

THE NONLINEAR SCHRÖDINGER EQUATION ON TORI: INTEGRATING HARMONIC ANALYSIS, GEOMETRY, AND PROBABILITY

ANDREA R. NAHMOD

ABSTRACT. The field of nonlinear dispersive and wave equations has undergone significant progress in the last twenty years thanks to the influx of tools and ideas from nonlinear Fourier and harmonic analysis, geometry, analytic number theory and most recently probability, into the existing functional analytic methods. In these lectures we concentrate on the semilinear Schrödinger equation defined on tori and discuss the most important developments in the analysis of these equations. In particular, we discuss in some detail recent work by J. Bourgain and C. Demeter proving the ℓ^2 decoupling conjecture and as a consequence the full range of Strichartz estimates on either rational or irrational tori, thus settling an important earlier conjecture by Bourgain.

1. INTRODUCTION

The nonlinear Schrödinger equation plays an ubiquitous role as a model for dispersive wave-phenomena in nature. Roughly speaking, dispersion means that when no boundary is present, waves of different wavelengths travel at different phase speeds: long wavelength components propagate faster than short ones. This is the reason why over time dispersive waves spread out in space as they evolve in time, while conserving some form of energy. This phenomenon is called broadening of the wave packet. Dispersive wave-phenomena should be contrasted with transport phenomena where all frequencies move at the same velocity or dissipative phenomena (heat equation) in which frequencies gradually taper to zero, that is they do not propagate.

The nonlinear Schrödinger equation serves as a mathematical model for the large class of so-called dispersive partial differential equations [1, 67]. It naturally arises in connection to a variety of different physical problems on flat space, tori, and other manifolds. One of them is nonlinear optics in a so-called Kerr medium where one considers electromagnetic waves in a material (e.g., glass fiber) whose time evolution are governed by Maxwell's equations on \mathbb{R}^3 . The nonlinear Maxwell equations however have disparate scales, and understanding their dynamics is a difficult problem. As a first attempt one looks for further simplifications: asymptotic methods then become useful. A natural *ansatz* is to write the electric field E as a Taylor series whose leading term is a small amplitude wave packet of the form

$$(1.1) \quad A(t, x)e^{i(\xi_0 \cdot x - \omega_0 t)} + \overline{A(t, x)}e^{-i(\xi_0 \cdot x - \omega_0 t)}$$

Received by the editors May 18, 2015.

2010 *Mathematics Subject Classification.* Primary 42-XX, 35-XX.

The author gratefully acknowledges support from NSF grants DMS 1201443 and DMS 1463714.

with wave vector $\xi_0 \in \mathbb{R}^3$, frequency $\omega_0 \in \mathbb{R}$, and A a small amplitude and slowly varying function. After inserting back into the nonlinear Maxwell equations, formal calculations, transformations, and multiple scale analysis yield a cubic nonlinear Schrödinger equation where time corresponds to the coordinate of the direction of propagation of the wave along the material for the (transformed) amplitude [1, 55]. Essentially the same type of approximations can be made in other problems such as, e.g., water waves. In this context one seeks solutions in which the interface of the fluid region is to leading order a wave packet of the same form as (1.1), i.e., with small $O(\epsilon)$ amplitude and slow spatial variation that are balanced. Lengthy formal calculations then suggest that the envelopes of these wave packets evolve on $O(\epsilon^{-2})$ time scales according to a version of a cubic nonlinear Schrödinger equation [25]. It often turns out that the nonlinear Schrödinger equation (approximately) describes the evolution of envelopes of wave packets on the appropriate nonlinear Schrödinger time scales; for a precise description, see [69, 70]. Other examples of nonlinear Schrödinger equations arising from other physical situations can be found in, for example, [61].

The nonlinear Schrödinger equations also arise as the equations governing Bose–Einstein condensates. Bose–Einstein condensation phenomena were predicted by S. N. Bose [3] and by A. Einstein [29] (1924–25); it is a fascinating phenomena predicted by quantum statistical mechanics. Bose–Einstein condensation however was experimentally achieved only in 1995 by Cornell and Wieman [24] and by W. S. Ketterle [49] who produced the first gaseous condensate. For this they were awarded the 2001 Nobel Prize in Physics.

A Bose–Einstein condensate is the state of matter of a gas of weakly interacting bosonic atoms confined by an external potential and cooled to temperatures very near *absolute zero* (0 Kelvin). In his 2001 Nobel lecture, W. S. Ketterle described *how profoundly the properties of a gas of bosonic atoms changes when you cool down the gas. Then the wave nature of matter tells us that the wave packets which describe an atom, this fuzzy object, becomes larger and larger and when the wave packet expands to the size that the waves of neighboring atoms overlap, then all atoms start to oscillate in concert and form what you may regard a giant matter wave. And this is the Bose–Einstein condensate* [51].

In other words, all bosons occupy the same quantum state and can thus be described by a single wave function $u(t, x)$. The pointwise density of this *gas* at time t is represented by $|u(x, t)|^2$. The interactions between the bosons lead to nonlinear contributions to the Schrödinger equation for this quantum system. Considering only binary collisions between the bosons, one sees that u satisfies a cubic nonlinear Schrödinger equation, which in this context is often called the *Gross–Pitaevskii equation* after work by Gross [39] and by Pitaevskii [58]. Physically, it makes sense to study the problem both in the periodic and the nonperiodic setting. Recently there has been intense activity and breakthrough results, particularly by L. Erdős, B. Schlein, and H. T. Yau in the (mathematically) rigorous derivation of the defocusing cubic nonlinear Schrödinger equation, both on \mathbb{R}^3 as well as \mathbb{T}^3 from the dynamics of many-body quantum systems. We refer the interested reader to [19–22, 30–36, 38, 53, 59] and references therein.

Bose–Einstein condensation is based on the wave nature of particles, which is at the heart of quantum mechanics. In a simplified picture, bosonic atoms in a gas may be regarded as quantum-mechanical wave packets with an extension of

their thermal *de Broglie wavelength* (the position uncertainty associated with the thermal momentum distribution). The lower the temperature, the longer is the de Broglie wavelength. When atoms are cooled to the point where the thermal de Broglie wavelength is comparable to the interatomic separation, then the atomic wave packets “overlap” and the *indistinguishability of particles* becomes important. Bosons undergo a phase transition and form a Bose–Einstein condensate, a dense and coherent cloud of atoms all occupying the same quantum mechanical state [49].

Graphically, we can visualize this as follows [49, 50]:

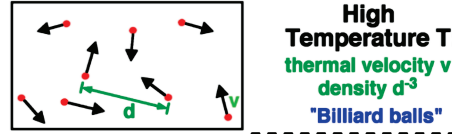


FIGURE 1. Gas at high temperature, treated as a system of billiard balls, with thermal velocity v and density d^{-3} , where d is the distance between bosonic particles.

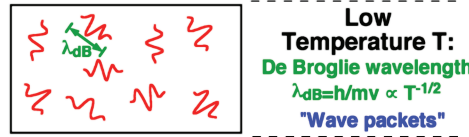


FIGURE 2. Simplified quantum description of gas at low temperature, in which the particles are regarded as wave packets with a spatial extent of the order of the *de Broglie wavelength*, λ_{dB} .



FIGURE 3. Gas at the transition temperature for Bose–Einstein condensation, when λ_{dB} becomes comparable to d . The wave packets overlap and a Bose–Einstein condensate forms.



FIGURE 4. Pure Bose condensate (giant matter wave), which remains as the temperature approaches absolute zero and the thermal cloud disappears.

As mentioned above, the nonlinear Schrödinger equation arising from many-body quantum bosonic atoms makes physical sense for bosons in a three-dimensional cube with periodic boundary conditions (or from an experimental perspective a rectangular box). The nonlinear Schrödinger equation, however, has very different

behavior on \mathbb{T}^d from that on \mathbb{R}^d since *dispersion is weaker* on (periodic) domains. Furthermore, by now we have several examples of nonlinear Schrödinger and wave equations defined on \mathbb{R}^d for which it is mathematically proven that dispersion sets in and, after a time long enough, solutions settle into a purely linear behavior. This phenomenon is often referred to as *scattering* (asymptotic stability). For linear solutions, energy at any given frequency does not migrate to higher or lower frequencies, that is there is no *forward* or *backward cascade*. Hence, as a consequence of scattering, certain nonlinear solutions in \mathbb{R}^d also will avoid these cascades. The situation is believed to be quite different for dispersive equations on compact domains. For example in the periodic case energy cascades and out-of-equilibrium dynamics are expected [23, 40, 45]. It is then not surprising that for understanding the time dynamics of solutions to the nonlinear Schrödinger on tori, one needs to bring to bear tools and ideas from many different other areas of mathematics, such as nonlinear Fourier and harmonic analysis, geometry, probability, analytic number theory, dynamical systems, and others. In these notes we will touch upon a few of these connections by explaining some key results and focus on the spectacular resolution of the ℓ^2 decoupling conjecture by J. Bourgain and C. Demeter [14, 15] (see also [12, 13]). Their results in turn (and in particular) solve a 1993 conjecture by Bourgain and yield the predicted full range of dispersive estimates (known as Strichartz estimates) for solutions to the Cauchy initial value problem for the nonlinear Schrödinger equation (p-NLS) on general rectangular d -dimensional tori,

$$(1.2) \quad \begin{cases} iu_t + \Delta u = \lambda |u|^{p-1}u, \\ u(0, x) = \phi(x), \quad x \in \Lambda_d(\theta), \end{cases}$$

where ϕ is the initial profile, $\lambda = \pm 1$, $p > 1$, $u : \mathbb{R} \times \Lambda_d(\theta) \rightarrow \mathbb{C}$ and for $\beta := (\beta_1, \beta_2, \dots, \beta_d)$, $\beta_j > 0$, $j = 1, \dots, d$ we define the d -dimensional tori by

$$\Lambda_d(\beta) := (\mathbb{R}/\beta_1\mathbb{Z}) \times (\mathbb{R}/\beta_2\mathbb{Z}) \times \dots \times (\mathbb{R}/\beta_d\mathbb{Z}).$$

When $\beta_j = 1$, we have $\Lambda_d(\beta) = \mathbb{T}^d := (\mathbb{R}/\mathbb{Z})^d$, the square d -dimensional torus \mathbb{T}^d of 1-periodic functions. If at least one of the ratios $\frac{\beta_j}{\beta_{j'}} \notin \mathbb{Q}$, $j \neq j'$, we call $\Lambda_d(\beta)$ an *irrational* torus,¹ otherwise we call $\Lambda_d(\beta)$ a *rational* torus. The point here is that if $\Lambda_d(\beta)$ is rational, then by tiling the space with copies of $\Lambda_d(\beta)$, we can construct some square torus which contains only full copies of $\Lambda_d(\beta)$. For irrational tori, one can never construct such a square torus. Irrational tori come up naturally experimentally as well as in KAM theory and Hamiltonian chaos ([10] and references therein).

In the context of nonlinear dispersive equations, the general rectangular tori $\Lambda_d(\beta)$ were first studied by Bourgain [4] where he noted that the methods from analytic number theory, previously employed in [5, 6] to obtain for the first time *some* dispersive estimates for the Schrödinger equation on the square \mathbb{T}^d , could not be used in the general rectangular case. It is fairly straightforward to see that the dispersive estimates known for \mathbb{T}^d imply those for the rational torus as well. This left open two questions: 1) obtaining the full range of expected dispersive estimates on \mathbb{T}^d and 2) proving dispersive estimates for irrational tori. The second was thought to be a much harder question from the harmonic analysis and analytic number theory point of view adopted in [4]. These questions were recently

¹Without any loss of generality we can assume $\frac{1}{2} < \beta_j < 2$, $j = 1, \dots, d$.

answered by Bourgain and Demeter [14, 15] who, rather than use analytic number theory, rely on decoupling inequalities (in discrete restriction phenomena) and sophisticated arguments from multilinear harmonic analysis (adapted wave packet decompositions, parabolic rescaling, bilinear square function estimates, multilinear Kakeya, multiscale bootstrap), and also implicitly on ideas from incidence theory. Their work is at the core of these notes, and it is the focus of Section 4. In Section 5 we discuss the interplay of deterministic and probabilistic approaches in the well-posedness theory of nonlinear Schrödinger equations. In Section 2 we start with some preliminaries.

Notation. We use $A \lesssim B$ to denote an estimate of the form $A \leq CB$ for some $C > 0$ which may depend on the underlying dimension as well as on fixed parameters such as p or s . However, we record dependence on variable parameters such as ε using the notation \lesssim_ε . The asymptotic notation $A \gtrsim B$ is defined analogously, and $A \sim B$ will mean $A \lesssim B$ and $B \lesssim A$. By H^s (resp., \dot{H}^s) we denote the usual inhomogeneous (resp., homogeneous) Sobolev spaces. Given a function $u = u(t, x)$ depending on time t and the space variable x , we denote by $\|u\|_{L_t^q L_x^r} := \|\|u\|_{L_x^r}\|_{L_t^q}$ the mixed space-time Lebesgue norm. For a fixed time interval I , the spaces $L^\infty(I; H^s)$ (resp., $C(I; H^s)$) denote the space of functions which are in L^∞ in t (resp., continuous in t) with values in the Banach space H^s .

2. PRELIMINARIES

Whether the underlying space is \mathbb{R}^d , a torus, or some other manifold, a basic question when studying the Cauchy initial value problem (1.2) is that of *well-posedness*—that is, i) *existence*, ii) *uniqueness*, and iii) *stability*—of solutions for initial data in a given Banach space, which in these notes we assume to be the Sobolev space H^s . To solve this question, the idea is to use a fixed point theorem on a space of functions whose norm is dictated by strong estimates for $v(t, x) := S(t)\phi(x)$, the solution of the associated linear problem,

$$(2.1) \quad \begin{cases} iv_t + \Delta v = 0, \\ v(0, x) = \phi(x). \end{cases}$$

One should of course recall that under reasonable regularity assumptions, (1.2) is equivalent to (2.2) below, thanks to the Duhamel principle. Formally, well-posedness is defined as follows.

Definition 2.1. We say that the Cauchy initial value problem (1.2) is locally well posed in H^s if for any ball \mathcal{B} in H^s there exists a time $T > 0$ and a Banach space of functions $X^s \subset L^\infty([-T, T]; H^s)$ such that for each initial data $\phi \in \mathcal{B}$, there exists a unique solution $u \in X^s \cap C([-T, T]; H^s)$ of the integral equation

$$(2.2) \quad u(x, t) = S(t)\phi(x) + c_\lambda \int_0^t S(t-t') |u(t', x)|^{p-1} u(t', x) dt'.$$

Moreover, the map $\phi \mapsto u$ is continuous from H^s into $C([-T, T]; H^s)$. If $T > 0$ can be taken arbitrarily large, then we say that the initial value problem is *globally* well posed.

Remark 2.2. Note that the definition above yields uniqueness on $X^s \cap C([-T, T]; H^s)$ but not necessarily on $C([-T, T]; H^s)$. Proving uniqueness on $C([-T, T]; H^s)$ requires additional work, and when it holds, the local well-posedness is said to be *unconditional*.

2.1. A tour onto \mathbb{R}^d . In these notes we primarily focus on the periodic setting, as in (1.2) where the problems associated to the nonlinear Schrödinger equation (NLS) are harder and less understood than when the underlying domain² is \mathbb{R}^d . Before doing so however let us recall a few important ideas about p-NLS on \mathbb{R}^d ,

$$(2.3) \quad \begin{cases} iu_t + \Delta u = \lambda |u|^{p-1}u, \\ u(0, x) = \phi(x), \quad x \in \mathbb{R}^d, \end{cases}$$

where $\phi \in H^s$, $\lambda = \pm 1$, $p > 1$, $u : \mathbb{R} \times \mathbb{R}^d \rightarrow \mathbb{C}$. The nonlinear Schrödinger equation (2.3) enjoys many symmetries (cf. [67]) among which we highlight the following.

Time-reversal symmetry:

$$(2.4) \quad \phi(x) \mapsto \overline{\phi(x)}, \quad u(t, x) \mapsto \overline{u(-t, x)};$$

Scaling symmetry:

$$(2.5) \quad \phi(x) \mapsto \mu^{-\frac{2}{(p-1)}} \phi\left(\frac{x}{\mu}\right) =: \phi_\mu(x), \quad u(t, x) \mapsto \mu^{-\frac{2}{(p-1)}} u\left(\frac{t}{\mu^2}, \frac{x}{\mu}\right),$$

for any dilation factor $\mu > 0$. From (2.5) we immediately notice that if the initial datum is in $\dot{H}^{s_c}(\mathbb{R}^d)$, $s_c := \frac{d}{2} - \frac{2}{p-1}$, then $\|\phi_\mu\|_{\dot{H}^{s_c}} = \|\phi\|_{\dot{H}^{s_c}}$ and (2.3) is scale invariant. The Sobolev regularity s_c is then called the *critical* scaling regularity. Note that the criticality of H^s depends on both the power p and the dimension d . In fact since we have that

$$\|\phi_\mu\|_{\dot{H}^s} \sim \mu^{s_c - s} \|\phi\|_{\dot{H}^s},$$

we can classify the difficulty of the p-NLS (2.3) above in terms of regularity of its data. When $s > s_c$, note that as $\mu \rightarrow \infty$, the norm of $\|\phi_\mu\|_{\dot{H}^s}$ gets smaller; the space H^s is called *subcritical* in this case. On the other hand if $s < s_c$, we have that as $\mu \rightarrow \infty$, the norm of $\|\phi_\mu\|_{\dot{H}^s}$ gets larger; the space H^s is then called *supercritical*. When $s = s_c$, the space \dot{H}^{s_c} is *critical* since as we noted above, as $\mu \rightarrow \infty$, the norm of $\|\phi_\mu\|_{\dot{H}^s}$ does not change. Accordingly, the local well-posedness theory for equation (2.3) is fairly well understood in the subcritical regime—when the equation can be treated as a perturbation of the linear one—while in the supercritical regime only nondeterministic results are available. We return to the latter in Section 5.

Remark 2.3. In proving local well-posedness for (2.2) in the *subcritical regime* $s > s_c$, one shows via a fixed point argument that the time of existence T is roughly like $\|\phi\|_{H^s}^{-\alpha}$ for some $\alpha > 0$.

We also note p-NLS conserves both mass and the Hamiltonian, i.e.,

$$(2.6) \quad \mathcal{M}(u(t)) := \int |u(t, x)|^2 dx = \int |\phi(x)|^2 dx =: \mathcal{M}(\phi(t)),$$

and the Hamiltonian, i.e.,

$$(2.7) \quad \begin{aligned} H(u(t)) &:= \frac{1}{2} \int |\nabla u(t, x)|^2 dx + \frac{2\lambda}{p+1} \int |u(t, x)|^{p+1} dx \\ &= \frac{1}{2} \int |\nabla \phi(x)|^2 dx + \frac{2\lambda}{p+1} \int |\phi(x)|^{p+1} dx =: H(\phi). \end{aligned}$$

²There is by now a substantial body of work on NLS on \mathbb{R}^d . The interested reader might want to consult [18, 67].

Remark 2.4. If $\lambda = 1$, (2.7) and (2.6) give a global in time bound for the H^1 norm of $u(t, x)$; p-NLS is called *defocusing* in this case. On the other hand, if $\lambda = -1$, the energy could be negative and blow up may occur; p-NLS is then called *focusing*.

In particular, we have that for those p and d for which the H^1 space is subcritical (e.g., cubic $p = 3$ in $d = 2$), Remark 2.3 and the conservation in time of the Hamiltonian (2.7) allow one to obtain global well-posedness by iterating the local theory.

To prove local well-posedness for p-NLS, one needs to find a suitable Banach space X^s as in Definition 2.1 on which to prove that the map

$$\Phi : u \longmapsto S(t)\phi + c_\lambda \int_0^t S(t-t') |u(t', x)|^{p-1} u(t', x) dt'$$

is a contraction, whence the solution $u(t, x)$ is as in (2.2). Determining a good choice of X^s is part of the problem. It is dictated by being able to have sufficiently good estimates for the linear evolution $S(t)\phi$ on such space so that then, at least in the subcritical regime and on short time intervals, one can show that $\Phi(u)$ —hence the solution u —satisfy similar estimates. The most basic and at the same time important space-time estimates that $S(t)\phi$, the solution to the linear problem (2.22), satisfy are the so-called Strichartz estimates.

2.1.1. The Strichartz estimates on \mathbb{R}^d . The Strichartz estimates are intimately related to the (L^p, L^2) restriction problem for the Fourier transform (to the paraboloid in the case of the Schrödinger equation). To understand the connection, it is illustrative to review Strichartz's original argument [56, 62]. We begin by briefly recalling the restriction question. For f a Schwartz function on \mathbb{R}^d , first recall that \widehat{f} , the Fourier transform of f is defined as

$$\widehat{f}(\xi) := \int_{\mathbb{R}^d} f(x) e^{-i x \cdot \xi} dx, \quad \xi \in \mathbb{R}^d.$$

By the Hausdorff–Young inequality we know that if f is an $L^p(\mathbb{R}^d)$ -integrable function, then for all $1 \leq p \leq 2$,

$$\|\widehat{f}\|_{L^{p'}(\mathbb{R}^d)} \leq \|f\|_{L^p(\mathbb{R}^d)},$$

where p' is the dual exponent defined by $\frac{1}{p'} + \frac{1}{p} = 1$. In fact one can easily show by scaling and the Kintchine inequality that if the inequality $\|\widehat{f}\|_{L^q(\mathbb{R}^d)} \leq C \|f\|_{L^p(\mathbb{R}^d)}$ holds, then one must have that $q = p'$ and $p \leq 2$.

If however we are interested in considering \widehat{f} as a function on B , the unit ball in \mathbb{R}^d for example (rather than on all of \mathbb{R}^d), then a larger range of estimates is available. Indeed, we have that for all $q \leq p'$ and $1 \leq p \leq 2$,

$$\|\widehat{f}\|_{L^q(B)} \leq C \|f\|_{L^p(\mathbb{R}^d)},$$

since $L^{p'}(B) \subseteq L^q(B)$ in this case. The range $q \leq p'$ and $p \leq 2$ can be shown similarly to be best possible.

The situation becomes more delicate when, instead of B , we consider \widehat{f} restricted onto a compact hypersurface in \mathbb{R}^d . This case is also more interesting given such restrictions arise quite naturally when considering distributional solutions to certain linear PDE, as we will be soon see. Consider then $1 \leq p \leq 2$, $f \in L^p(\mathbb{R}^d)$, and \mathcal{S} a compact C^2 smooth hypersurface on \mathbb{R}^d with nonzero Gaussian curvature. Let

$d\sigma$ denote the canonical measure³ associated to \mathcal{S} . The (L^p, L^2) -restriction for the Fourier transform asks whether the map

$$\mathcal{R} : f \mapsto \widehat{f}|_{\mathcal{S}}$$

extends to a bounded operator from $L^p(\mathbb{R}^d) \mapsto L^2(\mathcal{S}, d\sigma)$; i.e., whether

$$(2.8) \quad \|\widehat{f}|_{\mathcal{S}}\|_{L^2(\mathcal{S}, d\sigma)} \leq C_{d, \mathbf{p}, \mathcal{S}} \|f\|_{L^p(\mathbb{R}^d)}.$$

By duality, the restriction estimate (2.8) is equivalent to the *extension estimate*

$$(2.9) \quad \|(gd\sigma)^\vee\|_{L^{p'}(\mathbb{R}^d)} \leq C'_{d, \mathbf{p}, \mathcal{S}} \|g\|_{L^2(\mathcal{S}, d\sigma)}$$

for any smooth function g on \mathcal{S} where

$$(gd\sigma)^\vee(x) := \int_{\mathcal{S}} g(\xi) e^{i\xi \cdot x} d\sigma(\xi)$$

is the inverse Fourier transform of the measure $gd\sigma$.

If $\mathbf{p} = 1$, (2.8) is always true by the Riemann–Lebesgue lemma. But if $\mathbf{p} = 2$, then $\widehat{f}|_{\mathcal{S}}$ is meaningless since $\widehat{f} \in L^2(\mathbb{R}^d)$ and the d -dimensional Lebesgue measure of \mathcal{S} is zero. Moreover, if \mathcal{S} is a plane, then no $\mathbf{p} > 1$ is allowed, while if $\mathcal{S} = \mathbb{S}^{d-1}$, the unit sphere in \mathbb{R}^d , then P. Tomas and E. M. Stein gave an affirmative answer for $1 \leq \mathbf{p} \leq \frac{2(d+1)}{d+3}$ (and it can be shown to fail for any \mathbf{p} larger). In general, the answer to this question depends on the *curvature* of \mathcal{S} . Indeed, if \mathcal{S} is a compact hypersurface with nonvanishing Gaussian curvature, one can show using stationary phase methods that for every $z \in \mathbb{R}^d$,

$$(2.10) \quad |\check{\sigma}_{\mathcal{S}}(z)| \lesssim (1 + |z|)^{-\frac{d-1}{2}}.$$

By a standard T^*T argument,⁴ proving estimate (2.8) is equivalent to proving that $\|\check{\sigma} * f\|_{L^{p'}(\mathbb{R}^d)} \leq C_{d, \mathbf{p}} \|f\|_{L^p(\mathbb{R}^d)}$, which follows from (2.10) in conjunction with Littlewood–Paley and (complex) interpolation (or fractional integration in a direction transverse to \mathcal{S} and convex interpolation).

Finally, we note that one can show that for each $\mathbf{p}' \geq \frac{2(d+1)}{d-1}$ fixed, (2.9) is equivalent to the discretized estimate,

$$(2.11) \quad \left(\frac{1}{|B_R|} \int_{B_R} \left| \sum_{\xi \in \Lambda} a_\xi e^{2\pi i \cdot \xi} \right|^{p'} \right)^{1/p'} \lesssim \delta^{\frac{n}{2p'} - \frac{n-1}{4}} \|a_\xi\|_{\ell^2(\Lambda)},$$

for each $0 < \delta < 1$, $a_\xi \in \mathbb{C}$, ball $B_R \subset \mathbb{R}^n$ with $R \sim \delta^{-1/2}$ and $\delta^{1/2}$ -separated lattice $\Lambda \subset \mathcal{S}$.

We refer the reader to [56, 60, 65, 66, 72] and references therein for further details about the restriction problem and the more general (L^p, L^q) -restriction conjecture problem.

Consider now the linear Schrödinger equation on \mathbb{R}^d ,

$$(2.12) \quad iv_t + \Delta v = 0,$$

and let us assume that the initial datum $v(0, x) = \phi$ is a smooth function such that $\text{supp } \widehat{\phi} \subset \{|\xi| < 1\}$. By taking the Fourier transform in space of (2.12) and solving

³e.g., if $\mathcal{S} = \mathbb{S}^{d-1}$, $d\sigma$ is surface measure. Of interest for PDE is the case when $d\sigma$ is a measure supported on \mathcal{S} .

⁴i.e., $T : L^p \rightarrow L^2 \iff T^*T : L^p \rightarrow L^{p'}$, cf. [60].

the corresponding ODE, we have that the solution to the linear problem is defined by $\hat{v}(t, \xi) := e^{-i|\xi|^2 t} \hat{\phi}(\xi)$, whence we have that

$$(2.13) \quad v(t, x) = \int_{\mathbb{R}^d} e^{i(x \cdot \xi + h(\xi)t)} \hat{\phi}(\xi) d\xi = (\hat{\phi} d\sigma)^\vee(t, x),$$

where we have denoted by $h(\xi) = -|\xi|^2$, and $d\sigma$ is the measure in \mathbb{R}^{d+1} carried by the paraboloid hypersurface

$$\Sigma := \{(\xi, \tau) \in \mathbb{R}^d \times \mathbb{R} : \tau = h(\xi)\}$$

defined by

$$\int_{\mathbb{R}^{d+1}} \psi(\xi, \tau) d\sigma(\xi, \tau) = \int_{\mathbb{R}^d} \psi(\xi, h(\xi)) d\xi$$

for any ψ continuous on \mathbb{R}^{d+1} . In other words, $v(t, x) = \mathcal{R}^*(\phi)(x, t)$, where \mathcal{R}^* is the *extension operator* defined to be the adjoint of $\mathcal{R}(f) = \hat{f}|_\Sigma$, the operator that *restricts* the Fourier transform on \mathbb{R}^{d+1} to the the paraboloid Σ . Then by the Tomas–Stein endpoint estimate and duality, we obtain that

$$\|v(t, x)\|_{L^{\frac{2(d+2)}{d}}(\mathbb{R}^{d+1})} = \|(\hat{\phi} d\sigma)^\vee\|_{L^{\frac{2(d+2)}{d}}(\mathbb{R}^{d+1})} \lesssim \|\hat{\phi}\|_{L^2(\mathbb{R}^d)} = \|\phi\|_{L^2(\mathbb{R}^d)}.$$

Remark 2.5. Note that since the Tomas–Stein endpoint estimate is scale invariant by rescaling $\phi(x)$ and (parabolically) $v(t, x)$, one may remove the assumption made above that $\hat{\phi}$ is supported in the unit frequency ball.

The full range of Strichartz estimates can be derived via a shorter argument—essentially due to Ginibre and Velo and to Yajima—thanks to the explicit form of the linear semigroup (see [18, 67] and references therein). Indeed, from (2.13) we have that

$$(2.14) \quad v(t, x) = S(t)\phi(x) := e^{it\Delta}\phi(x) = K_t * \phi(x) = \frac{c}{t^{d/2}} \int_{\mathbb{R}^d} e^{i\frac{|x-y|^2}{2t}} \phi(y) dy,$$

whence we immediately obtain that

$$(2.15) \quad \|S(t)\phi\|_{L_x^\infty(\mathbb{R}^d)} \lesssim \frac{1}{t^{d/2}} \|\phi\|_{L_x^1(\mathbb{R}^d)},$$

called the *dispersive estimate*. On the other hand, given that the linear semigroup is unitary and commutes with other Fourier multipliers, we clearly have that for any s ,

$$(2.16) \quad \|S(t)\phi\|_{H_x^s(\mathbb{R}^d)} = \|\phi\|_{H_x^s(\mathbb{R}^d)}.$$

Interpolating (2.15) and (2.16) with $s = 0$, we have for any $1 \leq p \leq 2$ the *fixed time* estimates

$$(2.17) \quad \|S(t)\phi\|_{L_x^{p'}(\mathbb{R}^d)} \lesssim \frac{1}{t^{\frac{d}{p} - \frac{d}{2}}} \|\phi\|_{L_x^p(\mathbb{R}^d)}$$

where $\frac{1}{p} + \frac{1}{p'} = 1$. These estimates are not enough since the initial data is usually only assumed to be in an L^2 -based Sobolev space; however, by combining (2.17) with duality, T^*T arguments, fractional integration in time, and interpolation, we obtain the full range of Strichartz estimates (cf. [18, 56, 60, 62, 67] and references therein) which we are now ready to state.

For $d \geq 1$, we define \mathcal{A} the set of *admissible* exponents to be those pairs (q, r) such that $2 \leq q, r \leq \infty$ and

$$(2.18) \quad \frac{2}{q} = \frac{d}{2} - \frac{d}{r}, \quad (q, r, d) \neq (2, \infty, 2).$$

Theorem 2.6 (Strichartz estimates on \mathbb{R}^d). *For $(q, r) \in \mathcal{A}$, we have the homogeneous estimate*

$$\|S(t)\phi\|_{L_t^q L_x^r(\mathbb{R} \times \mathbb{R}^d)} \lesssim \|\phi\|_{L_x^2(\mathbb{R}^d)},$$

and for any other admissible pair (\tilde{q}, \tilde{r}) we also have the inhomogeneous estimate

$$\left\| \int_0^t S(t-t') N(u)(t') dt' \right\|_{L_t^q L_x^r(\mathbb{R} \times \mathbb{R}^d)} \lesssim \|N(u)\|_{L_t^{\tilde{q}'} L_x^{\tilde{r}'}(\mathbb{R} \times \mathbb{R}^d)},$$

where $\frac{1}{\tilde{q}} + \frac{1}{\tilde{q}'} = 1 = \frac{1}{\tilde{r}} + \frac{1}{\tilde{r}'}$ and $N(\cdot)$ is any Lipschitz continuous function.

Remark 2.7. It can be proved via a standard Knapp example and scaling arguments that the admissibility of (q, r) is a necessary condition. We also note that these estimates hold on $[-T, T] \times \mathbb{R}^d$ in lieu of $\mathbb{R} \times \mathbb{R}^d$.

2.2. Back to the periodic setting. We start by rescaling the tori $\Lambda_d(\beta)$ so that we can use coordinates based on the regular square torus \mathbb{T}^d and work with Fourier series based on the standard integer lattice \mathbb{Z}^d . In this way we incorporate the geometry of $\Lambda_d(\beta)$ into the Laplace operator, which after such rescaling is defined by

$$(2.19) \quad \Delta_\Theta := \Theta_1 \frac{\partial^2}{\partial x_1^2} + \Theta_2 \frac{\partial^2}{\partial x_2^2} + \cdots + \Theta_d \frac{\partial^2}{\partial x_d^2}, \quad \Theta_j = \frac{1}{\beta_j^2}, \quad j = 1, \dots, d.$$

In other words, for $\mathbf{k} = (k_1, \dots, k_d) \in \mathbb{Z}^d$,

$$(2.20) \quad \widehat{\Delta f}(\mathbf{k}) := -(\Theta_1 k_1^2 + \cdots + \Theta_d k_d^2) \widehat{f}(\mathbf{k}),$$

where as usual, we have denoted the Fourier transform and Fourier series as

$$\widehat{f}(\mathbf{k}) = \int_{\mathbb{T}^d} e^{-2\pi i \mathbf{k} \cdot x} f(x) dx, \quad (\mathbf{k} \in \mathbb{Z}^d) \quad \text{and} \quad f(x) = \sum_{\mathbf{k} \in \mathbb{Z}^d} e^{2\pi i \mathbf{k} \cdot x} \widehat{f}(\mathbf{k}), \quad (x \in \mathbb{T}^d).$$

Our problem (1.2) can then be rewritten on \mathbb{T}^d as

$$(2.21) \quad \begin{cases} iu_t + \Delta_\Theta u = \lambda |u|^{p-1} u, \\ u(0, x) = \phi(x), \quad x \in \mathbb{T}^d. \end{cases}$$

We note that the solution $u(t, x)$ to the linear Schrödinger equation,

$$(2.22) \quad \begin{cases} iu_t + \Delta_\Theta u = 0, \\ u(0, x) = \phi(x), \quad x \in \mathbb{T}^d, \end{cases}$$

is given by

$$(2.23) \quad u(t, x) = e^{it\Delta_\Theta} u_0 = \sum_{\mathbf{k} \in \mathbb{Z}^d} e^{2\pi i (\mathbf{k} \cdot x - t \sum_{j=1}^d \Theta_j k_j^2)} \widehat{u_0}(\mathbf{k}).$$

Of course, equation (2.21) also conserves mass (2.6) and the Hamiltonian (2.7).

Remark 2.8. The periodic cubic NLS equation

$$\begin{cases} iu_t + \Delta u = \lambda |u|^2 u, \\ u(x, 0) = u_0(x), \quad x \in \mathbb{T}^3, \end{cases}$$

is the one governing Bose–Einstein condensation alluded to in the introduction.

3. STRICHARTZ ESTIMATES ON TORI

In the periodic setting, the tools from harmonic analysis we used to establish the Strichartz estimates are no longer available. Even in the case of the square torus \mathbb{T}^d , obtaining some Strichartz estimates for $S(t)u_0(x)$ is highly nontrivial and required new ideas. They were introduced by Bourgain for the rational torus as a conjecture:

Conjecture 3.1 (Bourgain [6]). *Assume that \mathbb{T}^d is a rational torus, and for $N \geq 1$, let $\phi \in L^2(\mathbb{T}^d)$ be a smooth function such that the $\text{supp } \hat{\phi} \subset [-N, N]^d \subset \mathbb{Z}^d$. Then for any $\epsilon > 0$ the following estimates should hold:*

$$\begin{aligned} \|S(t)\phi\|_{L_t^q L_x^q(\mathbb{T}^{d+1})} &\lesssim C_q \|\phi\|_{L_x^2(\mathbb{T}^d)} \quad \text{if } q < \frac{2(d+2)}{d}, \\ \|S(t)\phi\|_{L_t^q L_x^q(\mathbb{T}^{d+1})} &\ll N^\epsilon \|\phi\|_{L_x^2(\mathbb{T}^d)} \quad \text{if } q = \frac{2(d+2)}{d}, \\ \|S(t)\phi\|_{L_t^q L_x^q(\mathbb{T}^{d+1})} &\lesssim C_q N^{\frac{d}{2} - \frac{d+2}{q}} \|\phi\|_{L_x^2(\mathbb{T}^d)} \quad \text{if } q > \frac{2(d+2)}{d}. \end{aligned}$$

In [6] Bourgain in fact partially proved these bounds in the following cases:

- i) $d = 1, 2$ and $q > \frac{2(d+2)}{d}$;
- ii) $d = 3$ and $q > 4$; and
- iii) $d \geq 4$ and $q > \frac{2(d+4)}{d}$.

His proof relies on Weyl's sum estimates, the Hardy–Littlewood circle method, and the Tomas–Stein restriction theorem. Partial improvements were obtained in [26, 47].

More recently, Bourgain [11] improved his results from [6] by establishing the above Conjecture 3.1 for $d \geq 4$ and $q > \frac{2(d+3)}{d}$ by relying on multilinear harmonic analysis techniques for restriction and the Kakeya problems developed by Bennet, Carbery, and Tao [2] and by Bourgain and Guth [16]. These techniques will once again mark the way for the resolution of the full conjecture by Bourgain and Demeter [14] as we will see in Section 4 below.

Remark 3.2. Bourgain also proved in [6] that dispersion is indeed weaker in the periodic setting by proving that when $d = 1$ the endpoint L^6 estimate which holds on \mathbb{R} with constant independent of N is false in the periodic setting. More precisely, he showed

$$\left\| \sum_{|k| \leq N} a_k e^{i(kx + k^2 t)} \right\|_{L^6(\mathbb{T} \times \mathbb{T})} \geq c(\log N)^{\frac{1}{6}} \left(\sum_{|k| \leq N} |a_k|^2 \right)^{1/2}.$$

The failure of the endpoint estimate $\|e^{it\Delta} \phi_N\|_{L^4(\mathbb{T}^2 \times \mathbb{T})} \lesssim \|\phi_N\|_{L^2(\mathbb{T}^2)}$ when $d = 2$ was established by Takaoka and Tzvetkov [64].

3.1. Strichartz estimates and the Fourier restriction. For any given $N \in \mathbb{N}$, let $S_{d,N}$ be the set

$$\{(k_1, \dots, k_d) \in \mathbb{Z}^d : |k_j| \leq N, 1 \leq j \leq d\}.$$

For $q > 1$, let $A_{q,d,N}$ represent the best constant satisfying

$$(3.1) \quad \sum_{\mathbf{k} \in S_{d,N}} \left| \hat{f}(\mathbf{k}, |\mathbf{k}|^2) \right|^2 \leq A_{q,d,N} \|f\|_{q'}^2,$$

where $\mathbf{k} = (k_1, \dots, k_d) \in S_{d,N}$, $|\mathbf{k}| = \sqrt{k_1^2 + \dots + k_d^2}$ and f is any $L^{q'}$ -function on \mathbb{T}^{d+1} and $q' = q/(q-1)$. As we mentioned above, Bourgain [6] obtained (in particular) the following estimate for the square torus \mathbb{T}^d :

$$(3.2) \quad A_{q,d,N} \leq CN^{d - \frac{2(d+2)}{q} + \epsilon} \text{ for } q > \frac{2(d+4)}{d}.$$

By duality, it is straightforward to see that the Strichartz estimates

$$(3.3) \quad \left\| \sum_{\mathbf{k} \in S_{d,N}} a_{\mathbf{k}} e^{i(\mathbf{k} \cdot \mathbf{x} + |\mathbf{k}|^2 t)} \right\|_{L^q(\mathbb{T}^{d+1})} \leq \sqrt{A_{q,d,N}} \left(\sum_{\mathbf{k} \in S_{d,N}} |a_{\mathbf{k}}|^2 \right)^{1/2}$$

are in fact equivalent to the discrete Fourier restriction estimates

$$(3.4) \quad \left(\sum_{\mathbf{k} \in S_{d,N}} |\hat{f}(\mathbf{k}, |\mathbf{k}|^2)|^2 \right)^{1/2} \leq \sqrt{A_{q,d,N}} \|f\|_{q'}.$$

To understand how the rational character of the torus enters in a basic fashion, let us review Bourgain's result for the square torus in the case where $q = 4$ and $d = 2$. We would like to show that $A_{4,2,N} < N^\epsilon$, $\epsilon > 0$. Bourgain [6] reduced the problem to estimating the number of representations of an integer as a sum of squares. Let

$$f(\mathbf{x}, t) = \sum_{|\mathbf{k}| < N} a_{\mathbf{k}} e^{i(\mathbf{k} \cdot \mathbf{x} + |\mathbf{k}|^2 t)} \text{ with } (\mathbf{x}, t) \in \mathbb{T}^2 \times \mathbb{T},$$

and for a given integer j and $\mathbf{p} \in \mathbb{Z}^2$ define

$$\mathcal{C}_{\mathbf{p},j} := \{\mathbf{k} \in \mathbb{Z}^2 : |\mathbf{k}| \leq N \text{ and } |\mathbf{k}|^2 + |\mathbf{p} - \mathbf{k}|^2 = j\},$$

and let $r_{\mathbf{p},j} = \#\mathcal{C}_{\mathbf{p},j}$. If we square our function f , we have that

$$f(\mathbf{x}, t)^2 = \sum_{\mathbf{p}} e^{i(\mathbf{p} \cdot \mathbf{x})} \left[\sum_{\mathbf{k}} a_{\mathbf{k}} a_{\mathbf{p}-\mathbf{k}} e^{i(|\mathbf{k}|^2 + |\mathbf{p}-\mathbf{k}|^2 t)} \right] = \sum_{\mathbf{p},j} \left(\sum_{\mathbf{k} \in \mathcal{C}_{\mathbf{p},j}} a_{\mathbf{k}} a_{\mathbf{p}-\mathbf{k}} \right) e^{i(\mathbf{p} \cdot \mathbf{x} + jt)},$$

so by taking the L^2 -norm we find that

$$(3.5) \quad \|f^2\|_{L^2(\mathbb{T}^2 \times \mathbb{T})}^2 \leq \left\{ \max_{\substack{|\mathbf{p}| \leq 2N \\ |j| \leq 2N^2}} r_{\mathbf{p},j} \right\} \left(\sum |a_{\mathbf{k}}|^2 \right)^2.$$

We can rewrite $|\mathbf{k}|^2 + |\mathbf{p} - \mathbf{k}|^2 = j$ as

$$(2k_1 - p_1)^2 + (2k_2 - p_2)^2 = 2j - |\mathbf{p}|^2,$$

so we have that $r_{\mathbf{k},j}$ is bounded by the number of solutions of

$$X_1^2 + X_2^2 = R^2,$$

where $R^2 = 2j - |\mathbf{p}|^2 \lesssim N^2$. Hence the right-hand side of (3.5) is bounded by the number of integer lattice points (X_1, X_2) that lie on the circle of radius R . Since there are at most $\exp(C \frac{\log R}{\log \log R}) \sim R^\epsilon$ such points, we get the desired estimate.

The case is much more difficult when generalizing to any given p and d . Hu and Li in [47] presented a variant of the proof of Bourgain's result (3.4), which makes use of the Hardy–Littlewood circle method and estimates on level sets just as Bourgain's does. Their proof of Bourgain's level set estimates is, however, somewhat simpler.

We briefly sketch their proof next. When q is large, the desired estimate (3.4) follows immediately from the following result (cf. [6, 47]).

Theorem 3.3. *For any $\sigma > 0$, any $d \geq 1$, and any $q > \frac{4(d+2)}{d}$, there exists a constant C , independent of N , such that*

$$\sum_{\mathbf{k} \in \mathbb{Z}^d} e^{-\frac{\sigma|\mathbf{k}|^2}{N^2}} \left| \hat{f}(\mathbf{k}, |\mathbf{k}|^2) \right|^2 \leq CN^{d-\frac{2(d+2)}{q}} \|f\|_{q'}^2,$$

for all $f \in L^{q'}(\mathbb{T}^{d+1})$.

The proof of this theorem in turn follows from the *Hardy–Littlewood circle method*, a tool to count the number of representations of a given integer as an arbitrary sum of powers. Let us recall some simple aspects about it.

For $A \subseteq \mathbb{Z}$, let $f(z) := \sum_{a \in A} z^a$, an analytic function converging in the open unit disc. Then we can write,

$$f(z)^s = \sum_{a_1 \in A} \cdots \sum_{a_s \in A} z^{a_1} \cdots z^{a_s} = \sum_{N=0}^{\infty} r_{A,s}(N) z^N,$$

where $r_{A,s}(N)$ is the number of representations of N as the sum of s elements of $A \subseteq \mathbb{Z}$. In other words, $r_{A,s}(N)$ is the number of solutions of the equation

$$N = a_1 + a_2 + \cdots + a_s$$

with $a_1, a_2, \dots, a_s \in A$.

We can now apply Cauchy's theorem to the summation above to obtain that

$$r_{A,s}(N) = \frac{1}{2\pi i} \int_{|z|=\rho} \frac{f(z)^s}{z^{N+1}} dz$$

for any $\rho \in (0, 1)$. This is the original form of the circle method. Note that the integral above counts the number of ways the number N can be written as a sum of arbitrary powers of s . The evaluation of such an integral is not a trivial task, and it requires breaking up our circle into *major arcs* and *minor arcs*.

Sketch of the proof of Theorem 3.3. For $r \in \mathbb{N}$, let

$$\mathcal{P}_r := \{y \in \mathbb{N} : 1 \leq y \leq r, (y, r) = 1\}.$$

For $a \in \mathcal{P}_r$, define the interval $J_{a/r}$ as

$$J_{a/r} = \left(\frac{a}{r} - \frac{1}{Nr}, \frac{a}{r} + \frac{1}{Nr} \right).$$

$J_{a/r}$ is referred to as a *minor arc* if $r \geq N/10$ and as a *major arc* if $r < N/10$. By the Dirichlet principle,⁵ we can then partition the interval $(0, 1]$ into a union of major and minor arcs as

$$(0, 1] = \bigcup_{1 \leq r \leq N, a \in \mathcal{P}_r} J_{a/r} = \mathcal{M}_1 \cup \mathcal{M}_2,$$

where \mathcal{M}_1 is the union of all major arcs and \mathcal{M}_2 is the union of all minor ones. If χ_J is the characteristic function on the set $J_{a/q}$, then set

$$K_{a/r}(\mathbf{x}, t) := K_\sigma(\mathbf{x}, t) \chi_{J_{a/r}}(t),$$

⁵Recall that the Dirichlet principle states that, given any $N \in \mathbb{N}$ and any $x \in (0, 1]$, there exist $a, r \in \mathbb{N}$ such that $|x - \frac{a}{r}| \leq \frac{1}{Nr}$, $1 \leq r \leq N$, $a \in \mathcal{P}_r$.

where

$$(3.6) \quad K_\sigma(\mathbf{x}, t) := \sum_{\mathbf{k} \in \mathbb{Z}^d} e^{-\frac{\sigma|\mathbf{k}|^2}{N^2}} e^{i|\mathbf{k}|^2 t} e^{i\mathbf{k} \cdot \mathbf{x}}.$$

The point is that now one can show, for $1 \leq r \leq N$, $a \in \mathcal{P}_r$, and $q > \frac{2(d+1)}{d}$,

$$\|K_{a/r}\|_{L^q} \leq C_{r,d,q} N^{d-\frac{d+2}{q}},$$

which then leads to the estimate for $q > \frac{2(d+2)}{d}$, that

$$\|K_\sigma\|_{L^q} \leq C_{q,\sigma} N^{d-\frac{d+2}{q}}.$$

Since

$$\sum_{\mathbf{n} \in \mathbb{Z}^d} e^{-\frac{\sigma|\mathbf{k}|^2}{N^2}} \left| \hat{f}(\mathbf{k}, |\mathbf{k}|^2) \right| = \langle K_\sigma * f, f \rangle,$$

if we apply Hölder's inequality and the Hausdorff–Young inequality, we have that

$$\langle K_\sigma * f, f \rangle \leq \|K_\sigma\|_{q/2} \|f\|_{q'}^2;$$

whence since $q > \frac{4(d+2)}{d}$ we have the desired conclusion. \square

The estimate for smaller cases of q , follow from level set estimates [6, 47] of the form:

Theorem 3.4 ([47]). *Let F_N be a periodic function on \mathbb{T}^{d+1} such that*

$$F_N(\mathbf{x}, t) = \sum_{\mathbf{k} \in S_{d,N}} a_{\mathbf{k}} e^{i\mathbf{k} \cdot \mathbf{x}} e^{i|\mathbf{k}|^2 t},$$

where $\{a_{\mathbf{k}}\}$ is a sequence with $\sum_{\mathbf{k}} |a_{\mathbf{k}}|^2 = 1$. For any $\lambda > 0$, let

$$E_\lambda = \{(\mathbf{x}, t) \in \mathbb{T}^{d+1} : |F_N(\mathbf{x}, t)| > \lambda\}.$$

Then for any $Q > 0$ such that $Q \geq N$, we have that

$$(3.7) \quad \lambda^2 |E_\lambda|^2 \leq C_1 Q^{d/2} |E_\lambda|^2 + \frac{C_2 N^\epsilon}{Q} |E_\lambda|$$

holds for all λ and $\epsilon > 0$. The constants C_1 and C_2 are independent of N and Q .

Without loss of generality assume Q is a positive integer and consider $N \leq Q \leq N^2$. The idea is to suitably decompose the kernel K_σ in (3.6) into the sum of two kernels $K_{1,Q} + K_{2,Q}$ such that

$$\|K_{1,Q}\|_{L^\infty} \leq C_1 Q^{d/2}$$

and

$$\|\widehat{K_{2,Q}}\|_{L^\infty} \leq \frac{C_2 N^\epsilon}{Q},$$

whence the estimates (3.7) follow. To find such decomposition, the key is to choose an appropriate function supported on $[0, 1]$ so that if we denote by $\Phi(t)$ its periodic extension, we define

$$K_{1,Q}(\mathbf{x}, t) = \frac{1}{\Phi(0)} K_\sigma(\mathbf{x}, t) \Phi(t) \quad \text{and} \quad K_{2,Q} = K_\sigma - K_{1,Q}.$$

The Φ that works is the periodic extension of the function

$$\sum_{Q \leq r < 2Q} \sum_{a \in \mathcal{P}_r} \varphi\left(\frac{t - a/r}{1/r^2}\right),$$

where φ is a bump function supported on a small interval; say $[2^{-8}, 2^{-7}]$. Full details can be found in [47].

The point is that one can then show as corollaries the following estimates in [6]:

- (1) If $\lambda \geq CN^{d/4}$ for some suitable $C > 0$, then the level set E_λ , defined in Theorem 3.4 above, satisfies $|E_\lambda| \lesssim N^\varepsilon \lambda^{-\frac{2(d+2)}{d}}$.
- (2) For each positive $\varepsilon > 0$, we have that $\sqrt{A_{q,d,N}} \leq C_\varepsilon N^{\frac{d}{2} - \frac{d+2}{q} + \varepsilon}$ provided $q > \frac{2(d+4)}{d}$; which in turn immediately yields (3.3) in this case.

3.2. The Strichartz estimates on general tori. As mentioned above, the study of the NLS on general rectangular tori was first started in the work of Bourgain [10] where it was shown that certain Strichartz estimates with a loss of derivative hold. Some other partial results for the NLS on irrational tori were obtained in [17, 28, 41, 63]. The combined range of estimates proved for irrational tori in these works are weaker than those proved by Bourgain in [6] due to number-theoretical difficulties. A completely different approach to the problem was recently taken in the work of Bourgain and Demeter [14, 15]; see also prior work by C. Demeter [27]. Such an approach has led to the full range of Strichartz estimates conjectured in [6] (cf. [10]) up to ε -loss for irrational tori. This ε -loss was removed in subsequent work by Killip and Viřan [52].

In [14] Bourgain and Demeter actually prove a stronger result than the Strichartz estimates. Namely they establish the ℓ^2 -**decoupling conjecture** (Theorem 4.1 below) whence, in addition to proving the Strichartz estimates on general (rational or irrational) tori, they also derive, perhaps somewhat surprisingly, new results in number theory and in incidence geometry theory. Our interest in these notes is in understating how Bourgain and Demeter establish the Strichartz estimates for general (rational or irrational) tori.

Relabeling the notation. From now through the end of Section 4 we follow the notation in [14] and relabel dimension d as $n - 1$. Hence \mathbb{T}^{d+1} will become \mathbb{T}^n . Furthermore, the L^q in Conjecture 3.1 and subsequent presentation above will become L^p (that is we will use p in lieu of q). This p should not be confused with the power nonlinearity of NLS.

Theorem 3.5 (Strichartz estimates for general tori). *Let $\phi \in L^2(\mathbb{T}^{n-1})$ with $\text{supp } \hat{\phi} \subset [-N, N]^{n-1}$. Then for each $\varepsilon > 0$, $p \geq \frac{2(n+1)}{n-1}$, and each interval $I \subset \mathbb{R}$ with $|I| \gtrsim 1$, we have*

$$(3.8) \quad \|e^{it\Delta_\Theta} \phi\|_{L^p(\mathbb{R}^{n-1} \times I)} \lesssim_\varepsilon N^{\frac{n-1}{2} - \frac{n+1}{p} + \varepsilon} |I|^{1/p} \|\phi\|_2.$$

The implicit constant does not depend on I, N or $\Theta := (\Theta_1, \dots, \Theta_{n-1})$ as in (2.19).

Remark 3.6. Theorem 3.5 in particular fully establishes Bourgain's Conjecture 3.1. Bourgain showed in [6, Proposition 3.113] how to remove⁶ the ε -loss for $p > \frac{2(n+1)}{n-1}$ in the case of square (rational) tori. Recall that the ε -loss is *necessary* when $q = \frac{2(n+1)}{n-1}$ as discussed above in Remark 3.2. Recent work by Killip and Viřan [52] shows how to remove the ε -loss for $p > \frac{2(n+1)}{n-1}$ in the case of irrational tori; in fact their argument works for either rational or irrational tori.

⁶i.e., obtain scale invariant estimates

Remark 3.7. As a consequence of Theorem 3.5 one can prove in particular that the p-NLS equation (2.21) on general tori is locally well posed in $H^s(\mathbb{T}^d)$ for any $s > \frac{d}{2} - \frac{2}{p-1}$.

The proof of Theorem 3.5 follows rather quickly once the ℓ^2 -decoupling theorem (Theorem 4.1 below) is proven. The idea is to use the discrete version of the ℓ^2 -decoupling theorem, as was done by Bourgain in [11], and apply a change of variables to (3.8) which puts us in the perfect position to apply the discrete estimate. In the next section we then focus on proving Theorem 4.1 and indicate in Section 4.5 how to obtain Theorem 3.5 from it.

4. ℓ^2 DECOUPLINGS

We provide a brief overview of the proof of the ℓ^2 -decoupling conjecture by J. Bourgain and C. Demeter in [14]. We borrow heavily from L. Guth's notes on the topic [43]. As mentioned above, in order to remain faithful to the literature, we relabel the spatial dimension d as $n - 1$ so that the space-time dimension will now be n . Hence $n = 2$ means 1 spatial dimension and so forth.

Throughout this section we take S to be a compact C^2 hypersurface in \mathbb{R}^n with positive definite second fundamental form. The example corresponding to the Schrödinger equation is the *truncated elliptic paraboloid*,

$$P^{n-1} := \{(\xi_1, \dots, \xi_{n-1}, \xi_1^2 + \dots + \xi_{n-1}^2) \in \mathbb{R}^n : |\xi_i| \leq 1/2\}.$$

We assume $n \geq 2$, and to fix ideas we will frequently give examples where $n = 2$.

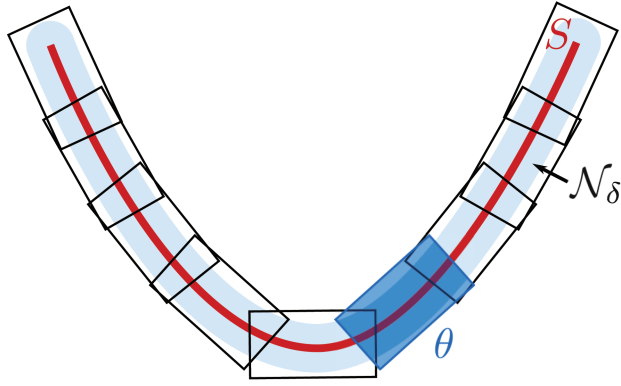


FIGURE 5. The setup for the truncated parabola P^1 . Each rectangular region represents a θ “slab”, and T is the collection of all such θ .

For a set θ and $f \in L^p(\mathbb{R}^n)$, we define f_θ as the Fourier restriction of f to θ ; that is $f_\theta := (\hat{f}|_\theta)^\vee$. The main result is the following ℓ^2 -decoupling theorem:

Theorem 4.1 (Theorem 1.1 in [14]). *Let S be a compact C^2 hypersurface in \mathbb{R}^n with positive definite second fundamental form. Let $\mathcal{N}_\delta S$ be the δ -neighborhood of S , and let T be a covering of $\mathcal{N}_\delta S$ by blocks θ of dimension $\delta^{1/2} \times \dots \times \delta^{1/2} \times \delta$. If*

$\text{supp}(\hat{f}) \subseteq \mathcal{N}_\delta S$, then for $p \geq \frac{2(n+1)}{n-1}$ and $\varepsilon > 0$,

$$(4.1) \quad \|f\|_p \leq C_{p,n,\varepsilon} \delta^{-\frac{n-1}{4} + \frac{n+1}{2p} - \varepsilon} \left(\sum_{\theta \in T} \|f_\theta\|_p^2 \right)^{1/2}.$$

Remark 4.2. Once Theorem 4.1 above is established, the subcritical estimate

$$\|f\|_p \lesssim_\varepsilon \delta^{-\varepsilon} \left(\sum_{\theta \in T} \|f_\theta\|_p^2 \right)^{1/2} \quad \text{where } 2 \leq p \leq \frac{2(n+1)}{n-1},$$

follows by a localization argument and interpolation between the $p = 2$ case and the endpoint $p_c = \frac{2(n+1)}{n-1}$ one in (4.1).

The endpoint p_c here is hinted at in prior discussions of this topic. G. Garrigós and A. Seeger proved in [37] that, up to the ε term, the exponent $-\frac{n-1}{4} + \frac{n+1}{2p} - \varepsilon$ of δ in Theorem 4.1 is optimal. Thus the clear breaking point for when this exponent is a constraint is precisely p_c .

Remark 4.3. In [11] Bourgain proved (4.1) for the subcritical $p = \frac{2n}{n-1}$. His proof relies on the multilinear theory developed in [2] and an induction on scales analysis as in [16]. We will see below that these same ingredients come into play in the proof of Theorem 4.1

4.1. Main steps. In this section we hope to adequately motivate and develop the tools used in Bourgain and Demeter's work; however, for brevity we will only be able to sketch the main ideas of their proof. We follow closely L. Guth's notes on the subject [43].

We will encounter below the norm $\|f\|_{L^{\frac{2(n+1)}{n}}}$, and we note that the algebraic properties of p_c allow the convenient bound

$$\|f\|_{L^{\frac{2(n+1)}{2}}} \leq \|f\|_{L^2}^{1/2} \|f\|_{L^{p_c}}^{1/2}$$

via the Hölder inequality.

4.1.1. Decoupling norms. We begin by inspecting the right-hand side of the ℓ^2 -decoupling inequality (4.1) further. For any f such that $\text{supp } \hat{f} \subseteq \mathcal{N}_\delta S$ and $\Omega \subseteq \mathbb{R}^n$ any domain, we fix a covering T of $\mathcal{N}_\delta S$ and define

$$(4.2) \quad \|f\|_{L^{p,\delta}(\Omega)} := \left(\sum_{\theta \in T} \|f_\theta\|_{L^p(\Omega)}^2 \right)^{1/2} = \|\|f_\theta\|_{L^p(\Omega)}\|_{\ell^2(T)}.$$

This turns out to be a norm with some similar properties to the L^p -norms; in particular it satisfies the Hölder-type inequality

$$(4.3) \quad \|f\|_{L^{q,\delta}(\Omega)} \leq \|f\|_{L^{q_1,\delta}(\Omega)}^{1-\alpha} \|f\|_{L^{q_2,\delta}(\Omega)}^\alpha$$

for $1 \leq q, q_1, q_2 \leq \infty$, $0 < \alpha < 1$, and $\frac{1}{q} = (1-\alpha)\frac{1}{q_1} + \alpha\frac{1}{q_2}$.

It is also useful to record the following super-additive property of the decoupling norms (4.2), which can be proven using the Minkowski inequality for the $\ell^{p/2}$ -norm.

Lemma 4.4. *If Ω is a disjoint union of Ω_j and $p \geq 2$, then for any δ and any f with $\text{supp } \hat{f} \subseteq \mathcal{N}_\delta S$, we have*

$$\sum_j \|f\|_{L^{p,\delta}(\Omega_j)}^p \leq \|f\|_{L^{p,\delta}(\Omega)}^p.$$

The benefit of this lemma is that it allows us to break Ω into disjoint pieces and reduce matters to proving a decoupling inequality on each piece. Indeed, as a corollary of Lemma 4.4 we have what Guth [43] calls *parallel decoupling*:

Lemma 4.5 (Parallel decoupling). *Suppose that Ω is a disjoint union of Ω_j , $\text{supp } \hat{f} \subseteq \mathcal{N}_\delta S$, and $p \geq 2$. Suppose that for each j we have the inequality*

$$\|f\|_{L^p(\Omega_j)} \leq M \|f\|_{L^{p,\delta}(\Omega_j)}.$$

Then we also have the inequality

$$\|f\|_{L^p(\Omega)} \leq M \|f\|_{L^{p,\delta}(\Omega)}.$$

4.1.2. *Decoupling constant.* We define the decoupling constant $D_p(R)$ as

$$D_p(R) := \inf\{C > 0 : \|f\|_{L^p(B_R)} \leq C \|f\|_{L^{p,1/R}(B_R)}\},$$

where the infimum is taken over all f with $\text{supp } \hat{f} \subset \mathcal{N}_{1/R}S$. We note that $D_p(R)$ also depends on S , but we will ignore this point for now. The claim is that, at the endpoint p_c ,

$$D_{p_c}(R) \lesssim R^\epsilon.$$

4.1.3. *Multiple scales.* We consider the problem at multiple scales in Fourier space. Instead of breaking $\mathcal{N}_{1/R}S$ into pieces at the scale of θ one asks what happens if one starts with a function supported in $\tau \subseteq \mathcal{N}_{1/R}S$ and then breaks τ into θ caps. The result is the following proposition.

Proposition 4.6. *If $\tau \subseteq \mathcal{N}_{1/R}S$ is an $r^{-1/2}$ cap for some $r \leq R$, $\text{supp } \hat{f} \subseteq \tau$, and $\theta \subseteq \mathcal{N}_{1/R}S$ are $R^{-1/2}$ caps as before, then*

$$\|f\|_{L^p(\Omega)} \lesssim D_p(R/r) \left(\sum_{\theta \subseteq \tau} \|f_\theta\|_{L^p(B_R)}^2 \right)^{1/2}.$$

The proof of this proposition is based on parabolic rescaling, in which we apply a linear transformation so that the region τ has diameter 1 and then use the parallel decoupling Lemma 4.5 above; see [43]. As a corollary of Proposition 4.6 we get the following estimate:

Corollary 4.7. *For any radii $R_1, R_2 \geq 1$, we have*

$$D_p(R_1 R_2) \lesssim D_p(R_1) D_p(R_2).$$

As a result, we see that there is a unique $\gamma = \gamma(n, p)$ such that for all R, ϵ we have

$$(4.4) \quad R^{\gamma-\epsilon} \lesssim D_p(R) \lesssim R^{\gamma+\epsilon}.$$

We want to show—in particular—that $\gamma = 0$ at the endpoint p_c . At this point the linear methods fail us when trying to establish (4.1). The key idea is to look then at multilinear versions of the decoupling problem. The crucial point being that *for the decoupling problem*, Bourgain and Demeter are able to show that the multilinear version and the linear version are essentially equivalent [14]. This is quite surprising, as other problems such as the linear Kakeya or restriction conjectures are currently out of reach and seem harder to prove than their multilinear formulations. We can thus attack the decoupling problem using multilinear methods, which we will leverage to our advantage (section 4.3). But before doing so, we present next an

overview of the multilinear Keakeya theory of Bennett, Carbery, and Tao [2] (see also [44]) whose techniques are at the core of the multilinear approach to the decoupling problem.

4.2. Multilinear Keakeya. We follow Guth [42] in which he succinctly states, “the multilinear Keakeya inequality is a geometric estimate about the overlap pattern of cylindrical tubes in \mathbb{R}^n pointing in different directions.” Bourgain and Demeter crucially rely on it to prove to prove Theorem 4.1.

Theorem 4.8 (Multilinear Keakeya). *Suppose that $\{\ell_{j,a}\}$ is a finite collection of lines in \mathbb{R}^n , where $j \in \{1, \dots, n\}$ and $a \in \{1, \dots, N_j\}$ such that each line $\ell_{j,a}$ makes an angle of at most $(10n)^{-1}$ with the x_j -axis. Let $T_{j,a}$ be the characteristic function of the 1-neighborhood of $\ell_{j,a}$, and let Q_S denote any cube of side length S . Then for any $\varepsilon > 0$ and any $S \geq 1$, the following integral inequality holds:*

$$(4.5) \quad \int_{Q_S} \prod_{j=1}^n \left(\sum_{a=1}^{N_j} T_{j,a} \right)^{\frac{1}{n-1}} \lesssim_\varepsilon S^\varepsilon \prod_{j=1}^n N_j^{\frac{1}{n-1}}.$$

Figure 6 is an example of a setup for multilinear Keakeya. The area being considered is simply that within the square Q_S . In addition, considering the values in the inequality (4.5) we note that $\sum_{a=1}^{N_j} T_{j,a}$ represents the color density of our overlayed transparencies (see Figure 7). Since there is a product on the left-hand side of (4.5), the only portion which is being counted at all are the areas where the

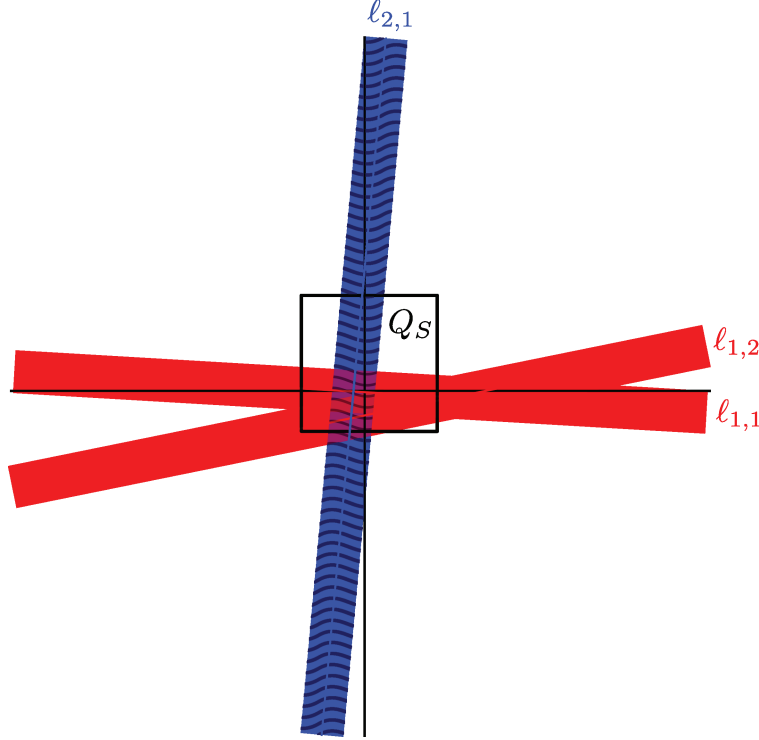
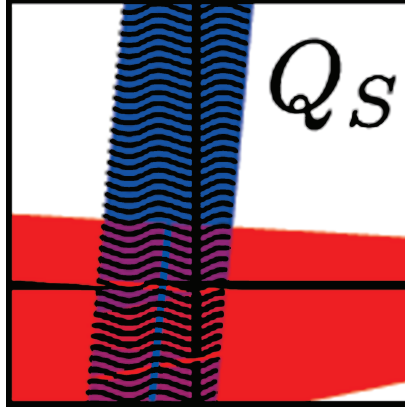


FIGURE 6. An example of the setup for multilinear Keakeya

FIGURE 7. Zooming in on the square Q_S .

tubes intersect transversally, i.e., the intersection of the $\ell_{2,1}$ tube with the $\ell_{1,1}$ or $\ell_{1,2}$ tubes.

4.2.1. *Nearly axis parallel.* The method of proving the multilinear Kakeya inequality (Theorem 4.8) which was established by Bennett, Carbery, and Tao in [2] and is also followed by Guth in [42], is to first reduce to nearly axis parallel tubes:

Theorem 4.9. *For every $\varepsilon > 0$ there is some $\delta > 0$ so that the following holds. Suppose that $\ell_{j,a}$ are lines in \mathbb{R}^n and that each line $\ell_{j,a}$ makes an angle of at most δ with the x_j -axis. Then for any $S \geq 1$ and any cube Q_S of side length S , we have*

$$\int_{Q_S} \prod_{j=1}^n \left(\sum_{a=1}^{N_j} T_{j,a} \right)^{\frac{1}{n-1}} \lesssim_\varepsilon S^\varepsilon \prod_{j=1}^n N_j^{\frac{1}{n-1}}.$$

The claim is that Theorem 4.9 implies Theorem 4.8. Suppose Theorem 4.9 is true, then if $\delta \geq (10n)^{-1}$ (for a given ε), then we are easily done. If $\delta < (10n)^{-1}$ however, we would like to stretch along an axis whose lines are not within δ , bringing the lines closer to the axis. The only problem with this idea is that doing so also inevitably pulls other lines away from their axes. Clearly if one axis has lines which make too much of an angle, and the other axes are well within δ , we may be able to stretch the space so that all the lines are within delta.

One problem with this idea, however, is that the amount we stretch relies on knowing information about the lines which we do not have. Obviously the other issue is that this does not help us if more than one set of lines makes an angle of more than δ . The technique to handle both problems will be to split up over all possible contributions from various possible directions of lines and scale them each independently.

Assume that for $\varepsilon > 0$, the corresponding $\delta > 0$ from Theorem 4.9 is less than $(10n)^{-1}$. Then we split the spherical cap S_j of radius $(10n)^{-1}$ into caps $S_{j,\beta}$ of radius $\delta/10$, and then apply a linear change of coordinates to each cap centering it on the standard unit vector e_j .

In this case, the specific angle each $\ell_{j,a}$ makes is not important as we know it is bounded by $(10n)^{-1}$, and so as we center each $S_{j,\beta}$ this linear change of coordinates has a controlled effect on lengths and areas, and we can bound the overall integral

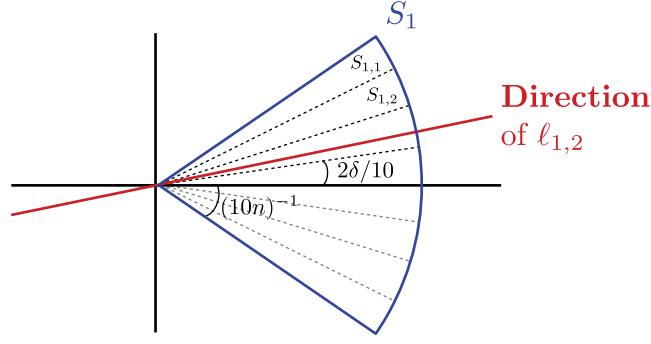


FIGURE 8. How to reduce multilinear Keakeya to the nearly axis parallel case: We split S_j into pieces, each of which has smaller radius than δ , and then sum over all contributions where the direction of $\ell_{j,a}$ is in $S_{j,\beta}$ (one from each S_j). Our angles here are not to scale.

by a sum of all combinations of contributions from these transformed systems, each of which is controlled by Theorem 4.9.

4.2.2. *Axis parallel (Loomis and Whitney)*. The idea of the rest of the argument will be to further simplify matters by zooming in sufficiently close so that nearly axis parallel tubes look almost like axis parallel tubes. In this case we can get the bound we want using the Loomis–Whitney inequality, proven in [54], which states

Theorem 4.10 (Loomis and Whitney). *Suppose that $f_j : \mathbb{R}^{n-1} \rightarrow \mathbb{R}$ are measurable functions, and let $\pi_j : \mathbb{R}^n \rightarrow \mathbb{R}^{n-1}$ be the linear map that forgets the j^{th} coordinate:*

$$\pi_j(x_1, \dots, x_n) = (x_1, \dots, x_{j-1}, x_{j+1}, \dots, x_n).$$

Then the following inequality holds:

$$\int_{\mathbb{R}^n} \prod_{j=1}^n f_j(\pi_j(x))^{\frac{1}{n-1}} \leq \prod_{j=1}^n \|f_j\|_{L^1(\mathbb{R}^{n-1})}^{\frac{1}{n-1}}.$$

The connection between this theorem and the axis-parallel case is that a line parallel to the x_j -axis can be written as $\pi_j(x) = y_a$ for some $y_a \in \mathbb{R}^{n-1}$. Then, as noted in [42] by Guth, $\sum_a T_{j,a}(x) = \sum_a \chi_{B(y_a,1)}(\pi_j(x))$, and applying Loomis–Whitney with $f_j = \sum_a \chi_{B(y_a,1)}$, we have

$$\int_{\mathbb{R}^n} \prod_{j=1}^n \left(\sum_{a=1}^{N_j} T_{j,a} \right)^{\frac{1}{n-1}} = \int_{\mathbb{R}^n} \prod_{j=1}^n (f_j(\pi_j(x)))^{\frac{1}{n-1}} \leq \prod_{j=1}^n \|f_j\|_{L^1(\mathbb{R}^{n-1})}^{\frac{1}{n-1}} \leq \omega_{n-1} N_j,$$

where ω_{n-1} is the volume of the $n-1$ dimensional unit ball. Therefore the axis parallel case does follow quickly from Loomis–Whitney, so we proceed to describe loosely the “zooming in” part of the argument.

Given a cube Q_S , we begin by splitting it up into small enough Q such that each tube $T_{j,a}$ which intersects a small Q can be covered by $\tilde{T}_{j,a,R}$, an axis-parallel tube with slightly larger radius R .⁷ Note that, since the $\tilde{T}_{j,a,R}$ actually cover the $T_{j,a}$

⁷Exact details for what constitutes sufficiently small and slightly larger are contained in [42].

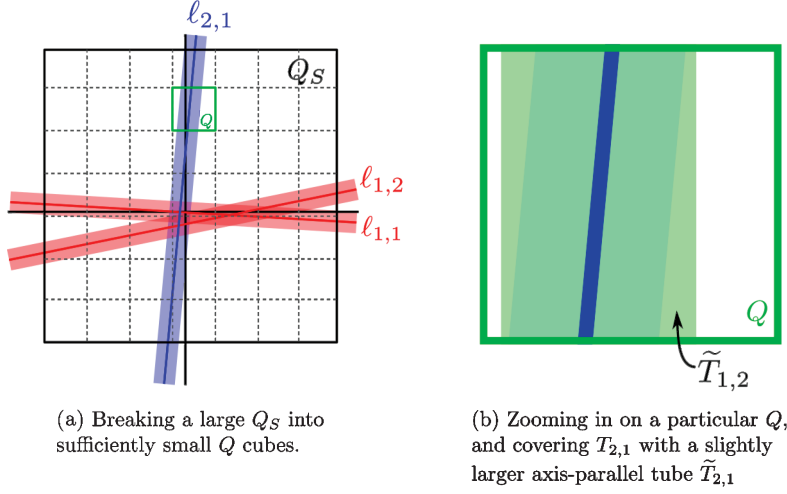


FIGURE 9. Zooming in

within Q , we have

$$\int_Q \prod_{j=1}^n \left(\sum_a T_{j,a} \right)^{\frac{1}{n-1}} \leq \int_Q \prod_{j=1}^n \left(\sum_a \tilde{T}_{j,a,R} \right)^{\frac{1}{n-1}} \lesssim R^n \prod_{j=1}^n N_j(Q)^{\frac{1}{n-1}},$$

where the last inequality follows from using Loomis–Whitney, and $N_j(Q)$ indicates the number of tubes $T_{j,a}$ intersecting Q . In fact, choosing Q small enough, we can make it so that if the tube $T_{j,a}$ intersects Q , the tube $T_{j,a,\delta^{-1}}$ of radius δ^{-1} around $\ell_{j,a}$ is identically 1 on Q . Therefore

$$R^n \prod_{j=1}^n N_j(Q)^{\frac{1}{n-1}} \lesssim \frac{R^n}{|Q|} \int_Q \prod_{j=1}^n \left(\sum_a T_{j,a,\delta^{-1}} \right)^{\frac{1}{n-1}}.$$

As Guth shows in [42], with the appropriate choice of $|Q|$ and R , we can make $R^n/|Q| \lesssim \delta^n$. Since we can then sum over all Q , this proves the following lemma:

Lemma 4.11. *Suppose that $\ell_{j,a}$ are lines with angle at most δ from the x_j -axis. Then if $S \geq \delta^{-1}$ and if Q_S is any cube of sidelength S , then*

$$\int_{Q_S} \prod_{j=1}^n \left(\sum_a T_{j,a} \right)^{\frac{1}{n-1}} \lesssim \delta^n \int_{Q_S} \prod_{j=1}^n \left(\sum_a T_{j,a,\delta^{-1}} \right)^{\frac{1}{n-1}}.$$

We have essentially traded off making the tubes larger for the δ^n factor. This can be seen as an exploit of the fact that a naive bound for the integrand is to assume that all tubes are identically 1 on Q_S , which yields $\prod_{j=1}^n N_j^{\frac{1}{n-1}}$, and so we lose nothing in the trade.

Without loss of generality, assume Q_S is centered at the origin. Now if $S \geq \delta^{-M}$, we begin *induction on the scales* $\delta^{-1}, \delta^{-2}, \dots, \delta^{-M}$:

$$(4.6) \quad \int_{Q_S} \prod_{j=1}^n \left(\sum_a T_{j,a}(x) \right)^{\frac{1}{n-1}} dx \leq C_n \delta^n \int_{Q_S} \prod_{j=1}^n \left(\sum_a T_{j,a,\delta^{-1}}(x) \right)^{\frac{1}{n-1}} dx$$

$$(4.7) \quad = C_n \int_{\delta Q_S} \prod_{j=1}^n \left(\sum_a T_{j,a,\delta^{-1}}(\delta^{-1}x) \right)^{\frac{1}{n-1}} dx,$$

where (4.7) follows by a change of variables. In the new coordinates, $T_{j,a,\delta^{-1}}(\delta^{-1}x)$ are just unit tubes again, and δQ_S is a cube with side lengths $\geq \delta^{-(M-1)}$, so we repeat the argument. After M repetitions we arrive at

$$\int_{Q_S} \prod_{j=1}^n \left(\sum_a T_{j,a}(x) \right)^{\frac{1}{n-1}} dx \leq C_n^M \int_{\delta^M Q_S} \prod_{j=1}^n \left(\sum_a T_{j,a,\delta^{-M}}(\delta^{-M}x) \right)^{\frac{1}{n-1}} dx,$$

and we can now use our naive bound to find

$$\int_{Q_S} \prod_{j=1}^n \left(\sum_a T_{j,a}(x) \right)^{\frac{1}{n-1}} dx \leq C_n^M (\delta^M S)^n \prod_{j=1}^n N_j^{\frac{1}{n-1}}.$$

If we had been working on a cube Q_S such that $S = \delta^{-M}$, at this point we would only need that $C_n^M \leq S^\varepsilon$ to be done. To accomplish this, we solve $S = \delta^{-M}$ for $M = -\log S / \log \delta$, thus we have

$$C_n^M = S^{-\frac{\log C_n}{\log \delta}}.$$

Therefore given $\varepsilon > 0$, we choose $\delta > 0$ such that $-\frac{\log C_n}{\log \delta} < \varepsilon$. Now we have proven the following lemma.

Lemma 4.12. *Given $\varepsilon > 0$, there exists $\delta > 0$ such that if $\ell_{j,a}$ are lines in \mathbb{R}^n which make an angle of at most δ with the x_j -axis, then, for every cube Q_S such that $S = \delta^{-M}$ for some integer M ,*

$$\int_{Q_S} \prod_{j=1}^n \left(\sum_a T_{j,a} \right)^{\frac{1}{n-1}} \leq S^\varepsilon \prod_{j=1}^n N_j^{\frac{1}{n-1}}.$$

This is enough to prove Theorem 4.9, since given $\varepsilon > 0$ we take δ as in the above lemma. Then for any S we take M to be the largest integer such that $S \geq \delta^{-M}$, and then we cover S by at most C cubes of sidelength δ^{-M} where a priori C depends on both S and δ . By proving Lemma 4.12 for all integers M , however, we have been able to remove the dependence on S , since we can simply cover the cube Q_S with one of side length $\delta^{-(M+1)}$ and then figure out how many cubes of side length δ^{-M} are needed to cover this cube. Consequently, the dependence of C is only on δ , which itself depends only on ε . Summing over these cubes yields

$$\int_{Q_S} \prod_{j=1}^n \left(\sum_a T_{j,a} \right)^{\frac{1}{n-1}} \lesssim_\varepsilon S^\varepsilon \prod_{j=1}^n N_j^{\frac{1}{n-1}}.$$

4.3. Multilinear versus linear decoupling. Inspired by the tractability of multilinear Keakeya and restriction over their linear counterparts, we now formulate a multilinear version of the decoupling problem. We first define the notion of *transversality* which has been useful in restriction theory. Roughly speaking, two smooth compact hypersurfaces S_1 and S_2 are transverse if the unit normals of S_1 are never oriented in the same direction as the unit normals of S_2 . This condition puts us in a good position to use almost orthogonality arguments. More generally we have:

Definition 4.13. A collection of $S_j \subset \mathbb{R}^n$ hypersurfaces are *transverse* if for any point $\omega \in S_j$, the normal vector $N_{S_j}(\omega)$ obeys

$$\text{Angle}(N_{S_j}(\omega), j^{\text{th}} \text{ coordinate axis}) \leq (10n)^{-1}.$$

Definition 4.14. We say that functions f_1, \dots, f_n on \mathbb{R}^n obey the *multilinear decoupling setup* (MDS)⁸ if

- for $i = 1, \dots, n$, $\text{supp } \hat{f}_i \subseteq \mathcal{N}_{1/R} S_i$;
- $S_i \subseteq \mathbb{R}^n$ are compact positively curved C^2 hypersurfaces;
- the surfaces S_j are transverse.

We define $\tilde{D}_{n,p}(R)$ to be the smallest constant so that whenever f_i obey (MDS),

$$\left\| \prod_{i=1}^n |f_i|^{1/n} \right\|_{L^p(B_R)} \leq \tilde{D}_{n,p}(R) \prod_{i=1}^n \|f_i\|_{L^{p,1/R}(B_R)}^{1/n}.$$

Bourgain and Demeter go on to prove the following relationship between linear decoupling and multilinear decoupling:

Theorem 4.15. *Suppose that in dimension $n - 1$, the decoupling constant $D_{n-1,p}(R) \lesssim R^\varepsilon$ for any $\varepsilon > 0$. Then for any $\varepsilon > 0$,*

$$D_{n,p}(R) \lesssim R^\varepsilon \tilde{D}_{n,p}(R).$$

Note that we always have $\tilde{D}_{n,p}(R) \leq D_{n,p}(R)$ for any n, p, R . Then, using induction on the dimension n , if the decoupling theorem holds in dimension $n - 1$ for the endpoint s , then we have essentially shown⁹ that

$$\tilde{D}_{n,p}(R) \sim D_{n,p}(R) \sim R^\gamma,$$

from (4.4) as well.

Bourgain and Demeter also use the *method of induction on scales*, similar to the argument in [16] and similar to the argument described above for multilinear Keakeya. The key requirement to apply induction on scales is to have a way of bounding a desired quantity of one scale by another. In multilinear Keakeya, this was Lemma 4.11. Bourgain and Demeter leverage the following multilinear restriction estimate which is a consequence of multilinear Keakeya:¹⁰

Lemma 4.16. *If f_1, \dots, f_n obey the multilinear decoupling setup, then*

$$\left\| \left(\prod_{i=1}^n |f_i| \right)^{1/n} \right\|_{L^{\frac{2n}{n-1}}(B_R)} \lesssim_\varepsilon R^{-1/2+\varepsilon} \left(\prod_{i=1}^n \|f_i\|_{L^2} \right)^{1/n}.$$

⁸This term is introduced in Guth's notes [43] on page 5, and the following definition is quoted directly from them.

⁹up to ε powers of R

¹⁰Lemma 4.16 is Theorem 6.1 in Bourgain and Demeter [14], and appears originally in [2].

Remark 4.17. By applying Bernstein's inequality at a sufficiently small scale r and then inductively working our way up using Lemma 4.16, we move through scales r^{2^m} until we reach R . This yields a weaker decoupling estimate for $2 \leq p \leq \frac{2n}{n-1}$, which appeared in [11].

4.4. Using curvature. The key component which allowed the improved estimates in Bourgain and Demeter's ℓ^2 decoupling paper [14] was the introduction of curvature.

Lemma 4.18. *Let f_1, \dots, f_n obey the multilinear decoupling setup with S_i compact positively curved transverse caps of P^{n-1} . Then for each $\frac{2n}{n-1} \leq p \leq \infty$, we have*

$$\left\| \left(\prod_{i=1}^n |f_i| \right)^{1/n} \right\|_{L^p(B_R)} \lesssim_\varepsilon R^{\frac{n-1}{4} - \frac{n^2+n}{2p(n-1)} + \varepsilon} \left(\prod_{i=1}^n \|f_i\|_{L^{\frac{p(n-1)}{n}, \frac{1}{R}}} \right)^{1/n},$$

and also

$$\left\| \left[\prod_{i=1}^n \left(\sum_{\theta \in T} |f_{i,\theta}|^2 \right)^{1/2} \right]^{1/n} \right\|_{L^p(B_R)} \lesssim_\varepsilon R^{\frac{-n}{(n-1)p} + \varepsilon} \left(\prod_{i=1}^n \|f_i\|_{L^{\frac{p(n-1)}{n}, \frac{1}{R}}} \right)^{1/n}.$$

Again, this estimate is perfectly suited to an induction on scales type argument. In order to prove this lemma, Bourgain and Demeter begin from Lemma 4.16 and show that we can reverse the previous Hölder inequality (4.3) if the function can be broken into a small number of “balanced” pieces, where each piece obeys a reverse Hölder inequality, i.e., if $1 \leq q, q_1, q_2 \leq \infty$ and $\frac{1}{q} = (1 - \alpha)\frac{1}{q_1} + \alpha\frac{1}{q_2}$, then

$$(4.8) \quad \|f\|_{L^{q_1, \delta}(\Omega)}^{1-\alpha} \|f\|_{L^{q_2, \delta}(\Omega)}^\alpha \lesssim \|f\|_{L^{q, \delta}(\Omega)}.$$

The proof of this fact relies on a wave packet decomposition of f and, essentially, interpolation with the $L^{p, \delta}$ -norms, which in turn facilitates the proof of Lemma 4.18. As mentioned at the beginning of this section, we will not be able to present these details in depth; however, the heart of the matter is that Lemma 4.18 allows Bourgain and Demeter to use induction on scales as well as parabolic rescaling (discussed in Section 4.1.3) to prove an estimate which yields the main result, Theorem 4.1, for the $n = 2$ and $p > p_c$ case. Higher dimensions then follow by induction.

To prove the endpoint $p = p_c = \frac{2(n+1)}{n-1}$, Bourgain and Demeter combine the ℓ^2 -decoupling result for $p > p_c$ with

$$\|f\|_{L^{p_c}(B_R)} \lesssim \|f\|_{L^p(B_R)} N^{\frac{n}{p_c} - \frac{n}{p}} \quad (\text{by Hölder's inequality})$$

and

$$\|f_\theta\|_{L^p(\mathbb{R}^n)} \lesssim N^{\frac{n+1}{2p} - \frac{n+1}{2p_c}} \|f_\theta\|_{L^p(\mathbb{R}^n)} \quad (\text{by Bernstein's inequality})$$

and then let $p \rightarrow p_c$.

4.5. Strichartz estimates for irrational tori. The idea is to use the discrete version of the ℓ^2 -decoupling theorem (Theorem 4.19 below), as was done in [11], and apply a change of variables to (3.8) which puts us in the perfect position to apply such a discrete estimate. We have previously discussed the connection between the Fourier restriction and Strichartz estimates in the case of the rational torus. More generally, in the case of irrational or rational tori, Bourgain and Demeter [14] prove the following discrete version of the restriction theorem which points to

the existence of stronger cancellations at the larger scale $R \gtrsim \delta^{-1}$ relative to the Tomas–Stein result (2.11), which is about oscillations at spatial scales $R \sim \delta^{-1/2}$.

Theorem 4.19 (Theorem 2.2 in [14]). *Let S be a compact C^2 hypersurface in \mathbb{R}^n with positive definite second fundamental form. Let $\Lambda \subset S$ be a $\delta^{1/2}$ -separated set, and let $R \gtrsim \delta^{-1}$. Then for each $\varepsilon > 0$,*

$$(4.9) \quad \left(\frac{1}{|B_R|} \int_{B_R} \left| \sum_{\xi \in \Lambda} a_\xi e(\xi \cdot x) \right|^p \right)^{1/p} \lesssim_\varepsilon \delta^{\frac{n+1}{2p} - \frac{n-1}{4} - \varepsilon} \|a_\xi\|_2$$

if $p \geq \frac{2(n+1)}{n-1}$.

In [11], Bourgain used weighted norms and a decomposition into caps to prove that Theorem 4.1 for a given p implies (4.9). A brief sketch of this argument is also given in [14].

The next step is to show that Theorem 3.5 follows from Theorem 4.19. To do this, we must prove that for $\phi \in L^2(\mathbb{T}^{n-1})$ with $\text{supp } \hat{\phi} \subset [-N, N]^{n-1}$, we have for each $\varepsilon > 0$, $p \geq \frac{2(n+1)}{n-1}$ and each interval $I \subset \mathbb{R}$ with $|I| \gtrsim 1$ that

$$(4.10) \quad \|e^{it\Delta_\Theta} \phi\|_{L^p(\mathbb{R}^{n-1} \times I)} \lesssim_\varepsilon N^{\frac{n-1}{2} - \frac{n+1}{p} + \varepsilon} |I|^{1/p} \|\phi\|_2,$$

where the implicit constant does not depend on I, N , or Θ , and where

$$(4.11) \quad e^{it\Delta_\Theta} \phi = \sum_{\mathbf{k} \in \mathbb{Z}^{n-1}} e^{2\pi i(\mathbf{k} \cdot x + t \sum_{j=1}^{n-1} \Theta_j k_j^2)} \hat{\phi}(\mathbf{k}).$$

Bourgain and Demeter show that by defining $\eta_i = \frac{\theta_i^{1/2} \xi_i}{4N}$, $a_\eta = \hat{\phi}(\mathbf{k})$, one can apply the change of variables

$$y_i = \frac{4N}{\theta_i^{1/2}} x_i \quad \text{for } 1 \leq i \leq n-1,$$

$$\tau = 16N^2 t,$$

and then use periodicity of the y_i variables to acquire the bound

$$\int_{\mathbb{T}^{n-1} \times I} |e^{it\Delta} \phi|^p \lesssim \frac{1}{N^{n+1} (N|I|)^{n-1}} \int_{B_{N^2|I|}} \left| \sum_{\eta} a_\eta e(\eta \cdot y + \tau \|\eta\|^2) \right|^p dy d\tau,$$

where $B_{N^2|I|}$ is some ball of radius on the order of $N^2|I|$. Finally, Bourgain and Demeter note that the points

$$(\eta_1, \dots, \eta_{n-1}, \|\eta\|^2)$$

are $\sim \frac{1}{N}$ separated on P^{n-1} , so Theorem 4.19 can be applied with $R \sim N^2|I|$.

5. THE NONLINEAR SCHRÖDINGER EQUATIONS: PROBABILISTIC METHODS

As we have seen in previous sections, the local well-posedness in the subcritical regime¹¹ is in place once the full range of Strichartz estimates (as stated in Theorem 3.5) are available. However, in certain critical and in supercritical regimes, there is no known deterministic local well-posedness theory in the periodic setting. Nevertheless, what has been within reach in recent years is the study of the local

¹¹This was defined on \mathbb{R}^d from the scaling symmetry of the equation. On tori it is still indicative of what to expect in terms of well-posedness, and so we transfer the same terminology into the periodic setting.

well-posedness of p-NLS (2.21) on the square/rational torus \mathbb{T}^d from a probabilistic point of view; that is, the almost sure well-posedness of p-NLS in the sense of probability. Such an approach was first used by Bourgain [7] in the mid 1990s to prove that the (Wick ordered) cubic nonlinear Schrödinger equation on \mathbb{T}^2 was almost sure locally well posed in $H^{-\epsilon}(\mathbb{T}^2)$. The latter was the first result in a supercritical regime, since $L^2(\mathbb{T}^2)$ is critical for this equation (see also [8, 9]). Recent work by A. Nahmod and G. Staffilani [57] established almost sure local well-posedness¹² in $H^{1-\alpha}(\mathbb{T}^3)$, $\alpha > 0$ small for the quintic nonlinear Schrödinger equation on \mathbb{T}^3 . This result is also in the supercritical regime since $H^1(\mathbb{T}^3)$ is critical for this equation.

It is worth noting that while for the quintic NLS on \mathbb{T}^3 , (deterministic) large data well-posedness at the critical $H^1(\mathbb{T}^3)$ regularity is known ([46] for local and [48] for global), to date there is no known (deterministic) large data well-posedness results available for the cubic NLS equation on \mathbb{T}^2 at critical $L^2(\mathbb{T}^2)$ regularity.

The results in [7] and [57] are for the square/rational torus \mathbb{T}^d and have not yet been established for irrational tori, despite the fact that we now know the Strichartz estimates on irrational tori by the work of Bourgain and Demeter. The reason for this lies in the fact that beyond the Strichartz estimates per se, precise integer lattice counting estimates used by Bourgain in [6] to *prove* Strichartz estimates on rational tori come into play in the works [7] and [57] in a crucial way, as we will illustrate below. Similar tools from analytic number theory are presently unknown in the context of irrational tori. The good news though is that, as proved by Bourgain and Demeter in [14], the decoupling theorem seems amenable to counting solutions of Diophantine inequalities (e.g., Theorems 2.18 and 2.19 in [14]). It would be interesting to explore further whether and how to use Bourgain and Demeter techniques to shed light into, or in lieu of, the necessary integer lattice counting estimates in the irrational setting.

We conclude by explaining some of the ideas behind the probabilistic approach on square tori below.

5.1. Random data: a nondeterministic approach. We start by giving an informal definition of almost sure well-posedness. Given μ a probability measure on the space of initial data X (e.g., $X = H^s$), we say that the Cauchy initial value problem (IVP) is almost sure locally well posed if there exists $Y \subset X$, with $\mu(Y) = 1$ and such that for any $\phi \in Y$ there exist $T > 0$ and a unique solution u to the IVP with data ϕ which is in $C([0, T], X)$ with data ϕ that is also stable in the appropriate topology.

The general idea is to consider the Cauchy initial data problem for rough but randomized initial data. To understand why randomization (of the initial data) helps, let us recall the following classical result going back to Rademacher, Kolmogorov, Paley, and Zygmund proving that random series on the torus enjoy better L^p -bounds than deterministic ones.¹³ For example, the consider the *Rademacher series*:

$$f(\tau) := \sum_{m=0}^{\infty} b_m r_m(\tau), \quad \tau \in [0, 1), \quad b_m \in \mathbb{C},$$

where

$$r_m(\tau) := \text{sign} \sin(2^{m+1} \pi \tau).$$

¹²In [57] the existence of infinite energy global in time solutions is also established.

¹³Akin to the Kintchine inequalities used to prove the Littlewood–Paley inequalities.

Note that if $b_m \in \ell^2$, the sum $f(\tau)$ converges a.e. The following is a classical result which can be found in Zygmund's book.

Theorem 5.1. *If $b_m \in \ell^2$, then the sum $f(\tau)$ belongs to $L^p([0, 1])$ for all $p \geq 2$. More precisely,*

$$\left(\int_0^1 |f|^p d\tau \right)^{1/p} \sim \|b_m\|_{\ell^2}.$$

The key point is that although randomized initial data live in the same (rough) space as the original (deterministic) data, their linear flow enjoys *almost surely* improved L^p -bounds. These bounds in turn yield improved nonlinear estimates *almost surely* in the analysis of the solution to the difference equation (obtained after subtracting from u the linear evolution of randomized data). More precisely, the general scheme is as follows. Consider the Cauchy initial value problem,

$$(5.1) \quad \begin{cases} iu_t + \Delta u = N(u) & x \in \mathbb{T}^d, t > 0, \\ u(0, x) = \phi(x), \end{cases}$$

and assume that $\phi \in X^s$. Then if we denote by $a_k := \widehat{\phi}(k)$, to solve (5.1) we proceed as follows:

- (1) Randomize ϕ : that is, consider $\phi^\omega := \sum_{k \in \mathbb{Z}^d} a_k g_k(\omega) e^{ix \cdot k}$, where $\{g_k(\omega)\}_k$ are i.i.d. standard (complex/real) centered (Gaussian) random variables on a probability space (Ω, \mathcal{F}, P) .
- (2) Let v^ω be the linear evolution with initial datum ϕ^ω .
- (3) Prove that v^ω satisfies “better estimates” than ϕ almost surely.
- (4) Show that $w := u - v^\omega$ solves a difference equation, and obtain for w a *deterministic* local well-posedness theory in $C([0, T]; X^{s'})$, $s' > s$. That is, prove that almost surely in ω the nonlinear part w is *smoother* than the linear part v^ω .

Remark 5.2. The difference equation that w solves is not an equation at the subcritical/smooth level but rather it is a *hybrid* equation with a nonlinearity containing a mixture of supercritical but random terms plus deterministic (smoother) ones.

For $\phi \in H^s$, $\phi^\omega(x)$ defines almost surely in ω a function in H^s but not in $H^{s'}$ for any $s' > s$. A representative example arises by considering $a_k = \frac{1}{|k|^\alpha}$. Then $\widehat{\phi^\omega}(k) = \frac{g_k(\omega)}{|k|^\alpha}$ gives rise almost surely in ω to a function in $H^{\alpha - \frac{d}{2} - \epsilon}$ but not in $H^{\alpha - \frac{d}{2}}$. Randomization does not improve regularity in terms of derivatives. The improvement is with respect to L^p -spaces *almost surely*. Another way to rephrase this and the classical Theorem 5.1 above is as follows: Let $\{g_m(\omega)\}$ be a sequence of complex i.i.d. zero mean Gaussian random variables on a probability space $(\Omega, \mathcal{A}, \mathbb{P})$ and $(c_m) \in \ell^2$. Define

$$F(\omega) := \sum_m c_m g_m(\omega).$$

Then, there exists $C > 0$ such that for every $\lambda > 0$, we have

$$\mathbb{P}(\{\omega : |F(\omega)| > \lambda\}) \leq \exp\left(\frac{-C\lambda^2}{\|F\|_{L^2(\Omega)}^2}\right).$$

As a consequence there exists $C > 0$ such that for every $q \geq 2$ and every $(c_m)_m \in \ell^2$,

$$\left\| \sum_m c_m g_m(\omega) \right\|_{L^q(\Omega)} \leq C \sqrt[q]{\sum_m c_m^2}^{\frac{1}{2}}.$$

More generally one uses the following result where k represents the number of random terms in the multilinear estimate at hand.

Proposition 5.3 (Large Deviation-type). *Let $d \geq 1$ and $c(m_1, \dots, m_k) \in \mathbb{C}$. Let $\{g_m\}_{1 \leq m \leq d} \in \mathcal{N}_{\mathbb{C}}(0, 1)$ be complex centered L^2 -normalized independent Gaussians. For $k \geq 1$ denote by $A(k, d) := \{(m_1, \dots, m_k) \in \{1, \dots, d\}^k, m_1 \leq \dots \leq m_k\}$ and*

$$F_k(\omega) = \sum_{A(k, d)} c(m_1, \dots, m_k) g_{m_1}(\omega) \cdots g_{m_k}(\omega).$$

Then for $p \geq 2$,

$$\|F_k\|_{L^p(\Omega)} \lesssim \sqrt[k]{k+1} (p-1)^{\frac{k}{2}} \|F_k\|_{L^2(\Omega)}.$$

As a consequence from Chebyshev's inequality for every $\lambda > 0$,

$$\mathbb{P}(\{\omega : |F_k(\omega)| > \lambda\}) \leq \exp\left(\frac{-C \lambda^{\frac{2}{k}}}{\|F_k\|_{L^2(\Omega)}^{\frac{2}{k}}}\right).$$

Remark 5.4. This result follows from the hyper-contractivity property of the Ornstein–Uhlenbeck semigroup by writing $G_n = H_n + iL_n$, where $\{H_1, \dots, H_d, L_1, \dots, L_d\} \in \mathcal{N}_{\mathbb{R}}(0, 1)$ are real centered independent Gaussian random variables with the same variance (cf. [68, 71]).

The key observation then is that for a given $\delta > 0$, the large deviation result above with, say,

$$\lambda = \delta^{-\frac{k}{2}} \|F_k\|_{L^2(\Omega)}$$

will allow us to replace $|F_k(\omega)|^2$ by $\|F_k\|_{L^2(\Omega)}^2$ on a set $\Omega_\delta \subset \Omega$ with $\mathbb{P}(\Omega_\delta^c) < e^{-\frac{1}{\delta}}$. Thus we use the independence and normalization of the random variables to reduce matters to geometric considerations and integer lattice counting, as we will illustrate below.

5.2. Almost sure local well-posedness results for the periodic NLS. Bourgain's almost sure local well-posedness result for the (Wick ordered) cubic NLS on the square/rational torus \mathbb{T}^2 reads as follows:

Theorem 5.5 (Bourgain [7]).

$$\begin{cases} iu_t + \Delta u = |u|^2 u - (\int |u|^2 dx) u, \\ u(0, x) = \phi(x), \quad x \in \mathbb{T}^2, \end{cases}$$

is almost sure locally well-posed below L^2 , that is for supercritical data $\phi \in H^{-\varepsilon}(\mathbb{T}^2)$.

Remark 5.6. The typical data considered is $\phi(x) = \sum_{n \in \mathbb{Z}^2} \frac{1}{\langle n \rangle} e^{in \cdot x} \in H^{-\varepsilon}(\mathbb{T}^2)$ and $\phi^\omega(x) = \sum \frac{g_n(\omega)}{\langle n \rangle} e^{in \cdot x} \in H^{-\varepsilon}(\mathbb{T}^2)$ defining almost surely in ω a function in $H^{-\varepsilon}(\mathbb{T}^2)$.

In [57] we considered the energy-critical quintic nonlinear Schrödinger equation on \mathbb{T}^3 ,

$$(5.2) \quad \begin{cases} iu_t + \Delta u = \lambda u|u|^4 & x \in \mathbb{T}^3, \\ u(0, x) = \phi(x) & \in H^\gamma(\mathbb{T}^3), \end{cases}$$

and established almost sure local well-posedness for random data in $H^\gamma(\mathbb{T}^3)$, $\gamma < 1$; that is, in the supercritical regime relative to scaling. The problem we considered is the analogue of Bourgain's Theorem 5.5 mentioned above. In our problem we consider data $\phi \in H^{1-\alpha-\varepsilon}(\mathbb{T}^3)$ for any $\varepsilon > 0$ of the form

$$\phi(x) = \sum_{n \in \mathbb{Z}^3} \frac{1}{\langle n \rangle^{\frac{5}{2}-\alpha}} e^{in \cdot x},$$

whose randomization is

$$(5.3) \quad \phi^\omega(x) = \sum_{n \in \mathbb{Z}^3} \frac{g_n(\omega)}{\langle n \rangle^{\frac{5}{2}-\alpha}} e^{in \cdot x},$$

where $(g_n(\omega))_n$ is a sequence of complex i.i.d centered Gaussian random variables on a probability space $(\Omega, \mathcal{A}, \mathbb{P})$. Let us denote by $\mathcal{X}^s([0, \delta])$ the solution space for the nonlinear part of the solution. Our result reads as follows:

Theorem 5.7 (Nahmod and Staffilani [57]). *Let $0 < \alpha < \frac{1}{12}$, let $s = s(\alpha) > 1$, and let ϕ^ω be as in (5.3). Then there exists $0 < \delta_0 \ll 1$ and $r = r(s, \alpha) > 0$ such that for any $\delta < \delta_0$, there exists $\Omega_\delta \in \mathcal{A}$ with*

$$\mathbb{P}(\Omega_\delta^c) < e^{-\frac{1}{\delta^r}},$$

and for each $\omega \in \Omega_\delta$ there exists a unique solution u of the quintic NLS (5.2) in the space

$$S(t)\phi^\omega + \mathcal{X}^s([0, \delta])$$

with initial condition ϕ^ω .

5.2.1. The difference equation. Heart of the matter. The approach to proving Theorems 5.5 and 5.7 is based on a study of the difference equation. More precisely, assume u solves our IVP, then we define $w := u - S(t)\phi^\omega$, where $S(t)\phi^\omega$ is the linear evolution of the randomized initial profile ϕ^ω . The aim is to study the IVP for w which solves a difference equation with nonlinearity,

$$(5.4) \quad \tilde{N}(w) := |w + S(t)\phi^\omega|^4(w + S(t)\phi^\omega),$$

and prove that w belongs to H^s for some $s > 1$. The heart of the matter then lies in proving the multilinear estimates to control $\tilde{N}(w)$ so we can set up a contraction argument and obtain well-posedness. The randomness coming from $(g_n(\omega))$ will allow us to say that in a certain space the nonlinearity increases its regularity so that it can hold a bit more than one derivative.

For the quintic NLS equation (5.2) however, multilinear estimates for $\tilde{N}(w)$ can be obtained only after having removed certain resonant terms involved in the nonlinear part of the equation. In Bourgain's case [7] the nonlinearity is *cubic* in two dimensions, and a Wick ordering of the Hamiltonian takes care of bad resonant terms. In our case the nonlinearity is quintic in three dimensions, and Wick ordering is not sufficient to remove the bad *resonant* terms. Instead, a suitable gauge transformation is required [57].

The Strichartz estimates on tori enter crucially in estimating the nonlinear interactions in (5.4) involving solely w . However, when faced with mixed *hybrid* terms or nonlinear interactions involving solely $S(t)\phi^\omega$, the Strichartz estimates are not enough, and we must also rely on probabilistic arguments and elementary analytic number theory as in Bourgain's Theorem 5.5. It is here that we must restrict our attention to rational tori as explained in the introduction to Section 5 above.

We conclude by illustrating how randomization, Proposition 5.3, and the ideas in Section 5 are used in a prototypical example that arises in the course of the proof of the nonlinear estimates (5.4) for the difference equation. To simplify matters we consider normalized random data of the form

$$(5.5) \quad \phi^\omega(x) = \sum_{n \in \mathbb{Z}^3} \frac{g_n(\omega)}{\langle n \rangle^{\frac{3}{2}}} e^{ix \cdot n}.$$

As an intermediate step in the course of the proof of the nonlinear estimates for the term (5.4), we must estimate in L^2 cubic nonlinear interactions in (5.4) involving only the free evolution of such random data. Let us denote by R_k , $k = 1, 2, 3$, the linear evolution of normalized random functions (5.5) which are localized at frequency $|n_k| \sim N_k$, with N_k a dyadic number. A particular case involves estimating $\|P_C \bar{R}_1 \bar{R}_2 R_3\|_{L^2}$, where we assume that the frequencies n_k , $k = 1, 2, 3$, are all different from each other, that $N_1 \geq N_2 \geq N_3$, and that P_C is a Fourier projection onto a cube C of sidelength N_2 .

By Plancherel,

$$\begin{aligned} \|P_C \bar{R}_1 \bar{R}_2 R_3\|_{L^2}^2 &= \sum_{m, n \in \mathbb{Z}^3} \left| \sum_{\substack{n=n_1+n_2-n_3, n_i \neq n_j \\ m=|n_1|^2+|n_2|^2-|n_3|^2}} \chi_C(n_1) \frac{\bar{g}_{n_1}(\omega)}{|n_1|^{\frac{3}{2}}} \frac{\bar{g}_{n_2}(\omega)}{|n_2|^{\frac{3}{2}}} \frac{g_{n_3}(\omega)}{|n_3|^{\frac{3}{2}}} \right|^2 \\ &=: \sum_{m, n \in \mathbb{Z}^3} |F_3(\omega)|^2. \end{aligned}$$

If we naively use Cauchy–Schwarz, we would obtain an estimate where we pick up the cardinality of the set

$$S_{(n,m)} := \{(n_1, n_2, n_3) : n = n_1 + n_2 - n_3, n_i \neq n_j, m = |n_1|^2 + |n_2|^2 - |n_3|^2\},$$

which unfortunately translates into a loss of derivatives.

Instead, we use Proposition 5.3 with

$$\lambda = \delta^{-\frac{3}{2}} \|F_3(\omega)\|_{L^2(\Omega)}$$

so that on a set Ω_δ with $\mathbb{P}(\Omega_\delta^c) < e^{-\frac{1}{\delta}}$ we can replace $|F_3(\omega)|^2$ by $\|F_3(\omega)\|_{L^2(\Omega)}^2$. Then using the independence and normalization of $g_n(\omega)$, we estimate

$$\|F_3(\omega)\|_{L^2(\Omega)}^2 \lesssim \delta^{-\frac{3}{2}} \sum_{S(n,m)} \chi_C(n_1) \prod_{i=1}^3 \frac{1}{|n_i|^3}.$$

All in all, we then have that

$$\|P_C \bar{R}_1 \bar{R}_2 R_3\|_{L^2}^2 \lesssim N_1 N_2 \sum_n |F_3(\omega)|^2 \lesssim \delta^{-\frac{3}{2}} N_1 N_2 N_1^{-3} N_2^{-3} N_3^{-3} \sup_m \#S(m),$$

where

$$S_m := \{(n, n_1, n_2, n_3) / n = n_1 + n_2 - n_3; m = |n_1|^2 + |n_2|^2 - |n_3|^2, n_1 \in C\},$$

and we have used that the variation for the time frequency m is of the order $N_1 N_2$. Since $\#S_m \lesssim N_3^3 N_2^3 N_1$, we finally obtain the bound

$$\|P_C \bar{R}_1 \bar{R}_2 R_3\|_{L^2}^2 \lesssim \delta^{-\frac{3}{2}} N_1^{-1} N_2,$$

which suffices for our purposes in Theorem 5.7.

ABOUT THE AUTHOR

Andrea R. Nahmod is professor of mathematics at the University of Massachusetts, Amherst. Her work belongs to nonlinear Fourier and harmonic analysis and to nonlinear partial differential equations, integrating into it tools from geometry and probability.

ACKNOWLEDGMENTS

The author would like to thank Larry Guth for making the insightful notes from his course [43] available online. We borrow heavily from them in Section 4 on ℓ^2 decouplings. The author also thanks her Ph.D. student Michael Boratko for stimulating discussions on multilinear Kakeya, the work of Bourgain and Demeter, as well as for his great help in the preparation of Section 4.

REFERENCES

- [1] Mark J. Ablowitz, *Nonlinear dispersive waves*, Cambridge Texts in Applied Mathematics, Cambridge University Press, New York, 2011. Asymptotic analysis and solitons. MR2848561
- [2] Jonathan Bennett, Anthony Carbery, and Terence Tao, *On the multilinear restriction and Kakeya conjectures*, Acta Math. **196** (2006), no. 2, 261–302, DOI 10.1007/s11511-006-0006-4. MR2275834 (2007h:42019)
- [3] S. N. Bose, *Plancks Gesetz und Lichtquantenhypothese*, Zeitschrift für Physik, **26** (1924), 178.
- [4] 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, DOI 10.1007/BF01896020. MR1209299 (95d:35160a)
- [5] J. Bourgain, *Eigenfunction bounds for the Laplacian on the n -torus*, Internat. Math. Res. Notices **3** (1993), 61–66, DOI 10.1155/S1073792893000066. MR1208826 (94f:58127)
- [6] 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, DOI 10.1007/BF01896020. MR1209299 (95d:35160a)
- [7] Jean Bourgain, *Invariant measures for the 2D-defocusing nonlinear Schrödinger equation*, Comm. Math. Phys. **176** (1996), no. 2, 421–445. MR1374420 (96m:35292)
- [8] Jean Bourgain, *Nonlinear Schrödinger equations*, Hyperbolic equations and frequency interactions (Park City, UT, 1995), IAS/Park City Math. Ser., vol. 5, Amer. Math. Soc., Providence, RI, 1999, pp. 3–157. MR1662829 (2000c:35216)
- [9] J. Bourgain, *Global solutions of nonlinear Schrödinger equations*, American Mathematical Society Colloquium Publications, vol. 46, American Mathematical Society, Providence, RI, 1999. MR1691575 (2000h:35147)
- [10] J. Bourgain, *On Strichartz’s inequalities and the nonlinear Schrödinger equation on irrational tori*, Mathematical aspects of nonlinear dispersive equations, Ann. of Math. Stud., vol. 163, Princeton Univ. Press, Princeton, NJ, 2007, pp. 1–20. MR2331676 (2008j:35165)
- [11] J. Bourgain, *Moment inequalities for trigonometric polynomials with spectrum in curved hypersurfaces*, Israel J. Math. **193** (2013), no. 1, 441–458, DOI 10.1007/s11856-012-0077-1. MR3038558
- [12] Jean Bourgain and Ciprian Demeter, *Improved estimates for the discrete Fourier restriction to the higher dimensional sphere*, Illinois J. Math. **57** (2013), no. 1, 213–227. MR3224568
- [13] J. Bourgain and C. Demeter, *New bounds for the discrete Fourier restriction to the sphere in four and five dimensions*, to appear in Int. Math. Res. Notices (2014). Preprint at <http://arxiv.org/pdf/1310.5244.pdf>

- [14] J. Bourgain and C. Demeter, *The proof of the ℓ^2 decoupling conjecture*. Annals of Math. (2) **182** (2015), no. 1, 351–389.
- [15] J. Bourgain, C. Demeter, *Decouplings for curves and hypersurfaces with nonzero Gaussian curvature*, Preprint at <http://arxiv.org/abs/1409.1634.pdf> (2014)
- [16] Jean Bourgain and Larry Guth, *Bounds on oscillatory integral operators based on multilinear estimates*, Geom. Funct. Anal. **21** (2011), no. 6, 1239–1295, DOI 10.1007/s00039-011-0140-9. MR2860188 (2012k:42018)
- [17] F. Catoire and W.-M. Wang, *Bounds on Sobolev norms for the defocusing nonlinear Schrödinger equation on general flat tori*, Commun. Pure Appl. Anal. **9** (2010), no. 2, 483–491, DOI 10.3934/cpaa.2010.9.483. MR2600446 (2011h:35270)
- [18] Thierry Cazenave, *Semilinear Schrödinger equations*, Courant Lecture Notes in Mathematics, vol. 10, New York University, Courant Institute of Mathematical Sciences, New York; American Mathematical Society, Providence, RI, 2003. MR2002047 (2004j:35266)
- [19] T. Chen, C. Hainzl, N. Pavlović, R. Seiringer, *Unconditional uniqueness for the cubic Gross-Pitaevskii hierarchy via quantum de Finetti*, preprint (2013), <http://arxiv.org/abs/1307.3168.pdf>.
- [20] Thomas Chen and Nataša Pavlović, *Recent results on the Cauchy problem for focusing and defocusing Gross-Pitaevskii hierarchies*, Math. Model. Nat. Phenom. **5** (2010), no. 4, 54–72, DOI 10.1051/mmnp/20105403. MR2662450 (2011c:35530)
- [21] Thomas Chen and Nataša Pavlović, *The quintic NLS as the mean field limit of a boson gas with three-body interactions*, J. Funct. Anal. **260** (2011), no. 4, 959–997, DOI 10.1016/j.jfa.2010.11.003. MR2747009 (2012b:82066)
- [22] Thomas Chen and Nataša Pavlović, *Derivation of the cubic NLS and Gross-Pitaevskii hierarchy from manybody dynamics in $d = 3$ based on spacetime norms*, Ann. Henri Poincaré **15** (2014), no. 3, 543–588, DOI 10.1007/s00023-013-0248-6. MR3165917
- [23] J. Colliander, M. Keel, G. Staffilani, H. Takaoka, and T. Tao, *Transfer of energy to high frequencies in the cubic defocusing nonlinear Schrödinger equation*, Invent. Math. **181** (2010), no. 1, 39–113, DOI 10.1007/s00222-010-0242-2. MR2651381 (2011f:35320)
- [24] M. H. Anderson, J. R. Ensher, M. R. Matthews, C. E. Wieman, and E. A. Cornell, *Observations of Bose–Einstein condensation in a dilute atomic vapor*, Science **269** (1995), 198–201.
- [25] W. Craig, C. Sulem, and P.-L. Sulem, *Nonlinear modulation of gravity waves: a rigorous approach*, Nonlinearity **5** (1992), no. 2, 497–522. MR1158383 (93k:76012)
- [26] Daniela De Silva, Nataša Pavlović, Gigliola Staffilani, and Nikolaos Tzirakis, *Global well-posedness for a periodic nonlinear Schrödinger equation in 1D and 2D*, Discrete Contin. Dyn. Syst. **19** (2007), no. 1, 37–65, DOI 10.3934/dcds.2007.19.37. MR2318273 (2008g:35191)
- [27] C. Demeter, *Incidence theory and discrete restriction estimates*, Preprint at <http://arxiv.org/pdf/1401.1873v1.pdf> (2014).
- [28] S. Demirbas, *Local Well-posedness for 2-D Schrödinger Equation on Irrational Tori and Bounds on Sobolev Norms*, preprint (2013), <http://arxiv.org/abs/1307.0051>.
- [29] A. Einstein, *Quantentheorie des einatomigen idealen Gases*, Sitzungsberichte der Preussischen Akademie der Wissenschaften 1: 3. (1925).
- [30] Alexander Elgart, László Erdős, Benjamin Schlein, and Horng-Tzer Yau, *Gross-Pitaevskii equation as the mean field limit of weakly coupled bosons*, Arch. Ration. Mech. Anal. **179** (2006), no. 2, 265–283, DOI 10.1007/s00205-005-0388-z. MR2209131 (2007b:81310)
- [31] László Erdős, Benjamin Schlein, and Horng-Tzer Yau, *Derivation of the Gross-Pitaevskii hierarchy for the dynamics of Bose–Einstein condensate*, Comm. Pure Appl. Math. **59** (2006), no. 12, 1659–1741, DOI 10.1002/cpa.20123. MR2257859 (2007k:82070)
- [32] László Erdős, Benjamin Schlein, and Horng-Tzer Yau, *Derivation of the cubic non-linear Schrödinger equation from quantum dynamics of many-body systems*, Invent. Math. **167** (2007), no. 3, 515–614, DOI 10.1007/s00222-006-0022-1. MR2276262 (2007m:81258)
- [33] L. Erdős, B. Schlein, H.-T. Yau, *Rigorous derivation of the Gross-Pitaevskii equation*, Phys. Rev. Lett. **98** (2007), no.4, 040404.
- [34] László Erdős, Benjamin Schlein, and Horng-Tzer Yau, *Rigorous derivation of the Gross-Pitaevskii equation with a large interaction potential*, J. Amer. Math. Soc. **22** (2009), no. 4, 1099–1156, DOI 10.1090/S0894-0347-09-00635-3. MR2525781 (2010g:82041)
- [35] László Erdős, Benjamin Schlein, and Horng-Tzer Yau, *Derivation of the Gross-Pitaevskii equation for the dynamics of Bose–Einstein condensate*, Ann. of Math. (2) **172** (2010), no. 1, 291–370, DOI 10.4007/annals.2010.172.291. MR2680421 (2011g:82068)

- [36] László Erdős and Horng-Tzer Yau, *Derivation of the nonlinear Schrödinger equation from a many body Coulomb system*, Adv. Theor. Math. Phys. **5** (2001), no. 6, 1169–1205. MR1926667 (2004c:82075)
- [37] Gustavo Garrigós and Andreas Seeger, *A mixed norm variant of Wolff’s inequality for paraboloids*, Harmonic analysis and partial differential equations, Contemp. Math., vol. 505, Amer. Math. Soc., Providence, RI, 2010, pp. 179–197, DOI 10.1090/conm/505/09923. MR2664568 (2011f:42012)
- [38] Philip Gressman, Vedran Sohinger, and Gigliola Staffilani, *On the uniqueness of solutions to the periodic 3D Gross-Pitaevskii hierarchy*, J. Funct. Anal. **266** (2014), no. 7, 4705–4764, DOI 10.1016/j.jfa.2014.02.006. MR3170216
- [39] E. P. Gross, *Structure of a quantized vortex in boson systems* (English, with Italian summary), Nuovo Cimento (10) **20** (1961), 454–477. MR0128907 (23 #B1944)
- [40] M. Guardia and V. Kaloshin, *Growth of Sobolev norms in the cubic defocusing nonlinear Schrödinger equation*, J. Eur. Math. Soc. (JEMS) **17** (2015), no. 1, 71–149, DOI 10.4171/JEMS/499. MR3312404
- [41] Zihua Guo, Tadahiro Oh, and Yuzhao Wang, *Strichartz estimates for Schrödinger equations on irrational tori*, Proc. Lond. Math. Soc. (3) **109** (2014), no. 4, 975–1013, DOI 10.1112/plms/pdu025. MR3273490
- [42] Larry Guth, *A short proof of the multilinear Kakeya inequality*, Math. Proc. Cambridge Philos. Soc. **158** (2015), no. 1, 147–153, DOI 10.1017/S0305004114000589. MR3300318
- [43] L. Guth, Mini-course Notes on *Decoupling and multilinear estimates in harmonic analysis*, MIT (2014) <http://math.mit.edu/~lguth/decouplingseminar.html>.
- [44] Larry Guth, *The endpoint case of the Bennett-Carbery-Tao multilinear Kakeya conjecture*, Acta Math. **205** (2010), no. 2, 263–286, DOI 10.1007/s11511-010-0055-6. MR2746348 (2012c:42027)
- [45] Z. Hani, B. Pausader, N. Tzvetkov and N. Visciglia, *Modified scattering for the cubic Schrödinger equation on product spaces and applications*, Preprint at <http://arxiv.org/pdf/1311.2275v3.pdf> (2013).
- [46] Sebastian Herr, Daniel Tataru, and Nikolay Tzvetkov, *Global well-posedness of the energy-critical nonlinear Schrödinger equation with small initial data in $H^1(\mathbb{T}^3)$* , Duke Math. J. **159** (2011), no. 2, 329–349, DOI 10.1215/00127094-1415889. MR2824485 (2012j:35392)
- [47] Yi Hu and Xiaochun Li, *Discrete Fourier restriction associated with Schrödinger equations*, Rev. Mat. Iberoam. **30** (2014), no. 4, 1281–1300, DOI 10.4171/RMI/815. MR3293434
- [48] Alexandru D. Ionescu and Benoit Pausader, *The energy-critical defocusing NLS on \mathbb{T}^3* , Duke Math. J. **161** (2012), no. 8, 1581–1612, DOI 10.1215/00127094-1593335. MR2931275
- [49] K. B. Davis, M. O. Mewes, M. R. Andrews, N. J. van Druten, D. S. Durfee, D. M. Kurn, and W. Ketterle, *Bose–Einstein Condensation in a Gas of Sodium Atoms*, Phys. Rev. Let. **75** (1995), no. 22, 3969–3973.
- [50] W. Ketterle, *When Atoms behave like Waves: Bose–Einstein Condensation and the Atom Laser*, Nobel Lecture (2001). Available at http://www.nobelprize.org/nobel_prizes/physics/laureates/2001/ketterle-lecture.pdf
- [51] W. Ketterle, *When Atoms behave like Waves: Bose–Einstein Condensation and the Atom Laser*, Video of Nobel Lecture (2001). Available at http://www.nobelprize.org/nobel_prizes/physics/laureates/2001/ketterle-lecture.html
- [52] R. Killip, M. Viřan, *Scale invariant Strichartz estimates on tori and applications*, preprint (2014), <http://arxiv.org/abs/1409.3603>.
- [53] Sergiu Klainerman and Matei Machedon, *On the uniqueness of solutions to the Gross-Pitaevskii hierarchy*, Comm. Math. Phys. **279** (2008), no. 1, 169–185, DOI 10.1007/s00220-008-0426-4. MR2377632 (2009a:35236)
- [54] L. H. Loomis and H. Whitney, *An inequality related to the isoperimetric inequality*, Bull. Amer. Math. Soc. **55** (1949), 961–962. MR0031538 (11,166d)
- [55] Jerome Moloney and Alan Newell, *Nonlinear optics*, Westview Press. Advanced Book Program, Boulder, CO, 2004. MR2018083 (2004g:78002)
- [56] Camil Muscalu and Wilhelm Schlag, *Classical and multilinear harmonic analysis. Vol. I*, Cambridge Studies in Advanced Mathematics, vol. 137, Cambridge University Press, Cambridge, 2013. MR3052498

- [57] Andrea Nahmod and Gigliola Staffilani, *Almost sure well-posedness for the periodic 3D quintic nonlinear Schrödinger equation below the energy space*, J. Eur. Math. Soc. (JEMS) **17** (2015), no. 7, 1687–1759, DOI 10.4171/JEMS/543. MR3361727
- [58] L. Pitaevskii, *Vortex lines in an imperfect Bose gas*, Sov. Phys. JETP **13** (1961), 451–454.
- [59] V. Sohinger, *A rigorous derivation of the defocusing cubic nonlinear Schrödinger equation on \mathbb{T}^3 from the dynamics of many-body quantum systems*, preprint (2014), <http://arxiv.org/abs/1405.3003>, to appear in Ann. Inst. H. Poincaré (C), Analyse Non Linéaire.
- [60] Elias M. Stein, *Harmonic analysis: real-variable methods, orthogonality, and oscillatory integrals*, Princeton Mathematical Series, vol. 43, Princeton University Press, Princeton, NJ, 1993. With the assistance of Timothy S. Murphy; Monographs in Harmonic Analysis, III. MR1232192 (95c:42002)
- [61] Catherine Sulem and Pierre-Louis Sulem, *The nonlinear Schrödinger equation. Self-focusing and wave collapse*, Applied Mathematical Sciences, vol. 139, Springer-Verlag, New York, 1999. MR1696311 (2000f:35139)
- [62] Robert S. Strichartz, *Restrictions of Fourier transforms to quadratic surfaces and decay of solutions of wave equations*, Duke Math. J. **44** (1977), no. 3, 705–714. MR0512086 (58 #23577)
- [63] Nils Strunk, *Strichartz estimates for Schrödinger equations on irrational tori in two and three dimensions*, J. Evol. Equ. **14** (2014), no. 4-5, 829–839, DOI 10.1007/s00028-014-0240-8. MR3276862
- [64] H. Takaoka and N. Tzvetkov, *On 2D nonlinear Schrödinger equations with data on $\mathbb{R} \times \mathbb{T}$* , J. Funct. Anal. **182** (2001), no. 2, 427–442, DOI 10.1006/jfan.2000.3732. MR1828800 (2002c:35251)
- [65] T. Tao, *Recent Progress on the Restriction Conjecture*, IAS/Park City Mathematics Series, 2003. Preprint at <http://arxiv.org/pdf/math/0311181.pdf>.
- [66] Terence Tao, *Some recent progress on the restriction conjecture*, Fourier analysis and convexity, Appl. Numer. Harmon. Anal., Birkhäuser Boston, Boston, MA, 2004, pp. 217–243. MR2087245 (2005i:42015)
- [67] Terence Tao, *Nonlinear dispersive equations. Local and global analysis*, CBMS Regional Conference Series in Mathematics, vol. 106, Published for the Conference Board of the Mathematical Sciences, Washington, DC; by the American Mathematical Society, Providence, RI, 2006. MR2233925 (2008i:35211)
- [68] Laurent Thomann and Nikolay Tzvetkov, *Gibbs measure for the periodic derivative nonlinear Schrödinger equation*, Nonlinearity **23** (2010), no. 11, 2771–2791, DOI 10.1088/0951-7715/23/11/003. MR2727169 (2011j:35228)
- [69] Nathan Totz and Sijue Wu, *A rigorous justification of the modulation approximation to the 2D full water wave problem*, Comm. Math. Phys. **310** (2012), no. 3, 817–883, DOI 10.1007/s00220-012-1422-2. MR2891875
- [70] Nathan Totz, *A justification of the modulation approximation to the 3D full water wave problem*, Comm. Math. Phys. **335** (2015), no. 1, 369–443, DOI 10.1007/s00220-014-2259-7. MR3314508
- [71] N. Tzvetkov, *Construction of a Gibbs measure associated to the periodic Benjamin-Ono equation*, Probab. Theory Related Fields **146** (2010), no. 3-4, 481–514, DOI 10.1007/s00440-008-0197-z. MR2574736 (2011b:35471)
- [72] Thomas H. Wolff, *Lectures on harmonic analysis*, University Lecture Series, vol. 29, American Mathematical Society, Providence, RI, 2003. With a foreword by Charles Fefferman and preface by Izabella Laba; Edited by Laba and Carol Shubin. MR2003254 (2004e:42002)

DEPARTMENT OF MATHEMATICS AND STATISTICS, UNIVERSITY OF MASSACHUSETTS AMHERST,
 710 NORTH PLEASANT STREET, LEDERLE GRT, AMHERST, MASSACHUSETTS 01003
E-mail address: nahmod@math.umass.edu