Nonlinear dispersive equations with random initial data

by

Dana Sydney Mendelson

B.Sc., McGill University (2010)

Submitted to the Department of Mathematics
in partial fulfillment of the requirements for the degree of

Doctor of Philosophy

at the

MASSACHUSETTS INSTITUTE OF TECHNOLOGY

June 2015

© Dana Sydney Mendelson, MMXV. All rights reserved.

The author hereby grants to MIT permission to reproduce and to distribute publicly
paper and electronic copies of this thesis document in whole or in part in any medium
now known or hereafter created.

Signature redacted

Author

Department of Mathematics

May 1, 2015

Signature réédacté

Certified by.

Gigliola Staffilani
Abby Rockefeller Mauzé Professor of Mathematics
Thesis Supervisor

Signature redacted

Accepted by...

Alexei Borodin
Chairman, Department Committee on Graduate Theses
Nonlinear dispersive equations with random initial data

by

Dana Sydney Mendelson

Submitted to the Department of Mathematics
on May 1, 2015, in partial fulfillment of the
requirements for the degree of
Doctor of Philosophy

Abstract

In the first part of this thesis we consider the defocusing nonlinear wave equation of power-type on $\mathbb{R}^3$. We establish an almost sure global existence result with respect to a suitable randomization of the initial data. In particular, this provides examples of initial data of supercritical regularity which lead to global solutions. The proof is based upon Bourgain’s high-low frequency decomposition and improved averaging effects for the free evolution of the randomized initial data.

In the second part of this thesis, we consider the periodic defocusing cubic nonlinear Klein-Gordon equation in three dimensions in the symplectic phase space $H^{\frac{1}{2}}(T^3) \times H^{-\frac{1}{2}}(T^3)$. This space is at the critical regularity for this equation, and in this setting there is no global well-posedness nor any uniform control on the local time of existence for arbitrary initial data. We prove several non-squeezing results: a local in time result and a conditional result which states that uniform bounds on the Strichartz norms of solutions for initial data in bounded subsets of the phase space implies global-in-time non-squeezing. As a consequence of the conditional result, we conclude non-squeezing for certain subsets of the phase space and, in particular, we obtain deterministic small data non-squeezing for long times.

To prove non-squeezing, we employ a combination of probabilistic and deterministic techniques. Analogously to the work of Burq and Tzvetkov, we first define a set of full measure with respect to a suitable randomization of the initial data on which the flow of this equation is globally defined. The proofs then rely on several approximation results for the flow, one which uses probabilistic estimates for the nonlinear component of the flow map and deterministic stability theory, and another which uses multilinear estimates in adapted function spaces built on $U^p$ and $V^p$ spaces. We prove non-squeezing using a combination of these approximation results, Gromov’s finite dimensional non-squeezing theorem and the infinite dimensional symplectic capacity defined by Kuksin.

Thesis Supervisor: Gigliola Staffilani
Title: Abby Rockefeller Mauzé Professor of Mathematics
Acknowledgments

I would like to thank my advisor Gigliola Staffilani, who has made this thesis possible and given me so much of her time over these past five years. She has taught me an incredible amount about mathematics and being a mathematician, all with an immense amount of kindness and patience.

I owe much gratitude to Andrea Nahmod for many mathematical conversations and plentiful sound advice, and to Michael Eichmair for support and guidance during some stressful times. I would also like to thank David Jerison and Jared Speck for serving on my thesis committee, and Jonas Lührmann for our extremely enjoyable collaboration (which has, in particular, given rise to the contents of Chapter 2 of this thesis).

My time at MIT would have been nowhere near as enjoyable without the friendship of my fellow classmates, particularly Michael Andrews, Nate Bottman, Saul Glasman, Jiayong Li and Alex Moll, and I am grateful to them for making these past years at the math department as colorful as they have been.

I would like to thank my family and friends for their support, including my grandparents for having set the best possible example for us all. My parents, Beverley and Morton, have given me so much throughout my life. I would like to specifically thank them for the confidence they have always had in me, and for believing that I was capable of finishing this thesis, even at times when I did not.

Finally, thank you, Michael, for being my family and my home.
# Table of Contents

1 Introduction .................................................. 9
   1.1 Overview .................................................. 9
      1.1.1 Random data Cauchy theory on Euclidean space ............. 12
      1.1.2 Symplectic non-squeezing for the cubic NLKG ............... 15
   1.2 Notation and conventions .................................. 21
   1.3 Outline of Thesis ......................................... 23

2 Random data Cauchy theory for NLW of power-type on $\mathbb{R}^3$ .... 25
   2.1 Introduction ............................................... 25
   2.2 Deterministic and Probabilistic estimates ...................... 31
      2.2.1 Large deviation estimate ................................ 32
      2.2.2 Averaging effects for the randomized initial data ........ 33
   2.3 Proof of Theorem 2.1 ....................................... 36
   2.4 Proof of Theorem 2.2 ....................................... 44
   2.5 Appendix A: Probabilistic estimates .......................... 46

3 An approximation result for the cubic NLKG in the critical space ... 49
   3.1 Introduction ............................................... 49
   3.2 Set-up .................................................... 51
   3.3 Strichartz estimates ....................................... 54
   3.4 Multilinear estimates ..................................... 58
   3.5 Stability theory in adapted function spaces .................... 64
   3.6 A low frequency equation ................................... 70
   3.7 Proof of Theorem 3.1 ....................................... 73

4 Symplectic non-squeezing for the cubic NLKG on $\mathbb{T}^3$ ........... 75
   4.1 Introduction ............................................... 75
      4.1.1 Overview of Proof ..................................... 78
      4.1.2 Conditional non-squeezing ............................... 85
      4.1.3 Probabilistic non-squeezing ............................. 86
      4.1.4 Organization of Chapter ................................. 87
Chapter 1

Introduction

1.1 Overview

Nonlinear dispersive Partial Differential Equations (PDEs) model wave propagation phenomena in many physical systems. For the last several decades, the study of dispersive equations has focused on questions about the existence and uniqueness of solutions, their asymptotic behaviour, and singularity formation. Many of these equations enjoy a scaling symmetry which gives rise to the notion of criticality and relatedly, that of subcritical and supercritical equations. While there is a well established local theory in the subcritical or critical regimes, there are ill-posedness results, for example [17], [39] and [31], which show that one cannot expect local well-posedness for all initial data at supercritical regularities. Moreover, in cases where the critical regularity does not correspond to a conserved quantity, global deterministic results still do not even reach critical regularities.

In recent years, probabilistic tools have been extremely useful in obtaining almost sure well-posedness theorems in super-critical regimes, as well as in closing the gap between the scaling prediction and existing deterministic results. One fruitful vein of research has been the study of invariant Gibbs measures for Hamiltonian PDEs. Such measures have been studied in the work of Zhidkov [73, 74, 75], Lebowitz, Rose and Speer [40] and subsequently by many others. In [6, 8], working with the Gibbs measure introduced in [40], Bourgain proved the existence of a well-defined Hamiltonian flow on the support of this measure for the nonlinear Schrödinger equation in one and two spatial dimensions. Bourgain then used the invariance of this measure to prove almost sure global well-posedness for these equations, for supercritical initial data. In [14, 15],
Burq and Tzvetkov consider the cubic nonlinear wave equation on a three-dimensional compact manifold. They construct large sets of initial data of supercritical regularities that give rise to local solutions, using a randomization procedure which relies on expansion of the initial data with respect to an orthonormal basis of eigenfunctions of the Laplacian. Together with invariant measure considerations, they prove almost sure global existence for the cubic nonlinear wave equation on the three-dimensional unit ball. Many further results in this direction have been obtained in recent years, see [5], [9], [15], [69], [70], [68], [51], [46], [47], [58], [11], [71] and references therein.

While it is desirable to establish the existence of an invariant Gibbs measure, in many situations there are serious technical difficulties associated with defining such a measure. Most notably, in dimension $d \geq 3$ it is only possible to define a Gibbs measure for initial data in very rough Sobolev spaces. In such spaces, the multilinear analysis necessary for well-posedness arguments is not available.

In the absence of an invariant measure, other approaches have been developed to prove almost sure global existence for supercritical equations via a suitable randomization of the initial data. Energy estimates are one such approach, which was used, for instance, by Nahmod, Pavlović and Staffilani [49] in the context of the periodic Navier-Stokes equation in two and three dimensions and by Burq and Tzvetkov [16] for the three-dimensional periodic defocusing cubic nonlinear wave equation. Another approach was employed by Colliander and Oh in [19], where they adapt Bourgain’s high-low frequency decomposition [10] to prove almost sure global existence of solutions to the one-dimensional periodic defocusing cubic nonlinear Schrödinger equation below $L^2(\mathbb{T})$.

In this thesis, we pursue the study of dispersive equations via probabilistic techniques, particularly in the absence of an invariant measure. We will focus mainly on two specific equations: the defocusing nonlinear wave equation with power type nonlinearity

$$
\begin{equation}
(NLW)_\rho \quad \begin{cases}
  u_{tt} - \Delta u + |u|^\rho u = 0, & u : \mathbb{R} \times \mathbb{R}^3 \to \mathbb{R} \\
  (u, u_t)|_{t=0} = (f_0, f_1) \in H^s(\mathbb{R}^3) \times H^{s-1}(\mathbb{R}^3),
\end{cases}
\end{equation}
$$

and the defocusing cubic nonlinear Klein-Gordon equation

$$
\begin{equation}
(NLKG) \quad \begin{cases}
  u_{tt} - \Delta u + \frac{m^2 c^2}{\hbar^2} u + u^3 = 0, & u : \mathbb{R} \times \mathbb{T}^3 \to \mathbb{R} \\
  (u, \partial_t u)|_{t=0} = (u_0, u_1) \in H^s(\mathbb{T}^3) \times H^{s-1}(\mathbb{T}^3),
\end{cases}
\end{equation}
$$
where \( m \) is the mass, \( c \) is the speed of light and \( h \) is Planck's constant. Here \( H^s \) denotes the usual inhomogeneous Sobolev space. We will normalize \( m = c = h = 1 \).

The nonlinear wave equation enjoys the scaling symmetry

\[
u(t, x) \mapsto \nu(\lambda t, \lambda x) := \lambda^{2/(\rho - 1)} \nu(t, x)
\]

and the scale invariant critical space corresponds to the homogeneous Sobolev space at regularity \( s_c := \frac{d}{2} - \frac{2}{\rho - 1} \). Additionally, (1.1) has a conserved Hamiltonian

\[
H_{NLW}(\nu(t)) = \frac{1}{2} \int |\nabla_x \nu|^2 + \frac{1}{2} \int |\partial_t \nu|^2 + \frac{1}{\rho + 1} \int |\nu|^{\rho + 1},
\]

alternatively called the energy, which controls the \( \dot{H}^1 \) Sobolev norm of solutions. We will use the terminology subcritical (respectively critical or supercritical) to refer to regularities above (respectively at or below) the scale invariant Sobolev space \( \dot{H}^{s_c} \). The Hamiltonian is also called the energy, and occasionally we will refer to the equation (NLW)\( \rho \) as energy subcritical (respectively critical and supercritical) if one is interested in the Cauchy problem (1.1) with \( \rho < 5 \) (respectively \( \rho = 5 \) and \( \rho > 5 \)) since the scale invariant Sobolev space for \( \rho = 5 \) is at regularity \( s_c = 1 \), which is controlled by the energy functional.

There is no scaling symmetry for the nonlinear Klein-Gordon equation due to the presence of the mass term. However, since ill-posedness and blowup are normally associated with high frequencies or short time scales, the nonlinear term dominates the mass term and one can still regard \( s_c \) as the critical regularity for this equation. The Hamiltonian for the nonlinear Klein-Gordon equation has the form

\[
H_{NLKG}(\nu(t)) = \frac{1}{2} \int |\nabla_x \nu|^2 + \frac{c}{2} \int |\nu|^2 + \frac{1}{2} \int |\partial_t \nu|^2 + \frac{1}{\rho + 1} \int |\nu|^{\rho + 1}.
\]

The presence of the \( L^2 \) norm provides control over low frequencies of solutions, which makes the nonlinear Klein-Gordon equation somewhat less singular than the nonlinear wave equation. On Euclidean space, a useful heuristic is that the Klein-Gordon behaves like the Schrödinger equation at frequencies \( \ll c \) where \( c \) is the speed of light, and the wave equation at frequencies \( \gtrsim c \). In certain cases, this heuristic can be made rigorous, and one can prove that in the non-relativistic limit as \( c \to \infty \) one recovers the nonlinear Schrödinger equation from (1.2), see for instance [43].
Solutions to both the linear and nonlinear Klein-Gordon equations exhibit finite speed of propagation, which means that information can only propagate at the speed of light. Because of this, there is usually no difference between compact and non-compact settings if one localizes in time. In the sequel, we will always consider solutions to the Klein-Gordon equation on bounded time intervals and hence we will often make use of this fact and carry over Strichartz estimates from the Euclidean to the periodic setting.

1.1.1 Random data Cauchy theory on Euclidean space

In the first part of this thesis, we will focus on the study of the nonlinear wave equation with power type nonlinearities on Euclidean space, without any radial symmetry assumption on the initial data. This is joint work with J. Lührmann, which has appeared in [42]. Many previous results on Euclidean space have involved first considering a related equation in a setting where an orthonormal basis of eigenfunctions of the Laplacian exists. This orthonormal basis is used to randomize the initial data for the related equation and an appropriate transform is then used to map solutions of the related equation to solutions of the original equation. Instead of using such transformations, we randomize functions directly on Euclidean space via a unit-scale decomposition in frequency space. More precisely, we fix a nonzero $\psi \in C_c^\infty(\mathbb{R}^3)$ with $\text{supp } \psi \subset B(0, 2)$, define the Fourier projection operator

$$P_k f(\xi) := \psi(\xi - k)\hat{f}(\xi), \quad k \in \mathbb{Z}^3.$$ 

Let $\{h_k, l_k\}_{k \in \mathbb{Z}^3}$ denote a sequence of iid, mean-zero Gaussian random variables on a probability space $(\Omega, \mathcal{F}, \mathbb{P})$ and consider the randomized initial data

$$f^\omega = (f_0^\omega, f_1^\omega) := \left( \sum_{k \in \mathbb{Z}^3} h_k(\omega)P_k f_0, \sum_{k \in \mathbb{Z}^3} l_k(\omega)P_k f_1 \right). \tag{1.3}$$

We crucially exploit that these unit projections satisfy a unit-scale Bernstein inequality. In Chapter 2, we prove the following theorem for energy subcritical nonlinearities.

**Theorem 2.1** (Lührmann-Mendelson, [42]). Let $3 \leq \rho < 5$ and let $\frac{\rho^3 + 5\rho^2 - 11\rho - 3}{9\rho^2 - 6\rho - 3} < s < 1$. Fix $f = (f_0, f_1) \in H_x^s(\mathbb{R}^3) \times H_x^{s-1}(\mathbb{R}^3)$ and let $f^\omega = (f_0^\omega, f_1^\omega)$ be the randomized initial data defined in (1.3), and $u^\omega$ the associated free evolution. For almost every $\omega \in \Omega$ there exists a unique global
solution

\[(u, u_t) \in (u^\omega, \partial_t u^\omega) + C(\mathbb{R}; H^1_x(\mathbb{R}^3) \times L^2_x(\mathbb{R}^3))\]

to the Cauchy problem for the nonlinear wave equation \((NLW)_p\)

\[
\begin{aligned}
- \Delta u - u_{tt} = |u|^{p-1} u & \text{ on } \mathbb{R} \times \mathbb{R}^3, \\
(u, u_t)|_{t=0} &= (f_0^\omega, f_1^\omega).
\end{aligned}
\]

Figure 1.1: The dashed line is the critical regularity \(s_c = \frac{3}{2} - \frac{2}{p-1}\). The solid line is the threshold for the exponent \(s\) in Theorem 2.1.

**Remark 1.** For \(\frac{1}{4} (7 + \sqrt{73}) \approx 3.89 < p < 5\), the range of allowable regularity exponents \(s\) in Theorem 2.1 contains supercritical exponents; see Figure 1.1. To the best of the authors’ knowledge, for \(4 \leq p < 5\), Theorem 2.1 is the first result which establishes the existence of large sets of initial data of supercritical regularity which lead to global solutions to \((NLW)_p\).

The proof of Theorem 2.1 combines a probabilistic local existence argument with Bourgain’s high-low frequency decomposition [10], an approach introduced by Colliander and Oh [19, Theorem 2] to show almost sure global existence of the one-dimensional periodic defocusing cubic nonlinear Schrödinger equation below \(L^2(\mathbb{T})\). The high-low method was previously used by Kenig, Ponce and Vega [34, Theorem 1.2] to prove global well-posedness of \((1.1)\) for \(2 \leq p < 5\) for initial data in a range of sub-critical spaces below the energy space (see also [1], [24] and [60] for \(p = 3\)). We will discuss this further in Chapter 2.
When \( p = 5 \), the nonlinear wave equation (NLW)_p is energy critical and the high-low argument, being an energy subcritical technique, no longer works. However, a similar argument to the one used to prove the key averaging effects for Theorem 2.1 yields bounds on the \( L^5_t L^1_x \) norm of the free evolution \( u^\omega_t \). Together with straightforward contraction and bootstrap arguments we obtain the following theorem.

**Theorem 2.2** (Lührmann-Mendelson, [42]). Let \( \frac{2}{3} < s < 1 \). Fix \( f = (f_0, f_1) \in H^s_x(\mathbb{R}^3) \times H^{s-1}_x(\mathbb{R}^3) \) and let \( f^\omega = (f_0^\omega, f_1^\omega) \) be the randomized initial data defined in (1.3), and \( u^\omega_t \) the associated free evolution. There exists \( \Omega_f \subset \Omega \) with

\[
P(\Omega_f) \geq 1 - C \exp\left(-c/\|f\|^2_{H^s_x(\mathbb{R}^3) \times H^{s-1}_x(\mathbb{R}^3)}\right)
\]

for some absolute constants \( C, c > 0 \), such that for every \( \omega \in \Omega_f \) there exists a unique global solution

\[
(u, u_t) \in (u^\omega_t, \partial_t u^\omega_t) + C(\mathbb{R}; H^1_x(\mathbb{R}^3) \times L^2_x(\mathbb{R}^3))
\]

to the quintic nonlinear wave equation

\[
\begin{aligned}
-u_{tt} + \Delta u &= |u|^4 u \text{ on } \mathbb{R} \times \mathbb{R}^3, \\
(u, u_t)|_{t=0} &= (f_0^\omega, f_1^\omega).
\end{aligned}
\]

Moreover, (1.4) scatters in the sense that for every \( \omega \in \Omega_f \) there exist \( (v_1, v_2) \in H^1_x(\mathbb{R}^3) \times L^2_x(\mathbb{R}^3) \) such that the free evolution \( v(t) = \cos(t|\nabla|)v_1 + \frac{\sin(t|\nabla|)}{|\nabla|}v_2 \) satisfies

\[
\|(u(t) - u^\omega_t(t) - v(t), \partial_t u(t) - \partial_t u^\omega_t(t) - \partial_t v(t))\|_{H^1_x(\mathbb{R}^3) \times L^2_x(\mathbb{R}^3)} \to 0 \text{ as } t \to \pm \infty.
\]

In [54], Pocovnicu proves almost sure global well-posedness for the energy critical defocusing nonlinear wave equation on \( \mathbb{R}^d \) for \( d = 4, 5 \). Her arguments are based on a probabilistic perturbation theory which relies on the fact that the cubic nature of the nonlinearity she considers allows one to treat the difference equation for the nonlinear component of the solution perturbatively and one can use the energy to obtain the necessary bounds for the perturbation argument. Recently, Oh and Pocovnicu have obtained control of an energy functional in [52] for the quintic equation, which similarly yields global existence of solutions when combined with the probabilistic perturbation theory. Such methods only yield a mild form of uniqueness in the energy critical setting.
1.1.2 Symplectic non-squeezing for the cubic NLKG

In the second part of the thesis, we study symplectic non-squeezing for the periodic defocusing cubic nonlinear Klein-Gordon equation

\[
\begin{cases}
  u_{tt} - \Delta u + u + u^3 = 0, & u : \mathbb{R} \times \mathbb{T}^3 \to \mathbb{R} \\
  (u, \partial_t u)|_{t=0} = (u_0, u_1) \in H^{1/2}(\mathbb{T}^3) \times H^{-3/2}(\mathbb{T}^3) =: \mathcal{H}^{1/2}(\mathbb{T}^3),
\end{cases}
\] (1.5)

where $H^{1/2}(\mathbb{T}^3)$ is the usual inhomogeneous Sobolev space.

Local strong solutions to (1.5) can be constructed by adapting the arguments from [41] to the compact setting. Due to the critical nature of this problem, however, the local time of existence for solutions depends not only on the norm of the initial data but also on its profile. Moreover, as the critical regularity for this equation does not correspond to a conserved quantity, global well-posedness for this equation remains open. The best known results on Euclidean space are for subcritical regularities $s > 3/4$, which was proved by Miao, Zhang and Fang [45] by working in Besov spaces and adapting Bourgain’s high-low argument [10] and arguments from [34]. Once again, these results can be adapted to the periodic setting. We recall that the $L_{t,x}^4$ Strichartz norm controls the global well-posedness for (1.5) and we have the standard finite time blow-up criterion, namely if $T_*$ denotes the maximal time of existence for a solution $u$ to (1.5) then

\[ T_* < \infty \quad \Longrightarrow \quad \|u\|_{L_{t,x}^4([0,T_*] \times \mathbb{T}^3)} = +\infty. \] (1.6)

We are interested in studying **symplectic non-squeezing** for (1.5), which is formally an infinite dimensional Hamiltonian system, as we will see in Section 4.2. In the finite dimensional setting, Gromov’s celebrated non-squeezing theorem states that there is no symplectic embedding of a ball into a cylinder unless the radius of the ball is less than or equal to that of the cylinder. Smooth finite-dimensional Hamiltonian flows are particular examples of symplectomorphisms and hence, they exhibit non-squeezing by Gromov’s theorem.

One may wonder which parts of this theory, if any, carry over to infinite dimensions. There is a natural symplectic structure on Sobolev spaces, and for infinite dimensional Hamiltonian equations, there are specific regularities where this structure is compatible with the flow of the equation, and Hamiltonian PDEs can be realized, at least formally, as a symplectic flow on these infinite dimensional phase spaces. Beyond the question of merely generalizing Gromov’s theorem to certain
infinite dimensional cases, symplectic non-squeezing for Hamiltonian equations is connected to the problem of weak turbulence, which examines, for example, whether the energy of a given solution concentrates at high frequencies over time, see Section 4.2 for a discussion of this interpretation.

Gromov proved his non-squeezing theorem by showing that a certain symplectic capacity, called the Darboux width, is invariant under the flow of a symplectomorphism via a sophisticated analysis using pseudoholomorphic curves. Symplectic capacities, which we define for symplectic Hilbert spaces in Subsection 4.2.1, are an important invariant of symplectic flows. Subsequent to Gromov’s proof, there were other (comparable but not necessarily equivalent) definitions of symplectic capacities, introduced by Ekeland-Hofer, Viterbo, and Hofer-Zehnder among others. We refer the reader [30] and references therein for a more thorough account of these developments.

The study of infinite dimensional symplectic capacities and non-squeezing for nonlinear Hamiltonian PDEs was initiated by Kuksin in [38]. There, he extended the definition of the Hofer-Zehnder capacity to infinite dimensional Hilbert spaces and proved the invariance of this capacity under the flow of certain Hamiltonian equations with flow maps of the form

$$\Phi(t) = \text{linear operator} + \text{compact smooth operator}. \quad (1.7)$$

This infinite dimensional symplectic capacity inherits the finite dimensional normalization

$$\text{cap}(B_r(u*)) = \text{cap}(C_r(z; k_0)) = \pi r^2,$$

where $B_r(u*)$ is the infinite dimensional ball centered at $u*$ in the Hilbert space, and $C_r(z; k_0)$, the infinite dimensional cylinder, see (1.14) for the precise definition of the cylinder in our context. The proof of this normalization in infinite dimensions is an adaptation of the original proof by Hofer and Zehnder which can be found in [30], see [38] for details of the infinite dimensional argument. Consequently, if a flow map $\Phi$ preserves capacities, one can conclude that squeezing is impossible, namely

$$\Phi(t)(B_R(u*)) \not\subset C_r(z; k_0) \quad \text{if } R < r.$$

Several examples of nonlinear Klein-Gordon equations with weak nonlinearities can readily be shown to be of the form (1.7), see [38]. Symplectic non-squeezing was later proved for certain subcritical nonlinear Klein-Gordon equations in [7] using Kuksin’s framework, see also [59]. Bourgain later extended these results to the cubic NLS in dimension one in [4], where the flow is not
a compact perturbation of the linear flow. There, the argument follows from approximating the full equation by a finite dimensional flow and applying Gromov’s finite dimensional non-squeezing result to this approximate flow. Symplectic non-squeezing was also proven for the KdV [20]. In this situation, there is a lack of smoothing estimates in the symplectic space which would allow the infinite dimensional KdV flow to be easily approximated by a finite-dimensional Hamiltonian flow. To resolve this issue, the authors of [20] invert the Miura transform to work on the level of the modified KdV equation, for which stronger estimates can be established.

As we are interested in studying non-squeezing for an equation at the critical regularity, we face substantial difficulties if we try to naively adapt the previous approaches to this setting. Additionally, there is no uniform control on the local time of existence and as the critical regularity is not controlled by a conserved quantity, the global well-posedness of (1.5) remains open. Ultimately, however, we are able to circumvent these difficulties, using a combination of probabilistic and deterministic techniques, which we combine to obtain several deterministic non-squeezing results. The first result, which we prove in Chapter 4 is a local-in-time non-squeezing theorem.

**Theorem 4.1.** Let $\Phi$ denote the flow of the cubic nonlinear Klein-Gordon equation (1.5). Fix $R > 0$, $k_0 \in \mathbb{Z}^3$, $z \in \mathbb{C}$, and $u_* \in \mathcal{H}^{1/2}(T^3)$. For all $0 < \eta < R$, there exists $N \equiv N(\eta, u_*, R, k_0)$ and $\sigma \equiv \sigma(\eta, N, u_*) > 0$ such that for all $0 < t < \sigma$,

$$\Phi(t)(\Pi_N B_R(u_*)) \not\subset C_r(z; k_0) \quad \text{for } r < R - \eta. \quad (1.8)$$

In the statement of this theorem, $\Pi_N$ is a projection onto frequencies $|k| \leq N$, see (1.13) for its precise definition.

**Remark 1.1.** The parameter $\eta$ which appears in Theorem 4.1 corresponds to the control we can obtain over the radius of the cylinder. If we demand better control over the radius, this theorem only holds for shorter time scales. See also Remark 4.4 for a discussion of the dependence on $u_*$. 

**Remark 1.2.** To prove this theorem, we combine a probabilistic approximation argument with the available stability theory to prove Theorem 4.1. As we have no control on the local time of existence in the critical space, a priori we cannot ensure the flow map $\Phi(t)$ is well-defined on any infinite set for any positive time. We fix the projection in (1.8) at frequency $N$ so that we have enough control to define the flow map $\Phi(t)$ for $t \in [0, \sigma]$. 

17
In order to state our global-in-time results, we need to introduce the following nonlinear Klein-Gordon equation with truncated nonlinearity

\[
\begin{cases}
(u_N)_{tt} - \Delta u_N + u_N + P_N(P_N u_N)^3 = 0, & u : \mathbb{R} \times \mathbb{T}^3 \to \mathbb{R} \\
(u_N, \partial_t u_N)|_{t=0} = (u_0, u_1) \in H^{1/2}(\mathbb{T}^3),
\end{cases}
\]  

(1.9)

where \( P_N = P_{\leq N} \) denotes the smooth projection operator defined in (1.12). We obtain the following global-in-time non-squeezing result.

**Theorem 4.2.** Let \( \Phi \) denote the flow of the cubic nonlinear Klein-Gordon equation (1.5). Fix \( R, T > 0, k_0 \in \mathbb{Z}^3, z \in \mathbb{C}, \) and \( u_* \in H^{1/2}(\mathbb{T}^3) \). Suppose there exists some \( K > 0 \) such that for all \( (u_0, u_1) \in B_R(u_*) \), the corresponding solutions \( u \) to (1.5) and \( u_N \) to (1.9) satisfy

\[
\|u\|_{L^4_t L^4_x([0,T) \times \mathbb{T}^3)} + \sup_N \|P_N u_N\|_{L^4_t L^4_x([0,T) \times \mathbb{T}^3)} \leq K.
\]  

(1.10)

Then

\[
\Phi(T)(B_R(u_*)) \not\subseteq C_r(z; k_0) \quad \text{for} \quad r < R.
\]

Moreover, if \( B_R(u_*) \subset B_{\rho_0} \) for some sufficiently small \( \rho_0(T) > 0 \), then non-squeezing holds without any additional assumptions on the initial data.

**Remark 1.3.** For small data, the additional assumptions on the Strichartz norm of solutions automatically hold by the standard arguments, see Lemma 4.31. By (1.6), the assumption (1.10) implies that in particular, the corresponding solutions \( u \) to (1.5) and \( u_N \) to (1.9) exist on \([0,T)\). We will frequently be implicitly making use of this fact throughout this thesis.

**Remark 1.4.** This theorem implies, in particular, that if one can define a global flow for the nonlinear Klein-Gordon and prove uniform Strichartz bounds for solutions to both (1.5) and (1.9) for initial data in bounded subsets, then one obtains the full, deterministic statement of non-squeezing theorem for this equation. In practice a global well-posedness result for the full equation will typically also yield estimates for the equation with truncated nonlinearity, so while the requirement for bounds on the solution to the equation with truncated nonlinearity may seem artificial, we do not believe it is too restrictive a requirement.

The main tool in the proof of Theorem 4.2, is an approximation result which we prove in Chapter 3. It says that in the critical space, solutions to (1.5) are stable at low frequencies under
high-frequency perturbations to the initial data. The proof uses the $U^p$ and $V^p$ function spaces, whose definition and properties we record in Appendix A. These spaces were applied for the first time by Koch and Tataru in [37]. They have previously been used in the context of critical problems by Hadac, Herr and Koch [28] for the KP-II equation, and by Herr, Tataru and Tzvetkov [29] for the quintic nonlinear Schrödinger equation on $T^3$, as well as by Nahmod and Staffilani [48] for probabilistic well-posedness of the quintic nonlinear Schrödinger equation on $T^3$ below the energy space. See [36] or [28] and references therein for a more complete overview of these function spaces.

**Theorem 3.1.** Let $\Phi$ denote the flow of the cubic nonlinear Klein-Gordon equation (1.5). Let $T > 0$ and $1 \leq N' < N_*$. Let $(u_0, u_1), (\tilde{u}_0, \tilde{u}_1) \in B_R \subset H^{1/2}(T^3)$ be such that $P_{\leq N}(u_0, u_1) = P_{\leq N}(\tilde{u}_0, \tilde{u}_1)$, and suppose there exists some $K > 0$ such that corresponding solutions $u$ and $\tilde{u}$ to (1.5) satisfy

$$
\|u\|_{L^4_t(L^4_x([0,T] \times T^3))} + \|\tilde{u}\|_{L^4_t(L^4_x([0,T] \times T^3))} \leq K.
$$

Then for sufficiently large $N_*$ depending on $R, T$ and $K$,

$$
\|P_{\leq N'}(\Phi(t)(u_0, u_1) - \Phi(t)(\tilde{u}_0, \tilde{u}_1))\|_{L^\infty_t H^{1/2}_x([0,T] \times T^3)} \lesssim \left( \log \frac{N_*}{N'} \right)^{-\theta},
$$

with implicit constant depending on $R, T, K$.

We prove Theorem 3.1 by demonstrating that under the above assumptions, the low frequency component $u_{t0}$ satisfies a perturbed cubic Klein-Gordon equation given by

$$
\square u_{t0} + u_{t0} = P_{t0}F(u_{t0}, u_{t0}, u_{t0}) + err,
$$

where $err$ is an error term which we can control by the well-posedness theory.

**Non-squeezing and Liouville's theorem**

We will now provide a heuristic interpretation of non-squeezing for Hamiltonian systems. We consider the simplest example of a symplectic phase space, given by the vector space $(\mathbb{R}^{2n}, \omega_0)$ representing the positions and momenta of $n$ particles, with coordinates $(x_i, y_i)$ corresponding to the $i$-th particle, symplectic form $\omega_0$ to be given by

$$
\omega_0 = dx_1 \wedge dy_1 + \ldots + dx_n \wedge dy_n.
$$
Given a smooth Hamiltonian $H : \mathbb{R}^{2d} \to \mathbb{R}$, Hamilton’s equations of motion can be written as

$$
\begin{align*}
\dot{x}_i &= \frac{\partial H}{\partial y_i} \\
\dot{y}_i &= -\frac{\partial H}{\partial x_i}
\end{align*}
$$
or equivalently, as the flow induced by the vector field $X_H$ defined implicitly by $\omega_0(X_H, \cdot) = -dH$.

Liouville’s well-known theorem in classical mechanics states that Hamiltonian flows preserve volume. Thus Liouville’s theorem exerts some control over how much uncertainty can arise in such a system. Namely, if you start with initial conditions in a certain region in phase space, you will not be able to determine the state of the system with any greater accuracy (in terms of phase space volume) after it undergoes a Hamiltonian evolution, nor will you lose accuracy completely.

It is thus natural to ask whether one will ever be able to determine the location of one particle very precisely, if one is willing to give up accuracy for the remaining particles. Say we start with initial data in a ball in the phase space, and after some time we wish to know the position and momentum of the $k$-th particle very precisely for some $1 \leq k \leq n$. Geometrically this corresponds to starting with initial conditions in a ball or radius $R$ in phase space and embedding this ball into a very thin cylinder in the $k$-the coordinate plane, defined by

$$
C_r(z; k) := \{(x, y) \in \mathbb{R}^{2n} : (x_k - z_0)^2 + (y_k - z_1)^2 < r^2\},
$$

for $z = (z_0, z_1) \in \mathbb{C}$. Recalling that Hamiltonian equations are a particular example of symplectic flows, Gromov’s non-squeezing theorem says this is impossible unless $R < r$.

**Non-squeezing and transfer of energy to higher Fourier modes**

Finally, we explain briefly the heuristics of how non-squeezing relates to the transfer of energy to higher Fourier modes for Hamiltonian equations. For a solution $u$ to a Hamiltonian equation with Fourier series

$$
u(x, t) = \sum_{k \in \mathbb{Z}^d} \hat{u}_k(t)e^{ikx}, \quad \hat{u}_k(t) \in \mathbb{C}.
$$

The energy transition problem investigates whether $|\hat{u}_k(t)|$ decays with $t$ for some fixed $k$. To reformulate this question, let $\mathcal{H}$ denote the phase space for the given equation, then this becomes
a question about whether for some ball $B \subset \mathcal{H}$, for $k \in \mathbb{Z}^3$ and for $\theta \ll 1$ one has

$$|\tilde{u}_k(t)| \leq \theta, \quad t = t_k \gg 1$$

(1.11)

for solutions with $u(0) \in B$. If the equation we consider has a conserved energy which controls a given Sobolev norm, as is the case for the nonlinear Klein-Gordon equation, we can interpret (1.11) as a statement that the higher Sobolev norms of a solution grow in time. This is due to the fact that if there is decay of certain Fourier modes, but a certain Sobolev norm remains bounded, then some other higher Sobolev norm must compensate by growing, since higher Sobolev norms weight higher Fourier modes more heavily. Now let $\Phi$ denote the flow map of the Hamiltonian equation. We can rewrite the condition from (1.11) as the requirement that $\Phi$ embeds the ball $B$ into a thin cylinder,

$$\Phi(t)(B) \subseteq \{ u \in \mathcal{H} : |u_k(t)| \leq \theta \}.$$

If the ball has radius $R > 0$, then non-squeezing says this is impossible unless $R \leq \theta$. Thus, non-squeezing implies a negative answer to the question of the uniform decay of Fourier modes.

### 1.2 Notation and conventions

We let $C > 0$ denote a constant that depends only on fixed parameters and whose value may change from line to line. Where relevant, we will explicitly record the dependence of $C$ on these fixed parameters. We write $X \lesssim Y$ to denote $X \leq CY$ and similarly $X \gtrsim Y$. If there exists some small constant $c > 0$ such that $X \leq cY$, then we write $X \ll Y$ and similarly for $X \gg Y$. For some constant $a$ we write $a+$ as a shorthand for $a + \varepsilon$, for $\varepsilon > 0$ some arbitrarily small, fixed parameter.

Let $\Lambda = \mathbb{T}$ or $\mathbb{R}$. Throughout we will be using space-time norms

$$\|u\|_{L^q_t L^r_x(\mathbb{R} \times \Lambda^d)} := \left( \int_{\mathbb{R}} \left( \int_{\Lambda^d} |u(t, x)|^r \right)^{q/r} \right)^{1/q}.$$

We use the usual convention of $L^2_x$ normalized Fourier transform with the notation

$$\hat{f}(n) = (2\pi)^{-\frac{n}{2}} \int_{\mathbb{T}^n} e^{-i\xi \cdot x} f(x) dx \quad \text{and} \quad \hat{f}(\xi) = (2\pi)^{-\frac{n}{2}} \int_{\mathbb{R}^n} e^{-i\xi \cdot x} f(x) dx.$$

We make use of the Japanese bracket $\langle \xi \rangle := \sqrt{1 + |\xi|^2}$. As mentioned above, $H^s_x(\Lambda^3)$ denotes the
usual inhomogeneous Sobolev space endowed with the norm

\[ \|u\|_{H^s_2(\mathbb{R}^d)} = \|\langle \xi \rangle^s \hat{u}(\xi)\|_{L^2_2(\mathbb{R}^d)} \quad \text{and} \quad \|u\|_{H^s_2(\mathbb{T}^d)} = \|\langle n \rangle^s \hat{u}(n)\|_{L^2_2(\mathbb{Z}^d)} \]

while \( \dot{H}^s_x(\mathbb{R}^3) \) denotes the corresponding homogeneous spaces with \(|\xi|\) or \(|n|\) in place of the Japanese bracket. We use the notation

\[ \mathcal{H}^s(\Lambda^3) := H^s_x(\Lambda^3) \times H^{s-1}_x(\Lambda^3), \]

where \( \mathcal{H}^s(\Lambda^3) \) is endowed with the euclidean norm on the product. We fix a nonzero \( \psi \in C_c^\infty(\mathbb{R}^3) \) with \( \text{supp} \Psi \subset B(-2, 2) \), and define the Littlewood-Paley projection operators

\[ \tilde{P}_N f(\xi) := \psi(N^{-1} \xi) \hat{f}(\xi), \quad N \in 2\mathbb{Z} \]

and note the \( P_N \) form a partition of unity. In most cases we will be using an inhomogeneous decomposition and take \( N \in 2^N \), a dyadic integer. We use the notation \( u_N = P_N u \) and we define

\[ P_{\leq N} u := \sum_{|M| \leq N} u_M, \quad P_{\geq N} = 1 - P_{\leq N}. \]

For more properties of these operators, see Appendix A in [66]. On the torus, \( P_N \) will similarly denote a smooth projection operator, defined by

\[ P_{\leq N}(u)(x) \equiv \psi(-N^{-2} \Delta)(u)(x) = \hat{u}(0) + \sum_{n \in \mathbb{Z}^d} \psi \left( \frac{|n|^2}{N^2} \right) \hat{u}(n) e^{i(n \cdot x)}, \quad (1.12) \]

for \( \psi \) a smooth cut-off as above. We will need the sharp Fourier projection operators

\[ \Pi_{\leq K}(u_0, u_1) := \left( \sum_{|k| \leq K} \hat{u}_0(k) e^{i k \cdot x}, \sum_{|k| \leq K} \hat{u}_1(k) e^{i k \cdot x} \right), \quad \Pi_{\geq K} = I - \Pi_K. \quad (1.13) \]

We define the ball of radius \( R \) centered at \( u_* \) in \( \mathcal{H}^{1/2} \) by

\[ B_R(u_*) = \{ u \in \mathcal{H}^{1/2} : \|u - u_*\|_{\mathcal{H}^{1/2}} < R \} \]

and we use the shorthand \( B_R := B_R(0) \). We define the cylinder in \( \mathcal{H}^{1/2}(\mathbb{T}^3) \) of radius \( r > 0 \) in the
$k$-th frequency, centered at $z = (z_0, z_1) \in \mathbb{C}$ by

$$C_r(z; k_0) := \{(u_1, u_2) \in \mathcal{H}^{1/2}(\mathbb{T}^3) : \langle k_0 \rangle |\widehat{u}_1(k_0) - z_0|^2 + \langle k_0 \rangle^{-1} |\widehat{u}_2(k_0) - z_1|^2 < r^2\}, \quad (1.14)$$

for real-valued Fourier coefficients $\widehat{u}_i(k)$. Since we only consider non-squeezing for the phase space of real-valued functions $u \in \mathcal{H}^{1/2}(\mathbb{T}^3)$, we can always identify such a function with real-valued Fourier coefficients.

### 1.3 Outline of Thesis

In Chapter 2 we study the random data Cauchy problem for the nonlinear wave equation with power type nonlinearity on $\mathbb{R}^3$. These results are joint work with Jonas Lührmann and have appeared in [42]. In Chapter 3, we prove an approximation result for the cubic nonlinear Klein-Gordon equation in the critical space. In Chapter 4, we prove symplectic non-squeezing for the cubic nonlinear Klein-Gordon equation on $\mathbb{T}^3$. Many of the results from Chapter 4 have appeared in the preprint [44]. Finally we record some background material from Harmonic analysis in Appendix A.

The author was supported in part by the U.S. National Science Foundation grant DMS-1068815 and by the NSERC Postgraduate Scholarships Program.
Chapter 2

Random data Cauchy theory for NLW of power-type on $\mathbb{R}^3$

2.1 Introduction

We consider the Cauchy problem for the defocusing nonlinear wave equation

$$\begin{cases}
-\partial_{tt} + \Delta u = |u|^{\rho-1}u & \text{on } \mathbb{R} \times \mathbb{R}^3, \\
(u, u_t)|_{t=0} = (f_0, f_1) \in H^s_x(\mathbb{R}^3) \times H^{s-1}_x(\mathbb{R}^3),
\end{cases} \quad (2.1)$$

where $3 \leq \rho \leq 5$ and $H^s_x(\mathbb{R}^3)$ is the usual inhomogeneous Sobolev space. The main result of this chapter establishes the almost sure existence of global solutions to (2.1) with respect to a suitable randomization of initial data in $H^s_x(\mathbb{R}^3) \times H^{s-1}_x(\mathbb{R}^3)$ for $3 \leq \rho < 5$ and $\frac{\rho^3+5\rho^2-11\rho-3}{9\rho^2-6\rho-3} < s < 1$. In particular, for $\frac{2}{7} (7 + \sqrt{73}) < \rho < 5$, this yields examples of initial data of super-critical regularity for which solutions to (2.1) exist globally in time. See Remark 7 for a discussion of uniqueness of these solutions.

A systematic investigation of the local well-posedness of (2.1) for initial data in homogeneous Sobolev spaces is undertaken by Lindblad and Sogge in [41], where local strong solutions to (2.1) are constructed for $s \geq \frac{3}{2} - \frac{2}{\rho-1}$ using Strichartz estimates for the wave equation. When $s < \frac{3}{2} - \frac{2}{\rho-1}$, this is the super-critical regime and the well-posedness arguments based on Strichartz estimates break down. Recently, several methods have emerged proving ill-posedness for (2.1) below the critical scaling regularity (see Lebeau [39], Christ-Colliander-Tao [17] and Ibrahim-
In this chapter we study almost sure global existence for nonlinear wave equations on Euclidean space without any radial symmetry assumption on the initial data. Previous results on Euclidean space have involved first considering a related equation in a setting where an orthonormal basis of eigenfunctions of the Laplacian exists. This orthonormal basis is used to randomize the initial data for the related equation and an appropriate transform is then used to map solutions of the related equation to solutions of the original equation. Burq, Thomann and Tzvetkov [12, Theorem 1.2] prove almost sure global existence and scattering for the defocusing nonlinear Schrödinger equation on \( \mathbb{R} \) by first treating a one-dimensional nonlinear Schrödinger equation with a harmonic potential and then invoking the lens transform [12, (10.2)]. Subsequently, similar approaches were used by Deng [22, Theorem 1.2], Poiret [56], [55] and Poiret, Robert and Thomann [57, Theorem 1.3] to study the defocusing nonlinear Schrödinger equation on \( \mathbb{R}^d \) for \( d \geq 2 \). In the context of the wave equation, Suzzoni [64], [21, Theorem 1.2] first considers a nonlinear wave equation on the three-dimensional sphere and then uses the Penrose transform [64, (3)] to obtain almost sure global existence and scattering for (2.1) for \( 3 \leq \rho < 4 \).

Instead of using such transforms, we will randomize functions directly on Euclidean space via a unit-scale decomposition in frequency space. More precisely, let \( \varphi \in C_0^\infty(\mathbb{R}^3) \) be a real-valued, smooth, non-increasing function such that \( 0 \leq \varphi \leq 1 \) and

\[
\varphi(\xi) = \begin{cases} 
1 & \text{for } |\xi| \leq 1, \\
0 & \text{for } |\xi| \geq 2.
\end{cases}
\]

For every \( k \in \mathbb{Z}^3 \) set \( \varphi_k(\xi) = \varphi(\xi - k) \) and define

\[
\psi_k(\xi) = \frac{\varphi_k(\xi)}{\sum_{l \in \mathbb{Z}^3} \varphi_l(\xi)}.
\]

Then \( \psi_k \) is smooth with support contained in \( \{ \xi \in \mathbb{R}^3 : |\xi - k| \leq 2 \} \). Note that \( \sum_{k \in \mathbb{Z}^3} \psi_k(\xi) = 1 \) for all \( \xi \in \mathbb{R}^3 \). For \( f \in L^2_x(\mathbb{R}^3) \) define the function \( P_k f : \mathbb{R}^3 \to \mathbb{C} \) by

\[
(P_k f)(x) = \mathcal{F}^{-1} \left( \psi_k(\xi) \hat{f}(\xi) \right)(x) \text{ for } x \in \mathbb{R}^3.
\]
If \( f \in H^s_x(\mathbb{R}^3) \) for some \( s \in \mathbb{R} \), then \( P_kf \in H^s_x(\mathbb{R}^3) \) and \( f = \sum_{k \in \mathbb{Z}^3} P_kf \) in \( H^s_x(\mathbb{R}^3) \) with

\[
\|f\|_{H^s_x(\mathbb{R}^3)} \sim \left( \sum_{k \in \mathbb{Z}^3} \|P_kf\|_{H^s_x(\mathbb{R}^3)}^2 \right)^{1/2}.
\]

In the proof of the almost sure existence of global solutions to (2.1), we will crucially exploit that these projections satisfy a unit-scale Bernstein inequality, namely that for all \( 2 \leq p_1 \leq p_2 \leq \infty \) there exists \( C \equiv C(p_1, p_2) > 0 \) such that for all \( f \in L^2_x(\mathbb{R}^3) \) and for all \( k \in \mathbb{Z}^3 \)

\[
\|P_kf\|_{L^{p_2}}(\mathbb{R}^3) \leq C\|P_kf\|_{L^{p_1}}(\mathbb{R}^3).
\]

(2.2)

For a formal statement of this inequality, see Lemma 2.3. Let now \( \{(h_k, l_k)\}_{k \in \mathbb{Z}^3} \) be a sequence of independent, zero-mean, real-valued random variables on a probability space \((\Omega, \mathcal{A}, \mathbb{P})\) with distributions \( \mu_k \) and \( \nu_k \). Assume that there exists \( c > 0 \) such that

\[
\left| \int_{-\infty}^{\infty} e^{\gamma x} \, d\mu_k(x) \right| \leq e^{c\gamma^2} \text{ for all } \gamma \in \mathbb{R} \text{ and for all } k \in \mathbb{Z}^3,
\]

and similarly for \( \nu_k \). The assumption (2.3) is satisfied, for example, by standard Gaussian random variables, standard Bernoulli random variables, or any random variables with compactly supported distributions. For a given \( f = (f_0, f_1) \in H^s_x(\mathbb{R}^3) \times H^{s-1}_x(\mathbb{R}^3) \) we define its randomization by

\[
f^\omega = (f_0^\omega, f_1^\omega) := \left( \sum_{k \in \mathbb{Z}^3} h_k(\omega) P_kf_0, \sum_{k \in \mathbb{Z}^3} l_k(\omega) P_kf_1 \right).
\]

(2.4)

The quantity \( \sum_{k \in \mathbb{Z}^3} h_k(\omega) P_kf_0 \) is understood as the Cauchy limit in \( L^2(\Omega; H^s_x(\mathbb{R}^3)) \) of the sequence \( \left( \sum_{|k| \leq N} h_k(\omega) P_kf_0 \right)_{N \in \mathbb{N}} \) and similarly for \( \sum_{k \in \mathbb{Z}^3} l_k(\omega) P_kf_1 \). Let

\[
u_t^\omega = \cos(t|\nabla|) f_0^\omega + \frac{\sin(t|\nabla|)}{|\nabla|} f_1^\omega
\]

(2.5)

be the free wave evolution of the initial data \( f^\omega \) defined in (2.4).

In the case of random variables such that there exists \( c > 0 \) for which their distributions satisfy

\[
\sup_{k \in \mathbb{Z}^3} \mu_k([-c, c]) < 1,
\]

one can show that if \( f \) does not belong to \( H^{s+\epsilon}_x(\mathbb{R}^3) \times H^{s-1+\epsilon}_x(\mathbb{R}^3) \), then the probability that \( f^\omega \)
belongs to \( H_x^{s+\varepsilon}(\mathbb{R}^3) \times H_x^{s-1+\varepsilon}(\mathbb{R}^3) \) is zero, see Lemma 2.12. Thus, our randomization procedure does not regularize at the level of Sobolev spaces.

A similar randomization procedure on Euclidean space was used by Zhang and Fang in [72, (1.12)] to study random data local existence and small data global existence questions for the generalized incompressible Navier-Stokes equation on \( \mathbb{R}^d \) for \( d \geq 3 \).

We are now in a position to state our results.

**Theorem 2.1.** Let \( 3 \leq \rho < 5 \) and let

\[
\frac{\rho^3 + 5\rho^2 - 11\rho - 3}{9\rho^2 - 6\rho - 3} < s < 1.
\]

Let \( f = (f_0, f_1) \in H_x^s(\mathbb{R}^3) \times H_x^{s-1}(\mathbb{R}^3) \). Let \( \{(h_k, l_k)\}_{k \in \mathbb{Z}^3} \) be a sequence of independent, zero-mean value, real-valued random variables on a probability space \((\Omega, \mathcal{A}, P)\) with distributions \( \mu_k \) and \( \nu_k \). Assume that there exists \( c > 0 \) such that

\[
\left| \int_{-\infty}^{+\infty} e^{\gamma x} d\mu_k(x) \right| \leq e^{c\gamma^2} \text{ for all } \gamma \in \mathbb{R} \text{ and for all } k \in \mathbb{Z}^3
\]

and similarly for \( \nu_k \). Let \( f^\omega = (f_0^\omega, f_1^\omega) \) be the associated randomized initial data as defined in (2.4) and let \( u^\omega(t) \) be the associated free evolution as defined in (2.5). For almost every \( \omega \in \Omega \) there exists a unique global solution

\[
(u, u_t) \in (u_x^+ \partial_t u_x^+ + C(\mathbb{R}; H^1_x(\mathbb{R}^3) \times L^2_x(\mathbb{R}^3)))
\]

to the nonlinear wave equation

\[
\begin{aligned}
- u_{tt} + \Delta u &=|u|^{\rho-1}u \text{ on } \mathbb{R} \times \mathbb{R}^3, \\
(u, u_t)|_{t=0} &= (f_0^\omega, f_1^\omega).
\end{aligned}
\]  

(2.6)

Here, uniqueness only holds in a mild sense; see Remark 7.

The proof of Theorem 2.1 combines a probabilistic local existence argument with Bourgain's high-low frequency decomposition [10], an approach introduced by Colliander and Oh [19, Theorem 2] to show almost sure global existence of the one-dimensional periodic defocusing cubic nonlinear Schrödinger equation below \( L^2(\mathbb{T}) \). The high-low method was previously used by Kenig, Ponce and Vega [34, Theorem 1.2] to prove global well-posedness of (2.1) for \( 2 \leq \rho < 5 \) for initial data.
in a range of sub-critical spaces below the energy space (see also [1], [24] and [60] for $\rho = 3$). We adopt arguments from [34].

Theorem 2.1 permits initial data at lower regularities than the deterministic result [34, Theorem 1.2]. This is mainly due to so-called averaging effects for the free evolution of randomized initial data (see Lemma 2.7 below), which are proven by combining the unit-scale Bernstein estimate (2.2) and Strichartz estimates for the wave equation on $\mathbb{R}^3$. In particular, we use here that the randomization is performed directly on Euclidean space.

Remark 2. For $3 \leq \rho < 4$, random data Cauchy theory for the nonlinear wave equation (2.1) on $\mathbb{R}^3$ has been addressed by Suzzoni [64, 21] using different approaches than in the proof of Theorem 2.1. In [64, Theorem 2], using methods from [14, 15], almost sure global existence and scattering for (2.1) for $3 \leq \rho < 4$ is established for radially symmetric initial data in a class of spaces of super-critical regularity related to the Penrose transform. For $\rho = 3$, almost sure global existence and scattering for (2.1) is proven in [21, Theorem 1.2], using methods from [16]. In both these cases, the spaces for the initial data do not coincide with $H_x^s(\mathbb{R}^3) \times H_x^{s-1}(\mathbb{R}^3)$. We do not address the question of scattering of the constructed solutions in Theorem 1.1. The main difficulty is that the high-low method does not yield bounds on any global space-time norm of the nonlinear component of the solution $u$.

Remark 3. During the final revision of this article, the preprints [2, 3] by Bényi-Oh-Pocovnicu and [54] by Pocovnicu appeared which use a similar randomization procedure. In [2, 3] almost sure well-posedness results are established for the cubic nonlinear Schrödinger equation on $\mathbb{R}^d$ for $d \geq 3$. In [54] almost sure global well-posedness is proven for the energy-critical defocusing nonlinear wave equation on $\mathbb{R}^d$ for $d = 4, 5$.

Remark 4. For $\frac{1}{4}(7 + \sqrt{13}) \approx 3.89 < \rho < 5$, the range of allowable regularity exponents $s$ in Theorem 2.1 contains super-critical exponents; see Figure 2.1. To the best of the authors’ knowledge, for $4 \leq \rho < 5$, Theorem 2.1 is the first result which establishes the existence of large sets of initial data of super-critical regularity which lead to global solutions to (2.1).

Remark 5. For the specific case $\rho = 3$, our randomization procedure allows us to straightforwardly adapt the proofs in [16] to the Euclidean setting. In particular, by modifying the proof of the averaging effects of Corollary A.4 of [16], we can use the arguments of Proposition 2.1 and Proposition
Figure 2.1: The dashed line is the critical regularity $s_c = \frac{3}{2} - \frac{2}{p-1}$. The solid line is the threshold for the exponent $s$ in Theorem 2.1.

2.2 of [16] to prove that if $0 < s < 1$ and $f = (f_0, f_1) \in H^s_x(\mathbb{R}^3) \times H^{s-1}_x(\mathbb{R}^3)$, then for almost every $\omega \in \Omega$ there exists a unique global solution

$$(u, u_t) \in (u^\omega_f, \partial_t u^\omega_f) + C(\mathbb{R}; H^1_x(\mathbb{R}^3) \times L^2_x(\mathbb{R}^3))$$

to the defocusing cubic nonlinear wave equation on $\mathbb{R}^3$ with initial data $f^\omega$. This is an improvement over the corresponding range $\frac{3}{5} < s < 1$ in Theorem 2.1. In contrast, these arguments which proceed via energy estimates cannot be applied directly in the case $3 < \rho < 5$. In such an argument, one would need bounds on $\|u|^{\rho-1}u\|_{L^1_t L^2_x(\mathbb{R} \times \mathbb{R}^3)} = \|u\|_{L^1_t L^{2\rho}_x(\mathbb{R} \times \mathbb{R}^3)}$ but one no longer has the necessary Sobolev embedding to close the argument.

When $\rho = 5$, the nonlinear wave equation (2.1) is energy-critical and the high-low argument no longer works. However, by establishing probabilistic a priori estimates on the $L^5_t L^{10}_x(\mathbb{R} \times \mathbb{R}^3)$ norm of the free evolution $u^\gamma_f$, we obtain the following probabilistic small data global existence result.

**Theorem 2.2.** Let $\frac{3}{5} \leq s < 1$ and $f = (f_0, f_1) \in H^s_x(\mathbb{R}^3) \times H^{s-1}_x(\mathbb{R}^3)$. Let $\{(h_k, l_k)\}_{k \in \mathbb{Z}^3}$ be a sequence of independent, zero-mean value, real-valued random variables on a probability space $(\Omega, \mathcal{A}, \mathbb{P})$ with distributions $\mu_k$ and $\nu_k$. Assume that there exists $c > 0$ such that

$$\left| \int_{-\infty}^{+\infty} e^{\gamma x} d\mu_k(x) \right| \leq e^{c\gamma^2} \text{ for all } \gamma \in \mathbb{R} \text{ and for all } k \in \mathbb{Z}^3$$

and similarly for $v_k$. Let $f^\omega = (f_0^\omega, f_1^\omega)$ be the associated randomized initial data as defined in (2.4)
and let \( u_f^\omega \) be the associated free evolution as defined in (2.5). There exists \( \Omega_f \subset \Omega \) with

\[
\mathbb{P}(\Omega_f) \geq 1 - Ce^{-c/\|f\|^2_{H^2_x \times H^{-1}_x}} 
\]

(2.7)

for some absolute constants \( C, c > 0 \), such that for every \( \omega \in \Omega_f \) there exists a unique global solution

\[
(u, u_t) \in (u_f^\omega, \partial_t u_f^\omega) + C(\mathbb{R}; \dot{H}^1_x(\mathbb{R}^3) \times L^2_x(\mathbb{R}^3))
\]

to the quintic nonlinear wave equation

\[
\begin{cases}
-u_{tt} + \Delta u = |u|^4 u \text{ on } \mathbb{R} \times \mathbb{R}^3, \\
(u, u_t)|_{t=0} = (f_0^\omega, f_1^\omega).
\end{cases}
\]

(2.8)

Moreover, we have scattering in the sense that for every \( \omega \in \Omega_f \) there exist \((v_1, v_2) \in \dot{H}^1_x(\mathbb{R}^3) \times L^2_x(\mathbb{R}^3)\) such that the free evolution \( v(t) = \cos(t|\nabla|)v_1 + \frac{\sin(t|\nabla|)}{|\nabla|}v_2 \) satisfies

\[
\| (u(t) - u_f^\omega(t) - v(t), \partial_t u(t) - \partial_t u_f^\omega(t) - \partial_t v(t)) \|_{\dot{H}^1_x(\mathbb{R}^3) \times L^2_x(\mathbb{R}^3)} \longrightarrow 0 \text{ as } t \rightarrow \pm \infty.
\]

**Remark 6.** The statement is only meaningful if

\[
\|f\|_{H^2_x \times H^{-1}_x} \leq \left( \frac{c}{\log(C)} \right)^{1/2},
\]

which reflects that Theorem 2.2 is a small data result.

### 2.2 Deterministic and Probabilistic estimates

The following unit-scale Bernstein estimate is a crucial ingredient in the proofs of the probabilistic a priori estimates on the randomized initial data in Section 2.2.2. Its advantage compared to the ordinary Bernstein estimate for the dyadic Littlewood-Paley decomposition is that there is no derivative loss.

**Lemma 2.3** ("Unit-scale Bernstein estimate"). Let \( 2 \leq p_1 \leq p_2 \leq \infty \). There exists a constant \( C \equiv C(p_1, p_2) > 0 \) such that for all \( f \in L^2_x(\mathbb{R}^3) \) and for all \( k \in \mathbb{Z}^3 \)

\[
\| P_k f \|_{L^2_x(\mathbb{R}^3)} \leq C \| P_k f \|_{L^p_x(\mathbb{R}^3)}.
\]

(2.9)
Proof. Let $\eta \in C^\infty_c(\mathbb{R}^3)$ be such that $0 \leq \eta \leq 1$ with $\eta(\xi) = 1$ for $|\xi| \leq 2$ and $\eta(\xi) = 0$ for $|\xi| \geq 3$. For $k \in \mathbb{Z}^3$ define

$$\eta_k(\xi) = \eta(\xi - k) \text{ for } \xi \in \mathbb{R}^3.$$ 

From Young’s inequality with $1 + \frac{1}{p_2} = \frac{1}{q} + \frac{1}{p_1}$ we then obtain

$$\|P_k f\|_{L^p_\mathbb{C}(\mathbb{R}^3)} = \|F^{-1}(\eta_k \hat{\psi} \hat{f})\|_{L^q_\mathbb{C}(\mathbb{R}^3)} = \|F^{-1}(\eta_k \psi_k \hat{f})\|_{L^q_\mathbb{C}(\mathbb{R}^3)} = \|\eta_k * (P_k f)\|_{L^p_\mathbb{C}(\mathbb{R}^3)} \leq \|\hat{\eta}\|_{L^q(\mathbb{R}^3)} \|P_k f\|_{L^p(\mathbb{R}^3)} = \|\hat{\eta}\|_{L^q(\mathbb{R}^3)} \|P_k f\|_{L^p(\mathbb{R}^3)}.$$  

2.2.1 Large deviation estimate

We will record a basic probabilistic result about the randomization procedure. Most of these estimates are consequences of the classical estimates of Paley-Zygmund for random Fourier series on the torus. These estimates were used heavily in the works of Burq and Tzvetkov, see especially [13] for proofs.

**Proposition 2.4** (Large deviation estimate; Lemma 3.1 in [14]). Let $\{h_n\}_{n=1}^\infty$ be a sequence of complex-valued independent random variables with associated distributions $\{\mu_n\}_{n=1}^\infty$ on a probability space $(\Omega, \mathcal{A}, \mathbb{P})$. Assume that the distributions satisfy the property that there exists $c > 0$ such that

$$\left|\int_{-\infty}^{+\infty} e^{\gamma x} d\mu_n(x)\right| \leq e^{c\gamma^2} \text{ for all } \gamma \in \mathbb{R} \text{ and for all } n \in \mathbb{N}.$$ 

Then there exists $c > 0$ such that for every $\lambda > 0$ and every sequence $\{c_n\}_{n=1}^\infty \in \ell^2(\mathbb{N})$ of complex numbers,

$$\mathbb{P}\left(\omega : \sum_{n=1}^\infty c_n h_n(\omega) > \lambda\right) \leq 2e^{-\frac{c^2\lambda^2}{\sum_n |c_n|^2}}.$$ 

As a consequence there exists $C > 0$ such that for every $p \geq 2$ and every $\{c_n\}_{n=1}^\infty \in \ell^2(\mathbb{N})$,

$$\left\|\sum_{n=1}^\infty c_n h_n(\omega)\right\|_{L^p(\Omega)} \leq C \sqrt{p} \left(\sum_{n=1}^\infty |c_n|^2\right)^{1/2}.$$ 

**Remark 2.1.** Gaussian iid or Bernoulli random variables satisfy the exponential moment assumption in the statement of Proposition 2.4.
2.2.2 Averaging effects for the randomized initial data

In the proof of Theorem 2.1 we decompose the randomized initial data into a low frequency and a high frequency component. More precisely, for $N \in \mathbb{N}$ define

$$f_{0, \leq N}^\omega = \sum_{|k| \leq N} h_k(\omega) P_k f_0$$

and $f_{0, > N}^\omega = f_0^\omega - f_{0, \leq N}^\omega$. Similarly, define $f_{1, \leq N}^\omega$ and $f_{1, > N}^\omega$. The next two lemmata establish probabilistic a priori estimates on the low frequency component. See also Proposition 4.4 in [14].

**Lemma 2.5.** Let $3 \leq \rho < 5$. Let $s > 0$ and $f = (f_0, f_1) \in \mathcal{H}^s(\mathbb{R}^3)$. For any $K > 0$ and $N \in \mathbb{N}$, let

$$A_{K,N} := \{ \omega \in \Omega : \| f_{0, \leq N}^\omega \|_{L_2^{\rho+1}(\mathbb{R}^3)}^{(\rho+1)/2} \leq K \}. \quad (2.10)$$

Then there exist constants $C \equiv C(\rho) > 0$ and $c \equiv c(\rho) > 0$ such that

$$\mathbb{P}(A_{K,N}^c) \leq Ce^{-cK^{4/(\rho+1)}/\| f \|_{\mathcal{H}^s(\mathbb{R}^3)}^2} \quad (2.11)$$

for every $K > 0$ and $N \in \mathbb{N}$.

**Proof.** For every $p \geq \rho + 1$ and every $N \in \mathbb{N}$, we bound

$$\| f_{0, \leq N}^\omega \|_{L_p^p(\Omega; L_2^{\rho+1}(\mathbb{R}^3))} = \left\| \sum_{|k| \leq N} h_k(\omega) P_k f_0 \right\|_{L_p^p(\Omega; L_2^{\rho+1}(\mathbb{R}^3))} \leq \left\| \sum_{|k| \leq N} h_k(\omega) P_k f_0 \right\|_{L_2^{\rho+1}(\mathbb{R}^3)} \quad (2.12)$$

$$\lesssim \sqrt{p} \left( \sum_{|k| \leq N} \| P_k f_0(x) \|^2 \right)^{1/2} \| P_k f_0 \|_{L_2^{\rho+1}(\mathbb{R}^3)} \quad (2.13)$$

$$\lesssim \sqrt{p} \left( \sum_{|k| \leq N} \| P_k f_0 \|_{L_2^{\rho+1}(\mathbb{R}^3)}^2 \right)^{1/2} \quad (2.14)$$

Using $p \geq \rho + 1$, we switched the order of integration in (2.12) and used the large deviation estimate from Proposition 2.4 in (2.13). We then used the unit-scale Bernstein estimate (2.9) in (2.14). The
The claim follows from Lemma 2.10.

**Lemma 2.6.** Let $s \in \mathbb{R}$ and $f = (f_0, f_1) \in \mathcal{H}^s(\mathbb{R}^3)$. For any $K > 0$ and $N \in \mathbb{N}$, let

$$B_{K,N} := \{ \omega \in \Omega : \|f_{0,N}^\omega\|_{H^s_2(\mathbb{R}^3)} + \|f_{1,N}^\omega\|_{H^{s-1}_2(\mathbb{R}^3)} \leq K \}. \quad (2.15)$$

Then there exist constants $C \equiv C(s) > 0$ and $c \equiv c(s) > 0$ such that

$$P(B_{K,N}^c) \leq Ce^{-cK^2/\|f\|^2_{\mathcal{H}^s(\mathbb{R}^3)}} \quad (2.16)$$

for every $K > 0$ and $N \in \mathbb{N}$.

**Proof.** For every $p \geq 2$ and every $N \in \mathbb{N}$, using the large deviation estimate Proposition 2.4 we have

$$\|(1 - \Delta)^{s/2} f_{0,N}^\omega \|_{L^p_\omega(\Omega; L^2_\omega(\mathbb{R}^3)))} = \left\| \sum_{|k| \leq N} h_k(\omega)((1 - \Delta)^{s/2} P_k f_0) \right\|_{L^p_\omega(\Omega; L^2_\omega(\mathbb{R}^3)))}$$

$$\lesssim \sqrt{p} \left( \sum_{|k| \leq N} \|((1 - \Delta)^{s/2} P_k f_0)(x)\|^2 \right)^{1/2} \|f_0\|_{H^s_2(\mathbb{R}^3)}$$

$$\lesssim \sqrt{p} \left( \sum_{|k| \leq N} \|((1 - \Delta)^{s/2} P_k f_0)(x)\|^2 \right)^{1/2}$$

$$\lesssim \sqrt{p} \|f_0\|_{H^s_2(\mathbb{R}^3)}.$$  

Similarly, we obtain

$$\|(1 - \Delta)^{(s-1)/2} f_{1,N}^\omega \|_{L^p_\omega(\Omega; L^2_\omega(\mathbb{R}^3)))} \lesssim \sqrt{p} \|f_1\|_{H^{s-1}_2(\mathbb{R}^3)}.$$  

The assertion then follows from Lemma 2.10.

Analogously to (2.5), define the free evolution of the high frequency component of the randomized initial data by

$$u_{f,>N}^\omega = \cos(t|\nabla|)f_{0,>N}^\omega + \frac{\sin(t|\nabla|)}{|\nabla|} f_{1,>N}^\omega.$$  

In the next lemma we prove probabilistic a priori estimates which exhibit a decay in $N$ on certain space-time norms of the free evolution $u_{f,>N}^\omega$ once one restricts to suitable subsets of the probability space $\Omega$. The decay is ultimately the reason why we obtain an improved range of regularity.
exponents $s$ in Theorem 2.1 compared to [34, Theorem 1.2] and results from the use of the unit-scale Bernstein estimate (2.9) and the Strichartz estimate (A.1).

**Lemma 2.7.** Let $T > 0$ and $3 \leq \rho < 5$. Let $0 < s < 1$ and $0 < \varepsilon < \min(\frac{s}{2}, \frac{1}{2}(1 - \frac{1}{\rho}))$. Let $f = (f_0, f_1) \in \mathcal{H}^s(\mathbb{R}^3)$. For any $K > 0$ and $N \geq 3$, let

$$D_{K,N,\varepsilon} := \{\omega \in \Omega : N^{s-2\varepsilon}||u^\omega_{f,j}>N||_{L_t^{1/\varepsilon}L_x^{2\rho}([0,T] \times \mathbb{R}^3)} \leq K\}.$$  

(2.17)

Then there exist constants $C \equiv C(s, \rho, \varepsilon) > 0$ and $c \equiv c(s, \rho, \varepsilon) > 0$ such that

$$\mathbb{P}(D_{K,N,\varepsilon}^c) \leq Ce^{-cK^2/\|f\|^2_{\mathcal{H}^s(\mathbb{R}^3)}}$$  

(2.18)

for every $K > 0$ and $N \geq 3$.

**Proof.** Set $r(\varepsilon) = \frac{2}{1-2\varepsilon}$. Then the exponent pair $(\frac{s}{\varepsilon}, r(\varepsilon))$ is wave-admissible and Strichartz-admissible at regularity $\gamma = 2\varepsilon$. For every $p \geq \max(\frac{1}{\varepsilon}, 2\rho)$ and every $N \geq 3$, we now estimate

$$\|u^\omega_{f,j}>N\|_{L_x^p(\Omega; L_t^{1/\varepsilon} L_x^{2\rho}([0,T] \times \mathbb{R}^3))} \leq \sqrt{P}\left(\sum_{|k| > N} \left\|\cos(t|\nabla|)P_k f_0\right\|_{L_t^{1/\varepsilon} L_x^{2\rho}([0,T] \times \mathbb{R}^3)}\right)^{1/2}$$

$$+ \sqrt{P}\left(\sum_{|k| > N} \left\|\sin(t|\nabla|)P_k f_1\right\|_{L_t^{1/\varepsilon} L_x^{2\rho}([0,T] \times \mathbb{R}^3)}\right)^{1/2}$$

$$\leq \sqrt{P}\left(\sum_{|k| > N} \left\|P_k f_0\right\|_{H_x^{s-\varepsilon}(\mathbb{R}^3)}\right)^{1/2} + \sqrt{P}\left(\sum_{|k| > N} \left\|P_k f_1\right\|_{H_x^{s-\varepsilon}(\mathbb{R}^3)}\right)^{1/2}$$

$$\leq \sqrt{\bar{P}}N^{-(s-2\varepsilon)} \left(\sum_{|k| > N} \left\|P_k f_0\right\|_{H_x^s(\mathbb{R}^3)}\right)^{1/2} + \sqrt{\bar{P}}N^{-(s-2\varepsilon)} \left(\sum_{|k| > N} \left\|P_k f_1\right\|_{H_x^{s-1}(\mathbb{R}^3)}\right)^{1/2}$$

Using $p \geq \max(\frac{1}{\varepsilon}, 2\rho)$, we switched the order of integration and used the large deviation esti-
mate Proposition 2.4 in (2.19). The assumption $f \in \mathcal{H}^{s}(\mathbb{R}^{3})$ together with Plancherel's theorem guarantees that

$$\left(\sum_{k \in \mathbb{Z}^{3}} |(\cos(t|\nabla|)P_{k}f_{0})(x)|^{2}\right)^{1/2} < \infty$$

and

$$\left(\sum_{k \in \mathbb{Z}^{3}} \left|\frac{\sin(t|\nabla|)}{|\nabla|}P_{k}f_{1}(x)\right|^{2}\right)^{1/2} < \infty$$

for almost every $x \in \mathbb{R}^{3}$ and for every $t \in [0, T]$, allowing us to apply Proposition 2.4. We use the unit-scale Bernstein estimate (2.9) in (2.20), noting that $r(\varepsilon) \leq 2\rho$, and then apply the Strichartz estimates (A.1) at regularity $\gamma = 2\varepsilon$ in (2.21). In (2.22) we may estimate $\|P_{k}f_{1}\|_{\mathcal{H}_{x}^{s-1}(\mathbb{R}^{3})} \lesssim \|P_{k}f_{1}\|_{\mathcal{H}_{x}^{s-1}(\mathbb{R}^{3})}$ uniformly for all $|k| > N \geq 3$ even though $s - 1 < 0$, since $\mathcal{F}(P_{k}f_{1})(\xi) = 0$ for $|\xi| < 1$ for all $|k| > N \geq 3$ due to the support properties of the unit-scale projections. The claim then follows from Lemma 2.10. □

2.3 Proof of Theorem 2.1

This section is devoted to the proof of the following proposition, which immediately implies Theorem 2.1.

Proposition 2.8. Let $T > 0$. Let $3 \leq \rho < 5$ and

$$\frac{\rho^{3} + 5\rho^{2} - 11\rho - 3}{9\rho^{2} - 6\rho - 3} < s < 1.$$  

Fix $f = (f_{0}, f_{1}) \in \mathcal{H}^{s}_{x}(\mathbb{R}^{3}) \times \mathcal{H}_{x}^{s-1}(\mathbb{R}^{3})$. Let $f^{\omega} = (f_{0}^{\omega}, f_{1}^{\omega})$ be the associated randomized initial data as defined in (2.4) and $u^{\omega}_{f}$ the corresponding free evolution as defined in (2.5). Then there exists $\Omega_{T} \subset \Omega$ with $P(\Omega_{T}) = 1$ such that for every $\omega \in \Omega_{T}$ there exists a unique solution

$$(u, u_{t}) \in (u^{\omega}_{f}, \partial_{t}u^{\omega}_{f}) + C([0, T]; \mathcal{H}^{1}_{x}(\mathbb{R}^{3}) \times L^{2}_{x}(\mathbb{R}^{3}))$$

(2.23)

to the nonlinear wave equation

$$\begin{cases}
-u_{tt} + \Delta u = |u|^{s-1}u & \text{on } [0, T] \times \mathbb{R}^{3}, \\
(u, u_{t})|_{t=0} = (f_{0}^{\omega}, f_{1}^{\omega}).
\end{cases}$$

(2.24)

Here, uniqueness only holds in a mild sense; see Remark 7.

Proof of Theorem 2.1. We only present the argument to construct solutions that exist for all pos-
itive times since the argument for negative times is similar. Define $T_j = j$ for $j \in \mathbb{N}$ and set $\Sigma = \bigcap_{j=1}^{\infty} \Omega_{T_j}$. By Proposition 2.8, we have that $\mathbb{P}(\Sigma) = 1$ and for every $\omega \in \Sigma$, we have global existence for (2.6) on the time interval $[0, \infty)$.

In the proof of Proposition 2.8 we will repeatedly use the following probabilistic low regularity local well-posedness result whose proof we defer to the end of this section. It is here that we invoke the crucial averaging effects for the free evolution $u^\omega_{f,N}$ from Lemma 2.7. We introduce the notation

$$q(\rho) = \frac{2\rho}{\rho - 3}, \quad \alpha(\rho) = \frac{5 - \rho}{2}.$$ 

**Lemma 2.9.** Let $3 \leq \rho < 5$ and $0 < s < 1$. Let $0 < \varepsilon < \min\left(\frac{3}{2}, \frac{1}{2}(1 - \frac{1}{\rho})\right)$ and $K > 0$ be fixed. For $N \geq 3$ and $0 < c < 1$ set $T_1 = c(KN^{1-s})^{(\rho-1)/\alpha(\rho)}$. Let $v : [0, T_1] \times \mathbb{R}^3 \to \mathbb{C}$ satisfy

$$\|v\|_{L^q_t L^{2\rho}_x([0,T_1] \times \mathbb{R}^3)} \leq CKN^{1-s}$$

and let $\omega \in \Omega$ be such that

$$\|u^\omega_{f,N}\|_{L^1_t L^{2\rho}_x([0,T_1] \times \mathbb{R}^3)} \leq KN^{-s+2\varepsilon}.$$  

(2.25)

(2.26)

For $0 < c < 1$ sufficiently small (independent of the size of $K$ and $N$) and $N \equiv N(K)$ sufficiently large, there exists a unique solution

$$(\tilde{w}, \tilde{w}_t) \in C([0, T_1]; L^1_x(\mathbb{R}^3)) \cap L^q_t L^{2\rho}_x([0, T_1] \times \mathbb{R}^3) \times C([0, T_1]; L^2_x(\mathbb{R}^3))$$

to the nonlinear wave equation

$$\begin{cases}
-\tilde{w}_{tt} + \Delta \tilde{w} = |v + u^\omega_{f,N} + \tilde{w}|^{\rho-1}(v + u^\omega_{f,N} + \tilde{w}) - |v|^{\rho-1}v \text{ on } [0, T_1] \times \mathbb{R}^3, \\
(\tilde{w}, \tilde{w}_t)|_{t=0} = (0, 0),
\end{cases}$$

(2.27)

satisfying

$$\|\tilde{w}(T_1)\|_{L^1_x(\mathbb{R}^3)} + \|\tilde{w}_t(T_1)\|_{L^2_x(\mathbb{R}^3)} + \|\tilde{w}(T_1)\|_{L^{p+1}_x(\mathbb{R}^3)} \lesssim T_1^{1 - \frac{\rho-1}{q(\rho)-\varepsilon}} KN^{2\varepsilon+(1-s)\rho-1}.$$  

(2.28)

We now present the proof of Proposition 2.8.
Proof of Proposition 2.8. The bulk of the proof is devoted to the construction of subsets $\Omega_{K,T} \subset \Omega$ for every $K \in \mathbb{N}$ such that for every $\omega \in \Omega_{K,T}$ there exists a unique solution of the form (2.23) to (2.24) and such that
\begin{equation}
\mathbb{P}(\Omega_{K,T}) \leq Ce^{-CK^4/(p+1)/\|f\|^2_{H^s(R^3)}}. \tag{2.29}
\end{equation}
We then set $\Omega_T = \bigcup_{K=1}^{\infty} \Omega_{K,T}$ and conclude from (2.29) that $\mathbb{P}(\Omega_T) = 1$, which completes the proof of Proposition 2.8.

In what follows, let $K \in \mathbb{N}$ be fixed. Let $N \equiv N(K, T) \in \mathbb{N}$ be sufficiently large, to be fixed later in the proof. We also make use of a fixed small parameter $0 < \varepsilon < \min(\frac{3}{2}, \frac{1}{2}(1 - \frac{1}{p}))$ whose value depends on $\rho$ and $s$, but is independent of $K$ and $N$, and is specified further below. We define
\begin{equation}
\Omega_{K,T} = A_{K,N} \cap B_{K,N} \cap D_{K,N,\varepsilon},
\end{equation}
where these sets are as in (2.10), (2.15) and (2.17). The estimate (2.29) then follows from (2.11), (2.16) and (2.18).

From now on we only consider $\omega \in \Omega_{K,T}$. It suffices to show that there exists a unique solution
\begin{equation}
(u, u_t) \in (u_{\tilde{q}}^\omega, \partial_t u_{\tilde{q}}^\omega) + C([0, T]; \dot{H}^1_x(R^3) \times L^2_x(R^3))
\end{equation}
since by a persistence of regularity argument, one has $u \in u_{\tilde{q}}^\omega + C([0, T]; H^1_x(R^3))$.

We first construct a solution $u^{(1)} = v^{(1)} + w^{(1)}$ to (2.24) on a small time interval $[0, T_1]$ with $0 < T_1 < 1$ to be fixed later and where $v^{(1)}$ solves the following nonlinear wave equation with low frequency initial data
\begin{equation}
\begin{cases}
-v^{(1)}_{tt} + \Delta v^{(1)} = |v^{(1)}|^{p-1}v^{(1)} & \text{on } [0, T_1] \times \mathbb{R}^3, \\
(v^{(1)}, v^{(1)}_t)|_{t=0} = (f_{0,0}^{\omega}, f_{T,1}^{\omega}).
\end{cases}
\tag{2.30}
\end{equation}
Note that the initial data $(f_{0,0}^{\omega}, f_{T,1}^{\omega})$ lies in $\dot{H}^1_x(R^3) \times L^2_x(R^3)$, since by (2.15)
\begin{equation}
\|f_{0,0}^{\omega}\|_{\dot{H}^1_x(R^3)} + \|f_{T,1}^{\omega}\|_{L^2_x(R^3)} \lesssim N^{1-s}(\|f_{0,0}^{\omega}\|_{H^s_x(R^3)} + \|f_{T,1}^{\omega}\|_{H^{s-1}_x(R^3)}) \lesssim KN^{1-s}. \tag{2.31}
\end{equation}
Thus, by the deterministic global existence theory [25, Proposition 3.2] there exists a unique solution $(v^{(1)}, v^{(1)}_t) \in C([0, T_1]; \dot{H}^1_x(R^3) \times L^2_x(R^3))$ to (2.30). Moreover, we have energy conservation
since \( \|f_{0,N}\|_{L_x^{p+1}(\mathbb{R}^3)}^p \leq K^2 \) by (2.10). Hence, for all \( t \in [0, T_1] \)

\[
E(v^{(1)}(t)) := \frac{1}{2} \|v^{(1)}(t)\|_{H^1_t(\mathbb{R}^3)}^2 + \frac{1}{2} \|v_t^{(1)}(t)\|_{L_x^p(\mathbb{R}^3)}^2 + \frac{1}{\rho + 1} \|v^{(1)}(t)\|_{L_x^{p+1}(\mathbb{R}^3)}^{p+1}
\]

\[
= \frac{1}{2} \|f_{0,N}^{(1)}\|_{H^1_t(\mathbb{R}^3)}^2 + \frac{1}{2} \|f_t^{(1)}\|_{L_x^p(\mathbb{R}^3)}^2 + \frac{1}{\rho + 1} \|f_{0,N}^{(1)}\|_{L_x^{p+1}(\mathbb{R}^3)}^{p+1}
\]

\[
\lesssim (KN^{1-s})^2.
\]

We note that the exponent pair \((q(\rho), 2\rho)\) is Strichartz-admissible at regularity \(\gamma = 1\). Using Strichartz estimates (A.1) and (2.31), we find

\[
\|v^{(1)}\|_{L_t^{q(\rho)}L_x^{2\rho}([0,T_1] \times \mathbb{R}^3)} \lesssim \|v^{(1)}(0)\|_{H^1_t(\mathbb{R}^3)} + \|v_t^{(1)}(0)\|_{L_x^p(\mathbb{R}^3)} + \|v^{(1)}|_{\rho-1} v^{(1)}\|_{L_t^1 L_x^p([0,T_1] \times \mathbb{R}^3)}
\]

\[
\lesssim KN^{1-s} + T_1^{\sigma(\rho)} \|v^{(1)}\|_{L_t^{q(\rho)}L_x^{2\rho}([0,T_1] \times \mathbb{R}^3)}.
\]

Hence, choosing \( T_1 = c(KN^{1-s})^{-\frac{q(\rho)-1}{\sigma(\rho)}} \) with \( 0 < c < 1 \) sufficiently small (independently of the size of \( K \) and \( N \)), we obtain

\[
\|v^{(1)}\|_{L_t^{q(\rho)}L_x^{2\rho}([0,T_1] \times \mathbb{R}^3)} \lesssim KN^{1-s}.
\]

Next, we consider the nonlinear wave equation that \( w^{(1)} = u^{(1)} - v^{(1)} \) must satisfy, namely

\[
\begin{cases}
-w_{tt}^{(1)} + \Delta w^{(1)} = |v^{(1)} + u^{(1)}|^{\rho-1}(v^{(1)} + u^{(1)}) - |v^{(1)}|^{\rho-1}v^{(1)} & \text{on } [0, T_1] \times \mathbb{R}^3, \\
(w^{(1)}, w_t^{(1)})|_{t=0} = (f_{0,N}^{(1)}, f_{t,N}^{(1)}).
\end{cases}
\]

We look for a solution of the form

\[
w^{(1)} = u^{(1)}_{f,N} + \tilde{w}^{(1)},
\]

where \( \tilde{w}^{(1)} \) solves the following initial value problem on \([0, T_1] \times \mathbb{R}^3\)

\[
\begin{cases}
-\tilde{w}_{tt}^{(1)} + \Delta \tilde{w}^{(1)} = |v^{(1)} + u^{(1)}|^{\rho} + \tilde{w}^{(1)}|^{\rho-1}(v^{(1)} + u^{(1)} + \tilde{w}^{(1)}) - |v^{(1)}|^{\rho-1}v^{(1)}, \\
(\tilde{w}^{(1)}, \tilde{w}_t^{(1)})|_{t=0} = (0, 0).
\end{cases}
\]

Using (2.33) and the averaging effects (2.17) for the free evolution \( u^{(1)}_{f,N} \), Lemma 2.9 yields a unique solution

\[
(\tilde{w}^{(1)}, \tilde{w}_t^{(1)}) \in C([0, T_1]; \dot{H}^1_t(\mathbb{R}^3)) \cap L_t^{q(\rho)}L_x^{2\rho}([0, T_1] \times \mathbb{R}^3) \times C([0, T_1]; L_x^2(\mathbb{R}^3)).
\]
to (2.34) provided $0 < c < 1$ in the definition of $T_1$ is chosen sufficiently small (independently of the size of $K$ and $N$) and $N$ is chosen sufficiently large. Moreover, we have
\[
\|\tilde{w}^{(1)}(T_1)\|_{\dot{H}_x^1(\mathbb{R}^3)} + \|\tilde{w}^{(1)}_t(T_1)\|_{L_x^2(\mathbb{R}^3)} + \|\tilde{w}^{(1)}(T_1)\|_{L_x^{p+1}(\mathbb{R}^3)} \\
\lesssim T_1^{-1} \frac{e^{\frac{1}{3}N}}{\rho^2} - \varepsilon K^p N^{2\varepsilon + (1-\varepsilon)\rho - 1}. \tag{2.35}
\]

In the next step we build a solution $u^{(2)} = v^{(2)} + w^{(2)}$ to (2.24) on the time interval $[T_1, 2T_1]$. As before, we would like to construct $v^{(2)}$ using the deterministic global existence theory at the energy level and construct $w^{(2)}$ through the probabilistic local well-posedness result from Lemma 2.9. To this end, we note that $\tilde{w}^{(1)}$ is comprised of the free evolution $\tilde{w}^{(1)}_{f, > N}$, which is at low regularity, and the nonlinear component $\tilde{w}^{(1)}$, which lies in the energy space by (2.35). As initial data for $v^{(2)}$ at time $T_1$ we therefore take the sum of $v^{(1)}(T_1)$ and $\tilde{w}^{(1)}(T_1)$, and consider
\[
\begin{cases}
-v^{(2)}_{tt} + \Delta v^{(2)} = |v^{(2)}|^{\rho-1} v^{(2)} & \text{on } [T_1, 2T_1] \times \mathbb{R}^3, \\
(v^{(2)}, v^{(2)}_t)|_{t=T_1} = (v^{(1)}(T_1) + \tilde{w}^{(1)}(T_1), v^{(1)}_t(T_1) + \tilde{w}^{(1)}_t(T_1)).
\end{cases} \tag{2.36}
\]

Once again, by the deterministic global theory, this initial value problem has a unique solution
\[
(v^{(2)}, v^{(2)}_t) \in C([T_1, 2T_1]; H_x^1(\mathbb{R}^3) \times L_x^2(\mathbb{R}^3)).
\]

Moreover, we obtain bounds on the $L_t^q(\rho)L_x^{2\rho}([T_1, 2T_1] \times \mathbb{R}^3)$-norm of $v^{(2)}$ as in (2.33). Using these bounds and the averaging effects (2.17), we apply Lemma 2.9 to solve the difference equation for $w^{(2)} = u^{(2)} - v^{(2)}$ on $[T_1, 2T_1] \times \mathbb{R}^3$,
\[
\begin{cases}
-w^{(2)}_{tt} + \Delta w^{(2)} = |v^{(2)} + w^{(2)}|^{\rho-1} (v^{(2)} + w^{(2)}) - |v^{(2)}|^{\rho-1} v^{(2)}, \\
(w^{(2)}, w^{(2)}_t)|_{t=T_1} = (u^{(f), > N}_f(T_1), \partial_t u^{(f), > N}_f(T_1)).
\end{cases} \tag{2.37}
\]

Note that Lemma 2.9 can also be applied on the time interval $[T_1, 2T_1]$ by time translation. We therefore find a solution $w^{(2)} = w^{(2)}_{f, > N} + \tilde{w}^{(2)}$ to (2.37) with
\[
(\tilde{w}^{(2)}, \tilde{w}^{(2)}_t) \in C([T_1, 2T_1]; H_x^1(\mathbb{R}^3)) \cap L_t^q(\rho)L_x^{2\rho}([T_1, 2T_1] \times \mathbb{R}^3) \times C([T_1, 2T_1]; L_x^2(\mathbb{R}^3)).
\]

In order to obtain a solution $u$ to (2.24) on the whole time interval $[0, T]$, we iterate this procedure on consecutive intervals for $\lceil \frac{T}{T_1} \rceil$ times. At every step we redistribute the data as in
To make the process uniform and thus reach the time $T$, we have to take into account that this redistribution increases the energy at each step. To estimate this growth, we invoke energy conservation for the solution to (2.30). We have

$$E(v^{(1)}(T_1) + \tilde{w}^{(1)}(T_1)) = E(v^{(1)}(0)) + \left( E(v^{(1)}(T_1) + \tilde{w}^{(1)}(T_1)) - E(v^{(1)}(T_1)) \right),$$

and

$$E(v^{(1)}(T_1) + \tilde{w}^{(1)}(T_1)) - E(v^{(1)}(T_1))$$

$$\lesssim \|v^{(1)}(T_1)\|_{H^1(\mathbb{R}^3)} \|\tilde{w}^{(1)}(T_1)\|_{H^1(\mathbb{R}^3)} + \|v^{(1)}(T_1)\|_{L^2(\mathbb{R}^3)} \|\tilde{w}^{(1)}(T_1)\|_{L^2(\mathbb{R}^3)}$$

$$+ \|v^{(1)}(T_1)\|_{L^{p+1}(\mathbb{R}^3)} \|\tilde{w}^{(1)}(T_1)\|_{L^{p+1}(\mathbb{R}^3)} + E(\tilde{w}^{(1)}(T_1))$$

$$\lesssim E(v^{(1)}(0))^\frac{1}{2} \left( \|\tilde{w}^{(1)}(T_1)\|_{H^1(\mathbb{R}^3)} + \|\tilde{w}^{(1)}(T_1)\|_{L^2(\mathbb{R}^3)} \right)$$

$$+ E(v^{(1)}(0))^\frac{1}{2} \|\tilde{w}^{(1)}(T_1)\|_{L^{p+1}(\mathbb{R}^3)} + E(\tilde{w}^{(1)}(T_1)).$$

The term $E(v^{(1)}(0))^\frac{1}{2} \|\tilde{w}^{(1)}(T_1)\|_{L^{p+1}(\mathbb{R}^3)}$ gives the largest contribution to the energy increment (2.38). In light of (2.32) and (2.35), we must ensure that

$$\frac{T}{T_1} \left( (KN^{1-s})^2 \right) \frac{\rho - 1}{\rho} \left( T_1^{1 - \frac{\rho - 1}{\rho}} K^\rho N^{2\epsilon + (1-s)\rho - 1} \right) \lesssim (KN^{1-s})^2.$$

Inserting $T_1 = c(KN^{1-s})^{-\frac{\rho - 1}{\rho}}$, this is equivalent to

$$T_c \frac{\rho - 1}{\rho} K \frac{\rho^2 - \rho - 3}{(5 - \rho)(\rho + 1)} \epsilon^{2(\rho - 1)} N^{\frac{\rho^2 + 2\rho^2 - 11\rho - 3}{(5 - \rho)(\rho + 1)} + \epsilon((1-s)2(\rho - 1) + 2)} \lesssim 1. \quad (2.39)$$

For any $s > \frac{\rho + 5\rho^2 - 11\rho - 3}{9\rho^2 - 6\rho - 3}$, we can make the exponent on $N$ in (2.39) negative by fixing $\epsilon > 0$ sufficiently small at the beginning (depending on the values of $s$ and $\rho$, which are fixed during the course of the argument). Hence, taking $N \equiv N(K, T)$ sufficiently large we can ensure that condition (2.39) is satisfied. This completes the proof of Proposition 2.8. \qed

We now present the proof of Lemma 2.9.

**Proof of Lemma 2.9.** In this proof we are working on $[0, T_1] \times \mathbb{R}^3$ and will omit this notation. For $\tilde{w} \in L_t^{q(\rho)} L_x^p$ define the map

$$\Phi(\tilde{w})(t) = - \int_0^t \frac{\sin((t-s)|\nabla|)}{|\nabla|} \left( |v + u_{M,>N} + \tilde{w}|^{p-1} (v + u_{M,>N} + \tilde{w}) - |v|^{p-1} v \right)(s) \, ds.$$
Using the general inequality

\[ |z_1|^{\rho-1}z_1 - |z_2|^{\rho-1}z_2| \lesssim |z_1 - z_2|(|z_1|^{\rho-1} + |z_2|^{\rho-1}) \text{ for any } z_1, z_2 \in \mathbb{C} \]

and the Strichartz estimates (A.1) at regularity \( \gamma = 1 \) we bound

\[
\|\Phi(\tilde{\eta})\|_{L^\infty_t H^s_x} + \|\partial_t \Phi(\tilde{\eta})\|_{L^\infty_t L^2_x} + \|\Phi(\tilde{\eta})\|_{L^q_t L^2_x} \\
\lesssim \left\| v + u_{j,>N} + \tilde{\eta} \right\|^{\rho-1} \left\| (v + u_{j,>N} + \tilde{\eta}) - |v|^{\rho-1}v \right\|_{L^1_t L^2_x} \\
\lesssim \left\| u_{j,>N} + \tilde{\eta} \right\|^{\rho-1} \left\| (v + u_{j,>N} + \tilde{\eta}) - |v|^{\rho-1}v \right\|_{L^1_t L^2_x} \\
\lesssim \left\| (u_{j,>N} + \tilde{\eta}) - |v|^{\rho-1}u \right\|_{L^1_t L^2_x} + \left\| u_{j,>N} - |v|^{\rho-1}u \right\|_{L^1_t L^2_x} \\
= I + II + III + IV + V.
\]

We now estimate the terms \( I - V \) separately.

**Term I:** By Hölder's inequality in time and (2.26), we obtain

\[
\left\| u_{j,>N} \right\|_{L^\rho_t L^2_x} \leq T_1^{1-\varepsilon} \left\| u_{j,>N} \right\|_{L^{\alpha(t)/2}_{t,\varepsilon} L^2_x} \leq T_1^{1-\varepsilon} (KN^{-s+2\varepsilon})^\rho.
\]

**Term II:** By Hölder's inequality in time, we have

\[
\left\| \tilde{\eta} \right\|_{L^\rho_t L^2_x} \leq T_1^{\alpha(t)} \left\| \tilde{\eta} \right\|_{L^{\alpha(t)} L^2_x}.
\]

**Term III:** Using (2.25) and (2.26), we find

\[
\left\| (u_{j,>N} - |v|^{\rho-1}u) \right\|_{L^1_t L^2_x} \leq \left\| u_{j,>N} \right\|_{L^\rho_t L^2_x} \left\| v \right\|^{\rho-1}_{L^\rho_t L^2_x} \\
\leq T_1^{1/\rho-\varepsilon} \left\| u_{j,>N} \right\|_{L^{\alpha(t)/2}_{t,\varepsilon} L^2_x} T_1^{\alpha(t)(\rho-1)/\rho} \left\| v \right\|^{\rho-1}_{L^{\alpha(t)} L^2_x} \\
\lesssim T_1^{1+\alpha(t)(\rho-1)/\rho-\varepsilon} (KN^{-s+2\varepsilon}) (KN^{-s})^{\rho-1} \\
\lesssim T_1^{1-\frac{\varepsilon}{\rho-1}-\varepsilon} K^{\rho} N^{2\varepsilon+(1-s)(\rho-1)}.
\]

42
Term IV: By (2.25) we have

\[
\|\tilde{w}\|^{\rho-1}_{L_1^1 L_2^2} \leq T_1^{\alpha(\rho)/\rho} \|\tilde{w}\|^{\rho-1}_{L_1^\alpha(\rho) L_2^2} T_1^{\alpha(\rho)(\rho-1)/\rho} \|\tilde{w}\|^{\rho-1}_{L_1^\alpha(\rho) L_2^2} \\
\leq T_1^{\alpha(\rho)} \|\tilde{w}\|^{\rho-1}_{L_1^\alpha(\rho) L_2^2} (KN^{1-s})^{\rho-1}.
\]

Term V: Using (2.26) we bound

\[
\|u_{\rho, \nu, N}^{\rho-1}\|^{\rho-1}_{L_1^1 L_2^2} \leq \|u_{\rho, \nu, N}^{\rho-1}\|^{\rho-1}_{L_1^{\alpha(\rho)} L_2^2} \|\tilde{w}\|^{\rho-1}_{L_1^\alpha(\rho) L_2^2} \\
\leq T_1^{(\rho-1)/(\rho-1)\epsilon} \|u_{\rho, \nu, N}^{\rho-1}\|^{\rho-1}_{L_1^{\alpha(\rho)} L_2^2} T_1^{\alpha(\rho)/\rho} \|\tilde{w}\|^{\rho-1}_{L_1^\alpha(\rho) L_2^2} \\
= T_1^{\rho-1+\alpha(\rho)-\rho(\rho-1)\epsilon} \|u_{\rho, \nu, N}^{\rho-1}\|^{\rho-1}_{L_1^{\alpha(\rho)} L_2^2} \|\tilde{w}\|^{\rho-1}_{L_1^\alpha(\rho) L_2^2} \\
\leq T_1^{1-\frac{1}{q(\rho)}-(\rho-1)\epsilon} (KN^{-s+2\epsilon})^{\rho-1} \|\tilde{w}\|^{\rho-1}_{L_1^\alpha(\rho) L_2^2}.
\]

Collecting terms we obtain

\[
\|\Phi(\tilde{w})\|_{L_1^{\infty} H_2^\infty} + \|\partial_t \Phi(\tilde{w})\|_{L_1^{\infty} L_2^2} + \|\Phi(\tilde{w})\|_{L_1^{\alpha(\rho)} L_2^2} \\
\leq T_1^{-\rho e} KN^{-\rho s+2\rho e} + T_1^{\alpha(\rho)} \|\tilde{w}\|^{\rho}_{L_1^\alpha(\rho) L_2^2} \\
+ T_1^{1-\frac{1}{q(\rho)}-\epsilon} KN^2\epsilon+(1-s)\rho-1 + T_1^{\alpha(\rho)} (KN^{1-s})^{\rho-1} \|\tilde{w}\|_{L_1^{\alpha(\rho)} L_2^2} \\
+ T_1^{1-\frac{1}{q(\rho)}-(\rho-1)\epsilon} (KN^{-s+2\epsilon})^{\rho-1} \|\tilde{w}\|_{L_1^{\alpha(\rho)} L_2^2}.
\]

It follows that by choosing \(0 < c < 1\) sufficiently small (independently of \(K\) and \(N\)) and \(N\) sufficiently large, \(\Phi\) maps a ball \(B\) of radius \(R(K, N) > 0\) with respect to the \(L_1^{\alpha(\rho)} L_2^2\)-norm into itself, where

\[
R(K, N) \lesssim T_1^{1-\frac{1}{q(\rho)}-\epsilon} KN^2\epsilon+(1-s)\rho-1.
\]

In a similar vein, we show that \(\Phi\) is a contraction on \(B\) with respect to the \(L_1^{\alpha(\rho)} L_2^2\)-norm for \(0 < c < 1\) sufficiently small and \(N\) sufficiently large. Thus, \(\Phi\) has a unique fixed point \(\tilde{w} \in B\) and \(\tilde{w}\) is the unique solution to (2.27).

In order to obtain (2.28) it remains to estimate the \(L_2^{\rho+1}(\mathbb{R}^3)\)-norm of \(\tilde{w}\) at time \(T_1\). By Sobolev
embedding and Minkowski's integral inequality we find

\[
\|\tilde{w}(T_1)\|_{L^{p+1}_x(R^3)} \\
\lesssim \|\tilde{w}(T_1)\|_{H^1_x(R^3)} \\
\lesssim \left\| \int_0^{T_1} \frac{\sin((T_1-s)|\nabla|)}{|\nabla|} \left( |v + u_{f,>N}^\omega + \tilde{w}|^{p-1}(v + u_{f,>N}^\omega + \tilde{w}) - |v|^{p-1}v(s) \right) ds \right\|_{H^1_x(R^3)} \\
\lesssim \int_0^{T_1} \left\| \frac{\sin((T_1-s)|\nabla|)}{|\nabla|} \left( |v + u_{f,>N}^\omega + \tilde{w}|^{p-1}(v + u_{f,>N}^\omega + \tilde{w}) - |v|^{p-1}v(s) \right) \right\|_{H^1_x(R^3)} ds \\
\lesssim (1 + T_1) \left\| |v + u_{f,>N}^\omega + \tilde{w}|^{p-1}(v + u_{f,>N}^\omega + \tilde{w}) - |v|^{p-1}v \right\|_{L^2_x(R^3)} \\
\lesssim \left\| |v + u_{f,>N}^\omega + \tilde{w}|^{p-1}(v + u_{f,>N}^\omega + \tilde{w}) - |v|^{p-1}v \right\|_{L^1_xL^2_x},
\]

where we used that \( T_1 \leq 1 \). From (2.40), (2.41) and (2.42) we infer

\[
\|\tilde{w}(T_1)\|_{L^{p+1}_x(R^3)} \lesssim R(K, N) \lesssim T_1^{1-\frac{\varepsilon}{8(\rho)} + \varepsilon} K^\rho N^{2\varepsilon + (1-s)\rho-1}.
\]

This completes the proof of Lemma 2.9. \( \square \)

Finally, we address the uniqueness statements in Theorem 2.1 and Proposition 2.8.

**Remark 7.** Analogously to [19, Remark 1.2], uniqueness for the "low frequency part" \((v(j), v_t^{(j)})\) in the \( j \)-th step holds in the space \( C([((j-1)T_1, jT_1]; H^1_x(R^3) \times L^2_x(R^3)) \). However, uniqueness for the "high frequency part" \((w(j), w_t^{(j)})\) in the \( j \)-th step only holds in the ball centered at \( v_t^{(j),>N}(t) \) of small radius in \( L^2_t(L^2_x) \) \((([j-1]T_1, jT_1] \times R^3) \).

### 2.4 Proof of Theorem 2.2

In this section we prove the almost sure global existence result for the quintic nonlinear wave equation from Theorem 2.2.

**Proof of Theorem 2.2.** We first derive probabilistic a priori estimates on the \( L^5_tL^{10}_x(R \times R^3) \) norm of the free evolution \( u^f \) and then use these to construct global solutions to (2.8) through a suitable fixed point argument.
Since we have by assumption that $\frac{2}{3} \leq s < 1$, there exists $\frac{10}{3} \leq r < 10$ such that the exponent pair $(5, r)$ is wave-admissible and Strichartz-admissible at regularity $s$. Similarly to the proof of Lemma 2.7, using the large deviation estimate Proposition 2.4, the unit-scale Bernstein estimate (2.9) and the Strichartz estimate (A.1) at regularity $s$, we obtain for any $p \geq 10$

$$\| u^\omega_f \|_{L^p_t L^r_x(\mathbb{R} \times \mathbb{R}^3)} \leq \sqrt{p} \left( \sum_{k \in \mathbb{Z}^3} \| \cos(t|\nabla|) P_k f_0 \|_{L^5_t L^{10}_x(\mathbb{R} \times \mathbb{R}^3)}^2 \right)^{1/2} + \sqrt{p} \left( \sum_{k \in \mathbb{Z}^3} \| \sin(t|\nabla|) P_k f_1 \|_{L^5_t L^{10}_x(\mathbb{R} \times \mathbb{R}^3)}^2 \right)^{1/2}$$

Then Lemma 2.10 implies that there exist absolute constants $C, c > 0$ such that

$$\mathbb{P}(\omega \in \Omega : \| u^\omega_f \|_{L^p_t L^r_x(\mathbb{R} \times \mathbb{R}^3)} > K) \leq C e^{-c K^2 \| f \|_{H^{1/2}_x \dot{H}^{3/2-1}_x}^2}$$

for every $K > 0$.

We now look for global solutions of the form $u = u^\omega_f + w$ to (2.8), where $w$ has to satisfy the nonlinear wave equation

$$\begin{cases} -w_{tt} + \Delta w = |u^\omega_f + w|^4 (u^\omega_f + w) & \text{on } \mathbb{R} \times \mathbb{R}^3, \\ (w, w_t)|_{t=0} = (0, 0). \end{cases}$$

It is straightforward to see that there exists $\varepsilon > 0$ such that if $\| u^\omega_f \|_{L^5_t L^{10}_x(\mathbb{R} \times \mathbb{R}^3)} \leq \varepsilon$, then the map

$$\Phi(w)(t) := -\int_0^t \frac{\sin((t-s)|\nabla|)}{|\nabla|} (|u^\omega_f + w|^4 (u^\omega_f + w))(s) \, ds$$

is a contraction on a ball of radius $\varepsilon$ with respect to the $L^5_t L^{10}_x(\mathbb{R} \times \mathbb{R}^3)$ norm. Its unique fixed point

$$(w, w_t) \in C(\mathbb{R}; \dot{H}^{1/2}_x(\mathbb{R}^3)) \cap L^5_t L^{10}_x(\mathbb{R} \times \mathbb{R}^3) \times C(\mathbb{R}; L^2_x(\mathbb{R}^3))$$

is the global solution to (2.44). The probability estimate (2.7) on the event $\Omega_f$ in the statement of Theorem 2.2 follows immediately from (2.43), while the scattering statement follows readily using
that \( \|u_f + w\|_{L^4_1 L^2_2(\mathbb{R} \times \mathbb{R}^3)} < \infty \) for every \( \omega \in \Omega_f \). 

### 2.5 Appendix A: Probabilistic estimates

We record several useful probabilistic facts. We begin with the following Lemma which is a variant of Lemma 4.5 in [69], and is often used to prove bounds on the probability of various events.

**Lemma 2.10.** Let \( F_f \) be a real valued measurable function on a probability space \((\Omega, \mathbb{P})\). Suppose that there exists \( \alpha > 0, N > 0, k \in \mathbb{N}^* \) and \( C > 0 \) such that for every \( p \geq r_0 \) one has

\[
\|F_f\|_{L^p(\Omega)} \leq C\|f\|_{\mathcal{H}^*} \sqrt{p}
\]

Then, there exists \( \delta \) and \( C_1 \) depending on \( C \) and \( r_0 \) but independent of \( f \) such that for \( \lambda > 0 \)

\[
\mathbb{P}(E_{\lambda,f}) := \mathbb{P}(\omega \in \Omega : |F_f(\omega)| > \lambda) \leq C_1 e^{-\delta \lambda^2/\|f\|_{\mathcal{H}^*}^2}.
\]

**Proof.** Note that by Chebychev's inequality,

\[
\mathbb{P}(E_{\lambda,f}) \leq \lambda^{-p}(C\|f\|_{\mathcal{H}^*} \sqrt{p})^p, \quad \text{for all } p \geq r_0,
\]

so for \( c > 0 \) to be specified later we consider the quantity \( e^{\delta \lambda^2/\|f\|_{\mathcal{H}^*}^2} \mathbb{P}(E_{\lambda,f}) \) and we will show that it is bounded by a constant \( C_1 \). We expand

\[
e^{\delta \lambda^2/\|f\|_{\mathcal{H}^*}^2} \mathbb{P}(E_{\lambda,f}) = \sum_{n=1}^{\infty} \frac{1}{n!} \left( \delta \frac{\lambda^2}{\|f\|_{\mathcal{H}^*}^2} \right)^n \mathbb{P}(E_{\lambda,f})
\]

\[
\leq \sum_{n=1}^{r_0} \frac{1}{n!} \left( \delta \frac{\lambda^2}{\|f\|_{\mathcal{H}^*}^2} \right)^n \mathbb{P}(E_{\lambda,f}) + \sum_{n=r_0}^{\infty} \frac{1}{n!} \left( \delta \frac{\lambda^2}{\|f\|_{\mathcal{H}^*}^2} \right)^n \mathbb{P}(E_{\lambda,f}).
\]

We start with the first term: when \( \|f\|_{\mathcal{H}^*} < eC\sqrt{r_0} \) we bound \( \mathbb{P}(E_{\lambda,f}) \leq 1 \) and let \( \delta = (C^2 r_0 e^2)^{-1} \). We obtain

\[
\sum_{n=1}^{r_0} \frac{1}{n!} \left( \delta \frac{\lambda^2}{\|f\|_{\mathcal{H}^*}^2} \right)^n \mathbb{P}(E_{\lambda,f}) \leq \sum_{n=1}^{r_0} \frac{1}{n!} \left( \delta \frac{\lambda^2}{\|f\|_{\mathcal{H}^*}^2} \right)^n \leq \sum_{n=1}^{r_0} \frac{1}{n!} \left( \|f\|_{\mathcal{H}^*}^2 e^2 C^2 r_0 \right)^n \leq \sum_{n=1}^{r_0} \frac{1}{n!}.
\]
When $eC\sqrt{r_0} \leq \frac{1}{\|f\|_{H^s}}$ we set $p = \left[ \frac{\lambda n}{eC} \right]^2$ which is $\geq r_0$ and with $\delta = (eC)^{-2}$ we bound

$$
\sum_{n=1}^{r_0} \frac{1}{n!} \left( \frac{\lambda^2}{\|f\|_{H^s}^2} \right)^n P(E_{\lambda,f}) \leq \sum_{n=1}^{r_0} \frac{1}{n!} \left( \frac{\lambda^2}{\|f\|_{H^s}^2} \right)^n \left( \frac{C\|f\|_{H^s}^2 \sqrt{p}}{\lambda} \right)^p 
$$

$$
= \sum_{n=1}^{r_0} \frac{1}{n!} \left( \frac{\lambda^2}{\|f\|_{H^s}^2 e^2 C^2} \right)^n \left( \frac{C\|f\|_{H^s}^2 \lambda}{\|f\|_{H^s}^2 e^2 C^2} \right)^2 
$$

$$
= \sum_{n=1}^{r_0} \frac{1}{n!} \left( \delta^2 \|f\|_{H^s}^2 e^2 C^2 \right)^n.
$$

which can be bounded by a constant. For the second term, we observe that

$$
\sum_{n=r_0}^{\infty} \frac{1}{n!} \left( \frac{\lambda^2}{\|f\|_{H^s}^2} \right)^n P(E_{\lambda,f}) \leq \sum_{n=r_0}^{\infty} \frac{1}{n!} \left( \frac{\lambda^2}{\|f\|_{H^s}^2} \right)^n \lambda^{-2n} \left( \frac{C\sqrt{2n}}{\|f\|_{H^s}} \right)^{2n} = \sum_{n=r_0}^{\infty} \frac{n^n}{n!} (2\delta C^2)^n.
$$

Using Stirling's approximation and choosing $\delta < (2eC^2)^{-1}$ yields the result.

We now turn to the proof of the fact that our randomization procedure does not regularize at the level of Sobolev spaces. We recall the definition of our randomization set-up.

**Definition 2.11 (H$^s$ randomization).** Let $\{h_k\}_{k \in \mathbb{Z}^3}$ be a sequence of independent, 0 mean value, real-valued random variables with associated distributions $\{p_k\}_{k \in \mathbb{Z}^3}$ on a probability space $(\Omega, \mathcal{A}, P)$. Assume that there exist $c > 0$ such that

$$
|\int_{-\infty}^{+\infty} e^{\gamma x} d\mu_k(x)| \leq e^{c\gamma^2} \text{ for all } \gamma \in \mathbb{R} \text{ and for all } k \in \mathbb{Z}^3. \tag{2.45}
$$

For $f \in H^s(\mathbb{R}^3)$ we define

$$
f^\omega(x) := \sum_{k \in \mathbb{Z}^3} h_k(\omega)(P_k f)(x) \text{ for every } x \in \mathbb{R}^3. \tag{2.46}
$$

Condition (2.45) ensures that (2.46) defines a measurable map $\omega \mapsto f^\omega$ from $(\Omega, \mathcal{A})$ to $H^s(\mathbb{R}^3)$ and that $f^\omega \in L^2(\Omega; H^s(\mathbb{R}^3))$. The following lemma is a variant of Lemma B.1 in [14] and demonstrates that the randomization procedure does not regularize functions at the level of Sobolev spaces.

**Lemma 2.12.** Let $f = \sum_{k \in \mathbb{Z}^3} P_k f \in H^s(\mathbb{R}^3)$ be such that for some $\varepsilon > 0$, $f$ does not belong in $H^{s+\varepsilon}(\mathbb{R}^3)$. Let $(h_k(\omega))_{k \in \mathbb{Z}^3}$ be a sequence of independent random variables with distributions $\mu_k$
such that there exists $c > 0$ satisfying

$$\sup_{k \in \mathbb{Z}^3} \mu_k([-c, c]) < 1$$

Consider $f^\omega$ as defined in (2.46). Then the probability that $f^\omega$ belongs to $H^{s+\epsilon}(\mathbb{R}^3)$ is zero.

Proof. Let $c, \delta > 0$ be such that $\mu_k([-c, c]) \leq (1 - \delta)$ and we compute

$$\int e^{-\|f^\omega\|_{H^{s+\epsilon}(\mathbb{R}^3)}^2} d\mathbb{P}(\omega) \leq \prod_{k \in \mathbb{Z}^3} \int e^{-x^2 \|P_k f\|_{H^{s+\epsilon}(\mathbb{R}^3)}^2} d\mathbb{P}(\omega)$$

$$= \prod_{k \in \mathbb{Z}^3} \int_{-c}^c e^{-x^2 \|P_k f\|_{H^{s+\epsilon}(\mathbb{R}^3)}^2} d\mathbb{P}(\omega) + \prod_{k \in \mathbb{Z}^3} \int_{|x| > c} e^{-x^2 \|P_k f\|_{H^{s+\epsilon}(\mathbb{R}^3)}^2} d\mathbb{P}(\omega)$$

$$\leq \prod_{k \in \mathbb{Z}^3} \left(1 - \delta + e^{-c^2 \|P_k f\|_{H^{s+\epsilon}(\mathbb{R}^3)}^2} \delta \right).$$

We know that $\sum_{k \in \mathbb{Z}^3} \|P_k f\|_{H^{s+\epsilon}(\mathbb{R}^3)}^2 = \infty$, and hence $\sum_{k \in \mathbb{Z}^3} \left(1 - e^{-c^2 \|P_k f\|_{H^{s+\epsilon}(\mathbb{R}^3)}^2}\right) = \infty$. Thus, since $1 - x \leq e^{-x}$ for all $x$,

$$\prod_{k \in \mathbb{Z}^3} \left(1 - \delta + e^{-c^2 \|P_k f\|_{H^{s+\epsilon}(\mathbb{R}^3)}^2} \delta \right) \leq \prod_{k \in \mathbb{Z}^3} \exp\left(-\delta (1 - e^{-c^2 \|P_k f\|_{H^{s+\epsilon}(\mathbb{R}^3)}^2})\right)$$

$$= \exp\left(-\delta \sum_{k \in \mathbb{Z}^3} (1 - e^{-c^2 \|P_k f\|_{H^{s+\epsilon}(\mathbb{R}^3)}^2})\right)$$

$$= 0.$$
Chapter 3

An approximation result for the cubic NLKG in the critical space

3.1 Introduction

In this chapter, we study the defocusing cubic nonlinear Klein-Gordon equation

\[
\begin{aligned}
& u_{tt} - \Delta u + u + u^3 = 0, \quad u : \mathbb{R} \times \mathbb{T}^3 \to \mathbb{R} \\
& \langle u, \partial_t u \rangle|_{t=0} = (u_0, u_1) \in H^{1/2}(\mathbb{T}^3) \times H^{-1/2}(\mathbb{T}^3) =: \mathcal{H}^{1/2}(\mathbb{T}^3),
\end{aligned}
\]

where $H^{1/2}(\mathbb{T}^3)$ is the usual inhomogeneous Sobolev space. If we consider the general power-type nonlinear Klein-Gordon equation

\[
u_{tt} - \Delta u + u + |u|^\rho u = 0,
\]

then $s_c := \frac{d}{2} - \frac{2}{\rho-1}$ defines the critical regularity, hence we are interested in studying the cubic nonlinear Klein-Gordon equation at the critical regularity for the cubic nonlinear Klein-Gordon in dimension three.

The purpose of this chapter is to prove that solutions of (3.1) are stable at low frequencies under high-frequency perturbations to the initial data. While this type of argument is often seen at subcritical regularities, there are additional difficulties at the critical regularity since the decay one usually needs for such arguments is not available in this setting. Nonetheless, one can manufacture some decay using refined bilinear Strichartz estimates which, roughly speaking, show
that for $M \lesssim N$, dyadic frequencies,

$$\| e^{it\nabla} \phi_N e^{it\nabla} \psi_M \|_{L^2_t L^2_x} \lesssim M \| \phi_N \|_{L^2_t L^2_x} \| \psi_M \|_{L^2_t L^2_x}$$

for $P_N \phi_N = \phi_N$, $P_M \psi_M = \psi_M$.

This corresponds to one half derivative loss on each function. Thus, if one can control the frequency separation between functions in certain key multilinear estimates, the bilinear Strichartz estimate demonstrates that it's possible to regain some decay even in the critical setting. Our main result is the following.

**Theorem 3.1.** Let $\Phi$ denote the flow of the cubic nonlinear Klein-Gordon equation (3.1). Let $T > 0$ and $1 \leq N' \ll N_*$. Let $(u_0, u_1), (\tilde{u}_0, \tilde{u}_1) \in B_R \subset H^{1/2}(\mathbb{T}^3)$ be such that $P_{\leq N_*} (u_0, u_1) = P_{\leq N_*} (\tilde{u}_0, \tilde{u}_1)$, and suppose there exists some $K > 0$ such that corresponding solutions $u$ and $\tilde{u}$ to (3.1) satisfy

$$\| u \|_{L^4_t L^4_x([0,T] \times \mathbb{T}^3)} + \| \tilde{u} \|_{L^4_t L^4_x([0,T] \times \mathbb{T}^3)} \leq K.$$

Then for sufficiently large $N_*$ depending on $R, T$ and $K$,

$$\| P_{\leq N'} (\Phi(t)(u_0, u_1) - \Phi(t)(\tilde{u}_0, \tilde{u}_1)) \|_{L^\infty_t H_x^{1/2}([0,T] \times \mathbb{T}^3)} \lesssim \left( \log \frac{N_*}{N'} \right)^{-\theta},$$

with implicit constant depending on $R, T, K$.

**Remark 3.1.** The same result holds in the Euclidean setting with almost no modifications to the arguments. We restrict the statements to the periodic case, however, for simplicity of exposition. Moreover, in the Euclidean setting we can eliminate the dependence on time in the implicit constants.

We prove Theorem 3.1 by demonstrating that under the above assumptions, the low frequency component of the solutions, $u_{lo} = P_{\leq M} u$ for some $M \in \mathbb{N}$ satisfies a perturbed cubic Klein-Gordon equation given by

$$\square u_{lo} + u_{lo} = P_{\leq M} F(u_{lo}, u_{lo}, u_{lo}) + err.$$  \hspace{1cm} (3.2)

where $err$ is an error term which we can control by the well-posedness theory. Such an expression will allow us to prove Theorem 3.1 by a stability type argument.

We will work with the $U^p$ and $V^p$ function spaces. These spaces have previously been used in the context of critical problems by Hadac, Herr and Koch [28] for the KP-II equation, and by
Herr, Tataru and Tzvetkov [29] for the quintic nonlinear Schrödinger equation on $\mathbb{T}^3$. See [36] or [28] and references therein for a more complete overview of these function spaces. We record the basic definitions and properties of these spaces in Appendix A.2.2.

The key benefit of these function spaces is that they recover the endpoint embeddings which fail in $X^{s,\frac{1}{2}}$ but still enable us to exploit the same type of multilinear estimate machinery, and are thus are a suitable setting for critical problems. Unfortunately, however, although one can treat small data theory in these spaces, they miss out on a key property of the Strichartz spaces which is exploited in the stability theory, namely that if a space-time norm is finite on a time interval, then one can isolate subintervals where the norm is small.

To overcome this difficulty, we introduce a weaker norm (3.6) which recovers the necessary properties in order to prove stability, but which still controls the well-posedness theory. This approach was used by Ionescu and Pausader in [32] to prove global well-posedness of the energy critical nonlinear Schrödinger equation on $\mathbb{T}^3$, although the norm we introduce is slightly different than the one used in that work.

We refer the reader to Chapter 4 for an application of Theorem 3.1 to the symplectic non-squeezing of the cubic nonlinear Klein-Gordon equation (3.1).

### 3.2 Set-up

In the sequel, we let $\langle \nabla \rangle$ be the operator with symbol $\sqrt{1 + |\xi|^2}$ and we let $F(u) = u^3$. We rewrite (3.1) as a system of first order equations by factoring

$$\partial_{tt} - \Delta + 1 = (\langle \nabla \rangle + i\partial_t)(\langle \nabla \rangle - i\partial_t).$$

For a sufficiently regular solution $u$ to the NLKG (3.1) we can define

$$u^\pm = \frac{(\langle \nabla \rangle \mp i\partial_t)}{2\langle \nabla \rangle} u,$$

then $u = u^+ + u^-$ and the functions $u^\pm$ solve the equations

$$(\langle \nabla \rangle \pm i\partial_t)u^\pm = -\frac{F(u^+ + u^-)}{2\langle \nabla \rangle}, \quad u^\pm(0) = \frac{1}{2} \left( u_0 \mp i\frac{u_1}{\langle \nabla \rangle} \right).$$ (3.3)
We set
\[ I^\pm(f) = \int_0^t e^{\pm i(t-s)(\nabla)} \frac{f}{2(\nabla)} ds \]
and we obtain a Duhamel's formula for solutions of (3.3) given by
\[ u^\pm = e^{\pm i t(\nabla)} u_0^\pm \pm i I^\pm(F(u)). \]

This formulation is equivalent to (3.1) and since \( u = u^+ + u^- \), we can reconstruct a solutions to (3.1) from this system and bounds for \( u^\pm \) imply the same bounds for \( u \). We will not be using the specific structure of \( u^\pm \) so we will often drop the notation where it has no impact on the argument.

Before proceeding, we recall the definition of the \( U^p \) and \( V^p \) spaces. Consider partitions given by a strictly increasing finite sequence \( -\infty < t_0 < t_1 < \ldots < t_K \leq \infty \). If \( t_K = \infty \) we use the convention \( v(t_K) := 0 \) for all functions \( v : \mathbb{R} \to H \). We will usually be working on bounded intervals \( I \subset \mathbb{R} \). For some additional details about these function spaces, see Appendix A.2.2. In what follows, \( B \) will denote an arbitrary Banach space.

**Definition 3.2 (\( U^p \) spaces).** Let \( 1 \leq p < \infty \). Consider a partition \( \{t_0, \ldots, t_K\} \) and let \( \{\varphi_k\}_{k=0}^{K-1} \subseteq \mathcal{B} \) with \( \sum_{k=0}^{K-1} \|\varphi_k\|^p_{L^2} = 1 \). We define a \( U^p \) atom to be a function
\[ a := \sum_{k=1}^K \mathbb{1}_{[t_{k-1}, t_k]} \varphi_{k-1} \]
and we define the atomic space \( U^p(\mathbb{R}, \mathcal{B}) \) to be the set of all functions \( u : \mathbb{R} \to \mathcal{B} \) such that
\[ u = \sum_{j=1}^\infty \lambda_j a_j, \]
for \( a_j \) \( U^p \) atoms, and \( \{\lambda_j\} \in \ell^1(\mathbb{C}) \), endowed with the norm
\[ \|u\|_{U^p} := \inf \left\{ \sum_{j=1}^\infty |\lambda_j|, u = \sum_{j=1}^\infty \lambda_j a_j : a_j \text{ is a } U^p \text{ atom} \right\}. \]

**Definition 3.3 (\( V^p \) spaces).** Let \( 1 \leq p < \infty \). We define \( V^p(\mathbb{R}, \mathcal{B}) \) as the space of all functions,
Let $v: \mathbb{R} \to B$, such that the norm

$$\|v\|_{V^p(R,B)} = \sup_{\text{partitions}} \left( \sum_{i=1}^{K} \|v(t_i) - v(t_{i-1})\|_B^p \right)^{1/p} < \infty$$

with the convention $v(\infty) = 0$.

We let $V_-(R, B)$ denote the subspace of all functions satisfying $\lim_{t \to -\infty} v(t) = 0$ and $V^p_{rc}(R, B)$ denote the subspace of all right continuous functions in $V_-(R, B)$, both endowed with norm defined above.

We will establish a local well-posedness theory for (3.3) via a contraction mapping argument in the adapted functions space

$$\|u\|_{X^s_{\pm}} = \left( \sum_k (k)^{2s} \|u_{\pm}^k(k)\|^2_{U^2} \right)^{1/2}, \quad \|f\|_{U^2_{\pm}} = \|e^{it\Delta} f\|_{U^2_{\pm}(R \times T^3)}.$$  (3.4)

Similarly we define

$$\|u\|_{Y^s_{\pm}} = \left( \sum_k (k)^{2s} \|u_{\pm}^k(k)\|^2_{V^2_{rc,\pm}} \right)^{1/2}, \quad \|f\|_{V^2_{rc,\pm}} = \|e^{it\Delta} f\|_{V^2_{rc,\pm}(R \times T^3)}. \quad (3.5)$$

We will simplify the notation and let $V^2_{\pm} := V^2_{rc,\pm}$. We define

$$X^s = X^s_{+} \times X^s_{-} \quad \text{and} \quad Y^s = Y^s_{+} \times Y^s_{-}$$

endowed with the obvious norm. These spaces are at a slightly finer scale than the spaces used in [28], and one should view these as analogous to the spaces used by Herr, Tataru and Tzvetkov in [29] for the quintic nonlinear Schrödinger equation on $T^3$. This choice of scale does not affect the multilinear estimates, as we only need the weaker Strichartz estimates at dyadic scales, however it allows us to make use of the following important orthogonality property.

**Corollary 3.4** (Corollary 2.9, [29]). Let $Z^3 = \cup C_k$ be a partition of $Z^3$ or $R^k$. Then

$$\left( \sum_k \|P_{C_k} u\|^2_{L^2 H^s(I \times T^3)} \right)^{1/2} \lesssim \|u\|_{Y^s_{\pm}(I \times T^3)}.$$
We will work with the restriction spaces

$$X^s(I) := \{ u \in C(I; H^s(\mathbb{T}^3) \times H^s(\mathbb{T}^3)) \mid \tilde{u}(t) = u(t) \text{ for } t \in I, \tilde{u} \in X^s \}$$

endowed with the norm

$$\|u\|_{X^s(I)} = \inf \{ \|\tilde{u}\|_{X^s} \mid \tilde{u}(t) = u(t) \text{ for } t \in I \},$$

and similarly for $Y^s(I)$. In our estimates we will often implicitly multiply functions by a sharp time cut-off. When $I = [0, T)$ we will use the notation $X^s_T$. Perhaps most importantly, bounds for solutions of (3.3) in these spaces implies the same bounds in $L^\infty_t H^s_x$ for solutions of (3.1), which was precisely the endpoint embedding we were missing in the $X^{s,1/2}$ spaces. Finally, for $I \subset J$, we have the embedding $X^s(I) \hookrightarrow X^s(J)$ which can be seen via extension by zero.

As mentioned above, we will need a weaker norm which controls the local well-posedness theory for the equation in order to prove the necessary stability theory. This norm is similar to the norm introduced in [32, Section 2] for the same purpose and should be thought of as the appropriate substitute for the $L^4_{t,x}$ Strichartz norm. We define the $Z(I)$ norm by

$$\|f\|_{Z(I)} = \sup_{J \subseteq I, |J| \leq 1} \left( \sum_N \|P_N f\|_{L^4_{t,x}(J \times \mathbb{T}^3)}^2 \right)^{1/2}.$$  \hspace{1cm} (3.6)

For $u = (u^+, u^-)$, we obtain as a consequence of Corollary 3.4 and the Strichartz estimates in Corollary 3.6 below that

$$\|u^\pm\|_{L^4_{t,x}(I \times \mathbb{T}^3)} \lesssim \|u^\pm\|_{Z(I)} \lesssim \|u^\pm\|_{Y^{1/2}_{t,x}(I)} \lesssim \|u^\pm\|_{X^{1/2}_{t,x}(I)},$$

hence this indeed defines a weaker norm. This norm only plays a role in the stability theory and is not necessary to prove the low-frequency component satisfies the perturbed equation (3.2).

### 3.3 Strichartz estimates

We record some Strichartz estimates for the Klein-Gordon equation in the $U^p$ and $V^p$ spaces. By finite speed of propagation, the Strichartz estimates for the torus are the same as those in Euclidean space, provided one localizes in time.
Lemma 3.5 (Strichartz). Let $2 < q, r < \infty$ with $\frac{1}{q} + \frac{3}{r} = \frac{3}{2} - s$. Then

$$\|e^{\pm it(\nabla)}\phi\|_{L^q_t L^r_x([0,T] \times \mathbb{T}^3)} \lesssim C_T \|\varphi\|_{H^s}.$$ 

We also obtain the following formulation of Strichartz estimates in $U^p$ spaces. The reasoning is identical to the proof of Corollary 2.21 in [28].

Corollary 3.6. Let $1 \leq p < 4$ and $2 \leq q \leq 4$. Let $T > 0$, and let $u_N = P_N u_N$. Then,

(i) $\|u_N\|_{L^4_{t,x}([0,T] \times \mathbb{T}^3)} \lesssim C_T N^{1/2} \|u_N\|_{U^1_{t,x}}$

(ii) $\|u_N\|_{L^4_{t,x}([0,T] \times \mathbb{T}^3)} \lesssim C_T N^{1/2} \|u_N\|_{V^p_{t,x}}$

(iii) $\|u_N\|_{L^4_{t,x}([0,T] \times \mathbb{T}^3)} \lesssim C_T N^{\frac{q-2}{q}} \|u_N\|_{U^q_{t,x}}$

Proof. The first claim follows from Lemma 3.5 and the transfer principle, Proposition A.13. The second claim follows from (A.3) and the fact that $u$ agrees with its right-continuous variant almost everywhere. The final claim follows from interpolation with the trivial $L^2_{t,x}$ estimate, since we are on bounded time intervals. \hfill $\square$

We have the following refined bilinear Strichartz estimate for the Klein-Gordon equation. Such estimates were treated in full for wave equations in [23]. For a proof of this statement for the Klein-Gordon equation on Euclidean space, see [61, Appendix A] which adapts the geometric proofs from [62] for the nonlinear wave equation, see also [35]. One can also adapt the proof from [62] to the compact setting to prove Proposition 3.7 directly, replacing the volume estimates with fairly straightforward lattice counting.

**Proposition 3.7.** Fix $T > 0$ and let $O, M, N$ be dyadic numbers and $\phi_M, \psi_N$ functions in $L^2(\mathbb{T}^3)$ localized at frequencies $M, N$ respectively. Define $u_M = e^{\pm \imath t(\nabla)}\phi_M$ and $v_M = e^{\pm \imath t(\nabla)}\psi_N$ Denote $L = \text{min}(O, M, N)$, and $H = \text{max}(O, M, N)$. Then

$$\|P_O(u_Mv_N)\|_{L^2_{t,x}} \lesssim C_T \begin{cases} L \|\phi_M\|_{L^2(\mathbb{T}^3)} \|\psi_N\|_{L^2(\mathbb{T}^3)} & \text{if } M \ll N \\ H^{1/2} L^{1/2} \|\phi_M\|_{L^2(\mathbb{T}^3)} \|\psi_N\|_{L^2(\mathbb{T}^3)} & \text{if } M \sim N. \end{cases}$$

In order to prove the multilinear estimates required for the stability theory, we will make use of the following refined Strichartz estimate which is proved using Proposition 3.7.
Proposition 3.8 (Proposition 10, [61]). Let \( n \geq 3 \) and let \( M \lesssim N \) and let \( u_{M,N} \) have Fourier support in a ball of radius \( \sim M \) centered at frequency \( N \). Then

\[
\|u_{M,N}\|_{L^4_t L^\infty_x([0,T) \times \mathbb{T}^3)} \lesssim C_T N^{1/4} M^{n/4} \|u_{M,N}\|_{U^1_\pm}.
\]

In particular if \( u_N = P_N u_N \) then the result holds with \( M = N \).

By the transfer principle, an orthogonality argument and Corollary 3.6, we obtain the following \( U^p \) version of the above estimates.

Proposition 3.9 (Proposition 7, [61]). Fix \( T > 0 \) and let \( L \), respectively \( H \), denote the highest and lowest frequencies of \( M, N, O \). Let \( u_M \in U^2_{\pm 1}, u_N \in U^2_{\pm 2} \). Then

\[
\|P_O(u_M u_N)\|_{L^2_t L^\infty_x([0,T))} \lesssim C_T \begin{cases} L \|u_M\|_{U^2_{\pm 1}} \|u_N\|_{U^2_{\pm 2}} & \text{if } M \ll N \\ H^{1/2} L^{1/2} \|u_M\|_{U^2_{\pm 1}} \|u_N\|_{U^2_{\pm 2}} & \text{if } M \sim N. \end{cases}
\]

Remark 3.2. One can convert these estimates to bounds in \( V^2_{\pm} \) with a logarithmic loss in the first estimate, and no loss in the second estimate, see Proposition A.14.

The following lemma follows immediately from the atomic structure of \( U^2_\pm \). See, for instance, [61, (17)] for the computation.

Lemma 3.10 (Linear solutions lie in \( X^s \)). Let \( s \geq 0, 0 < T \leq \infty \) and \( u_0^\pm \in H^s(\mathbb{T}^3) \). Then

\[
\|e^{\pm i t (\nabla)} u_0^\pm \|_{X^s(\mathbb{T}^3)} \leq \|u_0^\pm\|_{H^s(\mathbb{T}^3)}.
\]

We have the following duality estimate. The proof of this result is a straightforward adaptation of the proof of Proposition 2.11 in [29].

Proposition 3.11. Let \( s \geq 0 \) and \( T > 0 \). For \( f \in L^1_t H^s([0,T) \times \mathbb{T}^3) \) we have

\[
\|I^\pm(f)\|_{X^s_\pm} \leq \sup_{\omega \in L^1_{\pm s} : \|\omega\|_{L^1_{\pm s}} = 1} \left| \int_0^T \int_{\mathbb{T}^3} f(t,x) \overline{v(t,x)} dt dx \right|.
\]

Proof. We let \( (a_n)_{n \in \mathbb{Z}^3} \in \ell^2(\mathbb{Z}^3) \) be with \( \|a_n\|_{\ell^2} = 1 \) such that

\[
\|I^\pm(f)\|_{X^s_\pm} \leq \sum_{n \in \mathbb{Z}^3} a_n \langle n \rangle^s \left\| \int_0^t \chi_{[0,T)} e^{\pm i(t-s) \langle n \rangle} \frac{f(t,x)}{2\langle \nabla \rangle} dt \right\|_{U^2_{\pm}} + \varepsilon.
\]
We use the definition of $U^2_\pm$, and we use duality from Theorem A.9 and Proposition A.11 to estimate each piece by

$$\left\| \int_0^t X(t, x) e^{\pm is(n)} \frac{f(t, x)}{2(\nabla)} dt \right\|_{U^2_\pm} \leq \left| \int_0^T f(s) e^{\pm is(n)} \frac{v_n(s)}{2(n)} ds \right| + 2^{-|n|^2} \varepsilon$$

for a sequence $v_n \in V^2_{F\varepsilon}$ with $\|v_n\|_{V^2_{F\varepsilon}} = 1$ supported on $[0, T)$. We then define

$$v(t, x) \simeq \sum_{n \in \mathbb{Z}^3} \left( a_n(n)^{-1} e^{\pm it(n)} v_n(t) \right) e^{ix \cdot n},$$

and we observe that $v \in Y_{\pm}^{1-s}([0, T))$ with $\|v\|_{Y_{\pm}^{1-s}} \leq 1$. Hence

$$\|I^\pm(f)\|_{Y_{\pm}^{1/2}} \leq \sum_{n \in \mathbb{Z}^3} \left| \int_0^T f(t) e^{it(n)} v(t)(n) \right| + c \varepsilon,$$

and the claim follows by dominated convergence theorem and Plancherel. \qed

Finally, the following proposition demonstrates that a priori bounds on the Strichartz norm control the norm of solutions in the adapted function spaces (3.4).

**Proposition 3.12.** Let $u$ be a solution to the cubic nonlinear Klein-Gordon equation for initial data $(u_0, u_1) \in B_R \subset H^{1/2}$ which satisfies

$$\|u\|_{L^4_t L^{4/3}_{x}}([0, T) \times \mathbb{T}^3) \leq K. \quad (3.7)$$

Then $\|u\|_{X_T^{1/2}} \lesssim C(K, R)$.

**Proof.** Let $u$ solve the cubic nonlinear Klein-Gordon equation (3.3) and suppose that $u$ satisfies the uniform bound (3.7). Fix $T > 0$ then by Lemma 3.10 we estimate

$$\|u\|_{X_T^{1/2}} \lesssim \|(u_0, u_1)\|_{H^{1/2}} + \|I(F)\|_{X_T^{1/2}}, \quad F(u) = u^3.$$

We expand the nonlinear term and we deal with $I^+$ as the other term is handled analogously. By
Proposition 3.11 and Hölder’s inequality

\[ \| I^+(F) \|_{X^1_+}^{1/2} \leq \sup_{v \in Y^{1/2}_- : \|v\|_{Y^{1/2}_-} = 1} \left| \int_0^T \int_{T^3} F(t, x)v(t, x)dtdx \right| \]
\[ = \sup_{v \in Y^{1/2}_- : \|v\|_{Y^{1/2}_-} = 1} \sum_N \left| \int_0^T \int_{T^3} P_N F \cdot P_N v(t, x)dtdx \right| \]
\[ \leq \sup_{v \in Y^{1/2}_- : \|v\|_{Y^{1/2}_-} = 1} \sum_N \| P_N F \|_{L^{4/3}_{t,x}} \| P_N v(t, x) \|_{L^{4}_{t,x}}. \]

By complex interpolation, we have the dual square function type inequality

\[ \left( \sum_N \| P_N F \|_{L^{4/3}_{t,x}}^2 \right)^{1/2} \lesssim \| F \|_{L^{4/3}_{t,x}}. \]

Applying Cauchy-Schwarz, and noting that Remark A.4 applies to \( I^+ \), we use Corollary 3.6 part (ii) and Corollary 3.4, to obtain

\[ \sup_{v \in Y^{1/2}_- : \|v\|_{Y^{1/2}_-} = 1} \left( \sum_N \| P_N f \|_{L^{4/3}_{t,x}}^2 \right)^{1/2} \| v \|_{Y^{1/2}_-} \lesssim \| f \|_{L^{4/3}_{t,x}} \leq \| u \|_{L^4_{t,x}}^3, \]

which yields the result.

3.4 Multilinear estimates

Multilinear estimates for nonlinear Klein-Gordon equations in the form we need for well-posedness were proven in [61], where more general non-linearities were treated, but we state only the estimates we require. The statements are slightly modified for our setting and we include the proofs for completeness. Ultimately, however, we will need slightly stronger estimates for the stability theory, which we prove in Proposition 3.15.

Theorem 3.13 (Theorem 3, [61]). Suppose that the signs \( \pm_i \ (i = 0, 1, 2, 3) \) are arbitrary and \( H \sim H' \). Then

\[ \frac{1}{H} \left| \sum_{L_i \leq H} \left( \prod_{i=1}^2 u_{L_i} u_{H'} u_H dxdt \right) \right| \leq \prod_{i=1}^2 \left( \sum_{L_i \leq H} L_i \| u_{L_i} \|_{V_{x,t}^2}^2 \right)^{1/2} \| u_{H'} \|_{V_{x}^3} \| v_H \|_{V_{x}^2} \]  

(3.8)
Proof. We estimate

\[
\frac{1}{H} \int \left| \prod_{i=1}^{2} u_{L_i} u_{H_i} v_L dxdt \right| \lesssim \frac{1}{H} \|u_{L_i} u_{H_i}\|_{L^2_{i,x}} \|u_{L_2} v_L\|_{L^2_{1,x}} \lesssim \frac{1}{H} \left( \frac{H^2}{L_1 L_2} \right)^\delta L_1 L_2 \|u_{L_1}\|_{V^2_{1,\pm}} \|u_{L_2}\|_{V^2_{1,\pm}} \|u_{H_i}\|_{V^2_{2,\pm}} \|v_L\|_{V^2_{0,\pm}}
\]

where we used the improved bilinear estimates from Proposition 3.9 and Remark A.6 in each term. Summing over \(L_i \lesssim H\), we use Cauchy-Schwarz on the terms

\[
L_i^{1-\delta} \|u_{L_i}\|_{V^2_{1,\pm}} = L_i^{1-\delta} L_i^{\frac{1}{2}} \|u_{L_i}\|_{V^2_{1,\pm}}
\]

which yield the \(i = 1, 2\) factors in (3.8). Ultimately, it suffices to bound

\[
H^{-1+2\delta} \left( \sum_{L_1 \leq H} \sum_{L_2 \leq L_1} L_2^{1-2\delta} L_1^{1-2\delta} \right)^{1/2} \lesssim H^{-1+2\delta} \left( \sum_{L_1 \leq H} L_1^{2-4\delta} \right)^{1/2} \lesssim 1,
\]

and since \(1 - 2\delta > 0\) for \(\delta\) sufficiently small, this yields (3.8).

The second estimate is treated similarly, with the roles of \(u_{L_2}\) and \(v_L\) swapped. That is, we bound

\[
\left| \int \int u_{L_1} u_{H'} u_H v_L dxdt \right| \lesssim \|u_{L_1} u_H\|_{L^2_{i,x}} \|u_{H'} v_L\|_{L^2_{i,x}} \lesssim \left( \frac{H^2}{L_1 L} \right)^\delta L_1 \|u_{L_1}\|_{V^2_{1,\pm}} \|u_H\|_{V^2_{1,\pm}} \|u_{H'}\|_{V^2_{2,\pm}} \|v_L\|_{V^2_{0,\pm}},
\]

Collecting terms, summing in \(L_1, L \lesssim H\), and applying Cauchy-Schwarz yields

\[
(3.9) \lesssim H^{2\delta} \left( \sum_{L_1 \leq H} \sum_{L_2 \leq L_1} L_1^{1-2\delta} L_1^{1-2\delta} \right)^{1/2} \|u_{L_1}\|_{V^2_{1,\pm}} \|u_H\|_{V^2_{2,\pm}} \|u_{H'}\|_{V^2_{2,\pm}} \lesssim (HH')^{\frac{1}{2}} \|u_{L_1}\|_{V^2_{1,\pm}} \|u_H\|_{V^2_{2,\pm}} \|u_{H'}\|_{V^2_{2,\pm}},
\]

as required. \(\Box\)
We will now review the well-posedness theory for (3.3). We need to estimate the cubic nonlinearity $F(u) = u^3$ for $u = u^+ + u^-$. Hence we can decompose

$$F(u) = \sum_{\{i,j,k\} \in \{1,2,3\}} u^{(i)} u^{(j)} u^{(k)}$$

for $u^{(i)} = u^\pm$. Thus, it suffices to estimate these eight cubic terms, which we do in the sequel. Due to finite speed of propagation, the arguments from [61] apply if one allows implicit constants to depend on the time interval.

**Theorem 3.14** (Theorem 4, [61]). Fix $T > 0$. There exists a constant $C$ depending only on $T > 0$ such that

$$\|I(u^{(1)}, u^{(2)}, u^{(3)})\|_{X^1/2} \leq C \prod_{i=1}^{3} \|u^{(i)}\|_{Y^{-1/2}}. \tag{3.10}$$

In particular, since $X^s \hookrightarrow Y^s$, this implies

$$\|I(u^{(1)}, u^{(2)}, u^{(3)})\|_{X^1/2} \leq C \prod_{i=1}^{3} \|u^{(i)}\|_{X^{-1/2}}.$$  

**Remark 3.3.** The proof works similarly for any $s \geq 1/2$ with the obvious modifications, however we omit this generalization for simplicity of presentation. The well-posedness of (3.3) follows from these estimates via a straightforward contraction mapping argument.

**Proof.** We only treat $I^+$ as $I^-$ follows analogously, and $T > 0$ will be fixed throughout. In the usual manner, we take extensions of the $u^{(i)}$ to $\mathbb{R}$, which we still denote by $u^{(i)}$, and (3.10) follows by taking infimums over all such extensions. We also suppress the notation $\pm_i$ on each function. We do not repeat these considerations. By Proposition 3.11, we estimate

$$\left\| I^+(u^{(1)}, u^{(2)}, u^{(3)}) \right\|_{X^1/2} \leq \sup_{\|u\|_{X^{-1/2}} = 1} \sum_{\text{N\_i \approx N\_j}} \sum_{N_3 + N_2 + N_1 = N_0} \int u_{N_3} u_{N_2} u_{N_1} u_{N_0} \, dx \, dt.$$  

The convolution requirement, implies that the right-hand side above vanishes unless $N_i \sim N_j$ for some $i \neq j$, and hence we may assume without loss of generality that

$$N_3 \leq N_2 \leq N_1, \quad N_1 \sim \max\{N_0, N_2\}.$$
In the case that \( N_0 \sim N_1 \), we use (3.8), and we bound
\[
\left| \sum_{N_3+N_2+N_1=N_0} \int u_{N_3} u_{N_2} u_{N_1} v_{N_0} \, dx \, dt \right| \\
\lesssim \sum_{N_0 \sim N_1} \left( \sum_{N_i \leq N_0} N_i \| u_{N_i} \|_{Y_{\ell_i}^{1/4}}^2 \right)^{1/2} \| u_{N_1} \|_{Y_{\ell_1}^{1/2}} \| v_{N_0} \|_{Y_{\ell_0}^{1/2}}.
\]

By Corollary 3.4 and Cauchy-Schwarz in \( N_0 \sim N_1 \),
\[
\left\| I^+ (u^{(1)}, u^{(2)}, u^{(3)}) \right\|_{X^{1/2}_+} \lesssim \left\| u^{(1)} \right\|_{Y_{\ell_1}^{1/2}} \left\| u^{(2)} \right\|_{Y_{\ell_2}^{1/2}} \left\| u^{(3)} \right\|_{Y_{\ell_3}^{1/2}}.
\]

When \( N_2 \sim N_1 \), we use (3.9) and we can similarly bound
\[
\left\| I^+ (u^{(1)}, u^{(2)}, u^{(3)}) \right\|_{X^{1/2}_+} \leq \sup_{\| w \|_{Y_{1/2}}=1} \left| \sum_{N_3+N_2+N_1=N_0} \int u_{N_3} u_{N_2} u_{N_1} w \, dx \, dt \right| \\
\lesssim \left\| u_3 \right\|_{Y_{\ell_3}^{1/2}} \sum_{N_1 \sim N_2} (N_1 N_2)^{1/2} \| u_{N_1} \|_{Y_{\ell_1}^{1/2}} \left\| u_{N_2} \right\|_{Y_{\ell_2}^{1/2}}.
\]

Once again, we obtain the desired bound using Corollary 3.4 and Cauchy-Schwarz.

Finally, we need to following refinement to Theorem 3.14 which says that we can replace some of the factors with the mixed norm
\[
\| u^\pm \|_{Z'_+(I)} = \| u^\pm \|_{Z(I)}^{3/4} \| u^\pm \|_{X_{\ell}^{1/2}(I)}^{1/4},
\]
for \( Z \) the norm defined in (3.6), and we set \( Z' = Z_+ \times Z_- \).

As in the proof of Theorem 3.14, we need to consider cubic expressions in \( u^+ \) and \( u^- \). It will be clear from the proof that we could have, instead, relied only on bounding \( u^+ + u^- \) in the \( Z(I) \) norm, which is more consistent with the \( L^4_{t,x} \) Strichartz norm from the standard well-posedness theory. Ultimately, however, such considerations do not affect our arguments given Proposition 3.12 which allows us to control the \( Z(I) \) norm of solutions to (3.3) via Strichartz bounds.

**Proposition 3.15.** Under the above assumptions,
\[
\left\| I(u^{(1)}, u^{(2)}, u^{(3)}) \right\|_{X^{1/2}_+(I)} \lesssim \sum_{\{i,j,k\} \in \{1,2,3\}} \| u^{(i)} \|_{Y_{\ell_i}^{1/2}(I)} \| u^{(j)} \|_{Z'_+(I)} \| u^{(k)} \|_{Z'_+(I)}.
\]
Proof. We fix $T > 0$ and we only treat the $I^+$ term. By Proposition 3.11, we estimate

$$\left\| I^+ (u^{(1)}, u^{(2)}, u^{(3)}) \right\|_{L^1 \to \mathbb{R}} \leq \sup_{\|w\|_{L^1} = 1} \left| \sum_{N_3 + N_2 + N_1 = N_0} \int u_{N_3} u_{N_2} u_{N_1} v_{N_0} \, dx \, dt \right|.$$ 

The convolution requirement implies that the right-hand side above vanishes unless $N_i \sim N_j$ for some $i \neq j$, and hence we may without loss of generality assume that

$$N_3 \lesssim N_2 \lesssim N_1, \quad N_1 \sim \max\{N_0, N_2\}.$$ 

In the case that $N_0 \sim N_1$, we apply Hölder’s inequality and obtain

$$\left| \sum_{N_3 + N_2 + N_1 = N_0} \int u_{N_3} u_{N_2} u_{N_1} v_{N_0} \, dx \, dt \right| \lesssim \sum_{N_3 + N_2 + N_1 = N_0} \|u_{N_3} u_{N_1}\|_{L^2} \|u_{N_2} v_{N_0}\|_{L^2}.$$ 

Let us consider the first term. Let $C$ be a cube of size $N_3$ centered in frequency space at $\xi_0 \in \mathbb{Z}^3$ with $|\xi_0| \sim N_1$ and let $P_C$ denote the (sharp) Fourier projection onto this cube. Since the spatial Fourier support of $(P_C u_{N_1}) u_{N_3}$ is contained in a fixed dilate of $C$,

$$\|u_{N_3} u_{N_1}\|_{L^2} \lesssim \left( \sum_C \|P_C u_{N_1}\|^2_{L^2} \right)^{1/2},$$

and hence by Hölder’s inequality and Proposition 3.8 on the term with $P_C u_{N_1}$, we can bound

$$\|u_{N_3} u_{N_1}\|_{L^2}^2 \lesssim (N_3 N_1)^{1/2} \|u_{N_3}\|^2_{L^4(I \times \mathbb{T}^3)} \sum_C \|P_C u_{N_1}\|^2_{V^4_{\pm 1}}.$$ 

Thus for fixed $N_1$, we obtain

$$\sum_{N_3 \lesssim N_1} \|u_{N_3} u_{N_1}\|_{L^2} \lesssim \sum_{N_3 \lesssim N_1} (N_3 N_1)^{1/2} \|u_{N_3}\|^2_{L^4(I \times \mathbb{T}^3)} \left( \sum_C \|P_C u_{N_1}\|^2_{V^4_{\pm 1}} \right)^{1/2} \lesssim \sum_{N_3 \lesssim N_1} \left( \frac{N_3}{N_1} \right)^{1/4} \|u_{N_3}\|^2_{L^4(I \times \mathbb{T}^3)} \left( \sum_C \|P_C u_{N_1}\|^2_{V^4_{\pm 3}} \right)^{1/2},$$

62
hence by resumming over $C$ and using Cauchy-Schwarz in $N_3$, we can bound

$$\sum_{N_3 \leq N_1} \|u_{N_3} u_{N_1}\|_{L^2_{t,x}} \lesssim \sum_{N_3 \leq N_1} \left( \frac{N_3}{N_1} \right)^{\frac{1}{4}} \|u_{N_3}\|_{L^4_{t,x}(I \times T^3)} \|u_{N_1}\|_{Y^{1/2}}$$

which yields

$$\sum_{N_3 \leq N_1} \|u_{N_3} u_{N_1}\|_{L^2_{t,x}} \lesssim \|u_1\|_2 \|u_{N_1}\|_{Y^{1/2}}.$$

We perform the same analysis on the second term. By Cauchy-Schwarz in $N_0 \sim N_1$ and symmetrizing we obtain

$$\left\| I(u^{(1)}, u^{(2)}, u^{(3)}) \right\|_{X^{1/2}(I)} \lesssim \sum_{\{i,j,k\} \in \{1,2,3\}} \|u^{(i)}\|_{X^{1/2}(I)} \|u^{(j)}\|_{Z(I)} \|u^{(k)}\|_{Z(I)}.$$

Using the definition of the $Z'(I)$ norm, and combining this bound with the estimates from Theorem 3.14 yields the result in this case.

The case when $N_1 \sim N_2$ and $N_0 \lesssim N_2$ requires a bit more care. We estimate

$$\left| \sum_{N_3 + N_2 + N_1 = N_0} \int u_{N_3} u_{N_2} u_{N_1} v_{N_0} dx dt \right| \lesssim \sum_{N_3 + N_2 + N_1 = N_0} \|u_{N_3} u_{N_1}\|_{L^2_{t,x}} \|u_{N_2} v_{N_0}\|_{L^2_{t,x}}.$$

As above, for the first term, we obtain

$$\sum_{N_3 \leq N_1} \|u_{N_3} u_{N_1}\|_{L^2_{t,x}} \lesssim \sum_{N_3 \leq N_1} \left( \frac{N_3}{N_1} \right)^{\frac{1}{4}} \|u_{N_3}\|_{L^4_{t,x}(I \times T^3)} \left( \sum_{N_1} \|P_{C_1} u_{N_1}\|_{Y^{1/2}}^2 \right)^{1/2}$$

$$\lesssim \|u_{N_1}\|_{Y^{1/2}} \sum_{N_3 \leq N_1} \left( \frac{N_3}{N_1} \right)^{\frac{1}{4}} \|u_{N_3}\|_{L^4_{t,x}(I \times T^3)}.$$

For the second term we use Hölder's inequality to bound

$$\sum_{N_0 \leq N_2} \|u_{N_2} v_{N_0}\|_{L^2_{t,x}} = \sum_{N_0 \leq N_2} \|u_{N_2}\|_{L^4_{t,x}}^{1/4} \|u_{N_2} v_{N_0}\|_{L^4_{t,x}}^{3/4}.$$
and we use Propositions 3.9 and 3.8 to bound this by
\[
\sum_{N_0 \leq N_2} (N_0)^{1/4} \left( \frac{N_2}{N_0} \right)^{\delta} \|u_{N_2}\|_{V_{\pm 2}^1}^{1/4} \|v_{N_0}\|_{V_{\pm 2}^1}^{1/4} \|u_{N_2}\|_{L_{t,x}^4}^{3/4} N_0^{3/8} \|v_{N_0}\|_{V_{\pm 2}^1}^{3/4}.
\]
\[
\lesssim \sum_{N_0 \leq N_2} \left( \frac{N_0}{N_2} \right)^{\frac{1}{8} - \delta} (N_2)^{1/8} \|u_{N_2}\|_{V_{\pm 2}^1}^{1/4} \|u_{N_2}\|_{L_{t,x}^4}^{3/4} N_0^{1/2} \|v_{N_0}\|_{V_{\pm 2}^1}.
\]

We have split this term using Hölder’s inequality in order to gain some term which enables us to sum in $N_0$ without loss. Using that
\[
\sum_{N_0 \leq N_2} \left( \frac{N_0}{N_2} \right)^{\frac{1}{8} - \delta} N_0^{1/2} \|u_{N_0}\|_{V_{\pm 2}^1} \lesssim \|w\|_{V_{\pm}^{1/2}}^1,
\]
and
\[
\sum_{N_3 \leq N_1} \left( \frac{N_3}{N_1} \right)^{\frac{1}{4}} \|u_{N_3}\|_{L_{t,x}^4(I \times T^3)} \lesssim \|u_3\|_{Z(I)},
\]
we are left with
\[
\|u_{N_1}\|_{V_{\pm 1/2}^{1/2}} \sum_{N_3 \sim N_2} (N_2)^{1/8} \|u_{N_2}\|_{V_{\pm 2}^1} \|u_{N_2}\|_{L_{t,x}^4}^{3/4},
\]
and we once again conclude by Cauchy-Schwarz with $\frac{1}{8} + \frac{3}{8} + \frac{1}{2} = 1$ and Corollary 3.4.

### 3.5 Stability theory in adapted function spaces

In this section, we prove the necessary stability theory for the nonlinear Klein-Gordon equation in the adapted function spaces. As discussed in the introduction, the key difficulty in proving a satisfactory stability theory in this setting is that even if $X^{1/2}(I)$ is bounded, we cannot isolate a small interval on which the norm is small. Ultimately, however, using the intermediate $Z'(I)$ norm we are able to recover the desired stability theory. We record the following results which are, for the most part, straightforward adaptations of the analogous results for the nonlinear Schrödinger equation from [32, Section 3].

**Proposition 3.16.** Suppose that $R > 0$ is fixed and let $u_0 = (u_0^+, u_0^-)$ with $\|u_0^{\pm}\|_{H^{1/2}(I \times T^3)} \leq R$. Then there exists $\delta_0 = \delta_0(R) > 0$ such that if
\[
\|e^{\pm it\langle \nabla \rangle} u_0^{\pm}\|_{Z_{\pm}(I)} < \delta
\]
\[
64
\]
for some $\delta \leq \delta_0$, on some interval $I$ with $0 \in I$ and $|I| \leq 1$, then there exists a strong solution to (3.3) in $X^{1/2}(I)$ with initial data $u(0) = u_0$ and

$$\|u^\pm - e^{\pm it(\nabla)}u_0^\pm\|_{X^{1/2}(I)} \leq \delta^{3/2}.$$  

(3.11)

Remark 3.4. The choice $5/3$ is arbitrary and in fact the statement holds for any $\alpha$ with $1 < \alpha < 2$, and we merely require $\alpha > 1$ for our applications. This proposition can be thought of as an version of a small data result in the adapted function spaces which does not require that the initial data be small in the $H^{1/2}$ norm.

Proof. The statement about existence follows from a standard fixed point argument. Indeed, Let $R, \alpha > 0$ and consider

$$\mathcal{S} = \{u \in X^{1/2}(I) : \|u\|_{X^{1/2}(I)} \leq 4R, \|u\|_{Z'(I)} \leq 2\alpha\}$$

and the mapping

$$\Phi^\pm(u) = e^{\pm it(\nabla)}u_0^\pm \pm iI^\pm(F(u)).$$

By Proposition 3.15,

$$\|\Phi^\pm(u)\|_{X^{1/2}(I)} \leq \|u_0^\pm\|_{H^{1/2}} + CR\alpha^2 \leq R + CR\alpha^2$$

$$\|\Phi^\pm(u)\|_{Z'(I)} \leq \|e^{\pm it(\nabla)}u_0^\pm\|_{Z'(I)} + CR\alpha^2 \leq \delta + CR\alpha^2,$$

and similarly for the difference expression. Choosing $\alpha = 2\delta$ for $0 < \delta \leq \delta_0$ and $\delta_0 = \delta_0(R)$ small enough so that $4CR\delta < 1$, we find that $\Phi = \Phi^+ + \Phi^-$ possesses a unique fixed point $u$ in $\mathcal{S}$.

Finally, to prove (3.11), we obtain by another application of Proposition 3.15 that

$$\|u^\pm - e^{\pm it(\nabla)}u_0^\pm\|_{X^{1/2}(I)} \lesssim R\delta^2,$$

hence taking $\delta_0$ even smaller if necessary, we obtain the statement. \hfill \Box

In light of the fact that the $Z'$ norm controls the $L^{4,4}_{t,x}$ norm of solutions, the standard blow-up criterion in Strichartz spaces together with Proposition 3.12 imply that this weaker norm controls the global existence theory.
For simplicity, we define the norm
\[
\|h\|_{N(I)} = \left\| \int_a^t e^{+it\langle\nabla\rangle} \frac{h(s)}{2\langle\nabla\rangle} ds \right\|_{X^1(I)}.
\] (3.12)

The following is the main result of this section.

**Proposition 3.17.** Let \( I \subset \mathbb{R} \) a compact time interval and \( t_0 \in I \). Let \( v \) be a solution defined on \( I \times T^3 \) of the Cauchy problem

\[
\begin{aligned}
&v_{tt} - \Delta v + v + F(v) = e \\
&(v, \partial_t v)_{t=t_0} = (v_0, v_1) \in \mathcal{H}^{1/2}(T^3),
\end{aligned}
\]

and identify the solution \( v \) with \( (v^+, v^-) \). Suppose that

\[
\|v^\pm\|_{L^\infty T^3} + \|v^\pm\|_{L^t_t \mathcal{H}^{1/2}(I \times T^3)} \leq K.
\] (3.13)

Let \( (u, \partial_t u)_{t=t_0} = (u_0, u_1) \in \mathcal{H}^{1/2}(T^3) \) and suppose we have the smallness condition

\[
\|(v_0 - u_0, v_1 - u_1)\|_{\mathcal{H}^{1/2}(T^3)} + \|e\|_{N(I)} \leq \varepsilon < \varepsilon_1
\] (3.14)

for some \( 0 < \varepsilon < \varepsilon_1 \) where \( \varepsilon_1 \leq 1 \) is a small constant \( \varepsilon_1 = \varepsilon_1(K, I) > 0 \). Then there exists a unique solution \( (u, \partial_t u) \) to the cubic nonlinear Klein-Gordon equation on \( I \times T^3 \) with initial data \( (u_0, u_1) \) at time \( t_0 \) and \( C = C(K, I) \geq 1 \) which satisfies \( \|v - u\|_{X^{1/2}(I)} \leq C \varepsilon \).

**Remark 3.5.** In particular, (3.13) holds if we have \( X^{1/2}(I) \) bounds on the solution \( v \), and consequently, by Proposition 3.12, if we have \( L^{4/3}_{t,x}(I \times T^3) \) bounds. Additionally, the computations in Proposition 3.12 also imply that \( L^{4/3}_{t,x}(I \times T^3) \) bounds on the error imply the \( N(I) \) bounds on the error in (3.14). Hence Proposition 3.17 can be seen as a refined version of the long-time stability theory from Appendix 4.8.

We include a proof of this fact for completeness, although it follows almost identically to the corresponding statement for the NLS in [32, Section 3]. The main idea is to mimic the proof of the standard Strichartz space stability, exploiting the extra properties of the \( Z'(I) \) norm. Roughly speaking, we will work on small intervals where \( \|v\|_{Z'(I_k)} \) is sufficiently small and then, in spirit, the
computations which yield the standard stability theory yield the result. We only have to check that at each step we can guarantee that the assumptions still hold, namely that the difference between the solutions remains sufficiently small. This is possible since the number of steps depends only on $K$ and $\varepsilon_1$, hence we can iterate such an argument to cover the whole interval in order to obtain the result.

**Proof.** Without loss of generality, we may assume $|I| \leq 1$. As in the proof of Proposition 3.16, there exists some $\delta_1(K)$ such that if for some $J \ni t_0$,

$$\|e^{\pm i(t-t_0)}v(t_0)\|_{L_\pm(J)} + \|e\|_{N_\pm(J)} \leq \delta_1,$$

then there exists a unique solution $v$ to (3.3) on $J$ and

$$\|v^\pm - e^{\pm i(t-t_0)}v(t_0)\|_{X^{1/2}(J)} \leq \|e^{\pm i(t-t_0)}v(t_0)\|_{Z_\pm(J)}^3 + 2\|e\|_{N_\pm(J)}. \quad (3.15)$$

Next we claim that there exists $\varepsilon_1 = \varepsilon_1(K)$ such that if for some $I_k = (t_k, t_{k+1})$ it holds that

$$\|v^\pm\|_{Z(I_k)} \leq \varepsilon \leq \varepsilon_1 \quad (3.16)$$

then

$$\|e^{\pm i(t-t_k)}v^\pm(t_k)\|_{Z_\pm(I_k)} \leq C(1 + K)(\varepsilon + \|e\|_{N_\pm(I_k)})^3 \quad (3.17)$$

Indeed, we let $h(s) := \|e^{\pm i(t-t_k)}v^\pm(t_k)\|_{Z_\pm(t_k, t_{k+s})}$. Let $J_k = [t_k, t') \subset I_k$ be the largest interval such that $h(s) \leq \delta_1/2$, for $\delta_1(K)$ as above. Then by Duhamel's formula

$$\|e^{\pm i(t-t_k)}v^\pm(t_k)\|_{Z(t_k, t_k+s)} \leq \|v^\pm\|_{Z(t_k, t_k+s)} + \|v^\pm - e^{\pm i(t-t_k)}v^\pm(t_k)\|_{X^{1/2}(t_k, t_{k+s})}$$

$$\leq \varepsilon + h(s)^{3/2} + 2\|e\|_{N_\pm(I_k)}.$$

By definition,

$$h(s) \leq \|e^{\pm i(t-t_k)}v^\pm(t_k)\|_{Z(t_k, t_k+s)}^{3/2} \|e^{\pm i(t-t_k)}v^\pm(t_k)\|_{X^{1/2}(t_k, t_{k+s})}^{1/2}.$$
hence by (3.15), the boundedness of the free evolution in $X^{1/2}_\pm$ and (3.13),
\[
h(s) \leq \left( \varepsilon + h(s)^{\frac{3}{2}} + 2\|e\|_{N_\pm(I_k)} \right) K^{\frac{1}{2}}
\]
\[
\leq C(1 + K)(\varepsilon + \|e\|_{N_\pm(I_k)})^{3/4} + C(1 + K)h(s)^{3/2},
\]
and we can conclude the claim provided $\varepsilon_1$ is chosen sufficiently small. Let now $I_k$ be an interval such that
\[
\|e^{\pm i(t-t_k)}(\nabla) v^\pm(t_k)\|_{L^\infty(I_k)} + \|v^\pm\|_{L^\infty(I_k)} \leq \varepsilon \leq \varepsilon_0
\]
(3.18)
\[
\|e\|_{N_\pm(I_k)} \leq \varepsilon,
\]
it holds by the above considerations that $\|v^\pm\|_{X^{1/2}_\pm(I_k)} \leq K + 1$. Fix such an interval and let $u$ be a solution to (3.3) defined on an interval $J_u \ni t_k$ with
\[
\|(u(t_k) - v(t_k), v(t_k))\|_{H^{1/2}} \leq \varepsilon_0.
\]
Set $\phi = u - v$, then $\phi$ solves the difference equation
\[
\begin{cases}
  v_{tt} - \Delta v + \phi + (v + \phi)^3 - v^3 - e = 0, & u : \mathbb{R} \times \mathbb{T}^3 \to \mathbb{R} \\
  (\phi, \partial_t \phi)|_{t=t_k} = (u(t_k) - v(t_k), \partial_t u(t_k) - \partial_t v(t_k)) \in H^{1/2}(\mathbb{T}^3),
\end{cases}
\]
and we can identify $\phi$ with $(\phi^+ , \phi^-)$ as before. Let $J_k = [t_k , t_k + s] \cap I_k \cap J_u$ be the maximal interval such that
\[
\|\phi^\pm\|_{Z^\pm(I_k)} \leq 10C\varepsilon_0 \leq \frac{1}{10(K + 1)}.
\]
Such an interval exists and is non-empty since $s \mapsto \|\phi^\pm\|_{Z^\pm(t_0,t_0+s)}$ is a continuous function which vanishes at $s = 0$. Similarly to the argument used in the standard Strichartz stability theory, by Proposition 3.15 we can then estimate
\[
\|\phi^\pm\|_{X^{1/2}_\pm(J_k)} \leq \|e^{\pm i(t-t_k)}(\nabla) (u^\pm(t_k) - v^\pm(t_k))\|_{X^{1/2}_\pm(J_k)} + \|(v + \phi)^3 - v^3\|_{N_\pm(J_k)} + \|e\|_{N_\pm(J_k)} \leq \varepsilon_0 \|\phi^\pm\|_{X^{1/2}_\pm(J_k)} + \|e\|_{N_\pm(J_k)}.
\]
Thus, if \( \epsilon_0 \) is sufficiently small,

\[
\| \phi^\pm \|_{Z^\pm(I_k)} \leq C \| \phi^\pm \|_{X^{1/2}_\pm(I_k)} \leq 8C \epsilon_0,
\]

(3.20)

and in fact \( J_k = I_k \cap J_u \). Moreover, \( u \) can be extended to all of \( I_k \) by the remark after Proposition 3.16, and further (3.19) and (3.20) hold on all of \( I_k \). We conclude the proof by splitting \( I \) into subintervals such that

\[
\| v^\pm \|_{Z(I_k)} \leq \epsilon_2, \quad \| e \|_{N^\pm(I_k)} \leq \kappa \epsilon_2.
\]

On each interval (3.16) holds, hence by (3.17), (3.18) also holds, and as above we conclude that (3.19) and (3.20) hold. This concludes the proof of the proposition. \( \Box \)

**Remark 3.6.** If we consider the nonlinear Klein-Gordon equation with truncated nonlinearity,

\[
\begin{cases}
(u_N)^{tt} - \Delta u_N + u_N + P_{\leq N}(P_{\leq N} u_N)^3 = 0, & u : \mathbb{R} \times T^3 \to \mathbb{R} \\
(u_N, \partial_t u_N)|_{t=0} = (u_0, u_1) \in \mathcal{H}^{1/2}(T^3),
\end{cases}
\]

then all the estimates from Sections 3.4 and 3.5 go through with constants uniform in the truncation parameter. Indeed, set \( G(u) = F(P_N u) \) and with \( v_N = P_{\leq N} v_N \), we can estimate

\[
\| P_{\leq N} G(u) \|_{N(I)} \leq \sup_{\| v_N \|_{\gamma/2} = 1} \left| \int \int P_N G(u) v_N \right| = \sup_{\| v_N \|_{\gamma/2} = 1} \left| \int \int G(u) v_N \right|,
\]

and all of the previous multilinear estimates go through as before. Moreover, to repeat the stability argument above, we only need

\[
\| P_{\leq N} v \|_{Z(I)} + \| v \|_{L^\infty_t \mathcal{H}^{1/2}_\pm(\mathbb{R} \times T^3)} \leq K.
\]

Moreover, if one only needs to compare low frequencies, it suffices to require

\[
\| P_{\leq N}(v_0 - u_0, v_1 - u_1) \|_{\mathcal{H}^{1/2}(T^3)} + \| P_{\leq N} e \|_{N(I)} \leq \epsilon < \epsilon_1.
\]

This statement should be compared with Remark 4.20 and Lemma 4.35.
3.6 A low frequency equation

The next propositions should be compared with Proposition 5.1 in [20]. As we are at the critical level, we cannot hope to achieve the gain of derivatives for the Strichartz estimates, as was obtained in Theorem 4.3 in [20]. On the other hand, we can still obtain some decay by exploiting the improved Strichartz estimates from Proposition 3.9 provided we are able to create a scale separation between low and high-frequencies. In the sequel, where we denote errors of a given order, this is always understood to be in the \( N(I) \) norm defined in (3.12). In the sequel we will write \( F(u, v, w) = uvw \) as a means to record to various combinations in the nonlinearity.

**Proposition 3.18.** Let \( R, T > 0 \) and let \( u \) be a solution to the cubic nonlinear Klein-Gordon equation for initial data \((u_0, u_1) \in B_R \subset \mathcal{H}^{1/2}\) which satisfies

\[
\|u\|_{L^4_t(\mathbb{R}^2 \times \mathbb{R}^3)} \leq K. \tag{3.21}
\]

Let \( 1 \leq N' \ll N_* \). Then there exists \( M \in [N', N_*] \) such that the low frequency component \( u_{l0} = P_{<M}u \) satisfies the perturbed cubic nonlinear Klein-Gordon equation

\[
\Box u_{l0} + u_{l0} = P_{\leq M}F(u_{l0}, u_{l0}, u_{l0}) + O_{K, R, T} \left( \frac{\log N_*}{N'} \right)^{\theta}. \tag{3.22}
\]

**Proof.** In this proof we allow the value of the parameter \( \theta > 0 \) to change from line to line, and we allow implicit constants to depend on the various parameters involved. Given dyadic frequencies \( N', N_* \), the frequency interval \([N', N_*]\) contains \( O \left( \frac{\log (N_*/N')}{{\log}^2 (N_*/N')} \right) \) intervals of the form

\[
I_\alpha := [N'(\log(N_*/N'))^\alpha, N'(\log(N_*/N'))^\alpha+2] \quad \text{for} \quad \alpha = 2k, k \in \mathbb{N}.
\]

By definition of the \( Y^s \) norm (3.5), for any subset \( S \subset \mathbb{Z}^3 \),

\[
\|P_S u\|_{Y^s} \simeq \left( \sum_{Q_N \cap S \neq \emptyset} \langle N \rangle^{2s} \|P_Q u_{l0}\|_{L^{\infty}\times \mathbb{R}^3}^2 \right)^{1/2}.
\]

Since \( L^4_{t,x} \) bounds imply \( X^{1/2} \) bounds by Proposition 3.12, and \( X^{1/2} \hookrightarrow Y^{1/2} \), we obtain \( Y^{1/2} \) bounds for solutions of the nonlinear Klein-Gordon given the assumption (3.21). Consequently, we can decompose the sum in the definition of \( \|u\|_{Y^{1/2}} \) into the blocks \( I_\alpha \) and we conclude that there must be some frequency interval where
Fix this \( \alpha \). We introduce the notation \( M := N'((\log(N*/N'))^\alpha \), and we define

\[
\begin{align*}
    u_{lo} &= P_{\leq M}, \\
    u_{med} &= (P_{\leq M((\log(N*/N'))^2 - P_{\leq M})u,} \\
    u_{hi} &= (1 - P_{\leq M((\log(N*/N'))^2)u}.
\end{align*}
\]

Let \( a, b, c \in \{lo, med, hi\} \), then we can decompose the nonlinearity as

\[
P_{\leq M}F(u, u, u) = P_{\leq M}F(u_{lo}, u_{lo}, u_{lo}) + \sum_{\max(a, b, c) \geq med} P_{\leq M}F(u_a, u_b, u_c).
\]

By (3.23) and (3.10), any term involving \( u_{med} \) will be an error term. Hence, we only need to consider terms with \( u_{hi} \) and \( u_{lo} \) components. We will use largely the same analysis used to prove Theorem 3.14, but in this case, the dual function will always be localized to low frequencies. First we consider

\[
P_{\leq M}F(u_{lo}, u_{hi}, u_{hi}) = \sum_{N_1, N_2, N_3, N_0 \leq M} P_{N_0}P_{\leq M}F(u_{N_3}, u_{N_2}, u_{N_1}),
\]

which we will treat with similar estimates to those used to prove (3.9). By the convolution requirement, the term above will vanish unless \( N_i \sim N_j \) for some \( i \neq j \), hence we may assume without loss of generality that

\[
N_3 \lesssim N_2 \lesssim N_1, \quad N_1 \sim N_2,
\]

and we bound

\[
\left| \int \int u_{N_3} u_{N_2} u_{N_1} v_{N_0} dx dt \right| \leq \| u_{N_3} u_{N_1} \|_{L^2_t L^2_x} \| u_{N_1} v_{N_0} \|_{L^2_t L^2_x},
\]

or

\[
\lesssim \left( \frac{N_1 N_2}{N_3 N_0} \right)^\delta \| u_{N_1} \|_{V^1_{x_2}} \| u_{N_2} \|_{V^2_{x_2}} \| u_{N_3} \|_{V^3_{x_2}} \| v_{N_0} \|_{V^0_{x_0}}.
\]

Applying Cauchy-Schwarz in \( N_0, N_3 \lesssim M \), we obtain

\[
\begin{align*}
    \sup_{\| w \|_{Y^{1/2}_x} = 1} \sum_{N_0 \leq M} \sum_{N_1, N_2 \sim hi} \left( \frac{N_1 N_2}{N_3 N_0} \right)^\delta \| u_{N_1} \|_{V^1_{x_2}} \| u_{N_2} \|_{V^2_{x_2}} \| u_{N_3} \|_{V^3_{x_2}} \| v_{N_0} \|_{V^0_{x_0}} \\
    \lesssim \| u_3 \|_{Y^{1/2}_{x_3}} \sum_{N_1, N_2 \sim hi} \frac{M^{1-2\delta}}{(N_1 N_2)^{1/2-\delta}} \| u_{N_1} \|_{V^1_{x_2}} \| u_{N_2} \|_{V^2_{x_2}}.
\end{align*}
\]
In particular, by the restriction
\[ N_2, N_1 \gtrsim M(\log(N*/N'))^2, \] (3.24)
we can bound the above multilinear estimate by
\[ \|u_3\|_{Y^{1/2}} \sum_{N_1, N_2 \sim \sim} \frac{1}{(\log(N*/N'))^{2-4\delta}} (N_1 N_2)^{\frac{1}{2}} \|u_{N_1}\|_{V^2_{x_1}} \|u_{N_2}\|_{V^2_{x_2}}, \]
and by applying Cauchy-Schwarz in \( N_2 \sim N_1 \), we see this term is part of the error in (3.22).

For the term with high-frequencies
\[ P_{\leq M} F(u_{hi}, u_{hi}, u_{hi}) \]
we can once again assume without loss of generality that \( N_3 \lesssim N_2 \sim N_1 \). We obtain
\[ \left| \iint u_{N_3} u_{N_2} u_{N_1} v_{N_0} \, dx \, dt \right| \leq \|u_{N_3} u_{N_2}\|_{L^2_{t,x}} \|u_{N_1} v_{N_0}\|_{L^2_{t,x}} \leq \left( \frac{N_1 N_2}{N_0 N_3} \right)^{\delta N_3} N_0 \|u_{N_1}\|_{V^2_{x_1}} \|u_{N_2}\|_{V^2_{x_2}} \|u_{N_3}\|_{V^2_{x_3}} \|v_{N_0}\|_{V^2_{x_0}}. \]

By Cauchy-Schwarz in \( N_0 \) and \( N_3 \), we obtain
\[ \sup_{\|u\|_{Y^{1/2}}=1} \sum_{N_0 \leq M} \sum_{N_1, N_2, N_3 \sim \sim} \left( \frac{N_1 N_2}{N_0 N_3} \right)^{\delta N_3} N_0 \|u_{N_1}\|_{V^2_{x_1}} \|u_{N_2}\|_{V^2_{x_2}} \|u_{N_3}\|_{V^2_{x_3}} \|v_{N_0}\|_{V^2_{x_0}} \lesssim \|u_3\|_{Y^{1/2}} \sum_{N_1, N_2 \sim \sim} \frac{M^{\frac{1}{2} - \delta}}{N_1^{\frac{1}{2} - \delta}} (N_1 N_2)^{\frac{1}{2}} \|u_{N_1}\|_{V^2_{x_1}} \|u_{N_2}\|_{V^2_{x_2}}, \]
and once again by (3.24) and Cauchy-Schwarz in \( N_2 \sim N_1 \) we conclude that such expressions contribute to the error term in (3.22). Finally, we note that by the convolution requirement, the term
\[ P_{\leq M} F(u_{io}, u_{io}, u_{hi}) \]
vanishes provided \( \log(N*/N')^2 > 8 \), which concludes the proof. \( \square \)

**Remark 3.7.** The key difference between the proofs of Propositions 3.18 and 3.13 is that the low frequency is always bounded by \( M \) in Propositions 3.18, hence when we apply Cauchy-Schwarz in
the low frequency, we only lose a factor of \( M \), instead of the next largest frequency. This fact, together with the manufactured frequency separation, enables us to achieve the necessary decay.

### 3.7 Proof of Theorem 3.1

Finally we arrive at the proof of the main result of this chapter.

**Proof of Theorem 3.1.** Let \( u \) and \( \tilde{u} \) denote the solutions to (3.1) with initial data \((u_0, u_1)\) and \((\tilde{u}_0, \tilde{u}_1)\), respectively. Let \( K > 0 \) be such that

\[
\|u\|_{L^4_t x([0,T] \times \mathbb{T}^3)} + \|\tilde{u}\|_{L^4_t x([0,T] \times \mathbb{T}^3)} \leq K.
\]

Since \( u_0 = P \leq M u_0 \), and similarly for \( \tilde{u}_0 \), the same bounds hold for the low frequency components. By Proposition 3.18, the low frequency component satisfies the equation

\[
\Box u_0 + u_0 = P \leq M F(u_0, u_1, u_0) + O_{K, R, T}((\log(N_*/N'))^{-\theta})
\]

and similarly for \( \tilde{u}_0 \). Since \( (u_0 - \tilde{u}_0)|_{t=0} = 0 \), let \( N_* \) be chosen sufficiently large depending on \( R, K \) and \( T \) so that the smallness requirement of Proposition 3.17 is satisfied. Then the result follows from the stability theory, Remark 3.5 and the fact that the bounds for the equation with truncated nonlinearity are uniform in \( N \). \(\square\)
Chapter 4

Symplectic non-squeezing for the cubic NLKG on $\mathbb{T}^3$

4.1 Introduction

We consider the behaviour of solutions to the initial-value problem for the periodic defocusing cubic nonlinear Klein-Gordon equation

$$\begin{cases}
    u_{tt} - \Delta u + u + u^3 = 0, & u : \mathbb{R} \times \mathbb{T}^3 \rightarrow \mathbb{R} \\
    (u, \partial_t u)_{t=0} = (u_0, u_1) \in H^{\frac{1}{2}}(\mathbb{T}^3) \times H^{-\frac{1}{2}}(\mathbb{T}^3) =: \mathcal{H}^{1/2}(\mathbb{T}^3),
\end{cases}$$

(4.1)

where $H^{\frac{1}{2}}(\mathbb{T}^3)$ is the usual inhomogeneous Sobolev space. Recall that if we consider the general power-type nonlinear Klein-Gordon equation

$$u_{tt} - \Delta u + u + |u|^{\rho-1} u = 0,$$

then $s_c := \frac{d}{2} - \frac{2}{\rho-1}$ is the critical regularity for the equation. In (4.1) we have $\rho = d = 3$, hence we are interested in studying the Cauchy problem (4.1) at the critical regularity.

In this chapter, we will study the qualitative behaviour of solutions to (4.1) by investigating symplectic non-squeezing for the flow of this equation. The study of infinite dimensional symplectic capacities and non-squeezing for nonlinear Hamiltonian PDEs was initiated by Kuksin in [38]. There, he extended the definition of the Hofer-Zehnder capacity to infinite dimensional phase spaces and proved the invariance of this capacity under the flow of certain Hamiltonian equations.
with flow maps of the form

\[ \Phi(t) = \text{linear operator} + \text{compact smooth operator}. \tag{4.2} \]

This infinite dimensional symplectic capacity inherits the finite dimensional normalization

\[ \text{cap}(B_r(u_*)) = \text{cap}(C_r(z; k_0)) = \pi r^2, \]

where

\[ B_r(u_*) := \left\{ u \in \mathcal{H}^{1/2}(\mathbb{T}^3) : \| u - u_* \|_{\mathcal{H}^{1/2}} \leq r \right\}, \]

and for \( z = (z_0, z_0) \in \mathbb{C} \) and \( k \in \mathbb{Z}^3 \), the infinite dimensional cylinder

\[ C_r(z; k_0) := \left\{ (u_0, u_1) \in \mathcal{H}^{1/2}(\mathbb{T}^3) : \langle k \rangle |\hat{u}_0(k)| - z_0|^2 + \langle k \rangle^{-1} |\hat{u}_1(k)| - z_0|^2 \leq r^2 \right\}. \]

We will always be working with real-valued functions and the \( \hat{u}_i(k) \) are taken to be real-valued Fourier coefficients. The proof of this normalization in infinite dimensions is an adaptation of the original proof by Hofer and Zehnder which can be found in [30], see [38] for details of the infinite dimensional argument. Consequently, if a flow map \( \Phi \) preserves capacities, one can conclude that squeezing is impossible, namely

\[ \Phi(t)(B_R(u_*)) \not\subset C_r(z; k) \quad \text{if } R < r. \]

Several examples of nonlinear Klein-Gordon equations with weak nonlinearities can readily be shown to be of the form (4.2), see [38]. Symplectic non-squeezing was later proved for certain subcritical nonlinear Klein-Gordon equations in [7] using Kuksin’s framework, see also [59]. Bourgain later extended these results to the cubic NLS in dimension one in [4], where the flow is not a compact perturbation of the linear flow. There, the argument follows from approximating the full equation by a finite dimensional flow and applying Gromov’s finite dimensional non-squeezing result to this approximate flow. Symplectic non-squeezing was also proven for the KdV [20]. In this situation, there is a lack of smoothing estimates in the symplectic space which would allow the infinite dimensional KdV flow to be easily approximated by a finite-dimensional Hamiltonian flow. To resolve this issue, the authors of [20] invert the Miura transform to work on the level of
the modified KdV equation, for which stronger estimates can be established.

As we will see in Section 4.2, the symplectic phase space for any nonlinear Klein-Gordon equation is $H^{1/2}(\mathbb{T}^d)$ for any dimension $d \geq 1$. In particular, for the cubic nonlinear Klein-Gordon equation in dimension three, the symplectic phase space is at the critical regularity, which presents some serious obstructions to using simple modifications of the existing arguments. Kuksin’s approach requires some additional regularity in the compactness estimates. In light of ill-posedness results below the critical space, for instance the results of Christ-Colliander-Tao [17], [39] or [31] adapted to (4.1), there is no way to gain the additional regularity needed. Bourgain’s argument in [4] uses an iteration scheme in which one needs uniform control over time-steps of the iteration, and, once again, this seems to be a genuine obstruction to applying this argument at the critical regularity. Finally, the arguments of [20] depend heavily on the structure of the KdV equation. Additionally, the global well-posedness of (4.1) is not know and there is no uniform control on the local time of existence.

Ultimately, however, we are able to circumvent these difficulties, using a combination of probabilistic and deterministic techniques, which we combine to obtain several deterministic non-squeezing results. Our first result is a local-in-time non-squeezing theorem.

**Theorem 4.1.** Let $\Phi$ denote the flow of the cubic nonlinear Klein-Gordon equation (4.1). Fix $R > 0$, $k_0 \in \mathbb{Z}^3$, $z \in \mathbb{C}$, and $u_* \in H^{1/2}(\mathbb{T}^3)$. For all $0 < \eta < R$, there exists $N \equiv N(\eta, u_*, R, k_0)$ and $\sigma \equiv \sigma(\eta, N, u_*) > 0$ such that for all $0 \leq t \leq \sigma$,

$$\Phi(t)(\Pi_N B_R(u_*)) \not\subseteq C_r(z; k_0) \quad \text{for } r < R - \eta.$$ 

See Remarks 1.2 and 1.1 for a discussion of this result.

In order to state our global-in-time results, we need to introduce the following nonlinear Klein-Gordon equation with truncated nonlinearity

$$
\begin{cases}
(u_N)_t - \Delta u_N + u_N + P_N(P_N u_N)^3 = 0, & u : \mathbb{R} \times \mathbb{T}^3 \to \mathbb{R} \\
(u_N, \partial_t u_N)|_{t=0} = (u_0, u_1) \in H^{1/2}(\mathbb{T}^3),
\end{cases}
$$

(4.3)

where $P_N = P_{\leq N}$ denotes the smooth projection operator defined in (1.12). We obtain the following global-in-time non-squeezing result.
Theorem 4.2. Let \( \Phi \) denote the flow of the cubic nonlinear Klein-Gordon equation (4.1). Fix \( R, T > 0, k_0 \in \mathbb{Z}^3, z \in \mathbb{C}, \) and \( u_* \in \mathcal{H}^{1/2}(\mathbb{T}^3) \). Suppose there exists some \( K > 0 \) such that for all \((u_0, u_1) \in B_R(u_*), \) the corresponding solutions \( u \) to (4.1) and \( u_N \) to (4.3) satisfy
\[
\|u\|_{L^4_t([0,T) \times \mathbb{T}^3)} + \sup_N \|P_N u_N\|_{L^4_t([0,T) \times \mathbb{T}^3)} \leq K.
\]

Then
\[
\Phi(T)(B_R(u_*)) \not\subset C_r(z; k_0) \quad \text{for } r < R.
\]

In particular, if \( B_R(u_*) \subset B_{\rho_0} \) for some sufficiently small \( \rho_0(T) > 0, \) then non-squeezing holds without any additional assumptions on the initial data.

See Remarks 1.3 and 1.4 for some discussion on this result.

4.1.1 Overview of Proof

Almost sure global well-posedness

To prove Theorem 4.1, we rely on an adaptation of the almost sure global well-posedness result from [16]. This enables us to work on a set of full measure, \( \Sigma, \) with respect to a suitable randomization of the initial data, on which the nonlinear Klein-Gordon equation is globally well-posed. We will show that for a certain nested sequence of subsets \( \Sigma_\lambda \subset \Sigma, \) the flow of this equation can be seen as a compact perturbation of a linear flow in the sense used by Kuksin (4.2). We will return to this shortly.

We will now describe the randomization procedure for the initial data. Let \( \{(h_k, l_k)\}_{k \in \mathbb{Z}^3} \) be a sequence of zero-mean, complex-valued Gaussian random variables on a probability space \((\Omega, \mathcal{A}, \mathbb{P})\) with the condition \( h_{-k} = \overline{h_k} \) for all \( k \in \mathbb{Z}^3 \) and similarly for the \( l_k. \) We assume \( \{h_0, \text{Re}(h_k), \text{Im}(h_k)\}_{k \in \mathcal{I}} \) are independent, zero-mean, real-valued Gaussian random variables, where \( \mathcal{I} \) is such that we have a disjoint union \( \mathbb{Z}^3 = \mathcal{I} \cup (-\mathcal{I}) \cup \{0\}, \) and similarly for the \( l_k. \) This set-up ensures that the randomization of real-valued functions is real-valued.

Fix \( (f_0, f_1) \in \mathcal{H}^s(\mathbb{T}^3), \) and define a randomization map \( \Omega \times \mathcal{H}^s \to \mathcal{H}^s \) by
\[
(\omega, (f_0, f_1)) \mapsto (f'_0, f'_1) := \left( \sum_{k \in \mathbb{Z}^3} h_k(\omega) \hat{\phi}_0(k)e^{ik \cdot x}, \sum_{k \in \mathbb{Z}^3} l_k(\omega) \hat{\phi}_1(k)e^{ik \cdot x} \right).
\]
We could similarly take non-Gaussian random variables which satisfy suitable boundedness conditions on their distributions. For any \((f_0, f_1) \in \mathcal{H}^s\), the map (4.4) induces a probability measure on \(\mathcal{H}^s\), given by
\[
\mu_{(f_0, f_1)}(A) = \mathbb{P}(\omega \in \Omega : (f_0', f_1') \in A).
\]
We denote by \(\mathcal{M}^s\) the set of such measures:
\[
\mathcal{M}^s := \{\mu_{(f_0, f_1)} : (f_0, f_1) \in \mathcal{H}^s\}.
\]

**Remark 4.1.** The support of any \(\mu \in \mathcal{M}^s\) is contained in \(\mathcal{H}^s\) for all \(s \in \mathbb{R}\). Furthermore, if for some \(s_1 > s\) we have that \((f_0, f_1) \notin \mathcal{H}^{s_1}\) then the induced measure satisfies \(\mu_{(f_0, f_1)}(\mathcal{H}^{s_1}) = 0\). In other words, this randomization procedure does not regularize at the level of Sobolev spaces in the sense that almost surely, the randomization of a given function is no more regular than what you started with. Moreover, if all the Fourier coefficients of \((f_0, f_1)\) are nonzero, then the support of the corresponding measure \(\mu_{(f_0, f_1)}\) is all of \(\mathcal{H}^s\), that is, \(\mu_{(f_0, f_1)}\) charges every open set in \(\mathcal{H}^s\) with positive measure. As a consequence, for such a measure, sets of full \(\mu\) measure are dense. See [14] for details.

The arguments used to prove the almost sure global well-posedness of the defocusing cubic nonlinear wave equation by Burq and Tzvetkov [16, Theorem 2], apply to the defocusing cubic nonlinear Klein-Gordon equation, with the slight modification that one must consider the inhomogeneous energy functional
\[
\mathcal{E}(w) = \frac{1}{2} \int |\nabla w|^2 + w^2 + (wt)^2 + \frac{1}{2} w^4.
\]
(4.5)

We denote by \(S(t)\) the free evolution for (4.1), given by
\[
S(t)(u_0, u_1) = \cos(t(\nabla))u_0 + \frac{\sin(t(\nabla))}{(\nabla)}u_1,
\]
(4.6)
and we state [16, Theorem 2] adapted to our situation. Moreover, since the Hamiltonian for the nonlinear Klein-Gordon controls the \(L^2_x\) norm of solutions, we no longer need the projection away from constants which appears in [16].

**Theorem 4.3.** Let \(M = \mathbb{T}^3\) with the flat metric and fix \(\mu \in \mathcal{M}^s\), \(0 < s < 1\). Then there exists a full \(\mu\) measure set \(\Sigma \subset \mathcal{H}^s(\mathbb{T}^3)\) such that for every \((u_0, u_1) \in \Sigma\), there exists a unique global
solution \( u \) of the nonlinear Klein-Gordon equation

\[
\begin{aligned}
&\begin{cases}
  u_{tt} - \Delta u + u + u^3 = 0, & u : \mathbb{R} \times T^3 \to \mathbb{R} \\
  (u, \partial_t u) \big|_{t=0} = (u_0, u_1)
\end{cases} \\
\end{aligned}
\]  

(4.7)

satisfying

\((u(t), u_t(t)) \in (S(t)(u_0, u_1), \partial_t S(t)(u_0, u_1)) + C(\mathbb{R}_t; \mathcal{H}_x^1(T^3)).\)

Furthermore, if we denote by

\[
\Phi(t)(u_0, u_1) \equiv (u(t), \partial_t u(t))
\]

the flow thus defined, the set \( \Sigma \) is invariant under the flow \( \Phi(t) \):

\[
\Phi(t)(\Sigma) = \Sigma \quad \forall t \in \mathbb{R}.
\]

Finally, for any \( \varepsilon > 0 \), there exist \( C, c, \theta > 0 \) such that for every \((u_0, u_1) \in \Sigma \) there exists \( M = M(u_0, u_1) > 0 \) such that the global solution of (4.7) given by

\[
u(t) = S(t)(u_0, u_1) + w(t)
\]

satisfies

\[
\| (w(t), \partial_t w(t)) \|_{\mathcal{H}} \leq C(M + |t|)^{\frac{1-\varepsilon}{2} + \varepsilon},
\]

(4.8)

and furthermore, for \( M \) as in (4.8) and each \( \lambda > 0 \),

\[
\mu((u_0, u_1) \in \Sigma : M > \lambda) \leq Ce^{-c\lambda^\theta}.
\]

Let us introduce precisely the subset \( \Sigma \) of full measure we will work with. Let \( 0 < \gamma < \frac{1}{2} \) to be fixed later and define

\[
\begin{aligned}
\Theta_1 &:= \{(u_0, u_1) \in \mathcal{H}^{1/2} : \| S(t)(1 - \Delta)^{\gamma/2}(u_0, u_1) \|_{L^6_x(T^3)} \in L^1_t(\mathbb{R}_t) \} \\
\Theta_2 &:= \{(u_0, u_1) \in \mathcal{H}^{1/2} : \| S(t)(u_0, u_1) \|_{L^6_x(T^3)} \in L^1_t(\mathbb{R}_t) \}.
\end{aligned}
\]

(4.9)

Set \( \Theta := \Theta_1 \cap \Theta_2 \) and let \( \Sigma = \Theta + \mathcal{H}^1 \). The set specified in [16, Theorem 2] imposes an \( L^3_t L^6_x \) condition on the free evolution, however, it will be more convenient to work with the above definition and it is clear that the conditions in (4.9) are stronger. This choice will not change any of
the arguments from [16], and as we will see, $\Sigma$ also has full measure with respect to any $\mu \in \mathcal{M}^s$. In particular, by Remark 4.1, $\Sigma$ is not comprised of initial data which are smoother at the level of Sobolev spaces. We will work on the nested subsets $\Sigma_\lambda \subset \Sigma$, which we define in a following section. These subsets have the property that their union is all of $\Sigma$, and we prove in Proposition 4.22 that there exists $C, c, \theta > 0$ so that for any $\lambda > 0$, they satisfy

$$\mu(\Sigma_\lambda) \geq 1 - Ce^{-c\lambda^\theta}.$$  

Let $\tilde{\Phi}$ denote the nonlinear component of the flow map for the cubic nonlinear Klein-Gordon equation, given by

$$\tilde{\Phi}(t)(u_0, u_1) := \Phi(t)(u_0, u_1) - S(t)(u_0, u_1).$$  

(4.10)

On these subsets, we are able to prove a probabilistic version of the criteria needed for Kuksin's argument from [38], namely we prove bounds of the form

$$\|\tilde{\Phi}(t)(u_0, u_1)\|_{L^\infty_t H_{x}^{s_2}(\mathbb{R}^3)} \lesssim \|(u_0, u_1)\|_{H_{x}^{s_1}}, \quad (u_0, u_1) \in \Sigma_\lambda$$  

(4.11)

for some $s_1 < \frac{1}{2} < s_2$.

**Remark 4.2.** In [7], Bourgain proves the analogue of (4.11) for subcritical nonlinear Klein-Gordon equations via estimates in local-in-time $X^{s,b}$ spaces, see (4.16). The reason Bourgain's estimates fail at the critical regularity is because Strichartz estimates are not available at regularities $s_1 < \frac{1}{2}$, which one would need in order to obtain the smoothing bound (4.11). Generally speaking, $X^{s,b}$ spaces are typically ill-suited for critical problems, resulting in logarithmic divergences in the nonlinear estimates, and problems due to failure of the endpoint Sobolev embedding. The $U^p$ and $V^p$ spaces are a more suitable substitute for critical problem, and we use these to prove a conditional approximation result in Section 3.6. Nonetheless, we choose to prove the probabilistic convergence argument in these spaces, since they are slightly simpler to work with and they are sufficient to exploit the improved probabilistic bounds and to obtain (4.11).

**Probabilistic approximation of the flow map**

Once we have established the probabilistic bounds on the nonlinear component of the flow map, our argument has several key components. The first is an approximation estimate for the flow
map on the subsets $\Sigma_\lambda \subset \Sigma$. We consider (4.3) and we observe that this equation is defined on the whole space and it decouples into a nonlinear evolution on low-frequencies and a linear flow on high frequencies. This equation was used in [13] and we will show in Section 4.4 that this flow is a symplectomorphism when restricted to the finite dimensional subspaces $\Pi_N \mathcal{H}^{1/2}$. Consequently, we are able to show that this equation preserves infinite dimensional capacities. Moreover, we will show that this equation provides a good approximation to the full nonlinear Klein-Gordon equation. More precisely, we will prove the following proposition.

**Proposition 4.4.** Let $\Phi$ and $\Phi_N$ denote the flows of the cubic nonlinear Klein-Gordon equation with full (4.1) and truncated nonlinearities (4.3), respectively. Then for any $T > 0$ and for every $(u_0, u_1) \in \Sigma_\lambda \cap B_R$,

$$\sup_{t \in [0,T]} \|\Phi(t)(u_0, u_1) - \Phi_N(t)(u_0, u_1)\|_{\mathcal{H}^{1/2}(\mathbb{T}^3)} \leq C(\lambda, T, R) \varepsilon_1(N)$$

with $\varepsilon_1(N) \to 0$ as $N \to \infty$.

We remark that this is a global in time approximation result in the critical space, with no restriction on the size of the initial data, which can be viewed as a deterministic statement for initial data in certain subsets of the phase space.

**Remark 4.3.** The dependence of the constants on $R$ in Proposition 4.4 is somewhat artificial, and it can be removed by proving the statements from Section 4.5 with bounds in terms of the nonlinear components of the solutions. Since both $\Phi$ and $\Phi_N$ have the same free evolution, these bounds would suffice to prove Proposition 4.4. We do not undertake this here, however, since we cannot remove this dependence in our other convergence results, and thus it would not improve our main theorems.

An important ingredient in the proof of Theorem 4.1 is the following theorem which states that locally in time, if one restricts to initial data supported on finitely many frequencies, this approximation still holds uniformly.

**Theorem 4.5.** Let $\Phi$ denote the flow of the cubic nonlinear Klein-Gordon equation (4.1) and $\Phi_N$ the flow of (4.3). Fix $R > 0$, $u_* \in \mathcal{H}^{1/2}$ and $N', N \in \mathbb{N}$ with $N'$ sufficiently large, depending on $u_*$. Then there exists a constant $\varepsilon_0(u_*, R) > 0$ such that for all $0 < \varepsilon \leq \varepsilon_0$, there exists $\sigma \equiv \sigma(R, \varepsilon, N')$
such that for any \((u_0, u_1) \in B_R(u_*)\),

\[
\sup_{t \in [0, \sigma]} \| \Phi(t) \Pi_N'(u_0, u_1) - \Phi_N(t) \Pi_N'(u_0, u_1) \|_{H^{1/2}} \leq C(R, u_*) \left[ \epsilon_1(N) + \epsilon \right]
\]

(4.12)

with \(\lim_{N \to \infty} \epsilon_1(N) = 0\).

**Remark 4.4.** The dependence on \(u_*\) in this theorem arises in terms of the minimal \(\lambda\) such that the intersection \(\Sigma_\lambda \cap B_R(u_*)\) is non-empty and in terms of the frequency required so that

\[
\| \Pi_{\geq N} u_* \|_{H^{1/2}} < R.
\]

In particular, by using probabilistic techniques, we obtain a uniform statement at the critical regularity for all initial data in \(B_R(u_*)\). We suspect that the dependence of the constants on the various parameters is non-optimal, but we do not undertake improvements to this theorem here.

**Remark 4.5.** The projection operators in (4.12) allow us to gain the necessary control via the stability theory in order to obtain convergence on all of \(\Pi_N B_R\). It is the combination of Proposition 4.4 with the stability theory which enables us to get a deterministic non-squeezing theorem, regardless of the fact that Proposition 4.4 is a probabilistic result.

**Remark 4.6.** While we can prove global convergence for the flow maps for initial data in \(\Sigma_\lambda\), these sets are not open and have a rather complicated structure. This poses serious problems for the proof of the non-squeezing theorem for several reasons, among which is the fact that it is not at all clear how to compute the capacities \(\text{cap}(\Sigma_\lambda \cap B_R)\), or even ensure that they are positive since the subsets \(\Sigma_\lambda\) have empty interior. In order to upgrade the approximation to open subsets, one must localize in time or obtain a conditional result. We have no prediction at the moment for an estimate of this capacity using Kuksin's definition. Ultimately, it seems likely that an alternative definition of an infinite dimensional symplectic capacity might be a more suitable approach.

**Remark 4.7.** As an easy consequence of the type of arguments used to prove Theorem 4.5, we obtain well-posedness for long times on open subsets, as well as topological genericity for the set of initial data for which this equation is globally well-posed. Namely, we can write the set of initial data for which we can prove global well-posedness as the intersection of dense, open sets. Moreover, these arguments can be modified and combined with Gaussian measure considerations to obtain some qualitative information about the initial data which fail to lie in these open subsets.
Non-squeezing via Gromov’s Theorem

As in [4] and [20], we will prove non-squeezing for the full equation by using this approximation theorem and Gromov’s finite dimensional non-squeezing theorem, which we quote here for the symplectic space $(\mathbb{R}^{2N}, \omega_0)$, where $\omega_0$ is the standard symplectic form on $\mathbb{R}^{2N}$, see [30] for details.

**Theorem 4.6** (Gromov’s non-squeezing theorem, [27]). Let $R$ and $r > 0$, $z \in \mathbb{C}$, $0 \leq k_0 \leq N$, and $x_* \in \mathbb{R}^{2N}$. Let $\phi$ be a symplectomorphism defined on $B_R(x_*) \subset \mathbb{R}^{2N}$, then

$$\phi(B_R(x_*)) \not\subset C_r(z; k_0)$$

for any $r < R$.

We would like to take $N \to \infty$ in Theorem 4.6. To do so, we use Theorem 4.5 to obtain the necessary uniform local control.

The proof of Theorem 4.1 follows easily once we have proven Theorem 4.5. Indeed, fix parameters $R > 0$, $k_0 \in \mathbb{Z}^3$, $z = (z_0, z_0) \in \mathbb{C}$, $u_* \in \mathcal{H}^{1/2}$ and $0 < \eta < R$. Let $\varepsilon_0(R, u_*)$ be as in the statement of Theorem 4.5, and we fix $N > |k_0|$ sufficiently large and $0 < \varepsilon < \varepsilon_0$ sufficiently small so that for $\varepsilon_1(N)$ and $C = C(R, u_*)$ as in (4.12) we have

$$C\varepsilon_1(N) < \frac{\eta}{4} \quad \text{and} \quad C\varepsilon < \frac{\eta}{4}.$$  

Let $\sigma > 0$ be such that the conclusions of Theorem 4.5 hold with $N' = 2N$ and $\varepsilon_0$ as above. Note that, in particular, this choice ensures that $\Phi_N$ is a true symplectomorphism, which does not hold for $N' \ll N$, see Proposition 4.21. Then for any $(u_0, u_1) \in B_R(u_*)$,

$$\sup_{t \in [0, \sigma]} \|\Phi(t)\Pi_{2N}(u_0, u_1) - \Phi_N(t)\Pi_{2N}(u_0, u_1)\|_{\mathcal{H}^{1/2}} < \frac{\eta}{2}. \quad (4.13)$$

Let $r < R - \eta$ and define

$$\varepsilon_2 := \frac{R - r}{2} > \frac{\eta}{2}.$$  

The proof of Proposition 4.21 demonstrates that $\Phi_N$ is a symplectomorphism on $\Pi_{2N}\mathcal{H}^{1/2}$ with the symplectic structure which is compatible with that of the full flow on $\mathcal{H}^{1/2}$. Then we can find
initial data \((v_0, v_1) \in \Pi_2 N B_R(u_*) \subset B_{R_1}\) such that
\[
\left( (k_0) |\Phi_N(\sigma)(v_0, v_1)(k_0) - z_0|^2 + (k_0)^{-1} |\partial_0 \Phi_N(\sigma)(v_0, v_1)(k_0) - z_1|^2 \right)^{1/2} > r + \varepsilon_2. \tag{4.14}
\]
By triangle inequality, \((4.13)\) and \((4.14)\) we obtain
\[
\left( (k_0) |\Phi(\sigma)(v_0, v_1)(k_0) - z_0|^2 + (k_0)^{-1} |\partial_0 \Phi(\sigma)(v_0, v_1)(k_0) - z_1|^2 \right)^{1/2} > r + \varepsilon_2 - \frac{\eta}{2} > r,
\]
which concludes the proof.

Remark 4.8. As can be seen in this proof, the parameter \(\eta\) provides a lower bound for the accuracy needed in the approximation. This allows us to ensure that we can obtain some sort of uniform lower bound on the time in order to prove non-squeezing.

Remark 4.9. If we wish to use Theorem 4.2 to conclude non-squeezing for the flow in the case that a global flow is defined, we need to be a bit careful. There are several ways to go about this, first, we can perform the same argument taking \(N\) sufficiently large depending on the profile of \(u_*\), so as to guarantee that for some choice of \(\varepsilon > 0\)
\[
\|u_* - \Pi_2 N u_*\|_{\mathcal{H}^{1/2}} < \varepsilon,
\]
then we can find initial data \((u_0, u_1) \in \Pi_2 N B_{R-\varepsilon}(u_*),\) and hence
\[
\|u_* - (u_0, u_1)\|_{\mathcal{H}^{1/2}} \leq R.
\]
It is also the case that \(\Pi_2 N B_R(u_*) \subset B_{R_1}(u_*)\) for some \(R_1 > 0\), so we can only conclude that the flow does not squeeze some larger ball of initial data into the cylinder.

4.1.2 Conditional non-squeezing
In Section 3.6, we proved an approximation result, conditional on uniform Strichartz bounds for solutions. The proof is based on an approximation result for the nonlinear Klein-Gordon equation which quantifies the principle that perturbing initial data in high-frequencies does not affect the low frequencies of the corresponding solutions by too much. Ultimately, by a similar argument, we obtain the following approximation result.
Theorem 4.7. Let $\Phi$ and $\Phi_N$ denote the flows of the cubic nonlinear Klein-Gordon equation with full (4.1) and truncated (4.3) nonlinearities, respectively. Fix $T, R > 0$ and suppose there exists some $K > 0$ such that for all $(u_0, u_1) \in B_R(u_*)$, the corresponding solutions $u$ to (4.1) and $u_N$ to (4.3) satisfy
\[ \|u\|_{L^4(\mathbb{R}^3)} + \sup_N \|P_N u_N\|_{L^4(\mathbb{R}^3)} \leq K. \]
Then for all $\varepsilon > 0$ and any $N' \in \mathbb{N}$,
\[ \sup_{t \in [0, T]} \|P_N' (\Phi(t)(u_0, u_1) - \Phi_N(t)(u_0, u_1))\|_{H^{1/2}} < \varepsilon \]
for $N = N(N', \varepsilon, R, T, K)$ sufficiently large.

The ideas used in this proof are similar to those used in [20] however we mention again that in the current setting, the problem is at the critical regularity. This is also reflected in the fact that the constants in the above theorem depend on the Strichartz norm of the solutions and not just the Sobolev norm, as in [20, Theorem 1.3]. The point of Theorem 4.7 is that if one only needs to compare low frequencies, as is the case when proving non-squeezing, some decay can be regained even though we are in the critical setting.

The proof of Theorem 4.2 follows from this approximation result using similar argument to those in the proof of Theorem 4.1, once one uses Theorem 3.1 to restrict to a finite dimensional subspace of initial data. We refer the reader to the proof of Theorem 1.5 in [20] for a similar argument. Alternatively, to conclude non-squeezing from Theorem 4.7, we can use the fact that $\Phi_N$ preserves symplectic capacities, as we do for the proof of Theorem 4.8 in Section 4.7.2

4.1.3 Probabilistic non-squeezing

Finally, we present one last version of the non-squeezing theorem which is a direct application of Theorem 4.7. In contrast to Theorem 4.1, we do not need to consider the finite dimensional projection of the ball and the result we obtain is for large times. In exchange for this, however, we must restrict ourselves to initial data sufficiently close to elements of $\Sigma_\lambda$ and the control we obtain over the diameter of the cylinder is not as good.

Theorem 4.8. Fix $\mu \in \mathcal{M}^{1/2}$ and let $\Phi$ denote the flow of the cubic nonlinear Klein-Gordon equation (4.1). Let $R, T > 0$, $k_0 \in \mathbb{Z}^3$, $z \in \mathbb{C}$, and $u_* \in H^{1/2}(\mathbb{T}^3)$. Then there exists $\theta > 0$ such
that for every \( \varepsilon > 0 \) there exists an open set \( U_\varepsilon \) with

\[
\mu(U_\varepsilon) \geq 1 - e^{-1/\varepsilon}
\]

and such that

\[
\Phi(T)(U_\varepsilon \cap B_R(u_*)) \not\subset C_r(z; k_0),
\]

for all \( r > 0 \) with \( \pi r^2 < \text{cap}(U_\varepsilon \cap B_R(u_*)) \).

**Remark 4.10.** The capacity \( \text{cap}(U_\varepsilon \cap B_R(u_*)) \) is positive since it is the capacity of an open set. In practice, the sets \( U_\varepsilon \) will be constructed by taking the \( p \)-fattening of the subsets \( \Sigma_\lambda \), see Section 4.7 for details. At the moment, we do not have a better bound for this capacity than the trivial bound for open sets. In the case that the flow can be defined on the interval \([0, T]\) for all initial data in \( B_R(u_*) \). The proof of Theorem 4.8 is the only place where we need to use the infinite dimensional symplectic capacity. As a result, the critical stability theory alone is insufficient because the trivial lower bound for \( \text{cap}(U_\varepsilon \cap B_R(u_*)) \) depends on \( \varepsilon \) which is the parameter which yields the control for long-time approximations in the stability theory.

### 4.1.4 Organization of Chapter

In Section 4.2, we provide some background on Symplectic Hilbert spaces and the relation of non-squeezing to the energy transition problem and we introduce the capacity we will work with. In Section 4.3, we collect some deterministic and probabilistic facts. In Section 4.4, we prove local and global properties of solutions to the full equation (4.1) and similarly for the equation with truncated nonlinearity (4.3). In Section 4.5, we prove the boundedness assumptions on the flow maps of these equations. In Section 4.6, we prove Proposition 4.4. Finally in Section 4.7, we prove Theorem 4.5 and Theorem 4.8. Throughout this chapter, we will use the notation \( P_N = P_{\leq N} \).

### 4.2 Symplectic Hilbert spaces

We begin with some background on symplectic Hilbert spaces. We follow the exposition in [38, 20]. Consider a Hilbert space \( \mathcal{H} \) with scalar product \( \langle \cdot , \cdot \rangle \) and a symplectic form \( \omega_0 \) on \( \mathcal{H} \). Let \( J \) be an almost complex structure on \( \mathcal{H} \) which is compatible with the Hilbert space inner product, that is, a bounded self-adjoint operator with \( J^2 = -1 \) such that for all \( u, v \in \mathcal{H}, \omega_0(u, v) = \langle u, Jv \rangle \).
Definition 4.9. We say the pair \((\mathcal{H}, \omega_0)\) is the symplectic phase space for a PDE with Hamiltonian \(H[u(t)]\) if the PDE can be written as \(\dot{u}(t) = J\nabla H[u(t)]\).

Here, \(\nabla\) is the usual gradient with respect to the Hilbert space inner product, defined by

\[
\langle v, \nabla H[u] \rangle \equiv \frac{d}{d\varepsilon} H[u + \varepsilon v] \bigg|_{\varepsilon = 0}.
\]

Definition 4.9 is equivalent to the condition

\[
\omega_0(v, \dot{u}(t)) = \omega_0(v, J\nabla H[u(t)]) = -\langle v, \nabla H[u] \rangle = -\frac{d}{d\varepsilon} H[u + \varepsilon v] \bigg|_{\varepsilon = 0}.
\]

Let \(\langle \nabla \rangle := (1 - \Delta)^{1/2}\) and consider the Hilbert space \(\mathcal{H}^{1/2}(\mathbb{T}^3)\) with the usual scalar product

\[
\langle (u_1, u_2), (v_1, v_2) \rangle \frac{1}{2} := \int_{\mathbb{T}^3} u_1 \cdot \langle \nabla \rangle v_1 + \int_{\mathbb{T}^3} u_2 \cdot \langle \nabla \rangle^{-1} v_2.
\]

For \((u, u_\ell) = (u_1, u_2)\) we can rewrite (4.1) as the system of first order equations

\[
\begin{cases}
(u_1)_t = u_2 \\
(u_2)_t = -(1 - \Delta)u_1 - (u_1)^3.
\end{cases}
\] (4.15)

Define the skew symmetric linear operator

\[
J : \mathcal{H}^{1/2}(\mathbb{T}^3) \to \mathcal{H}^{1/2}(\mathbb{T}^3), \quad J = \begin{pmatrix} 0 & \langle \nabla \rangle^{-1} \\
-\langle \nabla \rangle & 0 \end{pmatrix},
\]

then \(J\) is an almost complex structure on \(\mathcal{H}^{1/2}(\mathbb{T}^3)\) compatible with the symplectic form

\[
\omega_\frac{1}{2}(u, v) := \int_{\mathbb{T}^3} u_1 \cdot v_2 - \int_{\mathbb{T}^3} u_2 \cdot v_1,
\]

that is, setting \(u := (u_1, u_2)\) and \(v = (v_1, v_2)\), we have \(\omega_\frac{1}{2}(u, v) = \langle u, Jv \rangle_\frac{1}{2}\). Then we can write \(\dot{u} = J\nabla H(u)\) for the Hamiltonian

\[
H(u) = \frac{1}{2} \int |\nabla u_1|^2 + \frac{1}{2} \int |u_1|^2 + \frac{1}{2} \int |u_2|^2 + \frac{1}{4} \int |u_1|^4.
\]

In particular, up to modifying the Hamiltonian, these computations holds for all nonlinearities and in all dimensions.
4.2.1 An infinite dimensional symplectic capacity

Kuksin’s construction of a symplectic capacity for an open set $\mathcal{O}$ is based on finite dimensional approximations of this set. It is an infinite dimensional analogue of the Hofer-Zehnder capacity [30]. Before defining this capacity, we first recall the definition of a symplectic capacity on a symplectic phase space $(\mathcal{H}, \omega)$.

**Definition 4.10.** A symplectic capacity on $(\mathcal{H}, \omega)$ is a function $\text{cap}$ defined on open subsets $\mathcal{O} \subset \mathcal{H}$ which takes values in $[0, \infty]$ and has the following properties:

1) Translational invariant: $\text{cap}(\mathcal{O}) = \text{cap}(\mathcal{O} + \xi)$ for $\xi \in \mathcal{H}$.

2) Monotonicity: $\text{cap}(\mathcal{O}_1) \geq \text{cap}(\mathcal{O}_2)$ if $\mathcal{O}_1 \supseteq \mathcal{O}_2$.

3) 2-homogeneity: $\text{cap}(\tau \mathcal{O}) = \tau^2 \text{cap}(\mathcal{O})$.

4) Non-triviality: $0 < \text{cap}(\mathcal{O}) < \infty$ if $\mathcal{O} \neq \emptyset$ is bounded.

In finite dimensions, the symplectic capacity is an important symplectic invariant, namely for a given symplectomorphism, $\varphi$, and an open subset, $\mathcal{O} \subset \mathcal{H}$, we have that $\text{cap}(\varphi(\mathcal{O})) = \text{cap}(\mathcal{O})$. In [38], Kuksin shows that the infinite dimensional symplectic capacity he constructs is invariant under the flow of certain nonlinear dispersive Hamiltonian equations. We will now define Kuksin’s infinite dimensional capacity. For a given Darboux\(^1\) basis of $\mathcal{H}$, let $\mathcal{H}_N$ denote the span of the first $N$ basis vectors. Similarly, we use the notation $\mathcal{O}_N$ for any subset $\mathcal{O}$ projected onto these basis vectors. We collect a few definitions.

**Definition 4.11 (Admissible function).** Consider a smooth function $f \in C^\infty(\mathcal{O})$ and let $m > 0$. The function $f$ is called $m$-admissible if

i) $0 \leq f \leq m$ everywhere.

ii) $f \equiv 0$ in a nonempty subdomain of $\mathcal{O}$.

iii) $f|_{\partial \mathcal{O}} \equiv m$ and the set $\{f < m\}$ is bounded and the distance from this set to $\partial \mathcal{O}$ is $d(f) > 0$.

**Definition 4.12 (Fast function).** Let $f_N := f|_{\mathcal{O}_N}$ and consider the corresponding Hamiltonian vector field $U_{f_N}$, that is, for $z, v \in \mathcal{H}_N$, we have

$$\omega(U_{f_N}(z), v) = \nabla f_N(z)v.$$\(^\text{1}\)

---

\(^{1}\)a Darboux basis of $\mathcal{H}$ is a basis $(u_1, \ldots, v_1, \ldots)$ such that $\omega(u_i, v_j) = \delta_{ij}$.
A periodic trajectory of \( U_{f_N} \) is called fast if it is not a stationary point and its period \( T \) satisfies \( T \leq 1 \). An admissible function \( f \) is called fast if there exists \( N_0(f) \) such that for all \( N \geq N_0 \), the vector field \( U_{f_N} \) has a fast trajectory.

**Remark 4.11.** In light of the fact that \( J^2 = -I \), we also have the representation

\[
U_{f_N}(z) = J \nabla f_N(z).
\]

With these definitions, we are now ready to state the definition of Kuksin's infinite dimensional capacity.

**Definition 4.13.** For an open, nonempty domain \( O \subset \mathcal{H} \), its capacity \( \text{cap}(O) \) equals

\[
\text{cap}(O) = \inf \{ m_* \mid \text{each } m\text{-admissible function with } m > m_* \text{ is fast} \}.
\]

In [38], it is shown that this definition satisfies the axioms of a capacity, that is the criteria of Definition 4.10, and while the construction of this capacity depends on the choice of Darboux basis, if one chooses another basis which is quadratically close\(^2\) to the first, then the capacity does not change.

### 4.3 Preliminaries

#### 4.3.1 Deterministic preliminaries

We recall the definition of the \( X^{s,b} \) spaces with norm

\[
\| u \|_{X^{s,b}(\mathbb{R} \times \mathbb{T}^3)} = \| (n)^s (|\tau| - (n))^b \tilde{u}(n, \tau) \|_{L_2^p L^q_\delta}.
\]

We will also work with the local-in-time restriction spaces \( X^{s,b,\delta} \), which are defined by the norm

\[
\| u \|_{X^{s,b,\delta}} = \inf \{ \| \tilde{u} \|_{X^{s,b}(\mathbb{R} \times \mathbb{T}^3)} : \tilde{u}|_{[-\delta, \delta]} = u \}.
\]

We refer the reader to Appendix A for more details.

We will record some facts about the projection operator (1.12). In [13], these operators were used to define an approximating equation for the cubic nonlinear wave equation and we similarly use\(^2\) that is, some other Darboux basis \( \{ \psi_j^\pm \} \) such that \( \sum_j \| \varphi_j^\pm - \psi_j^\pm \| < \infty \).
them to define the equation with truncated nonlinearity (4.3). We use this smoothed projection instead of the standard truncation because we will need to exploit the fact that this family of operators has uniform $L^p$ bounds.

**Lemma 4.14.** Let $M$ be a compact Riemannian manifold and let $\Delta$ be the Laplace Beltrami operator on $M$. Let $1 \leq p \leq \infty$. Then $P_N \equiv \psi(-N^{-2}\Delta) : L^p(K_1) \to L^p(K_1)$ is continuous and there exists $C > 0$ such that for every $N \geq 1$,

$$\|P_N\|_{L^p \to L^p} \leq C.$$ 

Moreover, for all $f \in L^p(K_1)$, $P_N f \to f$ in $L^p$ as $n \to \infty$.

**Proof.** See [63, Theorem 2.1].

Finally, we will need the following identity to prove the symplectic properties for the truncated nonlinear Klein-Gordon equation.

**Lemma 4.15.** Let $K$ be large enough so that $\Pi_K P_N = P_N$, then

$$\int_{T^3} P_N [(P_N u)^3] \Pi_K v = \int_{T^3} (P_N u)^3 P_N v.$$ 

### 4.3.2 Probabilistic preliminaries

We will record some of the basic probabilistic results about the randomization procedure. Most of these estimates are consequences of the classical estimates of Paley-Zygmund for random Fourier series on the torus. These estimates were used heavily in the works of Burq and Tzvetkov, see especially [13] for proofs.

The large deviation estimate from Proposition 2.4 is the key component in the proof of the following corollary, which states that the free evolution of randomized initial data satisfies almost surely better integrability properties. There is the minor modification in the following that we are dealing with complex random variables $\{h_k\}$ which satisfy the symmetry condition $h_{-k} = \overline{h_k}$. Given that the functions we are randomizing are real-valued, and thus have Fourier coefficients which satisfy an analogous symmetry condition, the arguments for the following results go through unchanged. Recall in the sequel that $S(t)$ denotes the free evolution for the nonlinear Klein-Gordon equation, defined in (4.6).
Corollary 4.16 (Corollary A.5, [16]). Fix $\mu \in \mathcal{M}^s$ and suppose $\mu$ is induced via the map (4.4) for $(f_0, f_1) \in \mathcal{H}^s$. Let $2 \leq p_1 < \infty$, $2 \leq p_2 < \infty$ and $\delta > 1 + \frac{1}{p_1}$ and $0 < \sigma \leq s$ Then there exist constants $C, c > 0$ such that for every $\lambda > 0$,

$$\mu \left( \{ (u_0, u_1) \in \mathcal{H}^s : \| (t)^{-\delta} S(t)(u_0, u_1) \|_{L_t^{p_1} L_x^{p_2}} > \lambda \} \right) \leq C e^{-c\lambda^2 / \| (f_0, f_1) \|_{H^s}^2}.$$ 

Remark 4.12. We can include the endpoint $p_2 = \infty$ if we restrict to bounding $0 < \sigma < s$ in the statement above, or equivalently to bounding.

$$\| (t)^{-\delta} (1 - \Delta)^{\gamma/2} S(t)(u_0, u_1) \|_{L_t^{p_1} L_x^{p_2}}$$

for any $0 < \gamma < s$. By Sobolev embedding, we are also able to include the endpoint $p_1 = \infty$ in this case. We use this in Section 4.9, see [13, Lemma 2.2].

Given these large deviation estimates, the following result, which enables us to construct the subset $\Sigma$ is a simple corollary.

Corollary 4.17. Let $T > 0$ fixed, $\mu \in \mathcal{M}^s$ and $2 \leq p < \infty$. Then there exists $C = C(T)$ such that for any $0 \leq \gamma < \gamma_1 \leq 1/2$,

$$\mu \left( \{ (v_0, v_1) \in \mathcal{H}^s : \| S(t)(u_0, u_1) \|_{L_t^{p_1} W_x^{2, p_2}} > \lambda \} \right) \leq C e^{-c\lambda^2 / \| (f_0, f_1) \|_{H^s}^2 \gamma_1}.$$ 

Additionally, Corollaries 4.16 and 4.17 imply that the set of initial data which satisfies good local $L^p$ bounds has full $\mu$ measure.

Corollary 4.18. Fix $\mu \in \mathcal{M}^s$ and let $2 \leq p < \infty$ and $0 < \gamma < s$. Then for a set of full $\mu$ measure,

$$\| (1 - \Delta)^{\gamma/2} S(t)(u_0, u_1) \|_{L_t^{p} W_x^{2, p} (\mathbb{R} \times \mathbb{T}^3)} \in L_{\text{loc}}^1 (\mathbb{R}) \quad \| S(t)(u_0, u_1) \|_{L_x^{p} (\mathbb{T}^3)} \in L_{\text{loc}}^1 (\mathbb{R}).$$

In particular, for $\Sigma$ as defined in (4.9), we have $\mu(\Sigma) = 1$.

4.4 Well-posedness theory

We record some global bounds on the solution to the cubic nonlinear Klein-Gordon (4.1). The only new component in this statement is the bounds on the $L^4$ norm of the solution, which follows by
a similar argument to the proof of [16, Proposition 4.1]. We include the proofs of these statements in Appendix 4.9.

**Proposition 4.19.** Let \( 0 < s < 1 \) and let \( \mu \in \mathcal{M}^s \). Then for any \( \varepsilon > 0 \), there exist \( C, c, \theta > 0 \) such that for every \( (u_0, u_1) \in \Sigma \), there exists \( M = M(u_0, u_1) > 0 \) such that the global solution \( u \) to the cubic nonlinear Klein-Gordon equation (4.1) satisfies

\[
\begin{align*}
    u(t) &= S(t)(u_0, u_1) + w(t) \\
    \|(w(t), \partial_t w(t))\|_{\mathcal{H}^1} &\leq C(M + |t|)^{1+\varepsilon} \\
    \|u(t)\|_{L^4(\mathbb{T}^3)} &\leq C(M + |t|)^{\frac{1}{2}+\varepsilon}
\end{align*}
\]

and furthermore \( \mu((u_0, u_1) \in \Sigma : M > \lambda) \leq C e^{-c\lambda^\theta} \).

We now turn to studying the global well-posedness and symplectic properties of the approximating equation (4.3). Let \( P_N \) be the smooth projection operator defined in (1.12). For \( K \) sufficiently large so that \( \|K P_N - P_N\| \leq 1 \), this equation is equivalent to the uncoupled system

\[
\begin{align*}
    \partial_t \Pi_K u_N + (1 - \Delta) \Pi_K u_N + P_N [(P_N u_N)^3] &= 0, \quad (t, x) \in \mathbb{R} \times \mathbb{T}^3 \\
    (\Pi_K u_N, \partial_t \Pi_K u_N)|_{t=0} &= (\Pi_K u_0, \Pi_K u_1) \\
    \Pi_{\geq K}(u_N) &= S(t)(\Pi_{\geq K} u_0, \Pi_{\geq K} u_1)
\end{align*}
\]

which we can write as a first order system as in (4.15).

**Remark 4.13.** As remarked in the introduction, (4.17) is a nonlinear flow for low frequencies and a decoupled linear evolution for high frequencies. In particular, the solution \( u_N \) is supported on all frequencies. We will nonetheless call this the truncated flow for simplicity even though this defines a flow on the whole space.

Global well-posedness for (4.17) follows from local well-posedness by observing that the linear evolution is globally defined and the energy functional

\[
    H_N(\Pi_K(u_N, (u_N)_t)) = \frac{1}{2} \int_{\mathbb{T}^3} |\nabla_x \Pi_K u_N|^2 + (\Pi_K u_N)^2 + (\Pi_K(u_N)_t)^2 + \frac{1}{4} \int_{\mathbb{T}^3} (P_N(u_N))^4
\]

is well defined and conserved under the flow of (4.17) for bounded frequency components. Note that the bounds on the solution depend on the energy of the initial data and consequently are not
uniform in the truncation parameter. Nonetheless, Burq-Tzvetkov [13] proved that if one restricts to initial data \((u_0, u_1) \in \Sigma\), then the nonlinear components of the solutions to the cubic nonlinear wave equation satisfy uniform bounds. As was the case in Theorem 4.3, the proof of these uniform bounds follow for the nonlinear Klein-Gordon equation from the arguments in [13] with only minor modifications.

**Proposition 4.20** (Proposition 3.1, [13]). Let \(0 < s < 1\) and let \(\mu \in \mathcal{M}^\theta\). Then for any \(\varepsilon > 0\), there exist \(C, c, \theta > 0\) such that for every \((u_0, u_1) \in \Sigma\), there exists \(M = M(u_0, u_1) > 0\) such that the family of global solutions \((u_N)_{N \in \mathbb{N}}\) to (4.17) satisfies

\[
\begin{align*}
    u_N(t) &= S(t)(u_0, u_1) + w_N(t) \\
    \|w_N(t), \partial_t w_N(t)\|_{\mathcal{H}^1} &\leq C(M + |t|)^{1+\varepsilon} \\
    \|P_N u_N(t)\|_{L^4(\mathbb{T}^3)} &\leq C(M + |t|)^{\frac{1}{2}+\varepsilon}
\end{align*}
\]

and furthermore \(\mu((u_0, u_1) \in \Sigma : M > \lambda) \leq Ce^{-c\lambda^\theta}\).

The truncated flow maps also preserve symplectic capacities. This is an easy consequence of the fact that, when restricted to bounded frequencies, these maps are finite dimensional symplectomorphisms.

**Proposition 4.21.** The flow maps \(\Phi_N(t)\) preserve symplectic capacities \(\text{cap}(O)\) for any domain \(O \subset \mathcal{H}^{1/2}(\mathbb{T}^3)\).

**Proof.** Let \(K\) be large enough so that \(\Pi_K P_N = P_N\) and consider the Hamiltonian (4.18). For \((v_1, v_2) \in \Pi_K \mathcal{H}^{1/2}(\mathbb{T}^3)\) and \((u_1, u_2)\) which solve (4.17), we have

\[
\begin{align*}
    \frac{d}{d\varepsilon} H_N(\Pi_K (u_1, u_2) + \varepsilon (v_1, v_2)) \bigg|_{\varepsilon=0} &= \frac{d}{d\varepsilon} \left[ \frac{1}{2} \int_{\mathbb{T}^3} (\Pi_K u_1 + \varepsilon v_1)^2 + |\nabla_x (\Pi_K u_1 + \varepsilon v_1)|^2 + (\Pi_K u_2 + \varepsilon v_2)^2 + \frac{1}{4} \int_{\mathbb{T}^3} (P_N (u_1 + \varepsilon v_1))^4 \right] \bigg|_{\varepsilon=0} \\
    &= \int_{\mathbb{T}^3} (\Pi_K u_1) v_1 + \nabla_x (\Pi_K u_1) \nabla_x v_1 + (\Pi_K u_2) v_2 + (P_N u_1)^3 P_N (v_1)
\end{align*}
\]
thus by Lemma 4.15 we obtain

\[ \int_{\mathbb{T}^3} (\Pi_K u_1) v_1 + \nabla_x (\Pi_K u_1) \nabla_x v_1 + (\Pi_K u_2) v_2 + P_N (P_N u_1)^3 \Pi_K v_1 \]

\[ = \int_{\mathbb{T}^3} -\partial_t (\Pi_K u_2) v_1 + \partial_t (\Pi_K u_1) v_2 \]

\[ = -\omega_{1/2} (v_1, v_2, (\partial_t \Pi_K u_1, \partial_t \Pi_K u_2)) \].

Thus the maps $\Phi_N$ are symplectomorphisms on $\Pi_K \mathcal{H}^{1/2}$, denote this restricted map by $\tilde{\Phi}_N$. Since the flow decouples low and high frequencies, we can write

\[ \Phi_N(u) = \tilde{\Phi}_N(\Pi_K u) + e^{tJ\omega} \Pi_{\geq K} u = e^{tJ\omega} (e^{-tJ\omega} \tilde{\Phi}_N(\Pi_K u) + \Pi_{\geq K} u), \]

where $\omega$ is the linear propagator for the Klein-Gordon equation on $\mathcal{H}^{1/2}$. The invariance of the symplectic capacity under the flow follows from [38, Lemma 51], since $e^{-tJ\omega} \tilde{\Phi}_N$ is also a symplectomorphism and that $e^{tJ\omega}$ is an isometry, hence it preserves admissible and fast functions. \hfill \Box

### 4.4.1 Definition and properties of $\Sigma_\lambda$

We recall the definition of $\Sigma$. We let $0 < \gamma < \frac{1}{2}$ to be fixed later and define

\[ \Theta_1 := \{(u_0, u_1) \in \mathcal{H}^{1/2} : \|S(t)(1 - \Delta)^{\gamma/2}(u_0, u_1)\|_{L^6_{\omega}(\mathbb{T}^3)} \in L^1_{\text{loc}}(\mathbb{R}_t)\} \]

\[ \Theta_2 := \{(u_0, u_1) \in \mathcal{H}^{1/2} : \|S(t)(u_0, u_1)\|_{L^\infty_{\omega}(\mathbb{T}^3)} \in L^1_{\text{loc}}(\mathbb{R}_t)\}. \]

Set $\Theta := \Theta_1 \cap \Theta_2$ and let $\Sigma = \Theta + \mathcal{H}^1$. For $(u_0, u_1) \in \Sigma$ we denote by $u(t) = S(t)(u_0, u_1) + w(t)$ the global solution to (4.1). Fix $\epsilon > 0$ and let $C > 0$ be as in the statements of Proposition 4.19 and Proposition 4.20. Define the subsets

\[ E_\lambda := \{(u_0, u_1) \in \Sigma : \|(w(t), \partial_t w(t))\|_{\mathcal{H}^1} \leq C(\lambda + |t|)^{1+\epsilon}\}, \]

\[ H_\lambda := \{(u_0, u_1) \in \Sigma : \|(w_N(t), \partial_t w_N(t))\|_{\mathcal{H}^1} \leq C(\lambda + |t|)^{1+\epsilon}\}, \]

\[ J_\lambda := \{(u_0, u_1) \in \Sigma : \|u\|_{L^4_\omega} \leq C(\lambda + |t|)^{1+\epsilon}\}, \]

\[ K_\lambda := \{(u_0, u_1) \in \Sigma : \|u_N\|_{L^4_\omega} \leq C(\lambda + |t|)^{1+\epsilon}\}, \]

95
and let $M_\lambda$ be the set of $(u_0, u_1) \in \Sigma$ such that for $\gamma$ as above we can find $C = C(T)$ as in Corollary 4.17 such that

$$
\| (1 - \Delta)^{\gamma/2} S(t)(u_0, u_1) \|_{L^6_x([0,T] \times \mathbb{T}^3)} \leq C\lambda.
$$

Setting

$$
\Sigma_\lambda := E_\lambda \cap H_\lambda \cap J_\lambda \cap K_\lambda \cap M_\lambda
$$

we have the following bounds for the measure of this set. The choice of $\gamma > 0$ does not affect the following proposition.

**Proposition 4.22.** Fix $\mu \in \mathcal{M}^{1/2}$ and let $\Sigma_\lambda$ be as defined in (4.19). Then there exists $C, c, \theta > 0$ such that for all $\lambda > 0$ we have

$$
\mu(\Sigma_\lambda) \geq 1 - Ce^{-c\lambda^\theta}. \tag{4.20}
$$

**Proof.** Suppose that $\mu$ is induced by the randomization of $(f_0, f_1) \in \mathcal{H}^{1/2}$. By Proposition 4.19 and Proposition 4.20, there exists $C, c > 0$ such that

$$
\mu(E_\lambda \cap H_\lambda \cap J_\lambda \cap K_\lambda) \geq 1 - Ce^{-c\lambda^\theta},
$$

and by Corollaries 4.16 and 4.17,

$$
\mu(M_\lambda) \geq 1 - Ce^{-c\lambda^\theta/\|F(f_0, f_1)\|^2_{\mathcal{H}^s}}.
$$

Taking intersections and using that these bounds are exponential yields (4.20).

**Remark 4.14.** In the above proof we are implicitly taking advantage of the fact that we have exponential bounds on the measure of the sets in question.

### 4.5 Probabilistic bounds for the nonlinear component of the flow

In this section, we will show boundedness for the nonlinear component of the cubic nonlinear Klein-Gordon equation on the subsets $\Sigma_\lambda \subset \Sigma$. In the sequel we let $F(u) = u^3$. We will begin with the proof of a local boundedness property for the nonlinearity $F(u)$. The argument is based on Strichartz estimates together with the improved averaging effects for the free evolution of initial
data \((u_0, u_1) \in \Sigma_\lambda\), where \(\Sigma_\lambda\) was defined in (4.19), as well as the uniform bounds on the nonlinear component of those global solutions from Proposition 4.19 and Proposition 4.20.

We wish to obtain bounds on \(\|F[u(t)]\|_{X^{s_2-1, -\frac{1}{2}+\delta}}\) for solutions to the cubic nonlinear Klein-Gordon equation (4.1) with initial data \((u_0, u_1) \in \Sigma_\lambda\). We will perform the estimates with \(b = \frac{1}{2} + \delta\) and hence, we need to estimate the expression

\[
\sum_n \int d\tau \frac{|\hat{F}(n, \tau)|^2}{\langle n \rangle^{2(1-s_2)}(\langle |\tau| - \langle n \rangle \rangle)^{2(1-b)}} \right)^{\frac{1}{2}}
\]

(4.21)

for \(F(u) = u^3\) and \(b > 1/2\). If we expand \(F(u)\) for \(u(t) = S(t)(u_0, u_1) + w(t)\), then we need to consider terms of the form \(u^{(1)}u^{(2)}u^{(3)}\) for \(u(t)\) either

(I) the free evolution \(S(t)(u_0, u_1)\) of initial data \((u_0, u_1) \in \Sigma_\lambda \subset \mathcal{H}^{1/2}\), or

(II) the nonlinear component, \(w(t)\), of a solution to (4.1) with initial data \((u_0, u_1) \in \Sigma_\lambda\).

We will refer to these as type (I) or type (II) functions. We define

\[
c_i(n_i, \tau_i) = \langle n_i \rangle^{s_1} \langle |\tau_i| - \langle n_i \rangle \rangle^b |\hat{u}^{(i)}(n_i, \tau_i)|,
\]

then \(\|c\|_{L^1_t \mathcal{F}_{n}^s(\mathbb{R} \times \mathbb{T}^3)} = \|u^{(i)}\|_{X^{s_1, b}(\mathbb{R} \times \mathbb{T}^3)}\). By duality, (4.21) can be estimated by

\[
\sum_{n=n_1+n_2+n_3} \int_{\tau=\tau_1+\tau_2+\tau_3} \prod_{i=1}^3 \frac{c_i(n_i, \tau_i)}{\langle n_i \rangle^{s_1} \langle |\tau_i| - \langle n_i \rangle \rangle^b} \frac{\langle n \rangle^{1-s_2} \langle |\tau| - \langle n \rangle \rangle^{1-b}}{\langle |\tau| - \langle n \rangle \rangle^{1-b}} d\tau
\]

(4.22)

where \(\|v\|_{L^1_t \mathcal{F}_{n}^s(\mathbb{R} \times \mathbb{T}^3)} \leq 1\). We remark that this notation should not be confused with the initial data for (4.1), and it will be clear from the context which we are considering.

We restrict the \(n_i\) and \(n\) to dyadic regions \(|n_i| \sim N_i\) and \(|n| \sim N\). We will implicitly insert a time cut-off with each function but we will omit the notation, since, in the usual way we can take extensions of the \(u_i\) and then take infinums. The ordering of the size of the frequencies will not play a role in this argument. We do not repeat these considerations. Letting

\[
\tilde{U}_{N_i}(n_i, \tau_i) = \frac{c_i(n_i, \tau_i)}{\langle |\tau_i| - \langle n_i \rangle \rangle^b} \chi_{|n_i| \sim N_i}, \quad \tilde{V}_N(n, \tau) = \frac{v(n, \tau)}{\langle |\tau| - \langle n \rangle \rangle^{1-b}} \chi_{|n| \sim N}
\]

(4.23)
we will need to estimate expressions of the form

$$\left( N_1 N_2 N_3 \right)^{-s_1} N^{-(1-s_2)} \int_{\mathbb{R}^4} \int_{T_3^2} \prod_{i=1}^{3} U_{N_i} \cdot V_N \, dx \, dt.$$  \hspace{1cm} (4.24)

We will use expressions (4.22) and (4.24) as starting points in proofs of the subsequent propositions. In the sequel, we will always take the constant $\gamma = s_1$ in our definition of $\Sigma$.

### 4.5.1 Boundedness of the flow map

**Proposition 4.23** (Local boundedness). Consider the cubic nonlinear Klein-Gordon equation (4.1). Then there exists $s_1 < \frac{1}{2} < s_2$ with $s_1, s_2$ sufficiently close to $1/2$ such that for any $\lambda, R, T > 0$, for every $(u_0, u_1) \in \Sigma \cap B_R$ and for any interval $I \subset [0, T]$ with $|I| = \delta$, the nonlinearity satisfies the bound

$$\|F(u)\|_{X^{s_2-1/2+\delta(\cdot \times T^3)}} \leq C(\lambda, R, T) \delta^{s_1} \left( 1 + \|u\|_{X^{s_1-1/2+\delta(\cdot \times T^3)}}^{9/4} \right).$$  \hspace{1cm} (4.25)

where $(u, \partial_t u)$ is the global solution to the cubic nonlinear Klein-Gordon equation (4.1) with initial data $(u_0, u_1)$.

**Proof.** For any solution $u(t) = S(t)(u_0, u_1) + w(t)$, our computations will yield (4.25) with the nonlinear component of the solutions, $w$, on the right-hand side instead of $u$. Now, for any $s, b \in \mathbb{R}$ and any interval $I \subset [0, T]$ with $|I| = \delta$ and $\inf I = t_0$, and $\eta(t)$ a Schwartz time-cutoff adapted to that interval we claim that we it sufficed to obtain (4.25) with the nonlinear components of the solutions on the right-hand side. Indeed,

$$\|\eta(t)w\|_{X^{s,b}(I \times T^3)} = \|\eta(t)(u - S(t)(u_0, u_1))\|_{X^{s,b}(I \times T^3)}$$

$$\lesssim \|\eta(t)S(t)(u_0, u_1)\|_{X^{s,b}(I \times T^3)} + \|\eta(t)u\|_{X^{s,b}(I \times T^3)}$$

$$\lesssim \|(u(t_0), \partial_t u(t_0))\|_{\mathcal{H}^s} + \|\eta(t)u\|_{X^{s,b}(I \times T^3)}.$$

Since the free evolution is bounded in $\mathcal{H}^s$, we have

$$\|(u(t_0), \partial_t u(t_0))\|_{\mathcal{H}^s} \lesssim \|(u_0, u_1)\|_{\mathcal{H}^s} + \|(w, \partial_t w)\|_{L^\infty \mathcal{H}^s([0, T] \times T^3)},$$

and since the terms on the right-hand side of (4.25) are computed with $s = s_1 < 1/2$, then by the
choice of $\Sigma_{\lambda}$, we have that for any such $I \subset [0, T],$

$$\|\eta(t)w\|_{X^{s,b}(I \times T^3)} \lesssim \|\eta(t)u\|_{X^{s,b}(I \times T^3)} + C(\lambda, R, T)$$

which will yield (4.25). The key point here is that on any interval $[0, T]$, the choice of $\Sigma_{\lambda}$ yields uniform control on bounds of the Sobolev norm of solutions. We will not repeat these considerations.

We analyze the different combinations of $u_i$ systematically. The argument only depends on the number of $u_i$ which are of type (I) or type (II), so we will only present one combination from each case.

• Case (A): All $u^{(i)}$ of type (II). Since

$$\|U_{N_i}^2 U_{N_i}^{1-\alpha}\|_{L^{4+\varepsilon_2}} \lesssim \|U_{N_i}\|_{L^{p_{\alpha}}}^\alpha \|U_{N_i}\|_{L^{1/(1-\alpha)}}^{1-\alpha}$$

for $p = \frac{(4+\varepsilon_2)^5}{1-\varepsilon_2}$, we use Hölder’s inequality to estimate (4.22) by

$$(N_1 N_2 N_3)^{-s_1} N^{-(1-s_2)} \int \int \prod_{i=1}^3 \int U_{N_i} \cdot V_N \, dx \, dt$$

$$\lesssim (N_1 N_2 N_3)^{-s_1} N^{-(1-s_2)} \prod_{i=1}^3 \|U_{N_i}\|_{L^{p_{\alpha}}}^\alpha \|U_{N_i}\|_{L^{1/(1-\alpha)}}^{1-\alpha} \|V_N\|_{L^{4-3\varepsilon_1}}.$$

Taking $\alpha = 1/4$, we have $(1-\alpha)5 < 4$ and for $\varepsilon_2$ sufficiently small, $p\alpha < 6$. Provided

$$1 - b > \frac{4 - 6\varepsilon_1}{2(4 - 3\varepsilon_1)} = \frac{\theta_2}{2},$$

we can bound

$$\frac{1}{(|\tau| - \langle n \rangle)^{1-b}} = \frac{1}{(|\tau| - \langle n \rangle)^{1-b-\frac{\varepsilon_2}{4}}} + \frac{1}{(|\tau| - \langle n \rangle)^{\frac{\varepsilon_2}{2}}} + \frac{1}{(|\tau| - \langle n \rangle)^{\frac{\varepsilon_2}{2}}}$$

and we can apply Strichartz estimates (A.2) with $r = \frac{15}{4}$ and $\theta_1 = 14/15$ for the $U_{N_i}$ and Strichartz estimates with $r = 4 - 3\varepsilon_1$ and $\theta_2 = \frac{4 - 6\varepsilon_1}{4 - 3\varepsilon_1}$ for $V_N$. By Sobolev embedding, accounting for the $N_i^{s_1}$ factors in the expressions for the $U_{N_i}$, we obtain

$$\lesssim (N_1 N_2 N_3)^{-s_1} N^{-(1-s_2)} \prod_{i=1}^3 \|u^{(i)}\|_{H^1(T^3)}^{1/4} \prod_{i=1}^3 \|c_i\|_{L^2}^{3/4}. $$
The expression on the right-hand side is summable for dyadic values of $N_i$ and $N$. By the definition of $\Sigma_\lambda$ (4.19), we obtain

\[
(4.22) \lesssim (\lambda + T)^{\frac{3}{4}} + \prod_{i=1}^{3} \|u^{(i)}\|_{X^{s_1, b, \delta}}^\frac{3}{4}.
\]  

**Case (B): $u^{(1)}$ of type (I) and $u^{(2)}$, $u^{(3)}$ of type (II).** In this case, we use Hölder’s inequality and Strichartz estimates (A.2) with $r = \frac{18}{5}$ and $\theta = \frac{8}{9}$, and provided

\[
1 - b > \frac{4}{9},
\]

which holds if (4.26) holds, we can bound

\[
(4.22) \lesssim (N_1 N_2 N_3)^{-s_1} N^{-(1-s_2)} \int U_{N_1} U_{N_2} U_{N_3} \cdot V_N \, dxdt
\]

\[
\lesssim (N_1 N_2 N_3)^{-s_1} N^{-(1-s_2)} \|U_{N_1}\|_{L^{6}} \|U_{N_2}\|_{L^{\frac{18}{5}}} \|U_{N_3}\|_{L^{\frac{8}{9}}} \|V_N\|_{L^{\frac{8}{9}}}
\]

\[
\lesssim (N_1)^{-s_1} (N_2 N_3)^{-s_1} N^{-(\frac{8}{5}-s_2)} \|U_{N_1}\|_{L^{6}} \|u^{(2)}\|_{X^{s_1, b, \delta}} \|u^{(3)}\|_{X^{s_1, b, \delta}}.
\]

Noting how we defined $U_{N_1}$, we use the fact that $s_1 < \frac{1}{2}$ and that $(u_0, u_1) \in \Sigma_\lambda$, to estimate that piece, exploiting the uniform boundedness of the frequency projections in $L^p_x$ spaces. Once again, we obtain an estimate which summable for dyadic $N$ and $N_i$, yielding

\[
(4.22) \lesssim \lambda \|u^{(2)}\|_{X^{s_1, b, \delta}} \|u^{(3)}\|_{X^{s_1, b, \delta}}.
\]  

**Case (C): $u^{(1)}$, $u^{(2)}$ of type (I) and $u^{(3)}$ of type (II).** In this case, we use Hölder’s inequality and Strichartz estimates (A.2) with $r = 3$ and $\theta = \frac{2}{3}$ and

\[
1 - b > \frac{1}{3},
\]

which holds if (4.26) holds and estimate

\[
(4.22) \lesssim (N_1 N_2 N_3)^{-s_1} N^{-(1-s_2)} \int U_{N_1} U_{N_2} U_{N_3} \cdot V_N \, dxdt
\]

\[
\lesssim (N_1 N_2 N_3)^{-s_1} N^{-(1-s_2)} \|U_{N_1}\|_{L^{6}} \|U_{N_2}\|_{L^{6}} \|U_{N_3}\|_{L^{3}} \|V_N\|_{L^{3}}
\]

\[
\lesssim (N_1 N_2)^{-s_1} (N_3)^{-s_1} N^{-(\frac{8}{3}-s_2)} \|U_{N_1}\|_{L^{6}} \|U_{N_2}\|_{L^{6}} \|u^{(3)}\|_{X^{s_1, b}}
\]
which is again summable for dyadic $N$ and $N_i$, yielding

\[(4.22) \lesssim \lambda^2 \|u^{(3)}\|_{X^{s_1, \delta}}.\]  \hspace{1cm} (4.29)

**Case (D): All $u^{(i)}$ of type (I).** In this case we estimate

\[(4.22) \lesssim (N_1 N_2 N_3)^{-s_1} N^{-(1-s_2)} \int U_{N_1} U_{N_2} U_{N_3} \cdot V_N \, dx \, dt\]

\[\lesssim (N_1 N_2 N_3)^{-s_1} N^{-(1-s_2)} \|U_{N_1}\|_{L^6} \|U_{N_2}\|_{L^6} \|U_{N_3}\|_{L^6} \|V_N\|_{L^2}\]

which is again summable for dyadic $N$ and $N_i$, yielding\[(4.22) \lesssim \lambda^3.\]  \hspace{1cm} (4.30)

Finally, we note that for $-\frac{1}{2} < -(1-b)$, we have

\[\|F(u)\|_{X^{s_2-\frac{1}{2}-(1-b)+\delta}} \lesssim \delta^\varepsilon \|F(u)\|_{X^{s_2-\frac{1}{2}-(1-b)+\varepsilon, \delta}}\]  \hspace{1cm} (4.31)

where the implicit constant depends only on $\varepsilon$. To obtain the small time factor in the estimates, we can perform the previous estimates replacing $b$ with $b+\varepsilon$ on the $V_N$ factor for some small $\varepsilon > 0$. Provided $b = \frac{1}{2} + \varepsilon$ is chosen sufficiently close to $\frac{1}{2}$ and $\varepsilon > 0$ is sufficiently small so as to ensure that (4.26) continues to hold for $b + \varepsilon$, we obtain the desired estimate. We will not repeat this consideration in the following Propositions.

Combining (4.27), (4.28), (4.29), (4.30) and (4.31) yields

\[\|F(u)\|_{X^{s_2-\frac{1}{2}+\delta}} \leq C(\lambda, R) \delta^\varepsilon \left(1 + \|u\|_{X^{s_1, \frac{1}{2}+\delta}}^{9/4}\right).\]  \hspace{1cm} $\Box$

**Remark 4.15.** If we consider $(u_0, u_1) \in \Sigma_{\lambda} \cap B_R$ and instead we look at solutions to the cubic nonlinear Klein-Gordon equation with truncated nonlinearity, $u_N$, then we obtain the same bounds

\[\|F(u_N)\|_{X^{s_2-\frac{1}{2}+\delta}} \leq C(\lambda, R, T) \delta^\varepsilon \left(1 + \|u_N\|_{X^{s_1, \frac{1}{2}+\delta}}^{9/4}\right),\]

and similarly for $P_N u_N$. Indeed, by the choice of $\Sigma_{\lambda}$, the nonlinear component of the solution $w_N$
will satisfy the same bounds as the nonlinear component, \( w \), of the solution to the full equation. Since the linear components of \( u \) and \( u_N \) are the same, certainly we have the same bounds on \( S(t)(u_0, u_1) \) and we can repeat the arguments in the previous proof to obtain (4.15).

### 4.5.2 Boundedness of the flow with truncated nonlinearity

We defined the truncated nonlinearity \( F_N(u_N) = P_N[(P_N u_N)^3] \). A direct consequence of Proposition 4.20, Proposition 4.23 and our choice of \( \Sigma_\lambda \) is the following local boundedness result for the truncated nonlinear Klein-Gordon equation. It is important to note that the bounds we obtain are uniform in the truncation parameter.

**Proposition 4.24** (Local boundedness for the truncated equation). Consider the cubic nonlinear Klein-Gordon equation with truncated nonlinearity (4.3). Then there exists \( s_1 < \frac{1}{2} < s_2 \) with \( s_1, s_2 \) sufficiently close to 1/2 such that for any \( \lambda, R, T > 0 \), for every \( (u_0, u_1) \in \Sigma_\lambda \cap B_R \) and for any interval \( I \subset [0, T] \) with \( |I| = \delta \), the truncated nonlinearity satisfies the bound

\[
\|F_N(u_N)\|_{X^{s_2-1, -\frac{1}{2}+\delta(I \times T^3)}} \leq C(\lambda, R, T) \delta^\epsilon \left( 1 + \|u_N\|_{X^{s_1, \frac{1}{2}+\delta(I \times T^3)}}^{9/4} \right),
\]

where \((u_N, \partial_t u_N)\) is the global solution to (4.3) with initial data \((u_0, u_1)\).

### 4.5.3 Continuity estimates for the flow map

We need the following continuity-type estimate when we compare the full nonlinearity on solutions of the full flow to solutions of the truncated equation.

**Proposition 4.25.** Consider the cubic nonlinear Klein-Gordon equation (4.1) and the cubic nonlinear Klein-Gordon equation with truncated nonlinearity (4.3). Then there exists \( s_1 < \frac{1}{2} < s_2 \) with \( s_1, s_2 \) sufficiently close to 1/2 such that for any \( \lambda, R, T > 0 \), for every \( (u_0, u_1) \in \Sigma_\lambda \cap B_R \) and for any interval \( I \subset [0, T] \) with \( |I| = \delta \), the nonlinearity satisfies the bound

\[
\|F(u) - F(u_N)\|_{X^{s_2-1, -\frac{1}{2}+\delta(I \times T^3)}} \leq c(\lambda, R, T) \delta^\epsilon \|u - u_N\|_{X^{s_1, \frac{1}{2}+\delta(I \times T^3)}} \left( 1 + \|u\|_{X^{s_1, \frac{1}{2}+\delta(I \times T^3)}}^{6/4} + \|u_N\|_{X^{s_1, \frac{1}{2}+\delta(I \times T^3)}}^{6/4} \right),
\]

where \((u, \partial_t u)\) and \((u_N, \partial_t u_N)\) to the full and truncated equations, respectively, with data \((u_0, u_1)\).
Proof. As in the proof of the Proposition 4.23, in light of (4.31) we will take \( b = \frac{1}{2} + \) for \( b \) sufficiently close to \( \frac{1}{2} \) so that the time localization yields the desired \( \delta^c \) factor. We first note that

\[
|(u)^3 - (u_N)^3| \lesssim |u - u_N| (|u|^2 + |u_N|^2),
\]

hence these estimates are similar to those in Proposition 4.23 but we will always estimate \( u_1 = |u - u_N| \) in \( X^{\frac{1}{2}, \frac{1}{2} + \delta} \). More precisely, once again we estimate the expression

\[
\sum_{n=n_1+n_2+n_3} \int_{\tau = \tau_1 + \tau_2 + \tau_3} d\tau \frac{c_1(n_1, \tau_1)}{(n_1)^{\frac{1}{2}}} \left( \frac{\langle n \rangle}{\langle n_1 \rangle} \right)^b \prod_{i=2}^3 \frac{c_i(n_i, \tau_i)}{(n_i)^{\frac{1}{2}}} \frac{v(n_i, \tau)}{(n_i)^{\frac{1}{2}} (\langle \tau \rangle - \langle n_i \rangle)^{1-b}}
\]

where \( \|v\|_{L^2(T_3)} \leq 1 \) and as before the functions \( u \) or \( u_N \) in the expression for \( c \) are either of type (I) or type (II), with \( u_1 \) always of type (II).

We define \( U_{N_i} \) and \( V_N \) as in (4.23). The key difference between this proof and the proof of Proposition 4.23 is that in each case, we will estimate \( u_1 = |u - u_N| \) in \( X^{\frac{1}{2}, \frac{1}{2} + \delta} \) instead of using \( X^{s_1, \frac{1}{2} + \delta} \) which we use for the other functions.

- **Case (A): All \( u^{(i)} \) of type (I).** Once again, we recall that since

\[
\|U_{N_i} U_{N_i}^{1-\alpha} \|_{L^{4+\epsilon_2}} \lesssim \|U_{N_i}\|_{L^p} \|U_{N_i}\|_{L^{(1-\alpha)s_1}}^{1-\alpha}
\]

for \( p = \frac{(4+\epsilon_2)5}{1-\epsilon_2} \), we use Hölder’s inequality to estimate (4.22) by

\[
N_1^{-\frac{1}{2}} (N_2 N_3)^{-s_1} N^{-\frac{1}{2}} \int_{R_1} \int_{T_3^2} \prod_{i=1}^3 U_{N_i} \cdot V_N \ dx \ dt \lesssim N_1^{-\frac{1}{2}} (N_1 N_2 N_3)^{-s_1} N^{-\frac{1}{2}} \|U_{N_1}\|_{L^{4-\epsilon_1}} \prod_{i=2}^3 \|U_{N_i}\|_{L^{p_\alpha}} \|U_{N_i}\|_{L^{(1-\alpha)s_1}}^{1-\alpha} \|V_N\|_{L^{4-\epsilon_1}}.
\]

Taking \( \alpha = 1/4 \), we have \( (1-\alpha)5 < 4 \) and for \( \epsilon_2 \) sufficiently small, \( p\alpha < 6 \). By Sobolev embedding and Strichartz estimates (A.2) with \( r = \frac{15}{4} \) and \( \theta_1 = 14/15 \) for the \( U_{N_2}, U_{N_3} \) and by Strichartz estimates with \( r = 4 - \epsilon_1 \) and \( \theta_2 = \frac{4-2\epsilon_1}{4-\epsilon_1} \) for \( U_{N_1} \) and \( V_N \) we obtain

\[
\lesssim (N_1 N)^{-\left(1-\frac{1}{2}\frac{4-2\epsilon_1}{4-\epsilon_1}\right)} (N_2 N_3)^{\frac{3}{8} - s_1} \|U_{N_1}\|_{L^{4-\epsilon_1}} \prod_{i=2}^3 \|u^{(i)}\|_{H^{1/4}(T^3)} \prod_{i=2}^3 \|c_i\|_{L^2}^{3/4}.
\]

The expression on the right-hand side of the inequality is summable for dyadic values of \( N_i \) and
provided $\varepsilon_1 > \theta$. By the definition of $\Sigma_\lambda$ (4.19), we obtain

$$\left|\left(\lambda + T\right)\frac{1}{2} + \|u^{(i)}\|_{X^{1, \frac{1}{2} + \delta}} \prod_{i=2}^{3} \|u^{(i)}\|_{X^{s_1, \frac{1}{2} + \delta}}\right|^3.$$ 

The other cases follow analogously.

Finally, the last continuity type estimate we will need demonstrates that if one of the functions in the multilinear estimates satisfies $\Pi_N u^{(i)} = 0$, that is if it is only supported on high frequencies, then one gains some additional decay in $N$.

**Proposition 4.26.** Consider the cubic nonlinear Klein-Gordon equation (4.1). Then there exists $s_2 > \frac{1}{2}$ with $s_2$ sufficiently close to $1/2$ and $\theta > 0$ such that for any $\lambda, R, T > 0$, for every $(u_0, u_1) \in \Sigma_\lambda \cap B_R$ and for any interval $I \subset [0, T]$ with $|I| = \delta$, the nonlinearity satisfies the bound

$$\|F(u) - F(P_N u)\|_{X^{s_2, \frac{1}{2} + \delta}(I \times \mathbb{T}^3)} \leq C(\lambda, R, T) \delta^\varepsilon N^{-\theta} \left(1 + \|u\|_{X^{\frac{3}{2}, \frac{1}{2} + \delta}(I \times \mathbb{T}^3)}^{3/4}\right).$$

where $(u, \partial_t u)$ is the global solution to the cubic nonlinear Klein-Gordon equation (4.1) with initial data $(u_0, u_1)$.

**Proof.** Once again, we use the inequality.

$$|F(u) - F(P_N u)| \lesssim |(I - P_N) u||u|^2 + |P_N u|^2).$$

Let $s_1$ be as in Proposition 4.23, and we repeat those arguments but we will always estimate the high-frequency term, $(I - P_N) u$, in $X^{s_1, \frac{1}{2} + \delta}$, even for the linear evolution. For instance, consider the case where all the $u^{(i)}$ are of type (I). We once again estimate the expression

$$\sum_{n_1 + n_2 + n_3} \int_{\tau_1 + \tau_2 + \tau_3} \prod_{i=1}^{3} \frac{c_i(n_i, \tau_i)}{\langle n_i \rangle^{s_1} \langle |\tau_i| - \langle n_i \rangle \rangle^{\frac{1}{2} + \delta}} \langle n \rangle^{\frac{1}{2}} \langle |\tau| - \langle n \rangle \rangle^{\frac{1}{2}} d\tau$$

and with the definitions for $U_N$ and $V_N$ as in (4.23), we obtain

$$\left(4.33\right) \lesssim \left(N_1 N_2 N_3\right)^{-s_1} N^{-\frac{1}{2}} \int U_{N_1} U_{N_2} U_{N_3} \cdot V_N \, dx \, dt$$

$$\lesssim \left(N_1 N_2 N_3\right)^{-s_1} N^{-\frac{1}{2}} \|U_{N_1}\|_{L^3} \|U_{N_2}\|_{L^6} \|U_{N_3}\|_{L^6} \|V_N\|_{L^3}$$

$$\lesssim N_1^{-s_1 + \frac{1}{2}} \left(N_2 N_3\right)^{-s_1} N^{-\frac{1}{2}} \|u^{(1)}\|_{X^{s_1, \frac{1}{2} + \delta}} \|U_{N_2}\|_{L^6} \|U_{N_3}\|_{L^6}.$$
which is summable for dyadic $N$ and $N_i$. By Strichartz estimates, recalling that we set

$$u^{(1)} = (1 - P_N)S(t)(u_0, u_1),$$

we obtain

$$
\|u^{(1)}\|_{X^{s_1, \frac{1}{2} + \delta}} = \|(1 - P_N)S(t)(u_0, u_1)\|_{X^{s_1, \frac{1}{2} + \delta}} \\
\lesssim N^{-\theta}\|S(t)(u_0, u_1)\|_{X^{\frac{1}{2}, \frac{1}{2} + \delta}} \lesssim N^{-\theta}\|(u_0, u_1)\|_{H^{1/2}(\mathbb{T}^3)},
$$

which yields the desired estimate. The other cases follow analogously to the previous propositions, with the modification that when we take $u^{(1)} = (1 - P_N)u$ we obtain

$$
\|u^{(1)}\|_{X^{s_1, \frac{1}{2} + \delta}} \lesssim N^{-\theta}\|u\|_{X^{\frac{1}{2}, \frac{1}{2} + \delta}},
$$

which yields the result. \qed

### 4.6 Probabilistic approximation of the flow of the NLKG

This section is devoted to the proof of the approximation of the flow map for the cubic nonlinear Klein-Gordon equation by the flow of the nonlinear Klein-Gordon equation with truncated nonlinearity. We will use here the probabilistic boundedness estimates from Section 4.5.

**Proposition 4.27.** Let $\Phi$ denote the flow of the cubic nonlinear Klein-Gordon equation (4.1), and $\Phi_N$ the flow of the cubic nonlinear Klein-Gordon equation with truncated nonlinearity (4.3). Fix $R, T, \lambda > 0$. Then for every $(u_0, u_1) \in \Sigma_\lambda \cap B_R$,

$$
\sup_{t \in [0, T]} \|\Phi(t)(u_0, u_1) - \Phi_N(t)(u_0, u_1)\|_{H^{1/2}(\mathbb{T}^3)} \leq C(\lambda, T, R) \varepsilon_1(N)
$$

with $\varepsilon_1(N) \to 0$ as $N \to \infty$.

**Proof.** Fix $R, T, \lambda > 0$ and let $\Sigma_\lambda$ be as defined in (4.19). We need to estimate the difference $\Phi - \Phi_N$ for initial data $(u_0, u_1) \in \Sigma_\lambda \cap B_R$. Fix such a $(u_0, u_1)$ and let $u(t)$ and $u_N(t)$ denote the
corresponding solutions to the full and truncated equations, respectively. By the choice of $\Sigma$, 

$$
\|(1 - \Delta)^{s_1/2} S(t)(u_0, u_1)\|_{L^0([0,T]; L^0(T^3))} < C \lambda
$$

$$
\|(w, \partial_t w)\|_{L^\infty([0,T]; H^1(T^3))} < C (\lambda + T)^{1+},
$$

(4.34)

$$
\|(w_N, \partial_t w_N)\|_{L^\infty([0,T]; H^1(T^3))} < C (\lambda + T)^{1+},
$$

where, as usual, $w(t)$ and $w_N(t)$ are the nonlinear components of the global solutions $u(t)$ and $u_N(t)$, respectively. Note in particular, for any subinterval $I \subset [0, T]$, these bounds hold uniformly. Furthermore, for $|I| = \delta$, Proposition 4.23, and the inhomogeneous estimate for the Klein-Gordon equation yield

$$
\|w\|_{X^{s_2, 1+\delta}} \lesssim \|F(u)\|_{X^{s_2-1, -\frac{1}{2}+\delta}} \leq C(\lambda, R, T) \delta^c \left(1 + \|u\|_{X^{s_1, \frac{1}{2}+\delta}}^{9/4}\right),
$$

hence if $\inf I = t_0$, by Lemma A.4 we obtain

$$
\|u\|_{X^{s_1, \frac{1}{2}+\delta}} \leq \|S(t)(u_0, u_1)\|_{X^{s_1, \frac{1}{2}+\delta}} + \|w\|_{X^{s_2, 1+\delta}} \lesssim \|(u(t_0), \partial_t u(t_0))\|_{H^{s_1}} + \|F(u)\|_{X^{s_2-1, -\frac{1}{2}+\delta}}.
$$

By the uniform bounds (4.34) and the boundedness of the free evolution on $H^s$,

$$
\sup_{t_0 \in [0, T]} \|(u(t_0), \partial_t u(t_0))\|_{H^{s_1}} \leq C(\lambda, R, T),
$$

hence

$$
\|u\|_{X^{s_1, \frac{1}{2}+\delta}} \leq C(\lambda, R, T) + C(\lambda, R, T) \delta^c \left(1 + \|u\|_{X^{s_1, \frac{1}{2}+\delta}}^{9/4}\right).
$$

Thus, taking $\delta \equiv \delta(\lambda, R, T) > 0$ sufficiently small we obtain that

$$
\|\Phi(t)(u_0, u_1)\|_{X^{s_1, \frac{1}{2}+\delta}(I \times T^3)} \leq C(\lambda, T, R).
$$

(4.35)

By Proposition 4.24, this argument yields the same result for $\Phi_N(t)$ with the same choice of $\delta > 0$. Note, too, that we can obtain the same bounds with $s_1$ replaced by $\frac{1}{2}$. Now, define

$$(\phi, f_t) := (u - u_N, \partial_t u - \partial_t u_N)$$

106
then we have

\[ \partial_t \phi - \Delta \phi + \phi = \left( F(u) - F(u_N) + F(u_N) - F(P_Nu_N) + F(P_Nu_N) - P_NF(P_Nu_N) \right). \]

Set

\[ \Delta_1 := F(u) - F(u_N), \quad \Delta_2 := F(u_N) - F(P_Nu_N), \quad \Delta_3 := (I - P_N)F(P_Nu_N), \]

and fix an interval \( I \subset [0, T] \) with \( |I| = \delta \). We will estimate

\[ \| \Delta_1 \|_{\dot{X}^{-\frac{1}{2}, -\frac{1}{2} + \theta}_{(I \times \mathbb{T}^3)}} + \| \Delta_2 \|_{\dot{X}^{-\frac{1}{2}, -\frac{1}{2} + \theta}_{(I \times \mathbb{T}^3)}} + \| \Delta_3 \|_{\dot{X}^{-\frac{1}{2}, -\frac{1}{2} + \theta}_{(I \times \mathbb{T}^3)}}. \]

By Proposition 4.25, we can bound \( \Delta_1 \) by

\[ \| \Delta_1 \|_{\dot{X}^{-\frac{1}{2}, -\frac{1}{2} + \theta}} \leq C(\lambda, R, T) \delta^c \| u - u_N \|_{\dot{X}^{\frac{1}{2}, \frac{1}{2} + \theta}} \left( 1 + \| u \|_{\dot{X}^{\frac{6}{4}, \frac{1}{2} + \theta}}^{6/4} + \| u_N \|_{\dot{X}^{\frac{6}{4}, \frac{1}{2} + \theta}}^{6/4} \right). \]

For the second term, Proposition 4.26 and Remark 4.15 yields

\[ \| \Delta_2 \|_{\dot{X}^{s_2-1, -\frac{1}{2} + \theta}} \leq C(\lambda, R, T) \delta^c N^{-\theta} \left( 1 + \| u_N \|_{\dot{X}^{\frac{9}{4}, \frac{1}{2} + \theta}}^{9/4} \right). \]

Finally, by Remark 4.15 we have

\[ \| \Delta_3 \|_{\dot{X}^{-\frac{1}{2}, -\frac{1}{2} + \theta}} \leq N^{-\theta} \| F(P_Nu_N) \|_{\dot{X}^{s_2-1, -\frac{1}{2} + \theta}} \leq C(\lambda, R, T) \delta^c N^{-\theta} \left( 1 + \| u_N \|_{\dot{X}^{\frac{9}{4}, \frac{1}{2} + \theta}}^{9/4} \right). \]

In the second and third terms, we used the observation that for any \( N \in \mathbb{N}, s \in \mathbb{R} \), we have

\[ \| P_Nv \|_{\dot{X}^{s, \frac{1}{2} + \theta}} \leq \| v \|_{\dot{X}^{s, \frac{1}{2} + \theta}} \]

for any \( v \) such that the right-hand side is finite. Now let \( I = [0, \delta] \). Then because \( u - u_N \) has zero initial data, the inhomogeneous estimate yields

\[ \| \phi \|_{\dot{X}^{\frac{1}{2}, \frac{1}{2} + \theta}} \lesssim \| \Delta_1 \|_{\dot{X}^{-\frac{1}{2}, -\frac{1}{2} + \theta}} + \| \Delta_2 \|_{\dot{X}^{-\frac{1}{2}, -\frac{1}{2} + \theta}} + \| \Delta_3 \|_{\dot{X}^{-\frac{1}{2}, -\frac{1}{2} + \theta}}, \quad (4.36) \]

107
and together with (4.35) and the similar result for $\Phi_N$, we can bound (4.36) by

$$\|\phi\|_{X^{3,\frac{1}{2}+\delta}} \leq C(\lambda, T, R) \delta^c \|\phi\|_{X^{3,\frac{1}{2}+\delta}} + C(\lambda, T, R) \delta^c N^{-\theta}, \tag{4.37}$$

and similarly for the time derivative component. Hence, for $\delta > 0$ sufficiently small,

$$\|(\phi, \partial_t \phi)\|_{L_t^{\infty} H_x^{3/2}(I \times \mathbb{T}^3)} \lesssim C(\lambda, R, T) \varepsilon_1(N), \tag{4.38}$$

with $\lim_{N \to \infty} \varepsilon_1(N) = 0$. On the next subinterval, $I_2 = [\delta, 2\delta]$ we bound

$$\|(\phi, \partial_t \phi)\|_{L_t^{\infty} H_x^{3/2}(I \times \mathbb{T}^3)} \lesssim C(\lambda, R, T) \varepsilon_1(N) + \left\| \int_\delta^t S(t-s) [F(u) - F_N(u_N)] \right\|_{L_t^{\infty} H_x^{3/2}}$$

and once again by the inhomogeneous estimate

$$\|\phi\|_{X^{3,\frac{1}{2}+\delta}(I_2 \times \mathbb{T}^3)} \lesssim C(\lambda, R, T) \varepsilon_1(N) + \| [F(u) - F_N(u_N)] \|_{X^{3,\frac{1}{2}+\delta}(I_2 \times \mathbb{T}^3)}.$$  

Applying the above argument, we obtain that

$$\|(\phi, \partial_t \phi)\|_{L_t^{\infty} H_x^{3/2}(I \times \mathbb{T}^3)} \lesssim C_2(\lambda, R, T) \varepsilon_1(N). \tag{4.39}$$

At each stage the coefficient of the nonlinear component is independent of the step number, the constants in (4.38) are independent of the subinterval and the bounds (4.35) are uniform, we can choose $\delta > 0$ sufficiently small to obtain the analogue of (4.39) at each stage uniformly for all subintervals. We do remark, however, that the bound that we obtain will grow with each iteration because the constant for the initial data is compounded. Since the number of steps is controlled by $\lambda, R$ and $T$, we obtain the desired result. \(\square\)

**Remark 4.16.** When the estimate is performed on the time derivative, the time localization may increase the left-hand side of (4.37) up to a factor of $\delta^{\frac{1}{2} - b}$ for $b = \frac{1}{2} + c$ as above. However, we recall that in Section 4.5, we the exponent $c$ which we obtain on $\delta$ is some fixed, small constant independent of $b$. By taking $b$ sufficiently close to $\frac{1}{2}$, we still obtain the necessary $\delta$ factor provided we ensure that

$$\frac{1}{2} - b + c > 0.$$  

**Remark 4.17.** The term $\Delta_2$ is a key reason why this argument will not work with probabilistic
energy estimates alone, as in [16], say. Indeed, this term requires us to bound the nonlinearity by a weaker norm and it does not seem possible to close the Gronwall argument if one needs to derive a bound with respect to some norm below $H^1$.

### 4.7 Approximation of the flow on open sets

The goal of this section is to prove the approximation results presented in the introduction.

#### 4.7.1 Proof of Theorem 4.5

One key component in our argument is the critical stability theory which allows us to upgrade the sets of large measure where the approximation holds to open sets, or at least to general initial data in $\Pi_{2N}B_R$ as is required for Theorem 4.5. Stability arguments first appeared in the context of the three-dimensional energy critical nonlinear Schrödinger equation in [18], see also [67].

For the nonlinear Klein-Gordon equation in periodic settings, some care is required as the Strichartz estimates need to be localized in time. Nonetheless, they follow in a similar manner from the Strichartz estimates of Proposition A.2, and we present the proofs in Appendix 4.8. One modification we will present are the stability arguments adapted to the nonlinear Klein-Gordon equation with truncated nonlinearity. Importantly, we will be able to choose the small parameters in these arguments uniformly in the truncation parameter. We recall the statement of this theorem.

**Theorem 4.5.** Let $\Phi$ denote the flow of the cubic nonlinear Klein-Gordon equation (4.1) and $\Phi_N$ the flow of (4.3). Fix $R > 0$, $u_* \in H^{1/2}$ and $N', N \in \mathbb{N}$ with $N'$ sufficiently large, depending on $u_*$. Then there exists a constant $\varepsilon_0(u_*, R) > 0$ such that for all $0 < \varepsilon \leq \varepsilon_0$, there exists $\sigma \equiv \sigma(R, \varepsilon, N')$ such that for any $(u_0, u_1) \in B_R(u_*)$,

$$\sup_{t \in [0, \sigma]} \|\Phi(t)\Pi_{N'}(u_0, u_1) - \Phi_N(t)\Pi_{N'}(u_0, u_1)\|_{H^{1/2}_{\varepsilon}} \leq C(R, u_*) [\varepsilon_1(N) + \varepsilon]$$

(4.40)

with $\lim_{N \to \infty} \varepsilon_1(N) = 0$.

**Proof of Theorem 4.5.** Fix $R > 0$, $I = [0, \sigma]$ for some $0 < \sigma \leq 1$ to be fixed, and let $\lambda > 0$ be sufficiently large so that we can find $(v_0, v_1) \in \Sigma_\lambda \cap B_R(u_*)$. The intersection is non-empty for some $\lambda > 0$ by density and for all $N' \in \mathbb{N}$ it holds that $P_{N'}(v_0, v_1) \in \Sigma_\lambda$ since (4.19) and (4.53) are invariant under smooth projections. Thus, there exists some constant $K_1 > 0$ such that the
corresponding global solutions \( v \) and \( v_N \) to equations (4.1) and (4.3) with initial data \( P_{N^r}(v_0, v_1) \) satisfy

\[
\|v\|_{L^4(I \times \mathbb{T}^3)} \leq K_1 \quad \text{and} \quad \|v_N\|_{L^4(I \times \mathbb{T}^3)} \leq K_1.
\]

Let \((u_0, u_1) \in B_R\), and let \(N'\) be sufficiently large so that

\[
\|\Pi_{\geq N'} u_u\|_{\mathcal{H}^{1/2}} < R.
\]

Then

\[
\|S(t)(P_{N^r}(v_0, v_1) - \Pi_{N^r}(u_0, u_1))\|_{L^4_{t,x}(I \times \mathbb{T}^3)} \\
\leq |I|^{1/4} \sup_{t \in I} \|S(t)(P_{N^r}(v_0, v_1) - \Pi_{N^r}(u_0, u_1))\|_{L^4_{x}(\mathbb{T}^3)} \\
\leq |I|^{1/4} (N')^{1/2} \left[ \|(P_{N^r}(v_0, v_1) - P_{N^r}u_u)\|_{H^{1/2}_{x}(\mathbb{T}^3)} + \|(P_{N^r}u_u - \Pi_{N^r}(u_0, u_1))\|_{H^{1/2}_{x}(\mathbb{T}^3)} \right] \\
\lesssim |I|^{1/4} (N')^{1/2} R.
\]

Let \( \rho_1 = \rho_1(K_1) \) be as in the stability lemma and recall that we can choose \( \rho_1 \) uniformly for all \( \sigma \leq 1 \). Let \( 0 < \varepsilon_0 < \rho_1 \), then setting

\[
\sigma \simeq (N')^{-2} R^{-4} \varepsilon_0,
\]

the smallness condition (4.45) of Lemma 4.33 is met and we conclude that for \( t \in I \), solutions

\[
u(t) := \Phi(t) \Pi_{N^r}(u_0, u_1) \\
u_N(t) := \Phi_N(t) \Pi_{N^r}(v_0, v_1)
\]

exist to equations (4.1) and (4.3), respectively. Moreover, we conclude from (4.46) that

\[
\|F(u) - F(v)\|_{L^{4/3}(I \times \mathbb{T}^3)} \leq C(K_1) \varepsilon_0.
\]

Hence, by Duhamel’s formula and Strichartz estimates, the nonlinear components \( \tilde{\Phi} \) of the solutions, which was defined in (4.10), satisfy

\[
\sup_{t \in I} \|\tilde{\Phi}(t) \Pi_{N^r}(v_0, v_1) - \tilde{\Phi}(t) \Pi_{N^r}(u_0, u_1)\|_{H^{1/2}_{x}(\mathbb{T}^3)} \lesssim \|F(u) - F(v)\|_{L^{4/3}_{t,x}(I \times \mathbb{T}^3)} \leq C(K_1) \varepsilon_0,
\]

110
and similarly for $\Phi_N$. We can estimate (4.40) using the triangle inequality by

$$\sup_{t \in I_0} \| \Phi(t) \Pi_{N'}(u_0, u_1) - \Phi_N(t) \Pi_{N'}(u_0, u_1) \|_{H^1_2}(T^3)$$

$$= \sup_{t \in I_0} \| \Phi(t) \Pi_{N'}(v_0, v_1) - \Phi_N(t) \Pi_{N'}(v_0, v_1) \|_{H^1_2}(T^3)$$

$$\leq \sup_{t \in I_0} \| \Phi(t) P_{N'}(v_0, v_1) - \Phi_N(t) P_{N'}(v_0, v_1) \|_{H^1_2}(T^3)$$

$$+ \sup_{t \in I_0} \| \Phi(t) P_{N'}(v_0, v_1) - \Phi_N(t) \Pi_{N'}(u_0, u_1) \|_{H^1_2}(T^3)$$

and hence we obtain that for all $(u_0, u_1) \in B_R(u_*)$,

$$\sup_{t \in [0, T]} \| \Pi_{N'}(u_0, u_1) - \Phi_N(t) \Pi_{N'}(u_0, u_1) \|_{H^1_2}(T^3) < C(R, u_*) [\varepsilon_1(N) + \varepsilon_0].$$

### 4.7.2 Proof of Theorem 4.8

We turn to the proof Theorem 4.8. We define the $\rho$-fattening of $\Sigma_\lambda$ by

$$\Sigma_{\lambda, \rho} := \bigcup_{u \in \Sigma_\lambda} B_{\rho}(u).$$

In this section, we will consider initial data $(u_0, u_1) \in \Sigma_{\lambda, \rho} \cap B_R$ for some sufficiently small $\rho \equiv \rho(\lambda, T) > 0$ for which we have uniform bounds for the corresponding solutions by the stability theory.

The following theorem demonstrates that for $\rho \equiv \rho(\lambda, T) > 0$ sufficiently small, we can obtain convergence uniformly on this open set. Our proof uses a combination of our probabilistic approximation result, Proposition 4.27, and the low-frequency stability result of Theorem 3.1. Using this theorem, the proof of Theorem 4.8 follows similarly to that of Theorem 4.1.

**Theorem 4.28.** Let $\Phi$ denote the flow of the cubic nonlinear Klein-Gordon equation (4.1) and $\Phi_N$ the flow of the cubic nonlinear Klein-Gordon equation with truncated nonlinearity (4.3). Fix $\lambda, T, R > 0$, then there exists some $\rho_1 = \rho_1(\lambda, T) > 0$ sufficiently small such that for any $0 < \rho < \rho_1$, any $(u_0, u_1) \in \Sigma_{\lambda, \rho} \cap B_R$, and any $\varepsilon > 0$ and $N' \in \mathbb{N}$,

$$\sup_{t \in [0, T]} \| P_{N'}(\Phi(t)(u_0, u_1) - \Phi_N(t)(u_0, u_1)) \|_{H^1_2} < \varepsilon$$
for $N = N(N', \varepsilon, \lambda, R, T) \gg N'$ sufficiently large.

Proof. Suppose that $(u_0, u_1) \in B_p(v_0, v_1)$ for some fixed $(v_0, v_1) \in \Sigma_{\lambda}$. Since $P_N(u_0, u_1) \in B_p P_N(v_0, v_1)$ for all $N_\ast \in \mathbb{N}$, the stability theory yields solutions on $[0, T]$ given by

$$u := \Phi(t)(u_0, u_1), \quad u_N := \Phi_N(t)(u_0, u_1)$$

$$\tilde{u} := \Phi(t) P_N(u_0, u_1), \quad \tilde{u}_N := \Phi_N(t) P_N(u_0, u_1)$$

which satisfy uniform $L^4_t L^4_x([0,T] \times \mathbb{T}^3)$ and $L^\infty_t H^1_x([0,T] \times \mathbb{T}^3)$ bounds depending only on $\lambda, R, T$. We apply the triangle inequality

$$\sup_{t \in [0,T]} \| P_N' (\Phi(t)(u_0, u_1) - \Phi_N(t)(u_0, u_1)) \|_{H^{1/2}_x} \leq \sup_{t \in [0,T]} \| P_N' (\Phi(t)(u_0, u_1) - \Phi(t) P_N(u_0, u_1)) \|_{H^{1/2}_x}$$

$$\quad \sup_{t \in [0,T]} \| P_N' (\Phi(t) P_N(u_0, u_1) - \Phi(t) P_N(u_0, u_1)) \|_{H^{1/2}_x}$$

and we estimate each term separately. For the second term, we observe that $P_N(u_0, u_1)$ is smooth, hence $P_N(u_0, u_1) \in \Sigma_{\lambda} \cap B_R$ for $\lambda = \lambda(R, K, N_\ast)$ by energy conservation and Sobolev embedding. Thus Proposition 4.27 yields the bound

$$\sup_{t \in [0,T]} \| P_N' (\Phi(t) P_N(u_0, u_1) - \Phi(t) P_N(u_0, u_1)) \|_{H^{1/2}_x} \lesssim C(N_\ast, R, T) \varepsilon_1(N).$$

By Theorem 3.1 and Remark 3.6, we can bound the first and last terms by

$$\sup_{t \in [0,T]} \| P_N' (\Phi(t)(u_0, u_1) - \Phi(t) P_N(u_0, u_1)) \|_{H^{1/2}_x} \lesssim \left( \frac{\log N_\varepsilon}{N'} \right)^{-\theta}$$

$$\sup_{t \in [0,T]} \| P_N' (\Phi_N(t)(u_0, u_1) - \Phi_N(t) P_N(u_0, u_1)) \|_{H^{1/2}_x} \lesssim \left( \frac{\log N_\varepsilon}{N'} \right)^{-\theta},$$

where the implicit constants depend on $\lambda, R$, and $T > 0$. Thus for fixed $N' \in \mathbb{N}$, choosing $N_\ast$ sufficiently large, and subsequently $N$ sufficiently large yields the result. \hfill $\Box$

Now we prove the following statement, from which we obtain Theorem 4.8 readily given the
bounds on the measure of the subsets $\Sigma_\lambda$ from Proposition 4.22.

**Theorem 4.29.** Let $\Phi$ denote the flow of the cubic nonlinear Klein-Gordon equation (4.1). Fix $T, R > 0$, $k_0 \in \mathbb{Z}^3$, $z \in \mathbb{C}$, and $u_* \in \mathcal{H}^{1/2}(\mathbb{T}^3)$ and let $\lambda > 0$ be such that $\Sigma_\lambda \cap B_R(u_*) \neq \emptyset$. Then for all $0 < \rho < \rho_1(\lambda, T)$ sufficiently small,

$$\Phi(T)(\Sigma_{\lambda, \rho} \cap B_R(u_*)) \not\subseteq C_r(z; k_0)$$

for all $r > 0$ with $\pi r^2 < \text{cap}(\Sigma_{\lambda, \rho} \cap B_R(u_*))$.

**Proof.** Fix $N' \in \mathbb{N}$ with $N' > |k_0|$ and let $R_1 := \|u_*\| + R$. Then by Theorem 4.28, we can find $N \in \mathbb{N}$ sufficiently large so that for any $(u_0, u_1) \in \Sigma_{\lambda, \rho} \cap B_{R_1}$ we have

$$\sup_{t \in [0, T]} \|P_{N'}(\Phi(t)(u_0, u_1) - \Phi(t)(u_0, u_1))\|_{\mathcal{H}^{1/2}} < \varepsilon. \quad (4.41)$$

Since $\Phi_N$ preserves capacities by Proposition 4.21, we have the equality

$$c(\Phi_N(t)(\Sigma_{\lambda, \rho} \cap B_R(u_*))) = c(\Sigma_{\lambda, \rho} \cap B_R(u_*))$$

for all $t \in \mathbb{R}$. Thus for $z = (z_0, z_0)$ we can find some $(u_0, u_1) \in \Sigma_{\lambda, \rho} \cap B_R(u_*)$ such that

$$\left(|k_0|\Phi_N(T)(u_0, u_1) - z_0|^2 + (k_0)^{-1}\left|\partial_t \Phi_N(T)(u_0, u_1) - z_0\right|^2\right)^{1/2} > r + \varepsilon,$$

and since $(u_0, u_1) \in \Sigma_{\lambda, \rho} \cap B_{R_1}$, we conclude by the triangle inequality and (4.41) that

$$\left(|k_0|\Phi(T)(u_0, u_1) - z_0|^2 + (k_0)^{-1}\left|\partial_t \Phi(T)(u_0, u_1) - z_0\right|^2\right)^{1/2} > r,$$

which completes the proof. \hfill \Box

### 4.7.3 Proof of Theorem 4.7

The goal of this subsection is to prove the conditional global result and the small-data non-squeezing result. The following result demonstrates that at low frequencies, the truncated flow is a good approximated to the full equation. The proof follows from the same arguments used to prove Theorem 3.1, which is unsurprising given that Theorem 3.1 essentially yields a decoupling between low and high frequencies. In this setting, we do not rely on the probabilistic estimates.
from Proposition 4.27, however, we are only able to compare the low frequency components of the corresponding solutions. The following proposition immediately yields the large data portion of Theorem 4.7.

**Proposition 4.30.** Let $\Phi$ and $\Phi_N$ denote the flows of the cubic nonlinear Klein-Gordon equation with full (4.1) and truncated (4.3) nonlinearities, respectively. Fix $T, R > 0$ and suppose there exists some $K > 0$ such that for all $(u_0, u_1) \in B_R(u_\star)$, the corresponding solutions $u$ to (4.1) and $u_N$ to (4.3) satisfy

$$
\|u\|_{L^4_t(L^6_x([0,T] \times \mathbb{T}^3))} + \sup_N \|P_N u_N\|_{L^4_t(L^6_x([0,T] \times \mathbb{T}^3))} \leq K.
$$

Then for all $\varepsilon > 0$ and any $N' \in \mathbb{N}$, and sufficiently large $N$ depending on $R, T$ and $K$,

$$
\sup_{t \in [0,T]} \|P_N^\varepsilon (\Phi(t)(u_0, u_1) - \Phi_N(t)(u_0, u_1))\|_{H^{1/2}_x} \lesssim \left(\log \frac{N}{N'}\right)^{-\theta},
$$

with implicit constants depending on $R, T$, and $K$.

**Proof.** Let $u$ and $u_N$ to the Cauchy problem (4.1) and the truncated equation (4.3) respectively on $[0,T)$. By Proposition 3.18, there exists some $M \in [N', N]$

$$
\Box u_0 + u_0 = P_M F(u_{0o}, u_{0o}, u_{0o}) + O_{K,R,T}((\log(N/N'))^{-\theta})
$$

for $u_{0o} = P_M u$. By the same reasoning,

$$
\Box u_{N,0} + u_{N,0} = P_M F(u_{N,0}, u_{N,0}, u_{N,0}) + O_{K,R,T}((\log(N/N'))^{-\theta})
$$

with a slightly different error term. Since $u$ and $u_N$ have the same initial data, the arguments used to prove Theorem 3.1 yield the result. 

**Remark 4.18.** Note that although any initial data in $P_{N'} B_R$ gives rise to global solutions of the relevant Cauchy problems, Proposition 4.30 is insufficient to prove Theorem 4.5 since the implicit constants would depend on the truncation parameter and thus we could not guarantee convergence uniformly.

To conclude, we present the following lemma which yields our small data result.

**Lemma 4.31.** Fix $T > 0$ and let $\Phi_N$ denote the flow of the cubic nonlinear Klein-Gordon equation with truncated nonlinearity (4.3). There exists some sufficiently small absolute constant $\rho_0 = \rho_0(T)$
such that for any $0 < \rho < \rho_0$ and for all $(u_0, u_1) \in B_\rho \subset \mathcal{H}^{1/2}(\mathbb{T}^3)$ there exists a unique solution $u_N := \Phi_N(u_0, u_1)$ on $[0, T]$ which satisfies

$$\|u_N\|_{L^4_{t,x}([0,T] \times \mathbb{T}^3)} \lesssim \rho.$$  

Proof. Fix $T > 0$ By the small data theory, we know that there exists some $\rho_1(T) > 0$ sufficiently small so that for any $0 < \rho < \rho_1$, and for all $(u_0, u_1) \in B_\rho$, a unique solution $u := \Phi(u_0, u_1)$ exists to the cubic nonlinear Klein-Gordon equation (4.1), which satisfies

$$\|u\|_{L^4_{t,x}([0,T] \times \mathbb{T}^3)} \lesssim \rho.$$  

Let $F(u) = u^3$, then we can expand

$$F(u) - P_N F(P_N u_N) = F(u) - P_N F(u) + P_N F(u) - P_N F(P_N u) + P_N F(P_N u) - P_N F(P_N u_N),$$  

hence by the boundedness of the smooth projections and Strichartz estimates we obtain

$$\|u - u_N\|_{L^4_{t,x}} \lesssim \|u\|_{L^4_{t,x}}^3 + \|P_N F(u) - F(P_N u_N)\|_{L^{4/3}_{t,x}}$$

$$\lesssim \|u\|_{L^4_{t,x}}^3 + \|u - u_N\|_{L^4_{t,x}}^3 + \|u\|_{L^4_{t,x}}^2 \|u - u_N\|_{L^4_{t,x}},$$

where the implicit constants may depend on time. By taking $\rho_0 = \rho_0(T) > 0$ smaller if necessary, we obtain the desired result by a standard continuity argument. \qed

4.8 Appendix A: Stability Arguments

This appendix is devoted the proofs of the critical stability lemmata for the cubic nonlinear Klein-Gordon equation. As usual, these statements are proved first for solutions which have sufficiently small Strichartz bounds, we will call these short-time stability arguments. In order to conclude the statement for arbitrarily large bounds, one needs to divide a given time interval into subintervals such that the norm of the solution is sufficiently small on each subinterval, then the statement follows from an iteration argument. We include these proofs as we would like to make explicit the dependence of the constants on the various parameters involved. We also point out that the
dependence on the time interval in the following estimates arises solely due to localizing Strichartz estimates and thus can be taken uniformly for all $I \subset [0,1]$. In the sequel we let $F(u) := u^3$ and $F_N(u_N) := P_N(P_Nu)^3$.

### 4.8.1 Stability theory for NLKG

**Lemma 4.32** (Short-time stability). Let $I \subset \mathbb{R}$ a compact time interval and $t_0 \in I$. Let $v$ be a solution defined on $I \times \mathbb{T}^3$ of the Cauchy problem

$$
\begin{align*}
\left\{ \begin{array}{l}
v_{tt} - \Delta v + v + F(v) = e \\
(v, \partial_t v) \big|_{t=t_0} = (v_0, v_1) \in \mathcal{H}^{1/2}(\mathbb{T}^3).
\end{array} \right.
\end{align*}
$$

Let $(u, \partial_t u) \big|_{t=t_0} = (u_0, u_1) \in \mathcal{H}^{1/2}(\mathbb{T}^3)$ be such that

$$
\|(v_0 - u_0, v_1 - u_1)\|_{\mathcal{H}^{1/2}(\mathbb{T}^3)} \leq R_1
$$

for some $R_1 > 0$. Suppose also that we have the smallness conditions

$$
\|v\|_{L^4(I \times \mathbb{T}^3)} \leq \rho_0
$$

$$
\|S(t-t_0)(v_0 - u_0, v_1 - u_1)\|_{L^4(I \times \mathbb{T}^3)} \leq \rho
$$

$$
\|e\|_{L^{q'}(L^{r'}(I \times \mathbb{T}^3))} \leq \rho,
$$

for some $0 < \rho < \rho_0(R_1)$ a small constant and $(q', r')$ a conjugate admissible pair. Then there exists a unique solution $(u(t), \partial_t u(t))$ to the cubic nonlinear Klein-Gordon equation on $I \times \mathbb{T}^3$ with initial data $(u_0, u_1)$ at time $t_0$ and $C = C(I) \geq 1$ which satisfies

$$
\|v - u\|_{L^4(I \times \mathbb{T}^3)} \leq C \rho
$$

$$
\|F(v) - F(u)\|_{L^{4/3}(I \times \mathbb{T}^3)} \leq C \rho
$$

$$
\|(v - u, \partial_t u - \partial_t v)\|_{L^{q\infty}_t H^{1/2}_x(I \times \mathbb{T}^3)} + \|v - u\|_{L^{q'}_t L^{r'}(I \times \mathbb{T}^3)} \leq CR_1
$$

for all admissible pairs $(q, r)$. Furthermore, the dependence of the constant $C$ on time arises only from the constant in the localized Strichartz estimates.

**Remark 4.19.** By Strichartz estimates, assumption (4.42) is redundant if $R_1 = O(\rho)$.

116
Proof. Without loss of generality, let \( t_0 = \inf I \) and let \( \phi := v - u \), then \( \phi \) satisfies the Cauchy problem

\[
\begin{aligned}
\phi_{tt} - \Delta \phi + \phi + F(v) - F(\phi + v) &= e \\
(\phi, \partial_t \phi)|_{t=t_0} &= (v_0 - u_0, v_1 - u_1).
\end{aligned}
\]

By Strichartz estimates, Hölder’s inequality and the assumptions above,

\[
\|\phi\|_{L^4(I \times \mathbb{T}^3)} \leq C \left( \|S(t)(v_0 - u_0, v_1 - u_1)\|_{L^4(I \times \mathbb{T}^3)} + \|F(\phi + v) - F(v)\|_{L^{4/3}(I \times \mathbb{T}^3)} + \|e\|_{L^4(I \times \mathbb{T}^3)} \right)
\]

\[
\leq C \left( 2 \rho + (\rho_0)^2 \|\phi\|_{L^4(I \times \mathbb{T}^3)} + \|\phi\|_{L^4(I \times \mathbb{T}^3)}^3 \right)
\]

hence a continuity argument yields (4.43) provided \( \rho_0 \) is sufficiently small. In particular

\[
\|F(\phi + v) - F(v)\|_{L^{4/3}(I \times \mathbb{T}^3)} \leq C \rho
\]

for such \( \rho_0 \). From (4.43) we have

\[
\|(\phi, \partial_t \phi)\|_{L^\infty_{t,x} H^{1/2}_{x}(I \times \mathbb{T}^3)} + \|\phi\|_{L^4_{t,x}(I \times \mathbb{T}^3)} + \|\phi\|_{L^4(I \times \mathbb{T}^3)}
\]

\[
\leq C \left( R_1 + \|F(\phi + v) - F(v)\|_{L^{4/3}(I \times \mathbb{T}^3)} + \|e\|_{L^4(I \times \mathbb{T}^3)} \right)
\]

\[
\leq C \left( R_1 + \rho + C \rho \right),
\]

hence we obtain (4.44) provided \( \rho_0 \equiv \rho_0(R_1) \) is chosen sufficiently small. \( \square \)

**Lemma 4.33** (Long-time stability). Let \( I \subset \mathbb{R} \) a compact time interval and \( t_0 \in I \). Let \( v \) be a solution defined on \( I \times \mathbb{T}^3 \) of the Cauchy problem

\[
\begin{aligned}
v_{tt} - \Delta v + v + F(v) &= e \\
(v, \partial_t v)|_{t=t_0} &= (v_0, v_1) \in H^{1/2}(\mathbb{T}^3).
\end{aligned}
\]

Suppose that

\[\|v\|_{L^4(I \times \mathbb{T}^3)} \leq L\]

for some constant \( L > 0 \). Let \( t_0 \in I \) and let \( (u, \partial_t u)|_{t=t_0} = (u_0, u_1) \in H^{1/2}(\mathbb{T}^3) \) be such that

\[\|(v_0 - u_0, v_1 - u_1)\|_{H^{1/2}(\mathbb{T}^3)} \leq R_1\]

117
for some $R_1 > 0$. Suppose also that we have the smallness conditions
\[
\|S(t - t_0)(v_0 - u_0, v_1 - u_1)\|_{L^4(I \times \mathbb{T}^3)} \leq \rho
\]
\[
\|e\|_{L^q_t L^r_x(I \times \mathbb{T}^3)} \leq \rho,
\]
for some $0 < \rho < \rho_1(R_1, I, L)$ a small constant and $(q', r')$ a conjugate admissible pair. Then there exists a unique solution $(u, \partial_t u)$ to the cubic nonlinear Klein-Gordon equation on $I \times \mathbb{T}^3$ with initial data $(u_0, u_1)$ at time $t_0$ and $C \equiv C(I, L) \geq 1$ which satisfies
\[
\|v - u\|_{L^4(I \times \mathbb{T}^3)} \leq C \rho
\]
\[
\|F(v) - F(u)\|_{L^{4/3}(I \times \mathbb{T}^3)} \leq C \rho
\]
\[
\|(v - u, \partial_t u - \partial_t v)\|_{L^\infty_t H_x^{1/2}(I \times \mathbb{T}^3)} + \|v - u\|_{L^{q'_r}_t L^{r'}_x(I \times \mathbb{T}^3)} \leq CR_1
\]
admissible pairs $(q, r)$. Moreover, the dependence of the constants $C$ and $\rho_1$ on time arises solely from the constant in the time localized Strichartz inequality.

**Proof.** Fix $I \subset \mathbb{R}$ and let $\rho_0 \equiv \rho_0(2R_1) > 0$ be as in Lemma 4.32. This will allow for some growth in the argument. We divide the time interval $I$ into $J \sim \left(1 + \frac{L}{\rho_0}\right)^4$ subintervals $I_j = [t_j, t_{j+1}]$ such that $\|v\|_{L^4(I_j)} \leq \rho_0$ letting $\phi := u - v$, we can apply the previous lemma on the first interval, yielding
\[
\|v - u\|_{L^4(I_0 \times \mathbb{T}^3)} \leq C \rho
\]
\[
\|F(v) - F(u)\|_{L^{4/3}(I_0 \times \mathbb{T}^3)} \leq C \rho
\]
\[
\|(v - u, \partial_t v - \partial_t u)\|_{L^\infty_t H_x^{1/2}(I_0 \times \mathbb{T}^3)} + \|v - u\|_{L^{q'_r}_t L^{r'}_x(I_0 \times \mathbb{T}^3)} \leq CR_1.
\]

We would like to apply this argument iteratively to claim that
\[
\|v - u\|_{L^4(I_j \times \mathbb{T}^3)} \leq C(j) \rho
\]
\[
\|F(v) - F(u)\|_{L^{4/3}(I_j \times \mathbb{T}^3)} \leq C(j) \rho
\]
\[
\|(v - u, \partial_t v - \partial_t u)\|_{L^\infty_t H_x^{1/2}(I_j \times \mathbb{T}^3)} + \|v - u\|_{L^{q'_r}_t L^{r'}_x(I_j \times \mathbb{T}^3)} \leq C(j) R_1.
\]

In order to do this, we need to ensure that for each $t_j$ we have
\[
\|(v(t_j) - u(t_j), \partial_t v(t_j) - \partial_t u(t_j))\|_{H_x^{1/2}(\mathbb{T}^3)} \leq 2R_1
\]
\[
\|S(t - t_j)(u(t_j) - v(t_j))\|_{L^4(I_j \times \mathbb{T}^3)} < C(I, L) \rho
\]
We prove these statements by induction. For (4.47) we use Strichartz estimates and we bound

\[
\|(\phi(t_j), \partial_t \phi(t_j))\|_{H^{1/2}_x(T^3)} \\
\leq \|(\phi(t_0), \partial_t \phi(t_0))\|_{H^{1/2}_x(T^3)} + \|F(\phi + v) - F(v)\|_{L^{4/3}_t L^{6}_x([0,t_j-1] \times T^3)} + \|e\|_{L^{q'}_t L^{p'}_x([0,t_j-1] \times T^3)} \\
\leq R_1 + \sum_{k=0}^{j-1} C(k) \rho + \rho
\]

and similarly for (4.48), we have

\[
\|S(t - t_j)(u(t_j) - v(t_j))\|_{L^4(I_j \times T^3)} \\
\lesssim \|S(t - t_j)(u(t_0) - v(t_0))\|_{L^4(I_0,t_j) \times T^3)} + \|F(u) - F(v)\|_{L^{q'}_t L^{p'}_x([0,t_j] \times T^3)} + \|e\|_{L^{q'}_t L^{p'}_x([0,t_j] \times T^3)} \\
\lesssim 2\rho + \sum_{k=0}^{j-1} C(k) \rho,
\]

so the conclusion follows by choosing \(\rho_1(R_1, I, L)\) sufficiently small.

4.8.2 Stability theory for the truncated NLKG

Lemma 4.34 (Short-time stability for the truncated equation). Let \(I \subset \mathbb{R}\) a compact time interval and \(t_0 \in I\). Let \(v_N\) be a solution defined on \(I \times T^3\) of the Cauchy problem

\[
\begin{align*}
(v_N)_{tt} - \Delta v_N + v_N + P_N(P_N v_N)^3 &= e \\
(v_N, \partial_t v_N)|_{t=t_0} &= (v_0, v_1) \in H^{1/2}(T^3).
\end{align*}
\]

Let \(t_0 \in I\) and let \((u_N, \partial_t u_N)|_{t=t_0} = (u_0, u_1) \in H^{1/2}(T^3)\) be such that

\[
\|(v_0 - u_0, v_1 - u_1)\|_{H^{1/2}(T^3)} \leq R_1
\]

for some \(R_1 > 0\). Suppose also that we have the smallness conditions

\[
\|P_N v_N\|_{L^4(I \times T^3)} \leq \rho_0 \\
\|S(t - t_0)(v_0 - u_0, v_1 - u_1)\|_{L^4(I \times T^3)} \leq \rho
\]

\[
\|e\|_{L^{q'}_t L^{p'}_x(I \times T^3)} \leq \rho,
\]

119
for some $0 < \rho < \rho_0(R_1)$ a small constant and $(\tilde{q}', \tilde{r}')$ a conjugate admissible pair. Then there exists a unique solution $(u_N, \partial_t u_N)$ to the truncated cubic nonlinear Klein-Gordon equation on $I \times \mathbb{T}^3$ with initial data $(u_0, u_1)$ at time $t_0$ and $C \equiv C(I) \geq 1$ which satisfies

\begin{align}
\|v_N - u_N\|_{L^4(I \times \mathbb{T}^3)} & \leq C \rho \\
\|F_N(v_N) - F_N(u_N)\|_{L^{4/3}(I \times \mathbb{T}^3)} & \leq C \rho \\
\|(v_N - u_N, \partial_t u_N - \partial_t v_N)\|_{L_x^6 H^{1/2}_x(I \times \mathbb{T}^3)} + \|v_N - u_N\|_{L_t^q x^r(I \times \mathbb{T}^3)} & \leq C R_1
\end{align}

(4.49) (4.50)

for any admissible pair $(q, r)$. Furthermore, the dependence of $C$ on time arises only from the constant in the localized Strichartz estimates.

Proof. Without loss of generality, let $t_0 = \inf I$. Let $\phi_N = v_N - u_N$, then $\phi_N$ satisfies the Cauchy problem

\begin{align}
\begin{cases}
(\phi_N)_{tt} - \Delta \phi_N + \phi_N + F_N(v_N) - F_N(\phi_N + v_N) = e \\
(\phi_N, \partial_t \phi_N)|_{t=0} = (v_0 - u_0, v_1 - u_1).
\end{cases}
\end{align}

By Strichartz estimates, Hölder's inequality, the boundedness of $P_N$ and the assumptions above,

\begin{align}
\|\phi_N\|_{L^4(I \times \mathbb{T}^3)} & \leq C \left( \|S(t)(v_0 - u_0, v_1 - u_1)\|_{L^4(I \times \mathbb{T}^3)} + \|F_N(v_N) - F_N(\phi_N + v_N)\|_{L^{4/3}_{t,x}(I \times \mathbb{T}^3)} + \|e\|_{L_{t,x}^{\tilde{q}' \tilde{r}'}(I \times \mathbb{T}^3)} \right) \\
& \leq C \left( 2\rho + (\rho_0)^2 \|\phi_N\|_{L^4(I \times \mathbb{T}^3)} + \|\phi_N\|_{L^6(I \times \mathbb{T}^3)}^3 \right),
\end{align}

hence a continuity argument yields (4.49) provided $\rho_0$ is sufficiently small. Similarly

\begin{align}
\|(\phi_N, \partial_t \phi_N)\|_{L_{t,x}^{\tilde{q}' \tilde{r}'}(I \times \mathbb{T}^3)} + \|\phi_N\|_{L_{t,x}^{\tilde{q}' \tilde{r}'}(I \times \mathbb{T}^3)} & \leq C \left( \|(v_0 - u_0, v_1 - u_1)\|_{H^{1/2}_x(I \times \mathbb{T}^3)} + \|F_N(v_N) - F_N(\phi_N + v_N)\|_{L^{4/3}_{t,x}(I \times \mathbb{T}^3)} + \|e\|_{L_{t,x}^{\tilde{q}' \tilde{r}'}(I \times \mathbb{T}^3)} \right) \\
& \leq C \left( R_1 + (\rho_0)^2 \|\phi_N\|_{L^4(I \times \mathbb{T}^3)} + \|\phi_N\|_{L^4(I \times \mathbb{T}^3)}^3 + \rho \right).
\end{align}

We conclude (4.50) by a continuity argument for $\rho_0 \equiv \rho_0(R_1) > 0$ sufficiently small. \qed

Remark 4.20. If one only requires bounds on the low frequency component $P_N(u_N - v_N)$, then

120
from the proof of the previous Lemma, it is clear that it suffices to assume that

\[ \|S(t - t_0)P_N(v_0 - u_0, v_1 - u_1)\|_{L^4(I \times \mathbb{T}^3)} \leq \rho \]

\[ \|P_N e\|_{L^4_t L^4_x(I \times \mathbb{T}^3)} \leq \rho. \]

Lemma 4.34 and the same proof as in Lemma 4.33 yields the long-time stability argument for the truncated equation with bounds uniform in the truncation parameter.

**Lemma 4.35 (Long-time stability).** Let \( I \subset \mathbb{R} \) a compact time interval and \( t_0 \in I \). Let \( v_N \) be a solution defined on \( I \times \mathbb{T}^3 \) of the Cauchy problem

\[
\begin{aligned}
\begin{cases}
(v_N)_{tt} - \Delta v_N + v_N + F_N(v_N) = e \\
(v_N, \partial_t v_N) \big|_{t=t_0} = (v_0, v_1) \in \mathcal{H}^{1/2}(\mathbb{T}^3).
\end{cases}
\end{aligned}
\]

Suppose that

\[ \|P_N v_N\|_{L^4(I \times \mathbb{T}^3)} \leq L \]

for some constant \( L > 0 \). Let \( (u_N, \partial_t u_N) \big|_{t=t_0} = (u_0, u_1) \in \mathcal{H}^{1/2}(\mathbb{T}^3) \) be such that

\[ \|(v_0 - u_0, v_1 - u_1)\|_{\mathcal{H}^{1/2}(\mathbb{T}^3)} \leq R_1 \]

for some \( R_1 > 0 \). Suppose also that we have the smallness conditions

\[ \|S(t - t_0)(v_0 - u_0, v_1 - u_1)\|_{L^4(I \times \mathbb{T}^3)} \leq \rho \]

\[ \|e\|_{L^4_t L^4_x(I \times \mathbb{T}^3)} \leq \rho, \]

for some \( 0 < \rho < \rho_1(R_1, I, L) \) a small constant and any \((\tilde{q}', \tilde{r}')\) a conjugate admissible pair. Then there exists a unique solution \((u_N, \partial_t u_N)\) to the truncated cubic nonlinear Klein-Gordon equation on \( I \times \mathbb{T}^3 \) with initial data \((u_0, u_1)\) at time \( t_0 \) and \( C = C(I) \geq 1 \) which satisfies

\[ \|v_N - u_N\|_{L^4(I \times \mathbb{T}^3)} \leq C \rho \]

\[ \|(v_N)^3 - (u_N)^3\|_{L^{4/3}(I \times \mathbb{T}^3)} \leq C \rho \]

\[ \|(v_N - u_N, \partial_t u_N - \partial_t v_N)\|_{L^\infty_t \mathcal{H}^{1/2}_x(I \times \mathbb{T}^3)} + \|v_N - u_N\|_{L^4_t L^4_x(I \times \mathbb{T}^3)} \leq C R_1. \]
4.9 Appendix B: Probabilistic bounds for the cubic NLKG

In this appendix, we record the proof of Proposition 4.19. The proof of Proposition 4.20 follows almost identically. We recall the statement.

**Proposition 4.20.** Let $0 < s < 1$ and let $\mu \in \mathcal{M}$. Then for any $\epsilon > 0$, there exist $C, c, \theta > 0$ such that for every $(u_0, u_1) \in \Sigma$, there exists $M > 0$ such that the family of global solution $u$ to cubic nonlinear Klein-Gordon equation (4.1) satisfies

\[
\begin{align*}
    u(t) &= S(t)(u_0, u_1) + w(t) \\
    \|(w(t), \partial_t w(t))\|_{\mathcal{H}^1} &\leq C(M + |t|)^{1+\epsilon} \\
    \|u(t)\|_{L^4(\mathbb{T}^3)} &\leq C(M + |t|)^{1/2+\epsilon}
\end{align*}
\]

and furthermore $\mu((u_0, u_1) \in \Sigma : M > \lambda) \leq Ce^{-c\lambda^\theta}$.

**Proof of Proposition 4.19.** Fix $\mu \in \mathcal{M}$. Following the proofs of Proposition 4.1 in [16] and Lemma 2.2 from [13], fix $\epsilon > 0$, $\rho > \frac{1}{2}$, $\bar{\rho} > 1/3$, and $\rho > 0$ and define

\[
\begin{align*}
    U_N &:= \{(v_0, v_1) \in \Sigma : \|\Pi_N(v_0, v_1)\|_{\mathcal{H}^1} \leq N^{1-s+\epsilon}\} \\
    G_N &:= \{(v_0, v_1) \in \Sigma : \|\Pi_N v_0\|_{L^4} \leq N^\epsilon\} \\
    H_N &:= \{(v_0, v_1) \in \Sigma : \|\langle t \rangle^{-\rho} \Pi_{\geq N} (v_0, v_1)\|_{L^2_{t,x} L^\infty_{t,x}} \leq N^{1-\epsilon-s}\} \\
    K_N &:= \{(v_0, v_1) \in \Sigma : \|\langle t \rangle^{-\bar{\rho}} \Pi_{\geq N} (v_0, v_1)\|_{L^2_{t,x} L^\infty_{t,x}} \leq N^{1-\epsilon-s}\} \\
    R_N &:= \{(v_0, v_1) \in \Sigma : \|\langle t \rangle^{-\rho} \Pi_{\geq N} S(t)(v_0, v_1)\|_{L^\infty_{t,x} L^4_{t,x}} \leq N^{1-\epsilon-s}\}.
\end{align*}
\]

We let

\[
E_N = U_N \cap G_N \cap H_N \cap K_N \cap R_N. \tag{4.53}
\]

The bounds on the measure of $E_N$ follow from Proposition 4.1 in [16] and Lemma 2.2 in [13]. We consider the inhomogeneous energy functional (4.5)

\[
\mathcal{E}(w(t)) = \frac{1}{2} \int |\nabla w|^2 + \frac{1}{2} \int |w|^2 + \frac{1}{2} \int |\partial_t w|^2 + \frac{1}{4} \int \langle w \rangle^4.
\]
Fix \((v_0, v_1) \in E_N\) and let \(w_N\) denote the solution to
\[
(w_N)_t - \Delta w_N + w_N + (w_N + S(t)\Pi\geq_N(v_0, v_1))^3 = 0, \quad (w_N, \partial_tw_N)|_{t=0} = \Pi_N(v_0, v_1)
\]
then
\[
\mathcal{E}(w_N(t))^{1/2} \leq CN^{1-s+\varepsilon}.
\]
Since \(\mathcal{E}\) controls the \(H^1\) norm, we no longer need to project away from constants to obtain (4.51).
To prove (4.52), note that by the definition of \(R_N\), we have
\[
\|u_N(t)\|_{L^4(T^3)} \leq \|w_N(t)\|_{L^4(T^3)} + \|S(t)\Pi\geq_N(v_0, v_1)\|_{L^4(T^3)}
\]
\[
\leq \mathcal{E}(w_N(t))^{1/4} + CN^{-s+\varepsilon}
\]
\[
\leq CN^{1-s+\varepsilon}.
\]
The conclusion then follows as in Proposition 4.1 in [16]. \(\square\)
Appendix A

Facts from Harmonic analysis

A.1 Strichartz Estimates

Definition A.1. A pair of real numbers \((q, r)\), is called \(\gamma\) admissible provided \(2 < q \leq +\infty\), \(2 \leq r < +\infty\),

\[
\frac{1}{q} + \frac{3}{r} = \frac{3}{2} - \gamma \quad \text{and} \quad \frac{1}{q} + \frac{1}{r} \leq \frac{1}{2}.
\]

We call a pair \((q', r')\) a conjugate admissible pair if

\[
\frac{1}{q'} + \frac{1}{r'} = 1 = \frac{1}{q} + \frac{1}{r}
\]

for some admissible pair \((q, r)\).

We state the following Strichartz estimates compactly for both the wave or Klein-Gordon equations on \(\Lambda = \mathbb{R}\) or \(\mathbb{T}\). These estimates are classical and due in parts to [65], [53], [26], [33].

Proposition A.2 (Strichartz estimates). Let \(u\) be a solution to the inhomogeneous equation

\[
(\partial_t^2 u - \Delta u + mu = F(u), \quad (u, \partial_t u)|_{t=0} = (u_0, u_1) \quad m = 0, 1
\]

on \(I \subset [0, T]\) for \(T\) fixed. Let \((q, r)\) be a \(\gamma\)-admissible pair, then for \(m = 1\)

\[
\|(u, \partial_t u)\|_{L_T^q L^{\gamma'}(I \times \Lambda)} + \|u\|_{L_T^q L^{\gamma'}(I \times \Lambda)} \lesssim \|(u_0, u_1)\|_{\mathcal{H}^\gamma(\Lambda)} + \|F\|_{L_T^{q'} L^{\gamma'}(I \times \Lambda)}, \quad (A.1)
\]
and similarly for \( m = 0 \) with homogeneous Sobolev spaces instead of inhomogeneous ones, and
\((\tilde{p}', \tilde{q}')\) is a conjugate admissible pair. When \( \Lambda = \mathbb{T} \) the implicit constant depends on the choice of \( T \) but is uniform for any subinterval \( I \subset [0, T] \). When \( \Lambda = \mathbb{R} \), the implicit constant is independent of \( T \) and we can take \( T = \infty \).

**Remark A.1.** In fact, a larger range of exponents are admissible for the nonlinear Klein-Gordon equation than the range stated above, which only coincides with the admissible exponents for the nonlinear wave equation. Since the above estimates are all we need in our arguments, we refrain from stating the full range of Strichartz exponents for simplicity. For a full formulation, see for instance [50], and references therein.

We use the following interpolation estimate for the Strichartz estimates for the nonlinear Klein-Gordon equation in Chapter 4 to estimate frequency localized functions in \( X^{s,b} \) spaces, see [7] for this formulation of the estimates.

**Proposition A.3 (Strichartz).** Let \( I \subset \mathbb{R} \). There exists a constant \( C \equiv C(I) > 0 \) such that

\[
\left\| \sum_{|n| \sim N} a(n) e^{i(x_n + t(n))} \right\|_{L^4(I \times \mathbb{T}^3)} \leq C(I) N^{1/2} \left( \sum |a(n)|^2 \right)^{1/2}.
\]

By Hölder’s inequality we have

\[
\left\| \sum_{|n| \sim N} \int d\tau \frac{c(n, \tau)}{(|\tau| - \langle n \rangle)^{1/2}} e^{i(x_n + \tau t)} \right\|_{L^4(I \times \mathbb{T}^3)} \leq C(I) N^{1/2} \left( \int \sum |c(n, \tau)|^2 d\tau \right)^{1/2}
\]

and by interpolating with Parseval’s identity we obtain that for \( 2 \leq r \leq 4 \)

\[
\left\| \sum_{|n| \sim N} \int d\tau \frac{c(n, \tau)}{(|\tau| - \langle n \rangle)^{1/2}} e^{i(x_n + \tau t)} \right\|_{L^r(I \times \mathbb{T}^3)} \leq C(I) N^{\theta/2} \left( \int \sum |c(n, \tau)|^2 d\tau \right)^{1/2} \tag{A.2}
\]

where \( \theta = 2 - \frac{4}{r} \).
A.2 Adapted Function spaces

A.2.1 $X^{s,b}$ spaces

For a good overview of these spaces, see Chapter 2.6 in [66]. For completeness, we recall the definition of $X^{s,b}(\mathbb{R} \times T^3)$ spaces, with norm

$$
\|u\|_{X^{s,b}(\mathbb{R} \times T^3)} = \|\langle n \rangle^s |\tau| - \langle n \rangle^b \widehat{u}(n, \tau)\|_{L^2_T L^2_x}.
$$

We also work with the local-in-time restriction spaces $X^{s,b,\delta}$, which are defined by the norm

$$
\|u\|_{X^{s,b,\delta}} = \inf\{\|\widehat{u}\|_{X^{s,b}(\mathbb{R} \times T^3)} : \widehat{u}|_{[-\delta, \delta]} = u\}.
$$

We have the obvious inclusions

$$
X^{s',b'} \subseteq X^{s,b}
$$

for $s \leq s'$ and $b \leq b'$. We remark that these spaces are not invariant under conjugation or modulation but they are invariant under translation.

Heuristically, these spaces measure how far a given function is from being a free solution. Additionally, free solutions lie in $X^{s,b}$, at least when time is localized.

**Lemma A.4.** Let $f \in H^s$ for $s \in \mathbb{R}$ and let $S(t)$ denote the free evolution for the Klein-Gordon equation. Then for any Schwartz time cutoff $\eta \in S_x(\mathbb{R})$,

$$
\|\eta(t)S(t)f\|_{X^{s,b}(\mathbb{R} \times T^3)} + \|\eta(t)\partial_t S(t)f\|_{X^{s-1,b}(\mathbb{R} \times T^3)} \leq c(\eta, b) \|f\|_{H^s(\mathbb{T}^3)}.
$$

We note that it does not hold that free solutions lie in $X^{s,b}$ globally hence these spaces are really only suitable for local theory. An important property of these spaces is the so-called transfer principle which allows one to convert bounds for free solutions into bounds for $X^{s,b}$ functions.

**Lemma A.5** (Lemma 2.9, [66]). Let $\Lambda = \mathbb{R}$ or $\mathbb{T}$ and let $L = iP(\nabla/i)$ for some polynomial $P : \mathbb{R}^d \rightarrow \mathbb{R}$, and let $s \in \mathbb{R}$ and let $Y$ be a Banach space of functions on $\mathbb{R} \times \Lambda^d$ such that

$$
\|e^{i\tau\sigma}e^{tL}f\|_Y \lesssim \|f\|_{H^s_x(\Lambda^d)}
$$
for all \( f \in H^s_x(\mathbb{R}^d) \) and \( \tau_0 \in \mathbb{R} \). Then for \( b > \frac{1}{2} \)

\[
\|u\|_Y \lesssim_b \|u\|_{X^{s,b}(\mathbb{R}^d)}.
\]

Letting \( Y \) be the Strichartz spaces from Proposition A.2, we obtain the following corollary immediately.

**Corollary A.6.** Let \((q,r)\) be Strichartz admissible pairs and let \( b > \frac{1}{2} \). Then

\[
\|u\|_{L^q_t L^r_x} \lesssim \|u\|_{X^{0,s}}.
\]

In particular, for \( b > \frac{1}{2} \), \( X^{s,b} \) embeds into \( C_t H^s_x \) (for both the full and restricted spaces on the appropriate domains). This embedding fails at the endpoint \( b = \frac{1}{2} \) and should be thought of analogously to the failure of the endpoint Sobolev embedding \( L^\infty \nsubseteq H^{\frac{3}{2}} \). It is precisely at the endpoint \( b = \frac{1}{2} \) that these spaces respect the scaling of \( C_t H^s \) and for critical problems where scale invariance is an issue, one no longer has the appropriate control in order to close the contraction mapping argument. It is possible to remedy this problem by including a Besov space type refinement, however we will focus instead on \( U^p \) and \( V^p \) spaces.

### A.2.2 \( U^p \) and \( V^p \) Function spaces

In this section, we introduce the basic facts we need about the \( U^p \) and \( V^p \) spaces used in Chapter 3. We follow the exposition in [28]. Consider partitions given by a strictly increasing finite sequence \(-\infty < t_0 < t_2 < \ldots t_K \leq \infty \). If \( t_K = \infty \) we use the convention \( v(t_K) := 0 \) for all functions \( v : \mathbb{R} \to H \). We usually work on bounded intervals \( I \subset \mathbb{R} \). A step functions associated to a partition is a function which is constant on each open sub-interval of the partition. In the sequel, we let \( \mathcal{B} \) denote an arbitrary Banach space.

**Definition A.7** (\( U^p \) spaces). Let \( 1 \leq p < \infty \). Consider a partition \( \{t_0, \ldots, t_K\} \) and let \( (\varphi_k)_{k=0}^{K-1} \subseteq \mathcal{B} \) with \( \sum_{k=0}^{K-1} \|\varphi_k\|_{L^2}^p = 1 \). We define a \( U^p \) atom to be a function

\[
a = \sum_{k=1}^{K} \mathbb{1}_{[t_{k-1}, t_k)} \varphi_{k-1}
\]
and we define the atomic space \( U^p(\mathbb{R}, \mathcal{B}) \) to be the set of all functions \( u : \mathbb{R} \to \mathcal{B} \) such that

\[
u = \sum_{j=1}^{\infty} \lambda_j a_j,
\]

for \( a_j \) \( U^p \) atoms, and \( \{\lambda_j\} \in \ell^1(\mathbb{C}) \), endowed with the norm

\[
\|u\|_{U^p} := \inf \left\{ \sum_{j=1}^{\infty} |\lambda_j|, \ u = \sum_{j=1}^{\infty} \lambda_j a_j : a_j \text{ is a } U^p \text{ atom} \right\}.
\]

**Remark A.2.** This yields a Banach space which satisfies the embeddings

\[
U^p(\mathbb{R}, \mathcal{B}) \hookrightarrow U^q(\mathbb{R}, \mathcal{B}) \hookrightarrow L^\infty(\mathbb{R}, \mathcal{B})
\]

for \( 1 \leq p < q < \infty \). Furthermore, every \( u \in U^p \) is right-continuous and \( \lim_{t \to -\infty} u(t) = 0 \).

**Definition A.8** (\( V^p \) spaces). Let \( 1 \leq p < \infty \). We define \( V^p(\mathbb{R}, \mathcal{B}) \) as the space of all \( \mathcal{B} \) valued functions, \( v \), such that the norm

\[
\|v\|_{V^p(\mathbb{R}, \mathcal{B})} = \sup_{\text{partitions}} \left( \sum_{i=1}^{K} \|v(t_i) - v(t_{i-1})\|_{\mathcal{B}}^p \right)^{1/p} < \infty
\]

with the convention \( v(\infty) = 0 \). We let \( V_-(\mathbb{R}, \mathcal{B}) \) denote the subspace of all functions satisfying \( \lim_{t \to -\infty} v(t) = 0 \) and we let \( V^p_{rc}(\mathbb{R}, \mathcal{B}) \) denote the subspace of all right continuous functions in \( V_-(\mathbb{R}, \mathcal{B}) \), endowing both these subspaces with the above norm.

**Remark A.3.** Note that for \( 1 \leq p < \infty \) we have the embeddings

\[
U^p(\mathbb{R}, \mathcal{B}) \hookrightarrow V^p_{rc}(\mathbb{R}, \mathcal{B}) \hookrightarrow L^\infty(\mathbb{R}, \mathcal{B}) \quad \text{and} \quad V^p(\mathbb{R}, \mathcal{B}) \hookrightarrow V^q(\mathbb{R}, \mathcal{B}).
\]

If further \( 1 < p < q < \infty \), then

\[
V^p_{rc}(\mathbb{R}, \mathcal{B}) \hookrightarrow U^q(\mathbb{R}, \mathcal{B}). \tag{A.3}
\]

A crucial property of the \( U^p \) and \( V^p \) spaces is the following duality relation.
Theorem A.9. Let $1 < p < \infty$ and $\frac{1}{p} + \frac{1}{p'} = 1$. Then

$$(U^p(\mathbb{R}, B))^* = V^{p'}(\mathbb{R}, B^*),$$

that is, there is a bounded bilinear form

$$T : V^{p'}(\mathbb{R}, B^*) \to (U^p(\mathbb{R}, B))^*, \quad T(v) := B(\cdot, v)$$

which is an isometric isomorphism.

Proposition A.10 (Proposition 2.9, [28]). For $1 < p < \infty$, let $u \in U^p$ be continuous and $v, v^* \in V^{p'}$ for $\frac{1}{p} + \frac{1}{p'} = 1$ such that $v(s) = v^*(s)$ except for at most countably many points. Then

$$B(u, v) = B(u, v^*).$$

Proposition A.11. Let $1 < p < \infty$, $u \in V^1(B)$ be absolutely continuous on compact intervals and $v \in V^{p'}(B^*)$ for $\frac{1}{p} + \frac{1}{p'} = 1$. Then,

$$B(u, v) = -\int_{-\infty}^{\infty} (v(t), u'(t)) g_*, g \ dt.$$

In particular $B(u, v) = B(u, \tilde{v})$ if $v(t) = \tilde{v}(t)$ almost everywhere. Consequently, $v$ can be replaced by its right-continuous version.

Proof. See [28] \hfill \square

Remark A.4 (Remark 2.11 in [28]). Let $1 < p < \infty$ and $u \in U^p$. Then for $\frac{1}{p} + \frac{1}{p'} = 1$ one clearly has

$$\|u\|_{U^p} = \sup_{v \in V^{p'} : \|v\|_{V^{p'}} = 1} |B(u, v)|.$$ 

However, in light of Proposition A.10, one can restrict to taking a supremum over right-continuous functions, which we do in the estimates in Chapter 3. The bilinear form above should be thought of as the corresponding the Stieltjes integral

$$\int f \, dg = \sum_{i=1}^{n} f(t_i)(g(t_{i+1}) - g(t_i)).$$
We define variants of the $U^p$ and $V^p$ spaces adapted to the linear propagator for the Klein-Gordon equation $e^{\pm it(\nabla)}$. In the sequel we will suppress the notational dependence on $I \subset \mathbb{R}$ and the Banach space $B$.

**Definition A.12.** Define $U^p_\pm$ the space $U^p$ adapted to the linear propagator equipped with the norm

$$\|u\|_{U^p_\pm} = \|e^{\pm it(\nabla)}u\|_{U^p}$$

and similarly for the $V^p$ spaces.

**Remark A.5.** The space $U^p_\pm$ is again an atomic space with atoms $\tilde{a} = e^{it(\nabla)}a$ for a $U^p$ atom, $a$. Further, in the case that $B = H^s$, we compare the above definition to that of the $X^{s,b}$ spaces associated to the Klein-Gordon equation given by $\|u\|_{X^{s,b}} = \|e^{-it(\nabla)}u\|_{H^s_t H^b_x}$.

A useful feature of the $U^p$ and $V^p$ spaces is that they satisfy a transfer principle, namely one can transfer multilinear estimates for free solutions to estimates for $U^2_\pm$ functions.

**Proposition A.13** (Transfer principle, Proposition 2.19 [28]). Let $T_0 : L^2_x \times \ldots \times L^2_x \to L^1_{loc}$ be an $m$-linear operator. Suppose that for some $1 \leq q, r \leq \infty$

$$\|T_0(S(t)\phi_1, \ldots, S(t)\phi_m)\|_{L^q_x L^r_x} \lesssim \prod_{i=1}^m \|\phi_i\|_{L^2_x}$$

then there exists an extension $T_0 : U^q_\pm \times \ldots \times U^q_\pm \to L^q_x L^r_x$ with

$$\|T(u_1, \ldots u_m)\|_{L^q_x L^r_x} \lesssim \prod_{i=1}^m \|u_i\|_{U^q_\pm}$$

which agrees with $T_0$ for almost every $t \in \mathbb{R}$.

**Remark A.6.** The transfer principle is the key tool which allows us to use the Strichartz estimates available for free solutions of the Klein-Gordon equation to derive bounds in $U^p$ and $V^p$ spaces. Because of the duality relation, typically one needs to put one function in $V^p$ when performing multilinear estimates. In these cases, the following proposition demonstrates that this is possible if one allows for a logarithmic loss.
Proposition A.14. Let $q_1, \ldots, q_m > 2$ where $m = 1, 2$ or $3$ and let $E$ be a Banach space and $T : U^{q_1} \times \cdots \times U^{q_m} \to E$ be a bounded $m$-linear operator with

$$
\| T(u_1, \ldots, u_m) \|_E \leq C \prod_{i=1}^{m} \| u_i \|_{U^{q_i}_+}.
$$

Suppose further there exists some $0 < C_2 \leq C$ such that the estimate

$$
\| T(u_1, \ldots, u_m) \|_E \leq C_2 \prod_{i=1}^{m} \| u_i \|_{U^{q_i}_+}
$$

holds. Then $T$ satisfies

$$
\| T(u_1, \ldots, u_m) \|_E \leq C_2 \left( \log \frac{C}{C_2} + 1 \right) \prod_{i=1}^{m} \| u_i \|_{V^{2,p}}, \quad u_i \in V^{p}_{rc, \pm}.
$$
Bibliography


134


