\perp + e_1$, this is less than

$$\int_{\lambda \ll 1} \int_{|y| \lesssim 1} \frac{1}{|y|} \, dy \, d\lambda < \infty.$$

Where $|x| \sim 1$, $|x + e_1| \gtrsim 1$, $|\lambda x^\perp + e_1| \gtrsim 1$, and $|x + \lambda x^\perp + e_1| \gtrsim 1$. This contribution is immediately seen to be finite.

Where $|x| \sim 1$ and $|\lambda x^\perp + e_1| \ll 1$. This region contributes less than

$$C \int_{|\lambda| \sim 1} \int_{|x| \sim 1, |\lambda x^\perp + e_1| \ll 1} \frac{1}{|\lambda x^\perp + e_1|} \, dx \, d\lambda \lesssim \int_{|\lambda| \sim 1} \int_{|y| \ll 1} \frac{1}{|y|} \, dy \, \frac{d\lambda}{\lambda^2} < \infty,$$

were we changed variables in the integral, setting $y = \lambda x^\perp + e_1$.

Where $|x| \sim 1$, and $|x + \lambda x^\perp + e_1| \ll 1$. By the relation

$$|e_1 + x|^2 + |e_1 + \lambda x^\perp|^2 = 1 + |e_1 + x + \lambda x^\perp|^2$$

We have that $both$ $|x + e_1|$ and $|\lambda x + e_1| \gtrsim 1$ (since $|x| \sim 1$), and therefore this contribution can be bounded by

$$C \int_{|\lambda| \sim 1} \int_{|x + \lambda x^\perp + e_1| \ll 1} \frac{1}{|x + \lambda x^\perp + e_1|} \, dx \, d\lambda < \infty$$

by arguments similar to the ones already employed.

We are thus left with the case when $|x| \sim 1$ and $|x + e_1| \ll 1$, which we split into three cases:

Where $|x + e_1| \ll 1$ and $|\lambda| > 4|x + e_1|$. The integral over this region can be bounded by

$$C \int_{|x+e_1| \ll 1} \int_{\substack{-1 < \lambda < 1 \\ |\lambda| > 4|x+e_1|}} \frac{1}{|x + e_1|} \frac{1}{|x + \lambda x^\perp + e_1|} \, d\lambda \, dx \lesssim \int_{|y| \ll 1} \int_{\substack{-1 < \lambda < 1 \\ |\lambda| > 4|y|}} \frac{1}{|y|} \frac{1}{|y + \lambda(y - e_1)^\perp|} \, d\lambda \, dy$$

upon changing variable $y = x + e_1$. Now notice that for fixed λ

$$|y + \lambda(y - e_1)^\perp| \geq \left| |y + \lambda y^\perp| - \lambda \right| = \left| \sqrt{1 + \lambda^2}|y| - |\lambda| \right|.$$

Therefore, the above can be bounded by

$$\cdots \lesssim \int_{|y| \ll 1} \int_{\substack{-1 < \lambda < 1 \\ |\lambda| > 4|y|}} \frac{1}{|y|} \frac{1}{\left| \sqrt{1 + \lambda^2}|y| - |\lambda| \right|} \, d\lambda \, dy \lesssim \int_{|y| \ll 1} \int_{\substack{-1 < \lambda < 1 \\ |\lambda| > 4|y|}} \frac{1}{|y|} \frac{1}{|\lambda|} \, d\lambda \, dy$$

$$\lesssim \int_{|y| \ll 1} \frac{|\log|y||}{|y|} \, dy < \infty.$$

Where $|x + e_1| \ll 1$ and $|\lambda| < \frac{1}{4}|x + e_1|$. Proceeding as in the previous case, we see that this contributes at most

$$C \int_{|y| \ll 1} \int_{|\lambda| < \frac{1}{4}|y|} \frac{1}{|y|} \frac{1}{\left| \sqrt{1 + \lambda^2}|y| - |\lambda| \right|} \, dy \, d\lambda \lesssim \int_{|y| \ll 1} \int_{|\lambda| < \frac{1}{4}|y|} \frac{1}{|y|^2} \, dy$$

$$\lesssim \int_{|y| \ll 1} \frac{dy}{|y|} < \infty.$$

Where $|x + e_1| \ll 1$ and $\frac{1}{4}|x + e_1| < |\lambda| < 4|x + e_1|$. For simplicity in the notations, we only treat the case $\lambda > 0$. This region contributes at most

$$\int_{|y| \ll 1} \int_{\frac{1}{4}|y| < \lambda < 4|y|} \frac{1}{|y|} \frac{1}{|y + \lambda(y - e_1)^\perp|} \, d\lambda \, dy$$

$$= 2\pi \int_{|r| \ll 1} \int_{\theta \in \mathbb{S}^1} \int_{\frac{1}{4} < \alpha < 4} \frac{1}{|\omega(\theta) + \alpha(r\omega(\theta) - e_1)^\perp|} \, d\alpha \, dr \, d\theta,$$

where we set $\lambda = \alpha r$, and parameterized y by $y = r\omega(\theta)$, with $\omega(\theta) = \begin{pmatrix} \cos\theta \\ \sin\theta \end{pmatrix}$, and $(r, \theta) \in \mathbb{R}^+ \times \mathbb{S}^1$. Now define

$$P(\alpha, r, \theta) = |\omega(\theta) + \alpha(r\omega(\theta) - e_1)^\perp|^2 = 1 + \alpha^2 r^2 + \alpha^2 - 2\alpha(\sin\theta + \alpha r \cos\theta).$$

The following properties are easily established:

- For fixed α and r, P is minimized at θ_0 such that $\sin\theta_0 = \frac{1}{\sqrt{1+\alpha^2 r^2}}$ and $\cos\theta_0 = \frac{\alpha r}{\sqrt{1+\alpha^2 r^2}}$. The minimal value of P is then $\left| \sqrt{1 + \alpha^2 r^2} - \alpha \right|^2$
- For $|\theta - \theta_0| < \frac{1}{10}$, $\partial_\theta^2 P(\alpha, r, \theta) \gtrsim \alpha$, thus $P(\lambda, r, \theta) \gtrsim \left| \sqrt{1 + \alpha^2 r^2} - \alpha \right|^2 + \alpha(\theta - \theta_0)^2$.
- For $|\theta - \theta_0| > \frac{1}{10}$, $P(\alpha, r, \theta) \gtrsim \alpha + \left| \sqrt{1 + \alpha^2 r^2} - \alpha \right|^2$.

The above can therefore be bounded by

$$\cdots \lesssim \int_{|r| \ll 1} \int_{\frac{1}{4} < \alpha < 4} \int_{|\theta - \theta_0| < \frac{1}{10}} \frac{1}{\sqrt{\left| \sqrt{1 + \alpha^2 r^2} - \alpha \right|^2 + \alpha(\theta - \theta_0)^2}} \, d\theta \, d\alpha \, dr$$

$$+ \int_{|r| \ll 1} \int_{\frac{1}{4} < \alpha < 4} \int_{|\theta - \theta_0| > \frac{1}{10}} \frac{1}{\sqrt{\alpha + (\sqrt{1 + \alpha^2 r^2} - \alpha)^2}} \, d\theta \, d\alpha \, dr$$

The second term on the above right-hand side is immediately seen to be bounded, thus we focus on the first one, which can be written as

$$\int_{|r|\ll 1}\int_{\frac{1}{4}<\alpha<4}\int_{|\theta|<\frac{1}{10}}\frac{1}{\sqrt{\left|\sqrt{1+\alpha^2 r^2}-\alpha\right|^2+\alpha\theta^2}}\,d\theta\,d\alpha\,dr$$
$$\lesssim \int_{|r|\ll 1}\int_{\frac{1}{4}<\alpha<4}\langle\log\left|\sqrt{1+\alpha^2 r^2}-\alpha\right|\rangle\,d\alpha\,dr,$$

(7.1)

where we used the following simple bound for $0\le u\le 10$:

$$\int_{|\theta|<1}\frac{1}{\sqrt{u+\theta^2}}\,d\theta=\int_{|\theta|\le u^{-1/2}}\frac{1}{\sqrt{1+\theta^2}}\,d\theta\lesssim\langle\log u\rangle.$$

Finally we note that the integral (7.1) is finite by the change of variable $\alpha\mapsto f(\alpha)=\alpha-\sqrt{1+\alpha^2 r^2}$ which satisfies $f'(\alpha)\sim 1$ in the considered range of α and r. $\qquad\square$

The final lemma in this subsection will be useful to prove the well-posedness of (CR) in the spaces X^σ that were used in the approximation results of Theorems 2.4 and 2.6.

Lemma 7.4. *The following tame bound holds: for any functions f, g and h, and $\sigma\ge 0$*

$$\|T(f,g,h)\|_{X^\sigma}\lesssim\|f\|_{X^\sigma}\|g\|_{L^2}\|h\|_{L^2}+\|f\|_{L^4}\|g\|_{L^{4/3}}\|h\|_{X^\sigma}$$

(7.2)

Proof. Without loss of generality, we assume that f,g, and h are non-negative. Using the orthogonality relation

$$|\xi|^2+|\xi+\lambda z+z^\perp|^2=|\xi+\lambda z|^2+|\xi+z^\perp|^2$$

we arrive at the conclusion that for any $\xi,z\in\mathbb{R}^2$ and $\lambda\in[-1,1]$,

$$\langle\xi\rangle^\sigma\lesssim_\sigma\langle\xi+\lambda z\rangle^\sigma+\langle\xi+z^\perp\rangle^\sigma.$$

As a result, we get that

$$\langle\xi\rangle^\sigma\mathcal{T}(f,g,h)\lesssim_\sigma\mathcal{T}(\langle\cdot\rangle^\sigma f,g,h)+\mathcal{T}(f,g,\langle\cdot\rangle^\sigma h)$$

and hence we arrive at

$$\|\mathcal{T}(f,g,h)\|_{X^\sigma}\lesssim_\sigma\|f\|_{X^\sigma}\sup_{\xi\in\mathbb{R}^2}\int_{-1}^1\int_{\mathbb{R}^2}g(\xi+\lambda z+z^\perp)h(\xi+z^\perp)dz\,d\lambda$$

$$+\|h\|_{X^\sigma}\sup_{\xi\in\mathbb{R}^2}\int_{-1}^1\int_{\mathbb{R}^2}f(\xi+\lambda z)g(\xi+\lambda z+z^\perp)dz\,d\lambda$$

$$\lesssim\|f\|_{X^\sigma}\|g\|_{L^2}\|h\|_{L^2}+\|h\|_{X^\sigma}\int_{-1}^1\underbrace{\|f(\xi+\lambda z)\|_{L_z^4}}_{=|\lambda|^{-1/2}\|f\|_{L^4}}\underbrace{\|g(\xi+\lambda z+z^\perp)\|_{L_z^{4/3}}}_{\lesssim\|g\|_{L^{4/3}}}d\lambda$$

$$\lesssim\|f\|_{X^\sigma}\|g\|_{L^2}\|h\|_{L^2}+\|h\|_{X^\sigma}\|f\|_{L^4}\|g\|_{L^{4/3}},$$

from which (7.2) follows. $\qquad\square$

7.2. **Local and global well-posedness for (CR).** We consider the Cauchy problem for (CR):

Theorem 7.5.   (i) *Local well posedness: For X being any of the spaces given in Proposition 7.1, the Cauchy problem* (CR) *is locally well-posed in X. That is, for any f_0 in X, there exists a time $T > 0$, and a solution in $\mathcal{C}^\infty([0,T], X)$, which is unique in $L^\infty([0,T], X)$, and depends continuously on f_0 in this topology.*
  (ii) *Global well-posedness: if $f_0 \in L^2$, the local solution can be prolonged into a global one. More precisely: there exists a unique solution in $\mathcal{C}^\infty([0,\infty), L^2)$. For any T, it is unique in $L^\infty([0,T], L^2)$, and depends continuously on f_0 in this topology.*
    Furthermore, the mass, kinetic energy, and first moment of f remain constant in time if they were finite at time 0.
  (iii) *Propagation of regularity: assume $f_0 \in L^2$, and let f be the solution given in (ii). If in addition $f_0 \in H^\sigma$ (respectively $L^{2,\sigma}$) for $\sigma \geq 0$ then $f \in \mathcal{C}^\infty([0,\infty), H^\sigma)$ (respectively $\mathcal{C}^\infty([0,\infty), L^{2,\sigma})$.*
  (iv) *The Cauchy problem* (CR) *is locally well-posed in the space X^σ for any $\sigma > 2$. If $g_0 \in H^1$, the solution can be extended globally in time in X^σ. Moreover, if $\nabla g(0) \in X^\sigma$, then $g, \nabla g \in C_{\mathrm{loc}}(\mathbb{R}; X^\sigma)$.*

Remark 7.6. *The restriction $\sigma > 2$ for local well-posedness in item (iv) could easily be improved to $\sigma \geq 1$, but global well-posedness beyond $\sigma > 2$ seems very difficult to prove.*

Proof. (i) Write (CR) as a fixed point problem:

$$f(t) = f_0 + i \int_0^t \mathcal{T}(f, f, f)(s)\, ds.$$

Proposition 7.1 immediately gives the a priori estimates

$$\begin{cases} \|f\|_{L^\infty([0,T],X)} \lesssim \|f_0\|_X + T\|f\|^3_{L^\infty([0,T],X)} \\ \|f - g\|_{L^\infty([0,T],X)} \lesssim \|f_0 - g_0\|_X + T\|f - g\|_{L^\infty([0,T],X)} \left(\|f\|^2_{L^\infty([0,T],X)} + \|g\|^2_{L^\infty([0,T],X)} \right), \end{cases}$$

for f and g two solutions of (CR), from which one deduces the local well-posedness statement by a standard fixed point argument. Notice that T can be chosen depending only on $\|f_0\|_X$ (namely $\sim \|f_0\|_X^{-2}$).

(ii) Combining the above local well-posedness with the conservation laws (Corollary 6.2) gives the global well-posedness (recall that the local well-posedness time T only depends on $\|u_0\|_2$).

(iii) is classical, so we only say a few words. Using the boundedness of $\mathcal{T}$ from $L^2 \times L^2 \times L^2 \to L^2$, one notices that the equation is locally well-posed in H^s with the time of existence depending on the conserved quantity $\|g(t)\|_{L^2}$. Consequently, one can iterate the local existence statement indefinitely to get global existence. The result for weighted spaces $L^{2,\sigma}$ follows from the invariance under Fourier transform.

(iv) Local well-posedness in X^σ is a direct consequence of Picard iteration and Lemma 7.4. This gives a time of existence bounded below by $C\|g_0\|_{X^\sigma}^{-2}$. Global existence follows from the same lemma, which gives due to the embeddings $H^1 \hookrightarrow L^4$ and $L^{2,1} \hookrightarrow L^{4/3}$ and the conservation of the H^1 and $L^{2,1}$ norms that

$$\|g(t)\|_{X^\sigma} \leq \|g_0\|_{X^\sigma} + \|g_0\|_{H^1} \|\langle x \rangle g_0\|_{L^2} \int_0^t \|g(s)\|_{X^\sigma} ds.$$

This yields an a priori bound on $\|g(t)\|_{X^\sigma}$ that allows extending the solution globally. Finally, if $\nabla g_0 \in X^\sigma$, then proceeding as above and using that ∇g satisfies

$$\nabla g(t, \xi) = \nabla g_0(\xi) + i \int_0^t \left(\mathcal{T}(\nabla g, g, g) + \mathcal{T}(g, \nabla g, g) + \mathcal{T}(g, g, \nabla g) \right) ds, \qquad (7.3)$$

one can obtain local-in-time existence of $\nabla g \in C([0, T]; X^\sigma)$. To extend this to a global statement, one has to show that $\|\nabla g\|_{X^\sigma}$ cannot blow up. This follows once we obtain an a priori bound for $\|\nabla g\|_{X^\sigma}$ that prohibits its blowup. Using Lemma 7.4 and (7.3), we have

$$\|\nabla g(t)\|_{X^\sigma} \lesssim \|\nabla g_0\|_{X^\sigma} + \int_0^t \Big(\|\nabla g(s)\|_{X^\sigma} \|g(s)\|_{L^2}^2 + \|g(s)\|_{X^\sigma} \|\nabla g\|_{L^4} \|g(s)\|_{L^{4/3}}$$

$$+ \|g(s)\|_{X^\sigma} \|\nabla g\|_{L^2} \|g\|_{L^2} + \|g(s)\|_{X^\sigma} \|\nabla g(s)\|_{L^{4/3}} \|g\|_{L^4} + \|\nabla g(s)\|_{X^\sigma} \|g\|_{L^4} \|g\|_{L^{4/3}} \Big) ds$$

$$\lesssim \|\nabla g_0\|_{X^\sigma} + \int_0^t \|\nabla g\|_{X^\sigma} \|g(s)\|_{X^\sigma}^2 \, ds$$

Since the solution g is bounded in X^σ on any compact time interval of $\mathbb{R}$, we get the needed a priori bound on $\|\nabla g\|_{X^\sigma}$ to finish the proof. $\qquad \square$

7.3. Weak continuity of $\mathcal{T}$ on L^2. Recall that

$$\mathcal{T}(f, g, h)(\xi) = \int_{-1}^1 \int_{\mathbb{R}^2} f(\xi + x) \overline{g(\xi + x + \lambda x^\perp)} h(\xi + \lambda x^\perp) \, dx \, d\lambda.$$

We introduce the following notation: $\mathcal{T}_{\substack{x \in I \\ \lambda \in J}}(f, g, h)$ is defined as $\mathcal{T}$ above, but with the integration domain changed from $(x, \lambda) \in \mathbb{R}^2 \times [-1, 1]$ to $(x, \lambda) \in I \times (J \cap [-1, 1])$. For instance:

$$\mathcal{T}_{\substack{|x| \sim 1 \\ |\lambda| < 1/4}}(f, g, h)(\xi) = \int_{|\lambda| < 1/4} \int_{|x| \sim 1} f(\xi + x) \overline{g(\xi + x + \lambda x^\perp)} h(\xi + \lambda x^\perp) \, dx \, d\lambda.$$

We start by proving more precise estimates on such localized versions of $\mathcal{T}$. Space and frequency localization operators will be needed; we start by defining them. Pick ψ a smooth function equal to 1 on $B(0, 1)$, 0 on $B(0, 2)^c$, and set

$$Q_{<M} f \stackrel{def}{=} \psi\left(\frac{x}{M}\right) f \quad , \quad Q_{>M} \stackrel{def}{=} f - P_{<M} f$$

$$P_{<M} \stackrel{def}{=} \mathcal{F}^{-1} \psi\left(\frac{\xi}{M}\right) \widehat{f}(\xi) \quad \text{and} \quad P_{>M} \stackrel{def}{=} f - Q_{<M} f$$

Proposition 7.7.   (i) *If f, g, h belong to L^2,*

$$\big\| \mathcal{T}_{|\lambda| < \epsilon}(f, g, h) \big\|_{L^2(\mathbb{R}^2)} \lesssim \epsilon \|f\|_{L^2(\mathbb{R}^2)} \|g\|_{L^2(\mathbb{R}^2)} \|h\|_{L^2(\mathbb{R}^2)}$$

(ii) *For $\epsilon > 0$, there exists a constant $C(\epsilon)$ such that if f, g, h belong to L^2,*

$$\left\| \mathcal{T}_{\substack{|x| \geq M \\ |\lambda| > \epsilon}}(f, g, h) \right\|_{L^\infty(\mathbb{R}^2)} \leq C(\epsilon) M^{-1} \|f\|_{L^2(\mathbb{R}^2)} \|g\|_{L^2(\mathbb{R}^2)} \|h\|_{L^2(\mathbb{R}^2)}.$$

(iii) *For $\epsilon > 0$, $R \ll M$, there exists a constant $C(\epsilon)$ such that if f, g, h belong to L^2,*

$$\big\| \mathcal{T}_{|\lambda| > \epsilon}(Q_{>M} f, g, h) \big\|_{L^\infty(B(0,R))} + \big\| \mathcal{T}_{|\lambda| > \epsilon}(f, Q_{>M} g, h) \big\|_{L^\infty(B(0,R))} \leq C(\epsilon) M^{-1} \|f\|_{L^2(\mathbb{R}^2)} \|g\|_{L^2(\mathbb{R}^2)} \|h\|_{L^2(\mathbb{R}^2)}$$

Proof. Proof of (i). The proof of (i) follows exactly that of (i) in Proposition 7.1, we do not repeat it here.

Proof of (ii). Observe that $\mathcal{T}_{\substack{|x|\sim 1 \\ |\lambda|>\epsilon}}(f,g,h)(0)$ can be written

$$\mathcal{T}_{\substack{|x|\sim 1 \\ |\lambda|>\epsilon}}(f,g,h)(0) = \int_{\substack{|x|\sim 1 \\ 1>|\lambda|>\epsilon}} f(\pi_1(x,\lambda))\overline{g}(\pi_2(x,\lambda))h(\pi_3(x,\lambda))\,dx\,d\lambda$$

with $\pi_1(x,\lambda) = x$, $\pi_2(x,\lambda) = x + \lambda x^\perp$, and $\pi_3(x,\lambda) = \lambda x^\perp$. This corresponds to the type of operators analyzed in Bennett, Carbery and Wright [3]. The non-degeneracy condition in Theorem 2 of this article can be checked, implying the nonlinear Loomis-Whitney inequality

$$\left| \mathcal{T}_{\substack{|x|\sim 1 \\ |\lambda|>\epsilon}}(f,g,h)(0) \right| \le C(\epsilon)\|f\|_{L^2}\|g\|_{L^2}\|h\|_{L^2}.$$

Since $\mathcal{T}_{\substack{|x|\sim 1 \\ |\lambda|>\epsilon}}(f,g,h)(z) = \mathcal{T}_{\substack{|x|\sim 1 \\ |\lambda|>\epsilon}}(f(z+\cdot),g(z+\cdot),h(z+\cdot))(0)$, we deduce that

$$\left\| \mathcal{T}_{\substack{|x|\sim 1 \\ |\lambda|>\epsilon}}(f,g,h) \right\|_{L^\infty} \le C(\epsilon)\|f\|_{L^2}\|g\|_{L^2}\|h\|_{L^2}.$$

By scaling,

$$\left\| \mathcal{T}_{\substack{|x|\sim 2^k \\ |\lambda|>\epsilon}}(f,g,h) \right\|_{L^\infty} \le C(\epsilon)2^{-k}\|f\|_{L^2}\|g\|_{L^2}\|h\|_{L^2}.$$

This gives the desired result since

$$\left\| \mathcal{T}_{\substack{|x|>M \\ |\lambda|>\epsilon}}(f,g,h) \right\|_{L^\infty} \le \sum_{2^k \gtrsim M} \left\| \mathcal{T}_{\substack{|x|\sim 2^k \\ |\lambda|>\epsilon}}(f,g,h) \right\|_{L^\infty} \le C(\epsilon)M^{-1}\|f\|_{L^2}\|g\|_{L^2}\|h\|_{L^2}.$$

Proof of (iii). We only prove that the first term on the left-hand side satisfies the inequality; the second being similar. Thus we wish to estimate

$$\mathcal{T}_{|\lambda|>\epsilon}(Q_{>M}f,g,h)(z) = \int_{1>|\lambda|>\epsilon}\int [Q_{>M}f]\,(x+z)\overline{g(x+\lambda x^\perp + z)}h(\lambda x^\perp + z)\,dx\,d\lambda,$$

where $|z| < R$. Since $Q_{>M}f$ is supported outside of a ball of radius $\sim M$ and center 0, we see that only x such that $|x+z| \gtrsim M$ contribute to the above integral. Since $M \gg R$, this implies $|x| \gtrsim M$. But then the estimate follows from (ii). $\square$

These estimates lead to the following theorem.

Theorem 7.8. *Suppose that (f^n), (g^n) and (h^n) are three sequences such that*

$$f^n \rightharpoonup f \qquad g^n \rightharpoonup g \qquad h^n \rightharpoonup h \qquad \text{in } L^2.$$

(we denote as is customary $f^n \to f$ for the strong convergence in L^2, and $f^n \rightharpoonup f$ for the weak convergence in L^2). Then

$$\mathcal{T}(f^n,g^n,h^n) \rightharpoonup \mathcal{T}(f,g,h) \qquad \text{in } L^2.$$

Remark 7.9. *It is tempting to mention here the div-curl lemma* [48] [60]*: if (u^n), (v^n) are sequences of functions from $\mathbb{R}^3$ to $\mathbb{R}^3$, in L^2, which are respectively divergence and curl free, and converge weakly to u and v, then their scalar product $u^n \cdot v^n$ converges to $u \cdot v$ in the sense of distribution. It was later realized by Coifman, Lions, Meyer, Semmes* [15] *that this particular structure also has implications in harmonic analysis. Namely, if u, v belong in L^2, and are respectively divergence and curl free, then their scalar product $u \cdot v$ not only belongs to L^1, but even to the Hardy space $\mathcal{H}^1$.*

One might wonder whether a similar phenomenon occurs for $\mathcal{T}$: does this operator map $L^2 \times L^2 \times L^2$ to a slightly smaller space than L^2?

Proof. First fix sequences f^n, g^n and h^n converging weakly to f, g, h in L^2. By the uniform boundedness principle, the functions f^n, g^n, h^n enjoy uniform bounds in L^2. Our aim will be to show that, for ϕ in the Schwartz class $\mathcal{S}$,

$$\langle \mathcal{T}(f^n, g^n, h^n), \phi \rangle \to \langle \mathcal{T}(f, g, h), \phi \rangle.$$

This will give the desired result by boundedness of $\mathcal{T}$ on L^2, and density of $\mathcal{S}$ in L^2. Thus we fix from now on $\phi \in \mathcal{S}$; all implicit constants are allows to depend on ϕ.

We will be using three parameters, ϵ, M and R, whose precise value will be fixed at the end of the proof, but which satisfy $\epsilon \ll 1$, $M \gg R \gg 1$.

Step 1: localization in space. Writing $f = Q_{<M} f^n + Q_{>M} f^n$, and similarly for g and h, we obtain that $\mathcal{T}(f, g, h)(x)$ can be written as the sum

$$\langle \mathcal{T}(f^n, g^n, h^n), \phi \rangle = \langle \mathcal{T}(Q_{>M} f^n, g^n, h^n), \phi \rangle \tag{7.4a}$$
$$+ \langle \mathcal{T}(Q_{<M} f^n, g^n, Q_{>M} h^n), \phi \rangle \tag{7.4b}$$
$$+ \langle \mathcal{T}(Q_{<M} f^n, Q_{>10M} g^n, Q_{<M} h^n), \phi \rangle \tag{7.4c}$$
$$+ \langle \mathcal{T}(Q_{<M} f^n, Q_{<10M} g^n, Q_{<M} h^n), \phi \rangle. \tag{7.4d}$$

To estimate (7.4a), we first split this term, and then use Proposition 7.7 as well as the fast decay of ϕ:

$$|(7.4a)| \leq \left| \langle \mathcal{T}_{|\lambda| < \epsilon}(Q_{>M} f^n, g^n, h^n), \phi \rangle \right| + \left| \langle \mathcal{T}_{|\lambda| > \epsilon}(Q_{>M} f^n, g^n, h^n), Q_{<R} \phi \rangle \right|$$
$$+ \left| \langle \mathcal{T}_{|\lambda| > \epsilon}(Q_{>M} f^n, g^n, h^n), Q_{>R} \phi \rangle \right|$$
$$\lesssim \epsilon \|f^n\|_2 \|g^n\|_2 \|h^n\|_2 \|\phi\|_2 + C(\epsilon) M^{-1} \|f^n\|_2 \|g^n\|_2 \|h^n\|_2 \|\phi\|_1 + \|f^n\|_2 \|g^n\|_2 \|h^n\|_2 R^{-10}$$
$$\lesssim \epsilon + C(\epsilon) M^{-1} + R^{-10}.$$
$$\tag{7.5}$$

The next term, (7.4b), can be treated identically to give

$$|(7.4b)| \lesssim \epsilon + C(\epsilon) M^{-1} + R^{-10}. \tag{7.6}$$

Finally, taking advantage of the localization properties of $\mathcal{T}$, (7.4c) can be bounded by

$$|(7.4c)| = |\langle \mathcal{T}(Q_{<M} f^n, Q_{>10M} g^n, Q_{<M} h^n), Q_{>R} \phi \rangle| \lesssim \|f^n\|_2 \|g^n\|_2 \|h^n\|_2 R^{-10} \lesssim R^{-10}.$$

We are left with (7.4d) which we will decompose further, in frequency this time.

Step 2: localization in frequency. By Plancherel's theorem and Lemma 6.5,

$$4\pi^2 (7.4d) = \langle \mathcal{F}\mathcal{T}(Q_{<M} f^n, Q_{<10M} g^n, Q_{<M} h^n), \widehat{\phi} \rangle = \langle \mathcal{T}(P_{<M} \widehat{f^n}, P_{<10M} \widehat{g^n}, P_{<M} \widehat{h^n}), \widehat{\phi} \rangle.$$

Performing the same decomposition as in Step 1, this is equal to

$$(7.4d) = \langle \mathcal{T}(Q_{>M}P_{<M}\widehat{f^n}, P_{<10M}\widehat{g^n}, P_{<M}\widehat{h^n}), \widehat{\phi} \rangle \tag{7.7a}$$

$$+ \langle \mathcal{T}(Q_{<M}P_{<M}\widehat{f^n}, P_{<10M}\widehat{g^n}, Q_{>M}P_{<M}\widehat{h^n}), \widehat{\phi} \rangle \tag{7.7b}$$

$$+ \langle \mathcal{T}(Q_{<M}P_{<M}\widehat{f^n}, Q_{>10M}P_{<10M}\widehat{g^n}, Q_{<M}P_{<M}\widehat{h^n}), \widehat{\phi} \rangle \tag{7.7c}$$

$$+ \langle \mathcal{T}(Q_{<M}P_{<M}\widehat{f^n}, Q_{<10M}P_{<10M}\widehat{g^n}, Q_{<M}P_{<M}\widehat{h^n}), \widehat{\phi} \rangle, \tag{7.7d}$$

and just like in Step 1 we can bound

$$|(7.7a)| + |(7.7b)| + |(7.7c)| \lesssim \epsilon + C(\epsilon)M^{-1} + R^{-10}.$$

Step 3: conclusion. Fix $\eta > 0$, we want to show that, for n big enough,

$$|\langle \mathcal{T}(f^n, g^n, h^n), \phi \rangle - \langle \mathcal{T}(f, g, h), \phi \rangle| < \eta.$$

First split the above as

$$4\pi^2 |\langle \mathcal{T}(f^n, g^n, h^n), \phi \rangle - \langle \mathcal{T}(f, g, h), \phi \rangle| \tag{7.8a}$$

$$= \left| \langle \mathcal{T}(\widehat{f^n}, \widehat{g^n}, \widehat{h^n}), \widehat{\phi} \rangle - \langle \mathcal{T}(\widehat{f}, \widehat{g}, \widehat{h}), \widehat{\phi} \rangle \right| \tag{7.8b}$$

$$\leq \left| \langle \mathcal{T}(\widehat{f^n}, \widehat{g^n}, \widehat{h^n}), \widehat{\phi} \rangle - \langle \mathcal{T}(Q_{<M}P_{<M}\widehat{f^n}, Q_{<10M}P_{<10M}\widehat{g^n}, Q_{<M}P_{<M}\widehat{h^n}), \widehat{\phi} \rangle \right| \tag{7.8c}$$

$$+ \left| \langle \mathcal{T}(Q_{<M}P_{<M}\widehat{f^n}, Q_{<10M}P_{<10M}\widehat{g^n}, Q_{<M}P_{<M}\widehat{h^n}), \widehat{\phi} \rangle \right.$$
$$\left. - \mathcal{T}(Q_{<M}P_{<M}\widehat{f}, Q_{<10M}P_{<10M}\widehat{g}, Q_{<M}P_{<M}\widehat{h}), \widehat{\phi} \rangle \right| \tag{7.8d}$$

$$+ \left| \langle \mathcal{T}(Q_{<M}P_{<M}\widehat{f}, Q_{<10M}P_{<10M}\widehat{g}, Q_{<M}P_{<M}\widehat{h}), \widehat{\phi} \rangle - \langle \mathcal{T}(\widehat{f}, \widehat{g}, \widehat{h}), \widehat{\phi} \rangle \right| \tag{7.8e}$$

Gathering the estimates of Step 1 and Step 2, we see that

$$|(7.8c)| + |(7.8e)| \lesssim \epsilon + C(\epsilon)M^{-1} + R^{-10}.$$

The weak convergence of f^n, g^n and h^n to, respectively, f, g and h in L^2 implies that, for M fixed

$$Q_{<M}P_{<M}\widehat{f^n} \xrightarrow{L^2} Q_{<M}P_{<M}\widehat{f} \;,\;\; Q_{<10M}P_{<10M}\widehat{g^n} \xrightarrow{L^2} Q_{<10M}P_{<10M}\widehat{g} \;,\;\; Q_{<M}P_{<M}\widehat{h^n} \xrightarrow{L^2} Q_{<M}P_{<M}\widehat{h}$$

as $n \to \infty$. Since $\mathcal{T}$ is bounded on L^2, we deduce that

$$|(7.8d)| \xrightarrow{n\to\infty} 0.$$

To conclude, it suffices to fix ϵ, M, R such that $|(7.8c)| + |(7.8e)| < \frac{1}{2}\eta$. Then

$$\mathrm{limsup}_{n\to\infty} |\langle \mathcal{T}(f^n, g^n, h^n), \phi \rangle - \langle \mathcal{T}(f, g, h), \phi \rangle| < \frac{1}{2}\eta.$$

$$\square$$

7.4. **Weak compactness of solutions in $L_t^\infty L_x^2$.** The preceding theorem on weak continuity of $\mathcal{T}$ on L^2 implies easily the following compactness result.

Corollary 7.10. *Assume (f^n) is a sequence of solutions of* (CR) *in $L^\infty([0,T], L^2)$. Then there exists a subsequence, which we still denote (f^n), and a solution $f \in L^\infty([0,T], L^2)$ of* (CR) *such that*

$$f^n \overset{n\to\infty}{\longrightarrow} f \qquad \text{weak-* in } L^\infty([0,T], L^2).$$

Proof. First notice that (f^n) is uniformly bounded in $\mathrm{Lip}([0,T], L^2)$. This means that we may define $(f^n(t))$ for all $t \in [0,T]$ and for each such t, there exists a weak limit $f(t) \in L_t^\infty([0,T]; L^2)$ such $f^n(t) \rightharpoonup f(t)$ weakly in $L^2(\mathbb{R}^2)$. One can easily verify that $f^n(t) \rightharpoonup f(t)$ weak-* in $L_t^\infty([0,T]; L^2)$ and that $f(t) \in \mathrm{Lip}([0,T], L^2)$.

It suffices to show now that f solves (*). Set

$$f_j^n(t) = f^n\left(\frac{k}{2^j}\right) \quad \text{and} \quad f_j(t) = f\left(\frac{k}{2^j}\right) \qquad \text{where } k \text{ minimizes } \left|t - \frac{k}{2^j}\right|.$$

We start from

$$\partial_t f_n = \mathcal{T}(f_n, f_n, f_n)$$

and want to pass to the limit $n \to \infty$ in the sense of distributions. For the linear part, it is automatic: $\partial_t f_n \to \partial_t f$. Now take ϕ in the Schwartz class $\mathcal{S}([0,T] \times \mathbb{R}^2)$. Then

$$
\begin{aligned}
\left|\langle \mathcal{T}(f^n, f^n, f^n), \phi \rangle - \langle \mathcal{T}(f,f,f), \phi \rangle\right| &\leq \left|\langle \mathcal{T}(f^n, f^n, f^n), \phi \rangle - \langle \mathcal{T}(f_j^n, f_j^n, f_j^n), \phi \rangle\right| \\
&+ \left|\langle \mathcal{T}(f_j^n, f_j^n, f_j^n), \phi \rangle - \langle \mathcal{T}(f_j, f_j, f_j), \phi \rangle\right| \\
&+ \left|\langle \mathcal{T}(f_j, f_j, f_j), \phi \rangle - \langle \mathcal{T}(f,f,f), \phi \rangle\right| \\
&\leq \left|\langle \mathcal{T}(f_j^n, f_j^n, f_j^n), \phi \rangle - \langle \mathcal{T}(f_j, f_j, f_j), \phi \rangle\right| + O(2^{-j}),
\end{aligned}
$$

where we used in the last line the uniform bound on (f^n) and f in $\mathrm{Lip}([0,T]; L^2)$. From this last inequality, by weak convergence of the f_j^n to f_j in L^2, and the weak continuity of $\mathcal{T}$ in L^2, we deduce that

$$\mathrm{limsup}_{n\to\infty} \left|\langle \mathcal{T}(f^n, f^n, f^n), \phi \rangle - \langle \mathcal{T}(f,f,f), \phi \rangle\right| \lesssim 2^{-j}$$

for any j; but this gives the desired result. $\qquad\square$

8. Variational properties

We examine here more precisely how the Hamiltonian $\mathcal{H}$ and the L^2 mass are related. Our main result is that Gaussians maximize the Hamiltonian for prescribed L^2 norm. This implies dynamical stability of the Gaussians.

The results in this section would follow immediately from the relation $\mathcal{H}(f) = \|e^{it\Delta}f\|_4^4$, and known results on the L^4 Strichartz norm, in particular [23, 34, 2]. Still, we chose to ignore this relation and give full proofs of Proposition 8.1 and Theorem 8.2; they turn out to be very short and provide a slightly different point of view.

8.1. **Positivity of the Hamiltonian.**

Proposition 8.1. *The Hamiltonian can be written*

$$\mathcal{H}(f) = \frac{1}{16} \int_{\mathbb{S}^1} \int_{\mathbb{R}} \int_{\mathbb{R}} \left| \int_{\mathbb{R}} f(u\omega + s\omega^\perp) \overline{f(t\omega + s\omega^\perp)} \, ds \right|^2 du \, dt \, d\omega. \tag{8.1}$$

In particular, it is non-negative, and zero only if $f = 0$.

Proof. The proof of the formula only consists of a change of integration variables in the formula (6.3). Parametrizing in this formula z by $v\omega$, with $(v, \omega) \in \mathbb{R} \times \mathbb{S}^1$ (at the expense of an extra factor of 2 in the demominator), and ξ by $t\omega + s\omega^\perp$, with $(s, t) \in \mathbb{R}^2$, it becomes

$$\mathcal{H}(f) = \frac{1}{16} \int_{\mathbb{S}^1} \int_{\mathbb{R}} \int_{\mathbb{R}} \int_{\mathbb{R}} \int_{\mathbb{R}} f\left((t+v)\omega + s\omega^\perp\right) f\left(t\omega + (s+\lambda v)\omega^\perp\right)$$
$$\overline{f\left((t+v)\omega + (s+\lambda v)\omega^\perp\right)} \, \overline{f\left(t\omega + s\omega^\perp\right)} |v| \, d\lambda \, ds \, dt \, dv \, d\omega.$$

Setting now $\lambda' = s + \lambda v$, $u = t + v$ gives

$$\mathcal{H}(f) = \frac{1}{16} \int_{\mathbb{S}^1} \int_{\mathbb{R}} \int_{\mathbb{R}} \int_{\mathbb{R}} \int_{\mathbb{R}} f\left(u\omega + s\omega^\perp\right) f\left(t\omega + \lambda'\omega^\perp\right)$$
$$\overline{f\left(u\omega + \lambda'\omega^\perp\right)} \, \overline{f\left(t\omega + s\omega^\perp\right)} \, ds \, d\lambda' \, dt \, du \, d\omega,$$

which is the desired result. It is then a small exercise to see that $\mathcal{H}$ vanishes only for the zero function. $\qquad\square$

8.2. Maxima of the Hamiltonian.
By Proposition 7.1, $\mathcal{H}$ is bounded on L^2. In other words,

$$0 \leq \mathcal{H}(f) \lesssim \|f\|_2^4.$$

Our next aim here is to understand for which functions the above inequality is saturated, and what is the best constant.

Theorem 8.2. *For fixed mass $\|f\|_2$, the Hamiltonian $\mathcal{H}(f)$ is maximized if f equals, up to the symmetries of the equation, $\|f\|_2 G$, where G is the L^2-normalized Gaussian $G(x) = \frac{1}{\sqrt{\pi}} e^{-\frac{|x|^2}{2}}$. Furthermore, all the maximizers are of the form $\|f\|_2 G$, up to the symmetries of the equation. Finally, the optimal constant in the bound for $\mathcal{H}$ is given by*

$$\mathcal{H}(f) \leq \frac{\pi}{8} \|f\|_2^4. \tag{8.2}$$

Remark 8.3. *As noticed earlier, the operator $\mathcal{T}$, or the Hamiltonian $\mathcal{H}$, correspond to a nonlinear generalization of the class studied by Brascamp and Lieb. Whether Brascamp-Lieb inequalities are saturated by Gaussians has been studied by many authors, we mention in particular [43, 11, 4]. Our proof relies on the heat flow, following an approach initiated in [11, 4].*

Proof. Step 1: if $g \in L^1$, convergence in L^1 of $\sqrt{t \left[e^{\frac{t}{4}\Delta} g\right](\sqrt{t}x)}$ to $\left(\int g\right)^{1/2} G(x)$. Recalling that the kernel of $e^{t\Delta}$ is $\frac{e^{-\frac{|x|^2}{4t}}}{4\pi t}$, it is easy to see that this statement is true if g is in addition in $\mathcal{C}_0^\infty$. The general case $g \in L^1$ follows by a density argument.

Step 2: if $f \in L^2$, $f \geq 0$, $\mathcal{H}\left(\sqrt{e^{\tau\Delta} f}\right)$ is increasing with τ. Notice that for $f \geq 0$, $\sqrt{e^{\tau\Delta} f}$ has a constant L^2-norm. It suffices to treat the case where f is smooth with rapid decay, the general case follows by an approximation argument using the boundedness of $\mathcal{H}$ on L^2. Furthermore, since $e^{\tau\Delta} f(x) > 0$ for $\tau > 0$, we assume from now on that f is positive, smooth,

and decays rapidly, which justifies all the manipulations which follow. Recall the formula (8.1) giving $\mathcal{H}(f)$. It can be written as

$$\mathcal{H}(f) = \int_{\mathbb{S}^1} \mathcal{H}_\omega(f) d\omega \quad \text{with} \quad \mathcal{H}_\omega(f) = \int_{\mathbb{R}} \int_{\mathbb{R}} \left| \int_s f(u\omega + s\omega^\perp) \overline{f(t\omega + s\omega^\perp)} \, ds \right|^2 du \, dt.$$

We will show that for any ω, $\frac{d}{d\tau}\mathcal{H}_\omega(\sqrt{e^{\tau\Delta}f}) \geq 0$. Due to the semi-group property of $e^{\tau\Delta}$, it suffices to prove it for $\tau = 0$, which we proceed to do. For convenience, we adopt the Cartesian coordinates given by ω, and denote simply $f(u,s) = f(u\omega + s\omega^\perp)$. Then

$$\frac{d}{d\tau}\mathcal{H}_\omega\left(\sqrt{e^{\tau\Delta}f}\right)(\tau = 0)$$

$$= 2 \int \int \int \int \int \sqrt{f}(u,\sigma)\sqrt{f}(t,\sigma)\sqrt{f}(t,s) \left[\partial_s^2 + \partial_u^2\right] f(u,s) \frac{1}{\sqrt{f}(u,s)} \, ds \, d\sigma \, du \, dt$$

$$\overset{def}{=} I + II.$$

Since I and II are symmetrical, it suffices to show that $II \geq 0$. But integrating by parts once in u, it can be written

$$II = \int \int \int \int \left[-\frac{\sqrt{f}(t,\sigma)\sqrt{f}(t,s)}{\sqrt{f}(u,\sigma)\sqrt{f}(u,s)} \partial_u f(u,\sigma) \partial_u f(u,s) \right.$$

$$\left. + \frac{\sqrt{f}(t,\sigma)\sqrt{f}(t,s)\sqrt{f}(u,\sigma)}{f(u,s)\sqrt{f}(u,s)} |\partial_u f(u,s)|^2 \right] \, du \, dt \, d\sigma \, ds$$

$$= \int \int \int \int \frac{1}{2} \left(\partial_u f(u,s) \frac{f^{1/2}(u,\sigma)}{f^{1/2}(u,s)} - \partial_u f(u,\sigma) \frac{f^{1/2}(u,s)}{f^{1/2}(u,\sigma)} \right)^2 \frac{\sqrt{f}(t,\sigma)\sqrt{f}(t,s)}{\sqrt{f}(u,\sigma)\sqrt{f}(u,s)} \, du \, dt \, d\sigma \, ds$$

$$\geq 0,$$

(8.3)

which concludes the proof.

Step 3: Gaussians are maximizers Let f in L^2. Then

$$\mathcal{H}(f) \leq \mathcal{H}(|f|)$$

$$\leq \mathcal{H}\left(\sqrt{e^{\frac{t}{4}\Delta}|f|^2}\right) \qquad \text{for } t \geq 0, \text{ by Step 2}$$

$$= \mathcal{H}\left(\sqrt{t\left[e^{\frac{t}{4}\Delta}|f|^2\right](\sqrt{t}\,\cdot)}\right) \qquad \text{by scaling invariance (Proposition 6.1)}$$

$$\overset{t\to\infty}{\longrightarrow} \|f\|_2^4 \, \mathcal{H}(G) \qquad \text{by Step 1 and continuity of } \mathcal{H} \text{ on } L^2 \text{ (Proposition 7.1)},$$

(8.4)

which gives the desired inequality.

Step 4: up to the symmetries, Gaussians are the unique non-negative maximizers. Assume that g a maximizer with $g \geq 0$. We will show that g is a Gaussian, up to the symmetries of the equation. Applying the heat flow, we might assume without loss of generality that g is positive and smooth. Then, if it is a maximizer, $\frac{d}{dt}\mathcal{H}(\sqrt{e^{t\Delta}g}) = 0$, which, by (8.3), implies (with

the notations of Step 2) that

$$\frac{\partial_u g(u, s)}{g(u, s)} = \frac{\partial_u g(u, \sigma)}{g(u, \sigma)} \quad \text{for any } u, s, \sigma.$$

Setting $h(u, s) = \log g(u, s)$, this becomes

$$\partial_u h(u, s) = \partial_u h(u, \sigma), \quad \text{for any } u, s, \sigma.$$

We can rewrite the above as $\partial_u h(u, s) = \partial_u h(u, s + s_0)$ for any u, s, s_0. Taking the Fourier transform of this identity in u, s and denoting ξ, η for their respective Fourier variables, we obtain

$$\xi \widehat{h}(\xi, \eta) = \xi \widehat{h}(\xi, \eta) e^{i\eta s_0} \quad \text{for any } \xi, \eta, s_0.$$

From this identity one deduces first that $\widehat{h}$ is localized at $(0, 0)$, and then that it is a linear combination of Dirac masses δ_0, first derivatives of Dirac masses $\partial_i \delta_0$, and $\Delta \delta_0$. Thus, h is a second order polynomial with an isotropic second order part. Coming back to g, and using that it belongs to L^2, it has to be a Gaussian up to symmetries.

Step 5: up to the symmetries, Gaussians are the unique complex-valued maximizers. Assume that f is a maximizer. So is $|f|$, which, by the previous step, is a Gaussian. We can thus assume that $f = Ge^{i\theta}$, where θ is real and smooth. By adding a constant factor to θ, we may also assume that $\mathcal{H}(f) = \mathcal{H}(G) > 0$. Using the original formulation of $\mathcal{H}(f)$ in (6.1), it is easy to see that we must then have that for any $\xi, z \in \mathbb{R}^2$ and $\lambda \in [-1, 1]$:

$$\theta(\xi + z^\perp) + \theta(\xi + \lambda z) = \theta(\xi) + \theta(\xi + z^\perp + \lambda z).$$

Taking the derivative with respect to λ we get that for any $\xi, z \in \mathbb{R}^2, \lambda \in [-1, 1]$:

$$z \cdot \nabla\theta(\xi + \lambda z) = z \cdot \nabla\theta(\xi + z^\perp + \lambda z).$$

which implies that

$$z \cdot \nabla\theta(\eta) = z \cdot \nabla\theta(\eta + z^\perp) \quad \text{for any } \eta, z \in \mathbb{R}^2.$$

As a result, for any $\omega \in \mathbb{S}^1$ and with the notation $\theta(u, s) = \theta(u\omega + s\omega^\perp)$, it holds that

$$\partial_u \theta(u, s) = \partial_u \theta(u, \sigma), \quad \text{for any } u, s, \sigma.$$

This identity already appeared in the previous step. Arguing as there, we obtain that θ is a quadratic polynomial with isotropic second order part, which gives the desired result. $\square$

8.3. Compactness of minimizing sequences and L^2 stability of Gaussians. Let $\mathcal{G}$ be the manifold obtained by letting the symmetries of (CR) act on G:

$$\mathcal{G} = \left\{ \gamma G\left(\frac{x - x_0}{|\gamma|}\right) e^{iv_0 \cdot x + i\beta|x|^2}, \ (\gamma, x_0, v_0, \beta) \in \mathbb{C} \times \mathbb{R}^2 \times \mathbb{R}^2 \times \mathbb{R} \right\}.$$

Theorem 8.4. *Let (f^n) be a sequence in L^2 such that, for some $Q > 0$,*

$$\|f^n\|_2 \to Q \quad \text{and} \quad \mathcal{H}(f^n) \to \frac{\pi}{8}Q^4 \qquad \text{as } n \to \infty.$$

Then

$$\mathrm{dist}_{L^2}(f^n, Q\mathcal{G}) \longrightarrow 0 \qquad \text{as } n \to \infty.$$

Proof. The proof follows by combining Theorem 8.2 with the profile decomposition in [47] via standard concentration compactness arguments. We omit the details. $\square$

This theorem has an easy consequence: dynamic stability of Gaussians by the flow of (CR).

Proposition 8.5. *Assume that (f_0^n) is a sequence in L^2 converging to $\mathcal{G}$. Denote $f_n(t)$ for the solution of* (CR) *with data f_0^n. Then for any time $t > 0$,*

$$\text{dist}_{L^2}(f^n(t), Q\mathcal{G}) \longrightarrow 0 \qquad as\ n \to \infty.$$

APPENDIX A. THE CASE OF DIMENSION 1

It is easy and instructive to describe in dimension 1 the weakly nonlinear ($\epsilon \to 0$), big box ($L \to \infty$) limit of (NLS). We sketch it here and start by considering

$$-i\partial_t u + \Delta u = \epsilon^2 |u|^2 u \qquad \text{set on } \mathbb{T}_L.$$

First expand u in Fourier series: $u = \frac{1}{L} \sum_{K \in \mathbb{Z}/L} a_K e^{2\pi i K x}$, and set $b_K(t) = a_K(L^2 t) e^{-i4\pi^2 t L^2 K^2}$. It solves

$$-i\partial_t b_K = \epsilon^2 \sum_{K_1 - K_2 + K_3 = K} b_{K_1} \bar{b}_{K_2} b_{K_3} e^{i4\pi^2 t L^2 \Omega} \qquad \text{with } \Omega = -K^2 + K_1^2 - K_2^2 + K_3^2.$$

In the above sum, keep only the resonant terms, for which Ω vanishes; the other ones have less impact on the dynamics due to the weak nonlinearity hypothesis. This gives

$$-i\partial_t b_K = -\epsilon^2 |b_K|^2 b_K + 2\epsilon^2 \underbrace{\left(\sum_{J \in \mathbb{Z}/L} |b_J|^2 \right)}_{L\|u\|^2_{L^2(\mathbb{T}_L)}} b_K.$$

Noticing that the L^2 norm of the solution is preserved for all times, and changing the dependent variables one last time to $c_K = b_K e^{i2t\epsilon^2 L \|u\|^2_2}$, we obtain the equation

$$-i\partial_t c_K = -\epsilon^2 |c_K|^2 c_K.$$

In the above equation, K ranges over $\mathbb{Z}/L$. As $L \to \infty$, it is clear that the above converges to the continuous equation

$$-i\partial_t g(t, \xi) = -\epsilon^2 |g(t, \xi)|^2 g(t, \xi), \qquad \text{where } (t, \xi) \in \mathbb{R}^2.$$

Rescaling time gives

$$i\partial_t g(t, \xi) = |g(t, \xi)|^2 g(t, \xi).$$

The corresponding Cauchy problem for an initial data $g(0, \xi) = g_0(\xi)$ is easily solved:

$$g(t, \xi) = g_0(\xi) e^{it|g_0(\xi)|^2}. \tag{A.1}$$

One notices that in this 1D setting, the continuous resonant dynamics are not as interesting as the 2D case. This is due to the absence of non-trivial cubic resonances in dimension 1. Indeed, an application of a normal forms transformation indicates that the nonlinear interactions responsible for any change in the dynamics of "action" variables $|\hat{u}(K)|^2$ are higher order (namely quintic). On the other hand, equation (A.1) suggest that the corresponding "angle" variables tend to become (generically) decorrelated as time increases.

References

[1] V. Arnold, *Mathematical methods of classical mechanics.* Translated from the Russian by K. Vogtmann and A. Weinstein. Second edition. Graduate Texts in Mathematics, 60. Springer-Verlag, New York, 1989.

[2] J. Bennett, N. Bez, A. Carbery, D. Hundertmark, *Heat-flow monotonicity of Strichartz norms.* Anal. PDE 2 (2009), no. 2, 147–158.

[3] J. Bennett, A. Carbery, J. Wright, *A non-linear generalisation of the Loomis-Whitney inequality and applications.* Math. Res. Lett. 12 (2005), no. 4, 443–457.

[4] J. Bennett, A. Carbery, M. Christ, T. Tao, *The Brascamp-Lieb inequalities: finiteness, structure and extremals.* Geom. Funct. Anal. 17 (2008), no. 5, 1343–1415.

[5] J. Bourgain, *Fourier transform restriction phenomena for certain lattice subsets and applications to nonlinear evolution equations.* I. Schrödinger equations. Geom. Funct. Anal. 3 (1993), no. 2, 107–156.

[6] J. Bourgain, *Invariant measures for the 2D-defocusing nonlinear Schrödinger equation*, Comm. Math. Phys. 176 (1996), no. 2, 421–445.

[7] J. Bourgain, *Quasi-periodic solutions of Hamiltonian perturbations of 2D linear Schrödinger equations*, Ann. of Math. (2) 148 (1998), no. 2, 363–439.

[8] J. Bourgain, *On Strichartz's inequalities and the nonlinear Schrödinger equation on irrational tori*, Mathematical aspects of nonlinear dispersive equations, 1–20, Ann. of Math. Stud., 163, Princeton Univ. Press, Princeton, NJ, 2007.

[9] H. Brascamp, E. Lieb, *Best constants in Young's inequality, its converse, and its generalization to more than three functions.* Advances in Math. 20 (1976), no. 2, 151–173.

[10] D. Cai, A. J. Majda, D. W. McLaughlin, E. G. Tabak, *Dispersive wave turbulence in one dimension.* Advances in nonlinear mathematics and science. Phys. D 152/153 (2001), 551–572.

[11] E. Carlen, E. Lieb, M. Loss, *A sharp analog of Young's inequality on SN and related entropy inequalities.* J. Geom. Anal. 14 (2004), no. 3, 487–520.

[12] R. Carles, E. Faou, *Energy cascades for NLS on the torus.* Discrete Contin. Dyn. Syst. 32 (2012), no. 6, 2063–2077.

[13] T. Cazenave and F. Weissler, *The Cauchy problem for the nonlinear Schrödinger Equation in H^s*, Nonlinear Analysis, 14:807–836, 1990.

[14] T. Cazenave, *Semilinear Schrödinger equations*, Courant Lecture Notes in Mathematics, 10. American Mathematical Society, Providence, RI, 2003.

[15] R. Coifman, P.L. Lions, Y. Meyer, S. Semmes, *Compensated compactness and Hardy spaces*, J. Math. Pures et Appl. 72 (1992), pp. 247–286.

[16] 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.

[17] C. Connaughton, S. Nazarenko, A. Pushkarev, *Discreteness and quasiresonances in weak turbulence of capillary waves*, Physical Reviews E, 63, 046306, (2001).

[18] P. Denissenko, S. Lukaschuk, S. Nazarenko, *Gravity wave turbulence in a laboratory flume.* Physical review letters 99.1 (2007): 014501.

[19] B. Dodson, *Global well-posedness and scattering for the defocusing, L^2-critical, nonlinear Schrödinger equation when $d = 2$*, preprint arXiv:1006.1375 [math.AP].

[20] B. Dodson, *Global well-posedness and scattering for the mass critical nonlinear Schrödinger equation with mass below the mass of the ground state*, preprint arXiv:1104.1114 [math.AP].

[21] L. H. Eliasson, S. Kuksin, *KAM for the nonlinear Schrödinger equation*, Ann. of Math. (2) 172 (2010), no. 1, 371–435.

[22] E. Faou, L. Gauckler, C. Lubich, *Sobolev stability of plane wave solutions to the cubic nonlinear Schrödinger equation on a torus.* Comm. PDE, to appear.

[23] D. Foschi, *Maximizers for the Strichartz inequality.* J. Eur. Math. Soc. (JEMS) 9 (2007), no. 4, 739–774.

[24] J. Geng, X. Xu and J. You, *An infinite dimensional KAM theorem and its application to the two dimensional cubic Schrdinger equation.* Adv. Math. 226 (2011), no. 6, 53615402.

[25] P. Gerard, S. Grellier, *The cubic Szegö equation.* Ann. Sci. Ec. Norm. Supér. (4) 43 (2010), no. 5, 761–810.

[26] P. Gérard, S. Grellier, *Invariant tori for the cubic Szegö equation.* Invent. Math. 187 (2012), no. 3, 707–754.

[27] P. Gérard, S. Grellier, *Effective integrable dynamics for a certain nonlinear wave equation*, Anal. PDE 5 (2012), no. 5, 1139–1155.

[28] M. Guardia and V. Kaloshin, *Growth of Sobolev norms in the cubic defocusing nonlinear Schrödinger equation.* arxiv.org:1205.5188

[29] Z. Hani, *Long-time instability and unbounded Sobolev orbits for some periodic nonlinear Schrödinger equations.* To appear in ARMA. Preprint arXiv:1210.7509 [math.AP].

[30] Z. Hani, B. Pausader, *On scattering for the quintic defocusing nonlinear Schrödinger equation on $\mathbb{R} \times \mathbb{T}^2$,* to appear in CPAM, arXiv:1205.6132.

[31] Hardy, G. H.; Wright, E. M., *An introduction to the theory of numbers*, Sixth edition. Oxford University Press, Oxford, 2008. xxii+621 pp. ISBN: 978-0-19-921986-5.

[32] K. Hasselmann,*On the non-linear energy transfer in a gravity-wave spectrum. I. General theory.* J. Fluid Mech. 12 (1962) 481–500.

[33] K. Hasselmann,*On the non-linear energy transfer in a gravity wave spectrum. II. Conservation theorems; wave-particle analogy; irreversibility.* J. Fluid Mech. 15 (1963) 273–281.

[34] D. Hundertmark, V. Zharnitsky, *On sharp Strichartz inequalities in low dimensions.* Int. Math. Res. Not. 2006.

[35] V. Jarnick, *Über die Gitterpunkte auf konvexen Curven*, Math. Z. 24 (1926), 500–518.

[36] E. Kartashova, *Wave resonances in systems with discrete spectra*, AMS Transl. (2), 182 (1998).

[37] E. Kartashova, *Exact and quasiresonances in discrete water wave turbulence.* Physical review letters 98.21 (2007): 214502.

[38] E. Kartashova, S. Nazarenko, O. Rudenko, *Resonant interactions of nonlinear water waves in a finite basin.* Physical Review E 78.1 (2008): 016304.

[39] E. Kartashova, *Discrete wave turbulence*, Europhysics Letters 87.4 (2009): 44001.

[40] E. Kartashova, *Nonlinear Resonance Analysis.* Cambridge University Press, 1 (2010).

[41] R. Killip, T. Tao, M. Visan, *The cubic nonlinear Schrödinger equation in two dimensions with radial data*, J. Eur. Math. Soc. (JEMS) 11 (2009), no. 6, 1203–1258.

[42] N. Kishimoto, *Remark on the periodic mass critical nonlinear Schrödinger equation*, preprint arXiv:1203.6375 [math.AP].

[43] E. Lieb, *Gaussian kernels have only Gaussian maximizers.* Invent. Math. 102 (1990), no. 1, 179–208.

[44] J. Lukkarinen, H. Spohn, *Weakly nonlinear wave equations with random initial data*, Invent. Math. 183 (2011), 79-188.

[45] V. S. L'vov, S. Nazarenko, *Discrete and mesoscopic regimes of finite-size wave turbulence*, Physical Review E 82, 056322 (2010).

[46] A. Majda, D. McLaughlin, E. Tabak, *A one-dimensional model for dispersive wave turbulence.* J. Nonlinear Sci. 7 (1997), no. 1, 9–44.

[47] F. Merle, L. Vega, *Compactness at blow-up time for L2 solutions of the critical nonlinear Schrödinger equation in 2D*, Int. Math. Res. Not. 8 (1998), 399–425.

[48] F. Murat, *Compacité par compensation*, Ann. Scuola Norm. Sup. Pisa Cl. Sci., 5(4), 1978.

[49] S. Nazarenko, *Wave turbulence*, Lecture notes in Physics 825, Springer, Heidelberg, 2011.

[50] S. Nazarenko,*Sandpile behaviour in discrete water-wave turbulence.* Journal of Statistical Mechanics: Theory and Experiment 2006.02 (2006): L02002.

[51] R. Peierls, *The kinetic theory of thermal conduction in crystals*, Annalen der Physik 3 (8): 1055–1101 (1929).

[52] O. Pocovnicu, *Traveling waves for the cubic Szegö equation on the real line*, Analysis & PDE, 4-3 (2011), 379–404.

[53] O. Pocovnicu, *Explicit formula for the solutions of the Szegö equation on the real line and applications*, Discrete Contin. Dynam. Syst. A 31 (2011) no. 3, 607–649.

[54] M. Procesi, C. Procesi, *A KAM algorithm for the resonant non–linear Schrödinger equation.* Preprint arXiv:1211.4242 [math.AP].

[55] H. Spohn, *The phonon Boltzmann equation, properties and links to weakly anharmonic lattice dynamics*, J. Stat. Physics 124 (2–4), 1041–1104 (2006).

[56] R. Strichartz, *Restrictions of Fourier transforms to quadratic surfaces and decay of solutions of wave equations.* Duke Math. J. 44 (1977), no. 3, 705–714.

[57] T. Tao, *Nonlinear dispersive equations: Local and global analysis*, CBMS Regional Conference Series in Mathematics, 106. Published for the Conference Board of the Mathematical Sciences, Washington, DC; by the American Mathematical Society, Providence, RI, 2006. xvi+373 pp. ISBN: 0-8218-4143-2.

[58] T. Tao, *Poincaré's legacies, pages from year two of a mathematical blog. Part I*, American Mathematical Society, Providence, RI, 2009. x+293 pp. ISBN: 978-0-8218-4883-8.

[59] M. Tanaka, N. Yokoyama, *Effects of discretization of the spectrum in water-wave turbulence*, Fluid Dynamics Research 34 (2004) 199–216.

[60] L. Tartar, *Compensated compactness and applications to partial differential equations.* Heriot-Watt Sympos., vol. IV, Pitman, New York, 1979.

[61] P. Tomas, *A restriction theorem for the Fourier transform.* Bull. Amer. Math. Soc. 81 (1975), 477–478.

[62] V. E. Zakharov, N. N. Filonenko, *Energy spectrum for the stochastic oscillations of a fluid surface.* Kohl. Acad. Nauk SSSR 170, 1292–1295 (1966) (translation: Sov. Phys. Dokl. 11, 881–884 (1967)).

[63] V. E. Zakharov, N. N. Filonenko, *Weak turbulence of capillary waves*, Journal of Applied Mechanics and Technical Physics, Volume 8, Issue 5, pp.37–40 (1967).

[64] V. E. Zakharov, V. L'vov, G. Falkovich, *Kolmogorov spectra of turbulence 1. Wave turbulence.* Springer, Berlin, 1992.

[65] V. E. Zakharov, F. Dias, A. Pushkarev, *One-dimensional wave turbulence.* Physics Reports 398.1 (2004): 1–65.

[66] V. E. Zakharov, A. O. Korotkevich, A. N. Pushkarev, A. I. Dyachenko, *Mesoscopic wave turbulence*, 82, 8, 487–491 (2005).

Erwan Faou, INRIA & ENS Cachan Bretagne, Campus de Ker Lann, Avenue Robert Schumann, 35170 Bruz, France.
  E-mail address: `Erwan.Faou@inria.fr`

Pierre Germain, Courant Institute of Mathematical Sciences, 251 Mercer Street, New York 10012-1185 NY, USA
  E-mail address: `pgermain@cims.nyu.edu`

Zaher Hani, Courant Institute of Mathematical Sciences, 251 Mercer Street, New York 10012-1185 NY, USA
  E-mail address: `hani@cims.nyu.edu`