Oberwolfach Seminars
45

Dispersive Equations and Nonlinear Waves

Herbert Koch
Daniel Tataru
Monica Vişan

Birkhäuser

Birkhäuser

Oberwolfach Seminars
Volume 45

Herbert Koch
Daniel Tataru
Monica Vişan

Dispersive Equations and Nonlinear Waves

Generalized Korteweg–de Vries, Nonlinear Schrödinger, Wave and Schrödinger Maps

Birkhäuser

Herbert Koch
Institute of Mathematics
University of Bonn
Bonn, Germany

Daniel Tataru
Department of Mathematics
University of California
Berkeley, CA, USA

Monica Vişan
Department of Mathematics
University of California
Los Angeles, CA, USA

ISSN 1661-237X
ISSN 2296-5041 (electronic)
ISBN 978-3-0348-0735-7
ISBN 978-3-0348-0736-4 (eBook)
DOI 10.1007/978-3-0348-0736-4
Springer Basel Heidelberg New York Dordrecht London

Library of Congress Control Number: 2014944740

Mathematics Subject Classification (2010): 35Q51, 35Q53, 35Q55, 35L70

Author portraits of Herbert Koch and Monica Vişan on the back cover: Photo collection of the Mathematisches Forschungsinstitut Oberwolfach, Germany.

Printed on acid-free paper

Springer Basel is part of Springer Science+Business Media (www.birkhauser-science.com)

Contents

Preface

Nonlinear wave equations are ubiquitous in physics and applied sciences; for example, they appear as effective equations in general relativity and elasticity, for water waves, nonlinear optics and superconductivity, in gauge theory and the motion of Bose-Einstein condensates. Despite encompassing a large class of equations, there are recurrent themes: dispersion, solitons and their stability, blow up and scattering. In these notes we want to pursue different coordinated threads: the nonlinear Schrödinger equation and generalized KdV, critical wave and Schrödinger equations, and geometric dispersive equations. Our goal is to introduce the equations and to describe an array of ideas and techniques used in their study, leading up to the current results and remaining open problems.

The first part is devoted to the nonlinear Schrödinger equation, the generalized Korteweg-de Vries equation, and the Kadomtsev-Petviashvili II equation. It introduces basic techniques, stationary phase and Strichartz estimates, the space of functions of bounded p-variation and their adaptation to dispersive equations, convolutions and bilinear estimates. The nonlinear Schrödinger equation and the (generalized) Korteweg-de Vries equation exhibit a fascinating and rich structure. They provide the simplest but nevertheless nontrivial context for many important techniques, as well as the simplest framework for open challenging questions. The last section of the first part describes a scheme for constructing solutions to dispersive equations, often in scale-invariant function spaces; this is demonstrated in the context of the generalized Korteweg-de Vries equation and the Kadomtsev-Petviashvili II equation.

Over the last decade, the induction on energy paradigm has grown into a powerful tool for the large-data analysis of evolution equations. It continues to develop in both depth and breadth, and has already proven useful over a wide range of equations, from semilinear wave and Schrödinger equations to fluid equations and geometric flows. While enjoying many parallels to the calculus of variations and often using its terminology, this new approach requires an independent set of techniques. In these notes we use the energy-critical nonlinear Schrödinger equation as a model to demonstrate these methods and their application to the question of large-data global well-posedness.

Within the field of nonlinear dispersive equations, a special role is played by the so-called geometric dispersive equations, which arise from the standard

Lagrangian or Hamiltonian formalism, but applied in a geometric context. The two simplest examples of such equations are the wave map and Schrödinger map equations. These are discussed in the third part. The emphasis is on wave maps, where even the small-data problem poses new challenges, both of technical nature (function spaces) and conceptually (renormalization). In addition, the large-data problem brings back techniques such as induction on energy and Morawetz estimates. All of this happens on top of a differential geometry layer which needs to be understood first. The corresponding elliptic and parabolic analogues, namely, harmonic maps and the harmonic map heat flow, also play a role. The last section concludes with a discussion of the small-data problem for Schrödinger maps; there the large-data problem is still open.

These notes grew out of an Oberwolfach seminar held in the fall of 2012 where each of the authors gave five 90 minutes lectures. We want to thank the Mathematisches Forschungsinstitut at Oberwolfach for the opportunity to organize this workshop, and the participants for lively discussions. H. Koch acknowledges the support of the Hausdorff Center of Mathematics and the SFBs 611 and 1060. D. Tataru was supported by the NSF grants DMS-0801261 and DMS-1266182, as well as by a Simons Fellowship and a Simons Investigator award from the Simons Foundation. M. Visan was supported by the Sloan Foundation and NSF grants DMS-0901166 and DMS-1161396.

Nonlinear Dispersive Equations

Herbert Koch

Chapter 1

Introduction

Nonlinearly interacting waves are often described by asymptotic equations. The derivation typically involves an ansatz for an approximate solution where higher order terms – the precise meaning of higher order term depends on the context and the relevant scales – are neglected. Often a Taylor expansion of a Fourier multiplier is part of that process.

There is an immediate consequence: This type of derivation leads to a huge set of asymptotic equations, and one should search for a general understanding of interacting nonlinear waves by asking for precise results for specific equations.

The most basic asymptotic equation is probably the nonlinear Schrödinger equation, which describes wave trains or frequency envelopes close to a given frequency, and their self-interactions. The Korteweg–de-Vries equation typically occurs as first nonlinear asymptotic equation when the prior linear asymptotic equation is the wave equation. It is one of the amazing facts that many generic asymptotic equations are integrable in the sense that there are many formulae for specific solutions, conserved quantities, Lax Pairs and Bi-Hamiltonian structures.

This text will focus on adapted function spaces and their recent application to a number of dispersive equations. They are build on functions of bounded p-variation, and their companion, the atomic space U^p. Combined with stationary phase resp. Strichartz estimates and bilinear refinements thereof, they provide an alternative to the Fourier restriction spaces $X^{s,b}$ which is better suited for scaling critical problems.

We discuss the method of stationary phase and dispersive estimates in Section 2, the application to the nonlinear Schrödinger equation in Section 3, the spaces U^p and V^p in Section 4, bilinear estimates in Section 5, and applications to nonlinear dispersive equations in Section 6.

In order to make these notes reasonably self-contained there are three appendices on Young's inequality, real and complex interpolation, on Bessel functions, and on the Fourier transform.

H. Koch et al., *Dispersive Equations and Nonlinear Waves: Generalized Korteweg–de Vries, Nonlinear Schrödinger, Wave and Schrödinger Maps*, Oberwolfach Seminars 45,
DOI 10.1007/978-3-0348-0736-4_1,

Chapter 2

Stationary phase and dispersive estimates

We begin with the evaluations of several integrals. Let m^d be the d-dimensional Lebesgue measure and define

$$I_d = \int_{\mathbb{R}^d} e^{-|x|^2} dm^d(x).$$

Then, with Fubini,

$$\begin{aligned}
I_{d_1+d_2} &= \int_{\mathbb{R}^{d_1}\times\mathbb{R}^{d_2}} e^{-|x_1|^2-|x_2|^2} dm^{d_1+d_2}(x) \\
&= \int_{\mathbb{R}^{d_1+d_2}} e^{-|x_1|^2} e^{-|x_2|^2} dm^{d_1+d_2}(x) \\
&= \int_{\mathbb{R}^{d_1}} e^{-|x_1|^2} \int_{\mathbb{R}^{d_2}} e^{-|x_2|^2} dm^{d_2} dm^{d_1} \\
&= I_{d_2} \int_{\mathbb{R}^{d_1}} e^{-|x_1|^2} dm^{d_1} \\
&= I_{d_1} I_{d_2},
\end{aligned}$$

hence

$$I_d = I_1^d.$$

Applying Fubini twice, we get

$$\begin{aligned}
I_d &= m^{d+1}(\{(x,t) : 0 < t < e^{-|x|^2}\}) \\
&= \int_0^1 m^d(\{x : e^{-|x|^2} > t\}) dt \\
&= \int_0^1 m^d(B_{(-\ln(t))^{1/2}}(0)) dt
\end{aligned}$$

H. Koch et al., *Dispersive Equations and Nonlinear Waves: Generalized Korteweg–de Vries, Nonlinear Schrödinger, Wave and Schrödinger Maps*, Oberwolfach Seminars 45,
DOI 10.1007/978-3-0348-0736-4_2, © Springer Basel 2014

$$
\begin{aligned}
&= m^d(B_1(0)) \int_0^1 (-\ln(t))^{d/2} dt \\
&= m^d(B_1(0)) \int_0^\infty s^{d/2} e^{-s} ds \\
&= m^d(B_1(0)) \Gamma(\frac{d}{2}+1)
\end{aligned}
$$

and hence $I_2 = \pi$, $I_1 = \sqrt{\pi}$, $I_d = \pi^{d/2}$,

$$
m^d(B_1(0)) = \frac{\pi^{d/2}}{\Gamma(\frac{d}{2}+1)},
$$

and

$$
\Gamma(\frac{1}{2}) = 2\Gamma(\frac{3}{2}) = \sqrt{\pi}.
$$

We proceed with

$$
I(\tau) := \int_{-\infty}^{\infty} e^{-\frac{\tau}{2}x^2} dx
$$

for $\operatorname{Re}\tau > 0$. Then

$$
\begin{aligned}
\frac{d}{dt}\sqrt{t+is} I(t+is) &= \frac{1}{2(t+is)}\sqrt{t+is} I(t+is) - \frac{1}{2}\sqrt{t+is}\int_{-\infty}^{\infty} e^{-\frac{t+is}{2}x^2} x^2 dx \\
&= \frac{\sqrt{t+is}}{2(t+is)} \left(I(t+is) + \int_{-\infty}^{\infty} \frac{d}{dx} e^{-\frac{t+is}{2}x^2} x dx \right) \\
&= \frac{\sqrt{t+is}}{2(t+is)} \left(I(t+is) - \int_{-\infty}^{\infty} e^{-\frac{t+is}{2}x^2} dx \right) \\
&= 0
\end{aligned}
$$

and similarly

$$
\frac{d}{ds}\sqrt{t+is} I(t+is) = 0
$$

Thus

$$
\sqrt{\tau} I(\tau) = \sqrt{2} I(2) = \sqrt{2\pi}
$$

and hence

$$
\int e^{-\frac{\tau}{2}x^2} dx = \sqrt{\frac{2\pi}{\tau}}. \tag{2.1}
$$

Now we fix τ and study

$$
\int e^{-\frac{\tau}{2}x^2} x^k dx.
$$

This vanishes when k is odd, since then the integrand is an odd function. Let

$$\begin{aligned} J(k) = \int e^{-\frac{\tau}{2}x^2} x^{2k} dx &= \frac{2k-1}{\tau} J(k-1) \\ &= 1 * 3 * \cdots * (2k-1)\tau^{-k}\sqrt{\frac{2\pi}{\tau}} \\ &= \frac{1}{2^k k!}\Big(\tau^{-1}\frac{d^2}{dx}\Big)^k x^{2k}\Big|_{x=0}\sqrt{\frac{2\pi}{\tau}}, \end{aligned}$$

where the second equality follows by an integration by parts. Let p be a polynomial. It is a sum of monomials and hence

$$\int e^{-\frac{\tau}{2}x^2} p(x)dx = \sqrt{\frac{2\pi}{\tau}} \sum_{k=0}^{\infty} \frac{1}{2^k k!}\Big(\tau^{-1}\frac{d^2}{dx}\Big)^k p(x)\Big|_{x=0}$$

The higher-dimensional case is contained in the following lemma. Let $A = A_0 + iA_1$ be a real symmetric $d \times d$ matrix with A_0 positive definite. This is equivalent to all eigenvalues λ_j being in $\{\lambda : \operatorname{Re}\lambda > 0\}$. Let (a_{ij}) be the inverse. By an abuse of notation, we set

$$\det(A)^{-1/2} = \prod \lambda_j^{-1/2},$$

where the λ_j are the eigenvalues of A.

Lemma 2.1. *Let p be a polynomial. Then*

$$\int e^{-\frac{1}{2}x^T Ax} p(x)\, dx = (2\pi)^{d/2}(\det A)^{-1/2} \sum_{k=0}^{\infty} \frac{1}{2^k k!}\Big(\sum_{i,j=1}^{d} a_{ij}\partial_{ij}^2\Big)^k p(x)\Big|_{x=0}. \quad (2.2)$$

The sum contains only finitely many non-vanishing terms.

Proof. We begin with a fact from linear algebra and claim that there exist a real $d \times d$ matrix B and a diagonal matrix D such that

$$A = BDB^T.$$

By the Schur decomposition, there exist an orthogonal matrix O and a diagonal matrix D_0 with non-negative entries such that

$$A_0 = OD_0O^T.$$

We set $B_0 = O\sqrt{D_0}$. Then

$$A_0 + iA_1 = B_0(1 + iB_0^{-1}A_1B_0^{-T})B_0^T$$

Again by the Schur decomposition, there exist an orthogonal matrix U and a diagonal matrix D_1 such that

$$B_0^{-1}A_1B_0^{-T} = UD_1U^T,$$

hence

$$A_0 + iA_1 = B(1 + iD_1)B^T,$$

with $B = B_0 U$. We set $D = 1_{\mathbb{R}^d} + iD_1$.

We change coordinates to $y = B^T x$. Then

$$\int e^{-\frac{x^T A x}{2}} p(x) dm^d(x) = (\det B)^{-1} \int e^{-\frac{y^T(1+iD_1)y}{2}} p(B^{-T} y) dm^d(y),$$

and by Fubini and the previous calculations,

$$\int e^{-\frac{y^T(1+iD_1)y}{2}} y^\alpha dm^d(y) = 0$$

if one of the indices is odd, and otherwise, with d_j denoting the diagonal entries of D_1,

$$\begin{aligned}\int e^{-\frac{y^T D y}{2}} y^{2\alpha} dm^d(y) &= (2\pi)^{d/2} \det(D)^{-1/2} \frac{1}{2^{|\alpha|}\alpha!} \prod ((1 + id_j)^{-1} \partial^2_{y_j y_j})^{\alpha_j} y_j^{2\alpha_j} \Big|_{y=0} \\ &= (2\pi)^{d/2} \det(D)^{-1/2} \frac{1}{2^{|\alpha|}|\alpha|!} \Big[\sum_{j=1}^d (1 + id_j)^{-1} \partial_j^2 \Big]^{|\alpha|} y^\alpha \Big|_{y=0}.\end{aligned}$$

Thus, for any polynomial q,

$$\int e^{-\frac{y^T D y}{2}} q(y) dm^d(y) = (2\pi)^{d/2} \det(D)^{-1/2} \sum_{k=0}^\infty \frac{1}{2^k k!} \Big[\sum_{j=1}^d (1 + id_j)^{-1} \partial_j^2 \Big]^k q(y) \Big|_{y=0}.$$

We complete the calculation by

$$(\det A)^{1/2} = (\det D)^{1/2} |\det B|$$

and, by the chain rule,

$$\sum a_{ij} \partial^2_{x_i x_j} p(x) = \Big[\sum_{j=1}^d (1 + id_j)^{-1} \partial_j^2 \Big] p(B^{-T} y). \qquad \square$$

Observe that the right-hand sides of the formulas have a limit as A tends to a purely imaginary invertible matrix. We call the integral on the left-hand side oscillatory integral in that limit.

Oscillatory integrals play a crucial role when studying dispersive equations. We consider

$$I(\tau) = \int_{\mathbb{R}^s} a(\xi) e^{i\tau\phi(\xi)} d\xi,$$

where a and ϕ are smooth functions. The simplest result is

Lemma 2.2. *Suppose that $a \in C_0^\infty(\mathbb{R}^d)$, $\phi \in C^\infty(\mathbb{R}^d)$ with $\operatorname{Im}\phi \geq 0$ and*

$$|\nabla \phi| + \operatorname{Im}\phi > 0$$

on supp a. Given $N > 0$, there exists c_N such that

$$|I(\tau)| \leq c_N \tau^{-N}.$$

The constant c depends only on N, the lower bound above, and derivatives up to order N.

Proof. By compactness, there is $\kappa > 0$ such that

$$|\nabla \phi| + \operatorname{Im}\phi > \kappa$$

on supp $a(\xi)$. Using a partition of unity we may restrict to the two cases:

(1) $\operatorname{Im}\phi > \kappa/2$ on supp a, in which case we get a bound $Ce^{-\kappa\tau/2}$;
(2) $|\nabla\phi| \geq \kappa/2$ on supp a, which we consider now.

We write

$$\begin{aligned}\int a(\xi)e^{i\tau\phi(\xi)}d\xi =&(i\tau)^{-1}\int a(\xi)|\nabla\phi|^{-2}\nabla\phi\nabla e^{i\tau\phi(\xi)}d\xi \\ =&-(i\tau)^{-1}\int \Big(\nabla\cdot\Big(\frac{a(\xi)\nabla\phi}{|\nabla\phi|^2}\Big)\Big)e^{i\tau\phi(\xi)}d\xi\end{aligned}$$

which is again an integral of the same type. Induction implies the full statement. □

In many cases these bounds hold even for non-compactly supported a.

Lemma 2.3. *Suppose that $A = A_0 + iA_1$ is invertible with A_0 positive semi-definite. Let $\eta \in C_0^\infty(\mathbb{R}^d)$ be identically 1 in a ball of radius 1, and supported in $B_2(0)$, and let a be a smooth function with uniformly bounded derivatives of order $M \geq N - \frac{d}{2}$ for some $M, N > 0$ and $0 < s < \frac{1}{2}$. Then*

$$\Big|\int e^{-\frac{\tau}{2}x^TAx}e^{-\varepsilon|x|^2}a(x)(1-\eta(x\tau^s))dm^d(x)\Big| \leq c_N\tau^{-N},$$

with c_N depending only on N, the norm of A and its inverse, and derivatives up to some order M of a, but not on $\varepsilon > 0$. The limit $\varepsilon \to 0$ exists.

We will use the formula with $\varepsilon = 0$.

Proof. We argue similarly as above. Each integration by parts adds a factor of size $(\tau|x|)^{-1}$ if the derivative falls on a, a factor of size $(\tau|x|^2)^{-1}$ if the derivative falls on $\frac{Ax}{|Ax|^2}$, and τ^s if the derivative falls on η. On the support of $\nabla\eta$

$$\tau^{-1}|x|^{-2} + \tau^{s-1}|x|^{-1} \leq \tau^{2s-1}.$$

We integrate by parts (and split the summands) until either

(1) M derivatives fall on a, or

(2) $\frac{N}{1-2s}+d$ derivatives fall on the other terms.

The integrand (after the integrations by parts) converges pointwise, with a majorant as above. This implies the statement on the limit as $\varepsilon \to 0$. $\square$

Similar statements hold for more general phase functions if

$$|\nabla\phi| \geq c|x|^{\delta} \qquad \text{for } |x| \geq R$$

and

$$|\partial^\alpha \phi| \leq |x|^{-\delta}|\nabla\phi| \qquad \text{for } |x| \geq R$$

some R and δ, and $|\alpha| \geq 2$.

Lemma 2.4. *Let A be invertible, symmetric, with real part positive semi-definite, and $\psi \in C^\infty$ with bounded derivatives of order $\geq M$. Given N with $0 \leq N \leq (M+d)/3$, there exist $c_n > 0$ such that for $\tau > 0$ and $L \geq N - \frac{d}{2}$,*

$$\left| \int e^{-\frac{\tau}{2}x^T A x}\psi(x)dx - (2\pi)^{d/2}\tau^{-d/2}(\det A)^{-1/2}\sum_{k=0}^{L}\tau^{-k}\frac{1}{2^k k!}\Big(\sum_{ij} a_{ij}\partial^2\Big)^k \psi\Big|_{x=0} \right| \leq c_N\tau^{-N}. \tag{2.3}$$

Proof. We subtract the Taylor expansion p of ψ at 0 up to some order L. We choose $0 < s = \frac{1}{3} < \frac{1}{2}$ and decompose the integral into

$$\int e^{-\frac{1}{2}x^T A x}\left[p(x) + \eta(x\tau^s)(\psi(x)-p(x)) + [1-\eta(x\tau^s)](\psi(x)-p(x))\right]dx.$$

The integral of the first summand has been evaluated in Lemma 2.1 The integral of the third summand is small by Lemma 2.3, and the one of the second summand is bounded by

$$\tau^{-(d+L)/3},$$

by a direct estimate. This gives estimate (2.3) for an L which may be too large. Inspection of the sum shows that we may omit terms which are smaller than the right-hand side. $\square$

Now we consider

$$I(\tau) = \int e^{i\tau\phi(x)}\psi(x)dx,$$

where ψ is compactly supported, 0 is the only point in the support where the imaginary part of ϕ and $\nabla\phi$ vanish, the imaginary part of ϕ is non-negative, and the Hessian of ϕ at 0 is invertible.

Lemma 2.5. *Let $\frac{1}{3} < s < \frac{1}{2}$. Then, with η as above, $\psi \in C_0^\infty$ and $N > 0$,*

$$\left| \int e^{i\tau\phi(x)}(1-\eta(x\tau^s))\psi(x)dx \right| \le c_N \tau^{-N}.$$

Proof. The proof is the same as for the quadratic phase. Again in this formula the compact support assumption on ψ can be weakened. □

We write

$$\phi(x) = \phi_0 + \frac{i}{2}x^T Ax + \psi(x),$$

where A is invertible and ψ is smooth, with $\psi(x) = O(|x|^3)$.

Theorem 2.6 (Stationary phase)**.** *Let a be a smooth compactly supported function on $\mathbb{R}^d$, and ϕ a phase function as above. Given $N > 0$ there exists c_N such that for $\tau > 1$*

$$\left| \int e^{i\tau\phi}a(x)dx - (2\pi)^{d/2}\tau^{-d/2}(\det A)^{-1/2}e^{i\tau\phi_0}\sum_{k=0}^{N}\frac{1}{2^k k! \tau^k}[(a_{ij}\partial^2)^k e^{i\tau\psi(x)}a(x)]_{x=0} \right| \le c_N \tau^{-d/2-\frac{N+1}{3}}.$$

Proof. We assume that the real part of A is positive definite. The general statement follows then by an obvious limit.

We choose M large and write $e^{i\tau\psi}\psi = p_M(x) + r_M(x)$, where p_M is the Taylor polynomial of degree M, and r_M is the remainder term. Clearly p_M depends on τ with typical terms being polynomials in τx^α, where α is a multi-index of length at least 3, and x_j. We write the term in the brackets as a sum of three terms, using that the second term on the left-hand side in Theorem 2.6 is a Gaussian integral evaluated in Lemma 2.1.

$$\int e^{i\tau\phi}\psi(x)(1-\eta(x\tau^s))dm^d(x),$$

$$\int e^{-\frac{\tau}{2}x^T Ax}p_M(x)(1-\eta(x\tau^s))dm^d(x),$$

and

$$\int \eta(x\tau^s)\left[e^{i\tau\phi}\psi(x) - e^{-\frac{\tau}{2}x^T Ax}p_M(x)\right]dm^d(x).$$

Lemma 2.3 and Lemma 2.5 control the first and the second term.

The integrand of the third term is bounded by

$$\tau^{\frac{M}{3}}\tau^{-Ms},$$

and hence the third term is bounded by a constant times

$$\tau^{-ds+M(\frac{1}{3}-s)}.$$

We choose s between $\frac{1}{3}$ and $\frac{1}{2}$ and M large. Finally, we omit the small terms in the sum. $\square$

In the one-dimensional setting the situation of the van der Corput Lemma provides an extremely useful and simple estimate.

Lemma 2.7. *Suppose that $d = 1$, ψ is of bounded variation with support in $[c, d]$, $\phi \in C^k(\mathbb{R})$ with $k \geq 1$, ϕ real, and $\phi^{(k)}(\xi) \geq \tau$ for $\xi \in [c, d]$. If $k = 1$ we assume in addition that ϕ' is monotone. Then*

$$I = \left| \int_c^d \psi(x) e^{i\phi(x)} dx \right| \leq 3k\tau^{-1/k} \int_c^d |\psi'| dx.$$

Proof. We begin with $k = 1$, assuming that ϕ' is monotone. Then

$$\begin{aligned}
\left| \int_c^d \psi e^{i\phi} dx \right| &= \left| \int_c^d \psi/\phi' \frac{d}{dx} e^{i\phi} dx \right| \\
&= \left| \int_c^d e^{i\phi} \frac{d}{dx}(\psi/\phi') dx + e^{i\phi(d)}\psi(d)/\phi'(dx) - e^{i\phi(c)}\psi(c)/\phi'(c) \right| \\
&\leq \left| \int_c^d \frac{\psi'}{\phi'} dx \right| + \sup|\psi| \left(\left| \int_c^d \frac{d}{dx}\frac{1}{\phi'} dx \right| + \frac{2}{\tau} \right) \\
&\leq \frac{1}{\tau} \left(\int_c^d |\psi'| dx + 3 \sup|\psi| \right).
\end{aligned}$$

We use induction on k on the inequality

$$\left| \int_c^d \psi(x) e^{i\phi(x)} dx \right| \leq \frac{1}{\tau^{1/k}} \left(\|\psi'\|_{L^1} + 3k \sup|\psi| \right).$$

Suppose that the estimate holds for $k - 1 \geq 1$ and we want to prove it for k. Suppose that there is point ξ_0 with $\phi^{(k-1)}(\xi_0) = 0$. We decompose the interval $[c, d]$ into $[c, \xi_0 - \delta]$, $[\xi_0 - \delta, \xi_0 + \delta]$, and $[\xi_0 + \delta, d]$. Then, by induction,

$$|I| \leq 2\delta\|\psi\|_{\sup}(\delta\tau)^{-1/(k-1)}(\|\psi'\|_{L^1} + 3(k-1)\|\psi\|_{\sup}).$$

We choose $\delta = \tau^{-\frac{1}{k}}$ to complete the induction. The argument is easier if there is no such point ξ_0. If ψ is supported in $[c, d]$, the fundamental theorem of calculus implies $\|\psi\|_{\sup} \leq \frac{1}{2}|\psi'\|_{L^1}$, which gives the desired inequality. $\square$

2.1 Examples and dispersive estimates

2.1.1 The Schrödinger equation

We consider the linear Schrödinger equation

$$i\partial_t u + \Delta u = 0.$$

A Fourier transform (see Appendix C), which we denote by $\mathcal{F}_x$), gives

$$i\partial_t \mathcal{F}_x u - |\xi|^2 \mathcal{F}_x u = 0$$

and hence the unique solution in the space of tempered distributions is given by its Fourier transform as

$$\mathcal{F}_x u(t,\xi) = e^{-it|\xi|^2} \mathcal{F}_x u(0,\xi).$$

Then

$$\frac{1}{(2\pi)^{d/2}} \int e^{-it|\xi|^2} d\xi = \frac{1}{\sqrt{2it}^d}.$$

Moreover, a change of coordinates shows that

$$\frac{1}{(2\pi)^{d/2}} \int e^{-i(t|\xi|^2 - x\xi)} d\xi = e^{i\frac{x^2}{4t}} \int e^{it\xi^2} d\xi = \frac{1}{\sqrt{2it}^d} e^{i\frac{x^2}{4t}}. \tag{2.4}$$

Again we omit the approximation by a positive definite real part, and the corresponding limit. We obtain the dispersive estimate

$$\|u(t,.)\|_{\sup} \leq \frac{1}{|4\pi t|^{d/2}} \|u(9,.)\|_{L^1}.$$

2.1.2 The Airy function and the Airy equation

We consider the Airy equation

$$u_t + u_{xxx} = 0.$$

The Fourier transform transforms it into

$$\mathcal{F}_x u_t = (ik)^3 \mathcal{F}_x u,$$

and hence, as above,

$$\mathcal{F}_x u(t,\xi) = e^{it\xi^3} \mathcal{F}_x u(0)(\xi).$$

The Airy function is defined by

$$\mathrm{Ai}(x) = \frac{1}{2\pi} \int e^{i\frac{1}{3}\xi^3 + ix\xi} d\xi,$$

where the right-hand side has to be understood (as usual) as

$$\lim_{\varepsilon\to 0}\frac{1}{2\pi}\int e^{i\frac{1}{3}\xi^3-\varepsilon|\xi|^2+ix\xi}d\xi.$$

As above for the quadratic phase function, we see that the limit exists at every point.

The phase function is

$$\phi(\xi)=\frac{1}{3}\xi^3+x\xi$$

and has as critical points the ξ which satisfy

$$\xi^2=-x.$$

If x is negative, there are two real critical points.

We choose $\rho\in C^\infty(\mathbb{R})$, supported in $[-1,\infty)$ and identically 1 in $[1,\infty]$, with $\rho(\xi)+\rho(-\xi)=1$. Then $\mathrm{Ai}(x)$ is the real part of

$$\frac{1}{2\pi}\int \rho(\xi)e^{i(\frac{1}{3}\xi^3+x\xi)}d\xi.$$

There is no harm from the non-compact interval of integration and we to apply the stationary phase, Theorem 2.6, for $x\to-\infty$. The Hessian at the stationary points is $2\tau:=2(-x)^{1/2}$ and we write

$$\phi(\xi)=\tau\phi_0(\xi-(-x)^{1/2}),$$

where

$$\phi_0(\eta)=\frac{1}{3\tau}\eta^3+\frac{1}{2}\eta^2,$$

which satisfies

$$\phi_0'(0)=0,\phi_0''(0)=1,\phi_0'''(0)=2[-x]^{-1/2}.$$

We write the integral as

$$\frac{1}{2\pi}e^{-i\frac{2}{3}|x|^{\frac{3}{2}}}\int\rho(\eta+(-x)^{1/2})e^{i\tau\phi_0(\eta)}d\eta.$$

The application of the stationary phase Theorem 2.6 gives

$$\left|\mathrm{Ai}(x)-\frac{1}{\sqrt{\pi}}|x|^{-1/4}\cos(\frac{2}{3}|x|^{\frac{3}{2}}-\frac{\pi}{4})\right|\le c|x|^{-\frac{7}{4}},$$

and there is even an asymptotic series. To see the error term we compute the next term, the sixth derivative of $e^{i\phi_0(\eta)}$, evaluated at 0. It gives an additional factor $\tau^{-3}=|x|^{-\frac{3}{2}}$.

For large positive x we need a different idea. For positive x there is fast decay and we want to determine the leading term. In this case the two critical points are purely imaginary, and we shift the contour of integration to

$$\xi + i\sqrt{x}.$$

To be more precise we define

$$\mathrm{Ai}_\sigma(x) = \frac{1}{2\pi}\int e^{i[\frac{1}{3}(\xi+i\sigma)^3 + x(\xi+i\sigma)]}d\xi.$$

We expand

$$i\Big[\frac{1}{3}(\xi+i\sigma)^3\Big] + x(\xi+i\sigma) = i\Big(\frac{1}{3}\xi^3 + x\xi - \xi\sigma^2\Big) - \sigma\Big(\xi^2 + x - \frac{1}{3}\sigma^2\Big).$$

We calculate, using the Cauchy–Riemann equations,

$$\frac{d}{d\sigma}\mathrm{Ai}_\tau(x) = \frac{1}{2\pi} = \int i\frac{\partial}{\partial\xi}e^{i(\frac{1}{3}\xi^3+x\xi-\xi\sigma^2)-\sigma(\xi^2+x-\frac{1}{3}\sigma^2)}d\xi = 0,$$

and hence, with $\sigma = \sqrt{x}$,

$$\mathrm{Ai}(x) = \frac{1}{2\pi}\int e^{i\frac{1}{3}\xi^3 - \sqrt{x}\xi^2 - \frac{2}{3}x^{\frac{3}{2}}}d\xi$$

with the critical point $\xi = 0$, at which point the Hessian is $2\sqrt{x}$. We argue as above and obtain

$$\left|\mathrm{Ai}(x) - \frac{1}{2\sqrt{\pi}}|x|^{-1/4}e^{-\frac{2}{3}x^{\frac{3}{2}}}\right| \le c|x|^{-\frac{7}{4}}e^{-\frac{2}{3}x^{\frac{3}{2}}}. \tag{2.5}$$

The van der Corput Lemma ensures that the function Ai is bounded. More is true: About half a derivative of the Airy function is bounded in the following sense:

Lemma 2.8.

$$\left|\int |\xi|^{1/2}e^{i(\frac{1}{3}\xi^3+x\xi)}d\xi\right| \le C.$$

This is left as an exercise.

The Airy function is the inverse Fourier transform of

$$\widehat{\mathrm{Ai}}(\xi) = (2\pi)^{-1/2}e^{i\frac{1}{3}\xi^3}.$$

Clearly

$$(\xi^2 + i\partial_\xi)e^{i\frac{1}{3}\xi^3} = 0,$$

and hence

$$\mathrm{Ai}'' + x\,\mathrm{Ai} = 0.$$

This however implies

$$(\partial_t + \partial^3_{xxx})((t/3)^{-1/3}\operatorname{Ai}(x(t/3)^{-1/3})) = 0$$

and (as oscillatory integral)

$$\int \operatorname{Ai}(x)dx = (2\pi)^{-1/2}.$$

The convolution by the Airy function gives a solution to the initial value problem

$$u_t + u_{xxx} = 0, \qquad u(0,x) = u_0(x),$$

namely

$$u(t,x) = (2\pi)^{1/2}\int (t/3)^{-1/3}\operatorname{Ai}((x-y)(t/3)^{-1/3})u_0(y)dy.$$

Again the equation defines unitary operators $S(t)$ which satisfy

$$\|S(t)u_0\|_{\sup} \le ct^{-1/3}\|u_0\|_{L^1}$$

and, in the sense of Lemma 2.8,

$$\||D|^{\frac{1}{2}}S(t)u_0\|_{\sup} \le ct^{-\frac{1}{2}}\|u_0\|_{L^1}. \tag{2.6}$$

2.1.3 Laplacian and related operators

Let $d > 2$. Then, by Lemma 9.5,

$$\widehat{|x|^{2-d}} = \frac{1}{2^{(d-4)/2}\Gamma(\frac{d-2}{2})}|\xi|^{-2}$$

and

$$-\Delta\frac{(4\pi)^{d/2}}{\Gamma(\frac{d-2}{2})}\int |x-y|^{2-d}f(y)dy = f(y).$$

The Fourier transform transforms higher partial derivatives into multiplication by monomial functions. For example

$$\mathcal{F}(u-\Delta u) = (1+|\xi|^2)\hat{u}$$

and hence

$$\hat{u} = (1+|\xi|^2)^{-1}\hat{f}$$

is the Fourier transform of a Schwartz function u (if f is a Schwartz function) which satisfies

$$-\Delta u + u = f.$$

Here $(1+|\xi|^2)^{-1}$ is a smooth function with bounded derivatives, but not a Schwartz function. Its inverse Fourier transform k allows to define a solution for a given function f by

$$u = (2\pi)^{d/2} k * f.$$

We compute k in one space dimension:

$$\int_{-\infty}^{\infty} e^{ix\xi}(1+\xi^2)^{-1} d\xi = \pi e^{-|x|} \tag{2.7}$$

using the residue theorem: The singular points are the zeroes of the polynomial $1+\xi^2$, which are $\pm i$. Consider the case $x > 0$ first. By the residue theorem

$$\int_{C_R} e^{ix\xi}(1+\xi^2)^{-1} d\vec{\xi} = \pi e^{-|x|},$$

where C_R is the union of the path from $-R$ to R and the upper semi-circle. The limit $R \to \infty$ implies the statement.

2.1.4 Gaussians, heat and Schrödinger equation

Lemma 2.9. *Let $A = A_0 + iA_1$ be an invertible symmetric matrix (A_0 and A_1 real) with A_0 positive semi-definite. Then*

$$\mathcal{F} e^{-\frac{1}{2}x^T A x}(\xi) = \det(A)^{-1/2} e^{-\frac{1}{2}\xi^T A^{-1}\xi}.$$

Proof. The formula is correct at $\xi = 0$ by Lemma 2.1 . We assume first that A_0 is positive definite. The general statement follows then by continuity of both sides. By definition,

$$\nabla e^{-\frac{1}{2}x^T A x} + e^{-\frac{1}{2}x^T A x} A x = 0.$$

The Fourier transform g is a Schwartz function which then satisfies

$$g\xi + A\nabla g = 0.$$

This is an ordinary differential equation on lines through the origin. There is a unique solution with the given value at $\xi = 0$, which has to coincide with the function on the right-hand side. □

With $A = 2t1_{\mathbb{R}^d}$ we obtain the formula for the fundamental solution to the heat equation. The inverse Fourier transform of $e^{-it|\xi|^2}$ is, as computed twice,

$$(2it)^{d/2} e^{-\frac{|x|^2}{4it}}$$

The solution to the Schrödinger equation

$$iu_t + \Delta u = 0$$

with initial data u_0 is given by

$$u(t,x) = \int_{\mathbb{R}^d} (4i\pi t)^{-d/2} e^{-\frac{|x-y|^2}{4it}} u_0(y)dy. \tag{2.8}$$

We denote the map $u(0,\cdot) \to u(t,\cdot)$ by $S(t)$. It is defined via the Fourier transform by

$$\widehat{S(t)u_0} = e^{-it|\xi|^2}\hat{u}_0(\xi).$$

It is a unitary operator:

$$\|S(t)u_0\|_{L^2} = \|\widehat{S(t)u_0}\|_{L^2} = \|e^{-it|\xi|^2}\hat{u}_0\|_{L^2} = \|\hat{u}_0\|_{L^2} = \|u_0\|_{L^2},$$

and it satisfies the so-called dispersive estimate

$$\|S(t)u_0\|_{\sup} \leq |4\pi t|^{-d/2}\|u_0\|_{L^1}.$$

2.1.5 The half-wave equation

The solution to the wave equation

$$u_{tt} - \Delta u = 0$$

with initial data

$$u(0,x) = u_0(x), \qquad u_t(0,x) = u_1(x)$$

is given for $d = 3$ by Kirchhoff's formula:

$$u(t,x) = \frac{1}{4\pi t^2}\int_{\partial B_t(x)} u_0 d\mathcal{H}^2 + \frac{1}{4\pi t}\int_{\partial B_t(x)} \partial_\nu u_0 d\mathcal{H}^2 + \frac{1}{4\pi t}\int_{\partial B_t(x)} u_1 d\mathcal{H}^2.$$

There are similar formulas in odd dimensions, and slightly more complicated ones in even dimensions.

The Fourier transform transforms the PDE to the ODE

$$\hat{u}_{tt} + |\xi|^2\hat{u} = 0$$

which factorizes as

$$(\partial_t - i|\xi|)(\partial_t + i|\xi|)\hat{u} = 0.$$

This motivated the study of the half-wave equation

$$(i\partial_t + |\xi|)\hat{u}(t,\xi) = 0,$$

which can easily be solved in the form

$$\hat{u}(t,\xi) = e^{it|\xi|}\hat{u}(0,\xi).$$

As above, we restrict to $t = 1$. Since $e^{it|\xi|}$ is radial,

$$\int e^{i(|\xi|+x\xi)} d\xi = dm^d(B_1(0))|x|^{-\frac{d-2}{2}} \int_0^\infty r^{d/2} e^{ir} J_{\frac{d-2}{2}}(|x|r)dr$$

provided the integrals exist as oscillatory integrals. They do, as we will see. By Lemma 8.1, we can write

$$z^{\frac{d-1}{2}} J(z) = \mathrm{Re}(e^{-iz}\phi(z))$$

for $z \geq 1$, with ϕ satisfying

$$|\phi^{(k)}(z)| \leq c_k z^{-k}.$$

We begin by considering the case $|x| \geq 2$. We decompose the integral above into two parts with a smooth cutoff function, one over $r \geq |x|^{-1}$, and one over $2|x|^{-1}$. In the first integral we can integrate by parts as often as we like:

$$\int_0^\infty (1-\eta(r|x|))e^{ir(1\pm|x|)} p(rx)dx = \frac{i}{1 \pm x} \int_0^\infty e^{ir(1\pm|x|)} (\frac{d}{dr}((1-\eta(r|x|))p(rx))dx$$

which gains a factor r in the integration, as well as a power $|x|^{-1}$. We repeat this as often as necessary. The second integral is bounded by $|x|^d$.

The same arguments apply as for $|x| \neq 1$, given bounds which depend only on $|x| - 1$. A careful calculation gives the first part of the following estimate.

Lemma 2.10. *There exist $c_d > 0$ and $c \in \mathbb{R}$ such that for $|x| \neq 1$*

$$\left| \int e^{i|\xi|+ix\xi} d\xi \right| \leq \begin{cases} c_d |1-|x||^{-\frac{d+1}{2}}, & \text{if } |x| \leq 2, \\ c_d |x|^{-d}, & \text{if } |x| \geq 2 \text{ and } d \text{ even,} \end{cases}$$

and

$$\left| \int |\xi|^{-\frac{d+1}{2}} e^{i|\xi|+ix\xi} d\xi - c \ln|1-|x|| \right| \leq c_d,$$

if $|x| \leq 2$ and

$$\left| \int |\xi|^{-\frac{d+1}{2}} e^{i|\xi|+ix\xi} d\xi \right| \leq c_d |x|^{-\frac{d-1}{2}},$$

for $|x| \geq 2$.

Proof. Only the second part remains to be shown. There is no difference in the argument for $|x| \leq 2$, unless $|x|$ is close to 1. In that case we decompose the integral into $r \leq 2$, $1 \leq r \leq |x| - 1$, and $r \geq |x| - 1$. The last part is bounded by the previous arguments. The first part is bounded because of the size $r \leq 1$. The second part is

$$\int_1^{||x|-1|} r^{-1} dr = \ln r$$

plus something bounded. $\square$

There is an important difference compared to the previous two examples: the group velocity $\nabla|\xi|$ depends only on the direction of ξ, not on the amplitude.

2.1.6 The Klein-Gordon half wave

Let

$$g(t,x) = \int e^{it\sqrt{1+|\xi|^2}+ix\xi}d\xi.$$

We report the analogue of Lemma 2.10. There is almost no change in the argument.

Lemma 2.11. *The following estimates hold for $t \geq 1$,*

$$|g(t,x)| \leq c \begin{cases} t^{-d/2}(1-|x|/t)^{-\frac{d+1}{2}}, & \text{if } |x| < t, \\ t^{-d}(|x|/t-1)^{-\frac{d+1}{2}}, & \text{if } t < |x| \leq 2t, \\ \frac{1}{|x|^d t^{d-1}}, & \text{if } |x| \geq 2t, \end{cases}$$

and if $0 < t < 1$,

$$|g(t,x)| \leq c \begin{cases} t^{-d}, & \text{if } |x| < t, \\ t^{-d}(|x|/t-1)^{-\frac{d+1}{2}}, & \text{if } t < |x| \leq 2t, \\ \frac{1}{|x|^d t^{d-1}}, & \text{if } |x| \geq 2t. \end{cases}$$

Moreover,

$$h = \int |\xi|^{-\frac{d+1}{2}} e^{it\sqrt{1+|\xi|^2}+ix\xi}d\xi$$

satisfies for $t \geq 1$ and $|x| \geq 2t$,

$$|h(t,x)| \leq C\frac{1}{|x|^{\frac{d-1}{2}} t^{\frac{d}{2}-3}},$$

$$\left| h(t,x) - ct^{\frac{1}{2}} |\ln ||1-|x|/t|| \right| \leq ct^{\frac{1}{2}},$$

for $1 \leq t$, $|x| \leq 2t$. Finally, if $0 < t \leq 1$, then

$$\left| \int |\xi|^{-\frac{d+1}{2}} e^{it|\xi|+ix\xi}d\xi - ct^{-\frac{d-1}{2}} \ln |1-|x|| \right| \leq c_d t^{-\frac{d-1}{2}},$$

if $|x| \leq 2t$, and

$$\left| \int |\xi|^{-\frac{d+1}{2}} e^{i|\xi|+ix\xi}d\xi \right| \leq c_d \frac{1}{|x|^{\frac{d-1}{2}} t^{\frac{d}{2}-3}},$$

for $|x| \geq 2$.

2.1.7 The Kadomtsev–Petviashvili equation

The linear parts of the Kadomtsev–Petviashvili equations are

$$u_t + u_{xxx} \pm \partial_x^{-1} u_{yy} = 0,$$

where $+$ is the linear KP II equation and $-$ the linear KP I equation. The equation should be understood as

$$\partial_x u_t + u_{xxxx} \pm u_{yy} = 0.$$

We denote the Fourier variables by ξ (of x) and η (of y). As above (for $+$, the argument for $-$ is very similar),

$$\mathcal{F}_{x,y} u(t,\xi,\eta) = e^{it(\xi^3 - \xi^{-1}\eta^2)} \mathcal{F}_{x,y} u(0,\xi,\eta)$$

and

$$\int e^{i[(\xi^3 - \xi^{-1}\eta^2) + x\xi + y\eta]} d\xi d\eta = (4\pi)^{-1/2} \int (-i\xi)^{\frac{1}{2}} e^{i[\xi^3 + \xi x + \xi y^2/4]} d\xi.$$

The stationary points of the phase function satisfy

$$3\xi^2 + x + y^2/4 = 0,$$

with zeroes

$$\xi = \pm\sqrt{-(x + y^2/4)/3},$$

provided

$$x < -\frac{1}{4} y^2.$$

The contribution from the Hessian compensates the factor $(-i\xi)^{\frac{1}{2}}$. A rigorous proof uses a smooth partition of unity, which decomposes the integral into one around $\xi = 0$, one over $\xi \geq 1$, and one with $\xi \leq -1$. The first integral is handled by the van der Corput Lemma, and the other two by stationary phase.

Otherwise, by the non-degeneracy of the phase

$$\left| \int e^{i[(\xi^3 - \xi^{-1}\eta^2) + x\xi + y\eta]} d\xi d\eta \right| \leq c_k |x + y^2/4|^{-k}.$$

The t dependence below is obtained by scaling.

Lemma 2.12.

$$\left| \int e^{it(\xi^3 \mp \eta^2/\xi) + ix\xi} d\xi d\eta \right| \leq c_k |t|^{-1} \Big(1 + \Big(\frac{x}{t^{\frac{1}{3}}} \pm \frac{y^2}{t^{\frac{2}{3}}}\Big)_+\Big)^{-k}.$$

There is an interesting interpretation:

- waves move to left for Kadomtsev–Petviashvili II,

- and to both sides for Kadomtsev–Petviashvili I (with respect to x)

This makes the study of Kadomtsev–Petviashvili I considerably harder than the study of Kadomtsev–Petviashvili II.

We define

$$\rho(x,y) = 2\pi\mathcal{F}^{-1}(e^{i(\xi^3-\eta^2/\xi)}).$$

Since $u(\lambda^3 t, \lambda x, \lambda^2 y)$ satisfies the linear KP equation for $\lambda > 0$ if and only if u does, we obtain the representation

$$u(t,x,y) = g_t * u(0,\cdot,\cdot)(x,y)$$

where

$$g_t(x,y) = t^{-1}\rho(x/t^{1/3}, y/t^{2/3}).$$

Hence, with $S(t)$ denoting the evolution operator,

$$\|S(t)u_0\|_{L^2} = \|u_0\|_{L^2}$$

and

$$\|S(t)u_0\|_{\sup} \leq c|t|^{-1}\|u_0\|_{L^1(\mathbb{R}^2)}.$$

Chapter 3

Strichartz estimates and small data for the nonlinear Schrödinger equation

3.1 Strichartz estimates for the Schrödinger equation

We return to the linear Schrödinger equation

$$i\partial_t u + \Delta u = 0$$

and the unitary operators $S(t) : u(0) \to u(t)$. They form a group: for $s, t \in \mathbb{R}$,

$$S(t+s) = S(t)S(s).$$

We claim that for $2 \le p \le \infty$ and p' with $\frac{1}{p} + \frac{1}{p'} = 1$

$$\|S(t)\|_{L^p} \le (4\pi|t|)^{-\frac{d}{2}(1-\frac{2}{p})} \|u_0\|_{L^{p'}}, \tag{3.1}$$

which follows by complex interpolation from

$$\|S(t)u_0\|_{L^2} = \|u_0\|_{L^2}$$

and the dispersive estimate

$$\|S(t)u_0\|_{L^\infty} \le (4\pi|t|)^{-\frac{d}{2}} \|u_0\|_{L^1}.$$

Let us be more precise. We put $p_0 = q_0 = 2$ and $p_1 = 1$, $q_1 = \infty$, $2 < \tilde{p} < \infty$ and determine λ so that

$$\frac{1-\lambda}{2} = \frac{1}{p},$$

i.e.,

$$\lambda = 1 - \frac{2}{p}.$$

H. Koch et al., *Dispersive Equations and Nonlinear Waves: Generalized Korteweg–de Vries, Nonlinear Schrödinger, Wave and Schrödinger Maps*, Oberwolfach Seminars 45,
DOI 10.1007/978-3-0348-0736-4_3, © Springer Basel 2014

Define
$$\frac{1-\lambda}{2}+\lambda=\frac{1}{q}.$$
We check easily
$$\frac{1}{p}+\frac{1}{q}=1,$$
and now (3.1) follows by the Riesz–Thorin interpolation Theorem 7.6.

The variation of constants formula resp. Duhamel's formula
$$u(t)=-i\int_{-\infty}^{t}S(t-s)f(s)ds$$
defines a solution to
$$i\partial_t u+\Delta u=f,$$
at least for Schwartz functions f in $d+1$ variables.

From the $L^{p'}$ to L^p estimate (3.1) one obtains
$$\|u(t)\|_{L^p}\le(4\pi)^{-\frac{d}{2}(1-\frac{2}{p})}\int_{-\infty}^{t}|t-s|^{\frac{d}{2}-\frac{d}{p}}\|f(s)\|_{L^{p'}}ds.$$
The right-hand side is a convolution $h*g$ where
$$h(t)=\begin{cases} 0, & \text{if } t\ge 0,\\ |4\pi t|^{-d(\frac{1}{2}-\frac{1}{p})}, & \text{if } t<0,\end{cases}$$
and
$$g(t)=\|f(t)\|_{L^{p'}(\mathbb{R}^d)}.$$
An immediate calculation gives $|t|^{-1/r}\in L^r_w(\mathbb{R})$, and by the weak Young inequality of Proposition 7.2
$$\|g*h\|_{L^q(\mathbb{R})}\le c\|g\|_{L^{q'}}\|h\|_{L^r_w},\tag{3.2}$$
where
$$\frac{1}{r}=d\Big(\frac{1}{2}-\frac{1}{p}\Big),\quad r>1,$$
and (p,q) are strict Strichartz pairs, i.e., numbers which satisfy
$$\frac{2}{q}+\frac{d}{p}=\frac{d}{2},\tag{3.3}$$
and $2<q\le\infty, 2\le p\le\infty$. The left-hand side of (3.2) controls
$$\|u\|_{L^q_tL^p_x}:=\Big(\int\|u(t)\|^q_{L^p(\mathbb{R}^d)}dt\Big)^{1/q}$$

with the obvious modification if $q = \infty$, and we obtain

$$\|u\|_{L^q_t L^p_x} \le c\|f\|_{L^{q'}_t L^{p'}_x}$$

for all strict Strichartz pairs. Here $L^q_t L^p_x$ consists of all equivalence classes of measurable functions such that the integral expression for the norm is finite.

It is not hard to see that u measurable implies

$$t \mapsto \|u(t,\cdot)\|_{L^p}$$

is measurable, the expression for the norm actually defines a norm, and the space is closed and hence a Banach space. The duality of the Lebesgue spaces extends to duality of this mixed norm spaces: The map

$$L^{q'}_t L^{p'}_x \ni f \mapsto (g \mapsto \int fg dm^d dt) \in (L^q_t L^p_x)^*$$

is an isometry if $1 \le p, q \le \infty$ and surjective if $p, q < \infty$. Complex interpolation extends to the mixed norm spaces – this is quite evident from the definition.

We claim

Theorem 3.1. *The variation of constants formula defines a function u which satisfies*

$$i\partial_t u + \Delta u = f, \qquad u(0) = u_0.$$

Let (q,p) be a strict Strichartz pair. Then

$$\|u\|_{C_b(\mathbb{R},L^2)} + \|u\|_{L^q_t L^p_x} \le c\Big(\|u(0)\|_{L^2} + \|f\|_{L^{q'}_t L^{p'}_x}\Big).$$

We will later improve this statement in several directions. Denote by T,

$$L^2 \ni v \mapsto Tv \in C([0,\infty), L^2),$$

the operator that maps the initial datum to the solution. Let (p,q) be Strichartz pairs. Then, with $L(X,Y)$ denoting the bounded linear operators from Banach space X to Banach space Y,

$$\|T\|^2_{L(L^2, L^{q'}_t L^{p'}_x)} = \|T^*\|^2_{L(L^q_t L^p_x, L^2)} = \|TT^*\|_{L(L^q_t L^p_x, L^{q'}_t L^{p'}_x)}$$

and

$$TT^* f(t) = \int_0^\infty S(t+s) f(s) ds = \int_{-\infty}^0 S(t-s) f(-s) ds$$

and the bound follows as above.

3.2 Strichartz estimates for the Airy equation

This section follows Kenig, Ponce, and Vega [16]. Scaling shows that the solution to the Airy equation satisfies

$$u(t,x) = \frac{1}{(t/3)^{1/3}} \int \mathrm{Ai}((x-y)/(t/3)^{\frac{1}{3}})u(0,y)dy$$

and we obtain the estimates

$$\|u(t)\|_{L^2} = \|u_0\|_{L^2},$$

$$\|u(t)\|_{L^\infty} \le ct^{-1/3}\|u_0\|_{L^1}$$

and

$$\||D|^{\frac{1}{2}}u(t)\|_{L^\infty} \le ct^{-\frac{1}{2}}\|u_0\|_{L^1}.$$

The Strichartz estimate is more complicated. Here we use complex interpolation to see for $2 < p \le \infty$ that

$$\|D^{\frac{1}{2}-\frac{1}{p}}S(t)v\|_{L^p} \le c|t|^{\frac{1}{p}-\frac{1}{2}}\|v\|_{L^{p'}}, \tag{3.4}$$

where D^s is defined through the Fourier multiplier. The multiplication on the Fourier side commutes with the evolution, and hence this estimates is equivalent to

$$\|D^{\frac{1}{q}}S(t)v\|_{L^p(\mathbb{R})} \le c|t|^{-\frac{2}{q}}\|D^{-\frac{1}{q}}v\|_{L^{p'}}.$$

The Strichartz estimates take the following form.

Theorem 3.2. *The variation of constants formula defines a function u which satisfies*

$$\partial_t u + u_{xxx} = f, \qquad u(0) = u_0$$

and

$$\|u\|_{C_b(\mathbb{R},L^2)} + \||D|^{\frac{1}{q}}u\|_{L^q_t L^p_x} \le c\left(\|u(0)\|_{L^2} + \||D|^{-\frac{1}{q}}f\|_{L^{q'}_t L^{p'}_x}\right)$$

for all Strichartz pairs (q,p).

Proof. It remains to prove (3.4).

We claim that it follows from

$$\left|\int |\xi|^{\frac{1}{2}+i\sigma}e^{i\xi^3+i\xi x}d\xi\right| \le C(1+|\sigma|) \tag{3.5}$$

uniformly in x. The integral has to be understood as oscillatory integral. We apply then complex interpolation with the family of operators

$$\widehat{T_\lambda u_0} = e^{\lambda^2}|D|^{\frac{\lambda}{2}}\widehat{S(t)u_0},$$

for which we easily see that

$$\|T_{i\sigma}u_0\|_{L^2} = e^{-\sigma^2}\|u_0\|_{L^2}$$

and

$$\|T_{i+\sigma}u_0\|_{L^\infty} \leq ct^{-1/2}(1+|\sigma|)e^{-\sigma^2}\|u_0\|_{L^1}.$$

Now (3.4) follows by complex interpolation. We next turn to (3.5).

There are three cases: $|x| \leq 10$, $x \geq 10$, and $x \leq -10$. The last one is the hardest, since there are large critical points $\pm\xi_c = \sqrt{-x/3}$ in the phase, and we restrict to it. We split the integration domain into the intervals

$$(-\infty, -\xi_c - |x|^{-1/4}), (-\xi_c - |x|^{-1/4}, \xi_c + |x|^{-1/4}), (-\xi_c + |x|^{-1/4}, -1), (-1,1),$$
$$(1, \xi_c - |x|^{-1/4}), (\xi_c - |x|^{-1/4}, \xi_c + |x|^{-1/4}), (\xi_c + |x|^{-1/4}, \infty).$$

The argument is immediate for the second, the fourth and the sixth integral, which we estimate by $3\xi_c^{1/2}|x|^{-1/4}$. Now

$$\int_{-\infty}^{-\xi_c-|x|^{-1/4}} |\xi|^{\frac{1}{2}+i\sigma} e^{i\xi^3+ix\xi} d\xi = i\int_{-\infty}^{-\xi_c-|x|^{-1/4}} e^{i\xi^3+ix\xi}\frac{d}{d\xi}\frac{|\xi|^{\frac{1}{2}+i\sigma}}{3\xi^2+x}d\xi$$
$$+\frac{(\xi_c+|x|^{-1/4})^{\frac{1}{2}+i\sigma}}{3(\xi_c+|x|^{-1/4})^2+x}e^{-i(\xi_c+|x|^{-1/4})^3-i(\xi_c+|x|^{-1/4})x}$$

and the direct estimate as for stationary phase gives the result. The largest term (in terms of σ) occurs when the derivative falls on $|\xi|^{\frac{1}{2}+i\sigma}$; all the others are estimated as when $\sigma = 0$. We recall that

$$3(\xi_c + |x|^{-1/4})^2 + x \sim |x|^{\frac{1}{4}}.$$

□

3.3 The Kadomtsev–Petviashvili equation

The symbol is $\xi^3 - \eta^2/\xi$ (for KP II, with similar arguments for KP I), with gradient

$$\begin{pmatrix} 3\xi^2 + \eta^2/\xi^2 \\ -2\eta/\xi \end{pmatrix},$$

Hessian matrix

$$\begin{pmatrix} 6\xi - 2\eta^2/\xi^3 & 2\eta/\xi^2 \\ 2\eta/\xi^2 & -2/\xi \end{pmatrix},$$

and Hessian determinant -12.

Lemma 3.3. *The following Strichartz estimate holds:*

$$\|u\|_{L_t^\infty L_x^2} + \|u\|_{L_t^p L_x^q} \leq c\left(\|u_0\|_{L^2} + \|f\|_{L_t^{p'}L_x^{q'}}\right).$$

The proof is the same (since the same dispersive estimate holds) as for the Schrödinger equation.

3.4 The (half-) wave equation and the Klein–Gordon equation

Here we only state the result. The proof requires a sharpening of complex interpolation, replacing L^∞ by BMO. The estimates for the wave equation (for $\sigma = 0$, with similar estimates for $\sigma \in \mathbb{R}$, compare the Airy equation) imply that

$$\| |D|^{-\frac{d+1}{2}+i\sigma} \|_{BMO} \le c(1+|\sigma|)|t|^{\frac{d-1}{2}} \|v\|_{L^1(\mathbb{R}^d)},$$

which in turn yields

$$\| |D|^{-\frac{d+1}{2}(1-\frac{2}{p})} S(t) v\|_{L^p} \le c t^{-\frac{d-1}{2}(1-\frac{2}{p})} \|v\|_{L^{p'}},$$

where the half-wave evolution operator $S(t)$ is defined by

$$S(t)v = \mathcal{F}^{-1}(e^{it|\xi|}\hat{v}).$$

As a consequence we obtain

Theorem 3.4. *Let $d \ge 2$. The variation of constants formula defines a function u which satisfies*

$$i\partial_t u + |D|u = f, \qquad u(0) = u_0,$$

and

$$\|u\|_{C_b(\mathbb{R},L^2)} + \| |D|^{-\frac{d+1}{4}(1-\frac{2}{p})} u\|_{L^q_t L^p_x} \le c\|u(0)\|_{L^2} + \| |D|^{\frac{d+1}{4}(1-\frac{2}{p})} f\|_{L^{q'}_t L^{p'}_x},$$

where q satisfies $2 < q < \infty$, $2 \le p \le \infty$, and

$$\frac{1}{q} + \frac{d-1}{p} = \frac{d-1}{2}.$$

3.5 The endpoint Strichartz estimate

Here we will prove the endpoint Strichartz estimate for the Schrödinger equation

$$iu_t + \Delta u = f, \qquad u(0) = u_0$$

for $d \ge 3$. The argument is due to Keel and Tao [15] and applies to much more general situations.

Theorem 3.5. *The solution defined by the variation of constants formula satisfies*

$$\|u\|_{L^\infty_t L^2_x} + \|u\|_{L^2_t L_x^{\frac{2d}{d-2}}} \le c\left(\|u_0\|_{L^2} + \|f\|_{L^2_t L_x^{\frac{2d}{d+2}}} \right). \tag{3.6}$$

Before we prove the statement we need a robust estimate for integral operators.

Lemma 3.6 (Schur's lemma)**.** *Let μ and ν be measures, and*

$$Tf(x) = \int K(x,y)f(y)d\mu(y),$$

where K satisfies

$$\sup_x \int |K(x,y)|d\mu(y) \le C_x, \quad \sup_y \int |K(x,y))d\nu(x) \le C_y.$$

Then

$$\|Tf\|_{L^p(\nu)} \le C_x^{1-\frac{1}{p}} C_y^{\frac{1}{p}} \|f\|_{L^p(\mu)}.$$

Proof. By duality, the claim is equivalent to

$$\left|\int f(x)g(y)K(x,y)d\mu(y)d\nu(x)\right| \le C_x^{1-\frac{1}{p}} C_y^{\frac{1}{p}} \|f\|_{L^{p'}(\mu)}\|g\|_{L^p(\mu)}.$$

This is obvious for $p = \infty$ and $p = 1$. Hence the operator satisfies the desired bounds on L^1 and L^∞. The general claim follows by complex interpolation. □

Proof. We denote by $S(t)$ the Schrödinger group. We first prove

$$\left|\int_{s<t} \langle S(-t)f(t), S(-s)g(s)\rangle\right| \le c\|f\|_{L_t^2 L_x^{\frac{2d}{d-2}}} \|g\|_{L_t^2 L_x^{\frac{2d}{d-2}}} \tag{3.7}$$

which by duality implies

$$\left\|\int_{-\infty}^t S(t-s)f(s)\right\|_{L_t^2 L_x^{\frac{2d}{d-2}}} \le c\|f\|_{L_t^2 L_x^{\frac{2d}{d+2}}}$$

and, by the TT^* argument, the full statement.

We define

$$T_j = \int_{t-2^{j+1}<s\le t-2^j} \langle S(-s)f(s), S(-t)g(t)\rangle dsdt$$

and claim that

$$|T_j| \le C2^{-j\beta(p,\tilde p)}\|f\|_{L_t^2 L_x^{p'}} \|g\|_{L_t^2 L_x^{\tilde p'}} \tag{3.8}$$

for $j \in \mathbb{Z}$, p and $\tilde p$ in a neighborhood of $\frac{2d}{d+2}$, and

$$\beta(p,\tilde p) = 1 - \frac{d}{2} + \frac{d}{2p} + \frac{d}{2\tilde p};$$

$\beta(p,\tilde p)$ vanishes for $p = \tilde p = \frac{2d}{d-2}$, as it should.

We set $\tilde t = t2^{-j}$, $\tilde s = s2^{-j}$, $\tilde x = 2^{-j/2}x$, and $\tilde y = 2^{-j/2}y$. This transformation of coordinates (which reflects the symmetry) reduces the estimate to the case $j = 0$.

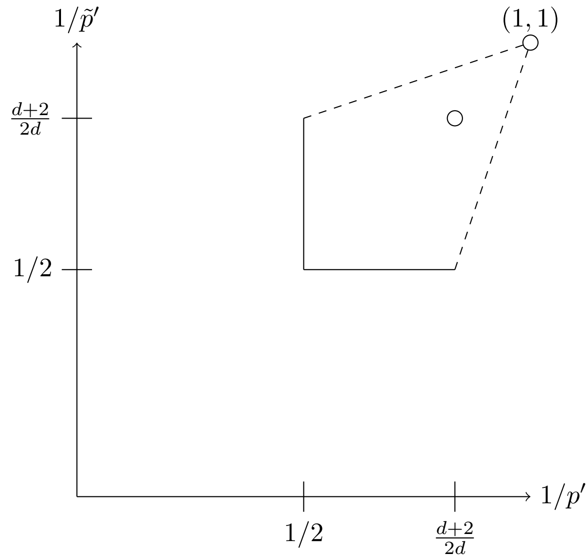
The estimate for $j=0$

$$|T_0| \leq C\|f\|_{L^2_t L^{p'}_x} \|g\|_{L^2_t L^{\tilde{p}'}_x} \tag{3.9}$$

holds for

1. $p = \tilde{p} = 1$ by the dispersive estimate,
2. $\tilde{p} = 2$ and $\frac{2d}{d+2} < p' \leq 2$,
3. $p = 2$ and $\frac{2d}{d+2} \leq \tilde{p}' \leq 2$.

Then the estimate (3.8) follows by complex interpolation and duality. It is convenient to draw a diagram:



Convex interpolation, this time for $L^{2,p'}$ spaces, gives the convex envelope, which contains the point $(\frac{d+2}{2d}, \frac{d+2}{2d})$ in its interior.

For the first case (which corresponds to $(1,1)$) observe that by the dispersive estimate, if $t-2 < s < t-1$, then

$$|\langle S(t-s)g(s), f(t)| \leq C\|f(t)\|_{L^1}\|g(s)\|_{L^1}.$$

Let $h_f(t) = \|f(t)\|_{L^1}$ and $h_g(t) = \|g(t)\|_{L^1}$. Then

$$|T_0(f,g)| \leq C \int\int K(t,s) h_g(s) ds h_f(t) dt,$$

where $K(t-s) = 1$ if $t-2 < s < t-1$, and 0 otherwise. The first estimate follows by Schur's lemma.

For the second estimate (which corresponds to the horizontal line) we use non-endpoint Strichartz estimate and finally Hölder's inequality to bound

$$\begin{aligned}\left|\int_{s+1}^{s+2}\langle f(t), S(t-s)g(s)\rangle dt\right| &\le \|f\|_{L_t^{q'}L_x^{p'}([s+1,s+2]\times\mathbb{R}^d)}\|S(t-s)g(s)\|_{L_t^qL_x^p}\\ &\le C\|f\|_{L_t^2L_x^{p'}([s+1,s+2]\times\mathbb{R}^d)}\|g(s)\|_{L^2},\end{aligned}$$

where (q,p) is a strict Strichartz pair.

Thus

$$\left|\int_k^{k+1}\int_{t-2}^{t-1}\langle S(-t)f(t), S(-s)g(s)\rangle dsdt\right| \le c\|f\|_{L_t^2L_x^{p'}([k,k+1]\times\mathbb{R}^d)}\|g\|_{L_t^2L_x^2([k-2,k]\times\mathbb{R}^d)}.$$

The statement follows by summation with respect to k, and the Cauchy–Schwarz inequality with respect to k.

The third estimate follows by the same argument. This completes the estimate (3.8) for $(p,\tilde p)$ close to $(\frac{2d}{d-2},\frac{2d}{d-2})$.

To make use of the flexibility we have, we decompose $f=\sum f_k$, $g=\sum g_k$ with

$$f_k(t,x)=c_k(t)\chi_{t,k}(x), \quad g_k(t,x)=d_k(t)\tilde\chi_{t,k}(x).$$

We define the decomposition as follows. Given $f:\mathbb{R}^d\to\mathbb{R}$, we define its distribution function for $s>0$

$$\lambda(s)=m^d\{x:|f(x)|>s\}.$$

It is monotonically decreasing and finite for $f\in L^p$. Let s_k be the infimum of all s such that $\lambda(s)<2^k$ (we allow $s=0$). We set $c_k=2^{k/p}s_k$ and

$$\chi_k(x)=c_k^{-1}\begin{cases} f, & \text{if } s_k<|f|<s_{k+1},\\ 0, & \text{otherwise .}\end{cases}$$

Then

$$f=\sum c_k\chi_k$$

and, for some $C>0$,

$$C^{-1}\|f\|_{L^p}\le\|(c_k)\|_{l^p}\le C\|f\|_{L^p}$$

which can be seen by comparing to

$$\|f\|_{L^p}^p=p\int m^d(\{|f|>s\})s^{p-1}ds.$$

By definition

$$m^d(\text{ supp }\chi_k)\le 2^k, \qquad |\chi_k|\le 2^{k/p}.$$

We apply this decomposition at every time t with $p=\frac{2d}{d+2}$. Then

$$f=\sum f_k,$$

where at most one is not 0.

We apply the first estimate (3.8):

$$
\begin{aligned}
|T_j(f_k, g_{k'})| &\le c2^{-\beta(p,\tilde p)} \|f_k\|_{L^2_t L^{p'}_x} \|g_{k'}\|_{L^2_t L^{\tilde p'}_x} \\
&\le c2^{-j(\frac{2-d}{2}+\frac{d}{2p}+\frac{d}{2\tilde p})-k(\frac{2-d}{2d}+\frac{1}{p})-k'(\frac{2-d}{2d}+\frac{1}{\tilde p})} \|f_k\|_{L^2_t L_x^{\frac{2d}{d+2}}} \|g_{k'}\|_{L^2_t L_x^{\frac{2d}{d+2}}} \\
&\le c2^{-(\frac{j}{2}-\frac{k}{d})(2-d+\frac{d}{p})-(\frac{j}{2}-\frac{k'}{d})(2-d+\frac{d}{\tilde p})} \|f_k\|_{L^2_t L_x^{\frac{2d}{d+2}}} \|g_{k'}\|_{L^2_t L_x^{\frac{2d}{d+2}}},
\end{aligned}
$$

where the second inequality follows from

$$
\|\chi_{t,k}\|_{L^{p'}} \le c2^{-k(\frac{1}{p'}-\frac{2d}{d+2})}.
$$

Given k, $\tilde k$ and j we choose $(p,\tilde p)$ so that the factor becomes (almost) minimal. Then there exists $\varepsilon > 0$ so that

$$
|T_j(f_k, g_{k'})| \le c2^{-\varepsilon(|\frac{jd}{2}-k|+|\frac{jd}{2}-k'|)} \|f_k\|_{L^2_t L_x^{\frac{2d}{d+2}}} \|g_{k'}\|_{L^2_t L_x^{\frac{2d}{d+2}}}.
$$

We sum with respect to j.

$$
\begin{aligned}
\sum_j |T_j| &\le C \sum_k \sum_{k'} (1+|k-k'|) 2^{-\varepsilon|k-k'|} \|f_k\|_{L^2_t L_x^{\frac{2d}{d+2}}} \|g_k\|_{L^2_t L^{\frac{2}{d}} d+2_x}. \\
&\le C \left(\sum_k \|f_k\|^2_{L^2_t L_x^{\frac{2d}{d+2}}} \right)^{1/2} \left(\sum_k \|g_k\|^2_{L^2_t L_x^{\frac{2d}{d+2}}} \right)^{1/2},
\end{aligned}
$$

by Schur's lemma. By Minkowski's inequality

$$
\begin{aligned}
\sum_k \int \left(\int_{\mathbb{R}^d} |g_k|^{\frac{2d}{d+2}} dm^d \right)^{\frac{d+2}{d}} dt &= \int \sum_k \left(\int_{\mathbb{R}^d} |g_k|^{\frac{2d}{d+2}} dm^d \right)^{\frac{d+2}{d}} dt \\
&\le \int \left(\int_{\mathbb{R}^d} \sum_k |g_k|^{\frac{2d}{d+2}} dm^d \right)^{\frac{d+2}{2}} dt \\
&= \|g\|^2_{L^2 L^{\frac{2d}{d+2}}},
\end{aligned}
$$

and hence we obtain (3.7). □

3.6 Small data solutions to the nonlinear Schrödinger equation

Most of this section can be found in [4].

We study the initial value problem with initial data $u_0 \in L^2$ for

$$iu_t + \Delta u = \pm |u|^\sigma u, \tag{3.10}$$

where $0 \le \gamma \le \frac{4}{d-2}$. The case of the plus sign is called defocusing and the case of the minus sign is called focusing. At least formally

$$M = \int_{\mathbb{R}^d} |u|^2 dx,$$

called mass,

$$\int_{\mathbb{R}^d} iu\partial_{x_j} \bar{u} dx,$$

called momentum,

$$E = \int_{\mathbb{R}} \frac{1}{2}|\nabla u|^2 \pm \frac{1}{\sigma+2}|u|^{\sigma+2} dx,$$

called energy, are conserved.

The argument will rely on the Strichartz estimates with $p = q = \frac{2(d+2)}{d}$ and $p' = q' = \frac{2(d+2)}{d+4}$.

The sign of the coefficients is of almost no importance in this section, and we choose + to cover both signs, indicating differences whenever necessary. This section establishes basic schemes that will be used over and over again. Simultaneously it provides a warm up, the set up, and the consequences of the key multilinear estimate. Later on we will often restrict ourselves to giving the estimates of the nonlinearity, and stating the properties.

The section also provides a playground for stability estimates, qualitative properties, criticality and subcriticality.

3.7 Initial data in L^2

Our approach will be based on the Strichartz estimates of Theorem 3.1 with $p = q = \frac{(2(d+2)}{d}$:

$$\|v\|_{L^{\frac{2(d+2)}{d}}(\mathbb{R}\times\mathbb{R}^d)} + \|v\|_{C_b(\mathbb{R};L^2(\mathbb{R}^d))} \lesssim \|v(0)\|_{L^2(\mathbb{R}^d)} + \|i\partial_t v + \Delta v\|_{L^{\frac{2(d+2)}{d+4}}(\mathbb{R}\times\mathbb{R}^d)}. \tag{3.11}$$

In order to prepare for variants and improvements we assume that there is a space X with

$$X \subset C(\mathbb{R}; L^2(\mathbb{R}^d)) \cap L^{\frac{2(d+2)}{d}}(\mathbb{R}\times\mathbb{R}^d) \tag{3.12}$$

and

$$\sup_t \|v(t)\|_{L^2} + \|v\|_{L^{\frac{2(d+2)}{d}}(\mathbb{R}\times\mathbb{R}^d)} \le c\|v\|_X$$

and

$$\|v\|_X \le c\left(\|v(0)\|_{L^2} + \|i\partial_t v + \Delta v\|_{L^{\frac{2(d+2)}{d+4}}(\mathbb{R}\times\mathbb{R}^d)}\right).$$

Clearly such a space exists: we could define X as the intersection in (3.12), and then the Strichartz estimates ensure that it has the desired properties. The choice of the function space is an important and nontrivial part of studying solutions to many different dispersive equations. Even though we do not need this flexibility here, and even though it complicates the notation a bit, we prefer to do it here to indicate possible modifications later on.

In the sequel we denote by v the solution to the homogeneous equation

$$i\partial_t v + \Delta v = 0, \qquad v(0) = u_0,$$

which we can write by using the unitary Schrödinger group $S(t)$ as

$$v(t) = S(t)u_0.$$

To approach the question of existence and uniqueness, we make the ansatz $u = v + w$, where v satisfies the linear Schrödinger equation with initial data u_0, and w satisfies $w(0) = 0$ and

$$\begin{aligned} iw_t + \Delta w &= \chi_{(0,T)}(t)|v+w|^\sigma(v+w) \quad \text{in } \mathbb{R}\times\mathbb{R}^d, \\ w(0,x) &= 0 \quad \text{in } \mathbb{R}^d, \end{aligned} \tag{3.13}$$

where $T \in (0,\infty]$ will be chosen later. We will construct a unique w in X by a fixed point argument. It is obvious that $u = v + w$ is the unique solution up to time T. Then $u = v + w$ is the searched for solution on the time interval $(0,T)$.

We rewrite the problem as a fixed point problem: Given $\tilde{w}$, we write $w = J(\tilde{w})$, where J maps $\tilde{w}$ to the function w which satisfies

$$iw_t + \Delta w = \chi_{(0,T)}(t)|v+\tilde{w}|^\sigma(v+\tilde{w}), \quad w(0) = 0. \tag{3.14}$$

Suppose first that $\frac{2(d+2)}{d+4}(1+\sigma) \geq 2$ and $\sigma \leq \frac{4}{d}$. By Hölder's inequality,

$$\|f\|^{1+\sigma}_{L^{(1+\sigma)\frac{2(d+2)}{d+4}}(\mathbb{R}^d)} \leq \|f\|^{\frac{4-d\sigma}{2}}_{L^2(\mathbb{R}^d)} \|f\|^{\frac{d+2}{2}\sigma-1}_{L^{\frac{2(d+2)}{d}}(\mathbb{R}^d)}$$

Observe that the exponent of $\|f\|_{L^2}$ is non-negative if $\sigma < \frac{4}{d}$ and vanishes if $\sigma = \frac{4}{d}$.

If $0 < \frac{2(d+2)}{d+4}(1+\sigma) \leq 2$ we estimate again by Hölder's inequality:

$$\|f\|^{1+\sigma}_{L^{(1+\sigma)2}(\mathbb{R}^d)} \leq \|f\|^{1-\frac{d\sigma}{4}}_{L^2(\mathbb{R}^d)} \|f\|^{(1+\frac{d}{4})\sigma}_{L^{\frac{2(d+2)}{d}}(\mathbb{R}^d)}.$$

In the first case we obtain the space-time estimate

$$\|\chi_{(0,T)}|u|^{1+\sigma}\|_{L^{\frac{2(d+2)}{d+4}}} \leq T^{1-\frac{d\sigma}{4}} \|u\|^{\frac{4-d\sigma}{2}}_{L^\infty_t L^2_x} \|u\|^{\frac{d+2}{2}\sigma-1}_{L^{\frac{2(d+2)}{d}}}, \tag{3.15}$$

and in the second case

$$\|\chi_{(0,T)}|v|^{1+\sigma}\|_{L^1_t L^2_x(\mathbb{R}^d)} \leq T^{1-\frac{d\sigma}{4}} \|u\|^{1-\frac{d\sigma}{4}}_{L^\infty L^2} \|u\|^{(1+\frac{d}{4})\sigma}_{L^{\frac{2(d+2)}{d}}([0,T]\times\mathbb{R}^d)}. \tag{3.16}$$

If $\sigma < \frac{4}{d}$, T carries a positive power and we call this situation L^2 subcritical. This power becomes zero if $\sigma = \frac{4}{d}$, which we call L^2 or mass critical.

In the both cases

$$\|J(\tilde{w})\|_X \le cT^{1-\frac{d\sigma}{4}}(\|\tilde{w}\|_X + \|v\|_X)^{1+\sigma}$$

which we complement by the similar estimate

$$\|J(w) - J(\tilde{w})\|_X \le cT^{1-\frac{d\sigma}{4}}(\|\tilde{w}\|_X + \|w\|_X + \|v\|_X)^{\sigma}\|w - \tilde{w}\|_X.$$

We set up the problem for an application of the contraction mapping principle Let $R = \|v\|_X$. If $\|\tilde{w}\|_X \le R$ then, for some $c > 0$,

$$\|w\|_X \le cT^{1-\frac{d\sigma}{4}}(2R)^{1+\sigma} \le R,$$

where the last inequality holds provided

$$T \le (2c(2R)^{\sigma})^{-\frac{4}{4-d\sigma}} := T_0,$$

which we assume in the sequel. Moreover, if w and $\tilde{w}$ have norm at most R, then

$$\|J(w) - J(\tilde{w})\|_X \le cT^{1-\frac{d\sigma}{4}}R^{\sigma}\|w - \tilde{w}\|_X$$

We obtain a contraction after decreasing T if necessary.

The critical case requires slightly different arguments, and it yields different conclusions. This time we cannot gain a small power of T and the smallness must have a different source.

In the mass critical case we assume that $\|\chi_{(0,T)}v\|_{L_t^{\frac{2(d+2)}{d}} L_x^{\frac{2(d+2)}{d}}} \le \varepsilon$ for some small ε.

This is true for all T by Lemma 3.11 if $\|u_0\|_{L^2}$ is sufficiently small. Moreover, for all initial data $u_0 \in L^2$ we have by dominated convergence

$$\|\chi_{(0,T)}v\|_{L^q L^p} \to 0 \qquad \text{as } T \to 0 \tag{3.17}$$

for all Strichartz pairs with $q < \infty$.

It is obvious from the argument above (where we replace

$$\|\chi_{(0,T)}v\|_X \text{ by } \|\chi_{(0,T)}v\|_{L^{\frac{2(d+2)}{2}}}$$

for the mass critical case) that the iteration argument applies if ε is sufficiently small. We obtain local existence under the smallness assumption, and hence global existence provided the initial data are sufficiently small.

We collect the results in a theorem.

Theorem 3.7. *There exists $\varepsilon > 0$ such that the following is true. Suppose that $0 < \sigma \le \frac{4}{d}$, $u_0 \in L^2$, and*

$$T^{1-\frac{d\sigma}{4}} \|\chi_T v\|_X^\sigma < \varepsilon,$$

resp. $\sigma = \frac{4}{d}$ and

$$\|\chi_T v\|^\sigma_{L^{\frac{2(d+2)}{2d}}(\mathbb{R}\times\mathbb{R}^d)} < \varepsilon.$$

Then there is a unique solution in X up to time T which satisfies

$$\|\chi_T(u-v)\|_X \le \|u-v\|_X \lesssim T^{1-\frac{d\sigma}{4}} \|v\|_X^{1+\sigma} \tag{3.18}$$

resp., if $\sigma = \frac{4}{d}$,

$$\|\chi_T(u-v)\|_X \le \|u-v\|_X \lesssim T^{1-\frac{d\sigma}{4}} \|v\|^{1+\sigma}_{L^{\frac{2(d+2)}{d}}}. \tag{3.19}$$

There is a unique global solution

$$u \in L^{\frac{2(d+2)}{d}}((-T,T)\times\mathbb{R}^d) \cap C((-T,T);L^2(\mathbb{R}^d))$$

for all T if either $0 \le \sigma < \frac{d}{4}$, or, if $\|u_0\|_{L^2} \le \varepsilon$ and $\sigma = \frac{d}{4}$. In the last case we have (3.19) with $T = \infty$. If $0 \le k < 1+\sigma$, then

$$(u_0 \to u) \in C^k(L^2(\mathbb{R}^d);X)$$

There is a stability estimate. Suppose that $\tilde u \in X$ satisfies

$$T^{1-\frac{d\sigma}{4}} \|\tilde u\|_X < \varepsilon$$

$$\|\tilde u - u_0\|_{L^2} + \|i\partial_t \tilde u + \Delta \tilde u - |\tilde u|^\sigma \tilde u\|_{L^{\frac{2(d+2)}{d+4}}} < \varepsilon.$$

Then there exists a unique solution up to time T with

$$\|u-\tilde u\|_X \le c\left(\|\tilde u - u_0\|_{L^2} + \|i\partial_t \tilde u + \Delta \tilde u - |\tilde u|^\sigma \tilde u\|_{L^{\frac{2(d+2)}{d+4}}}\right). \tag{3.20}$$

If $\sigma = \frac{4}{d}$, it suffices to replace $\|u\|_X < \varepsilon$ by

$$\|\chi_{(0,T)}\tilde u\|_{L^{\frac{2(d+2)}{d}}} < \varepsilon.$$

Proof. Local existence in the subcritical case has been shown above. The fixed point formulation leads, via the contraction mapping theorem, to existence on a time interval whose length depends only on $\|u_0\|_{L^2}$. We claim that the L^2 norm (mass) is conserved. Indeed, for sufficiently regular and decaying $\tilde u = v + \tilde w$ and $u = v + w$ with $w = J(\tilde w)$ we have

$$\frac{1}{2}\|u(t)\|^2_{L^2} = \frac{1}{2}\|u_0\|^2_{L^2} + \operatorname{Re} i \int_{(0,t)\times\mathbb{R}^d} |\tilde u|^\sigma \tilde u \bar u \, dx dt,$$

which remains true for general $\tilde{u}$ and initial data by an approximation argument. By then it also holds for the fixed point, for which the second term on the right-hand side is the real part of something purely imaginary.

Thus we can extend the solution to a global solution in the subcritical case.

It follows from the construction by the contraction mapping principle that the solution depends Lipschitz continuously on the initial data.

The map

$$L^{\frac{2(d+2)}{d}}(\mathbb{R} \times \mathbb{R}^d) \ni w \mapsto \chi_{(0,T)}|w|^\sigma w \in L^{\frac{2(d+2)}{d+4}}(\mathbb{R} \times \mathbb{R}^d)$$

is k times continuously differentiable for $k < 1 + \sigma$, and $\sigma \leq \frac{4}{d}$.

Thus J is k times continuously Fréchet differentiable. Moreover, by the very same estimates as for the contraction, the derivative of J with respect to $\tilde{w}$ is invertible, and by the implicit function theorem, the map from the initial data to the solution is k times continuously differentiable. Checking the norms implies the stability estimate. $\square$

We also have

$$\lim_{T\to\infty} \|\chi_{(T,\infty)} v\|_{L^{\frac{2(d+2)}{d}}(\mathbb{R}\times\mathbb{R}^d)} = 0.$$

Suppose that $u \in X$ is a solution for $T = \infty$ and $\sigma = \frac{4}{d}$. One can deduce that the limit

$$\lim_{t\to\infty} S(-t)u(t)$$

exists in L^2. Let w_0 be this limit, and w the solution to the homogeneous equation with initial data w_0. Then the convergence statement can be formulated as

$$\lim_{t\to\infty} \|u(t) - w(t)\|_{L^2} = 0.$$

This is called scattering. The map $u_0 \to w_0$ is called wave operator.

3.8 Initial data in $\dot{H}^1$ for $d \geq 3$

Consider

$$iu_t + \Delta u = \pm|u|^\sigma u \tag{3.21}$$

with initial data $u_0 \in \dot{H}^1$, by which we mean the space with the norm $\| \, |\nabla u_0| \, \|_{L^2}$. We want to use Strichartz spaces for the derivative and we define the function spaces X by

$$\|u\|_X := \sup_t \|\nabla u(t)\|_{L^2} + \|\nabla u\|_{L^{\frac{2(d+2)}{d}}}.$$

Then the Strichartz estimate (3.11) combined with Sobolev's estimate gives

$$\|u\|_X \leq c\left(\|\nabla u_0\|_{L^2} + \|\nabla f\|_{L^{\frac{2(d+2)}{d+4}}}\right)$$

for a solution u to the inhomogeneous linear problem.

Then, if $\sigma \le \frac{4}{d-2}$, by Hölder's and Sobolev's inequality

$$\| \nabla |f|^\sigma f \|_{L^{\frac{2(d+2)}{d+4}}(\mathbb{R}^d)} \lesssim \|f\|^{\frac{4-(d-2)\sigma}{2}}_{L^{\frac{2d}{d-2}}(\mathbb{R}^d)} \|\nabla f\|^{-1+\frac{4-d}{2}\sigma}_{L^{\frac{2(d+2)}{d}}(\mathbb{R}^d)},$$

provided σ is not too small. For small σ we argue as for the case of L^2. We obtain in both cases

$$\|J(w)\|_X \lesssim T^{1-\frac{(d-2)\sigma}{4}} \left(\|v\|_X + \|w\|_X \right)^{1+\sigma}, \tag{3.22}$$

and, checking the same argument for differences,

$$\begin{aligned} \|J(w^2) - J(w^1)\|_{L^{\frac{2(d+2)}{d}}} &+ \|J(w^2) - J(w^1)\|_{L^\infty L^2} \\ &\lesssim T^{1-\frac{(d-2)\sigma}{4}} \left(\|v\|_X + \|w^1\|_X + \|w^2\|_X \right)^\sigma \qquad (3.23) \\ &\quad \times \left(\|w^2 - w^1\|_{L^{\frac{2(d+2)}{d}}} + \|w^2 - w^1\|_{L^\infty L^2} \right) \end{aligned}$$

Theorem 3.8 (Local existence and uniqueness in energy space)**.** *Suppose that* $0 < \sigma \le \frac{4}{d-2}$. *There exists* $\varepsilon > 0$ *such that the following is true. Let* v *be the solution to the homogeneous linear Schroedinger equation. Suppose that*

$$T^{1-\frac{(d-2)\sigma}{4}} \|v\|_X^\sigma \le \varepsilon.$$

Then there exists a unique solution $u = v + w$ *with*

$$\|\nabla w\|_{L^\infty L^2} + \|\nabla w\|_{L^{\frac{2(d+2)}{d}}} \lesssim T^{1-\frac{(d-2)\sigma}{4}} \|v\|_X^{1+\sigma}.$$

Again we may replace $\|v\|_X$ *by* $\|\chi_{0,T} \nabla v\|_{L^{\frac{2(d+2)}{d}}}$. *In the defocusing case, the solution is global if* $\sigma < \frac{4}{d-2}$. *In the energy critical case* $\sigma = \frac{4}{d-2}$, *there is global existence for small data, and local existence for all data in* $\dot H^1$.

Proof. Again we characterize the solution as the fixed point of the same map as above, but now with respect to the norm X. By (3.22), we obtain a map of a closed ball in X to itself, but a contraction only in the metric of $L^{\frac{2(d+2)}{d}}$ in a ball in X, at least for large space dimensions and small σ. We change the space X slightly by replacing $C(\mathbb{R}; L^2)$ by $L^\infty(\mathbb{R}; L^2)$. We claim that sequences that are bounded in X and converge in $L^{\frac{2(d+2)}{d}}$ have a limit in X. There is a weak* converging subsequence in X, and the limits have to coincide.

It is not hard to complete the argument for initial data additionally in $L^2(\mathbb{R}^d)$: then $v \in L^{\frac{2(d+2)}{2}}$, and this remains true for the fixed point map. In general, we define iteratively $v_{j+1} = J(v_j)$. We claim that there exists j so that

$$v_{j+1} - v_j \in L^{\frac{2(d+2)}{d}}(\mathbb{R} \times \mathbb{R}^d).$$

The contraction argument then completes the proof. This argument gives uniqueness in the set

$$v_j + X \cap L^{\frac{2(d+2)}{d}}(\mathbb{R} \times \mathbb{R}^d).$$

The proof of the claim is technical and omitted.

The remaining arguments are adaptations of similar arguments in Theorem 3.7. □

3.9 Initial data in $H^1(\mathbb{R}^d)$

In this case we combine the arguments. We obtain global well-posedness in the defocussing subcritical case $\sigma < \frac{4}{d-2}$, local existence in the subcritical and the critical case ($\sigma \leq \frac{4}{d-2}$), and global existence in the critical case $\sigma = \frac{4}{d-2}$ and small initial data.

Chapter 4

Functions of bounded p-variation

The study of p-variation of functions of one variable has a long history. Function of bounded p-variation have been studied by Wiener in [33]. The generalization of the Riemann–Stieltjes integral to functions of bounded p-variation against the derivative of a function of bounded q-variation, $1/p + 1/q > 1$, is due to Young [34]. Much later Lyons developed his theory of rough paths [23] and [24], building on Young's ideas, but going much further.

In parallel D. Tataru realized that the spaces of functions of bounded p-variation, and their close relatives, the U^p spaces, allow a powerful sharpening of Bourgain's technique of function spaces adapted to the dispersive equation at hand. These ideas were applied for the first time in the work of the author and Tataru in [18]. Since then there have been a number of questions in dispersive equations where these function spaces have been used. For example, they play a crucial role in [19], but there they could probably be replaced by Bourgain's Fourier restriction spaces $X^{s,b}$. On the other hand, for well-posedness for the Kadomtsev–Petviashvili II in a critical function space (see [12]) the $X^{s,b}$ spaces seem to be insufficient. The theory of the spaces U^p and V^p and some of their basic properties like duality and logarithmic interpolation have been worked out for the first time in [12]. The developments in stochastic differential equations and dispersive equations have been largely independent.

We will introduce and study functions from an interval (a, b) to $\mathbb{R}$, $\mathbb{R}^n$, a Hilbert space, or a Banach space X, and spaces of such functions which are invariant under continuous monotone reparametrizations of the interval. For the most part of this section there are no more than the obvious modifications when considering Banach space valued functions. We allow $a = -\infty$ and $b = \infty$.

We call a function f ruled function if at every point (including the endpoints, which may be $\pm\infty$) left and right limits exist. The set of ruled functions is closed with respect to uniform convergence. We denote the Banach space of ruled functions equipped with the supremum norm by $\mathcal{R}$.

A step function is a function f for which there exists a partition so that f is constant on every interval (a, t_1), (t_i, t_{i+1}), and (t_n, b). We do not require that

H. Koch et al., *Dispersive Equations and Nonlinear Waves: Generalized Korteweg–de Vries, Nonlinear Schrödinger, Wave and Schrödinger Maps*, Oberwolfach Seminars 45,
DOI 10.1007/978-3-0348-0736-4_4, © Springer Basel 2014

the value at a point coincides with the limit from either side. Step functions are dense in $\mathcal{R}$ (Aumann [1], Dieudonné [6]). We denote the set of step functions by $\mathcal{S}$.

Let $\mathcal{R}_{rc} \subset \mathcal{R}$ be the closed subspace of right continuous functions f with $\lim_{t\to a} f(t) = 0$. Similarly, if $A \subset \mathcal{R}$ we denote by A_{rc} the intersection with $\mathcal{R}_{rc}$.

Let X be a Banach space and X^* its dual. We consider functions with values in X, resp. X^*, and we denote the corresponding spaces by $\mathcal{R}(X)$, resp. $\mathcal{S}(X)$.

There is a bilinear map B from $\mathcal{S}(X)_{rc} \times \mathcal{R}(X^*)$ to $\mathbb{R}$ resp. $\mathbb{C}$, defined by

$$B(u,v) = \sum_{i=1}^{n} v(t_i)(u(t_i) - u(t_{i-1})), \tag{4.1}$$

where $a = t_0 < t_1 < \cdots < t_n < b$ is the partition. In the sequel we will omit the space X and X^* from the notation unless there is some ambiguity. Similarly, the formula above defines a bilinear map on $\mathcal{R}(X^*) \times \mathcal{S}(X)$.

It will be convenient to extend every function on $[a,b)$ by zero to $[a,b]$, i.e. we will always set $f(b) = 0$, even if $a = -\infty$ or $b = \infty$. Similarly, we extend every function by 0 to $\mathbb{R}$ whenever this is convenient.

Definition 4.1. *For $u \in \mathcal{R}$ and a partition*

$$\tau = (t_1, t_2, \ldots, t_n), \quad a < t_1 < t_2 < t_3 \cdots < t_n < b,$$

we define (denoting the limit from the right by $f(t+)$)

$$u_\tau(t) = \begin{cases} u(t), & \text{if } t = t_j \text{ for a } j, \\ u(a+), & \text{if } a < t < t_1, \\ u(t_i+), & \text{if } t_i < t < t_{i+1}, \\ u(t_n+), & \text{if } t_n < t. \end{cases}$$

We observe that u_τ is a step function, and it is right continuous if u is right continuous.

Lemma 4.2. *Let $u \in \mathcal{R}_{rc}$ and $v \in \mathcal{R}$. Then*

$$B(u_\tau, v) = B(u_\tau, v_\tau).$$

Moreover, if in addition v is left continuous, then

$$B(u, v_\tau) = B(u_\tau, v_\tau).$$

If $u, v \in \mathcal{S}_{rc}$, then, with t_i a partition containing all points of discontinuity of u and v,

$$B(u,v) + B(v,u) = \sum_i (v(t_i) - v(t_i-)(u(t_i) - u(t_i-)) + \lim_{t\to b} v(t)u(t).$$

Proof. This follows immediately from the definitions. $\square$

4.1 Functions of bounded p-variation and the spaces U^p and V^p

Unless explicitly stated otherwise, we consider $p \in (1, \infty)$.

In later chapters we use U^p and V^p to study well-posedness questions for several dispersive PDEs, where we select a number of relevant and representative problems.

A partition τ of (a, b) is a strictly increasing finite sequence

$$a < t_1 < t_2 < \cdots < t_{n+1} < b,$$

where we allow $b = \infty$ and $a = -\infty$.

Definition 4.3. *Let I be an interval, X a Banach space, $1 \leq p < \infty$, and $f : I \to X$. We define*

$$\omega_p(v, I) := \sup_\tau \left(\sum_{i=1}^{n-1} \|v(t_{i+1}) - v(t_i)\|_X^p \right)^{1/p} \in [0, \infty].$$

There are obvious properties. The function $t \mapsto \omega_p(v, [a, t))$ is monotonically increasing. The same is true if we consider closed or open intervals.

Lemma 4.4. *Suppose that $a < b < c$. Then*

$$\omega_p(v, [a, b)) \leq \omega_p(v, [a, c)) \leq 2^{1-1/p} \Big(\omega_p(v, [a, b]) + \omega_p(v, [b, c)) \Big).$$

Proof. Consider a partition τ. If b is a point of τ, then the p-th power of the τ-variation in the large interval is the sum of the p powers of the parts. If not, we add the point b. The factor $2^{1-1/p}$ follows from

$$|a + b|^p \leq 2^{p-1}(|a|^p + |b|^p).$$ □

The p-variation can sometimes be explicitly estimated.

Lemma 4.5. *For bounded monotone functions we have*

$$\omega_p(v, [a, b)) = \sup v - \inf v.$$

We denote by $\dot{C}^s(I)$ the homogeneous Hölder norm:

$$\|f\|_{\dot{C}^s(I)} = \sup_{t \neq \tau} \frac{|u(t) - u(\tau)|}{|t - \tau|^s}.$$

Lemma 4.6. *We have*

$$\omega_p(v,(a,b)) \le \|v\|_{\dot C^{1/p}}(b-a)^{1/p}. \tag{4.2}$$

Suppose that

$$\omega_p(v,(a,b)) < \infty.$$

Then v has left and right limits at every point. Moreover

$$\omega_p(\lambda v,(a,b)) = |\lambda|\omega_p(v,(a,b)),$$

$$\omega_p(v+w,(a,b)) \le \omega_p(v,(a,b)) + \omega_p(w,(a,b)).$$

Proof. Let $t_0 < t_1 < \cdots < t_N$. Then

$$\sum_j \|v(t_{i+1}) - v(t_i)\|_X^p \le \sum_i (t_{i+1}-t_i)\|v\|^p_{\dot C^{1/p}}.$$

The other statement follow from a straightforward calculation. □

The p-variation is continuous at points where v is continuous, provided the p-variation is finite.

Lemma 4.7. *Suppose that $\omega_p(v,[a,b)) < \infty$ and v is continuous at $c \in [a,b)$. Then*

$$\lim_{t\to c} \omega_p(v,[a,t)) = \omega_p(v,[a,c]).$$

Proof. Suppose that

$$\lim_{t\downarrow c} \omega(v,(a,t)) - \omega(v,(a,c)) = 2\delta > 0.$$

Then there is a sequence of points $c < t_1 < t_2 \cdots < t_n < b$ with

$$\sum \|v(t_{i+1}) - v(t_i)\|_X^p \ge \frac{\delta^p}{p\omega_p(v,[a,b))^{p-1}}.$$

Similarly, there is such a sequence in (c, t_1) and recursively we get an arbitrarily large number of such sequences. Putting N of them together, we see that

$$\omega_p(v,(c,b)) \ge Nc\delta,$$

which would bound N. This is a contradiction. We argue similarly for the limit from below. □

Definition 4.8. *Let X be a Banach space , $1 \le p < \infty$, and $v : (a,b) \to X$. We define*

$$\|v\|_{V^p((a,b),X)} = \max\{\|v\|_{\sup}, \omega_p(v,(a,b))\}.$$

Let $V^p = V^p((a,b)) = V^p(X) = V^p((a,b);X)$ be the set of all functions for which this expression is finite. We omit the interval and/or the Banach space in the notation when this seems appropriate.

The interval will usually be of minor importance. The following properties are immediate:

1. $V^p(I)$ is closed with respect to this norm and hence $V^p(I)$ is a Banach subspace of $\mathcal{R}$. Moreover, $V^p_{rc}(I)$ is a closed subspace.
2. We set $V^\infty = \mathcal{R}$ with $\|.\|_{V^\infty} = \|.\|_{\sup}$.
3. If $1 \le p \le q \le \infty$, then
$$\|v\|_{V^q} \le \|v\|_{V^p}.$$
4. Let X_i be Banach spaces, $T : X_1 \times X_2 \to X_3$ a bounded bilinear operator, $v \in V^p(X_1)$, and $w \in V^p(X_2)$. Then $T(v,w) \in V^p(X_3)$ and
$$\|T(v,w)\|_{V^p(X_3)} \le 2\|T\|\|v\|_{V^p(X_1)}\|w\|_{V^p(X_2)}.$$
5. We embed $V^p((a,b))$ into $V^p(\mathbb{R})$ by extending v by 0.
6. The space V^1 has some additional structure: every bounded monotone function is in V^1, and functions in V^1 can be written as the difference of two bounded monotone functions.

The space of bounded p-variation is build on the sequence space l^p. We may also replace it by the weak space l^p_w, with

$$\|(a_j)\|_{l^p_w} = \sup_{\lambda>0} \lambda(\#\{j : |a_j| > \lambda\})^{\frac{1}{p}}.$$

This does not satisfy the triangle inequality, but if $p > 1$, there is an equivalent norm, which makes l^p_w a Banach space. We set $l^\infty_w = l^\infty$.

Definition 4.9. *Let $1 \le p < \infty$. The weak V^p_w space consists of all functions such that*
$$\|v\|_{V^p_w} = \max\left\{ \sup_{t_1<\cdots<t_n} \|(v(t_{i+1}) - v(t_i))_{1\le i\le n-1}\|_{l^p_w}, \|v\|_{\sup}\right\}$$
is finite.

By Tschebycheff's inequality

$$\|v\|_{V^p_w} \le \|v\|_{V^p}.$$

The spaces of bounded p-variation are of considerable importance in probability and harmonic analysis. We shall see that V^p is the dual space of a space U^q, $1/p + 1/q = 1$, $1 < p < \infty$, with a duality pairing closely related to the Stieltjes integral, and its variant, the Young integral [34].

Definition 4.10. *A p-atom a is a step function in $\mathcal{S}_{rc}$,*

$$a(t) = \sum_{i=1}^n \phi_i \chi_{[t_i,t_{i+1})}(t),$$

where $\tau = (t_1, \dots, t_n)$ is a partition, $t_{n+1} = b$, with $\sum |\phi_i|^p \le 1$. A p-atom a is called a strict p-atom if
$$\max_i \|\phi_i\|_X (\#\tau)^{1/p} \le 1.$$
It is important that atoms are right continuous, zero in a neighborhood of a, but the limit as $t \to b$ may be different from 0.

Let a_j be a sequence of atoms and let λ_j be a summable sequence. Then
$$u = \sum \lambda_j a_j$$
is a U^p function. This is well defined, since the right-hand side converges in $\mathcal{R}$. We define U^p as the set of functions having such a representation and give it the norm
$$\|u\|_{U^p} := \inf \Big\{ \sum |\lambda_j| : u = \sum \lambda_j a_j \Big\}.$$
The strict space U^p_{strict} is defined in the same fashion using strict p-atoms.

We collect a number of elementary properties.

1. If a is a p-atom, then $\|a\|_{U^p} \le 1$. The norm of an atom may be less than 1. Determining the norm of an atom is a difficult task.
2. Functions in U^p are continuous from the right. The limit as $t \to a$ vanishes.
3. The expression $\|.\|_{U^p}$ defines a norm on U^p, and U^p is closed with respect to this norm. Moreover $U^p \subset \mathcal{R}_{rc}$ is a subspace with $\|.\|_{\sup} \le \|.\|_{U^p}$.
4. If $p < q$, then $U^p \subset U^q$ and
$$\|u\|_{U^q} \le \|u\|_{U^p}.$$
5. If $1 \le p < \infty$, then for all $u \in U^p$ we have $u \in V^p_{rc}$ and
$$\|u\|_{V^p} \le 2^{1/p} \|u\|_{U^p}.$$
6. Let Y be a Banach space, and let the linear operator $T : \mathcal{S}_{rc} \to Y$ satisfy
$$\|Ta\|_Y \le C \tag{4.3}$$
for every p-atom. Then T has a unique extension to a bounded linear operator from U^p to Y which satisfies
$$\|Tf\|_Y \le C \|f\|_{U^p}. \tag{4.4}$$
7. Let X_i be Banach spaces, $T : X_1 \times X_2 \to X_3$ a bounded bilinear operator, $v \in U^p(X_1)$, and $w \in U^p(X_2)$. Then $T(v, w) \in U^p(X_3)$ and
$$\|T(v, w)\|_{U^p(X_3)} \le \|T\| \|v\|_{U^p(X_1)} \|w\|_{U^p(X_2)}.$$

8. We consider $U^p([a,b))$ as a subset of $U^p(\mathbb{R})$ by extending the function u by zero to the left, and by $\lim_{t\to b} u(t)$ to the right.

The following decomposition is crucial for most of the following. It is related to Young's generalization of the Stieltjes integral, and it deals with a crucial point in the theory. We denote the number of points in a partition τ by $\#\tau$.

Lemma 4.11. *There exists $\delta > 0$ such that for v right continuous with $v(a+) = 0$ and $\|v\|_{V^p_w} = \delta$ there are strict p-atoms a_i with*

$$\|a_j(t)\|_{\sup} \le 2^{1-j} \quad \textit{and} \quad \#\tau_j \le 2^{jp},$$

such that in the sense of uniform convergence

$$v = \sum a_j.$$

Proof. We set $v_0 = v$, and we search for a recursive decomposition with

$$v_j = a_j + v_{j+1},$$

such that

$$\|v_j\|_{\sup} \le 2^{-j}, \quad \|a_j\|_{\sup} \le 2^{-j}$$

and, with τ_j the partition related to a_j,

$$\#\tau_j \le 2^{pj}.$$

Suppose we have constructed v_i for $i \le j$ and a_i for $i \le j-1$. We construct the a_j, which also defines v_{j+1}. We choose the unique partition τ so that

$$\sup_t \|v_j(t)\|_X < 2^{-1-j} \text{ in } [a, t_1), \quad \|v_j(t_1)\|_X \ge 2^{-1-j},$$

$$\|v_j(t) - v_j(t_i)\|_X < 2^{-1-j} \text{ in } t \in [t_i, t_{i+1}),$$

and

$$\|v_j(t_{i+1}) - v_j(t_i)\|_X \ge 2^{-1-j}.$$

We define a_j as the step function adapted to the partition τ_j (recall Definition 4.1)

$$a_j = (v_j)_\tau$$

Then, by construction,

$$\|a_j\|_{\sup} \le \|v_j\|_{\sup} \le 2^{-j},$$

$$\|v_{j+1}\|_{\sup} \le 2^{-1-j},$$

and since either $(t_j, t_{j+1}]$ contains no points of an earlier partition, in which case we estimate the number of such points using the V^p_w norm of v, or it does, and then we simply add the number of those terms, and iterate, we get

$$\begin{aligned}
\#\tau_j &\le 2^p \|v\|^p_{V^p_w} 2^{jp} + \sum_{i=0}^{j-1} \#\tau_i \\
&\le 2^p \|v\|^p_{V^p_w} \sum_{i=0}^{j} (j+1-i) 2^{ip} \\
&\le c_p \|v\|^p_{V^p_w} 2^{jp}
\end{aligned} \tag{4.5}$$

We choose $\delta = c_p^{-1/p}$. □

There are a number of simple interesting and useful consequences.

Lemma 4.12. *Let $1 < p < q < \infty$. There exists $\kappa > 0$, depending only on p and q, such that for all $v \in V^p_{w,rc}$ and $M \ge 1$ there exist $u \in U^p_{\text{strict}}$ and $w \in U^q_{\text{strict}}$ with*

$$v = u + w$$

and

$$\frac{\kappa}{M} \|u\|_{u^p_{\text{strict}}} + e^M \|w\|_{U^q_{\text{strict}}} \le \|v\|_{V^p_w}.$$

Observe that we may replace U^p_{strict} by U^p (since $U^p_{\text{strict}} \subset U^p$) and V^p_w by V^p (since $V^p \subset V^p_w$).

Proof. Multiplying v by $\delta/\|v\|_{V^p_{w,rc}}$ we may assume that $\|v\|_{V^p_w} = \delta$ as in Lemma 4.11, and setting $\tilde u = \sum_{j=1}^m a_j$ for some m to be chosen later, we have

$$\|\tilde u\|_{U^p_{\text{strict}}} \le m.$$

By construction, $2^{j(1-p/q)} a_j$ is a strict q-atom and hence, with $\tilde w = \sum_{j=m+1}^\infty a_j$,

$$\|\tilde w\|_{U^q_{\text{strict}}} \le \sum_{j=m+1}^\infty \|a_j\|_{U^q_{\text{strict}}} \le c_{p,q} 2^{(\frac{p}{q}-1)m},$$

hence, with $u = \frac{\|v\|_{V^p_w}}{\delta} \tilde u$ and $w = \frac{\|v\|_{V^p_w}}{\delta} \tilde w$,

$$u + w = v$$

and, with $\delta = -\ln 2(\frac{p}{q} - 1)$, there exists c depending only on p and w such that

$$\frac{1}{m} \|u\|_{U^p} + e^{\delta m} \|v\|_{U^q} \le c \|v\|_{V^p_{rc}}.$$

We choose $m = (M + \ln 2c)/\delta$ and, for $M \ge \ln 2c$, $\kappa = \delta/2$ to obtain the claimed estimate. □

We obtain the following embedding

Lemma 4.13. *Let $1 < p < q < \infty$. Then*

$$V^p_{rc} \subset V^p_{w,rc} \subset U^q_{\text{strict}} \subset U^q.$$

Proof. Apply Lemma 4.12 with $M = 1$. □

4.2 Duality and the Riemann–Stieltjes integral

The Riemann–Stieltjes integral defines

$$\int f dg = \int f g_t dt$$

for $f \in \mathcal{R}$ and $g \in V^1$. If $g \in \mathcal{S}_{rc}$ then, with the obvious partition,

$$\int f g_t dt = \sum f(t_i)(g(t_i) - g(t_{i-1})). \tag{4.6}$$

This formula was the definition of the bilinear map B. We shall see that it uniquely defines an 'integral' for $f \in V^p$ and $g \in U^q$, for $1/p + 1/q = 1$, $p > 1$. Results become much cleaner when we use an equivalent norm in V^p,

$$\|v\|_{V^p} = \sup_{a<t_1<\cdots<t_n<b} \left(\sum_{j=1}^{n-1} |v(t_{j+1}) - v(t_j)|^p + |v(t_n)|^p \right)^{1/p} \tag{4.7}$$

which we do in the sequel. We also set $v(b) = 0$ and, for any partition, $t_{n+1} = b$.

Theorem 4.14. *The bilinear map B defines a unique continuous bilinear map*

$$B : U^q(X) \times V^p(X^*) \to \mathbb{R}$$

which satisfies (with $t_0 = a$ and $u(t_0) = 0$)

$$B(u, v) = \sum_{i=1}^{n} v(t_i)(u(t_i) - u(t_{i-1}))$$

for $v \in V^p$ and $u \in \mathcal{S}_{rc}$, with associated partition $(t_1, \dots, t_n)$ and $v(t_i)(.)$ the evaluation of $v(t_i) \in X^*$ on the argument in X. It satisfies

$$|B(u, v)| \leq \|u\|_{U^q(X)} \|v\|_{V^p(X^*)}. \tag{4.8}$$

The map

$$V^p(X^*) \ni v \mapsto (u \mapsto B(u, v)) \in (U^q(X))^*$$

is a surjective isometry if $1 \le q < \infty$. *Moreover*

$$\|v\|_{V^p(X^*)} = \sup_{u \in U^q(X), \|u\|_{U^q(X)}=1} B(u,v) = \sup_{a \text{ is a } q\text{-atom}} B(a,v). \tag{4.9}$$

The same statements up to constants are true if we replace U^p *by* U^p_{strict} *and* V^q *by* V^q_w.

Proof. Let $v \in V^p$. The expression

$$F_v(u) = \sum_{i=1}^n v(t_i)(u(t_i) - u(t_{i-1})) = -\sum_{i=1}^n (v(t_{i+1}) - v(t_i))u(t_i)$$

is clearly defined for $v \in V^q$ and $u \in \mathcal{S}_{rc}$ with partition $\tau = (t_i)$. The product is an abuse of notation for the duality pairing between X and X^* which we suppress in the notation. The map is linear in v and u and satisfies for every atom (by Hölder's inequality, and using the right-hand side of the equation for $F_v(u)$)

$$\begin{aligned} |F_v(a)| &\le \sum_{i=1}^n \|v(t_{i+1}) - v(t_i)\|_{X^*} \|a(t_i)\|_X \\ &\le \left(\sum_{i=1}^n \|v(t_{i+1}) - v(t_i)\|^p_{X^*} \right)^{1/p} \left(\sum_{i=1}^n \|a(t_i)\|^q_X \right)^{1/q} . \end{aligned}$$

The first factor is bounded by $\|v\|_{V^p}$, and the second, by the definition of a q-atom, by 1.

Existence of a unique extension to U^q follows from this estimate and (4.4). Linearity in v and estimate (4.8) are immediate consequences. Clearly B defines a map from V^p to the dual of U^q with norm at most 1. Let us prove that it defines an isometry and choose $v \in V^p$, $\varepsilon > 0$, and a partition $t_0 < t_1 < \cdots < t_n$ with

$$\|v\|_{V^p} \le \left(\sum_{j=1}^n \|v(t_{j+1}) - v(t_j)\|^p_{X^*} \right)^{1/p} + \varepsilon.$$

Here we set again $t_{n+1} = b$ and $v(b) = 0$. We choose $x_i \in X$ of norm 1 with

$$(v(t_{i+1}) - v(t_i))(x_i) \ge (1 - \varepsilon)\|\|v(t_{i+1}) - v(t_i)\|_{X^*}$$

and

$$\phi_j := \mu \|v(t_{j+1}) - v(t_j)\|^{p-1}_{X^*} x_j,$$

where $\mu = \|v\|^{1-p}_{V^p}$. Then

$$\sum_{j=1}^n \|\phi_j\|^{p'}_x \le \mu^{-p} \sum_{j=1}^n \|v(t_{j+1}) - v(t_j)\|^p_{X^*} \le 1.$$

Thus the partition and the ϕ_j define a q-atom a, and

$$\|v\|_{V^p} \geq B(a,v) - C\varepsilon.$$

The map is an isometry since ε is arbitrary. We turn to surjectivity. Let $F \in (U^q)^*$ and define the element $v(t) \in X^*$ by

$$v(t)(x) := F(x\chi_{[t,\infty)}) \quad \text{for } x \in X.$$

Let a be an atom. Then

$$\begin{aligned} F(a) &= \sum_i F(\phi_i \chi_{[t_i,b)}) - F(\phi_i \chi_{[t_{i+1},b)}) = -\sum \phi_i (v(t_{i+1}) - v(t_i)) \\ &= \sum v(t_i)(a(t_i) - a(t_{i-1})) = B(a,v). \end{aligned}$$

By the previous estimate,

$$\|v\|_{V^p} \leq \|F\|_{(U^q)^*}.$$

Hence both sides coincide on U^q. The remaining claims are simple consequences. $\square$

The previous results show that $U^p \subset V^p_{rc}$, and that the two spaces are very close. They are, however, not equal. The following example goes back to Young [34] with the same intention, but in a slightly different context.

Lemma 4.15. *Let ϕ be a smooth function with compact support, $1 < q < \infty$. Then*

$$u_q(t) = \phi(t) \sum_{j=1}^{\infty} 2^{-j/q} \cos(2^j t) \in V^q_{rc},$$

but not in U^q.

Proof. Let p be the Hölder dual exponent of q and

$$v_p^N(t) = \phi \sum_{j=1}^{N} 2^{-j/p} \sin(2^j t),$$

where we allow $N = \infty$. Then, with $M = [\ln_2(|t-s|)] \leq N$, $[\,]$ the Gauss bracket,

$$\begin{aligned} |v_p^N(t) - v_p^N(s)| &\leq \sum_{j=1}^{M} 2^{-j/p} |\phi(t)\sin(2^j t) - \phi(s)\sin(2^j s)| + c_1 \sum_{j=M+1}^{N} 2^{-j/p} \\ &\leq c_2 \left(\sum_{j=1}^{M} 2^{-j/p+j} |t-s| + 2^{-j/M} \right) \\ &\leq c_3 \left(2^{-M/p+M} |t-s| + 2^{-j/M} \right) \\ &\leq c_4 |t-s|^{\frac{1}{p}} \end{aligned}$$

and hence, by Lemma 4.6

$$\sup_N \|v_p^N\|_{V^p} < \infty,$$

and similarly $u_q \in V^q_{rc}$. Now, assuming that $u_q \in U^q$, with $1/p+1/q=1$, we claim

$$\|u_q\|_{U^p}\|v_p^N\|_{V^q} \geq \left|\int (u_q)' v_p^N dx\right| = N/2 \int \phi^2 dx + O(1) \tag{4.10}$$

which is unbounded, hence a contradiction, and so $V^q_{rc} \ni u_q^\infty \notin U^q$. It remains to verify (4.10). The first inequality is a consequence of the duality theorem. We expand both factors in the integral and claim for $j \neq l$ that, by stationary phase,

$$\left|\int \phi(t) 2^{-j/p-l/q} \cos(2^j t)(\phi(t)\sin(2^l t))' dt\right| \leq c_M 2^{-j}|2^j - 2^l|^{-M}$$

for every $M \in \mathbb{N}$. Thus

$$\sum_{j\neq l, l\leq N} \left|\int \phi(t) 2^{-j/p-l/q} \cos(2^j t)(\phi(t)\sin(2^l t))' dt\right| \leq c \sum_{j=1}^\infty 2^{-j} \sum_{l=1, l\neq j}^N 2^{-l},$$

which is bounded independent of N. Next

$$\left|\int \phi(t) 2^{-j/p-j/q} \sin(2^j t)\cos(2^j t)\phi'(t)) dt\right| \leq c_1 2^{-j}$$

and

$$\begin{aligned}\left|\int \phi^2(t) 2^{-j/p-j/q+j} \cos^2(2^j t) dt\right| &= \left|\int \phi^2(t)\frac{1}{2}(1+\cos(2^{j+1}t))dt\right| \\ &= \frac{1}{2}\int \phi^2(t)dt + c_2^{-j}.\end{aligned}$$

We expand (4.10). Only the diagonal terms contribute. This completes the proof. □

4.3 Step functions are dense

Lemma 4.16. *For all $v \in V^p$ and all partitions τ we have (recall Definition 4.1)*

$$\|v_\tau\|_{V^p} \leq \|v\|_{V^p}, \tag{4.11}$$

and for all $u \in U^p$

$$\|u_\tau\|_{U^p} \leq \|u\|_{U^p}. \tag{4.12}$$

For $v \in V^p$ and $\varepsilon > 0$ there is a partition τ so that

$$\|v - v_\tau\|_{V^p} < \varepsilon. \tag{4.13}$$

Given $u \in U^p$ and $\varepsilon > 0$, there exists τ with

$$\|u - u_\tau\|_{U^p} < \varepsilon. \tag{4.14}$$

In particular, the step functions $\mathcal{S}$ are dense in V^p and $\mathcal{S}_{rc}$ is dense in U^p.

Proof. When we take the supremum over partitions for v_τ we may restrict to subsets of τ and the first statement becomes obvious. For U^p it suffices to check p-atoms a,

$$\|a_\tau\|_{U^p} \le 1.$$

Density of step functions in U^p follows from the atomic definition of the space: Let $u \in U^p$ and $\varepsilon > 0$. By definition, there exists a finite sum of atoms (which is a right continuous step function u_{step}) such that

$$\|u - u_{\text{step}}\|_{U^p} < \varepsilon/2.$$

Let τ be the partition associated to u_{step}. Then

$$\begin{aligned}\|u - u_\tau\|_{U^p} &\le \|u_{\text{step}} - u_\tau\|_{U^p} + \|u - u_{\text{step}}\|_{U^p}\\ &< \|(u_{\text{step}} - u)_\tau\|_{U^p} + \varepsilon/2\\ &< \varepsilon,\end{aligned}$$

which is the claim for U^p. Let $\tilde V^p$ be the closure of the step functions in V^p. Suppose there exists $v \in V^p$ with distance > 1 to $\tilde V^p$, and $\|v\|_{V^p} < 1 + \varepsilon$. Such a function exists when $\tilde V^p$ is not V^p. Let $D \subset U^q$ be the subset such $B(u, v) = 0$ whenever $u \in D$ and $v \in \tilde V^p$. Then u is continuous. Since the dual space of D is naturally given by $D^* = V^p/\tilde V^p$, and since v defines an element in D^* of norm > 1 there exists $u \in D$ with $B(u, v) = 1$, and a partition τ so that $\|u - u_\tau\|_{U^p} < \varepsilon$. However,

$$0 = B(u, v_\tau) = B(u_\tau, v) = B(u, v) + B(u_\tau - u, v) \ge 1 - \varepsilon(1 + \varepsilon),$$

which is a contradiction if $\varepsilon < \frac{1}{2}$. Hence the step functions are dense in V^p and, given $v \in V^p$ and $\varepsilon > 0$, there is a step function v_{step} with $\|v - v_{\text{step}}\|_{V^p} < \varepsilon$ and partition τ. Then

$$\begin{aligned}\|v - v_\tau\|_{V^p} &\le \|v_{\text{step}} - v_\tau\|_{V^p} + \|v - v_{\text{step}}\|_{V^p}\\ &< \|(v_{\text{step}} - v)_\tau\|_{V^p} + \varepsilon/2\\ &< \varepsilon,\end{aligned}$$

which is the density assertion. □

4.4 Convolution and regularization

Convolution by an L^1 function defines a bounded operator on U^p and V^p. Ruled functions are in L^∞ and hence the product of a function in U^p or V^p with an L^1 function can be integrated. In particular, the convolution of a ruled function and an L^1 function is well defined.

Lemma 4.17. *Let $a = -\infty$ and $b = \infty$, $v \in V^p$ and $\phi \in L^1$. Then*

$$\|v * \phi\|_{V^p(X)} \le \|\phi\|_{L^1} \|v\|_{V^p(X)}$$

and

$$\|u * \phi\|_{U^p(X)} \le \|\phi\|_{L^1} \|u\|_{U^p(X)}.$$

Proof. Let τ be a partition. It suffices to consider ϕ non-negative and with integral 1. Then, by convexity and Jensen's inequality,

$$\sum |\phi * v(t_{i+1}) - \phi * v(t_i)|^p \le \int |\phi(h)| \sum_i |v(t_{i+1} + h) - v(t_i + h)|^p dh \le \|v\|_{V^p}^p.$$

The statement for U^p follows by duality: We have

$$B(\phi * a, v) = B(a, \tilde{\phi} * v)$$

with $\tilde{\phi}(t) = \phi(-t)$. □

The first part of the next result it due to Hardy and Littlewood [13]. The Besov spaces of the lemma will be explained in the proof. We include the third statement for completeness, but it will not be used later on.

Lemma 4.18. *Let $I = \mathbb{R}$, $h > 0$ and $v \in V^p$. Then*

$$\|v(. + h) - v(.)\|_{L^p} \le (2h)^{1/p} \|v\|_{V^p}. \tag{4.15}$$

In particular, if $1 < p < \infty$,

$$\|v\|_{\dot{B}^{1/p}_{p,\infty}} \le c\|v\|_{V^p}$$

and

$$\|u\|_{U^p} \le c\|u\|_{\dot{B}^1_{p,1}}.$$

Proof. Let $I_j = [jh, (j+1)h]$ where

$$|v(t+h) - v(t)| \le \max\{ \sup_{[jh,(j+1)h]} v - \inf_{[(j+1)h,(j+2)h]} v, \sup_{[(j+1)h,(j+2)h]} v - \inf_{[jh,(j+1)h]} v\}.$$

For $\varepsilon > 0$ there exist two points $t_{j,0} \in I_j$ and $t_{j,1} \in I_{j+1}$ with

$$\sup_{t \in I_j} |v(t + h) - v(t)| \le (1 + \varepsilon)|v(t_{j+1}) - v(t_j)|.$$

For simplicity we assume that v is continuous, in which case we may choose $\varepsilon = 0$, which is the only use we will make of the continuity assumption. Hence

$$\begin{aligned}\int |v(t + h) - v(t)|^p dt &\le h\Big(\sum_j |v(t_{2j+1,1}) - v(t_{2j+1,0})|^p + \sum_j |v(t_{2j,1}) - v(t_{2j,0})|^p \Big) \\ &\le 2h\|v\|_{V^p}^p.\end{aligned}$$

All partial sums on the right-hand side are bounded by $2h\|v\|_{V^p}^p$ and hence the same is true for the sum. There are many equivalent norms on the homogeneous Besov space, one of them being

$$\|v\|_{\dot{B}^{1/p}_{p,\infty}} = \sup_{h>0} h^{-1/p}\|v(.+h)-v\|_{L^p},$$

and the bound follows from the estimate for the difference. The last statement follows by duality: the bilinear map

$$\dot{B}^{\frac{1}{p}}_{p,\infty} \times \dot{B}^{1-\frac{1}{p}}_{\frac{p}{p-1},1} \ni (f,g) \mapsto \int f dg$$

defines an isomorphism $\dot{B}^{\frac{1}{p}}_{p,\infty} \to \left(\dot{B}^{1-\frac{1}{p}}_{\frac{p}{p-1},1}\right)^*$. Here for $0<s<1$ and $1 \le q < \infty$

$$\|v\|_{\dot{B}^s_{p,q}} = \left(\int_0^\infty (h^{-1}\|v(.+h)-v\|_{L^p})^q \frac{dh}{h}\right)^{1/q}.$$

See Triebel [32] for the theory of these spaces. □

Let $\phi \in C_0^\infty$ with $\int \phi = 0$. Then it is an immediate consequence that

$$\begin{aligned} \|v*\phi\|_{L^p} &= \|\int (v(t+h)-v(t))\phi(h)dh\|_{L^p} \\ &\le \sup_h h^{-1/p}\|v(t+h)-v(t)\|_{L^p}\int h^{1/p}|\phi(h)|dh \\ &\le c\|v\|_{V^p} \end{aligned} \tag{4.16}$$

and, by duality, for $\phi \in C_0^\infty$,

$$\begin{aligned} \|u*\phi\|_{U^p} \le & \sup_{\|v\|_{V^q}\le 1} B(\phi*u,v) \\ = & \sup_{\|v\|_{V^q}\le 1} \int \phi' * uvdt \\ = & \sup_{\|v\|_{V^q}\le 1} \int u\tilde{\phi}'vdt \\ \le & \sup_{\|v\|_{V^q}\le 1} \|u\|_{L^p}\|\tilde{\phi}' * v\|_{L^q} \\ \le & C\|u\|_{L^p}. \end{aligned} \tag{4.17}$$

Clearly $C_0^\infty \subset V^1_{rc}$. Let $\tilde{V}^p \subset V^p$ be the closed subspace of functions with $f(t) = \frac{1}{2}(\lim_{h\to 0}(f(t+h)+f(t-h)))$. We consider functions on $\mathbb{R}$. If $v \in V^p$ is continuous, then

$$B(\phi_h * a, v) \to B(a,v) \text{ as } h \to 0$$

for all atoms a. Here $\phi \in L^1$ with $\int \phi dx = 1$ and $\phi_h(x) = h^{-1}\phi(x/h)$. If, moreover, ϕ is symmetric, then

$$\phi_h * v \to v$$

pointwise for all $v \in \tilde{V}^p$ and $B(\phi_h * u, v) = B(u, \phi_h * v)$ for all $u \in U^q$ and $v \in V^p$.

Lemma 4.19. *We have*

$$B(\phi_h * u, v) \to B(u, v)$$

for $u \in U^p(\mathbb{R})$ and $v \in V^q \cap C$, and

$$\phi_h * v \to v$$

in the weak * topology for $v \in \tilde{V}^p(\mathbb{R})$ for $1 \leq p < \infty$.

Proof. Only the last statement needs a proof. By definition and the pointwise convergence, $B(u, \phi_h * v) \to B(u, v)$ for all $u \in \mathcal{R}_{rc}$. This implies weak star convergence. □

4.5 More duality

The space $U^q \cap C(X)$ is a closed subspace of U^q.

Lemma 4.20. *The bilinear map B defines a surjective isometry*

$$\tilde{V}^p(X^*)_{rc} \to (U^q \cap C(X))^*, \quad \frac{1}{p} + \frac{1}{q} = 1, \quad 1 < p, q < \infty.$$

Proof. The kernel of the duality map composed with the inclusion $(U^p \cap C) \subset U^p$ consists exactly of those elements of V^q which are nonzero at at most countably many points. We claim that the duality map is an isometry. Let $v \in \tilde{V}^p$, and let a be an atom such that

$$\|v\|_{V^p} \leq (1 + \varepsilon) B(a, v).$$

If ϕ_h is a symmetric mollifier, then, if h is sufficiently small,

$$B(a, \phi_h * v) = B(\phi_h * a, v),$$

which shows that the duality map is an isometry.

It remains to prove surjectivity. Let $L : U^p \cap C(X) \to \mathbb{R}$ be linear and continuous. By the Hahn–Banach theorem, there is a extension with the same norm to U^p, and by duality there is $v \in V^q$ with $\|v\|_{V^q} = \|L\|$ and $L(u) = B(u, v)$ for all $u \in U^p$. Changing v on a countable set does not change the image in $(U^p \cap C(X))^*$, hence we may choose $v \in V^p_{rc}$. □

In the sequel we identify $u(a)$, resp. $u(b)$, with the limit from the right, resp. the left.

Lemma 4.21. *Let $u \in U^q$ and $v \in U^p$, $1/p + 1/q = 1$ and let $(t_j)_{j \to 1}$ be the points where both v and u have jumps, and denote the size of the jumps by $\Delta u(t_j)$. Then*

$$B(u,v) + B(v,u) = \sum_j \Delta u(t_j)\Delta v(t_j) + u(b)v(b). \tag{4.18}$$

Proof. The right-hand side of (4.18) is continuous with respect to $u \in V^q$ and $v \in V^p$, with the jump understood as the difference between the limit from the right and the left; the sum over the jumps to the power p is bounded by the V^p norm. The left-hand side is continuous with respect to $u \in U^q$ and $v \in V^p$, and it suffices to verify the formula for $u, v \in \mathcal{S}_{rc}$ with joint partition (where we add $t_0 = a$) $a = t_0 < t_1 < \cdots < t_N < b$. Then the statement follows from Lemma 4.2. □

Lemma 4.22. *Test functions are weak* dense in V^p.*

Proof. Step functions are dense in V^p, and it suffices to verify that step functions can be approximated by C_0^∞ functions in the weak$*$ sense. Moreover it suffices to consider test functions with a partition consisting of a single point, which we choose to be 0. Hence we reduce the problem to a proof for three functions. We fix $\phi \in C_0^\infty(\mathbb{R})$, identically 1 in $[-1,1]$, and $\eta \in C^\infty(\mathbb{R})$ supported in $(0,\infty)$ and identically 1 for $t \geq 1$. Then for $u \in \mathcal{S}_{rc}$ checking the definition shows

$$B(u, \phi(t/j)) \to B(u,1),$$

and with $v(t) = 0$ for $t \neq 0$ and $v(0) = 1$,

$$B(u, \phi(jt)) \to B(u,v),$$

and, with $v(t) = 0$ for $t \leq 0$ and 1 for $t > 0$,

$$B(u, \phi(t/j)\eta(jt)) \to B(u,v)$$

when $j \to \infty$. □

We define

$$V_C^q = \{v \in V^q(a,b) \cap C(a,b) : \lim_{t\to a} v(t) = \lim_{t \to b} v(t) = 0\}. \tag{4.19}$$

Lemma 4.23. *The map*

$$U^p(X^*) \to (V_C^q(X))^*,$$

$$u \mapsto (v \to B(u,v)),$$

is a surjective isometry.

Proof. By the duality estimates the duality map is defined, and it is an isometry since the space V^q_C is weak star dense in V^q. Let $L : V^q_C \to \mathbb{R}$ be linear and continuous. By Hahn–Banach, L can be extended to a continuous linear form on $\tilde{L} \in (V^q)^*$. Since $U^q \subset V^q_{rc}$, by an abuse of notation $L \in (U^q)^*$ and there exists $\tilde{u} \in V^p$ such that

$$B(w, -\tilde{u}) = \tilde{L}(w)$$

for all $w \in U^q$. We define (with $t\pm$ the limit from the left resp. the right)

$$u(t) = \tilde{u}(t+) - \tilde{u}(a).$$

Then $u \in \bigcap_{\tilde{p}>p} U^{\tilde{p}}$, and by Lemma 4.21, for all $v \in V^p_C$,

$$\begin{aligned} L(v) &= B(v, -\tilde{u}) \\ &= B(v, -u) - \tilde{u}(a)B(v, 1) \\ &= B(u, v) - \tilde{u}(a)(v(b)) \\ &= B(u, v), \end{aligned}$$

where we used that $v(b) = 0$ and that v is continuous.

For every partition we have $u_\tau \in U^p$, with

$$\|u_\tau\|_{U^p} \le \sup_{v \in V^p_C, \|v\|_{V^p} \le 1} B(v, u_\tau) = \sup_{\|v_\tau\|_{V^p}=1} L(v_\tau).$$

Since $u \in V^p_{rc}$, there is a sequence of partitions τ_i so that $u_{\tau_i} \to u \in V^p$ and hence the sequence converges uniformly. Thus for every step function v

$$B(u_{\tau_i}, v) \to B(u, v).$$

Since step functions are dense in V^q even

$$B(u_{\tau_i}, v) = B(u, v_{\tau_i}) \to B(u, v)$$

for all $v \in V^q$. Let U^{p**} be the bidual space of U^p , which we consider as isometric closed subspace of X^{**}. By an abuse of notation, we consider u as element of U^{p**}. Then

$$B(u_{\tau_i}, v) \to u(v)$$

for all $v \in V^q$ and the distance between u and U^p in U^{p**} is zero, and hence $u \in U^p$. $\square$

Corollary 4.24. *We have*

$$\|u\|_{U^p(X)} = \sup\{B(u, v) : v \in C_0^\infty(X),\ \|v\|_{V^q(X^*)} = 1\}.$$

and

$$\|v\|_{V^p_{rc}(X)} = \sup\{B(u, v) : u \in C_0^\infty,\ \|u\|_{U^q(X^*)}\} = 1\}.$$

Proof. Clearly C_0^∞ is weak dense in $V^p(X^*)$. This implies the first statement. Given $\varepsilon > 0$, there exists a q-atom in $U^q(X^*)$ with

$$B(a, v) \geq \|v\|_{V^p} - \varepsilon.$$

Since

$$B(x\chi_{[t,b)}, v) \to 0$$

as $t \to a$, we may assume that $a(b) = 0$. A standard regularization implies the full statement. $\square$

4.6 Consequences of Minkowski's inequality

For a Banach space Y, we denote by $L^p(Y)$ the weakly measurable maps with values in Y, for which the norm is p-integrable.

Lemma 4.25. *We have for $1 < p \leq q < \infty$,*

$$\|u\|_{L^q_x(U^p)} \leq \|u\|_{U^p(L^q_x)} \tag{4.20}$$

and

$$\|v\|_{V^p(L^q_x)} \leq \|v\|_{L^q_x(V^p)}. \tag{4.21}$$

Proof. It suffices to verify the first inequality for a p-atom

$$a(t,x) = \sum \chi_{[t_i,t_{i+1})}(t)\Phi_i(x)$$

with values in L^q. This is a function of x and t. Then $t \mapsto a(t,x)$ is a step function. Let

$$f(x) = \left(\sum_i |\Phi_i(x)|^p\right)^{1/p}.$$

Then

$$\begin{aligned}
\|a\|_{L^q_x(U^p)} &= \left(\int f(x)^q dx\right)^{1/q} \\
&= \left(\int \left(\sum_j |\Phi_j(x)|^p\right)^{q/p}\right)^{1/q} \\
&\leq \left(\sum_j \|\Phi_j\|^p_{L^q}\right)^{1/p} \\
&\leq 1,
\end{aligned}$$

where we use Minkowski's inequality for the first inequality. The argument for the space V^p is similar. $\square$

The argument works the same way if we consider Banach space valued functions in U^pL^q etc.

4.7 The bilinear form as integral

Here we consider scalar-valued functions. We consider functions on different intervals and denote the quadratic form on the interval (s,t) by $B_{s,t}$.

Definition 4.26. *Let $v \in V^p(a,c)$ and $u \in U^q(a,c)$. We define for $a \le s < t \le b$*

$$\int_s^t v du := B_{s,t}(u - u(s), v) + (u(t) - u(t-))v(t) \tag{4.22}$$

and

$$\begin{aligned}\int_s^t u dv := -\int_s^t v du + \sum_j (u(t_j) - u(t_j-))(v(t_j) - v(t_j-)) \\ + u(t-)v(t-) - u(s)v(s+) + u(t)(v(t+) - v(t-)) + v(t)(u(t) - u(t-)),\end{aligned} \tag{4.23}$$

where the sum is taken over all joint jumps in (s,t).

The second definition is partly motivated by

1. The integration by parts formula (4.18). It should reduce to integration by parts if $v \in U^q$, and if there are no jumps at t.
2. The desire to have a certain symmetry with respect to time reversal if v is continuous from the left and u is continuous from the right. In general the notation is ambiguous, and one has to pay attention whether the integrand is supposed to be in V^p or U^p.
3. We want the integral to be additive in the interval.

Lemma 4.27. *For $u \in U^q$ and $v \in V^p$, $1/p + 1/q = 1$ and $a < b < c$, we have*

$$\int_a^c v du = \int_a^b v du + \int_b^c v du$$

and

$$\int_a^c u dv = \int_a^b u dv + \int_b^c u dv.$$

With the obvious notation,

$$\Big\| \int_a^t u dv \Big\|_{V^p(a,t)} \le \|u\|_{U^q(a,t)} \|v\|_{V^p(a,t)} \tag{4.24}$$

and

$$\Big\| \int_a^t v du \Big\|_{U^q(a,t)} \le \|u\|_{U^q(a,t)} \|v\|_{V^p(a,t)}. \tag{4.25}$$

Proof. It suffices to check the first formula for atoms u. Suppose that $t_j < b \le t_{j+1}$. On both sides we have a sum of terms of the form

$$v(t_{j+1})(u(t_{j+1}) - u(t_j)).$$

For the second formula we see from the definition that

$$\int_a^c udv = \int_a^b udv + \int_b^c udv,$$

where we have to check the contribution at $t = b$.

Formally, for smooth functions

$$\begin{aligned} B(\int_a^t vdu, w) &= \int_a^b w(t)v(t)u'(t)dt \\ &= B(u, vw) \\ &\le \|vw\|_{V^q}\|u\|_{U^p} \\ &\le 2|v\|_{V^q}\|w\|_{V^q}\|u\|_{U^p}, \end{aligned} \tag{4.26}$$

which formally implies (4.25).

For a rigorous proof we verify the formula in the case when u is an atom, and v and w are step functions; with a common partition the proof is done for general functions. Then $\int_a^t vdu$ is a right continuous step function and

$$\sum_j (v(t_j)(u(t_j) - u(t_{j-1}))w(t_j) = \sum_j [v(t_{j+1})w(t_{j+1}) - v(t_j)w(t_j)]u(t_j),$$

where we neglect the boundary terms. We apply Hölder's inequality to bound the expression by

$$\Big(\sum |v(t_{j+1})w(t_{j+1}) - v(t_j)w(t_j)|^q\Big)^{1/q} \Big(\sum |u(t_j)|^p\Big)^{1/p}.$$

Again formally for smooth functions

$$\begin{aligned} B(w, \int_a^t udv) &= -\int_a^b vwu'dt + (w(b) - w(a))\int_a^b uv'dt \\ &= \int_a^b v(uw)'dt - (w(b) - w(a))\int_a^b vu'dt \\ &\quad - u(b)v(b)w(b) + u(a)v(a)w(a) \\ &\quad + (w(b) - w(a))(u(b)v(b) - u(a)v(a)) \\ &= B(uw, v) - (w(b) - w(a))B(v, u) \end{aligned} \tag{4.27}$$

if $u(a) = w(a) = 0$. This formally implies (4.25). For a rigorous proof we apply integration by parts several times. First

$$\begin{aligned}\int_{t-}^{t+} udv &= (u(t)-u(t-))(v(t)-v(t-)) + u(t)v(t+) - u(t-)v(t-) \\ &\quad - v(t)(u(t)-u(t-)) \\ &= u(t)(v(t+)-v(t-))\end{aligned}$$

and

$$\int_{t}^{t+} udv = u(t)(v(t+)-v(t)),$$

and hence the l^p sum over the jumps is bounded. Thus the bound reduces to the bound for

$$B(w, \int_a^t vdu),$$

and by the same token to

$$B(\int_a^t vdu, w)$$

which we have proven above. □

Sometimes it is convenient to have a notation for spaces of derivatives of functions in U^p, resp. V^p.

Definition 4.28. *We define dU^p as the space of all distributions f for which there exists an antiderivative in U^p, equipped with the norm in U^p. Similarly, let dV^p be the space of all distributions which have an antiderivative in $\tilde{V}^p_{rc}$, equipped with the obvious norm.*

4.8 Differential equations with rough paths

This type of study was initiated by Lyons [23]. We will only scratch the surface. We observe that the duality mapping extends the Young integral.

We consider the differential equation

$$\dot{y} = F(y,x)\dot{x}, \quad y(0) = y_0,$$

where $x \in U^2$ and F is a bounded Lipschitz function continuously Fréchet differentiable with respect to y, and d_yF is uniformly Lipschitz continuous. To simplify, we denote the bound for F by $\|F\|_{\sup}$, the Lipschitz bound with respect to y by $\|D_YF\|_{\sup}$, and the homogeneous Hölder bound with respect to y by $\|F\|_{C^s(Y)}$.

Suppose that y is a solution, i.e.,

$$y(t) = y(a) + \int_a^t F(y,x)dx.$$

Then, by (4.25),

$$\begin{aligned}\|y(t)-y(a)\|_{U^2} &\le \|F(y,x)\|_{V^2}\|x\|_{U^2}\\ &\le (\|F\|_{\sup}+\|D_yF\|_{\sup}\|y\|_{V^2})+\|D_xF\|_{\sup}\|x\|_{V^2})\|x\|_{U^2}.\end{aligned} \tag{4.28}$$

It is trivial that there is a unique solution if x is a step function in $\mathcal{S}_{rc}$ – for that we consider a finite number of differences. We shall construct a solution to the initial value problem for $\|x\|_{U^p}$ small. This implies existence of a unique solution, since we may first approximate x by a step function, and then solve the differential equation on each of the intervals of the step function.

We want to construct a solution as a fixed point of

$$y(t) = y_0 + \int_0^t F(y(s),x(s))\dot{x}ds.$$

We claim that there is a unique solution y with $y - y(a) \in U^2$ provided

$$\|x\|_{U^2} < \varepsilon,$$

with ε sufficiently small. Let

$$y(t) = y_0 + \int_0^t F(\tilde{y}(s),x(s))\dot{x}ds.$$

Now, by (4.28),

$$\|y-y(a)\|_{U^2} \le (\|F\|_{\sup}+\|D_yF\|_{\sup}\|\tilde{y}\|_{V^2}+\|D_xF\|_{\sup}\|x\|_{V^2})\|x\|_{U^2}$$

and we obtain a uniform bound R on the iteration provided $\|D_yF\|_{\sup}\|x\|_{U^2} \le \frac{1}{4}$. If $\tilde{y}_1, \tilde{y}_2 \in U^2$ and y_i is defined by the Young integral above, we get, by considering scalar-valued functions to simplify the notation,

$$\begin{aligned}\|y_2-y_1\|_{U^2} &\le 2\|F(\tilde{y}_2,x)-F(\tilde{y}_1,x)\|_{V^2}\|x\|_{U^2}\\ &\le \Big(\|D_yF\|_{\sup}\|\tilde{y}_2-\tilde{y}_1\|_{V^2}+\|D^2_{yy}F\|_{\sup}\|\tilde{y}_2-\tilde{y}_1\|_{\sup}\|\tilde{y}_2-\tilde{y}_1\|_{V^2}\\ &\qquad +\|D^2_{yx}F\|_{\sup}\|\tilde{y}_1-\tilde{y}_1\|_{\sup}\|x\|_{U^2}\Big)\|x\|_{U^2}.\end{aligned}$$

We easily construct a unique solution by a standard contraction argument provided

$$\left(\|D_yF\|_{\sup}+\|D^2_{yy}F\|_{\sup}\|R+\|D^2_{xy}\|x\|_{U^2}\right)\|x\|_{U^2} < \frac{1}{2},$$

where R is the uniform bound from above.

The modifications for U^p, $p<2$, are as follows. The differentiability requirements on F are weaker: Let $1<p<2$ and $\frac{1}{p}+\frac{1}{q}=1$. The apriori estimate requires few changes and we concentrate on the contraction, for which we consider

$$\begin{aligned}\|F(\tilde{y}_2,x)-F(\tilde{y}_1,x)\|_{V^q} \le \|D_yF\|_{\sup}\|\tilde{y}_2-\tilde{y}_1\|_{V^q}+\|D_yF\|_{C^{p/q}}\Big(\|y_2-\tilde{y}_1\|^{p/q}_{V^p}\\ +\|x\|^{p/q}_{V^p}\Big)\|\tilde{y}_2-\tilde{y}_1\|_{\sup}.\end{aligned}$$

We recall that $p-1=p/q$. We obtain the contraction as above.

Theorem 4.29. *Let $1 < p \le 2$, $F : X \times Y \to Y$ be bounded, uniformly Lipschitz continuous, Frechet differentiable with respect to X and Y, and such that dF is Hölder continuous with respect to y with Hölder exponent $p-1$. We study*

$$dy = F(x,y)dx, \quad y(a) = y_0.$$

Then there exists a unique solution $y \in U^p(Y)$ if $x \in U^p$ if $1 \le p \le 2$ and $y \in V^p$ if $x \in V^p$ and dF is Hölder continuous with exponent $p-1 < s \le 1$.

4.9 The Brownian motion

The Brownian motion is almost surely in V^p for $p > 2$. We denote by $B_t(\omega)$ the path of the Brownian motion as a function of t and the element of the probability space ω. If the Brownian motion would be in U^2 with positive probability we could solve stochastic differential equations in a pointwise sense. The 2-variation however is almost certainly infinite.

The regularity of the Brownian motion is characterized by the following fairly sharp result of Taylor [31] , see also [8].

Theorem 4.30. *Let*

$$\psi_{2,1}(h) = \begin{cases} h^2, & \text{if } h \ge e^{-e}, \\ \frac{h^2}{\ln\ln(1/h)}, & \text{if } h < e^{-e}. \end{cases}$$

There exists $\eta > 0$ such that

$$\mathbb{E}\Big(\exp\Big(\frac{\eta}{T}\|B\|^2_{\psi_{2,1};[0,T]}\Big)\Big) < \infty,$$

where

$$\|B\|_{\psi_{2,1};[0,T]} = \inf\{M > 0 : \sup_{\tau} \sum \psi_{2,1}(|B_{t_{i+1}} - B_{t_i}|/M) \le 1\}.$$

Moreover, if

$$\frac{h^2}{\psi(h)\ln\ln(1/h)} \to 0 \text{ as } h \to 0,$$

then

$$\sup_{\tau_T} \sum \psi(|B_{t_{i+1}} - B_{t_i}|) = \infty.$$

See Theorem 13.15 and Theorem 13.69 in [8]. This result deviates from the V^p spaces by an iterated logarithm.

Let (Ω, μ) be a probability space with a filtration μ_t, $t \in \mathbb{R}$, $f \in L^p$ and $f_t = \mathbb{E}(f, \mu_t)$. Then

$$\|f_t\|_{L^p(\Omega, V^2_w)} \le c_p \|f\|_{L^p} \tag{4.29}$$

is a consequence of Doob's oscillation lemma for martingales [25], see also Bourgain's proof of p-variation estimate [2]. A weaker version is due to Lepingle [21].

For the Brownian motion B_t we obtain

Theorem 4.31.

$$\|B_t\|_{L^p(\Omega,V^2_w([0,1)))} \le c_p.$$

This has been a motivation to introduce V^p_w.

4.10 Adapted function spaces

Given a distribution T, we want to construct an element in U^p or V^p which has T as derivative. This is done in the next lemma. Again $\frac{1}{p} + \frac{1}{q} = 1$.

Lemma 4.32. *Suppose that T is a distribution supported in $[0,\infty)$ so that*

$$\sup\{T(\phi) : \phi \in C_0^\infty,\ \|\phi\|_{U^q} \le 1\} = C_1 < \infty.$$

Then there exists a unique $v \in V^p_{rc}$ with

$$T(\phi) = -B(\phi, v),$$

$$C_1 \le \|v\|_{V^p} \le 2C_1,$$

and $v_t = T$ in the sense of distributions. Suppose that T is a distribution supported in $[0,\infty)$ so that

$$\sup\{T(\phi) : \phi \in C_0^\infty, \|\phi\|_{V^q} \le 1\} = C_2 < \infty.$$

Then there exists a unique $u \in U^p$ with

$$T(\phi) = B(u, \phi),$$

$$\|u\|_{U^p} = C_2,$$

and $u_t = T$ in the sense of distributions.

Proof. There exists a unique distribution V supported in $[0,\infty)$ with $\partial_t V = T$ which is defined as follows. We fix a function $\eta \in C^\infty$ supported in $[-2,\infty)$ and identically 1 in $[-1,\infty)$. Then

$$V(\phi) := T(\eta \int_t^\infty \phi),$$

which does not depend on the choice of η. Then

$$V(\partial_t \phi) = -T(\eta\phi) = -T(\phi)$$

by definition. The difference of two such distributions has zero derivative, hence it is constant, and by the assumption on the support it is unique.

Next we choose a function $\psi \in C_0^\infty(\mathbb{R})$ supported in $(-1,1)$ and satisfying $\int \psi dx = 1$, and define for $h > 0$ and $s \in \mathbb{R}$

$$\phi(t) = \eta(t) h^{-1} \int_t^\infty \psi((x-s)/h) dx.$$

Then by the support property,

$$V(h^{-1}\psi((t-s)/h)) = V(\partial_t \phi) = T(\phi)$$

and, since, for nonnegative ψ

$$\|\phi\|_{U^q} \leq 1$$

and hence

$$|V(h^{-1}\psi((t-s)/h)| \leq C_1,$$

we have

$$\sup_s |V * h^{-1}\psi((.-s)/h)| \leq C_1.$$

Thus, there exists a bounded and measurable function v with

$$V(\phi) = \int v \phi dt,$$

and moreover v is supported in $[t_0, \infty)$. At Lebesgue points

$$|V(h^{-1}\psi((t-s)/h))| = h^{-1} \int v(t)\psi((t-s)/h) dt \to v(s)$$

as $h \to 0$. Similarly, if τ is partition for which all points are Lebesgue points, and arguing as for duality, we see that

$$\Big(\sum |v(t_j) - v(t_{j-1})|^p \Big)^{\frac{1}{p}} \leq C_1.$$

In particular, left and right limits at $t \in \mathbb{R}$ exist if we restrict the approach to Lebesgue points. Hence we may assume that v is a right continuous ruled function, supported in $[0, \infty)$. But then the very same argument shows (since we have to include the supremum in the norm) that

$$\|v\|_{V^p_{rc}} \leq 2C.$$

By construction, the weak derivative of v is T. We conclude that T defines an element of $(U^q)^*$ which is represented by some function which is v. This completes the argument in this case.

In the second part we construct the function u as above. Then

$$T(\phi) = -\int u \partial_t \phi = -B(\phi, u) = B(u, \phi).$$

In particular, for every partition, since C_0^∞ is weak star dense,

$$\|u_\tau\|_{U^p} \le C$$

We conclude, as for the duality, that

$$\|u\|_{U^p} \le C. \qquad \square$$

We observe that only obvious changes are required when we consider Hilbert spaces valued functions, and if we replace the product by the inner product.

We briefly survey constructions going back to Bourgain, which have become standard. The following situation will be of particular interest. Let $t \to S(t)$ be a continuous unitary group on a Hilbert space H. We define U^p_S and V^p_S by

$$\|v\|_{V^p_S(H)} = \|S(-t)v(t)\|_{V^p(H)},$$

or, to put it differently, we say that $v \in V^p_S$ if $S(-t)v \in V^p$. Similarly we define U^p_S. Alternatively, we could define U^p_S by U^p_S atoms. Such an atom is given by a partition $t_1 < t_2 < \cdots < t_n$ and n elements $\phi_j \in H$, with $\sum \|\phi_j\|^p \le 1$, and $a(t) = 0$ if $t < t_1$, and $a(t) = S(t-t_j)\phi_j$ if $t_j \le t < t_{j+1}$, with the obvious modification if $t \ge t_n$.

By Stone's theorem, unitary groups are in one-one correspondence with self-adjoint operators, in the sense that

$$i\partial_t u = AU$$

with a self-adjoint operator A defines unitary a group $S(t)$ and vice versa. At least formally

$$i\partial_t(S(-t)u(t)) = S(-t)(i\partial_t u - Au),$$

and hence the duality assertion is

$$\|u\|_{U^q_S} = \sup_{\|v\|_{V^p_S}\le 1} B(S(-t)u(t), S(-t)v(t)).$$

Now suppose, that again formally

$$i\partial_t u + Au = f.$$

Then, if we use Duhamel's formula, the solution is given by

$$u(t) = \int_{-\infty}^t S(t-s)f(s)ds.$$

A related construction goes back to Bourgain. He defines

$$\|u\|_{X^{0,b}_S} = \|S(-t)u(t)\|_{H^bL^2}, \tag{4.30}$$

where the Sobolev space H^b is defined by the Fourier transform

$$\|f\|_{H^b} = \|(1+|\tau|^2)^{b/2}\hat{f}\|_{L^2}.$$

Clearly,

$$X_S^{0,b} \subset X_S^{0,b'}$$

whenever $b \geq b'$. We may use a Besov refinement of the right-hand side of (4.30), i.e.,

$$\|u\|_{\dot{X}^{s,b,q}} = \left(\sum_{N\in 2^{\mathbb{Z}}} N^{sq}\|u_N\|_{H^b(L^2)}^q\right)^{1/q},$$

where we choose a disjoint partition $A_N = \{(\tau,\xi) : 2^N \leq |\tau+\phi(\xi)| \leq 2^{1+N}\}$ and define u_N by the Fourier multiplication by the characteristic function of A_N.

Then

$$\dot{X}_S^{0,\frac{1}{2},1} \subset U_S^2 \subset V_{S,rc}^2 \subset \dot{X}^{0,\frac{1}{2},\infty}$$

follows from Lemma 4.18.

There is an obvious generalization to the case of time dependent operators $A(t)$. Definitions are simple, but this often leads to technical questions.

Now consider A given by a Fourier multiplier $-\phi(\xi)$.

$$\mathcal{F}_{t,x}(S(-t)u)(\tau,\xi) = \mathcal{F}_t e^{-it\phi(\xi)}\hat{u}(t,\xi) = \mathcal{F}_{t,x}u(\tau+\phi(\xi),\xi),$$

and hence by the Plancherel formula and a translation in the τ variable,

$$\|u\|_{X^{0,b}} = \|(1+\tau^2)^{b/2}\mathcal{F}_{t,x}(u)(\tau+t\phi(\xi),\xi)\|_{L^2} = \|(1+(\tau-\phi(\xi))^2)^{b/2}\mathcal{F}_{t,x}(u)\|_{L^2}.$$

4.10.1 Strichartz estimates

We want to use this construction for dispersive equations. There A is often defined by a Fourier multiplier, most often even by a partial differential operator with constant coefficients.

We consider the Schrödinger equation

$$i\partial_t u + \Delta u = 0 \quad \text{in } [0,\infty),$$

$$u(0) = u_0 \quad \text{on } \mathbb{R}^d.$$

Let $u(t) = 0$ for $t < 0$ and the solution otherwise. Then

$$\|u\|_{U_S^1} = \|u_0\|_{L^2(\mathbb{R}^d)}.$$

One of the Strichartz estimates states that

$$\|u\|_{L_t^p L_x^q} \leq \|u_0\|_{L^2} \tag{4.31}$$

whenever

$$\frac{2}{p} + \frac{d}{q} = \frac{d}{2}, \qquad 2 \le p, q, \quad (p, q, d) \ne (2, \infty, 2).$$

We claim that this implies

$$\|u\|_{L^p_t L^q_x} \le c\|u\|_{U^p}.$$

It suffices to verify this if $S(-t)u$ is an atom with partition $(t_1, t_2, \dots, t_n)$. Then, with $t_{n+1} = \infty$, by the Strichartz estimate,

$$\|u\|_{L^p_t((t_j,t_{j+1});L^q_x)} \le c\|u(t_j)\|_{L^2}.$$

We raise this to the pth power, and sum over j. Then

$$\|u\|_{L^q_t L^q_x} \le c\left(\sum \|u(t_j)\|^p_{L^2}\right)^{1/p} \le c,$$

since $S(-t)u$ is a p-atom.

Consider $v(t) = \int_{-\infty}^t S(t-s)f(s)ds$ and let $\tau = (t_j)$ be a partition. Then

$$v(t_j) - S(t_j - t_{j-1})v(t_{j-1}) = \int_{t_{j-1}}^{t_j} S(t_j - t)f(t)dt$$

and by the Strichartz estimate,

$$\|S(-t_j)v(t_j) - S(-t_{j-1})v(t_{j-1})\|_{L^2} \le c\|f\|_{L^{p'}_t L^{q'}_x}$$

and

$$t \mapsto S(-t)v(t)$$

is continuous.

We take the power p' and sum over j to reach the conclusion

$$\|v\|_{V^{p'}_S} \le c\|f\|_{L^{q'}_t L^{p'}_x}$$

This implies the dual estimate to (4.31). If $p > 2$ we can combine the estimates with an embedding to obtain the full Strichartz estimate. In particular we arrive at the asymmetric improvement of the Strichartz estimate:

$$\|u\|_{L^\infty(L^2)} + \|u\|_{L^{q_0,p_0}} \le c\left(\|u_0\|_{L^2} + \|f\|_{L^{q_1'p_1'}}\right)$$

if both (q_1, p_1) and (q_0, p_0) are Strichartz pairs, but not necessarily the same ones.

We prove this estimate over the interval $(0, \infty)$ and extend u by 0 to negative t. Then

$$\|u\|_{L^\infty(L^2)} + \|u\|_{L^{q_0}_t L^{p_0}_x} \le c\|u\|_{U^{p_0}} \le c\|u\|_{V^{p_1'}} \le c\|u_0\|_{L^2} + \|f\|_{L^{q_1'}_t L^{p_1'}_x}.$$

Lemma 4.33. *The following estimates hold for Strichartz pairs:*

$$\|u\|_{L^q_t L^p_x} \le c\|u\|_{U^p}$$

and

$$\left\|S(t)u_0 + \int_0^t S(t-s)f(s)ds\right\|_{V^{p'}} \le c(\|u_0\|_{L^2} + \|f\|_{L^{q'}_t L^{p'}_x}).$$

4.10.2 Estimates by duality

We return to duality questions and calculate formally

$$\begin{aligned}
\|u\|_{U^q_S} &= \sup_{\|v\|_{V^p_S}\le 1} |B(S(-t)u(t), S(-t)v(t))| \\
&= \sup_{\|v\|_{V^p_S}\le 1} \left| \int_{\mathbb{R}} \langle \partial_t S(-t)u(t), S(-t)v(t)\rangle dt \right| \\
&= \sup_{\|v\|_{V^p_S}\le 1} |-i\langle S(-t)(i\partial_t u - Au), S(-t)v\rangle dt| \\
&= \sup_{\|v\|_{V^p_S}\le 1} \int_{\mathbb{R}} \langle f, v\rangle dt,
\end{aligned} \tag{4.32}$$

with a similar statement for V^p_S. This observation will be crucial for nonlinear dispersive equations.

Lemma 4.34. *Let $\phi \in C^\infty(\mathbb{R}^d)$ be a real polynomial and let S be the unitary group defined by the Fourier multiplier $e^{it\phi(\xi)}$. Let $1 < p, q < \infty$ and $\frac{1}{p} + \frac{1}{q} = 1$. Let T be a tempered distribution in $(a,b) \times \mathbb{R}^d$ which satisfies*

$$\sup\{|T(\bar u)| : u \in C_0^\infty((a,b)\times\mathbb{R}^d), \|u\|_{U^p_S} \le 1\} = C_1 < \infty$$

Then there is a unique $v \in V^q_{S,rc}(a,b)$ with

$$T(\bar u) = \int_{(a,b)\times\mathbb{R}} v\, \overline{iu_t + \phi(D)u}\, dx dt$$

and $\|v\|_{V^q} = C_1$. Let T be a distribution in space time which satisfies

$$\sup\{|T(\bar v)| : v \in C_0^\infty((a,b)\times\mathbb{R}^d),\ \|v\|_{V^p_S}\} = C_2 < \infty$$

Then there is a unique $u \in U^q_S$ with

$$T(\bar v) = \int_{(a,b)\times\mathbb{R}} u\, \overline{iv_t + \phi(D)v}\, dx dt$$

and $\|u\|_{U^q} = C_2$.

Proof. This is a consequence of Lemma 4.32. □

The theorem implies existence of a weak solution to

$$i\partial_t u + \phi(D)u = f, \quad u(a) = 0,$$

together with an estimate for u.

4.10.3 High modulation estimates

We denote by $f(D)$ the Fourier multiplier defined by a function f. Let

$$f = 1 - \chi(\tau/\Lambda),$$

where τ is the Fourier variable corresponding to t and χ is an approximate characteristic function, i.e., χ is supported on a ball of radius 2, and identically 1 on a ball of radius 1.

Lemma 4.35. *The following estimate holds:*

$$\|f(D)v\|_{L^2} \le c\Lambda^{-1/2}\|v\|_{V^2}.$$

Suppose the group $S(t)$ is defined by the Fourier multiplier $e^{it\phi(\xi)}$. Then, with

$$f(D) = 1 - \chi((\tau - \phi(\xi))/\Lambda),$$

$$\|f(D)u\|_{L^2} \le c\Lambda^{-1/2}\|v\|_{V^2_S}.$$

Proof. We have

$$\mathcal{F}_t(e^{-it\phi(\xi)}\hat{u}(t,\xi)) = \mathcal{F}_{x,t}u(\tau - \phi(\xi),\xi)$$

and the second claim follows from the first one. Let

$$g = \mathcal{F}^{-1}\chi(\xi/\Lambda).$$

Then

$$g(t) = \Lambda^{-1}(\mathcal{F}^{-1}\chi)(\Lambda\xi)$$

and

$$\begin{aligned}
&\left\| \int (v(t+h) - v(t))g(h)dh \right\|_{L^2} \\
&\le \sup_h |h|^{-1/2}\|v(t+h) - v(t)\|_{L^2} \int |h|^{1/2}\Lambda^{-1/2}|\mathcal{F}^{-1}\chi(h\Lambda)|dh \\
&\le c\|u\|_{V^2}\Lambda^{-1/2} \int |h|^{1/2}|\mathcal{F}^{-1}\chi|dh.
\end{aligned}$$

□

Chapter 5

Convolution of measures on hypersurfaces, bilinear estimates, and local smoothing

The contents of this section developed in discussions with S. Herr, T. Schottdorf and J. Li. Related results have been proven by Foschi and Klainerman [7] and by Grünrock for the Airy equation [10] and the Kadomtsev–Petviashvili II equation [11]. The bilinear estimates for the Kadomtsev–Petviashvili equation have been influenced by the careful work of M. Hadac. Bilinear estimates are standard tools in dispersive equations. Here we attempt to streamline arguments and sharpen the results. In particular, the bilinear estimates for the KP II seem to be new.

The transformation formula for a diffeomorphism $\phi : U \to V$ $U, V \subset \mathbb{R}^d$, states

$$\int_V f dm^d = \int_U f \circ \phi |\det D\phi| dm^d.$$

Its relative, the area formula for $n \geq d$,

$$\phi : U \to S \subset \mathbb{R}^n,$$

ϕ continuously differentiable and injective, reads

$$\int_S f d\mathcal{H}^d = \int_U f \circ \phi (\det D\phi^T D\phi)^{1/2} dm^d,$$

where $\mathcal{H}^d$ denotes the Hausdorff measure. The coarea formula deals with the opposite situation $d \geq n$ and

$$\phi : U \to V \subset \mathbb{R}^n$$

surjective. It states that for $f : U \to \mathbb{R}$ measurable

$$\int_V \int_{\phi^{-1}(y)} f d\mathcal{H}^{d-n} dm^n(y) = \int_U f \det(D\phi D\phi^T)^{1/2} dm^d.$$

H. Koch et al., *Dispersive Equations and Nonlinear Waves: Generalized Korteweg–de Vries, Nonlinear Schrödinger, Wave and Schrödinger Maps*, Oberwolfach Seminars 45,
DOI 10.1007/978-3-0348-0736-4_5,

Often it is useful to write it in the form

$$\int_V \int_{\phi^{-1}(y)} \det(D\phi(x)D\phi^T(x))^{-1/2} f(x) d\mathcal{H}^{d-n}(x) dm^n(y) = \int_U f dm^d. \tag{5.1}$$

The Fourier transform maps a product into a convolution, and vice versa. Let Σ_1 and Σ_2 be two $(d-1)$-dimensional hypersurfaces in $\mathbb{R}^d$ such that for all $x_i \in \Sigma_i$ the tangent spaces of Σ_i at x_i are transverse, for $i = 1, 2$.

Let Σ_1 and Σ_2 be non degenerate level sets of functions ϕ_1 and ϕ_2. Let h be a continuous function. Then, by the coarea formula,

$$\int_{\mathbb{R}^d} f(x) h \circ \phi_1(x) dm^d(x) = \int_{\mathbb{R}} h(s) \int_{\phi_1^{-1}(s)} f(x) |\nabla\phi_1|^{-1}(x) d\mathcal{H}^{d-1}(x) ds.$$

This motivates the notation

$$\delta_\phi = |\nabla\phi|^{-1} d\mathcal{H}^{d-1}\Big|_{\phi=0}.$$

We study the convolution of two measures supported on the hypersurfaces Σ_1 and Σ_2.

Theorem 5.1. *Let $\Sigma_i \subset \mathbb{R}^d$ be hypersurfaces and ϕ_i be as above, and f_i be square integrable functions on Σ_i with respect to δ_{ϕ_i}. Then*

$$\|f_1\delta_{\phi_1} * f_2\delta_{\phi_2}\|_{L^2(\mathbb{R}^d)} \le L \|f_1|\nabla\phi_1|^{-1/2}\|_{L^2(\Sigma_1)} \|f_2|\nabla\phi_2|^{-1/2}\|_{L^2(\Sigma_2)},$$

where with the notation $\Sigma(x,y) = \{y + \Gamma_1\} \cap \{x + \Gamma_2\}$,

$$L = \sup_{x\in\Sigma_1, y\in\Sigma_2} L(x,y),$$

and where $L(x,y)$ is the square root of

$$\int_{\Sigma(x,y)} \left[|\nabla\phi_1(z-y)|^2 |\nabla\phi_2(z-x)|^2 - \langle\nabla\phi_1(z-x), \nabla\phi_2(z-y)\rangle^2 \right]^{-1/2} d\mathcal{H}^{d-2}.$$

Proof. Let f_i be measurable functions in a neighborhood of Σ_i, let h be continuous and non-negative, and $g_i = h \circ \phi_i$. Then, by Cauchy–Schwarz and Fubini,

$$\begin{aligned}
&\|f_1 g_1 * f_2 g_2\|_{L^2}^2 \\
&= \int_{\mathbb{R}^d} \left(\int_{\mathbb{R}^d} f_1(x) g_1^{\frac{1}{2}}(x) g_2^{\frac{1}{2}}(z-x) f_2(z-x) g_1^{\frac{1}{2}}(x) g_2^{\frac{1}{2}}(z-x) dm^d(x) \right)^2 dm^d(z) \\
&\le \int_{\mathbb{R}^d} \int_{\mathbb{R}^d} |f_1(x)|^2 g_1(x) g_2(z-x) dm^d(x) \int_{\mathbb{R}^d} |f_2(y)|^2 g_2(y) g_1(z-y) dm^d(y) dm^d(z) \\
&= \int_{\mathbb{R}^{2d}} |f_1(x)|^2 g_1(x) |f_2(y)|^2 g_2(y) \int g_2(z-x) g_1(z-y) dm^d(z) dm^{2d}(x,y).
\end{aligned}$$

By the coarea formula,

$$\int g_2(z-x)g_1(z-y)dm^d = \int_{\mathbb{R}^2} h(s)h(t)I(s,t)dsdt,$$

where, with the notation

$$\Sigma_{s,t} = \{z : \phi_1(y+z) = s, \phi_2(x+z) = t\}$$

and

$$\rho(s,t,z) = \left||\nabla\phi_1(z-y)|^2|\nabla\phi_2(z-x)|^2 - (\nabla\phi_1(z-y)\cdot\nabla\phi_2(z-x))^2\right|^{-1/2},$$

we have

$$I(s,t) = \int_{\Sigma_{s,t}} \rho(s,t,z)d\mathcal{H}^{d-2}(z).$$

Here we suppress the dependence on x and y, but we set

$$\gamma(x,y) = I(0,0).$$

Using again the coarea formula, we get

$$\int_{\mathbb{R}^d} |f_1(x)|^2 g_1(x)dm^d(x) = \int_{\mathbb{R}} h(s)\int_{\phi_1^{-1}(s)} |f_1(x)|^2\,|\nabla\phi_1(x)|^{-1}d\mathcal{H}^{d-1}(x)ds.$$

There is a similar formula for the second integral. We assume that f_i is continuous and choose a Dirac sequence for h to obtain the estimate. The statement for measurable functions on the surfaces follows by a standard approximation argument. $\square$

Using the coarea formula we obtain a more explicit formula for the convolution:

$$\begin{aligned} f_1h\circ\phi_1 * f_2h\circ\phi_2(z) &= \int (f_1h\circ\phi_1)(z-y)(f_2h\circ\phi_2)(y)dm^d(y) \\ &= \int\limits_{\mathbb{R}}\int\limits_{\mathbb{R}} h(s)h(t)\int_{\Sigma(s,t)} f_1(z-y)f_2(y)\rho(s,t,z)d\mathcal{H}^{d-2}(y)dsdt. \end{aligned}$$

Hence,

$$\begin{aligned} &f_1\delta_{\phi_1} * f_2\delta_{\phi_2}(x) \\ &= \int_{\Gamma_1\cap(x-\Gamma_2)} \left||\nabla\phi_1(y)|^2|\nabla\phi_2(x-y)|^2 - (\nabla\phi_1(y)\cdot\nabla\phi_2(x-y))^2\right|^{-1/2} d\mathcal{H}^{d-2}(y). \end{aligned} \tag{5.2}$$

The following is a trivial and useful improvement of the convolution estimate of Theorem 5.1:

$$\begin{aligned}\Big\| \int_{\Gamma_1\cap(z-\Gamma_2)} \gamma^{-1/2}(x,y)\rho(0,0,z) f_1(x) f_2(y) d\mathcal{H}^{d-2} \Big\|_{L^2} \\ \le \|f_1\|_{L^2(\Sigma_1,\delta_{\phi_1})} \|f_2\|_{L^2(\Sigma_2,\delta_{\phi_2})}.\end{aligned} \tag{5.3}$$

It follows from the same arguments as Theorem 5.1. Here $L^2(\Sigma_i,\delta_{\phi_i})$ denotes the space of square integrable functions on the hypersurface with respect to the measure δ_{ϕ_i}.

We use the convolution estimate to bound products of solutions to dispersive equations. Consider

$$iu_t + \psi(D)u = 0,$$

where the operator $\psi(D)$ is defined as the multiplication of the Fourier transform by the real function ψ. The characteristic surface Σ is defined in $\mathbb{R}^{d+1}$ as the zero level set of the function

$$\phi(\tau,\xi) = \tau - \psi(\xi).$$

Let u be the solution with initial data u_0. Then

$$\mathcal{F}_x u(t,\xi) = e^{it\psi(\xi)} \mathcal{F}_x u_0(\xi)$$

and, for any Schwartz function $f \in \mathcal{S}(\mathbb{R}^{d+1})$ with Fourier transform g, by Plancherel

$$\begin{aligned}\int_{\mathbb{R}\times\mathbb{R}^d} u\overline{f} dm^{d+1}(t,x) &= \int_{\mathbb{R}}\int_{\mathbb{R}^d} \mathcal{F}_x u(t,\xi)\overline{\mathcal{F}_x f(t,\xi)} dm^d(\xi)dt \\ &= \int_{\mathbb{R}}\int_{\mathbb{R}^d} e^{it\psi(\xi)}\hat{u}_0(\xi)\overline{\mathcal{F}_x f(t,\xi)} dm^d(\xi)dt \\ &= \int_{\mathbb{R}^d} \hat{u}_0(\xi) \int \overline{e^{-it\psi(\xi)}\mathcal{F}_x f(t,\xi)} dt dm^d(\xi) \\ &= \sqrt{2\pi}\int_{\mathbb{R}^d} \hat{u}_0(\xi)\overline{g(\psi(\xi),\xi)} d\xi \\ &= \sqrt{2\pi}\int_{\tau=\psi(\xi)} |\nabla_{\tau,\xi}\phi(\tau,\xi)|^{-1} u_0(\xi)\bar{g}(\tau,\xi) d\mathcal{H}^d(\tau,\xi) \\ &=: \sqrt{2\pi}\int \bar{g}(\tau,\xi)\hat{u}_0(\xi)\delta_\phi.\end{aligned}$$

This calculation implies the following lemma.

Lemma 5.2. *Let $\mathcal{F}_x u(t,x) = e^{it\psi}\mathcal{F}_x u_0$. Then the space-time Fourier transform of u is the the measure $\sqrt{2\pi}\hat{u}_0\delta_\phi$.*

Let ψ_1 and ψ_2 be real smooth functions and, as above,

$$\phi_1(\tau,\xi) = \tau - \psi_1(\xi), \text{ resp. } \phi_2(\tau,\xi) = \tau - \psi_1(\xi).$$

The product uv of two solution of the linear equations

$$iu_t + \psi_1(D)u = 0, \quad iv_t + \psi_2(D)v = 0$$

is the convolution of the Fourier transforms in Lemma 5.2, which in turn can be estimated by Theorem 5.1. We identify the terms occurring in Theorem 5.1.

The integration set for given ξ_j is

$$M = \{(\tau, \xi) : \tau = \psi_2(\xi_2) + \psi_1(\xi - \xi_2) = \psi_1(\xi_1) + \psi_2(\xi - \xi_1)\}.$$

The most important case will be $\psi_i = \psi$. We express the integrand in terms of $\nabla\psi_i$ using

$$\begin{aligned} |\nabla_{\tau,\xi}\phi_1|^2|\nabla_{\tau,\xi}\phi_2|^2 - (\nabla_{\tau,\xi}\phi_1 \cdot \nabla_{\tau,\xi}\phi_2)^2 \\ = |\nabla\psi_1 - \nabla\psi_2|^2 + |\nabla\psi_1|^2|\nabla\psi_2|^2 - (\nabla\psi_1 \cdot \nabla\psi_2)^2. \end{aligned} \tag{5.4}$$

The first term is the square of the distance of the gradients, and the second is the square of product of length multiplied by $\sin^2$ of the angle between them. Here we did not write the arguments. With them the integrand reads

$$\Big[|\nabla\psi_1(\xi - \xi_2) - \nabla\psi_2(\xi - \xi_1)|^2 + |\nabla\psi_1(\xi - \xi_2)|^2|\nabla\psi_2(\xi - \xi_1)|^2 \\ - (\nabla\psi_1(\xi - \xi_2) \cdot \nabla\psi_2(\xi - \xi_1))^2\Big]^{-\frac{1}{2}}. \tag{5.5}$$

The proof of bilinear estimates reduces to bounding the integral of this expression over M.

We first consider the case of one space dimension, where the sum of the second and the third term on the right-hand side of (5.4) vanishes. The set $(x + \Sigma_1) \cap (y + \Sigma_2)$ consists generically of a discrete set of points and we obtain a sum of $|\psi_1'(z-x) - \psi_2'(z-y)|^{-1}$ over the points of the intersection. Often the intersection consists of one point, as for the Schrödinger equation, or up to two points, as for the Airy equation. We consider the more general case of $\psi(\xi) = \xi^N$ for an even integer N. Then the equation

$$\xi_1^N + (\xi - \xi_1)^N = \xi_2^N + (\xi - \xi_2)^N$$

has the obvious and unique solution $\xi = \xi_2 + \xi_1$, unless $\xi_1 = \xi_2$. If N is odd there are the exactly two solutions, $\xi = \xi_1 + \xi_2$ and $\xi = 0$, unless $\xi_2 = \xi_1$.

At these points

$$|\psi'(\xi - \xi_1) - \psi'(\xi - \xi_2)| = |\psi'(\xi_1) - \psi'(\xi_2)|$$

and we obtain from inequality (5.3):

Theorem 5.3. *With the notation introduced above*

$$\Big\| \int |N[(\xi - \eta)^{N-1} - \eta^{N-1}]|^{1/2} e^{it(\xi-\eta)^N + it\eta^N} \hat{u}_0(\xi - \eta)\hat{u}_1(\eta)d\eta \Big\|_{L^2(\mathbb{R}_t \times \mathbb{R}_\xi)} \\ \leq 2\pi\|u_0\|_{L^2(\mathbb{R})}\|u_1\|_{L^2(\mathbb{R})}, \tag{5.6}$$

if N is even, and

$$\Big\| \int |N[(\xi-\eta)^{\frac{N-1}{2}} - \eta^{\frac{N-1}{2}}]|^{1/2} e^{it(\xi-\eta)^N + it\eta^N} \hat u_0(\xi-\eta)\hat u_1(\eta)d\eta \Big\|_{L^2(\mathbb{R}_t\times\mathbb{R}_\xi)} \le \sqrt{2}\, 2\pi \|u_0\|_{L^2(\mathbb{R})}\|u_1\|_{L^2(\mathbb{R})}, \tag{5.7}$$

if N is odd.

We will use this estimate often via the following corollary. Given $\lambda \in (0,\infty)$, we define

$$u_{>\lambda} = \mathcal{F}^{-1}(\chi_{|\xi|>\lambda}\hat u),$$

and similarly for $u_{<\lambda}$.

Corollary 5.4. *Let $0<\mu<\lambda$ and $u(t,x) = S(t)u_0(x)$, $v(t,x) = S(t)v_0(x)$, where S is the unitary group defined by $\phi = \xi^N$. Then*

$$\|u_{<\mu}v_{>\lambda}\|_{L^2(\mathbb{R}^2)} \le \frac{4\pi}{\big[N|\lambda^{N-1}-\mu^{N-1}|\big]^{\frac12}}\|u_0\|_{L^2(\mathbb{R})}\|v_0\|_{L^2(\mathbb{R})}$$

and

$$\|(u_{>\lambda}v_{>\lambda})_{>\mu}\|_{L^2} \le \frac{4\pi}{\big[N|\lambda^{N-1}-(\lambda-\mu)^{N-1}|\big]^{\frac12}}\|u_0\|_{L^2(\mathbb{R})}\|v_0\|_{L^2(\mathbb{R})}.$$

There is an interesting special case of the bilinear estimate: Local smoothing corresponds to $\Sigma_1 = \{(\xi^N,\xi)\}$ and Σ_0 given by $\tau = 0$.

Theorem 5.5. *Let $\psi(\xi) = \xi^N$ be as above. Then*

$$\||ND^{N-1}|^{1/2}S(t)u_0\|_{L^\infty_x L^2_t} \le 4\pi^2\|u_0\|_{L^2(\mathbb{R})},$$

if N is odd, and

$$\||ND^{N-1}|^{1/2}S(t)u_0\|_{L^\infty_x L^2_t} \le \sqrt{2}\,4\pi^2\|u\|_{L^2(\mathbb{R})},$$

if N is even.

Proof. We apply the convolution estimate with $\psi_1(\xi) = \xi^N$ and $\psi_0 = 0$. The set M is given by

$$\tau = (\xi-\xi_0)^N = \xi_1^N$$

which has the unique solution $\xi = \xi_1 - \xi_0$ if N is odd, and $\xi = \xi_0 \pm \xi_1$ if N is even, and the integrand is

$$|\psi'(\xi-\xi_0)|^{-1} = N|\xi_1|^{1-N}.$$

Thus, if N is odd, then

$$\sqrt{N}\int |(|D|^{\frac{N-1}{2}}S(t)u_0)v(x)|^2 dxdt \le 2\pi\|u_0\|^2_{L^2(\mathbb{R})}\|v\|^2_{L^2(\mathbb{R})}$$

and we choose v so that $|v|^2$ is a Dirac sequence. Only obvious adaptations are needed if N is even. $\square$

In particular, if u satisfies the Airy equation, then

$$\|\partial_x u\|_{L^\infty_x L^2_t(\mathbb{R})} \le 2\pi \|u_0\|_{L^2} \tag{5.8}$$

and u has locally square integrable derivatives for almost all t.

We continue with a case by case study of several linear dispersive equations in several space dimensions. The first is the Schrödinger equation in higher space dimensions. Here the characteristic set Σ is a standard parabola. The set

$$\{(\tau_1, \xi_1) + \Sigma\} \cap (\tau_2, \xi_2) + \Sigma\}$$

is the intersection of two paraboloids, and hence a paraboloid of dimension $d-1$. It is given by the equations

$$\tau = |\xi_1|^2 + |\xi - \xi_1|^2 = |\xi_2|^2 + |\xi - \xi_2|^2.$$

The first equality determines τ, which is of minor importance, and the second is equivalent to

$$\langle \xi, \xi_2 - \xi_1 \rangle = |\xi_2|^2 - |\xi_1|^2,$$

resp.

$$\langle \xi - (\xi_2 + \xi_1), \xi_2 - \xi_1 \rangle = 0 \tag{5.9}$$

which describes a hyperplane with normal $\xi_2 - \xi_1$, if $\xi_2 \neq \xi_1$. We restrict to this non-degenerate situation. This suffices for the estimate.

Let w be the closest point of the hyperplane defined by (5.9) to ξ_1, resp. ξ_2. With this notation the intersection is given by

$$\{(\tau, w+v) : \tau = \xi_1^2 + |w-\xi_1|^2 + |v|^2 = \xi_2^2 + |w-\xi_2|^2 + |v|^2, \langle v, \xi_2 - \xi_1 \rangle = 0\}. \tag{5.10}$$

If we integrate with respect to v, we obtain by the coarea formula an integral

$$\int \dots \sqrt{1 + 4|v|^2} dv.$$

For $\xi = w + v$,

$$\nabla |\xi - \xi_1|^2 = 2v + 2(w - \xi_1),$$

and similarly

$$\nabla |\xi - \xi_2|^2 = 2v + 2(w - \xi_2).$$

Thus the square of the difference is given by

$$4|\xi_2 - \xi_1|^2$$

and

$$(|v|^2 + |w - \xi_1|^2)(|v|^2 + |w - \xi_2|^2) - (|v|^2 + (w - \xi_1)(w - \xi_2))^2 = |v|^2 |\xi_2 - \xi_1|^2$$

and the integrand is

$$(|\xi_2 - \xi_1| \sqrt{4 + 4|v|^2})^{-1}.$$

We will choose Σ_1 to be the part of the parabola above $|\xi| \ge \lambda$ and Σ_2 the part of the parabola above the ball of radius μ.

Lemma 5.6 (Schrödinger, d dimensions). *Let $d \geq 2$, $u(t,x) = S(t)u_0$, $v(t) = S(t)v_0$, where S denotes the Schrödinger group. Let $\mu \leq \frac{1}{2}\lambda$. Then*

$$\|u_{>\lambda} v_{<\mu}\|_{L^2} \leq c_d \mu^{\frac{d-1}{2}} \lambda^{-1/2} \|u_{0,>\lambda}\|_{L^2} \|v_{0,<\mu}\|_{L^2}$$

and

$$\|(uv)_{<\mu}\|_{L^2(\mathbb{R}^2)} \leq c_d \mu^{\frac{d-2}{2}} \|u_0\|_{L^2} \|v_0\|_{L^2}.$$

Proof. In the first case $|\xi_2 - \xi_1| \geq \lambda/2$, and we integrate over a ball of radius μ. The factor from the area formula cancels the one from the integrand, hence the first estimate. It is not difficult to determine the constant c_d.

The second estimate could probably be proven with the arguments here. We derive it from the Strichartz estimate

$$\|u\|_{L^4_t L_x^{\frac{2d}{d-1}}} \leq c \|u_0\|_{L^2(\mathbb{R}^d)}.$$

We combine it with Bernstein's inequality for $p \leq q$,

$$\|v_{<\mu}\|_{L^q} \leq c \mu^{\frac{d}{p} - \frac{d}{q}} \|v_{<\mu}\|_{L^p}.$$

With a smooth truncation (instead of the Fourier multiplication by a characteristic function) we obtain for fixed t

$$\|(uv)_{<\mu}(t)\|_{L^2(\mathbb{R}^d)} \leq c \mu^{\frac{d-2}{2}} \|uv\|_{L^{\frac{d}{d-1}}} \leq c \mu^{\frac{d-2}{2}} \|u\|_{L^{\frac{2d}{d-1}}(\mathbb{R}^d)} \|v\|_{L^{\frac{2d}{d-1}}(\mathbb{R}^d)},$$

and we complete the argument by taking the L^2 norm with respect to t. □

The case of the Kadomtsev–Petviashvili II equation is considerably more intricate. We study

$$u_t + u_{xxx} + \partial_x^{-1} u_{yy} = 0.$$

The symbol resp. Fourier multiplier is

$$\psi(\xi, \eta) = \xi^3 - \eta^2/\xi.$$

Here the formal notation ∂_x^{-1} has to be understood as Fourier multiplier. Here is it useful to first apply Fubini's theorem for the integration over $\Sigma((\tau_1, \xi_1, \eta_1), (\tau_2, \xi_2, \eta_2))$, or more precisely in its derivation, and to integrate first with respect to ξ.

For fixed ξ the intersection consists of at most two points η and the considerations in one space dimension show that the integrand for the integration with respect to ξ is the following to the power $-1/2$:

$$\begin{aligned} &\left|\partial_\eta[(\xi-\xi_1)^3 - (\eta-\eta_1)^2/(\xi-\xi_1) - (\xi-\xi_2)^3 - (\eta-\eta_2)^2/(\xi-\xi_2)]\right| \\ &\qquad = 2\left|\frac{\eta-\eta_1}{\xi-\xi_1} - \frac{\eta-\eta_2}{\xi-\xi_2}\right|. \end{aligned} \tag{5.11}$$

The integration curve is described by the equations

$$\tau = \xi_1^3 - \frac{\eta_1^2}{\xi_1} + \tau - \xi_1^3 + (\xi-\xi_1)^3 - \frac{(\eta-\eta_1)^2}{\xi-\xi_1} = \xi_2^3 - \frac{\eta_2^2}{\xi_2} + (\xi-\xi_2)^3 - \frac{(\eta-\eta_2)^2}{\xi-\xi_2}.$$

We reorganize the second identity to

$$\begin{aligned}
&\xi_1^3 - \frac{\eta_1^2}{\xi_1} + \left[(\xi_2-\xi_1)^3 - \frac{(\eta_2-\eta_1)^2}{\xi_2-\xi_1}\right] - \xi_2^3 + \frac{\eta_2^2}{\xi_2} \\
&= \left[(\xi-\xi_2)^3 - \frac{(\eta-\eta_2)^2}{\xi-\xi_2}\right] + \left[(\xi_2-\xi_1)^3 - \frac{(\eta_2-\eta_1)^2}{\xi_2-\xi_1}\right] - \left[(\xi-\xi_1)^3 - \frac{(\eta-\eta_1)^2}{\xi-\xi_1}\right]
\end{aligned}$$

and use the algebraic resonance relation

$$(\xi_1+\xi_2)^3 - \frac{(\eta_1+\eta_2)^2}{\xi_1+\xi_2} - (\xi_1^3-\eta_1^2/\xi_1) - (\xi_2^3-\eta_2^2/\xi_2) = \xi_1\xi_2(\xi_1+\xi_2)\left[3 + \frac{\left|\frac{\eta_1}{\xi_1} - \frac{\eta_2}{\xi_2}\right|^2}{|\xi_1+\xi_2|^2}\right] \tag{5.12}$$

to arrive at

$$\begin{aligned}
\omega &:= \xi_1\xi_2(\xi_1-\xi_2)\left(3 + \frac{\left|\frac{\eta_1}{\xi_1} - \frac{\eta_2}{\xi_2}\right|^2}{(\xi_2-\xi_1)^2}\right) \\
&= (\xi-\xi_2)(\xi-\xi_1)(\xi_1-\xi_2)\left(3 + \frac{\left|\frac{\eta-\eta_1}{\xi-\xi_1} - \frac{\eta-\eta_2}{\xi-\xi_2}\right|^2}{|\xi_2-\xi_1|^2}\right).
\end{aligned} \tag{5.13}$$

Here we used the elementary identities which show a high degree of symmetry:

$$\frac{\left|\frac{\eta_1}{\xi_1} - \frac{\eta_2}{\xi_2}\right|^2}{|\xi_1+\xi_2|^2} = \frac{|\xi_1\eta_2 - \xi_2\eta_1|^2}{(\xi_1\xi_2(\xi_1+\xi_2))^2} = \frac{|(\xi_1+\xi_2)\eta_2 - \xi_2(\eta_1+\eta_2)|^2}{(\xi_1\xi_2(\xi_1+\xi_2))^2} = \frac{\left|\frac{\eta_1+\eta_2}{\xi_1+\xi_2} - \frac{\eta_2}{\xi_2}\right|^2}{|\xi_1|^2}.$$

The left-hand side of (5.13) is the called modulation. Assuming neither $\xi_1 = 0$, nor $\xi_2 = 0$, nor $\xi_1 = \xi_2$, there is only a solution if $\xi_1\xi_2$ has the same sign as $(\xi-\xi_2)(\xi-\xi_1)$. Below we neglect the question whether there is a solution and we rewrite the identity (5.3) as

$$\left|\frac{\eta-\eta_1}{\xi-\xi_1} - \frac{\eta-\eta_2}{\xi-\xi_2}\right|^2 = \frac{\xi_2-\xi_1}{(\xi-\xi_1)(\xi-\xi_2)}(\omega - 3(\xi-\xi_1)(\xi-\xi_2)(\xi_2-\xi_1)) = f(\xi), \tag{5.14}$$

which is useful for determining η as a function of ξ. The left-hand side coincides with (5.11) and allows us to determine the integrand as a function of ξ.

Algebraic manipulation allow a fairly explicit determination of the solutions to the polynomial equation (5.13). To shorten the notation, we write $\tilde{\xi} = \xi_2 - \xi_1$ in the sequel. Then

$$0 = 3\tilde{\xi}^2(\xi-\xi_1)^2(\xi-\xi_2)^2 + \big((\eta-\eta_1)(\xi-\xi_2)-(\eta-\eta_2)(\xi-\xi_1)\big)^2 + \omega\tilde{\xi}(\xi-\xi_1)(\xi-\xi_2), \tag{5.15}$$

which we rewrite in terms of

$$\hat{\xi} = \xi - \frac{1}{2}(\xi_1 + \xi_2)$$

and

$$\begin{aligned}\hat{\eta} &= (\eta-\eta_1)(\xi-\xi_2) - (\eta-\eta_2)(\xi-\xi_1) \\ &= \eta(\xi_1-\xi_2) + \xi(\eta_2-\eta_1) + \eta_1\xi_2 - \eta_2\xi_1.\end{aligned}$$

We observe that

$$f(\xi) = \frac{\hat{\eta}^2}{\hat{\xi}^2 - \frac{1}{4}\tilde{\xi}^2}, \tag{5.16}$$

since

$$(\xi-\xi_1)(\xi-\xi_2) = \hat{\xi}^2 - \left(\frac{\xi_1-\xi_2}{2}\right)^2,$$

and we obtain from (5.3)

$$3\tilde{\xi}^2\Big(\hat{\xi}^2 - \frac{1}{4}\tilde{\xi}^2\Big)^2 + \omega\tilde{\xi}\big(\hat{\xi}^2 - \frac{1}{4}\tilde{\xi}^2\big) + \hat{\eta}^2 = 0. \tag{5.17}$$

We arrive at

$$\Big[\sqrt{3}\tilde{\xi}(\hat{\xi}^2 - \frac{1}{4}\tilde{\xi}^2) + \frac{\omega}{2\sqrt{3}}\Big]^2 + \hat{\eta}^2 = \frac{\omega^2}{12}. \tag{5.18}$$

It remains to partly undo and interpret the formulas and transformations. For simplicity, we assume that $\xi_1 < \xi_2$. All solutions of the polynomial equation satisfy

$$\left|\sqrt{3}\tilde{\xi}(\hat{\xi}^2 - \frac{1}{4}\tilde{\xi}^2) + \frac{\omega}{2\sqrt{3}}\right| \le \frac{|\omega|}{2\sqrt{3}},$$

resp.

$$\frac{1}{6}(\omega - |\omega|) \le (\xi_2-\xi_1)(\xi-\xi_1)(\xi-\xi_2) \le \frac{1}{6}(\omega + |\omega|), \tag{5.19}$$

which we could have read from (5.13). Clearly,

$$|\xi_1-\xi_2|(\xi_1-\xi)(\xi-\xi_2) \le \frac{1}{4}|\xi_1-\xi_2|^3$$

with equality if $\rho = 0$ resp. $\xi = \frac{\xi_1+\xi_2}{2}$. This set always contains the points $\xi = \xi_1$, $\eta = \eta_1$ and $\xi = \xi_2$, $\eta = \eta_2$. We list the geometric cases.

1. If $\xi_1\xi_2 < 0$ and, for simplicity, $\xi_1 < 0 < \xi_2$, then (5.19) describes two intervals, the length of which is at least $\min\{|\xi_1|, |\xi_2|\}$ The set is a union of two topological circles contained in $\{\xi \le \xi_1\} \cup \{\xi \ge \xi_2\}$. The size of the circles is given by ω.

2. If $\xi_1\xi_2 > 0$ and, for simplicity, $0 < \xi_1 < \xi_2$ and
$$|\omega| < \frac{4}{3}|\xi_1 - \xi_2|^3,$$
then there are again two topological circles, but this time contained in $\{\xi_1 \le \xi < \frac{\xi_2+\xi_1}{2}\}$ and $\{\frac{\xi_1+\xi_2}{2} < \xi \le \xi_2\}$. Again their size is at least $\min\{|\xi_1|, |\xi_2|\}$.

3. If $\xi_1\xi_2 > 0$ and
$$|\omega| = \frac{4}{3}|\xi_1 - \xi_2|^3,$$
then the intersection is a topological figure 8 contained in $\xi_1 \le \xi \le \xi_2$. The center of the figure 8 is at $\xi = \frac{\xi_1+\xi_2}{2}$ and $\eta = \frac{\eta_1+\eta_2}{2}$.

4. If $|\omega| > \frac{4}{3}|\xi_1-\xi_2|^3$, then the intersection is a topological sphere in $\xi_1 < \xi < \xi_2$. In this case
$$f(\xi) \sim \frac{\xi_2 - \xi_1}{(\xi - \xi_1)(\xi - \xi_2)}\omega.$$

The set expressed in terms of $\hat\eta$ and ξ is always symmetric with respect to the reflection at $\frac{\xi_1+\xi_2}{2}$ and $\hat\eta = 0$. We choose various subsets of the characteristic surface. Let $\mu \le \lambda$, $\Sigma_1 = \Sigma \cap \{\mu/2 \le |\xi| \le \mu\}$, and $\Sigma_2 = \Sigma \cap \{\lambda \le |\xi|\}$.

If $\mu \le \lambda/10$, then we obtain only the parts of the curves with $|\xi_2 - \xi| \sim \mu$ and $\eta \sim \eta_2$. In particular, we stay away from $\xi = \frac{\xi_1+\xi_2}{2}$. If $\left|\frac{\eta_1}{\xi_1} - \frac{\eta_2}{\xi_2}\right| \ge 5\lambda$, then $|f| \ge \frac{\omega}{\mu}$, the ξ integral is over an interval of length μ, and
$$\int_I |f|^{-\frac{1}{2}} d\xi \sim \frac{\mu^{\frac{3}{2}}}{\sqrt{\omega}}. \tag{5.20}$$

If $\mu \sim \lambda$ and $\left|\frac{\eta_1}{\xi_1} - \frac{\eta_2}{\xi_2}\right| \le 5\lambda$ we apply the L^4 Strichartz estimate. In the opposite case we argue as above.

Let
$$A_{\mu,\Lambda,k} = \Big\{(\xi,\eta) : \mu \le |\xi| \le 2\mu, k\mu - \frac{\Lambda}{\mu} \le \frac{\eta}{\xi} < k\mu + \frac{\Lambda}{\mu}\Big\}.$$
We use the Strichartz estimate for $\mu \sim \lambda$.

Theorem 5.7. *The following estimate holds with suggestive notation and if $\mu \le \lambda$:*
$$\left\| \int \left| \Big(3 + \frac{\left|\frac{\eta_1}{\xi_1} - \frac{\eta_2}{\xi_2}\right|^2}{|\xi_1 - \xi_2|^2}\Big)\right|^{1/4} \hat u_{<\mu}(t,\xi_1)\hat v_{>\lambda}(t,\xi_2) \right\|_{L^2} \le c\Big(\frac{\mu}{\lambda}\Big)^{\frac{1}{2}} \|v_\mu(0)\|_{L^2}\|u_\lambda(0)\|_{L^2} \tag{5.21}$$

where the inner integral is a two-dimensional integral with respect to ξ_1 and η_1, and $\xi_2 = \xi - \xi_1$, resp. $\eta_2 = \eta - \eta_1$. Similarly,

$$\left\| \int \left| \left(3 + \frac{\left| \frac{\eta_1}{\xi_1} - \frac{\eta_2}{\xi_2} \right|^2}{|\xi_1 - \xi_2|^2} \right) \right|^{1/2} \hat{u}_{A_{\mu,\Lambda,k}}(t,\xi_1) \hat{v}_{>\lambda}(t,\xi_2) \right\|_{L^2} \leq c \frac{\sqrt{\Lambda}}{\lambda} \|v_\mu(0)\|_{L^2} \|u_{>\lambda}(0)\|_{L^2}. \tag{5.22}$$

Proof. The first estimate follows from the previous estimates. If $|\mu| \sim |\lambda|$ and $|\omega| \leq c\lambda^3$, the first estimate follows from the L^4 Strichartz estimate.

Only the second estimate remains to be shown. We prove the estimate first for $k = 0$. The curve described by (5.18) lies on one side of ξ_1, resp. ξ_2, and hence it is vertical there. Assuming $\eta_1 = 0$, we expand equation (5.18) to

$$\begin{aligned} 3(\xi_2 - \xi_1)^2(\xi - \xi_1)^2(\xi - \xi_2)^2 + \omega(\xi_2 - \xi_1)(\xi - \xi_1)(\xi - \xi_2) \\ + \eta^2(\xi_2 - \xi_1)^2 - 2\eta\eta_2(\xi - \xi_1)(\xi - \xi_2) + \eta_2^2(\xi - \xi_1)^2. \end{aligned}$$

We consider the situation where $\Lambda \leq \mu^{\frac{1}{2}}\omega^{\frac{1}{2}}$; in the complementary case estimate (5.21) is stronger.

The dominant terms are the second and the third term and hence in that range

$$|\xi - \xi_1| \leq C\frac{\eta^2}{\omega}. \tag{5.23}$$

This bounds the interval of integration in (5.20) and implies the estimates.

The bound (5.23) follows from our discussion above, which controls the global geometry, and a continuity argument from $\xi = \xi_1$ and $\eta = \eta_1$:

$$3(\xi_2-\xi_1)^2(\xi-\xi_1)^2(\xi-\xi_2)^2 = \left(\frac{(\xi_2 - \xi_1)(\xi - \xi_1)(\xi - \xi_2)}{\omega} \right) \omega(\xi_2-\xi_1)(\xi-\xi_1)(\xi-\xi_2)$$

where the bracket is small compared to the next term provided $|\xi - \xi_1| << \mu|\omega|$. Similarly,

$$\eta_2^2(\xi - \xi_1)^2 = \left(\frac{\eta_2^2(\xi - \xi_1)}{\omega(\xi_2 - \xi_1)(\xi - \xi_2)} \right) \omega(\xi_2 - \xi_1)(\xi - \xi_1)(\xi - \xi_2)$$

is small by a continuity argument. The restriction $k = 0$, resp. $\eta_1 = 0$ is possible due to the Galilean symmetry

$$(t, x, y) \to (t, x + ct + cy, y + 2ct)$$

which is a symmetry of the linear and nonlinear KP II equation, and it respects the bilinear estimate. On the Fourier side, this corresponds to

$$(\tau, \eta, \xi) \to (\tau - 2c\eta + c\xi, \eta - c\xi, \xi).$$

If we neglect τ, then the lines through the origin in the (ξ,η)-plane are mapped into such lines, and the lines $\xi = c$ are preserved.

The center of the figure 8 never plays a role unless $\mu \sim \lambda$ and $|\frac{\eta_1}{\xi_1} - \frac{\eta_2}{\xi_2}| \le \lambda$, but then its contribution is not hard to control. □

We conclude this section by explaining the relation to U^2 spaces. Let, as above, u_A resp. u_B be the projection on the Fourier side to sets A resp. B.

Theorem 5.8. *Suppose that*

$$\|S(t)u_{0,A}S(t)v_{0,B}\|_{L^2} \le c_{A,B}\|u_{0,A}\|_{L^2}\|u_{0,B}\|_{L^2}.$$

Then with the same constant we have

$$\|u_A v_B\|_{L^2} \le c_{A,B}\|u_A\|_{U^2_S}\|u_B\|_{U^2_S}.$$

Proof. As for the Strichartz estimates, the assertion reduces to the assumption and a summation for 2-atoms. We first write the second term as a sum of atoms to obtain the statement when the first factor is a an atom, and the second factor is in U^2, and then we expand the first factor to obtain the full statement. □

Chapter 6

Well-posedness for nonlinear dispersive equations

In this section we will study local and global well-posedness for a number of different equations where the techniques developed so far are relevant. The first example describes the interaction of three waves of different velocities. It is elementary and displays the role of adapted function spaces on an elementary level. The limitations of our current understanding become obvious as well: The result should remain true under small perturbations of the system, but I have no idea how to approach perturbed equations.

Next we turn to generalized KdV equations and establish global well-posedness and scattering in a large scale invariant Besov space for the quartic and the quintic equation, and local existence for modified KdV and KdV in the spaces $B^{\frac{1}{4}}_{2,\infty}$ and $B^{-\frac{3}{4}}_{2,\infty}$ using the U^2-V^2 spaces, bilinear estimates, Strichartz estimates and, for KdV, modulation arguments. This is basically well known, but for KdV and mKdV slightly stronger than available results in the literature. Going from $H^{-\frac{3}{4}}$ to $B^{-\frac{3}{4}}_{2,\infty}$ for the initial data for Korteweg–de-Vries requires a new technique, which also allows to treat low frequencies similarly to high frequencies.

Next we turn to higher dimensional non-resonant derivative Schrödinger equations, following the dissertation of T. Schottdorf, and conclude with a discussion of the two-dimensional Kadomtsev–Petviashvili II equation.

6.1 Adapted function spaces approach for a model problem

To motivate the relevance of adapted function spaces, we begin with a self-contained study of a simple toy problem, where a nonstandard choice of adapted function spaces leads to global well-posedness for small data in L^2, and where I know of no other technique which allows to prove this result. Consider the three

H. Koch et al., *Dispersive Equations and Nonlinear Waves: Generalized Korteweg–de Vries, Nonlinear Schrödinger, Wave and Schrödinger Maps*, Oberwolfach Seminars 45,
DOI 10.1007/978-3-0348-0736-4_6, © Springer Basel 2014

wave interaction

$$\begin{aligned} u_t + u_x &= vw, \\ v_t + v_y &= uw, \\ w_t &= -2uv. \end{aligned} \tag{6.1}$$

It is easy to solve the linear equation for given initial data. We define the evolution operator

$$S(t)[u_0, v_0, w_0](x, y) = [u_0(x - t, y), v_0(x, y - t), w(x, y)]$$

and the operator-adapted function space with norm

$$\begin{aligned} \|[u, v, w]\|_X = \max\big\{ &\| \sup_t |u(t, x + t, y)| \|_{L^2(\mathbb{R}^2)}, \| \sup_t |v(t, x, y + t)| \|_{L^2(\mathbb{R}^2)}, \\ &\| \sup_t |w(t, x, y)| \|_{L^2(\mathbb{R}^2)} \big\} \end{aligned}$$

or, written differently, with the equivalent norm

$$\|[u, v, w]\|_X \sim \| \sup_t S(-t)[u(t, x, y), v(t, x, y), w(t, x, y)] \|_{L^2(\mathbb{R}^2)}.$$

Theorem 6.1. *If*

$$\max\{\|u_0\|_{L^2}, \|v_0\|_{L^2}, \|w_0\|_{L^2}\} \leq \frac{1}{4},$$

then there exists a unique global solution $[u, v, w] \in X$ *which satisfies*

$$\|[u, v, w] - S(t)[u_0, v_0, w_0]\|_X \leq 2 \max\{\|u_0\|_{L^2}, \|v_0\|_{L^2}, \|w_0\|_{L^2}\}^2.$$

Proof. The assertion follows by an easy duality argument from the trilinear estimate

$$\left| \int uvw \, dx \, dy \, dt \right| \leq \| \sup_t |u(t, x+t, y)| \|_{L^2} \| \sup_t |v(t, x+y+t)| \|_{L^2} \| \sup_t |w(t, x, y)| \|_{L^2}. \tag{6.2}$$

To prove this estimate we denote

$$\tilde{u}(x, y) = \sup_t |u(t, x + t, y)|, \quad \tilde{v} = \sup_t |v(t, x + y + t)|, \quad \tilde{w}(x, y) = \sup_t |w(t, x, y)|.$$

Then

$$\int |uvw| dxdydt \leq \int \tilde{u}(x - t, y)\tilde{v}(x, y - t)\tilde{w}(x, y) dtdxdy \leq \|\tilde{u}\|_{L^2} \|\tilde{v}\|_{L^2} \|\tilde{w}\|_{L^2},$$

by a multiple application of the Cauchy–Schwarz inequality.

It is not difficult to set up an iteration argument to construct a global solution for small data, which depends analytically on the initial datum. $\square$

6.2 The (generalized) KdV equation

For integers $p \geq 1$ we consider the initial value problems

$$u_t + u_{xxx} + (u^p u)_x = 0, \tag{6.3}$$

$$u(0) = u_0; \tag{6.4}$$

(the case $p = 1$ is the Korteweg–de-Vries equation, $p = 2$ yields the modified Korteweg–de-Vries equation) and

$$u_t + u_{xxx} + (|u|^p u)_x = 0, \tag{6.5}$$

$$u(0) = u_0, \tag{6.6}$$

for positive real p.

Both equations have soliton solutions

$$u(x,t) = c^{\frac{1}{p}} Q(c^{1/2}(x - ct)),$$

with

$$Q_p = \left(\frac{p+1}{2}\right)^{2/p} \cosh^{2/p}\left(\frac{2}{p}x\right).$$

The equation is invariant with respect to scaling: $\lambda^{2/p} u(\lambda x, \lambda^3 t)$ is a solution if u satisfies the equation. The mass $\int u^2 dx$ and energy $\int (\frac{1}{2}u_x^2 - \frac{1}{p+2}u^{p+2})dx$ for (6.5) are conserved. The energy, however, is not bounded from below.

The space $\dot{H}^{\frac{1}{2}-\frac{2}{p}}$ (with norm $\|u\|_{\dot{H}^s} = \||\xi|^s \hat{u}\|_{L^2}$) is invariant with respect to scaling and it is not hard to see that the generalized KdV equation is globally well posed in H^1 if $p < 4$. For $p \geq 4$ one expects blow-up. This has been proven in series of seminal papers by Martel, Merle and Martel, and Merle and Raphael.

Using the Fourier transform we see that

$$v_t + v_{xxx} = 0, \quad v(0,x) = v_0(x)$$

defines a unitary group on L^2. We denote

$$S(t)v_0 = v(t)$$

for $t \geq 0$ and $v(t) = 0$ otherwise, and define the adapted function spaces by

$$\|u\|_{U^p_{\mathrm{KdV}}} = \|S(-t)u(t)\|_{U^p}, \qquad \|u\|_{V^p_{\mathrm{KdV}}} = \|S(-t)u(t)\|_{V^p}.$$

The Strichartz estimates are

$$\|u\|_{L^p_t L^q_x} \leq c\||D|^{-1/p} u_0\|_{L^2} \tag{6.7}$$

for $p \geq 4$ and

$$\frac{2}{p} + \frac{1}{q} = \frac{1}{2}.$$

We have seen in Sections 2.1.2 and 4.10.1 that they imply the embedding estimates

$$\| |D|^{1/p} u \|_{L^p_t L^q_x} \leq c \|u\|_{U^p_{\mathrm{KdV}}} \tag{6.8}$$

in the same range.

For $\lambda > 0$ we denote

$$u_\lambda = \chi_{\lambda \leq |\xi| \leq 1.01\lambda}(D) u,$$

the projection of the Fourier transform. Then the Strichartz embedding applied to $g(D)u$, $g(\xi) = |\xi|^{-\frac{1}{p}}$ gives

$$\|u_\lambda\|_{L^p_t L^q_x} \leq c \lambda^{-1/p} \|u\|_{U^p_{\mathrm{KdV}}} \tag{6.9}$$

(checking atoms one sees that Fourier multipliers act nicely on U^p and V^p).

The bilinear estimates for $\mu \leq \frac{9}{10}\lambda$,

$$\|S(t)u_{0,\lambda} S(t) v_{0,\mu}\|_{L^2} \leq c \lambda^{-1} \|u_{0,\lambda}\|_{L^2} \|v_{0,\mu}\|_{L^2},$$

are a direct consequence of the bilinear estimate of the last section. Hence

$$\|u_\lambda v_\mu\|_{L^2} \leq c\lambda^{-1} \|u_\lambda\|_{U^2_{\mathrm{KdV}}} \|v_\mu\|_{U^2_{\mathrm{KdV}}}. \tag{6.10}$$

After these preparations we turn to the cases $p = 4$ and $p = 3$. There is a number of aspects which are the same for both cases, and also for many other equations. We discuss them in detail for the case $p = 4$ and only sketch them for other p later.

We begin with the L^2 critical case

$$u_t + u_{xxx} + u_x^5 = 0, \tag{6.11}$$

and choose the norm

$$\|u_0\|_{\dot{B}^0_{2,\infty}} = \sup_{\lambda \in 1.01^{\mathbb{Z}}} \|u_{0,\lambda}\|_{L^2(\mathbb{R})}$$

for the initial data, and, with $I = [0, T)$, $T \in (0, \infty]$,

$$\|u\|_X = \sup_{\lambda \in 1.01^{\mathbb{Z}}} \|u_\lambda\|_{V^2_{\mathrm{KdV}}(I)}.$$

We will usually omit I in the notation.

Theorem 6.2. *There exists $\varepsilon > 0$ such that if*

$$\|u_0\|_{\dot{B}^0_{2,\infty}} < \varepsilon$$

there is a unique global weak solution u in X with

$$\|u - S(t)u_0\|_X \le c\|u_0\|^5_{\dot{B}^0_{2,\infty}}.$$

We need Bernstein's inequality for the proof. For $q \ge p$,

$$\|u_\lambda\|_{L^q(\mathbb{R})} \le \lambda^{\frac{1}{p}-\frac{1}{q}}\|u_\lambda\|_{L^p(\mathbb{R})}. \tag{6.12}$$

Bernstein's inequality is easy to prove. Scaling reduces the question to $\lambda = 1$. So we consider u with Fourier transform supported in $[-2, 2]$. We choose a Schwartz function η with $\hat{\eta}(\xi) = 1$ for $|\xi| \le 2$. Then

$$\eta * u_1 = u_1$$

and Young's inequality gives the bound.

Proof of Theorem 6.2. We consider $T = \infty$ and claim that the assertion follows from the estimate

$$\int u_1 u_2 u_3 u_4 u_5 \partial_x v_\lambda dx dt \le c \prod_{i=1}^{5} \|u_i\|_X \|v_\lambda\|_{V^2_{\mathrm{KdV}}}. \tag{6.13}$$

Suppose that this estimate is true. We search for a solution $u = S(t)u_0 + w$, where

$$w_t + w_{xxx} + (S(t)u_0 + w)^4_x = 0,$$

with initial value $w(0) = 0$, which we formulate as a fixed point problem for the map $w \mapsto \tilde{w}$, where

$$\tilde{w}_t + \tilde{w}_{xxx} = -(S(t)u_0 + w)^5_x.$$

This equation has to be understood as follows: $\tilde{w}_\lambda$ satisfies

$$\tilde{w}_{\lambda,t} + \tilde{w}_{\lambda,xxx} = (-(S(t)u_0 + w)^5_x)_\lambda$$

in the sense of Lemma 4.34 with $a = 0$ and $b = \infty$. The derivative can be replaced by the multiplication by λ after the frequency localization.

By Lemma 4.34, there exists a unique such $\tilde{w}_\lambda \in U^2_{\mathrm{KdV}}$ with

$$\|\tilde{w}_\lambda\|_{U^2_{\mathrm{KdV}}} \le c\|S(t)u_0 + w\|^5_X$$

and, for the difference for two different data,

$$\|\tilde{w}^2_\lambda - \tilde{w}^1_\lambda\|_{U^2_{\mathrm{KdV}}} \le c\big(\|S(t)u_0 + w^1\|_X + \|S(t)u_0 + w^2\|_X\big)^4\|w^2 - w^1\|_X.$$

We take the supremum with respect to λ and arrive at, denoting the map from w to $\tilde{w}$ by J,

$$\|J(w)\|_X \le c\big(\|w\|_X + \|u_0\|_{\dot{B}^0_{2,\infty}}\big)^5,$$

$$\|J(w^2) - J(w^1)\|_X \le c\big(\|w^2\|_X + \|w^1\|_X + \|u_0\|_{\dot{B}^0_{2,\infty}}\big)^4 \|w^2 - w^1\|_X.$$

Thus J maps a ball of radius R to a ball of radius

$$c\big(R + \|u_0\|_{\dot{B}^0_{2,\infty}}\big)^4 < R$$

provided

$$\max\big\{R^3, \|u_0\|^3_{\dot{B}^0_{2,\infty}}\big\} < \frac{1}{16c}.$$

Then

$$\|J(w^2) - J(w^1)\|_X \le \frac{1}{2}\|w^2 - w^1\|_X$$

provided $\|w^j\|_X \le R$, $\|u_0\| < \frac{c^{1/3}}{10}$, and $R < \frac{1}{10c^{1/3}}$. We choose $R = \delta = \frac{1}{10c^{1/3}}$. Then J defines a contraction on the closed ball of radius R in X. The contraction mapping theorem implies existence of a unique fixed point, which by Lemma 4.34 is the unique weak solution in X. The map J is a polynomial, and hence analytic. The map J is a contraction, and this implies that its derivative is invertible. Now the analytic implicit function theorem in Banach spaces implies analytic dependence on the initial data.

These arguments remain valid, with small differences, for most dispersive equations, some wave equations, parabolic equations, and even ordinary differential equations.

It remains to prove (6.13). We expand the terms and claim that

$$\int \prod_{i=1}^{6} u_{i,\lambda_i}\, dxdt \le c\lambda_6^{-1+\frac{1}{10}} \lambda_1^{\frac{1}{2}-\frac{1}{10}} (\lambda_3\lambda_4\lambda_5)^{-\frac{1}{6}} \prod \|u_{i,\lambda_i}\|_{V^2_{\mathrm{KdV}}} \tag{6.14}$$

for $\lambda_1 \le \lambda_2 \le \lambda_3 \le \lambda_4 \le \lambda_5 \le \lambda_6$.

Let us check that this gives (6.13) by summation. We break the sum according to the relative size of λ compared to λ_i. We begin with the case $\lambda = \lambda_6$. Then necessarily $\lambda_6 \sim \lambda_5$, otherwise the frequencies cannot add up to 0, and it remains to sum over λ_i for fixed λ, taking into account that the derivative contributes a factor λ:

$$\sum_{\lambda_1 \le \lambda_2 \le \lambda_3 \le \lambda_4 \le \lambda} \lambda_1^{\frac{1}{2}-\frac{1}{10}} \lambda_3^{-\frac{1}{6}} \lambda_4^{-\frac{1}{6}} \lambda^{-\frac{1}{6}+\frac{1}{10}}$$

and to verify that this is uniformly bounded. This is done by summing first over λ_1, then λ_2, λ_3, and λ_4.

Next consider $\lambda = \lambda_4$, which leads to the sum

$$\sum_{\lambda_1 \le \lambda_2 \le \lambda_3 \le \lambda \le \lambda_6} \lambda_1^{\frac{1}{2}-\frac{1}{10}} \lambda_3^{-\frac{1}{6}} \lambda^{\frac{5}{6}} \lambda_6^{-\frac{7}{6}+\frac{1}{10}}.$$

We obtain a uniform bound by first summing with respect to λ_1, then λ_2, λ_3, and $\lambda_5 \sim \lambda_6$.

If $\lambda = \lambda_3$, we are led to

$$\sum_{\lambda_1 \le \lambda_2 \le \lambda \le \lambda_4 \le \lambda_6} \lambda_1^{\frac{1}{2}-\frac{1}{10}} \lambda^{\frac{5}{6}} \lambda_4^{-\frac{1}{6}} \lambda_6^{-\frac{7}{6}+\frac{1}{10}},$$

if $\lambda = \lambda_2$, we get

$$\sum_{\lambda_1 \le \lambda \le \lambda_3 \le \lambda_4 \le \lambda_6} \lambda_1^{\frac{1}{2}-\frac{1}{10}} \lambda \lambda_3^{-\frac{1}{6}} \lambda_4^{-\frac{1}{6}} \lambda_6^{-\frac{7}{6}+\frac{1}{10}},$$

and finally, if $\lambda = \lambda_1$,

$$\sum_{\lambda \le \lambda_2 \le \lambda_3 \le \lambda_4 \le \lambda_6} \lambda_1^{\frac{3}{2}-\frac{1}{10}} \lambda_3^{-\frac{1}{6}} \lambda_4^{-\frac{1}{6}} \lambda_6^{-\frac{7}{6}+\frac{1}{10}}.$$

None of the summations poses difficulties. We observe that $\lambda \sim \lambda_6$ has been the most difficult case, and in later proofs we often point out the most difficult case, and neglect the others. This has to be done with care.

We turn to the proof of (6.14). The Strichartz estimate (6.8) gives

$$\int \prod_{j=1}^{6} u_{j,\lambda_j}\, dxdt \le \prod \lambda_j^{-1/6} \|u_{j,\lambda_j}\|_{U^6_{\mathrm{KdV}}}.$$

The product $\prod_{j=1}^6 \lambda_j^{1/6}$ compensates for the derivative if the output frequency is λ_1, which in particular is the case if all frequencies are of the same size.

Now suppose that λ_1 is much smaller than λ_6. Then the integral vanishes unless

$$\lambda_6 - \lambda_2 \ge \frac{1}{5}\lambda_6$$

since otherwise no frequencies in the Fourier supports can add up to zero. We assume that this inequality holds and estimate using Bernstein's inequality on the first factor and the bilinear estimate (6.10) for the second and third factor:

$$\begin{aligned}
\int \prod_{j=1}^{6} u_{j,\lambda_j}\, dxdt &\le \|u_{2,\lambda_2} u_{6,\lambda_6}\|_{L^2} \|u_{1,\lambda_1}\|_{L^\infty} \prod_{j=3}^{5} \|u_{j,\lambda_j}\|_{L^6} \\
&\le \lambda_1^{1/2} (\lambda_3\lambda_4\lambda_5)^{-1/6} \lambda_6^{-1} \|u_{2,\lambda_2}\|_{U^2_{\mathrm{KdV}}} \|u_{6,\lambda_6}\|_{U^2_{\mathrm{KdV}}} \|u_{1,\lambda_1}\|_{V^\infty_{\mathrm{KdV}}} \prod_{j=3}^{5} \|u_{j,\lambda_j}\|_{U^6_{\mathrm{KdV}}}.
\end{aligned}$$

We recall the embedding $U^p \subset V^2_{rc}$ if $p > 2$. This is almost good enough, upon replacing U^2 by V^2. Let $\mu \le \frac{9}{10}\lambda$. Then

$$\begin{aligned}
\|S(t)u_{0,\mu} S(t)v_{0,\lambda}\|_{L^{\frac{25}{12}}} &\le c \|S(t)u_{0,\mu} S(t)v_{0,\lambda}\|_{L^2}^{\frac{21}{25}} \|S(t)u_{0,\mu}\|_{L^6}^{\frac{4}{25}} \|S(t)v_{0,\lambda}\|_{L^6}^{\frac{3}{25}} \\
&\le c\lambda^{-\frac{21}{25}} \lambda^{-\frac{2}{75}} \mu^{-\frac{2}{75}} \|u_{0,\mu}\|_{L^2} \|v_{0,\lambda}\|_{L^2}
\end{aligned}$$

and hence, if $\mu \le \frac{9}{10}\lambda$, then

$$\|u_\mu v_\lambda\|_{L^{\frac{25}{12}}} \le c\lambda^{-\frac{61}{75}}\mu^{-\frac{2}{75}}\|u_\mu\|_{V^2_{\mathrm{KdV}}}\|v_\lambda\|_{V^2_{\mathrm{KdV}}}. \tag{6.15}$$

Consequently,

$$\begin{aligned}\int \prod_{j=1}^{6} u_{j,\lambda_j}\,dxdt \le & \|u_{2,\lambda_2}u_{6,\lambda_6}\|_{L^{\frac{25}{12}}}\|u_{1,\lambda_1}\|_{L^{50}}\prod_{j=3}^{5}\|u_{j,\lambda_j}\|_{L^6} \\ & \le \lambda_1^{\frac{22}{50}}\|u_1\|_{L^{50}_t L^{\frac{50}{23}}{}_x}\lambda_6^{-\frac{61}{75}}\lambda_2^{-\frac{2}{75}}(\lambda_3\lambda_4\lambda_5)^{-1/6}\prod_{i=2}^{6}\|u_{i,\lambda_i}\|_{V^2_{\mathrm{KdV}}} \\ & \le \lambda_1^{\frac{21}{50}}\lambda_2^{-\frac{2}{75}}\lambda_6^{-\frac{61}{75}}(\lambda_3\lambda_4\lambda_5)^{-1/6}\prod_{i=1}^{6}\|u_{i,\lambda_i}\|_{V^2_{\mathrm{KdV}}}\end{aligned}$$

This is slightly stronger than the claimed estimate. It completes the proof. □

A variant yields local existence for large data. There are two key observations. First we may expand

$$\prod(S(t)u_0+w)_{\lambda_j} = \prod(S(t)u_0)_{\lambda_j} + \dots + \prod w_{\lambda_j};$$

there is one term without w, a term linear in w, and higher order terms in w. If w is small, then the higher order terms are even smaller. So we need some smallness of the first and the second term. We do not want to assume that the initial data are small, but we are willing to choose a small time.

Theorem 6.3. *There exists $\delta > 0$ such that, if $R > 0$,*

$$\|u_0\|_{\dot B^0_{2,\infty}} \le R$$

and with $v = S(t)u_0$,

$$(1+R^3)\sup_\lambda \lambda^{-\frac{1}{6}}\|v_\lambda\|_{L^6([0,T]\times\mathbb{R})} \le \delta, \tag{6.16}$$

then there is a unique solution u to

$$u_t + u_{xxx} + \partial_x(\chi_{[0,T]}(t)u^5) = 0$$

with initial data u_0 which satisfies

$$\|u - S(t)u_0\|_X \le cR^3\sup_\lambda \lambda^{-\frac{1}{6}}\|v_\lambda\|^2_{L^6([0,T]\times\mathbb{R})}$$

and which depends analytically on the initial data.

Proof. By the discussion above it suffices to consider integrals

$$\int_0^T \int_{\mathbb{R}} (S(t)u_0)^5 v dx dt.$$

and

$$\int_0^T \int_{\mathbb{R}} (S(t)u_0)^4 wv dx dt.$$

We observe that here we may always estimate two $S(t)u_0$ factors in L^6. Thus

$$\|w\|_X \leq cR^3\delta^2$$

which is small provided δ is sufficiently small. The rest of the proof works with virtually no change in the argument. □

The assumptions and statement of Theorem 6.3 are uniform with respect to T. Here $T = \infty$ is allowed even for large initial data. In that case the solution is in U^2_{KdV} and hence

$$w_\lambda = \lim S(-t)u_\lambda(t)$$

exists, since all one-sided limits exist. Equivalently,

$$u_\lambda(t) - S(t)w_\lambda \to 0$$

in L^2 and the solution to the nonlinear equation is for large t close to a solution to the linear equation. This is called scattering.

Suppose that

$$\lim_{\lambda\to\infty} \|u_{0,\lambda}\|_{L^2} = 0. \tag{6.17}$$

Since, by dominated convergence,

$$\lim_{T\to 0} \lambda^{-1/6}\|v_\lambda\|_{L^6([0,T]\times\mathbb{R})} = 0$$

whenever $v_\lambda \in L^6$, there exists T such that

$$\sup_{\lambda\geq 1} \lambda^{-1/6}\|v_\lambda\|_{L^6([0,T]\times\mathbb{R})} < \delta.$$

Trivially

$$\|v_\lambda\|_{L^6([0,T];L^2)} \leq cT^{1/6}\|u_{0,\lambda}\|_{L^2}$$

and, together with Bernstein's inequality for the case $\lambda \leq 1$,

$$\|v_\lambda\|_{L^6([0,T]\times\mathbb{R})} \leq \lambda^{\frac{1}{2}} T^{\frac{1}{6}} (\lambda^{-\frac{1}{6}}\|u_{0,\lambda}\|_{L^2}),$$

which is much stronger than needed to ensure the smallness assumption (6.16) for sufficiently small time. As a consequence we obtain existence of unique local solutions provided (6.17) is satisfied.

Since there are solitons, in general solutions are not in L^6 of space-time. Solitons clearly do not scatter. This version of well-posedness has been proven by Strunk [28]. The result in L^2 is due to Kenig, Ponce, and Vega.

We next turn to

$$u_t + u_{xxx} + u_x^4 = 0. \tag{6.18}$$

Here $\dot{H}^{-1/6}$ is the critical Sobolev space. We choose a slightly larger space,

$$\|u\|_X = \sup_{\lambda \in 1.01^{\mathbb{Z}}} \lambda^{-1/6} \|u_\lambda\|_{V^2_{\mathrm{KdV}}(0,\infty)}$$

for the solution and use

$$\|u_0\|_{\dot{B}^{-1/6}_{2,\infty}} = \sup_{\lambda \in 1.01^{\mathbb{Z}}} \lambda^{-1/6} \|u_{0,\lambda}\|_{L^2}.$$

Then

$$\sup_\lambda \lambda^{-1/6} \|S(t)u_{0,\lambda}\|_{V^2_{\mathrm{KdV}}} \sim \sup_\lambda \lambda^{-1/6} \|u_{0,\lambda}\|_{L^2}.$$

Theorem 6.4. *There exists $\delta > 0$ such that for all u_0 with*

$$\|u_0\|_{\dot{B}^{-\frac{1}{6}}_{2,\infty}} < \delta.$$

there is a unique global solution u which satisfies

$$\|u - S(t)u_0\|_X \le c\|u_0\|^4_{\dot{B}^{-1/6}_{2,\infty}}$$

and which depends analytically on the initial data.

Proof. We claim that

$$\left| \int u_1 u_2 u_3 u_4 v_\lambda dxdt \right| \le \lambda^{-\frac{5}{6}} \prod \|u_i\|_X \|v_\lambda\|_{V^2}. \tag{6.19}$$

The theorem follows from this estimate in the same fashion as for $p = 4$. As in that case, (6.19) follows from

$$\left| \int u_1 u_2 u_3 u_4 v_\lambda dxdt \right| \le \lambda^{-\frac{5}{6}} \prod \|u_i\|_X \|v_\lambda\|_{V^2}. \tag{6.20}$$

To prove it we expand the left-hand side into a dyadic sum and we try to bound

$$I = \left| \int \prod_{i=1}^{5} u_{i,\lambda_i} dxdt \right|$$

where (by symmetry) $\lambda_1 \le \lambda_2 \le \lambda_3 \le \lambda_4 \le \lambda_5$. We claim that

$$\left| \int \prod_{i=1}^{5} u_{i,\lambda_i} dxdt \right| \le c_\varepsilon \lambda_5^{-1} (\lambda_2\lambda_3\lambda_4)^{-1/6} (\lambda_5/\lambda_1)^\varepsilon \prod \|u_{i,\lambda_i}\|_{V^2}. \tag{6.21}$$

We assume that (6.21) holds. The integral with respect to x vanishes unless there are frequencies in the support of the Fourier transform which add up to zero. Since, if $|\xi_1| \le |\xi_2| \le \cdots \le |\xi_5|$, the frequencies can only add up to zero, $\xi_1+\xi_2+\xi_3+\xi_4+\xi_5 = 0$, if $|\xi_5| - |\xi_1| \ge \frac{1}{10}|\xi_5|$, which we restrict to in the sequel. We observe that we may restrict to $\lambda_4 \ge \lambda_5/8$ – otherwise the integral vanishes. The summations is done as for $p = 4$.

It remains to prove (6.21). We have seen that we may assume that $\lambda_1 \le 4\lambda_5/5$ and $\lambda_4 \ge \lambda_5/8$. The first attempt is

$$\begin{aligned} I &\le \|u_{1,\lambda_1} u_{5,\lambda_5}\|_{L^2} \prod_{j=2}^{4} \|u_{j,\lambda_j}\|_{L^6} \\ &\le (\lambda_2\lambda_3\lambda_4)^{-1/6} \lambda_5^{-1} \|u_{1,\lambda_1}\|_{U^2_{\mathrm{KdV}}} \|u_{5,\lambda_5}\|_{U^2_{\mathrm{KdV}}} \prod_{j=2}^{4} \|u_{j,\lambda_j}\|_{U^6_{\mathrm{KdV}}}, \end{aligned} \tag{6.22}$$

where we used Hölder's inequality for the first inequality, the bilinear estimate for the first factor, and the L^6 Strichartz embedding for the remaining factors. This is almost what we need – we still have to replace the norm U^2_{KdV} by V^2_{KdV}.

The Strichartz estimates imply

$$\|S(t)u_{0,\lambda} S(t)u_{0,\mu}\|_{L^3} \le c(\lambda\mu)^{-1/6} \|u_{0,\mu}\|_{L^2} \|u_{0,\lambda}\|_{L^2}$$

and the bilinear estimate is, for $\mu \le \lambda/1.03$,

$$\|S(t)u_{0,\lambda} S(t)u_{0,\mu}\|_{L^2} \le c\lambda^{-1} \|u_{0,\mu}\|_{L^2} \|u_{0,\lambda}\|_{L^2}.$$

Thus, for $2 \le p \le 3$,

$$\|S(t)u_{0,\lambda} S(t)u_{0,\mu}\|_{L^p} \le c\lambda^{-6(\frac{1}{p}-\frac{1}{3})} (\lambda\mu)^{-(\frac{1}{2}-\frac{1}{p})} \|u_{0,\mu}\|_{L^2} \|u_{0,\lambda}\|_{L^2},$$

and hence, by Hölder's inequality,

$$\|u_\lambda u_\mu\|_{L^p} \le c\lambda^{2-\frac{5}{p}} \mu^{\frac{1}{p}-\frac{1}{2}} \|u_\mu\|_{U^p_{\mathrm{KdV}}} \|u_\lambda\|_{U^p_{\mathrm{KdV}}}.$$

With this argument we may replace the U^2 by V^2 norms, but now the remaining terms are no longer square integrable. We use this modified bilinear estimate twice if there are two pairs of λ_i with ratio at least $\ge 1.01^2$. Oversimplifying slightly this leaves us with $\lambda_2 = \lambda_3 = \cdots = \lambda_5$ and $\lambda_1 = \lambda_2 = \lambda_3 = \lambda_4$ and $\lambda_5 \sim 3\lambda_1$. The second case is easier, and we focus on the first. We again turn our attention to

$$\xi_1 + \xi_2 + \xi_3 + \xi_4 + \xi_5 = 0,$$

assuming $|\xi_1| \le |\xi_2| \le |\xi_3| \le |\xi_4| \le |\xi_5|$. We have already seen that $|\xi_1| \le 0.9|\xi_5|$. We decompose the set $\{\xi : \lambda_j \le |\xi| < 1.01\lambda_j\}$ for $2 \le j \le 5$ into symmetric unions of intervals of length $\lambda_1/100$. We label these intervals by μ_{ij} with $2 \le i \le 5$ and $j \le \lambda_5/\lambda_1$, omit the index i for simplicity and expand:

$$\int u_{1,\lambda_1} u_{2,\lambda_5} u_{3,\lambda_5} u_{4,\lambda_5} u_{5,\lambda_5}\, dxdt = \sum_{90 \le \left| \sum_{j=2}^{5} \mu_j \right| \le 110} \int u_{1,\lambda_1} u_{2,\mu_2} u_{3,\mu_3} u_{4,\mu_4} u_{5,\mu_5}\, dxdt;$$

the sum contains at most $\sim (\lambda_5/\lambda_1)^4$ terms. We fix μ_j and assume that they are ordered. Then $\mu_5 - \mu_2 \ge 2$ and we estimate

$$\left| \int u_{1,\lambda_1} u_{2,\mu_2} u_{3,\mu_3} u_{4,\mu_4} u_{5,\mu_5}\, dxdt \right| \le \| u_{\lambda_1} u_{4,\mu_4} \|_{L^p} \| u_{\mu_2} u_{\mu_5} \|_{L^q} \| u_{\lambda_3} \|_{L^6},$$

and hence (changing indices if necessary, or summing over similar terms)

$$\left| \int u_{1,\lambda_1} u_{2,\mu_2} u_{3,\mu_3} u_{4,\mu_4} u_{5,\mu_5}\, dxdt \right| \le c \lambda_5^{-1} (\lambda_2\lambda_3\lambda_4)^{-1/6} (\lambda_5/\lambda_1) \prod \| u_{i,\lambda_i} \|_{U^p}, \tag{6.23}$$

since p is the smallest exponent. This is almost good, but $(\lambda_5/\lambda_1)^5$ is too big.

We recall Lemma 4.12, which allows us to write for $M \ge 1$,

$$u = v + w,$$

with

$$\frac{\kappa}{M} \| w \|_{U^2_{\mathrm{KdV}}} + e^M \| v \|_{U^p_{\mathrm{KdV}}} \le \| u \|_{V^2_{\mathrm{KdV}}}.$$

We expand all the u_i. This yields, by (6.22),

$$\left| \int v_{1,\lambda_1} v_{2,\mu_2} v_{3,\mu_3} v_{4,\mu_4} v_{5,\mu_5}\, dxdt \right| \le c M^5 \lambda_5^{-1} (\lambda_2\lambda_3\lambda_4)^{-\frac{1}{6}} \prod \| u_{i,\lambda_i} \|_{V^2_{\mathrm{KdV}}}$$

and

$$\begin{aligned} &\left| \int w_{1,\lambda_1} v_{2,\lambda_2} v_{3,\lambda_3} v_{4,\lambda_4} v_{5,\lambda_5}\, dxdt \right| \\ &\le c \lambda_5^{-1} (\lambda_2\lambda_3\lambda_4)^{-1/6} (\lambda_5/\lambda_1)^5 \| w_{1,\lambda_1} \|_{u^p_{\mathrm{KdV}}} \prod_{i=2}^{5} \| v_{i,\lambda_i} \|_{U^p_{\mathrm{KdV}}} \\ &\le e^{-M} \lambda_5^{-1} (\lambda_2\lambda_3\lambda_4)^{-1/6} (\lambda_5/\lambda_1)^5 \prod \| u_{i,\lambda_i} \|_{V^2_{\mathrm{KdV}}}. \end{aligned}$$

Similarly we estimate all the other terms in the expansion. Then

$$\begin{aligned} \left| \int u_{1,\lambda_1} u_{2,\mu_2} u_{3,\mu_3} u_{4,\mu_4} u_{5,\mu_5}\, dxdt \right| &\le c(M^5 + e^{-M}(\lambda_5/\lambda_1)^5) \lambda_5^{-1} (\lambda_2\lambda_3\lambda_4)^{-1/6} \\ &\qquad \times \prod \| u_{i,\lambda_i} \|_{V^2_{\mathrm{KdV}}}. \\ &\le c \ln(1 + (\lambda_5/\lambda_1))^5 \lambda_5^{-1} (\lambda_2\lambda_3\lambda_4)^{-1/6} \prod \| u_{i,\lambda_i} \|_{V^2_{\mathrm{KdV}}}. \end{aligned}$$

if we choose $M = 5 \ln(\lambda_5/\lambda_1)$. This completes the proof of estimate (6.19), and hence the proof of the theorem. □

Again there are similar refinements, as for the critical gKdV equation. Well-posedness in slightly smaller spaces has been proven by Grünrock [10] and Tao [30] based on a modification of the Fourier restriction spaces of Bourgain at the critical level.

The statement and proof are based on [20], where it was one step in proving stability of the soliton in $\dot{B}^{-1/6}_{2,\infty}$, and scattering, which is probably the first stability result of solitons for gKdV which is not based on Weinstein's convexity argument.

Next we turn to the modified KdV equation

$$u_t + u_{xxx} + u_x^3 = 0. \tag{6.24}$$

The space $\dot{H}^{-1/2}$ is scaling invariant, but we are not able the reach the critical space. Instead we construct global in time solutions to

$$u_t + u_{xxx} + \partial_x(\chi_{[0,T]} u^3) = 0$$

for given initial data u_0 and $T > 0$. We aim for a scale invariant formulation. Given $T > 0$, we define the equivalent norms on $B^{\frac{1}{4}}_{2,\infty}$,

$$\|u_0\|_E = \max\left\{ T^{\frac{1}{6}} \|u_{0,<T^{-\frac{1}{3}}}\|_{L^2}, \sup_{\lambda \geq T^{-\frac{1}{3}}} (\lambda T)^{\frac{1}{4}} \|u_{0,\lambda}\|_{L^2} \right\}$$

and

$$\|u\|_X = \max\left\{ T^{\frac{1}{6}} \|u_{<T^{-\frac{1}{3}}}\|_{V^2_{\mathrm{KdV}}}, \sup_{\lambda \geq T^{-\frac{1}{3}}} (\lambda T)^{\frac{1}{4}} \|u_\lambda\|_{V^2_{\mathrm{KdV}}} \right\}.$$

Well-posedness by different arguments has been shown by [17] in a slightly smaller space of initial data.

Theorem 6.5. *There exists $\varepsilon > 0$ such that for $u_0 \in B^{\frac{1}{4}}_{2,\infty}$ with*

$$\|u_0\|_E \leq \varepsilon$$

there is a unique weak solution $u \in X$ with

$$\|u - S(t)u_0\|_X \leq c\|u_0\|_E^3.$$

Proof. We want to construct a fixed point of

$$v = \int_0^t S(t-s)\chi_{[0,T]}(s)\partial_x(w+v)^3 ds.$$

The key estimate (for small data) is

$$\lambda^{\frac{1}{4}} \left| \int_{\mathbb{R}\times R} \chi_{[0,T]} u_1 u_2 u_3 \partial_x v_\lambda dxdt \right| \leq c \prod_{j=1}^{3} \|u_j\|_X \|v_\lambda\|_{V^2_{\mathrm{KdV}}}. \tag{6.25}$$

The theorem follows from it by repeating the arguments for the L^2 critical case.

To prove (6.25) we expand the left-hand side into a dyadic sum. The pieces are estimated by

$$\left| \int_0^T \int_{\mathbb{R}} \prod_{j=1}^4 u_{i,\lambda_i} dxdt \right| \le cT^{\frac{1}{2}} \prod_{j=1}^4 \lambda_j^{-1/8} \|u_{j,\lambda_j}\|_{U^8_{\mathrm{KdV}}} \tag{6.26}$$

if $\lambda_1 \ge 2$ using the Strichartz embedding of Theorem 3.2

$$\|u_{j,\lambda_j}\|_{L^8_t L^4_x} \le c\lambda_j^{-1/8} \|u_{j,\lambda_j}\|_{U^8_{\mathrm{KdV}}}.$$

This is good enough if $\lambda_1 \sim \lambda_4$ and $\lambda_1 \ge 2$. If $\mu \le \lambda/4$, one has the bilinear estimate

$$\begin{aligned} \|S(t)u_{0,\mu}S(t)v_{0,\lambda}\|_{L_t^{\frac{8}{3}} L^2_x} &\le \|S(t)u_{0,\mu}S(t)v_{0,\lambda}\|^{\frac{1}{2}}_{L^4_t L^2_x} \|S(t)u_{0,\mu}S(t)v_{0,\lambda}\|^{\frac{1}{2}}_{L^2_t L^2_x} \\ &\le c\lambda^{-\frac{9}{8}} \mu^{-\frac{1}{16}} \|u_{0,\mu}\|_{L^2} \|v_{0,\lambda}\|_{L^2} \end{aligned}$$

and hence, if $\lambda_1 \le \lambda_3/4$ and $\lambda_2 \le \lambda_4/4$,

$$\begin{aligned} \left| \int_0^T \int \prod_{j=1}^4 u_{i,\lambda_i} dxdt \right| &\le cT^{\frac{1}{4}} \|u_{1,\lambda_1} u_{3,\lambda_3}\|_{L^{\frac{8}{3},2}} \|u_{2,\lambda_2} u_{4,\lambda_4}\|_{L^{\frac{8}{3},2}} \\ &\le cT^{\frac{1}{4}} \lambda_4^{-\frac{9}{8}} \lambda_1^{-\frac{1}{16}} \lambda_2^{-\frac{1}{16}} \prod_j \|u_{j,\lambda_j}\|_{V^2_{\mathrm{KdV}}}. \end{aligned} \tag{6.27}$$

If $\lambda_4 \le 2T^{\frac{1}{3}}$, we estimate

$$\begin{aligned} \int_0^T \int_{\mathbb{R}} \prod_{i=1}^4 u_{i,\lambda_i} dxdt &\le T \|u_{1,\lambda_1}\|_{L^\infty} \|u_{2,\lambda_2}\|_{L^\infty} \|u_{3,\lambda_3}\|_{L^\infty L^2} \|u_{4,\lambda_4}\|_{L^\infty L^2} \\ &\le c \prod \|u_{i,\lambda_i}\|_{V^\infty_{\mathrm{KdV}}}. \end{aligned}$$

Checking the support we see that the integral vanishes unless either $\lambda_1 \ge \lambda_4/16$ or $\lambda_1 \le \lambda_3/4$ and $\lambda_2 \le \lambda_4/4$ or $\lambda \le 16$.

We turn to the summation.

1. $\lambda > \lambda_4/16$, $\lambda_4 \ge 16T^{-\frac{1}{3}}$. The sum can be bounded using (6.26) for $\lambda_1 \ge \lambda_4/16$ and (6.27) for $\lambda_1 \le \lambda_4/16$ and $\lambda_4 \ge 16$, where the sum takes the form

$$\begin{aligned} \Big(\sum_{1 \le \lambda_1 \le \lambda_2 \le \lambda_4/4} & (T^{\frac{1}{3}}\lambda_1)^{-\frac{5}{16}} (T^{\frac{1}{3}}\lambda_2)^{-\frac{5}{16}} (T^{\frac{1}{3}}\lambda_4)^{-\frac{1}{8}} \\ & \times (T\lambda_1)^{\frac{1}{4}} \|u_{1,\lambda_1}\|_{V^2_{\mathrm{KdV}}} (T\lambda_2)^{\frac{1}{4}}) \|u_{2,\lambda_2}\|_{V^2_{\mathrm{KdV}}} \Big) \|u_{4,\lambda_4}\|_{V^2_{\mathrm{KdV}}} \|v_\lambda\|_{V^2_{\mathrm{KdV}}}. \end{aligned}$$

The bound is obvious.

2. $\max\{T^{-\frac{1}{3}}, \lambda\} \le \lambda_4/16$. Here we use (6.27). The uniform bound for the sum is immediate.
3. $\lambda_4 \le 16T^{-1/3}$. Now the estimate follows from the last estimate.

The proof is complete. $\square$

The proof could easily be simplified by first rescaling to $T = 1$. The advantage of the presented proof above is that it makes the behavior of all terms with respect to scaling transparent.

Finally, we study the Korteweg–de-Vries equation

$$u_t + u_{xxx} + u_x^2 = 0.$$

The well-posedness result in $H^{-\frac{3}{4}}$ is due to Christ, Colliander, and Tao [5], who also prove that below $-\frac{3}{4}$ some sort of ill-posedness must occur. Despite this there are uniform global a-priori estimates in H^{-1}, see [3]. Uniqueness between $-\frac{3}{4}$ and -1 is entirely open.

We seek a solution u to

$$u_t + u_{xxx} + \partial_x(\chi[0,1](t)u^2) = 0$$

with the given initial data. We again make the ansatz

$$u = v + w,$$

where $v = S(t)u_0$ and

$$w_t + w_{xxx} + \partial_x(\chi(t)(v + w)^2) = 0.$$

The identity

$$(\xi_1 + \xi_2)^3 - \xi_1^3 - \xi_2^3 = 3\xi_1\xi_2(\xi_1 + \xi_2)$$

describes the vertical distance of the sum of two points (τ_j, ξ_j) from the characteristic set. We will make use of this property through 'high modulation' L^2 estimates. For this purpose we fix a smooth function ϕ supported in $[-2, 2]$, and identically 1 in $[-1, 1]$, and define $u^\Lambda(t)$ by the Fourier multiplier $1 - \phi(\tau/\Lambda)$. The Fourier multiplier $\phi(\tau/\Lambda)$ defines a convolution. Let ψ be the inverse Fourier transform. Then up to a power of $\sqrt{2\pi}$,

$$(1 - \phi(\tau/\Lambda))u = u - \Lambda\psi(\Lambda t) * u$$

and hence, by Lemma 4.17,

$$\|u^\Lambda\|_{U^p} \le c\|u\|_{U^p},$$

$$\|u^\Lambda\|_{V^p} \le c\|u\|_{V^p}.$$

Theorem 6.6. *There exists $\delta > 0$ such that for all initial data satisfying*

$$\|u_0\|_{B^{-\frac{3}{4}}_{2,\infty}} < \delta$$

there is a unique function $u \in X$ where X will be defined below, with

$$\|\chi_T u - S(t)u\|_X \le c\|u\|^2_{B^{-\frac{3}{4}}_{2,\infty}}$$

which solves the equation up to time 1*. It depends analytically on the initial data.*

Proof. We define the sets

$$A(0) = \{(\tau,\xi) | |\xi| \le 1, |\tau - \xi^3| \le 1\},$$

$$A(\lambda) = \{(\tau,\xi) | \lambda \le |\xi| \le 2\lambda, |\tau - \xi^3| \le \lambda^3\},$$
$$B(\lambda) = \{(\tau,\xi) | |\xi| \le \lambda, 1 \le |\tau - |\xi|^3| \le |\xi|\lambda^2\}.$$

Then, denoting the Fourier projections to a set D simply by an index D,

$$\begin{aligned}\Big\| |D_x|^{\frac{1}{2}} \int S(t-s)\rho(s)\partial_x(u_{A(\lambda)}u_{A(\lambda)})dt_{B(\lambda)}\Big\|_{L^2} &\le \lambda^{-2}\| |D_x|^{\frac{1}{2}} u_{A(\lambda)}u_{A(\lambda)}\|_{L^2} \\ &\le \lambda^{-\frac{5}{2}}\|u_{A(\lambda)}\|^2_{U^2_{\mathrm{KdV}}},\end{aligned}$$

which is scale invariant. Alternatively, we may estimate

$$\begin{aligned}&\Big\| |D_x|^{-\frac{1}{2}} \int S(t-s)\rho(s)\partial_x(u_{A(\lambda)}u_{A(\lambda)})dt_{\mu,B(\lambda)}\Big\|_{L^2} \\ &\qquad\le \lambda^{-1}\Big\| \int S(t-s)\rho(s)u_{A(\lambda)}u_{A(\lambda)}ds_{\mu,B(\lambda)}\Big\|_{V^2_{\mathrm{KdV}}} \\ &\qquad\le \lambda^{-2}\|u_{A(\lambda)}\|_{U^2_{\mathrm{KdV}}}\|u_{A(\lambda)}\|_{U^2_{\mathrm{KdV}}}.\end{aligned}$$

Observe that the two terms are of the same size for $\mu = \lambda^{-1/2}$. More precisely, the L^2 norm is of unit size:

$$\Big\| \int S(t-s)\rho(s)\partial_x(u_{A(\lambda)}u_{A(\lambda)})_{\lambda^{-1/2},B(\lambda)}\Big\|_{L^2} \le c\lambda^{-\frac{3}{4}}\big(\lambda^{-\frac{3}{4}}\|u_{A(\lambda)}\|_{U^2_{\mathrm{KdV}}}\big)^2.$$

There is nothing to loose, and hence we need to control $u_{A(\lambda)}$ in U^2_{KdV}. Similarly

$$\Big\| \int S(t-s)\rho(s)\partial_x(u_{A(\lambda)}u_{B(\lambda)})_{A(\lambda)}ds\Big\|_{V^2_{\mathrm{KdV}}} \le c\lambda^{\frac{1}{2}}\|u_{A(\lambda)}\|_{U^2_{\mathrm{KdV}}}\| |D_x|^{-\frac{1}{2}}u_{B(\lambda)}\|_{L^2}$$

and

$$\Big\| \int S(t-s)\rho(s)\partial_x(u_{A(\lambda)}u_{\mu,B(\lambda)})_{A(\lambda)}ds\Big\|_{U^2_{\mathrm{KdV}}} \le c\|u_{A(\lambda)}\|_{U^2_{\mathrm{KdV}}}\| |D_x|^{\frac{1}{2}}u_{\mu,B(\lambda)}\|_{L^2}.$$

This is the only place which does not allow us to go beyond $B^{-\frac{3}{4}}_{2,\infty}$.
We define the function space X by the norm

$$\begin{aligned}\|u\|_X = \sup_{\lambda\ge 1}\{&\lambda^{-\frac{3}{4}}\|u_{A(\lambda)}\|_{U^2_{KdV}}, \lambda^{3/4}\|(\lambda^{-\frac{1}{2}}|D_x| + \lambda^{\frac{1}{2}}|D|^{-1})^{\frac{1}{4}} u_{B(\lambda)}\|_{L^2}\}\\ &+ \|u_{A(0)}\|_{L^2_x L^\infty_t}.\end{aligned}$$

We only sketch the estimates, and choose the most instructive ones, using dyadic decompositions on the Fourier side, bilinear estimates and modulation estimates. We have chosen the most instructive estimates for a sketch of the proof. We used a dyadic decomposition on the Fourier side, bilinear estimates and modulation estimates. Similarly,

$$\begin{aligned}&\left\|\left(\int_0^\infty S(t-s)\chi(s)\partial_x(u_{A(\lambda)}u_{B(\lambda)})ds\right)_{A(\lambda)}\right\|_{V^2_{KdV}}\\ &\le c\lambda^{1/2}\||D_x|^{-1/2}u_{B(\lambda)}\|_{L^2}\|u_{A(\lambda)}\|_{U^2_{KdV}}\end{aligned}$$

and for $\mu \le \lambda$,

$$\begin{aligned}\left\|\left(\int_0^\infty S(t-s)\chi(s)\partial_x(u_{A(\lambda)}u_{B(\lambda),\mu})ds\right)_{A(\lambda)}\right\|_{U^1_{KdV}} &\le c\lambda\|u_{A(\lambda)}u_{B(\lambda),\mu}\|_{U^2_{KdV}}\\ &\le c\|u_{B(\lambda),\mu}\|_{L^2}\|u_{A(\lambda)}\|_{U^2_{KdV}},\end{aligned}$$

hence, by summation and logarithmic interpolation

$$\begin{aligned}&\left\|\left(\int_0^\infty S(t-s)\chi(s)\partial_x(u_{A(\lambda)}u_{B(\lambda)})ds\right)_{A(\lambda)}\right\|_{U^2_{KdV}}\\ &\le c\|u_{A(\lambda)}\|_{U^2_{KdV}}\lambda^{-1/8}\|D^{-1/4}u_{B(\lambda)}\|_{L^2}.\end{aligned}$$

There are many more terms, but they are easier to deal with, which we omit. $\square$

The interest in this setup is twofold: First it shows how to go beyond $H^{-\frac{3}{4}}$. Second, X is not a subset of $L^\infty(\mathbb{R}; B^{-\frac{3}{4}}_{2,\infty})$, and one has to use energy estimates to see that the solution is bounded and weakly continuous as a map to $B^{-\frac{3}{4}}_{2,\infty}$. This difficulty is related to the classical ill-posedness results: the flow map does not extend to a differentiable map from the initial data to $u(t) \in \mathcal{S}$ below $-\frac{3}{4}$.

6.3 The derivative nonlinear Schrödinger equation

We consider

$$iu_t + \Delta u = \bar{u}\partial_1\bar{u}. \tag{6.28}$$

This equation has no significance for applications as far as I know. The choice of the non-linearity is crucial. If u satisfies (6.28) then the same is true for

$$\lambda u(\lambda^2 t, \lambda x)$$

and the underlying critical space is $\dot H^{\frac{d-2}{2}}$.

The Strichartz with $\frac{2}{4} + \frac{d}{p} = \frac{d}{2}$ and Bernstein give for $d \geq 2$

$$\|u_\lambda\|_{L^4(\mathbb{R}\times\mathbb{R}^d)} \leq \lambda^{\frac{d-2}{4}} \|u_{i,\lambda_i}\|_{L^4_t L^p_x(\mathbb{R}^d)} \leq \lambda^{\frac{d-2}{4}} \|u_{i,\lambda_i}\|_{U^4_{i\Delta}}. \tag{6.29}$$

The corresponding bilinear estimates are

$$\|u_\lambda v_\mu\|_{L^2} \leq c\mu^{\frac{d-1}{2}} \lambda^{-1/2} \|u_\lambda\|_{U^2_{i\Delta}} \|v_\mu\|_{U^2_{i\Delta}} \tag{6.30}$$

and

$$\|(u_\lambda v_\lambda)_\mu\|_{L^2} \leq c\mu^{\frac{d-2}{2}} \|u_\lambda\|_{U^4_{i\Delta}} \|v_\lambda\|_{U^4_{i\Delta}} \tag{6.31}$$

if $\mu < \lambda/4$.

This time we need the complex inner product. The modulation relation is

$$\xi_1^2 + \xi_2^2 + (-\xi_1 - \xi_2)^2 \geq \xi_1^2 + \xi_2^2,$$

which is a particularly pleasant situation.

The dyadic estimates become, for $\lambda_1 \ll \lambda_2 \sim \lambda_3$, with u^h denoting the part with modulation at least $(|\xi_1|^2 + |\xi_2|^2 + |\xi_3|^2)/10$,

$$\left| \int u^h_{\lambda_1} u_{\lambda_2} u_{\lambda_3} dx\, dt \right| \leq c\lambda_3^{-1} \lambda_1^{\frac{d-2}{2}} \|u_{1,\lambda_1}\|_{V^2_{i\Delta}} \|u_{2,\lambda_2}\|_{V^2_{i\Delta}} \|u_{3,\lambda_3}\|_{V^2_{i\Delta}} \tag{6.32}$$

and

$$\left| \int u_{\lambda_1} u^h_{\lambda_2} u_{\lambda_3} dx\, dt \right| \leq c\lambda_3^{-\frac{3}{2}} \lambda_1^{\frac{d-1}{2}} \|u_{1,\lambda_1}\|_{U^2_{i\Delta}} \|u_{2,\lambda_2}\|_{V^2_{i\Delta}} \|u_{3,\lambda_3}\|_{U^2_{i\Delta}}, \tag{6.33}$$

whence

$$\left| \int \prod u_{1,\lambda_1} u^h_{2,\lambda_2} u_{3,\lambda_3} dx dt \right| \leq c\lambda_3^{-\frac{3}{2}} \lambda_1^{\frac{d-1}{2}} (\lambda_3/\lambda_1)^{\varepsilon} \prod_{i=1}^3 \|u_{i,\lambda_i}\|_{V^2_{i\Delta}}.$$

Theorem 6.7. *Let $d = 2$. There exists $\varepsilon > 0$ such that if*

$$\|u_0\|_{L^2} < \varepsilon,$$

then there is a unique solution to

$$iu_t + \Delta u = \bar u \partial_{x_1} \bar u$$

with

$$\|u\|_X := \left(\sum_{\lambda \in 2^{\mathbb{Z}}} \|u_\lambda\|^2_{U^2_{i\Delta}} \right)^{1/2} \leq c\|u_0\|_{L^2}.$$

If $d \geq 3$, there exists $\varepsilon > 0$ such that if

$$\|u_0\|_{\dot{B}^{\frac{d-2}{2}}_{2,1}} = \sum_\lambda \lambda^{\frac{d-2}{2}} \|u_{0,\lambda}\|_{L^2} < \varepsilon,$$

then there is a unique weak solution with

$$\|u\|_X := \sum_\lambda \lambda^{\frac{d-2}{2}} \|u_\lambda\|_{U^2} \leq c\|u_0\|_{\dot{B}^{\frac{d-2}{2}}_{2,1}}.$$

Proof. The key estimates are again

$$\left| \int_{\mathbb{R}\times\mathbb{R}^d} (\partial_{x_1} \bar{u}_1) \bar{u}_2 \bar{v} dx dt \right| \leq \|u_1\|_X \|u_2\|_X \left(\sum_\lambda \|v_\lambda\|^2_{V^2_{\mathrm{KdV}}} \right)^{1/2},$$

resp.

$$\left| \int_{\mathbb{R}\times\mathbb{R}^d} (\partial_{x_1} \bar{u}_1) \bar{u}_2 \bar{v} dx dt \right| \leq \|u_1\|_X \|u_2\|_X \sup_\lambda \lambda^{-\frac{d-2}{2}} \|v_\lambda\|_{V^2_{\mathrm{KdV}}}$$

if $d \geq 3$. We abuse the notation and set $\lambda_2 = \lambda_3 = \lambda$ and compute for $d = 2$

$$\begin{aligned} \sum_{\mu<\lambda} \lambda \left| \int \bar{u}^h_\mu \bar{u}_{2,\lambda} \bar{v}_\lambda dx dt \right| &\leq \sum_{\mu \leq \lambda} \lambda \|u^h_\mu\|_{L^2} \|(u_{2,\lambda} v_\lambda)_\mu\|_{L^2} \\ &\leq \left(\sum_{\mu\leq\lambda} \|u_{1,\mu}\|_{V^2_{i\Delta}})^2 \right)^{1/2} \|u_\lambda v_\lambda\|_{L^2(\mathbb{R}^2)} \\ &\leq \|u_1\|_X \|u_\lambda\|_{U^4_{i\Delta}} \|v_\lambda\|_{U^4_{i\Delta}}. \end{aligned}$$

The factor λ^{-1} compensates for the derivative. The summation with respect to λ is trivial. The estimate is easier if the high modulation falls on other terms:

$$\begin{aligned} \sum_{\mu<\lambda} \lambda \left| \int \bar{u}_\mu \bar{u}^h_{2,\lambda} \bar{v}_\lambda dx dt \right| &\leq \sum_{\mu\leq\lambda} \lambda \|u^h_{2,\lambda}\|_{L^2} \|u_{1,\mu} v_\lambda\|_{L^2} \\ &\leq \mu^{1/2} \|u_{1,\mu}\|_{V^2_{i\Delta}} \lambda^{-1/2} \|u_{1,\mu}\|_{U^2_{i\Delta}} \|v_\lambda\|_{U^2_{i\Delta}}. \end{aligned}$$

By logarithmic interpolation (Lemma 4.12),

$$\begin{aligned} \sum_{\mu<\lambda} \lambda \left| \int \bar{u}_\mu \bar{u}^h_{2,\lambda} \bar{v}_\lambda dx dt \right| &\leq \sum_{\mu\leq\lambda} \lambda \|u^h_{2,\lambda}\|_{L^2} \|u_{1,\mu} v_\lambda\|_{L^2} \\ &\leq \sum_{\mu\leq\lambda} (\mu/\lambda)^{\frac{1}{2}-\varepsilon} \|u_{1,\mu}\|_{V^2_{i\Delta}} \lambda^{-1/2} \|u_{1,\mu}\|_{V^2_{i\Delta}} \|v_\lambda\|_{V^2_{i\Delta}} \end{aligned}$$

and the summation is straightforward.

The modification for $d \geq 3$ is simple: We give up orthogonality and sum for the first estimate

$$\begin{aligned}\sum_{\mu<\lambda} \lambda \left| \int \bar{u}^h_\mu \bar{u}_{2,\lambda} \bar{v}_\lambda dxdt \right| &\leq \sum_{\mu\leq\lambda} \lambda \|u^h_\mu\|_{L^2} \|(u_{2,\lambda} v_\lambda)_\mu\|_{L^2} \\ &\leq \sum_{\mu\leq\lambda} \mu^{\frac{d-2}{2}} \|u_{1,\mu}\|_{V^2_{i\Delta}} \|u_\lambda\|_{V^2_{i\Delta}} \|v_\lambda\|_{V^2_{i\Delta}}.\end{aligned}$$

For the second estimate we put in powers of μ resp. λ. □

6.4 The Kadomtsev–Petviashvili II equation

The Kadomtsev–Petviashvili II (KP II) equation

$$\begin{aligned}\partial_x(\partial_t u + \partial_x^3 u + u\partial_x u) + \partial_y^2 u = 0 \quad &\text{in } (0,\infty)\times\mathbb{R}^2, \\ u(0,x,y) = u_0(x,y), \quad &(x,y)\in\mathbb{R}^2,\end{aligned} \tag{6.34}$$

has been introduced by B.B. Kadomtsev and V.I. Petviashvili [14] to describe weakly transverse water waves in the long wave regime with small surface tension. It generalizes the Korteweg–de Vries equation, which is spatially one-dimensional and thus neglects transversal effects. The KP II equation has a remarkably rich structure.

Here we describe a setup leading to global well-posedness and scattering for small data. The Hilbert space will be denoted by $\dot{H}^{-\frac{1}{2},0}$, and is defined by the norm

$$\|u_0\|_{\dot{H}^{-\frac{1}{2},0}} = \||\xi|^{-1/2}\hat{u}_0\|_{L^2},$$

where ξ is the Fourier multiplier with respect to x. The Fourier multiplier $|\xi|^{-\frac{1}{2}}$ defines an isomorphism from L^2 to $\dot{H}^{-\frac{1}{2},0}$.

For $\lambda > 0$ we define the projection to the range $1 \leq |\xi|/\lambda < 2$ by

$$\mathcal{F}(u_\lambda) = \chi_{\lambda\leq|\xi|\leq 2\lambda}\mathcal{F}u,$$

where $\mathcal{F}$ denotes the Fourier transform. Usually we choose $\lambda \in 2^{\mathbb{Z}}$, the set of integer powers of 2.

Let $u(t) = S(t)u_0$. The Strichartz estimate is

$$\|u\|_{L^4(\mathbb{R}^3)} \leq c\|u(0)\|_{L^2}$$

which implies the embedding $U^4_{\mathrm{KP}} \subset L^4(\mathbb{R}^3)$ and the inequalities (see Section 3.3)

$$\|u\|_{L^4(\mathbb{R}^3)} \leq c\|u\|_{U^4_{\mathrm{KP}}} \leq c\|u\|_{V^2_{\mathrm{KP}}}. \tag{6.35}$$

One has the bilinear improvement of Theorem 5.7 which implies

$$\|u_\lambda v_\mu\|_{L^2} \le c(\lambda/\mu)^{1/2}\|u_\lambda(0)\|_{L^2}\|v_\mu(0)\|_{L^2}, \tag{6.36}$$

thus,

$$\|u_\lambda v_\mu\|_{L^2} \le c(\mu/\lambda)^{1/2}\|u_\lambda\|_{U^2_{\mathrm{KP}}}\|v_\mu\|_{U^2_{\mathrm{KP}}}, \tag{6.37}$$

and together with the logarithmic interpolation of Lemma 4.12 one gets

$$\|u_\lambda v_\mu\|_{L^2} \le c(\mu/\lambda)^{1/2}(\ln(2+\lambda/\mu))^2\|u_\lambda\|_{V^2_{\mathrm{KP}}}\|v_\mu\|_{V^2_{\mathrm{KP}}}. \tag{6.38}$$

Formally the L^2 norm is constant.

We use the norm

$$\|u\|_X = \left(\sum_{\lambda\in 2^{\mathbb{Z}}}\|u_\lambda\|^2_{V^2_{\mathrm{KP}}}\right)^{1/2}.$$

Theorem 6.8. *There exists $\varepsilon > 0$ such that for $u_0 \in \dot H^{-1/2,0}(\mathbb{R}^2)$ there exists a unique solution $u \in X$ with*

$$\|u\|_X \le c\|u_0\|_{H^{-1/2,0}}(\mathbb{R}^2).$$

If $u_0 \in L^2$, then there is a unique solution in $C(\mathbb{R}; L^2)$ with

$$\|\chi_{[k,k+1]}(t)u\|_{U^2_{\mathrm{KP}}} < C(\|u_0\|_{L^2}).$$

Proof. By definition,

$$\|S(t)u_0\|_X \le c\|u_0\|_{\dot H^{-1/2}}.$$

We claim that

$$\|\int_0^t S(t-s)\partial_x(uv)ds\|_X \le c\|u\|_X\|v\|_X. \tag{6.39}$$

With this information we set up the fixed point argument and obtain a unique fixed point which is the solution. By duality, (6.39) follows from

$$\left|\int uvw\, dxdydt\right| \le c\|u\|_X\|v\|_X\|w\|_X. \tag{6.40}$$

We expand all factors and consider

$$\int u_{\lambda_1}v_{\lambda_2}w_{\lambda_3}dxdydt.$$

The integral is symmetric with respect to the factors and we may assume that $\lambda_1 \le \lambda_2 \le \lambda_3$. If there are no $\lambda_1 \le |\xi_1| \le 2\lambda_1$, $\lambda_2 \le |\xi_2| \le 2\lambda_2$, and $\lambda_3 \le |\xi_3| \le 2\lambda_3$ which add up to zero, then the integral vanishes. Thus

$$\lambda_3 \le 4\lambda_2$$

The integral vanishes unless there are such ξ_i, η_i, and τ_i which add up to zero. Now, if $0 = \xi_1 + \xi_2 + \xi_3 = \eta_1 + \eta_2 + \eta_3$, then

$$\xi_1^3 + \xi_2^3 + \xi_3^3 - \frac{\eta_1^2}{\xi_1} - \frac{\eta_2^2}{\xi_2} - \frac{\eta_3^2}{\xi_3} = 3\xi_1\xi_2\xi_3\Big(1 + \frac{|\eta_1\xi_2 - \eta_2\xi_1|^2}{\xi_1\xi_2\xi_3}\Big).$$

We define Q_H by the Fourier multiplier $\chi_{|\tau-\xi^3+\eta^2/\xi|>|\xi_1||\xi_2||\xi_1+xi_2|/10}$ and $Q_L = 1 - Q_H$. Then by the consideration of the supports

$$\int Q_L u_{\lambda_1} Q_L v_{\lambda_2} Q_L w_{\lambda_3} dxdydt = 0.$$

It follows from Lemma 4.35 that

$$\|Q_H u\|_{L^2} \le c(|\xi_1||\xi_2||\xi_1 + \xi_2|)^{-1/2} \|u\|_{V^2_{\mathrm{KP}}}$$

and

$$\|Q_H u\|_{V^2_{\mathrm{KP}}} \le c\|u\|_{V^2_{\mathrm{KP}}}.$$

We estimate

$$\begin{aligned}
&\left|\int (u_{\lambda_1}) v_{\lambda_2} Q_H w_{\lambda_3} dxdydt\right| \\
&\le \|u_{\lambda_1} v_{\lambda_2}\|_{L^2} \|Q_H w_{\lambda_3}\|_{L^2} \\
&\le c\left(\frac{\lambda_{\min}}{\lambda_{\max}}\right)^{1/2} (1 + \ln(\lambda_2/\lambda_1))^2 \lambda_{\max}^{-1} \lambda_{\min}^{-1/2} \|v_{\lambda_1}\|_{V^2_{\mathrm{KP}}} \|v_{\lambda_2}\|_{V^2_{\mathrm{KP}}} \|w_{\lambda_3}\|_{V^2_{\mathrm{KP}}} \\
&\le c\left(\frac{\lambda_{\min}}{\lambda_{\max}}\right)^{1/2} (1 + \ln(\lambda_2/\lambda_1)) \|v_{\lambda_1}\|_X \|v_{\lambda_2}\|_X \|w_{\lambda_3}\|_X.
\end{aligned}$$

This is easy to sum with respect to all indices. Th case with $Q_H u_{\lambda_1}$ is different since we don't gain a factor for the summation over the small frequencies. Here we need some orthogonality:

$$\begin{aligned}
\left|\sum_{\lambda_1<\lambda_2} \int Q_H u_{\lambda_1} u_{\lambda_2} w_{\lambda_3} dxdydt\right| &\le \Big(\sum_{\lambda_1<\lambda_2} \|Q_H u_{\lambda_1}\|_{L^2}^2\Big)^{1/2} \|v_{\lambda_2} w_{\lambda_3}\|_{L^2} \\
&\le \Big(\sum \lambda_1^{-1} \|u_{\lambda_1}\|_{V^2_{\mathrm{KP}}}\Big)^{\frac{1}{2}} \lambda_{\max}^{-1} \|v_{\lambda_2}\|_{V^2_{\mathrm{KP}}} \|w_{\lambda_3}\|_{V^2_{\mathrm{KP}}}
\end{aligned}$$

which can be summed.

Now consider data in $u_0 \in L^2$ with $\|u_0\|_{L^2} \le 1$. Let v be the solution to linear KP with initial data u_0. We seek a solution in the form $u = v + w$. We need in addition the following two estimates. The arguments are simpler than in the homogeneous case, and we leave them to the reader. Observe that the L^2 norm is formally conserved.

$$\|\chi_{[0,1]} \int_0^t S(t-s)\partial_x(uv)ds\|_{U^2} \le c\|u_{>1}\|_{U^2_{\mathrm{KP}}} \|v_{>1}\|_{U^2_{\mathrm{KP}}}$$

and

$$\|\chi_{[0,1]} \int_0^t S(t-s)\partial_x(u_{<1}v_{>1})ds\|_{U^2} \leq c\|u\|_{U^2_{\mathrm{KP}}}\|v\|_{U^2_{\mathrm{KP}}}. \qquad \square$$

Chapter 7

Appendix A: Young's inequality and interpolation

Young's inequality bounds convolutions in Lebesgue spaces. It is part of the statement that the integral exists for almost all arguments of the convolution. Let m^d denote the d-dimensional Lebesgue measure.

Lemma 7.1. *Let* $1 \le p, q, r \le \infty$ *satisfy*

$$\frac{1}{p} + \frac{1}{q} + \frac{1}{r} = 2,$$

and let

$$f \in L^p(\mathbb{R}^d), \quad g \in L^q(\mathbb{R}^n), \quad h \in L^r(\mathbb{R}^d).$$

Then

$$\int_{\mathbb{R}^d \times \mathbb{R}^d} f(x)g(x-y)h(y)dm^{2d}(x,y) \le \|f\|_{L^p}\|g\|_{L^q}\|h\|_{L^r}.$$

We assume that the Lemma holds and choose $f(x) = e^{-|x|^2} \in L^r(\mathbb{R}^d)$. It follows by Fubini's theorem that $g(x-y)h(y)$ is integrable with respect to y for almost all x. The estimate of the lemma shows that

$$L^p(\mathbb{R}^d) \ni f \mapsto \int_{\mathbb{R}^d} \left(\int_{\mathbb{R}^d} h(y)g(x-y)dm^d(y) \right) f(x)dm^d(x) \in \mathbb{R}$$

defines a linear form on L^r of norm $\le \|g\|_{L^q}\|g\|_{L^r}$. Thus

$$\|g * h\|_{L^{p'}} \le \|g\|_{L^q}\|h\|_{L^r}$$

for

$$\frac{1}{q} + \frac{1}{r} = 1 + \frac{1}{p'}.$$

H. Koch et al., *Dispersive Equations and Nonlinear Waves: Generalized Korteweg–de Vries, Nonlinear Schrödinger, Wave and Schrödinger Maps*, Oberwolfach Seminars 45,
DOI 10.1007/978-3-0348-0736-4_7, © Springer Basel 2014

Proof of Lemma 7.1*, as in [22].* Set

$$\frac{1}{\gamma_1} = 1 - \frac{1}{p}, \quad \frac{1}{\gamma_2} = 1 - \frac{1}{q}, \quad \frac{1}{\gamma_3} = 1 - \frac{1}{r}.$$

Then $1 \leq \gamma_1 \leq \infty$,

$$\frac{1}{\gamma_2} + \frac{1}{\gamma_3} = \frac{1}{p}, \quad \frac{1}{\gamma_1} + \frac{1}{\gamma_3} = \frac{1}{q}, \quad \frac{1}{\gamma_1} + \frac{1}{\gamma_2} = \frac{1}{r}$$

and

$$\frac{1}{\gamma_1} + \frac{1}{\gamma_2} + \frac{1}{\gamma_3} = 1.$$

Let

$$a(x,y) = |f(x)|^{p/\gamma_3} |g(x-y)|^{q/\gamma_3}, \quad b(x,y) = |g(x-y)|^{q/\gamma_1} |h(y)|^{r/\gamma_1},$$
$$c(x,y) = |f(x)|^{p/\gamma_2} |h(y)|^{r/\gamma_2}.$$

Then

$$|f(x)g(x-y)h(y)| = a(x,y)b(x,y)c(x,y)$$

and, by applying Hölder's inequality twice,

$$\int |f(x)g(x-y)h(y)| dm^{2d} \leq \|a\|_{L^{\gamma_3}} \|b\|_{L^{\gamma_1}} \|c\|_{L^{\gamma_2}} = \|f\|_{L^p} \|g\|_{L^q} \|h\|_{L^r}. \qquad \square$$

There is an improvement: the weak Young inequality. Let (X, μ) be a measure space. We will often omit space and measure in the notation. The weak L^p spaces are defined by the quasi-norm

$$\|f\|_{L^p_w} = \sup_{t>0} t \left(\mu(\{x : |f(x)| > t\})\right)^{1/p}.$$

If $1 < p < \infty$, then there is an equivalent norm on L^p_w:

$$\|f\|_{L^p_w} \sim \sup_{t>0} t \left(\int_{\{x:|f(x)|>t\}} |f(y)| d\mu(y) \right)^{1/p}.$$

It is not hard to see the equivalence, and that the term on the right-hand side defines a norm.

Proposition 7.2. *Suppose that*

$$1 < p, q, r < \infty, \quad \frac{1}{p} + \frac{1}{q} = 1 + \frac{1}{r},$$

$f \in L^p$ *and* $g \in L^q_w$*. Then* $f(x)g(x-y)$ *is integrable with respect to* x *for almost all* y *and*

$$\|f * g\|_{L^r} \leq c_{p,q} \|f\|_{L^p} \|g\|_{L^q_w}.$$

This is a consequence of the Marcinkiewicz interpolation theorem. We state and prove the following version.

Let X and Y be normed linear spaces. We denote by $L(X, Y)$ the normed space of bounded linear operators from X to Y.

Lemma 7.3 (Marcinkiewicz interpolation). *Let (X, μ) and (Y, ν) be measure spaces and $1 \leq p_1 < p_2 \leq \infty$, $1 \leq q_1, q_2 \leq \infty$, $q_1 \neq q_2$, $0 < \lambda < 1$,*

$$\frac{1}{p} = \frac{\lambda}{p_1} + \frac{1-\lambda}{p_2}, \qquad \frac{1}{q} = \frac{\lambda}{q_1} + \frac{1-\lambda}{q_2}.$$

Suppose that

$$T \in L(L^{p_1}(\mu), L^{q_1}_w(\nu)) \cap L(L^{p_2}(\mu), L^{q_2}_w(\nu)).$$

Then $T \in L(L^p_w(\mu), L^q_w(\nu))$, and

$$\|T\|_{L(L^p_w(\mu), L^q_w(\nu))} \leq c\|T\|^{\lambda}_{L(L^{p_1}(\mu), L^{q_1}_w(\nu))} \|T\|^{1-\lambda}_{L(L^{p_2}(\mu), L^{q_2}_w(\nu))}$$

and, if $p \leq q$, then $T \in L(L^p(\mu), L^q(\nu))$ and

$$\|T\|_{L(L^p(\mu), L^q(\nu))} \leq c\|T\|^{\lambda}_{L(L^{p_1}(\mu), L^{q_1}_w(\nu))} \|T\|^{1-\lambda}_{L(L^{p_2}(\mu), L^{q_2}_w(\nu))}$$

with a constant c depending only on the exponents.

Proof of Proposition 7.2. Let $f \in L^p$ and $Tg : L^q \to L^r$ be the convolution with g. We interpolate the estimate with $p_1 = 1$, $p_2 = p'$, $q_1 = q$, and $q_2 = \infty$ to get the estimate in weak spaces:

$$\|f * g\|_{L^r_w} \leq \|g\|_{L^q_w} \|f\|_{L^p}.$$

Now we fix g and consider $T : f \mapsto f * g$, and get

$$\|f * g\|_{L^r} \leq c\|f\|_{L^p} \|g\|_{L^q_w}$$

by the second part of the Lemma. □

It is useful to generalize and sharpen the Marcinkiewicz interpolation estimates before proving them.

Definition 7.4 (Lorentz spaces). *Let (A, μ) be a measure space and $1 \leq p, q \leq \infty$. We define*

$$\|f\|_{L^{p,q}(\mu)} = \left(q \int_0^\infty \left(\mu(\{x : |f(x)| > t\})^{1/p} t \right)^q \frac{dt}{t} \right)^{1/q},$$

with the obvious modification for $q = \infty$. We denote by $L^{pq}(\mu)$ the set of all measurable functions f for which $\|f\|_{L^{pq}(\mu)} < \infty$.

Properties:

1. Since

$$\{x : |f(x) + g(x)| > t\} \subset \{x : |f(x)| > t/2\} \cup \{x : |g(x)| > t/2\},$$

it follows that

$$\mu(\{x : |f(x) + g(x)| > t\}) \leq \mu(\{x : |f(x)| > t/2\}) + \mu(\{x : |g(x)| > t/2\}),$$

and hence

$$\|f + g\|_{L^{pq}} \leq c\left(\|f\|_{L^{pq}} + \|g\|_{L^{pq}}\right).$$

2. For $q_1 \leq q_2$

$$\|f\|_{L^{pq_2}} \leq c\|f\|_{L^{pq_1}}.$$

We begin the proof with

$$\mu(\{|f| \geq t\})t^q = q\int_0^t \mu(\{|f| \geq t\})s^{q-1}ds \leq q\int_0^t \mu(\{|f| \geq s\})s^{q-1}ds \leq \|f\|_{L^{pq}}^q.$$

Now, if $q_1 < q_2$,

$$q_2 \int_0^\infty [\mu(\{|f| \geq t\})^{1/p}t]^{q_2}\frac{dt}{t} \leq \frac{q_2}{q_1}\|f\|_{L^{p,\infty}}^{q_2-q_1}\|f\|_{L^{p,q_1}}^{q_1} \leq \frac{q_2}{q_1}\|f\|_{L^{p,q_1}}^{q_2}.$$

3. If $1 < p < \infty$ and $\frac{1}{p} + \frac{1}{p'} = \frac{1}{q} + \frac{1}{q'} = 1$, then there exists $c > 0$ such that

$$\left|\int fgd\mu\right| \leq c\|f\|_{L^{pq}}\|g\|_{L^{p'q'}}.$$

For the proof we define $f^* : (0,\infty) \to \mathbb{R}^+$ to be the unique function with

$$m^1(\{\tau : f^*(\tau) > t\}) = \mu(\{x : f(x) > t\})$$

for all $t > 0$. Then, using Fubini several times (with the Lebesgue measure $\mu = m^d$ for definiteness, but the argument holds for general measures)

$$\begin{aligned}
\int |fg|dm^d &= m^{d+2}(\{(x,s,t) \in \mathbb{R}^d \times \mathbb{R} \times \mathbb{R} : 0 < s < |f(x)|, 0 < t < |g(x)|\}) \\
&= \int_{\mathbb{R}^+\times\mathbb{R}^+} m^d(\{x : |f(x)| > s\} \cap \{x : |g(x)| > t\})dsdt \\
&\leq \int_{\mathbb{R}^+\times\mathbb{R}^+} \min\{m^d(\{x : |f(x)| > s\}), m^d(\{x : |g(x)| > t\})\}dsdt \\
&= \int_{\mathbb{R}^+\times\mathbb{R}^+} m^1(\{\sigma : |f^*(\sigma)| > s\} \cap \{\tau : |g^*(\tau)| > t\})dsdt \\
&= \int_0^\infty f^*(\tau)g^*(\tau)d\tau.
\end{aligned}$$

Therefore,

$$\begin{aligned}\int fg d\mu &\le \int_0^\infty f^*(t) g^*(t) dt \\ &= \int_0^\infty (t^{1/p} f^*)(t^{1/p'} g^*(t)) dt/t \\ &\le \left(\int_0^\infty t^{(q/p)-1} (f^*)^q dt \right)^{1/q} \left(\int_0^\infty t^{(q'/p')-1} (g^*)^{q'} dt \right)^{1/q'}.\end{aligned}$$

The last inequality is an application of Hölder's inequality. The proof of the third part is completed by the equality

$$\frac{q}{p} \int_0^\infty t^{(q/p)-1} (f^*(t))^q dt = q \int_0^\infty (\mu(|f(x)| > s))^{q/p} s^{q-1} ds \tag{7.1}$$

in one-dimensional calculus. We observe that

$$s \mapsto m^1(\{\tau : f^*(\tau) > s\})$$

is the inverse of f^*. Both functions are monotonically decreasing.

Let f and f^{-1} be mutually inverse non-negative monotonically decreasing functions, and g and h non-negative monotonically increasing functions with antiderivatives G and H with

$$H(t) G \circ f(t) \to 0$$

as $t \to \infty$ and $t \to 0$. Then, by an integration by parts and one substitution

$$\int_0^\infty h G \circ f dt == - \int_0^\infty H g \circ f f' dt = \int_0^\infty H \circ f^{-1}(s) g(s) ds.$$

This specializes to (7.1). Moreover, checking the inequalities shows that

$$\|f\|_{L^{pq}} \le c \sup\{ \int fg d\mu : \|g\|_{L^{p'q'}} \le 1 \}.$$

4. This pairing defines a duality isomorphism if $1 < p < \infty$ and $1 \le q < \infty$:

$$L^{p'q'} \ni g \mapsto (f \mapsto \int fg d\mu) \in (L^{pq})^*.$$

In particular all spaces L^{pq} with $1 < p$ are Banach spaces. To prove this we choose B to be a measurable set of positive finite measure. There exists $\tilde p > p$ so that $L^{\tilde p}(B) \subset L^{pq}$. If l is a bounded linear functional on L^{pq}, then it defines a bounded linear functional on $L^{\tilde p}$, which in turn is represented by a function $g \in L^{\tilde{p}'}(\mu)$. The previous step gives a bound for $\|g\chi_B\|_{L^{p'q'}}$ in terms of l.

We order the measurable subsets of A by inclusion up to sets of measure zero. This defines a partial order on the subsets on which the duality statement holds. Every chain has an upper bound, the union of the chain. By the Zorn lemma there is a maximal element. The procedure above allows to show that the maximal set is necessarily the full space.

In particular, duality allows one to define an equivalent norm on $L^{pq}(\mu)$ for $1 \le p \le \infty$ and $1 \le q \le \infty$. Completeness of dual spaces is obvious. Completeness of $L^{p1}(\mu)$ is left as an exercise.

Lemma 7.5. *Suppose that* $1 \le p_1, p_2, q_1, q_2 \le \infty$,

$$T \in L(L^{p_1 1}(\mu), L^{q_1 \infty}(\nu)) \cap L(L^{p_2 1}(\mu), L^{q_2 \infty}(\nu)),$$

$p_1 \ne p_2$, $q_1 \ne q_2$, $0 < \lambda < 1$, *and*

$$\frac{1}{p} = \frac{1-\lambda}{p_1} + \frac{\lambda}{p_2}, \quad \frac{1}{q} = \frac{1-\lambda}{q_1} + \frac{\lambda}{q_2},$$

and $1 \le r \le \infty$.

Then the operator T can be continuously extended to $T \in L(L^{pr}(\mu), L^{qr}(\nu))$. Moreover,

$$\|T\|_{L(L^{pr}(\mu),L^{qr}(\nu))} \le c\|T\|^{\lambda}_{L(L^{p_1}(\mu),L^{q_1}_w(\nu))}\|T\|^{1-\lambda}_{L(L^{p_2}(\mu),L^{q_2}_w(\nu))}.$$

Proof. An easy calculation shows that

$$\frac{1-\frac{p}{p_2}}{1-\frac{p}{p_1}} = \frac{\lambda - 1}{\lambda}. \tag{7.2}$$

This will be useful later on. Let $t > 0$ and

$$f_t(x) = \begin{cases} f(x), & \text{if } |f(x)| \le t, \\ tf(x)/|f(x)|, & \text{if } |f(x)| > t, \end{cases}$$

and let $f^t = f - f_t$. Then

$$f = f_t + f^t$$

and, if $p_1 < p < p_2$, which we assume in the sequel, then

$$\|f^t\|_{L^{p_1}} \le (p - p_1)^{1/p_1} t^{1-\frac{p}{p_1}} \|f\|^{\frac{p}{p_1}}_{L^p_w}$$

and

$$\|f_t\|_{L^{p_2}} \le (p_2 - p)^{1/p_2} t^{1-\frac{p}{p_2}} \|f\|^{\frac{p}{p_2}}_{L^p_w},$$

with obvious modifications if $p_2 = \infty$.

Moreover, by the triangle inequality,

$$\{|Tf| > t\} \subset \{Tf^s > t/2\} \cup \{Tf_s > t/2\}.$$

Let

$$a_1 = \|T\|_{L(L^{p_1}, L^{q_1}_w)} \quad a_2 = \|T\|_{L(L^{p_2}, L^{q_2}_w)},$$

$$\tau = \frac{q_2 - q_1}{q_2(1 - \frac{p}{p_2}) - q_1(1 - \frac{p}{p_1})}$$

and

$$s = t^\tau a_1^{\frac{(1-\lambda)\frac{q}{q_1} - 1}{1 - \frac{p}{p_1}}} a_2^{\frac{\lambda \frac{q}{q_2} - 1}{1 - \frac{p}{p_2}}}.$$

Step 1. The bound in weak L^p space. We want to prove

$$\lambda \nu(\{|Tf(x)| > t\})^{1/q} \le c a_1^{1-\lambda} a_2^\lambda$$

for $\|f\|_{L^p_w} = 1$, with c depending only on the exponents. We estimate

$$\begin{aligned}
\lambda^q \mu(\{|Tf| > t\}) &\le c \left(t^{q-q_1} \|Tf^s\|^{q_1}_{L^{q_1}_w} + t^{q-q_2} \|Tf_s\|^{q_2}_{L^{q_2}_w} \right) \\
&\le c \left(t^{q-q_1} a_1^{q_1} \|f^s\|^{q_1}_{L^{p_1}} + t^{q-q_2} a_2^{q_2} \|f_s\|^{q_2}_{L^{p_2}} \right) \\
&= c \left(t^{q-q_1} s^{q_1 - q_1 p/p_1} \|f\|^{pq_1/p_1}_{L^p_w} + t^{q-q_2} s^{q_2 - q_2 p/p_2} \|f\|^{pq_2/p_2}_{L^p_w} \right) \\
&= c \left(t^{q - q_1 - \frac{q_1(q_2 - q_1)}{q_2 \frac{1-\lambda}{\lambda} + q_1}} + t^{q - q_2 - \frac{q_2(q_1 - q_2)}{q_1 \frac{\lambda}{1-\lambda} + q_2}} \right) a_1^{q(1-\lambda)} a_2^{q\lambda} \\
&= c \left(t^{q_1[\frac{q}{q_1} - 1 - (\frac{q}{q_1} - \frac{q}{q_2})\lambda]} + t^{q_2[\frac{q}{q_2} - 1 - (\frac{q}{q_2} - \frac{q}{q_1})(1-\lambda)]} \right) a_1^{q(1-\lambda)} \\
&\quad \times a_2^{q\lambda} \\
&= c a_1^{q(1-\lambda)} a_2^{q\lambda}.
\end{aligned}$$

This completes the proof of the weak type estimate.

Step 2: The endpoints $L(L^{p_1}, L^{q_1})$ and $L(L^{p_\infty}, L^{q_\infty})$. We assume that $1 < p_1, p_2, q_1, q_2 < \infty$, which can be achieved by the first step.

By duality, with constant changing from line to line

$$\begin{aligned}
\|Tf\|_{L^{qr}} &\le c \sup\{\int (Tf) g d\nu : \|g\|_{L^{q'r'}} \le 1\} \\
&= c \sup\{\int f T^* g d\nu : \|g\|_{L^{q'r'}} \le 1\} \\
&= c \|f\|_{L^{pq}} \|T^*\|_{L(L^{q',r'}(\nu), L^{p',q'}(\mu))}
\end{aligned}$$

and hence, for $1 < p < \infty$,

$$\|T\|_{L(L^{pr}, L^{qr})} \le c \|T^*\|_{L(L^{q'r'}, L^{p'r'})}.$$

We apply this with $L^{p_1 1} \to L^{q_1 \infty}$ to see that

$$\|T^*\|_{L(L^{q_i' 1}, L^{p_i' \infty})} \le c\|T\|_{L(L^{p_i 1}, L^{q_i \infty})}$$

for $i = 1, 2$. From Step 1

$$\|T^*\|_{L(L^{q' \infty}, L^{p' \infty})}$$

satisfies the desired bounds. Duality again gives the statement for $r = 1$.

Step 3: Interpolation in L^p. Suppose that $T \in L(L^1(\mu), L^1(\nu)) \cap L(L^\infty(\mu), L^\infty(\mu))$ with norm $\le \frac{1}{2}$. Then

$$\|Tf\|_{L^p(\nu)} \le \left(\frac{p}{p-1}\right)^{1/p} \|f\|_{L^p(\mu)}.$$

We begin the proof by observing that

$$\{|Tf| > t\} \subset \{Tf_t > t/2\} \cup \{Tf^t > t/2\}.$$

The first set is empty by the assumption on the norm of T. Hence

$$\begin{aligned}
p \int \nu(\{|Tf| > t\}) t^{p-1} dt &\le p \int \nu(\{Tf^t > t/2) t^{p-1} dt \\
&\le p \int_0^\infty \|f^t\|_{L^1} t^{p-2} dt \\
&= p \int_0^\infty \int_t^\infty \mu(\{|f| \ge s\}) ds t^{p-2} dt \\
&= p \int_0^\infty \int_0^s t^{p-2} dt \mu(\{|f| \ge s\}) ds \\
&= \frac{p}{p-1} \|f\|_{L^p}^p.
\end{aligned}$$

Step 4: Conclusion. We have proven the bounds for $\|T\|_{L(L^{p,\infty}, L^{q,\infty})}$ and for $\|T\|_{L(L^{p,1}, L^{q,1})}$. We will argue similarly to the previous step. We decompose

$$f_j(x) = \begin{cases} f(x) & \text{if } 2^j \le |f_j(x)| < 2^{j+1} \\ 0 & \text{otherwise} \end{cases}$$

and define A_j as the set where f_j is not zero. Then f is the sum over the f_j. Let $t > 0$,

$$B_t = \{j \in \mathbb{Z} : (\mu(A_j))^{1/p} 2^j < t\}.$$

We define

$$A(t) = \bigcup_{j \in B_t} A_j,$$

$$f_t(x) = \begin{cases} f(x) & \text{if } x \in A(t) \\ 0 & \text{otherwise} \end{cases}$$

and $f^t = f - f_t$. Then,

$$\{x : |f_t| > s\} \subset \bigcup_{j \in B_t, 2^{j+1} \geq s} A_j$$

and hence

$$\|f_t\|_{L^{p,\infty}} \leq t \sup_{s>0} s \left(\sum_{j \geq \ln_2 s} 2^{-jp} \right)^{1/p} \leq 2t.$$

It suffices to consider T with $\|T\|_{L(L^{p,\infty}, L^{p,\infty})} < \frac{1}{4}$. With

$$g(t) := m^1(\{s : \nu(\{|Tf(y)| > s\})^{1/q} s > t\})$$

we obtain as in Step 3

$$\begin{aligned} g(t) &\leq m^1(\{s : \nu(\{|Tf^{t/2}(y)| > s\})^{1/q} s > t\}) \\ &\leq ct^{-1} \|Tf^{t/2}\|_{L^{q,1}} \leq ct^{-1} \|f^{t/2}\|_{L^{p,1}}. \end{aligned}$$

We define $h(t) = \#(\mathbb{Z} \backslash A_t)$. Then

$$\|f^{t/2}\|_{L^{p,1}} \leq \int_{t/2}^{\infty} h(s) ds$$

and we conclude for $1 < r < \infty$ as in Step 3. $\square$

7.1 Complex interpolation: The Riesz–Thorin theorem

The Riesz–Thorin interpolation theorem states the following. For notational simplicity we omit the measures in the notation.

Theorem 7.6. *Let $1 \leq p_1, p_2, q_1, q_2 \leq \infty$. Let T_λ, $0 \leq \operatorname{Re} \lambda \leq 1$, be an operator from $L^1 \cap L^\infty \to L^1 + L^\infty$. Suppose that*

$$\lambda \mapsto \int T_\lambda f g$$

is continuous in $0 \leq \operatorname{Re} \lambda \leq 1$ and holomorphic inside the strip, for all $f \in L^1 \cap L^\infty$ and $g \in L^1 \cap L^\infty$. Suppose that

$$\sup_{\operatorname{Re} \lambda = 0} \|T_\lambda\|_{L(L^{p_0}, L^{q_0})} = C_0 < \infty$$

and

$$\sup_{\operatorname{Re} \lambda = 1} \|T_\lambda\|_{L(L^{p_1}, L^{q_1})} = C_1 < \infty.$$

Then

$$\|T_\lambda\|_{L(L^p, L^q)} \leq C_0^{1 - \operatorname{Re} \lambda} C_1^{\operatorname{Re} \lambda}$$

if

$$\frac{1-\operatorname{Re}\lambda}{p_0}+\frac{\operatorname{Re}\lambda}{p_1}=\frac{1}{p},\qquad \frac{1-\operatorname{Re}\lambda}{q_0}+\frac{\operatorname{Re}\lambda}{q_1}=\frac{1}{q}$$

.

The proof relies on the three lines theorem in complex analysis:

Lemma 7.7 (Three lines theorem)**.** *Suppose that v is a bounded holomorphic function on the strip $C=\{z=x+iy:0<x<1\}$ and that it is continuous on the closure. Then*

$$|v(x)|\le(\sup_y|v(iy)|)^{1-x}(\sup_y|v(1+iy)|)^x.$$

Proof. By the maximum principle for harmonic functions, any harmonic function on a bounded open set which is continuous on the closure, assumes the maximum of the modulus at the boundary. This is true for

$$u_\varepsilon(x,y)=e^{\varepsilon(x+iy)^2}u(x,y)$$

on $C\cap\overline{B_R(0)}$ for every R. This function tends to 0 as $y\to\infty$ hence

$$|u_\varepsilon(x+iy)|\le\max\{(\sup_y|u(iy)|)^{1-x},(\sup_y|u(1+iy)|)^x\}$$

and letting $\varepsilon\to0$ gives the result. □

Proof of Theorem 7.6. Let $f\in L^1(\mu)\cap L^\infty(\mu)$ and $g\in L^1(\nu)\cap L^\infty(\nu)$. Then, by assumption,

$$v(\lambda)=\int T_\lambda fgd\nu$$

is a bounded analytic function. By the three lines theorem 7.7 we have

$$|v(\lambda)|\le\sup_t\max\{|v(it)|,|v(1+it)|\}.$$

Now

$$\left|\int T_{it}fgd\nu\right|\le\|T_{it}f\|_{L^{q_0}}\|g\|_{L^{q_0'}}\le C_0\|f\|_{L^{p_0}}\|g\|_{L^{q_0'}}$$

and

$$\left|\int T_{1+it}fgd\nu\right|\le\|T_{1+it}f\|_{L^{q_1}}\|g\|_{L^{q_1'}}\le C_0\|f\|_{L^{p_1}}\|g\|_{L^{q_1'}}.$$

Thus,

$$\left|\int(T_\lambda f)gd\mu\right|\le\max\{C_0,C_1\}\left(\|f\|_{L^{p_0}}\|g\|_{L^{q_0'}}+\|f\|_{L^{p_1}}\|g\|_{L^{q_1'}}\right)$$

and we could derive that

$$\|T\|_{L(L^{p_0}\cap L^{p_1},L^{q_0}+L^{q_1})}\le\max\{C_0,C_1\},$$

but we will avoid this step. Let $f \in L^p$ and $g \in L^{q'}$. We want to prove that

$$\left| \int g T_\lambda f \right| \le \|f\|_{L^p} \|g\|_{L^{q'}} \sup_y \|T_{iy}\|^{1-\lambda}_{L(L^{p_1},L^{q_1})} \sup_y \|T_{1+iy}\|^{\lambda}_{L(L^{p_2},L^{q_2})} \tag{7.3}$$

for $f \in L^p$ and $g \in L^{q'}$. The theorem follows then by an duality argument. Moreover, it suffices to consider a dense set of functions, which are measurable, bounded, and for which there is $\varepsilon > 0$ such that either the functions vanish at a point, or else they are at least of size ε. Also, we may restrict to f and g with $\|f\|_{L^p} = \|g\|_{L^{q'}} = 1$.

Let

$$f_z(x) = \frac{f(x)}{|f(x)|} |f(x)|^{(1-z)\frac{p}{p_0} + z\frac{p}{p_1}},$$

$$g_z(x) = \frac{g(x)}{|g(x)|} |g(x)|^{(1-z)\frac{q'}{q'_0} + z\frac{q'}{q'_1}}$$

and

$$v(z) = \int g_z(y) T_z f_z(y) d\nu(y).$$

This is a bounded holomorphic map from the strip to $L^1 \cap L^\infty$ with values in $\mathbb{C}$. We claim that it is continuous on the closure of the strip. Let λ be an arbitrary point of the closure. We write

$$v(z) - v(\lambda) = \int g_\lambda (T_z - T_\lambda) f_\lambda d\nu + \int [(g_z - g_\lambda) T_z f_\lambda + g_z T_z (f_z - f_\lambda)] d\nu.$$

The first term tends to zero as $z \to \lambda$ by assumption. Then

$$g_z - g_\lambda \to 0 \quad \text{and } g_z - f_\lambda \to 0 \quad \text{as } z \to \lambda$$

in $L^1 \cap L^\infty$. Continuity follows by the uniform bound above.

We turn to complex differentiability at an arbitrary point λ in the interior. Indeed

$$\frac{v(z) - v(\lambda)}{z - \lambda} = \frac{\int g_\lambda (T_z - T_\lambda) f_\lambda d\nu}{z - \lambda} + \int \frac{g_z - g_\lambda}{z - \lambda} T_z f_\lambda d\nu + \int g_z T_z \frac{f_z - f_\lambda}{z - \lambda} d\nu.$$

The first term converges to a complex number by assumption. Moreover,

$$\frac{g_z - g_\lambda}{z - \lambda}$$

converges to a function g'_λ in $L^1 \cap L^\infty$ as $z \to \lambda$. Let $\tilde{g}$ denote the difference between the difference quotient and g'_λ. Then

$$\int \frac{g_z - g_\lambda}{z - \lambda} T_z f_\lambda d\nu = \int g' T_\lambda f_\lambda d\nu + \int \tilde{g} T_z f_\lambda d\nu + \int g' (T_z - T_\lambda) f_\lambda d\nu.$$

The second term tends to zero since $\tilde{g}$ tends to zero in $L^1 \cap L^\infty$ and the third one by the continuity assumption as $z \to \lambda$. Similarly we deal with the remaining term.

We turn to the behavior at the boundary:

$$|v(it)| = \int T_{it} f_{it} g_{it} d\nu \leq \|T_{it}\|_{L(L^{p_0}, L^{p_1})} \|f_{it}\|_{L^{p_0}} \|g_{it}\|_{L^{q_0}}$$

and

$$\|f\|_{L^{p_0}} = \|f\|_{L^p}^{p_0/p} = 1 = \|g_{it}\|_{L^{q_0'}} = \|g\|_{L^{q'}}^{q_0'/q'}.$$

We apply the three lines theorem 7.7. This yields

$$|v(z)| \leq \sup_y \|T_{iy}\|_{L(L^{p_1}, L^{q_1})}^{1-x} \sup_y \|T_{1+iy}\|_{L(L^{p_2}; L^{q_2})}^x.$$

Evaluated at $z = \lambda$, this gives inequality (7.3). □

Chapter 8

Appendix B: Bessel functions

The Bessel functions are confluent hypergeometric functions. They are solutions to confluent hypergeometric differential equations. Here is a very brief introduction. Consider a complex differential equation

$$x^{(n)} = \sum_{j=0}^{n-1} a_j(z) z^{(j)}$$

with initial data

$$x^{(j)}(z_0) = y_j$$

for $j = 0, 1, \ldots, n-1$ and given complex numbers z_0 and y_j. If the coefficients are holomorphic in a neighborhood of z_0, then there is a unique solution which is holomorphic in z and the y_j.

Consider the scalar equation

$$\dot{x} = \frac{\lambda}{z - z_0} x.$$

The space of solutions is at most one-dimensional. Formally, a solution is given by $x = (z - z_0)^\lambda$, which, unless z is an integer, is only defined in a set of the type $\mathbb{C} \backslash (-\infty, z_0]$, called slit domain. Similarly, if

$$\dot{x} = \Big(\frac{\lambda}{z - z_0} + \phi(z) \Big) x,$$

with a holomorphic function ϕ near z_0, then there is a unique solution of the type

$$(z - z_0)^\lambda \left[1 + \sum_{k=1}^{\infty} a_k (z - z_0)^k \right],$$

again defined in the slit domain as above unless λ is an integer. The number λ is called characteristic number. It is not hard to see that there is a unique such

H. Koch et al., *Dispersive Equations and Nonlinear Waves: Generalized Korteweg–de Vries, Nonlinear Schrödinger, Wave and Schrödinger Maps*, Oberwolfach Seminars 45,
DOI 10.1007/978-3-0348-0736-4_8, © Springer Basel 2014

solution, and the power series can be defined iteratively. The point z_0 is called a regular singular point. A point is called irregular singular point if the Laurent series of the coefficients contains terms of order below $(z-z_0)^{-1}$.

We call ∞ regular point, resp. regular singular, resp. irregular singular point for

$$\dot{x} = a(z)x$$

if, when we express z in terms of z^{-1}, 0 is a regular, resp. regular singular, resp. irregular singular point of

$$\dot{x} = -z^{-2}a(z^{-1})x.$$

We use the same notation for systems of equations. The eigenvalues of A in

$$\dot{x} = \frac{1}{z-z_0}A(z)x + f(z)x$$

are called characteristic values. They play a very similar role as for scalar equations. Multiple characteristic values and/or resonances (a resonance refers to the situation when eigenvalues of A are linearly dependent over the integers) may lead to logarithmic terms.

We are interested in second-order scalar equations

$$a(z)\ddot{x} + b(z)\dot{x} + c(z)x = 0$$

with meromorphic functions a, b, and c. We may rewrite them as a 2×2 system, which we use to define the notion of a regular, regular singular, and irregular singular point. The point z_0 is regular if $b(z)/a(z)$ and $c(z)/a(z)$ have a holomorphic extension near z_0. It is a regular singular point if the Laurent expansion of $b(z)/a(z)$ begins with c_0z^{-1} and the one of $c(z)/a(z)$ begins with $c_1z^{-2}+c_2z^{-1}$. The characteristic numbers can be calculated in terms of the Laurent series. If they are independent over the integers then there are unique solutions of the type

$$z^{\lambda}\sum a_j z^j,$$

where λ is one of the characteristic numbers.

Of particular importance is the case when there are only regular singular points. In that case there are exactly three of them, and applying a Moebius transform we may choose them to be 0, 1, and ∞. Moreover, multiplying by $z^{\lambda}(1-z)^{\mu}$ we can ensure that one of the characteristic values at 0 and at 1 is 0. Then we are in the case of hypergeometric differential equations

$$z(1-z)\frac{d^2}{dz^2}w + [c-(a+b+1)z]\frac{dw}{dz} - abw = 0.$$

The characteristic numbers at $z=0$ are 0 and $1-c$, the ones at $z=1$ are 0 and $c-a-b$, and the ones at infinity are $-a$ and $-b$.

The regular solution near 0 with value 1 at zero is the hypergeometric function

$$ {}_2F_1(a, b; c; z). $$

The Bessel differential equation is

$$ z^2\ddot{w} + w\dot{w} + (z^2 - \nu^2)w = 0. $$

It has a regular singularity at $z = 0$ with indices $\pm\nu$, and an irregular singularity at $z = \infty$. The Bessel function of the first kind is

$$ J_\nu = \left(\frac{1}{2}z\right)^\nu \sum_{k=0}^{\infty} (-1)^k \frac{\left(\frac{1}{4}z^2\right)^k}{k!\Gamma(\nu + k + 1)}. $$

We have, unless ν is a negative integer,

$$ J_\nu(z) - \left(\frac{1}{2}z\right)^\nu / \Gamma(\nu + 1) = O(|z|^{\operatorname{Re}\nu + 1}) \text{ near } 0, $$

$$ J_\nu(z) - \sqrt{\frac{2}{\pi z}} \cos\left(z - \frac{1}{2}\nu\pi - \frac{1}{4}\pi\right) + e^{|\operatorname{Im} z|} o(1) $$

for $z \to \infty$ and $\nu \in \mathbb{R}$.

There are integral representation for $\nu > -\frac{1}{2}$:

$$ \begin{aligned} J_\nu(z) &= \frac{2(\frac{1}{2}z)^\nu}{\pi^{1/2}\Gamma(\nu + \frac{1}{2})} \int_0^1 (1 - t^2)^{\nu - \frac{1}{2}} \cos(zt) dt \\ &= \frac{(\frac{1}{2}z)^\nu}{\pi^{1/2}\Gamma(\nu + \frac{1}{2})} \int_0^\pi \cos(z\cos(\theta)) \sin(\theta)^{2\nu} dt \end{aligned} $$

and if the absolute value of the argument of z is bounded by $\frac{1}{2}\pi$, the Schläfli–Sommerfeld formulas hold:

$$ J_\nu(z) = \frac{1}{2\pi i} \int_{-\infty - \pi i}^{\infty + \pi i} e^{z \sinh t - \nu t} dt, $$

$$ \begin{aligned} J_\nu(z) &= \frac{2(\frac{1}{2}z)^\nu}{\pi^{1/2}\Gamma(\nu + \frac{1}{2})} \int_0^1 (1 - t^2)^{\nu - \frac{1}{2}} \cos(zt) dt \\ &= \frac{(\frac{1}{2}z)^\nu}{2\pi i} \int_{-\infty}^{0+} \exp(t - \frac{z^2}{4t}) t^{-\nu + 1} dt. \end{aligned} $$

The Bessel functions satisfy

$$ \left(\frac{d}{xdx}\right)^m (x^\nu J_\nu) = x^{\nu - m} J_{\nu - m}. $$

See [26] for more information. We want to evaluate (with the Hausdorff measure of dimension s denoted by $\mathcal{H}^s$)

$$\begin{aligned} H(\xi) = \int_{\mathbb{S}^{d-1}} e^{ix\xi} d\mathcal{H}^{d-1} &= \int_0^\pi \mathcal{H}^{d-2}(\mathbb{S}^{d-2}) \sin^{d-2}(\theta) e^{i|x|\cos(\theta)} d\theta \\ &= J_{\frac{d-2}{2}}(|x|) \pi^{\frac{d-1}{2}} (\frac{1}{2}|x|)^{-\frac{d-2}{2}}. \end{aligned}$$

This is seen by a substitution reducing the one-dimensional integral to the formula of Schläfli–Sommerfeld. The function $H(\xi)$ is real and radial. We choose a real function $\eta \in C^\infty(\mathbb{R})$, supported in $[-\frac{1}{2}, \infty)$, with $\eta(x) + \eta(-x) = 1$. Then $H(\xi)$ is the real part of

$$\int_{-\pi}^{\pi} \mathcal{H}^{d-2}(\mathbb{S}^{d-2}) \eta(\cos\theta) \sin^{d-2}(\theta) e^{i|x|\cos(\theta)} d\theta$$

An application of stationary phase gives

Lemma 8.1. *For all r, $H(r)$ is the real part of a function $e^{-ir}\phi$ which satisfies*

$$\Big| \Big(\frac{d}{dr}\Big)^k \phi \Big| \le c_k r^{-\frac{d-1}{2} - k}.$$

Proof. Exercise. □

Chapter 9

Appendic C: The Fourier transform

Let f be an integrable complex-valued function. We define its Fourier transform by

$$\hat{f}(\xi) = \frac{1}{(2\pi)^{d/2}} \int e^{-ix\cdot\xi} f(x) dm^d(x). \tag{9.1}$$

9.1 The Fourier transform in L^1

Properties are

1) The Fourier transform of an integrable function is a bounded continuous function which converges to 0 as $|\xi| \to \infty$. It satisfies

$$\|\hat{f}\|_{\sup} \leq (2\pi)^{-d/2} \|f\|_{L^1}.$$

The estimate is obvious, as is the continuity if f is compactly support. The limit as $|\xi| \to \infty$ follows by an integration by parts if the integrand is compactly supported and differentiable. Those functions are dense, and we obtain continuity and vanishing of the limit for compactly supported functions. The limit

$$\lim_{R\to\infty} \int_{B_R(0)} e^{-ix\cdot\xi} f(x) dm^d(x)$$

is uniform, and hence the Fourier transform is continuous and converges to 0 as $|\xi| \to \infty$.

2) For all η and y in $\mathbb{R}^d$,

$$\hat{f}(\xi + \eta) = \widehat{e^{-i\eta\cdot x} f} \tag{9.2}$$

and

$$\widehat{f(\cdot + y)} = e^{iy\xi} \hat{f}(\xi). \tag{9.3}$$

This follows by an simple calculation.

H. Koch et al., *Dispersive Equations and Nonlinear Waves: Generalized Korteweg–de Vries, Nonlinear Schrödinger, Wave and Schrödinger Maps*, Oberwolfach Seminars 45,
DOI 10.1007/978-3-0348-0736-4_9, © Springer Basel 2014

3) For $f, g \in L^1(\mathbb{R})$
$$\widehat{f * g}(\xi) = (2\pi)^{n/2}\hat{f}(\xi)\hat{g}(\xi).$$
This follows by application of Fubini's theorem:
$$\begin{aligned}
&\frac{1}{(2\pi)^{d/2}}\int e^{-ix\xi}\int f(y)g(x-y)dm^d(y)dm^d(\xi)\\
&= \frac{1}{(2\pi)^{d/2}}\int\int e^{-iy\xi}f(y)e^{-i(x-y)\xi}g(x-y)dm^d(y)dm^d(x)\\
&= \frac{1}{(2\pi)^{d/2}}\int\int e^{-iy\xi}f(y)e^{-iz\xi}g(z)dm^d(z)dm^d(y)\\
&= (2\pi)^{d/2}\hat{f}(\xi)\hat{g}(\xi).
\end{aligned}$$

4) For f and $g \in L^1$,
$$\int f\hat{g}dm^d(x) = \int \hat{f}gdm^d. \tag{9.4}$$
This is seen by applying Fubini to
$$\int\int e^{-iy\xi}f(y)e^{-i(x-y)\xi}g(y)dm^d(y)dm^d(x).$$

5)
$$\widehat{e^{-\frac{1}{2}|x|^2}} = e^{-\frac{1}{2}|\xi|^2}.$$
We calculate, as for the Airy function,
$$(2\pi)^{-d/2}\int e^{-ix\xi-\frac{1}{2}|x|^2}dm^d(x) = (2\pi)^{-d/2}\int e^{-i(x-i\eta)\xi-\frac{1}{2}(x-i\eta)^2}dm^d(x)$$
for $\eta \in \mathbb{R}^n$. We set $\eta = \xi$ and get
$$(2\pi)^{-d/2}e^{-\frac{|\xi|^2}{2}}\int e^{-\frac{1}{2}|x|^2}dx = e^{-\frac{|\xi|^2}{2}}.$$

9.2 The Fourier transform of Schwartz functions

Definition 9.1. *We say $f \in C^\infty(\mathbb{R}^d)$ is a Schwartz function, and write $f \in \mathcal{S}(\mathbb{R}^d)$, if for all multi-indices α and β*
$$\|x^\alpha\partial^\beta f\|_{\sup} < \infty.$$
We say that $f_j \to f$ in $\mathcal{S}$ if for all multi-indices α and β
$$x^\alpha\partial^\beta f_j \to x^\alpha\partial^\beta f$$
uniformly.

We collect below a number of elementary properties.

1) $f \in \mathcal{S}$ if and only if $x^\alpha \partial^\beta f \in \mathcal{S}$ for all α and β.
2) $f \in \mathcal{S}$ implies f integrable.
3) $f \in \mathcal{S}$ and $g \in C^\infty$ with bounded derivatives implies $fg \in \mathcal{S}$.
4) $f \in \mathcal{S}$ and A an invertible $d \times d$ matrix implies $f \circ A \in \mathcal{S}$.
5) $f \in \mathcal{S}$ and $x_0 \in \mathbb{R}^d$ implies $f(\cdot + x_0) \in \mathcal{S}$.
6) We say that a distribution T has compact support, if there exists a ball $B_R(0)$ such that for all functions f in $C_0^\infty(\mathbb{R}^d)$ with support disjoint from $B_R(0)$, $Tf = 0$. We can easily extend such distributions to Schwartz functions (exercise).

   We define the convolution with a Schwartz function by

$$T * f(x) = T(f(x - \cdot))$$

   This is well defined and $T * f$ is a Schwartz function whenever f is a Schwartz function. To see this we recall that, by the definition of a distribution, there exist $C > 0$ and $N > 0$ such that (since f has compact support)

$$|T(f)| \le c_N \|f\|_{C^N}.$$

   Taking difference quotients shows that $x \mapsto T * f(x)$ is differentiable and

$$\partial_i T * f = T * \partial_i f.$$

   Recursively we see that $Tf \in C^\infty$. Morever

$$\|f(x - \cdot)\|_{C^N(B_R(0))} \le c_M (1 + |x|)^{-M}$$

   for Schwartz functions, and hence $T * f$ is a Schwartz function.
7) $f, g \in \mathcal{S}$ implies $f * g \in \mathcal{S}$ and

$$\widehat{f * g} = (2\pi)^{d/2} \hat{f} \hat{g}. \tag{9.5}$$

   If $f \in \mathcal{S}$ and S is a distribution with compact support, then

$$S * f(x) := S(f(x - \cdot)) \in \mathcal{S}.$$

8) All the operations above are continuous.

Theorem 9.2. *If $f \in \mathcal{S}$, then $\hat{f} \in \mathcal{S}$, and vice versa,*

$$\widehat{x_j f} = -i \partial_{\xi_j} \hat{f},$$

$$\widehat{-i\partial_{x_j} f} = \xi_j \hat{f},$$

and the Fourier inversion formula

$$f(x) = (2\pi)^{-d/2} \int e^{ix\xi} \hat{f}(\xi) dm^d(\xi)$$

and the Plancherel formula

$$\int \hat{f}\, \bar{\hat{g}} dm^d(\xi) = \int f\, \overline{g} dm^d(x)$$

hold. If A is a real invertible $d \times d$ matrix, then

$$\widehat{f \circ A}(\xi) = (\det |A|)^{-1} \hat{f}(A^{-T}\xi).$$

Proof. The first two formulas formally follow from a simple calculation. According to property 1)

$$x^\alpha \partial^\beta f \in \mathcal{S},$$

and hence $x^\alpha \partial^\beta f$ is integrable. With the first calculation

$$\mathcal{F}(x^\alpha(-i\beta^\beta f)) = -i\partial^\alpha \xi^\beta \hat{f},$$

which is bounded by property 2). Thus $\hat{f} \in \mathcal{S}$. We calculate

$$\mathcal{F}((2\pi)^{-d/2} \tau^{d/2} e^{-\frac{\tau}{2}x^2} * f) = e^{-\frac{1}{2\tau}\xi^2} \hat{f}(\xi)$$

and, letting $\tau \to \infty$,

$$f(0) = (2\pi)^{-d/2} \int \hat{f} d\xi.$$

Together with the formulas (9.3) we obtain the inversion formula

$$f(x) = (2\pi)^{-d/2} \int e^{ix\xi} \hat{f}(\xi) d\xi.$$

The Plancherel formula follows by (9.4). The last formula follows from

$$(2\pi)^{-d/2} \int e^{-ix\cdot\xi} f(Ax) dm^d(x) = (2\pi)^{-d/2} |\det A|^{-1} \int e^{-i(A^{-1}y)\cdot\xi} f(y) dm^d(y).$$

$\square$

9.3 Tempered distributions

Definition 9.3. *A tempered distribution T is a linear map*

$$T : \mathcal{S} \to \mathbb{C}$$

which is continuous, i.e., $f_j \to f \in \mathcal{S}$ implies

$$Tf_j \to Tf.$$

We denote the set of tempered distributions by $\mathcal{S}^*$. We say that T_j converges to T if $T_j f \to Tf$ for all $f \in \mathcal{S}$.

We list a number of properties.

1) We call T bounded if there exists N such that

$$|Tf| \le C \sup_{|\alpha|+|\beta|\le N} \sup_x |x^\alpha \partial_x^\beta f|.$$

The linear $T : \mathcal{S} \to \mathbb{C}$ is bounded if and only if it is continuous.

2) Distributions with compact support are tempered distributions.

3) Let $T \in \mathcal{S}^*$ and $\phi \in C^\infty$ with bounded derivatives. We define

$$\phi T(f) = T(\phi f).$$

4) The derivative of a tempered distribution $\partial_j T$ is defined by

$$\partial_j T(f) = -T(\partial_j f).$$

5) Let $T \in \mathcal{S}^*$ and $\phi \in \mathcal{S}$. Then

$$T * \phi \in C^\infty(\mathbb{R}^d),$$

where we define $T * \phi$ as for distributions with compact support.

6) Let $T \in \mathcal{S}^*$ and S be a distribution with compact support. We define

$$S * T(f) = T(\tilde{S} * f),$$

where $\tilde{S}(f) = S(\tilde{f})$, $\tilde{f}(x) = f(-x)$. Then $S * T \in \mathcal{S}^*$.

7) Let $g \in L^p$ for one $1 \le p \le \infty$. Then g defines a unique distribution by

$$T_g(f) = \int g f dm^d.$$

The operations commute with this representation. We have for the Schwartz functions ϕ:

$$T_{\phi g} = \phi T_g,$$

and we identify L^p with its image via the embedding.

8) We define the Fourier transform $\hat{T} \in \mathcal{S}^*$ by

$$\hat{T}(f) = T(\hat{f}).$$

The inverse Fourier transform is defined similarly.

This is compatible with the interpretation for functions.

9)

$$\hat{\delta}_0 = (2\pi)^{d/2}$$

and

$$\hat{1} = (2\pi)^{d/2}\delta_0.$$

The Euler relation

$$x \cdot \nabla f = mf$$

holds for every homogeneous function of degree m. We want to define homogeneous distributions.

Definition 9.4. *A tempered distribution is called homogeneous of degree $m \in \mathbb{C}$ if*

$$T(\phi) = \lambda^{d+m} T(\phi(\lambda x)).$$

Let $\operatorname{Re} m > -d$. Then $|x|^m$ is a tempered distribution. Its Fourier transform is again a tempered distribution of homogeneity $-d-m$. This can be seen from the Euler relation $x \cdot \nabla f = mf$.

Lemma 9.5. *Let $0 < \operatorname{Re} m < d$. The following identity holds:*

$$\mathcal{F}\Big(\frac{1}{2^{m/2}\Gamma(m/2)}|x|^{m-d}\Big) = \frac{1}{2^{(d-m)/2}\Gamma(\frac{d-m}{2})}|x|^{-m}.$$

Proof. We claim that the Fourier transform of a homogeneous distribution of degree $m \in \mathbb{C}$ is a homogeneous distribution of degree $-d-m$. We denote by T_λ the distribution

$$T_\lambda(f) = \lambda^{-d} T f_{\lambda^{-1}}.$$

Here $f_\lambda(x) = f(x/\lambda)$. Then

$$\hat{T}_\lambda(f) = T_\lambda(\hat{f}) = T(\lambda^{-d}\hat{f}(\lambda \cdot)) = T(\widehat{f(\cdot/\lambda)}) = \lambda^{-m-d}T(\hat{f}) = \lambda^{-m-d}\hat{T}(f).$$

Let f be a homogeneous function of degree m such that T_f is a homogeneous distribution. Let O be an orthogonal matrix with $f \circ O = f$. Then

$$\widehat{T_f} \circ O^T = \widehat{T_f}$$

where the term on the left-hand side is defined by the action on Schwartz functions. In particular, the Fourier transform of $|x|^{-m}$ is radial in the sense that it is invariant under the action of orthogonal matrices. This is equivalent to

$$Tf = T\left(\mathcal{H}^{d-1}(\mathbb{S}^{d-1})^{-1}\int_{\mathbb{S}^{d-1}} f(|x|\sigma)\,\mathcal{H}^{d-1}(\sigma))\right)$$

(a rigorous justification requires either an approximation, or a symmetrization argument). We denote the symmetrization operator by S.

Let T be a radial homogeneous distribution of degree m. We fix a nonnegative function h with integral 1 and with compact support and observe that $Tf = T(Sf) = \lambda T(Sf(\lambda x))$. Let $e \in \mathbb{R}^d$ be a unit vector.

$$\begin{aligned}
T(f) &= \int_0^\infty \lambda^{-1} h(\ln\lambda) T(Sf) d\lambda \\
&= \int_0^\infty \lambda^{m+d-1} h(\ln\lambda) T(Sf(\lambda x)) d\lambda \\
&= T\left(\int_0^\infty (Sf)(\mu e)(\mu/|x|)^{d+m}\mu^{-1} h(\ln(\mu/|x|)) d\mu\right) \\
&= \int_0^\infty \mu^{-1} Sf(\mu e) T((\mu/|x|)^{d+m} h(\ln(\mu/|x|))) d\mu \\
&= -T(|x|^{-d-m} h(\ln|x|))(d\, m^d(B_1(0)))^{-1}\int_{\mathbb{R}^d} |y|^m f(y)) d\mu.
\end{aligned}$$

This shows that a rotational symmetric distribution of homogeneity $m > -d$ is given by $c|x|^m$. Below we determine c for the Fourier transform of $|x|^m$ with $-d < m < 0$.

By the consideration above,

$$\widehat{|x|^{-m}} = c(n,m)|x|^{m-d}$$

and we have to determine $c(n,m)$. The Gaussian is its own Fourier transform. Let $T = |x|^m$ and denote by $\hat{T}$ its Fourier transform. Then, by the definition,

$$T(e^{-\frac{|x|^2}{2}}) = \hat{T}(e^{-\frac{|\xi|^2}{2}}).$$

We calculate

$$\begin{aligned}
\int |x|^m e^{-\frac{|x|^2}{2}} dm^d(x) &= d\, m^d(B_1(0)) \int_0^\infty e^{-r^2/2} r^{d-1+m} dr \\
&= d\, m^d(B_1(0)) 2^{-\frac{d+m}{2}-1} \int_0^\infty t^{\frac{d+m}{2}-1} e^{-t} dt \\
&= d\, m^d(B_1(0)) 2^{-\frac{d+m}{2}-1} \Gamma(\frac{d+m}{2}).
\end{aligned}$$

Comparison with the calculation for $|x|^{-d-m}$ gives the formula. $\square$

The formula extends to all $m \in \mathbb{C}\backslash(-\infty, -d] \cup [0, \infty)$. This requires however a proper definition of the homogeneous tempered distribution.

Bibliography

[1] Georg Aumann. *Reelle Funktionen.* Die Grundlehren der mathematischen Wissenschaften in Einzeldarstellungen mit besonderer Berücksichtigung der Anwendungsgebiete, Bd LXVIII. Springer-Verlag, Berlin, 1954.

[2] Jean Bourgain. Pointwise ergodic theorems for arithmetic sets. *Inst. Hautes Études Sci. Publ. Math.*, (69):5–45, 1989. With an appendix by the author, Harry Furstenberg, Yitzhak Katznelson and Donald S. Ornstein.

[3] Tristan Buckmaster and Herbert Koch. The Korteweg-de-Vries equation at H^{-1} regularity . *arXiv:1112.4657v2*, 2012.

[4] Thierry Cazenave. *Semilinear Schrödinger equations*, volume 10 of *Courant Lecture Notes in Mathematics.* New York University Courant Institute of Mathematical Sciences, New York, 2003.

[5] Michael Christ, James Colliander, and Terrence Tao. Asymptotics, frequency modulation, and low regularity ill-posedness for canonical defocusing equations. *Amer. J. Math.*, 125(6):1235–1293, 2003.

[6] J. Dieudonné. *Foundations of modern analysis.* Pure and Applied Mathematics, Vol. X. Academic Press, New York, 1960.

[7] Damiano Foschi and Sergiu Klainerman. Bilinear space-time estimates for homogeneous wave equations. *Ann. Sci. École Norm. Sup. (4)*, 33(2):211–274, 2000.

[8] Peter K. Friz and Nicolas B. Victoir. *Multidimensional stochastic processes as rough paths*, volume 120 of *Cambridge Studies in Advanced Mathematics.* Cambridge University Press, Cambridge, 2010. Theory and applications.

[9] Loukas Grafakos. *Classical Fourier analysis*, volume 249 of *Graduate Texts in Mathematics.* Springer, New York, second edition, 2008.

[10] Axel Grünrock. A bilinear Airy-estimate with application to gKdV-3. *Differential Integral Equations*, 18(12):1333–1339, 2005.

[11] Axel Grünrock. On the Cauchy-problem for generalized Kadomtsev-Petviashvili-II equations. *Electron. J. Differential Equations*, pages No. 82, 9, 2009.

H. Koch et al., *Dispersive Equations and Nonlinear Waves: Generalized Korteweg–de Vries, Nonlinear Schrödinger, Wave and Schrödinger Maps*, Oberwolfach Seminars 45,
DOI 10.1007/978-3-0348-0736-4, © Springer Basel 2014

[12] M. Hadac, S. Herr, and H. Koch. Well-posedness and scattering for the KP-II equation in a critical space. *Annales de l'Institut Henri Poincare / Analyse non lineaire*, 2008.

[13] G. H. Hardy and J. E. Littlewood. A convergence criterion for Fourier series. *Math. Z.*, 28(1):612–634, 1928.

[14] B. B. Kadomtsev and V. I. Petviashvili. On the stability of solitary waves in weakly dispersing media. *Sov. Phys., Dokl.*, 15:539–541, 1970.

[15] Markus Keel and Terence Tao. Endpoint Strichartz estimates. *Amer. J. Math.*, 120(5):955–980, 1998.

[16] Carlos E. Kenig, Gustavo Ponce, and Luis Vega. Oscillatory integrals and regularity of dispersive equations. *Indiana Univ. Math. J.*, 40(1):33–69, 1991.

[17] Carlos E. Kenig, Gustavo Ponce, and Luis Vega. Well-posedness and scattering results for the generalized Korteweg-de Vries equation via the contraction principle. *Comm. Pure Appl. Math.*, 46(4):527–620, 1993.

[18] H. Koch and D. Tataru. Dispersive estimates for principally normal pseudodifferential operators. *Comm. Pure Appl. Math.*, 58(2):217–284, 2005.

[19] H. Koch and D. Tataru. A Priori Bounds for the 1D Cubic NLS in Negative Sobolev Spaces. *Int. Math. Res. Not.*, 2007:Article ID rnm053, 2007.

[20] Herbert Koch and Jeremy L. Marzuola. Small data scattering and soliton stability in $\dot{H}^{16}$ for the quartic KdV equation. *Analysis and PDE*, 5(1):145–198, 2012.

[21] D. Lépingle. La variation d'ordre p des semi-martingales. *Z. Wahrscheinlichkeitstheorie und Verw. Gebiete*, 36(4):295–316, 1976.

[22] Elliott H. Lieb and Michael Loss. *Analysis*, volume 14 of *Graduate Studies in Mathematics*. American Mathematical Society, Providence, RI, second edition, 2001.

[23] Terry Lyons. Differential equations driven by rough signals. I. An extension of an inequality of L. C. Young. *Math. Res. Lett.*, 1(4):451–464, 1994.

[24] Terry J. Lyons. Differential equations driven by rough signals. *Rev. Mat. Iberoamericana*, 14(2):215–310, 1998.

[25] Jacques Neveu. *Martingales à temps discret*. Masson et Cie, éditeurs, Paris, 1972.

[26] Frank W. J. Olver, Daniel W. Lozier, Ronald F. Boisvert, and Charles W. Clark, editors. *NIST handbook of mathematical functions*. U.S. Department of Commerce National Institute of Standards and Technology, Washington, DC, 2010. With 1 CD-ROM (Windows, Macintosh and UNIX).

[27] Elias M. Stein. *Harmonic analysis: real-variable methods, orthogonality, and oscillatory integrals*, volume 43 of *Princeton Mathematical Series*. Princeton University Press, Princeton, NJ, 1993. With the assistance of Timothy S. Murphy, Monographs in Harmonic Analysis, III.

[28] Nils Strunk. Well-posedness for the supercritical gKdV equation. *Communications on Pure and Applied Analysis*, 13(2), 2014.

[29] Terence Tao. *Nonlinear dispersive equations*, volume 106 of *CBMS Regional Conference Series in Mathematics*. Published for the Conference Board of the Mathematical Sciences, Washington, DC, 2006. Local and global analysis.

[30] Terence Tao. Scattering for the quartic generalised Korteweg-de Vries equation. *J. Differential Equations*, 232(2):623–651, 2007.

[31] S. J. Taylor. Exact asymptotic estimates of Brownian path variation. *Duke Math. J.*, 39:219–241, 1972.

[32] Hans Triebel. *Theory of function spaces. II*, volume 84 of *Monographs in Mathematics*. Birkhäuser Verlag, Basel, 1992.

[33] N. Wiener. The quadratic variation of a function and its Fourier coefficients. *Journ. Mass. Inst. of Technology*, 3:73–94, 1924.

[34] L. C. Young. An inequality of the Hölder type, connected with Stieltjes integration. *Acta Math.*, 67(1):251–282, 1936.

Geometric Dispersive Evolutions

Daniel Tataru

Chapter 1

Introduction

Among the nonlinear dispersive equations, a distinguished class is that of geometric evolutions. Unlike the models seen earlier where nonlinear interactions are added to an underlying linear dispersive flow, here the nonlinear structure arises from the curvature of the state space itself. Precisely, our geometric evolutions are obtained by applying the standard linear Lagrangian or Hamiltonian formalism to a state space consiting of maps into (curved) manifolds.

The simplest geometric pde's are the elliptic and parabolic ones, namely the harmonic map equation and the harmonic heat flow. While these still play a role in our exposition, in these notes we are primarily concerned with the dispersive evolutions, the wave map equation and the Schrödinger map equation.

Both the short and the long time behavior of wave and Schrödinger maps are dependent on the curvature properties of the target manifold. Because of these, the model cases of maps into the sphere $\mathbb{S}^m$ and into the hyperbolic space play an important role.

Compared with other dispersive pde's, an additional structure present here is that of "gauge invariance". The simplest way this arises is in the choice of coordinates on the target manifold; also, in a more subtle way, in the choice of frames in the tangent space of the target manifold. Often a more favourable nonlinear structure is obtained by making a suitable choice of gauge. This is also related to the notion of "renormalization", which here represents a paradifferential version of choosing a good gauge.

The dimension of the underlying space-time affects the scaling and criticality properties of our equations. Our primary target here is the case of 2+1 dimensions, which is arguably the most interesting. This is the energy critical case, i.e., for which the energy is invariant with respect to the natural scaling of the equations.

We begin these notes with a brief description of the state space of maps into manifolds, followed by an introduction of the four main pde's, namely harmonic maps, the harmonic heat flow, wave maps, and finally Schrödinger maps. Our main interest is in wave maps, where a series of developments in the last 15 years have led to a reasonably complete theory. We first discuss the small data case,

H. Koch et al., *Dispersive Equations and Nonlinear Waves: Generalized Korteweg–de Vries, Nonlinear Schrödinger, Wave and Schrödinger Maps*, Oberwolfach Seminars 45, DOI 10.1007/978-3-0348-0736-4_10, © Springer Basel 2014

where the emphasis is on function spaces and renormalization. Then we consider the large data problem, where in addition we bring in the concept of induction on energy, and study energy concentration using Morawetz estimates. Finally, the last section is concerned with the small data problem for Schrödinger maps, where the difficulties revolve around the gauge choice and function spaces. The large data problem for Schrödinger maps is still open.

Chapter 2

Maps into manifolds

Instead of working with real or complex-valued functions, our main objects of study here are evolutions whose state space, in the simplest setting, consists of maps from the Euclidean space $\mathbb{R}^n$ into a Riemannian manifold (M, g). More generally, one can consider maps whose domains are also Riemannian manifolds.

In terms of the target manifold (M, g), the most common situation we will consider is that of compact manifolds without boundary. Among these, the sphere $\mathbb{S}^2$ or its higher dimensional counterparts $\mathbb{S}^m$ will play the role of a model positively curved manifold. On such manifolds one often does not have a nice global coordinate chart. Thus, in order to describe global objects it is often convenient to view such manifolds, via Nash's theorem, as isometrically embedded into a higher dimensional Euclidean space,

$$(M, g) \hookrightarrow (\mathbb{R}^m, e).$$

We call this the **extrinsic setting**. The simplest such example is the unit sphere representation

$$\mathbb{S}^2 = \{x \in \mathbb{R}^3;\ |x| = 1\} \subset \mathbb{R}^3.$$

Among negatively curved manifolds, the model is the hyperbolic space $\mathbb{H}^2$ or, more generally, $\mathbb{H}^m$. While this is not compact, it can be viewed globally as embedded in the Minkowski space $(\mathbb{M}^{2+1}, m)$, with metric $m = ds^2 = -d\phi_0^2 + d\phi_1^2 + d\phi_2^2$:

$$\mathbb{H}^2 = \{\phi \in \mathbb{M}^{2+1};\ |\phi|_m^2 = -1\} \subset \mathbb{M}^{2+1}.$$

Alternatively, one can also use compact quotients of $\mathbb{H}^m$ as surrogates for $\mathbb{H}^m$. This is convenient if, for instance, one wants to adapt $\mathbb{H}^m$ to the extrinsic setting.

2.1 The tangent bundle and covariant differentiation

Given a differentiable map

$$\phi : \mathbb{R}^n \to (M, g)$$

H. Koch et al., *Dispersive Equations and Nonlinear Waves: Generalized Korteweg–de Vries, Nonlinear Schrödinger, Wave and Schrödinger Maps*, Oberwolfach Seminars 45,
DOI 10.1007/978-3-0348-0736-4_11, © Springer Basel 2014

we can define its partial derivatives with respect to the $\mathbb{R}^n$ coordinates at a point $x \in \mathbb{R}^n$, $\partial_i\phi(x) \in T_{\phi(x)}M$. These can be viewed as sections of a vector bundle E^ϕ over $\mathbb{R}^n$, where the fiber is given by $E^\phi_x = T_{\phi(x)}M$. Precisely, E^ϕ is a metric bundle, where the metric is inherited from TM.

On TM one has the Levi-Civita connection induced by the metric. Its pull-back to $\mathbb{R}^n$ is a connection on E^ϕ. The easiest way to describe it is by using a local coordinate chart on M. If in a chart ϕ is given by $\phi = (\phi^1, \dots \phi^m)$ and a section of E^ϕ is given by $v = (v^1, \dots, v^m)$, then the covariant derivatives of v are given by

$$\mathbf{D}_j v^k(x) = \partial_j v^k(x) + \Gamma^k_{il}\partial_j\phi^i v^l(x). \tag{2.1}$$

Here Γ^k_{il} represent the Riemann-Christoffel symbols on M. This is a metric connection, i.e. $\mathbf{D}g = 0$. Another way to express this property is via the relation

$$\mathbf{D}_j\langle v, w\rangle_g = \langle \mathbf{D}_j v, w\rangle_g + \langle v, \mathbf{D}_j w\rangle_g.$$

In particular, one can consider the covariant derivatives of $\partial_j\phi$; then it is easy to establish that

$$\mathbf{D}_i\partial_j\phi = \mathbf{D}_j\partial_i\phi. \tag{2.2}$$

Of course, the covariant derivatives themselves do not commute; instead, the curvature $\mathbf{R}$ of the connection $\mathbf{D}$ is related to the curvature tensor R of M. Precisely, for any two sections v, w of E^ϕ we have the relation

$$\langle [D_i, D_j]v, w\rangle_g = R(\partial_i\phi, \partial_j\phi, v, w). \tag{2.3}$$

Another way to express the covariant differentiation is in the context of the extrinsic setting. For this we assume that (M, g) is a submanifold of the Euclidean space $\mathbb{R}^m$. Then one can define the normal bundle NM. The second fundamental form $\mathcal{S}$ is a symmetric quadratic form

$$\mathcal{S} : TM \times TM \to NM,$$

given by

$$\langle \mathcal{S}(X, Y), \nu\rangle = \langle \nabla_X Y, \nu\rangle = -\langle X\nu, Y\rangle.$$

Here $X\nu$ is the X-derivative of ν since the Euclidean space is flat. In this context, the connection $\mathbf{D}$ can be expressed in terms of the second fundamental form $\mathcal{S}$ as

$$\mathbf{D}_j v^k(x) = \partial_j v^k(x) + \mathcal{S}^k_{il}\partial_j\phi^i v^l(x). \tag{2.4}$$

By the Gauss–Codazzi equations, the curvature of the connection takes the form

$$\langle [D_i, D_j]v, w\rangle_g = \langle \partial_i\phi, v\rangle_g\langle \partial_j\phi, w\rangle_g - \langle \partial_j\phi, v\rangle_g\langle \partial_i\phi, w\rangle_g. \tag{2.5}$$

2.2 Special targets

For the most part, the work so far in geometric dispersive equations is devoted to special targets, namely the sphere $\mathbb{S}^2$ (or $\mathbb{S}^m$) and the hyperbolic space $\mathbb{H}^2$ (or $\mathbb{H}^m$). The advantage is that the algebra is simpler, while one hopes that nothing fundamental is lost in the process. In both cases the preferred setting is the extrinsic one.

Consider first the sphere $\mathbb{S}^2 \subset \mathbb{R}^3$, and a map $\phi : \mathbb{R}^n \to \mathbb{S}^2$. By a slight abuse of notation, we also use ϕ for the coordinates in $\mathbb{R}^3$. Then ϕ represents the unit outer normal to the sphere. The second fundamental form of the sphere is

$$\mathcal{S}(u, v) = -\langle u, v\rangle, \quad u, v \perp \phi.$$

The sections of E are $\mathbb{R}^3$-valued vector fields u with the property that $\langle u, \phi\rangle = 0$. The covariant derivatives are given by

$$D_j u = \partial_j u - \langle u, \partial_j \phi\rangle\phi, \tag{2.6}$$

and their commutator is

$$[D_i, D_j]u = \langle \partial_i\phi, u\rangle\partial_j\phi - \langle \partial_j\phi, u\rangle\partial_i\phi.$$

The case of $\mathbb{H}^2$ is almost identical. Representing it as the space-like hyperboloid

$$-\phi_0^2 + \phi_1^2 + \phi_2^2 = -1$$

in the Minkowski space $(\mathbb{M}^{2+1}, m)$, the upward normal is still given by ϕ and the above formulas for covariant differentiation remain unchanged provided that the inner products are now taken with respect to the Minkowski metric.

2.3 Sobolev spaces

The question of characterizing the Sobolev regularity of maps between manifolds is not fully understood at this time, and many open problems exist. The discussion below is confined to the specific setting that is needed later in these notes. For further references we refer the reader to the survey paper [27].

The issue at hand is primarily to understand the H^s regularity of maps $\phi : \mathbb{R}^n \to (M, g)$. There is a natural scaling law associated with such maps,

$$\phi(x) \to \phi(\lambda x).$$

In terms of L^2 based Sobolev norms, the one with exactly this scaling law is the $\dot{H}^{\frac{n}{2}}$ norm. The problems which we will discuss later all have $\dot{H}^{\frac{n}{2}}$ as a critical (scale invariant) Sobolev norm. Hence most of our discussion will revolve around $\dot{H}^{\frac{n}{2}}$. We also care about higher regularity; to study that we will consider the spaces $\dot{H}^s \cap \dot{H}^{\frac{n}{2}}$ for $s > \frac{n}{2}$. Finally, in various contexts we need to measure the regularity of sections of the vector bundle E^ϕ. For this we will still use homogeneous Sobolev spaces $\dot H^s$, but here we will allow a range of s below $\frac{n}{2}$.

A key feature of the space $\dot H^{\frac{n}{2}}$ is that it is a threshold in terms of Sobolev embeddings. Precisely, the embedding $\dot H^{\frac{n}{2}} \subset L^\infty$ barely fails, and instead we have $\dot H^{\frac{n}{2}} \subset VMO$, the space of functions with vanishing mean oscillation. So while $\dot H^{\frac{n}{2}}$ functions are not continuous, they are almost localized in the sense that on small sets they vary very little in average.

As it turns out, VMO is a borderline space as far as the topological properties of maps are concerned. Precisely, the homotopy of VMO maps is well defined, and one can use the homotopy classes in order to partition VMO (and also $\dot H^{\frac{n}{2}}$) into connected components.

Another consequence of working with $\dot H^{\frac{n}{2}}$ is that it is not possible to confine the range of a map to the domain of a local chart on M, not even locally. Thus the extrinsic setting seems far more desirable from this perspective.

The space of maps $\phi : \mathbb{R}^n \to (M, g)$ is not a linear space, so one cannot endow it with a norm. There are two main methods to define the class of $\dot H^{\frac{n}{2}}$ maps:

In the extrinsic setting, where we have a uniform isometric embedding $(M, g) \hookrightarrow (\mathbb{R}^m, e)$. There one can simply view maps $\phi : \mathbb{R}^n \to M$ as maps $\phi : \mathbb{R}^n \to \mathbb{R}^m$ which happen to take values in M. Then their regularity, as well as the regularity of sections of E^ϕ, is computed on components as real-valued functions.

This is the most convenient setting to use in the analysis. The disadvantage is that it is not at all obvious whether this definition is geometric or it depends on the embedding at hand.

In the geometric setting. The easier case is when n is even. Then for ϕ smooth and constant outside a compact set one can define the homogeneous $\dot H^k$ Sobolev size of ϕ by

$$\|\phi\|^2_{\dot H^k} = \sum_j \sum_{|\alpha|=k-1} \|\mathbf{D}^\alpha \partial_j \phi\|^2_{L^2}, \quad k \geq \frac{n}{2}.$$

Then one can define the set of $\dot H^{\frac{n}{2}}$ maps by taking, say, L^2_{loc} limits of sequences which have bounded size in the above sense.

One can also endow the vector bundle E with a related norm. Precisely, for $v \in E$ we set

$$\|v\|^2_{\dot H^k} = \sum_{|\alpha|=k} \|\mathbf{D}^\alpha v\|^2_{L^2}, \quad 0 \leq k \leq \frac{n}{2}.$$

In the case of odd n one needs to work with fractional spaces, and for that it is necessary to consider a more roundabout route. This is based on the Littlewood-Paley theory. To describe the idea, we begin with a complex-valued function $\phi : \mathbb{R}^n \to \mathbb{C}$. To ϕ we associate its Littlewood-Paley truncations $\phi_{\leq k}$ to frequencies less than 2^k, as well as its dyadic pieces $\phi_k = \frac{d}{dk}\phi_{<k}$, where k is a real dyadic frequency parameter. Then for any large N we have

$$\|\phi\|^2_{\dot H^s} = c_{s,k} \int_{-\infty}^{\infty} 2^{2sk} \|\phi_k\|^2_{L^2} + 2^{2(s-N)k} \|\phi_k\|^2_{\dot H^N}\, dk.$$

If, instead of taking $\phi_{<k}$ to be the exact Littlewood-Paley localization of ϕ, one takes an arbitrary smooth function which decays to 0 as $k \to -\infty$ and converges to ϕ as $k \to \infty$, then the above equality becomes an inequality,

$$\|\phi\|_{\dot{H}^s}^2 \lesssim \int_{-\infty}^{\infty} 2^{2sk}\|\phi_k\|_{L^2}^2 + 2^{2(s-N)k}\|\phi_k\|_{\dot{H}^N}^2 \, dk.$$

Then the $\dot{H}^s$ norm of ϕ can be defined by minimizing the right hand side with respect to all extensions $\phi_{<k}$ of ϕ as above,

$$\|\phi\|_{\dot{H}^s}^2 \approx \inf_{\phi_{<k}} \int_{-\infty}^{\infty} 2^{2sk}\|\frac{d}{dk}\phi_{<k}\|_{L^2}^2 + 2^{2(s-N)k}\|\frac{d}{dk}\phi_{<k}\|_{\dot{H}^N}^2 \, dk.$$

The above definition involves only integer H^s norms, and it carries over easily to our context. Precisely, given a measurable map

$$\phi_0 : \mathbb{R}^n \to M$$

we call a smooth function

$$\phi : \mathbb{R} \times \mathbb{R}^n \to M$$

an admissible extension of ϕ_0 if $\lim_{k\to\infty} \phi(k) = \phi_0$ in L^2, and $\lim_{k\to-\infty} \nabla\phi(k) = 0$. Then we set

$$\|\phi_0\|_{\dot{H}^s} = \inf_{\phi \text{ admissible}} \int_{-\infty}^{\infty} 2^{2sk}\|\partial_k\phi(k)\|_{L^2}^2 + 2^{2(s-N)k}\|\partial_k\phi(k)\|_{\dot{H}^N}^2 \, dk.$$

A similar definition applies to sections of E^{ϕ_0}. There one needs to consider also extensions to sections of E^{ϕ}.

An alternate route is to consider a distinguished extension rather that all possible extensions. A suitable one is given for instance by the harmonic heat flow described below.

To compare the above $\dot{H}^s$ classes of maps we have the following:

Theorem 2.1 ([48]). *The extrinsic $\dot{H}^{\frac{n}{2}}$ class and the geometric $\dot{H}^{\frac{n}{2}}$ class are equivalent for small $\dot{H}^{\frac{n}{2}}$ sizes. In the same context, the higher regularity classes of maps $\dot{H}^s \cap \dot{H}^{\frac{n}{2}}$ are also equivalent.*

Likely this correspondence extends to all maps in the zero homotopy class. Unfortunately the geometric definition, as stated, applies only to homotopy zero maps.

2.4 $\mathbb{S}^2$ and targets: homotopy classes and equivariance

As mentioned before, the family of $\dot{H}^{\frac{n}{2}}$ maps is divided into connected components, indexed by the homotopy class. One model case of interest is that of maps $\phi : \mathbb{R}^2 \to \mathbb{S}^2$. There the homotopy class is indexed by integers m, computed via the formula

$$\int_{\mathbb{R}^2} \phi \cdot (\partial_1\phi \times \partial_2\phi) dx = 4\pi m.$$

Here the integrand is exactly the pull-back of the volume form on $\mathbb{S}^2$, and the integral is finite for all finite energy maps by the Cauchy–Schwarz inequality. Intuitively, this measures the number of times the map ϕ wraps around the sphere.

We remark that in the case of the $\mathbb{H}^2$ target all the finite energy maps have homotopy zero, and the direct analogue of the above integral vanishes.

In many difficult nonlinear pde's one can gain insights by studying classes of solutions which have additional symmetries. Often one uses the class of radial solutions. In our case, spherically symmetric maps are less useful, in part because they have homotopy zero (the integrand above is in fact identically zero). Instead, the interesting class of maps is the **equivariant** class.

The **equivariant** maps are maps which, when expressed in polar coordinates, satisfy

$$\phi(r,\theta) = (u(r), k\theta + \theta_0(r)), \quad u \in [0,\pi], \tag{2.7}$$

where k is the equivariance class. Another interpretation of this is the relation

$$\phi(Rx) = R^k \phi(x),$$

where R stands at the same time for a rotation around the origin in $\mathbb{R}^2$, respectively a rotation around the N-S axis in $\mathbb{S}^2$.

Here $k = 0$ corresponds to radial symmetry. If $k \neq 0$, then all $\dot{H}^{\frac{n}{2}}$ equivariant maps must have a limit at 0 and at infinity, which can be either pole, S or N. The homotopy index is then a multiple of the equivariance class.

We also remark that, in a more restrictive interpretation, sometimes one defines equivariant maps as maps of the form

$$\phi(r,\theta) = (u(r), k\theta + \theta_0). \tag{2.8}$$

This works for harmonic maps, the harmonic heat flow, and wave maps. However, this restricted class is not invariant with respect to the Schrödinger map flow.

2.5 Frames and gauge freedom

This approach to the study of maps from $\mathbb{R}^n$ into manifolds begins with a choice of an orthonormal frame $\{e_k(\phi)\}$ in $T_\phi M$. Then the idea is to describe the map ϕ via its gradient expressed in this frame. We obtain the **differentiated fields** ψ_α, given by

$$\psi_{\alpha,k} = \langle \partial_\alpha \phi, e_k \rangle_g.$$

To start with, these satisfy the compatibility conditions

$$\mathbf{D}_\alpha \psi_\beta = \mathbf{D}_\beta \psi_\alpha, \tag{2.9}$$

where the new covariant differentiation operators $\mathbf{D}_\alpha$, expressed in the frame, have the form

$$\mathbf{D}_\alpha = \partial_\alpha + A_\alpha.$$

Here the **connection coefficients** A_α are antisymmetric matrices given by

$$(A_\alpha)_{jk} = \langle e_j, D_\alpha e_k \rangle_g.$$

A priori the coefficients A_α satisfy the curl system

$$(\partial_\alpha A_\beta - \partial_\beta A_\alpha)_{jk} = R(\partial_\alpha \phi, \partial_\beta \phi, e_j, e_k) = \psi_{\alpha,i}\psi_{\beta,l} R(e_i, e_l, e_j, e_k), \tag{2.10}$$

where R is the Riemann curvature tensor on (M, g). This is not yet a well determined system because the orthonormal frame has not been specified. Varying the frame choice leads to the gauge invariance

$$\psi_\alpha \to \mathcal{O}\psi_\alpha, \quad A_\alpha \to \mathcal{O}A_\alpha\mathcal{O}^{-1} - \partial_\alpha \mathcal{O}\mathcal{O}^{-1}, \quad \mathcal{O} \in SO(m).$$

Specifying an orthonormal frame is called fixing the gauge.

Assuming that M is parallelizable, one natural option would be to consider a fixed frame which is tied to M. However, this does not improve at all the analysis, and defeats the purpose of trying to express all equations exclusively in terms of the differentiated fields ψ_α. Indeed, the main advantage of the frame method is that one can produce equations with a better structure by choosing a favorable frame which depends not only on M, but also on the map ϕ.

Another obstruction to the above goal has to do with the fact that in general the curvature tensor in (2.10) depends on the original map ϕ. However, there is one interesting case when we do obtain a self-contained system, namely when M has constant curvature κ. Then the system (2.10) can be rewritten in the simpler form

$$\partial_\alpha A_\beta - \partial_\beta A_\alpha = \kappa(\psi_\alpha \otimes \psi_\beta - \psi_\beta \otimes \psi_\alpha). \tag{2.11}$$

For this reason, the frame method has been primarily used so far in the case when M is either the sphere or the hyperbolic space.

An obvious way to complete this system and uniquely determine A is to add the divergence relation

$$\partial_\alpha A_\alpha = 0. \tag{2.12}$$

This is called the **Coulomb gauge**. Then A_α are uniquely determined by (2.10) and (2.12), namely

$$A_\alpha = -\frac{1}{2}\kappa\Delta^{-1}\partial_\beta(\psi_\alpha \otimes \psi_\beta - \psi_\beta \otimes \psi_\alpha). \tag{2.13}$$

A further simplification occurs when the target manifold is two-dimensional. Then $\psi_\alpha \in \mathbb{R}^2$, which we identify with $\mathbb{C}$. On the other hand, A_α can be viewed as real rotation coefficients. Then the ψ_α belong to a complex vector bundle over $\mathbb{R}^n$ endowed with the connection

$$D_\alpha = \partial_\alpha + iA_\alpha.$$

The curl relations (2.10) become

$$\partial_\alpha A_\beta - \partial_\beta A_\alpha = \kappa \Im(\psi_\alpha \bar{\psi}_\beta) \tag{2.14}$$

and the gauge freedom translates to

$$\psi_\alpha \to e^{i\chi}\psi_\alpha, \quad A_\alpha \to A_\alpha + \partial_\alpha \chi,$$

where χ is any real-valued function. In the Coulomb gauge the connection coefficients are given by

$$A_\alpha = -\frac{1}{2}\kappa \Delta^{-1} \partial_\beta \Im(\psi_\alpha \bar{\psi}_\beta). \tag{2.15}$$

As a final remark here, the Coulomb gauge works well in high dimensions (say $n \geq 4$). In low dimensions, however, there are issues associated to high $\times$ high $\to$ low frequency interactions in the above expression for A, and new gauge choices are needed. The situation improves somewhat if one considers maps with extra symmetries (e.g., equivariant).

Chapter 3

Geometric pde's

3.1 Harmonic maps

We first review the linear Laplace equation. For functions

$$\phi : \mathbb{R}^n \to \mathbb{R}$$

we define the Lagrangian

$$L^e(\phi) = \frac{1}{2} \int_{\mathbb{R}^n} |\nabla_x \phi|^2 \, dx = \frac{1}{2} \int_{\mathbb{R}^n} \partial_\alpha \phi \cdot \partial_\alpha \phi \, dx, \tag{3.1}$$

with the Einstein summation convention. Local critical points solve the corresponding Euler-Lagrange equation, which is the Laplace equation.

$$-\Delta \phi = 0, \quad \text{or} \quad -\partial_j \partial_j \phi = 0.$$

We now repeat the above process, but with the key difference that instead of considering maps ϕ which take real or complex values, we consider maps which take values in a Riemannian manifold (M, g). The analogue of the elliptic Lagrangian in (3.1) is

$$L^e(\phi) = \frac{1}{2} \int_{\mathbb{R}^n} \langle \partial_\alpha \phi, \partial_\alpha \phi \rangle_g \, dx. \tag{3.2}$$

The associated Euler-Lagrange equation is called the **harmonic map** equation, and is similar to the Laplace equation, namely

$$-\mathbf{D}_\alpha \partial_\alpha \phi = 0, \tag{3.3}$$

where $\mathbf{D}_j$ are the covariant differentiation operators introduced in the previous section. Thus the above equation is no longer a linear equation; instead, as we shall see in a moment, it becomes a semilinear elliptic equation.

Expressed in local coordinates on the target manifold, the above equation takes the form

$$-\Delta \phi^i = \Gamma^i_{jk}(\phi) \partial_\alpha \phi^j \partial_\alpha \phi^k.$$

H. Koch et al., *Dispersive Equations and Nonlinear Waves: Generalized Korteweg–de Vries, Nonlinear Schrödinger, Wave and Schrödinger Maps*, Oberwolfach Seminars 45,
DOI 10.1007/978-3-0348-0736-4_12, © Springer Basel 2014

This problem is invariant with respect to the dimensionless scaling

$$\phi(x) \to \phi(\lambda x),$$

therefore a natural translation invariant setting to study this problem is that of the Sobolev space $\dot{H}^{\frac{n}{2}}$. On the other hand, the Lagrangian is invariant with respect to this scaling only if $n = 2$. We call that the **energy critical** problem. The higher dimensional case $n > 2$ is energy supercritical.

As mentioned before, another fact to consider is that $\dot{H}^{\frac{n}{2}}$ functions are not necessarily bounded. Hence there is no guarantee that any such map will stay locally within the domain of a local chart on M. This emphasizes the global aspects of the problem, and effectively eliminates the use of local coordinates in the study of the equation.

Switching to the extrinsic setting, the harmonic map equation takes the form

$$-\Delta\phi^i = \mathcal{S}^i_{jk}(\phi)\partial_\alpha\phi^j\partial_\alpha\phi^k.$$

While just considering the above equation involves no additional structure, one has to also keep in mind the geometric properties of the second fundamental form. In particular, we have the relation

$$\mathcal{S}^k_{ji}(\phi)\partial_\alpha\phi^k = 0,$$

as one is a normal vector and the other is a tangent vector to M. Thus one can rewrite the equation in the form

$$-\Delta\phi^i = (\mathcal{S}^i_{jk}(\phi) - \mathcal{S}^k_{ji}(\phi))\partial_\alpha\phi^j\partial_\alpha\phi^k$$

which leads to the study of more general equations of the form

$$-\Delta\phi = \Omega_\alpha\partial_\alpha\phi$$

with the key property that $\Omega_\alpha \in \dot{H}^{\frac{n}{2}-1}$ are antisymmetric matrices.

From the perspective of geometric dispersive equations, harmonic maps are interesting as the steady states of the evolution problems. Thus it is useful to us to discuss the existence and regularity of harmonic maps. We begin with the local regularity question. In two dimensions this is provided by the following result for finite energy maps:

Theorem 3.1 (Hélein [19]). *Harmonic maps with locally finite energy are smooth in the energy critical case $n = 2$.*

The frame method and the Coulomb gauge have played a critical role in Heléin's approach. Their role is roughly to produce an elliptic equation with a perturbative nonlinearity. However, an alternate, more recent approach by Rivière [33] uses the extrinsic formulation of the problem. The higher dimensional counterpart of the above result is as follows[1]:

[1] Their results are actually stronger than stated here.

Theorem 3.2 (Evans [14], Bethuel [7]). *Local $\dot{H}^{\frac{n}{2}}$ harmonic maps are smooth.*

Secondly, we discuss the issue of existence of nontrivial finite energy harmonic maps in dimension $n = 2$. This is relevant since such maps are stationary solutions for wave and Schrödinger maps. The answer to this question depends on the geometry of the target manifold. We consider two opposite examples. The first is the hyperbolic space $\mathbb{H}^2$, where we have the following Liouville type result:

Theorem 3.3 (Lemaire [24]). *There are no nontrivial finite energy harmonic maps from $\mathbb{R}^2$ into*[2] $\mathbb{H}^2$.

By contrast, the class of finite energy harmonic maps $\phi : \mathbb{R}^2 \to \mathbb{S}^2$ is quite rich. To describe it, we first recall that the class of all finite energy maps $\phi : \mathbb{R}^2 \to \mathbb{S}^2$ consists of infinitely many connected components, indexed by their homotopy class $k \in \mathbb{Z}$ defined by

$$4\pi k = \int_{\mathbb{R}^2} \phi \cdot (\partial_1 \phi \times \partial_2 \phi) dx.$$

This is finite since by Cauchy–Schwarz we have

$$4\pi |k| \lesssim \frac{1}{2} \int_{\mathbb{R}^2} |\partial_1 \phi|^2 + |\partial_2 \phi|^2 dx = E(\phi).$$

Within each homotopy class one can look for energy minimizers which turn out to have energy exactly $4\pi|k|$. In order for equality to hold above the two derivatives $\partial_1 \phi$ and $\partial_2 \phi$ must be orthogonal and of equal size. This means that ϕ must be conformal. Such maps are nonunique due to the many symmetries of the problem. To remove some of the degrees of freedom, we turn our attention to k-equivariant maps which take 0 to the south pole and infinity to the north pole. Then, for $k \neq 0$, one can find a k-equivariant harmonic map with energy $4\pi k$, namely

$$Q^k(r, \theta) = (2 \tan^{-1}(r^k), k\theta), \quad k \geq 1,$$

which is unique modulo scaling and rotations.

3.2 The harmonic heat flow

Starting again with the Euclidean case, consider the gradient flow associated to the Lagrangian (3.1). We obtain the heat equation in $\mathbb{R} \times \mathbb{R}^n$, namely

$$(\partial_t - \Delta)\phi = 0 \quad \text{or} \quad (\partial_t - \partial_\alpha \partial_\alpha)\phi = 0, \quad \phi(0) = \phi_0.$$

The geometric analogue of this, namely the **harmonic heat flow**, is the gradient flow associated to the geometric Lagrangian (3.2). The equation has the form

$$\partial_t \phi - \mathbf{D}_\alpha \partial_\alpha \phi = 0, \quad \phi(0) = \phi_0 : \mathbb{R}^n \to M. \tag{3.4}$$

[2] The same result holds for any negatively curved target.

This is a semilinear parabolic equation for which L^e is a Lyapunov functional,

$$\frac{d}{dt} L^e(\phi) = -\int_{\mathbb{R}^n} \langle \mathbf{D}_i \partial_i \phi, \mathbf{D}_j \partial_j \phi \rangle_g dx.$$

The associated scaling is

$$\phi(t,x) \to \phi(\lambda^2 t, \lambda x)$$

As before, this makes the problem energy critical in dimension $n = 2$, and supercritical in higher dimension.

In the extrinsic formulation the harmonic heat flow takes the form

$$(\partial_t - \Delta)\phi^i = \mathcal{S}^i_{jk}(\phi)\partial_\alpha \phi^j \partial_\alpha \phi^k, \quad \phi(0) = \phi_0. \tag{3.5}$$

This is a semilinear parabolic equation with a nonlinear constraint, namely that $\phi(t,x) \in M$ for all $(x,t) \in \mathbb{R}^{n+1}$. Extending $\mathcal{S}$ in any fashion outside M one may also interpret this equation as a parabolic equation for $\mathbb{R}^m$-valued functions, where the above constraint is dynamically preserved.

We begin with the small data problem, for which one can directly use perturbative techniques to solve the equation:

Theorem 3.4 (Chen-Ding [9]). *Assume that the initial data u_0 for the harmonic heat flow is small in the critical Sobolev space $\dot H^{\frac{n}{2}}$. Then there is a unique global solution, which is smooth for $t > 0$.*

A similar result holds for data which are small in the larger space BMO, see [25].

Consider now the large data problem. In supercritical dimensions $n \geq 3$, blow up can occur in finite time in a self-similar manner. However, in the critical dimension $n = 2$ the self-similar blow up is disallowed, and the only possibility for blow up is the "bubbling off" of harmonic maps, where a portion of the energy concentrates at a point close to a rescaled harmonic map, see Chen and Struwe [10] and Topping [49]. Precisely, we have the following result for energies below $E_{crit}(M)$, the lowest energy of a nontrivial harmonic map $\phi : \mathbb{R}^n \to M$:

Theorem 3.5 (Struwe [39], Qing and Tian [30], Smith [36]). *Let $n = 2$. Assume that the energy of the initial data u_0 for the harmonic heat flow is below $E_{\text{crit}}(M)$. Then there is a unique global solution, which is smooth for $t > 0$.*

In the particular case of the $\mathbb{H}^m$ target space, there are no nontrivial harmonic maps so there is a large data global well-posedness result. The case of the sphere $\mathbb{S}^2$ as a target is much richer. There we have at our disposal the equivariant harmonic maps Q_k described in the previous section, and a natural question is what happens for data that are close in energy to these. A result in [17] asserts that within the equivariant class the Q_k's are stable for $|k| \geq 3$. For $|k| = 2$ instability can occur, but there is no finite time blow up [16]. Finally, one can have finite time blow up for $k = 1$, see [31].

This seems to indicate that the generic blow-up pattern should be the bubbling off of single spheres, associated by a corresponding decrease in the homotopy class.

3.3 Wave maps

Formally, wave maps can be described by replacing the domain $\mathbb{R}^n$ used for harmonic maps by the the Minkowski space $\mathbb{M}^{n+1}$. For real-valued functions $\mathbb{M}^{n+1}$, the corresponding Lagrangian is

$$L^m(\phi) = \frac{1}{2}\int_{\mathbb{M}^{n+1}} -|\partial_t\phi|^2 + |\nabla\phi|^2 \; dxdt = \frac{1}{2}\int_{\mathbb{M}^{n+1}} \partial^\alpha\phi\partial_\alpha\phi \; dxdt, \tag{3.6}$$

where indices are lifted with respect to the Minkowski metric. The associated Euler-Lagrange equation is the wave equation in $\mathbb{M}^{n+1}$,

$$\Box\phi = 0, \quad \phi(0) = \phi_0, \quad \partial_t\phi(0) = \phi_1,$$

where the d'Alembertian is given by

$$\Box = \partial_t^2 - \Delta_x = -\partial^\alpha\partial_\alpha.$$

For functions with values in a Riemannian manifold (M, g) we can consider a similar Lagrangian to the above one,

$$L^m(\phi) = \frac{1}{2}\int_{\mathbb{M}^{n+1}} \langle\partial^\alpha\phi, \partial_\alpha\phi\rangle_g \;\; dxdt,$$

The associated Euler-Lagrange equation is called the **wave map** equation, and has the form

$$\mathbf{D}^\alpha\partial_\alpha\phi = 0, \quad \phi(0) = \phi_0, \quad \partial_t\phi(0) = \phi_1. \tag{3.7}$$

This is a semilinear wave equation, for which the initial position and velocity are maps

$$\phi_0 : \mathbb{R}^n \to M, \quad \phi_1 : \mathbb{R}^n \to T_{\phi_0}M,$$

with $\phi_1 \in E^{\phi_0}$. The steady states of this evolution are precisely the harmonic maps discussed before.

A feature which is shared with the linear wave equation is the conservation of the energy and momentum,

$$E(\phi) = \frac{1}{2}\int_{\mathbb{R}^n} |\partial_x\phi|^2 + |\partial_t\phi|^2 \; dx, \quad M_i(\phi) = \int_{\mathbb{R}^n} \partial_i\phi \cdot \partial_t\phi \; dx.$$

The scaling associated to this problem is

$$\phi(t, x) \to \phi(\lambda t, \lambda x),$$

so the scale invariant initial data space is $\dot{H}^{\frac{n}{2}} \times \dot{H}^{\frac{n}{2}-1}$. Again, the most interesting case is the energy critical case, $n = 2$.

In addition, the wave map problem inherits the full Lorentz group of symmetries from the linear wave equation. Thus, in addition to steady states (harmonic maps), we also have their Lorentz transforms, which are waves with a fixed profile and constant velocity (less than 1). It is worth noting that taking a Lorentz transform of a harmonic map leads to an increase in energy.

In the extrinsic formulation the wave map equation is:

$$\Box \phi^i = -\mathcal{S}^i_{jk}(\phi)\partial^\alpha \phi^j \partial_\alpha \phi^k, \quad \phi(0) = \phi_0, \quad \partial_t \phi(0) = \phi_1. \tag{3.8}$$

In the case of the $\mathbb{S}^m$ target this equation takes a very simple form,

$$\Box \phi = -\phi \, \langle \partial^\alpha \phi, \partial_\alpha \phi \rangle.$$

A very similar formula holds for maps into $\mathbb{H}^m$,

$$\Box \phi = \phi \, \langle \partial^\alpha \phi, \partial_\alpha \phi \rangle_m.$$

This problem is quite different from the corresponding heat flow, in that it is a **dispersive** equation. In other words, one has, on one hand, energy conservation, while, on the other hand linear waves travel (with speed one) in different directions and disperse. Hence, one does not expect, as in the parabolic case, a pure decay to a harmonic map pattern, but instead a more plausible picture is that of a splitting into one or more solitons (Lorentz transforms of harmonic maps) plus a dispersive part. While such a complete picture is not proved at the moment, considerable progress was made in recent years.

The first aim of the present notes is to describe the proof of the small data result:

Theorem 3.6 (Tao [45]: $\mathbb{S}^m$, Krieger [22]: $\mathbb{H}^2$, Tataru [48]: (M, g))**.** *The wave map equation is globally well-posed for initial data which are small in* $\dot{H}^{\frac{n}{2}} \times \dot{H}^{\frac{n}{2}-1}$.

This is done in the next section. The result is briefly stated above. A more precise formulation requires the introduction of a suitable function space S for the solutions, associated to the initial data space $\dot{H}^{\frac{n}{2}} \times \dot{H}^{\frac{n}{2}-1}$. This is done later, but for now we mention the embedding

$$S \subset C(\mathbb{R}; \dot{H}^{\frac{n}{2}}) \cap \dot{C}^1(\mathbb{R}; \dot{H}^{\frac{n-1}{2}}).$$

Expressed in terms of S, the above result includes:

- Existence: solutions exist in S.
- Uniqueness: solutions are unique in S.
- Continuous dependence: the map

  $$(\dot{H}^{\frac{n}{2}} \times \dot{H}^{\frac{n}{2}-1}) \cap (\dot{H}^{\frac{n}{2}-\delta} \times \dot{H}^{\frac{n}{2}-1-\delta}) \ni (\phi_0, \phi_1) \mapsto \phi \in S, \quad \delta > 0,$$

  is continuous.

- Regularity: If in addition the datum is in $\dot{H}^s \times \dot{H}^{s-1}$ for some $s > \frac{n}{2}$, then the solution stays uniformly bounded in the same norm.
- Scattering: after a suitable renormalization, the solutions approach a free wave at infinity.

The next question to ask is to what extent are the results in the small data case still valid for large data. One key difference in that regard occurs between the critical dimension $n = 2$ and supercritical dimensions $n \geq 3$. In two space dimensions the energy coincides with the critical Sobolev norm, and is a conserved quantity. In higher dimensions, on the other hand, there is no known mechanism to keep the critical Sobolev norm bounded; the energy is too weak for that purpose. Hence, if $n \geq 3$ it makes sense to try to study solutions for which an uniform a priori critical Sobolev bound is known.

An obstruction to having global scattering solutions comes from known solutions which either blow up, or do not decay as time goes to infinity. Such examples include:

- Self-similar solutions $\phi(t, x) = \phi\left(\frac{x}{t}\right)$ blow up in finite time; many examples are known if $n \geq 3$, but such solutions cannot exist and have finite energy if $n = 2$.
- Solitons (harmonic maps and their Lorentz transforms) do not blow up, but cause scattering to fail.
- Soliton-like concentration; this can indeed occur even if $n = 2$, and is discussed in Section 4.9.

On the positive side, we do have the finite speed of propagation: if blow up occurs, it has to happen via critical Sobolev norm concentration at the tip of a light cone. This severely limits the possible blow-up geometries.

We begin our discussion with the two-dimensional case, where the primary enemies for global solutions are the solitons, which correspond to harmonic maps. Then it is natural to introduce the following heuristic classification of target manifolds (M, g):

- No nonconstant harmonic maps $\Rightarrow$ **defocusing**, $E_{\text{crit}} = \infty$, e.g., $M = \mathbb{H}^m$.
- Nontrivial harmonic maps $\Rightarrow$ **focusing**, $E_{\text{crit}} < \infty$, e.g., $M = \mathbb{S}^m$.

In the defocusing case, one expects global well-posedness for large data. In the focusing case, global well-posedness should hold at least for data with energy below the ground state energy E_{crit}, i.e., the energy of the smallest nontrivial harmonic map. This has been known as the **Threshold Conjecture**, but is now a theorem:

Theorem 3.7 (Sterbenz and Tataru [37],[38]). *The following hold for the wave map equation in dimension $n = 2$:*

a) *In the defocusing case we have global well-posedness and scattering for large data in* $\dot{H}^1 \times L^2$.

b) *In the focusing case we have global well-posedness and scattering for all data in* $\dot{H}^1 \times L^2$ *below the ground state energy* E_{crit}.

The main ideas of the proof of this theorem are also presented in the next section. Prior to this, the same result was established in the equivariant case by Cote, Kenig, and Merle [13]. Independently, the case $M = \mathbb{H}^m$ was treated by Tao, see [43] and further references therein, and the case $M = \mathbb{H}^2$ was treated by Krieger and Schlag [23].

3.4 Schrödinger maps

The Schrödinger equation is closely related to the heat equation, and can be obtained by allowing complex-valued solutions for the heat equation and then extending those analytically in the half-space $\Re t \geq 0$. Restricting these solutions to the imaginary axis one obtains

$$(i\partial_t - \Delta)\phi = 0 \quad \text{or} \quad (i\partial_t - \partial_\alpha \partial_\alpha)\phi = 0, \quad \phi(0) = \phi_0.$$

The situation is slightly more complicated in the case of the Schrödinger maps. For that to make sense in the above context, we need a complex structure on the tangent space TM. Thus the natural setting is to have a Kähler manifold (M, g, J, ω) as a target. Even then, the Schrödinger map equation can no longer be obtained by taking a holomorphic extension of the harmonic heat flow in a half-space; indeed, the two flows no longer commute.

To introduce the Schrödinger map equation it is convenient to use the Hamiltonian formalism. In the case of the linear Schrödinger equation, the Hamiltonian is

$$H(\phi) = \frac{1}{2} \int_{\mathbb{R}^n} |\nabla \phi|^2 \, dx$$

and the symplectic form is

$$\underline{\omega}(u, v) = \Im \int_{\mathbb{R}^n} u \bar{v} \, dx.$$

For the Schrödinger map equation the Hamiltonian stays essentially unchanged,

$$H(\phi) = \frac{1}{2} \int_{\mathbb{R}^n} |\nabla \phi|_g^2 \, dx, \tag{3.9}$$

while the symplectic form becomes

$$\underline{\omega}(u, v) = \int_{\mathbb{R}^n} \langle u, Jv \rangle_g \, dx = \int_{\mathbb{R}^n} \omega(u, v) \, dx, \quad u, v \in E^\phi. \tag{3.10}$$

The associated Hamilton flow is the **Schrödinger map** equation

$$\phi_t = JD^\alpha \partial_\alpha \phi, \quad \phi(0) = \phi_0, \tag{3.11}$$

where J is the complex structure on TM.

The associated scaling law is the parabolic scaling

$$\phi(t,x) \to \phi(\lambda^2 t, \lambda x),$$

and the scale invariant space for the initial data is again $\dot{H}^{\frac{n}{2}}$.

While the above form of the equation is fairly general, most of the work so far has been done for special targets, namely the sphere $\mathbb{S}^2$ and the hyperbolic space $\mathbb{H}^2$. In the case of the sphere the form of the equation is

$$\partial_t \phi = \phi \times \Delta\phi,$$

where the cross product's purpose is twofold: to eliminate the component of $\Delta\phi$ which is normal to the sphere, and to rotate the remaining part by $\pi/2$. In the $\mathbb{H}^2$ case the equation looks identical except for a sign twist in the definition of the cross product.

The equation (3.11) admits one conserved quantity which is the counterpart of the usual energy functional for the linear Schrödinger equation:

$$E(\phi) = \frac{1}{2}\int_{\mathbb{R}^n} |\nabla\phi|_g^2 \, dx.$$

This is also the Hamiltonian; we use the terminology interchangeably.

In general there seems to be no direct counterpart of the conservation of mass and momentum; see however [15]. This can be related to the loss of the Galilean invariance.

The aim of the last section of these notes is to describe the proof of the small data result in critical Sobolev spaces:

Theorem 3.8 (Bejenaru, Ionescu, Kenig and Tataru [4]). *Consider the Schrödinger map equation with values into $\mathbb{S}^2$. Then global well-posedness holds for initial data which are small in the space $\dot{H}^{\frac{n}{2}}$.*

As for wave maps, this result includes existence, uniqueness, regularity, scattering, as well as continuous dependence on the initial data. The first result of this type was proved in [3] in high dimension $n \geq 4$ using the Coulomb gauge and suitable dispersive type estimates for the linear Schrödinger equation. The more difficult lower dimensional case $n = 2$ was proved in [4]. This requires a Schrödinger type counterpart[3] of the null frame spaces, as well as the caloric gauge. The corresponding result for the $\mathbb{H}^2$ target, though not explicitly spelled out in [4], follows by an almost identical argument.

The Schrödinger map counterpart of the large data problem result for wave maps in Theorem 3.7 is still open. However, we have the following partial result:

[3]Considerably simpler than for wave maps, though.

Theorem 3.9 (Bejenaru, Ionescu, Kenig and Tataru [1],[2]). *The following hold for the Schrödinger map equation in dimension $n = 2$ in the* 1-equivariant *class:*

a) *For the $\mathbb{H}^2$ target we have global well-posedness and scattering for all large data in the energy space $\dot{H}^1$.*

- [b)] *For the $\mathbb{S}^2$ target we have global well-posedness and scattering for all large data in the energy space $\dot{H}^1$ below the ground state energy.*

Chapter 4

Wave maps

4.1 Small data heuristics

Here we outline the main difficulties encountered in the study of the small data problem, and describe the ideas needed to overcome these difficulties. For simplicity we confine ourselves to the most interesting case of dimension two. Some simplifications arise in higher dimension, but the principles remain the same.

4.2 A perturbative set-up

In a first approximation, suppose that we are trying to view the wave map equation in the extrinsic formulation, namely

$$\Box \phi^i = -\mathcal{S}^i_{jk}(\phi)\partial^\alpha \phi^j \partial_\alpha \phi^k, \quad \phi(0) = \phi_0, \quad \partial_t \phi(0) = \phi_1, \tag{4.1}$$

as a small perturbation of the constant coefficient wave equation. This will not actually work, but it provides very useful insight. For this we would need two function spaces; one, call it S, for solutions, and a second, call it N, for the nonlinearity. For these spaces we would like to have two estimates:

a) a linear bound,

$$\|\phi\|_S \lesssim \|\phi[0]\|_{\dot{H}^1 \times L^2} + \|\Box \phi\|_N; \tag{4.2}$$

b) an estimate for the nonlinearity,

$$\|N(\phi)\|_N \lesssim \|\phi\|_S, \quad N(\phi) = \mathcal{S}(\phi)\partial^\alpha \phi \, \partial_\alpha \phi. \tag{4.3}$$

Further digesting the estimate for the nonlinearity, it would seem natural to break this into three parts:

b1) The algebra property for S.

H. Koch et al., *Dispersive Equations and Nonlinear Waves: Generalized Korteweg–de Vries, Nonlinear Schrödinger, Wave and Schrödinger Maps*, Oberwolfach Seminars 45,
DOI 10.1007/978-3-0348-0736-4_13, © Springer Basel 2014

b2) The null form bilinear estimate

$$\|\partial^\alpha \phi \, \partial_\alpha \phi\|_N \lesssim \|\phi\|_S^2. \tag{4.4}$$

b3) The product bound $S \cdot N \to N$.

4.2.1 The Strichartz norms

A key ingredient in the study of semilinear wave equations is the Strichartz estimates. Here we can easily incorporate the estimates in the structure of our function spaces by setting, in dimension $n = 2$,

$$S \subset |D|^{-1} L^\infty L^2 \cap |D|^{-\frac{1}{4}} L^4 L^\infty, \quad N \supset L^1 L^2 + |D|^{\frac{3}{4}} L^{\frac{4}{3}} L^1. \tag{4.5}$$

However, one sees that the Strichartz estimates cannot suffice to estimate the bilinear expression in (4.4). There are two reasons for that:

(i) The balance of the exponents. This is worst in two dimensions and improves as the dimension increases, up to the point where, in $5 + 1$ dimensions, it becomes favorable.

(ii) The balance of the derivatives. Because of the form of (4.4), one actually cannot use the full range of Strichartz exponents for each factor. This limitation is independent of the dimension.

Thus, by themselves, Strichartz estimates will not solve the problem. To remedy that, one needs to take advantage of the structure of the nonlinearity.

4.2.2 The null structure

We denote by τ the time Fourier variable and by ξ the space Fourier variable. We will refer to ξ as the frequency. An important role is played by the null cone $\tau^2 = \xi^2$, which is the characteristic set of $\Box$. The distance to the null cone, which has size $||\tau| - |\xi||$, will be referred to as modulation.

The symbol of the bilinear form $\partial^\alpha \phi \, \partial_\alpha \phi$ is $\tau s - \xi \eta$. As it is easy to see, this symbol vanishes if (τ, ξ) and (s, η) are parallel and located on the null cone. This is what we call the **null condition**. The geometric interpretation of this is that the nonlinear interaction of waves traveling in the same direction is killed in the nonlinearity, leaving the bulk of the nonlinear interaction to come from transversal waves. Heuristically that should be better behaved, because transversal waves have a short interaction time.

As the null condition depends on location of waves in the Fourier space, it cannot be handled via Strichartz estimates, which are invariant with respect to Fourier translations. Instead, one needs to take advantage of the $X^{s,b}$ type structure. The homogeneous $X^{s,b}$ spaces associated to the homogeneous wave equation are defined using the size of the Fourier transform:

$$\|u\|_{X^{s,b}} = \|\hat{u}(\tau, \xi) |\xi|^s ||\tau| - |\xi||^b\|_{L^2}.$$

Scaling considerations would dictate that we choose

$$S = X^{1,\frac{1}{2}}, \quad N = X^{0,-\frac{1}{2}}.$$

Unfortunately, this is just outside the range of indices for which these spaces are well defined.

To avoid the above difficulty one may use the $U^2_\Box$ and $V^2_\Box$ type spaces associated to the wave equation. These were first introduced in unpublished work of the author in connection to wave maps, and are described in detail elsewhere in these notes. They can be associated separately to each half wave and then combined using suitable multiplier. They are close to the above $X^{s,b}$ spaces, in the sense that

$$X^{1,\frac{1}{2},\infty} \subset V^2_\Box \dot{H}^1 \subset U^2_\Box \dot{H}^1 \subset X^{1,\frac{1}{2},1}, \tag{4.6}$$

where the third index in the $X^{s,b}$ notation is a Besov index with respect to modulation.

For the moment we neglect what happens far away from the null cone, which will turn out to be easier to deal with anyway. Then one would roughly have to choose

$$S \subset U^2_\Box \dot{H}^1, \quad N \supset DU^2_\Box L^2. \tag{4.7}$$

In view of Strichartz type embeddings associated to the U^2 and V^2 spaces, this is stronger than (4.5). With this choice we would have to prove a bound of the type

$$\|\partial^\alpha \phi^1 \partial_\alpha \phi^2\|_{DU^2_\Box L^2} \lesssim \|\phi^1\|_{U^2_\Box \dot{H}^1} \|\phi^2\|_{U^2_\Box \dot{H}^1}. \tag{4.8}$$

By the duality $(DU^2_\Box L^2)^* = V^2_\Box L^2$, this becomes

$$\left| \int \partial^\alpha \phi^1 \partial_\alpha \phi^2 \phi^3 \; dxdt \right| \lesssim \|\phi^1\|_{U^2_\Box \dot{H}^1} \|\phi^2\|_{U^2_\Box \dot{H}^1} \|\phi^3\|_{V^2_\Box L^2}. \tag{4.9}$$

To test this theory, we consider the usual Littlewood–Paley trichotomy. In order to be able to work with U^2 atoms, we also neglect for now the difference between V^2L^2 and U^2L^2. Then we can prove the following sharp dyadic estimate:

Lemma 4.1. *Assume that $j \leq k$. Then the following dyadic estimates hold:*

$$\left| \int \partial^\alpha \phi^1_k \partial_\alpha \phi^2_j \phi^3_k \; dxdt \right| \lesssim 2^{j+k} \|\phi^1_k\|_{U^2_\Box L^2} \|\phi^2_j\|_{U^2_\Box L^2} \|\phi^3_k\|_{U^2_\Box L^2}, \tag{4.10}$$

respectively

$$\left| \int \partial^\alpha \phi^1_k \partial_\alpha \phi^2_k \phi^3_j dxdt \right| \lesssim 2^{\frac{k+3j}{2}} \|\phi^1_k\|_{U^2_\Box L^2} \|\phi^2_k\|_{U^2_\Box L^2} \|\phi^3_j\|_{U^2_\Box L^2}. \tag{4.11}$$

Proof. The proof of the lemma is fairly simple. First of all, it suffices to prove the result for U^2 atoms. Secondly, by considering the nesting of the steps in each atom, one sees that it suffices to assume that two of the three atoms are free waves. Remembering the relation between U^2 and $X^{s,b}$ spaces, we are left with having to prove bilinear L^2 estimates for free waves. We need to consider two cases, depending on the frequency balance of the two free waves:

a) high × low free wave interactions. Denoting by 2^k, respectively 2^j the size of two frequencies, we will prove the estimate

$$\|\partial^\alpha \phi_j \partial_\alpha \phi_k\|_{X^{0,-\frac{3}{4}}} \lesssim 2^{k+\frac{3}{4}j} \|\phi_k(0)\|_{L^2} \|\phi_j(0)\|_{L^2}, \tag{4.12}$$

where the output modulation is at most 2^j. Let ξ, respectively η be the frequencies for the two inputs. The output frequency $\xi + \eta$ will have size 2^k, but we also need to compute its distance d from the null cone. This distance turns out to be related to the angle θ between ξ and η. Precisely, we have

$$2^k d \approx |(\tau + s, \xi + \eta)|_m^2 = 2\langle(\tau, \xi), (s, \eta)\rangle_m \approx \pm 2^{k+j}\theta^2,$$

where the sign depends on the relative orientation of the two input cones. Fixing the angle θ between the two waves we can reduce the problem to the following L^2 estimate for two free waves at angle $\theta \in [0, 1]$:

$$\|\phi_k \phi_j\|_{L^2} \lesssim \theta^{-\frac{1}{2}} 2^{\frac{j}{2}} \|\phi_k(0)\|_{L^2} \|\phi_j(0)\|_{L^2}. \tag{4.13}$$

This estimate no longer has anything to do with the curvature of the cone, instead it is based on the *transversality* of the two sectors of the cone. Thus it follows by general principles (see the exposition in [46], though such estimates had been known before, e.g. [20], [8]) since the angle of the two cone sections is θ and the size of the intersection of two translates of them is 2^j.

From here one arrives to (4.12) by adding the size of the symbol of the null form $\tau s - \xi\eta \approx \pm 2^{k+j}\theta^2$. There is an additional orthogonality argument which is needed in order to gain the square summability with respect to θ, but we skip it since it plays no role in the sequel.

b) high × high free wave interactions. Denoting by 2^k the size of two input frequencies , and by 2^j the size of the output frequency, we will prove the estimate

$$\|P_j(\partial^\alpha \phi_k^1 \partial_\alpha \phi_k^2)\|_{X^{0,-\frac{3}{4}}} \lesssim 2^{\frac{1}{2}k+\frac{5}{4}j} \|\phi_k^1(0)\|_{L^2} \|\phi_k^2(0)\|_{L^2}, \tag{4.14}$$

where the output modulation is at most 2^j. As before let ξ, respectively η be the Fourier variables for the two inputs. The output frequency $\xi + \eta$ is restricted to a 2^j cube, so by orthogonality we can also restrict ξ and η to 2^j cubes.

This time the distance of $\xi + \eta$ from the null cone is related to the angle θ between ξ and η by the relation

$$2^j d \approx \pm 2^{2k}\theta^2,$$

where the sign depends on the relative orientation of the two input cones. Fixing the angle θ between the two waves we can reduce the problem to the following L^2 estimate for two free waves localized in 2^j cubes at frequency 2^k and at angle θ:

$$\|\phi_k^1 \phi_k^2\|_{L^2} \lesssim \theta^{-\frac{1}{2}} 2^{\frac{j}{2}} \|\phi_k^1(0)\|_{L^2} \|\phi_k^2(0)\|_{L^2}. \tag{4.15}$$

This is again a transversality estimate which follows by general principles. From here (4.14) is obtained by adding the size of the symbol of the null form $2^{2k}\theta^2$.

□

Compare the needed bound (4.8) with what is actually proved in Lemma 4.1. On the positive side, we have

- extra gains in the high $\times$ high $\to$ low interactions
- extra gains at small interaction angles.

On the negative side, we have

- possible losses in the transition from U^2 to V^2 in (4.9);
- lack of dyadic summation with respect to low frequencies in low $\times$ high $\to$ high interactions.

Both of these difficulties are nontrivial, and will be successively discussed in what follows.

4.2.3 The null frame spaces

As mentioned above, one of the difficulties in the direct approach above is the need to transition from V^2 to U^2 spaces in bilinear estimates. This venue was initially pursued by the author, and, on the positive side, it led to the introduction of the U^p and V^p type spaces to the field of dispersive equations. Unfortunately, this attempt was not entirely successful, and a more radical reworking of the function spaces S and N was eventually introduced in [47]. We remark that at this point we do have a well established mechanism for transitioning from V^2 to U^2 spaces in estimates, see [18]. However, this transition entails logarithmic frequency losses of one type or another, which seem to be too much for this particular problem.

Backtracking to the proof of the estimates (4.13) and (4.15), the key idea is that one would like to have a version of that which also applies to inhomogeneous waves. We focus on the first bound, and revisit its proof. Rather than thinking of it as a convolution of two surface carried distributions in the Fourier space, of the form, say,

$$\|f_j(\xi)\delta_{\tau=\pm|\xi|} * f_k(\xi)\delta_{\tau=\pm|\xi|}\|_{L^2} \lesssim \theta^{-\frac{1}{2}} 2^{\frac{j}{2}} \|f_j\|_{L^2} \|f_k\|_{L^2}, \tag{4.16}$$

where $\hat{\phi}_j = f_j(\xi)\delta_{\tau=\pm|\xi|}$ and $\hat{\phi}_k = f_k(\xi)\delta_{\tau=\pm|\xi|}$, we instead take advantage of the extra dimension that we have available to foliate the frequency μ waves with respect to null rays in frequency,

$$f_j(\xi)\delta_{\tau=\pm|\xi|} = \int_\omega f_j^\omega \, d\omega, \quad f_j^\omega = f_j(r\omega)\delta_{\tau=\pm|\xi|}\delta_{\xi=r\omega}.$$

For each f_j^ω we have the bilinear estimate

$$\|f_j^\omega * f_k(\xi)\delta_{\tau=\pm|\xi|}\|_{L^2} \lesssim \theta^{-1}\|f_\mu(\omega r)\|_{L^2_r}\|f_\lambda\|_{L^2}, \tag{4.17}$$

simply due to the fact that the incidence angle is θ^2 (compare this with the angle θ of the two surfaces!). Then (4.16) follows easily from (4.17) by Cauchy–Schwartz with respect to ω after also accounting for the change in the surface measure.

So far all we have is an alternate proof of (4.13). The key observation now is that we can rework the proof of (4.17) in terms of mixed L^p norms as follows. If $\phi_j^\omega = \widehat{f_j^\omega}$, then by Plancherel we have the estimate

$$\|\phi_j^\omega\|_{L^2_\gamma L^\infty_{\gamma^\perp}} = \|f_j(\omega r)\|_{L^2_r}, \quad \gamma = (\omega, \pm|\omega|).$$

On the other hand, using the fact that ω is at angle θ from the support of f_k, we also have the characteristic energy estimate

$$\|\phi_\lambda\|_{L^\infty_\gamma L^2_{\gamma^\perp}} \approx \theta^{-1}\|f_\lambda\|_{L^2}.$$

Then (4.17) follows from the last two relations. This suggests that the space S should include, beside the standard Strichartz norm and the U^2 structure, the following two components associated to null frames:

- characteristic energy norms $\cap_\omega L^\infty_\gamma L^2_{\gamma^\perp}$;
- foliated norms $\sum_\omega L^2_\gamma L^\infty_{\gamma^\perp}$.

By duality considerations, the space N also needs to include

- dual characteristic energy norms $\sum_\omega L^1_\gamma L^2_{\gamma^\perp}$.

Fortunately, the second set of dual spaces $\bigcap_\omega L^2_\gamma L^1_{\gamma^\perp}$ turns out not to be needed.

Both of these have to be introduced carefully, with suitable frequency, modulation, and angular localizations. An additional difficulty occurs when applying this idea to *high* $\times$ *high* interactions, where one needs to either allow for radial frequency localizations below the frequency scale, or to admit some losses in the interaction angle or in the high-low frequency balance in the estimates. Fortunately this is not a crucial issue, since there is sufficient room there to allow for some flexibility.

4.2.4 The paradifferential equation and renormalization

Suppose now that we have good function spaces S and N for which the dyadic versions of the null form estimates hold:

$$\|\partial^\alpha \phi_j \partial_\alpha \phi_k\|_N \lesssim \|\phi_j\|_S \|\phi_k\|_S, \quad j < k, \tag{4.18}$$

$$\|P_j(\partial^\alpha \phi_k \partial_\alpha \phi_k)\|_N \lesssim 2^{-\delta|j-k|} \|\phi_k\|_{S^1} \|\phi_k\|_{S^1}, \quad j \lesssim k. \tag{4.19}$$

While the second has some extra room, the first one is tight, and does not allow for a favorable j summation, since the norms of ϕ_j are only square summable. This suggests that the nonlinearity in the wave map equation is actually nonperturbative. If that is the case, then the next best thing to do is to understand exactly what is the nonperturbative part. That immediately leads to the paradifferential formulation of the problem, namely

$$\Box \phi_k^i = -2\mathcal{S}_{jl}^i(\phi)_{<k} \partial^\alpha \phi_{<k}^j \partial_\alpha \phi_k^l + \text{perturbative}(N)$$

One advantage in doing this is that now we only need to study a linear equation, where the coefficients have lower frequency. The above equation is closely linked to the linearized wave map equation; indeed, it largely represents a high frequency linearized wave evolving on a low frequency background.

A generic equation of the above form does not seem to have enough structure to allow for good linear estimates. However, so far we have not used at all the geometry of the problem. To take advantage of that we begin with the orthogonality relation

$$\mathcal{S}_{ji}^l(\phi)\partial_\alpha \phi^l = 0.$$

Transitioning to the paradifferential form of this and combining it with the previous paradifferential equation we arrive at a more favorable equation,

$$\Box \phi_k = -2A_i(\phi)_{<k} \partial^\alpha \phi_{<k}^i \partial_\alpha \phi_k + \text{perturbative}(N), \tag{4.20}$$

where the matrices $(A_i)_l^j = \mathcal{S}_{il}^j - \mathcal{S}_{ij}^l$ are antisymmetric. This antisymetry adds some conservation structure to the paradifferential equation; this is closely linked to the question of getting good energy estimates for solutions to (4.20).

Tao [45]'s approach to the above equation in the $\mathbb{S}^2$ case was to develop a renormalization procedure which transforms the nonlinearity into a a perturbative nonlinearity in the context of the null frame spaces. This is reminiscent of Heléin's work on harmonic maps, and is achieved in a multiplicative way in the paradifferential setting. Precisely, one seeks a linear transformation

$$w_k = \mathcal{O}_{<k} \psi_k$$

which transforms the previous equation into the flat wave equation

$$\Box w_k = \text{perturbative}(N). \tag{4.21}$$

In the context of the frame method introduced earlier, this corresponds to studying high frequency solutions to the linearized wave map equation, represented in a favorable frame in the tangent space TM.

Substituting into the equation and neglecting some lower order terms, one sees that this works provided that the (orthogonal or almost orthogonal) matrix-valued function $\mathcal{O}_{<k}$ is a reasonably good approximate solution for the system of equations

$$\partial_\alpha \mathcal{O}_{<k} = \mathcal{O}_{<k} \partial_\alpha \phi^i_{<k}.$$

The construction of such a renormalization matrix $\mathcal{O}$ is a key idea of Tao [45]. This construction was further refined and simplified in Tataru [48] and later in Sterbenz and Tataru [37]. One choice that needs to be made here is between the frequency localization and the orthogonality of $O_{<k}$; both are desirable, but seem mutually exclusive. Frequency localization is easier to work with and was the preferred choice in the small data problem in [45], [48]. However, for large data the orthogonality losses become unmanageable, and instead one must sacrifice frequency localization, see [37].

An alternate approach, based on the frame method with the Coulomb gauge, was developed by Krieger [22] for the case of an $\mathbb{H}^2$ target.

4.3 Function spaces

Here we define the function spaces S and N, following Sterbenz and Tataru [37]. The space N is essentially as originally introduced in Tataru [47]; there the space $\Box^{-1} N$ was used in place of S, along with the key embedding $\Box^{-1} N \subset S$. Tao [45] observed that using S instead of $\Box^{-1} N$ as the main function space helps with the algebra type properties. Tao's version of S was then strengthened to some extent in Sterbenz and Tataru [37]. A related but somewhat different modification of S was proposed by Krieger [22].

We recall that P_k denote Littlewood-Paley localization with respect to the spatial frequency. For modulation localizations we use the space-time multipliers Q_j with symbol

$$q_j(\tau, \xi) = \varphi\big(2^{-j}\big||\tau| - |\xi|\big|\big),$$

where φ truncates smoothly on a unit annulus. We denote by $Q_j^{\pm}$ the restriction of this multiplier to the upper or lower time frequency space.

Beside the frequency and modulation decompositions, we also need to deal with the angular decompositions which are needed for the proof of the bilinear estimates. We denote by $\kappa \in K_l$ a collection of caps of diameter $\sim 2^{-l}$ providing a finitely overlapping cover of the unit sphere. According to this decomposition, we cut up the spatial frequency domain according to

$$P_k = \sum_{\kappa \in K_l} P_{k,\kappa}.$$

These decompositions usually occur in conjunction with modulation cutoffs up to 2^j, where $j = k - 2l$. This is related to the discussion in Section 4.2.2; another interpretation of this scale choice is that it corresponds to the thinnest angular slabs of angle 2^{-l} on the null cone which are well approximated by a parallelipiped, i.e., have no curvature.

For each integer k we define the following frequency localized norm:

$$\| \phi \|_{S_k} := \| \nabla_{t,x}\phi_k \|_{L_t^\infty(L_x^2)} + \| \nabla_{t,x}\phi_k \|_{X_\infty^{0,\frac{1}{2}}} + \| \phi_k \|_{\underline{S}} + \sup_{j<k-20} \| \phi \|_{S[k;j]}, \quad (4.22)$$

with components as follows:

- The fixed frequency space $X_p^{s,b}$ is defined as
$$\| P_k\phi \|_{X_p^{s,b}}^p := 2^{psk} \sum_j 2^{pbj} \| Q_j P_k \phi \|_{L_t^2(L_x^2)}^p,$$
with the obvious definition for $X_\infty^{s,b}$.
- The "physical space Strichartz" norms are given by
$$\| \phi_k \|_{\underline{S}} := \sup_{(q,r):\ \frac{2}{q}+\frac{1}{r}\leqslant\frac{1}{2}} 2^{(\frac{1}{q}+\frac{2}{r}-1)k} \| \nabla_{t,x}\phi_k \|_{L_t^q(L_x^r)}. \quad (4.23)$$
- The "modulational Strichartz" norms are
$$\| \phi \|_{S[k;j]} := \sup_{\pm} \Big(\sum_{\kappa\in K_l} \| Q^{\pm}_{<k-2l} P_{k,\pm\kappa}\phi \|_{S[k,\kappa]}^2 \Big)^{\frac{1}{2}}, \quad l = \frac{k-j}{2} > 10. \quad (4.24)$$
- The "angular Strichartz" space is defined in terms of the three components:
$$\| \phi \|_{S[k,\kappa]} := 2^k \sup_{\omega\notin 2\kappa} \mathrm{dist}(\omega,\kappa) \| \phi \|_{L_{t_\omega}^\infty(L_{w_\omega}^2)} + 2^k \| \phi \|_{L_t^\infty(L_x^2)}$$
$$+ 2^{\frac{1}{2}k} |\kappa|^{-\frac{1}{2}} \inf_{\sum_\omega \phi^\omega = \phi} \sum_\omega \| \phi^\omega \|_{L_{t_\omega}^2(L_{x_\omega}^\infty)}. \quad (4.25)$$

The first component on the RHS above will often be referred to as NFA^*.

We define S as the space of functions ϕ in $\mathbb{R}^{2+1}$ with $\nabla_{x,t}\phi \in C(\mathbb{R}; L_x^2)$ and finite norm
$$\| \phi \|_S^2 = \| \phi \|_{L_t^\infty(L_x^\infty)}^2 + \sum_k \| \phi \|_{S_k}^2.$$

Two other norms related to S play an auxiliary role in the study of the large data problem, namely

- The null frame energy:
$$\| \phi \|_{\underline{E}} := \| \nabla_{t,x}\phi \|_{L_t^\infty(L_x^2)} + \sup_\omega \| \not{\nabla}^\omega_{t,x}\phi \|_{L_{t_\omega}^\infty(L_{x_\omega}^2)}. \quad (4.26)$$

- The high modulation L^2 norm:

$$\| \phi \|_{\underline{X}_k} := 2^{-\frac{1}{2}k} \| \Box P_k \phi \|_{L_t^2(L_x^2)}. \tag{4.27}$$

We also define $\underline{X}$ as the square sum of $\underline{X}_k$. Notice that there are no square sums or frequency localizations in the norm $\underline{E}$. This makes proving $\underline{E}$ bounds amenable to energy estimates techniques, bypassing the more difficult bilinear and multilinear estimates. The $\underline{X}$ bounds are also easier to obtain and provide stronger high modulation bounds than what is included in the S norm.

In the same manner as in the case of the S space, for each integer k we define the dyadic versions of the N norm by

$$\begin{aligned}\| F \|_{N_k} := \inf_{F_A + F_B + \sum_{l,\kappa} F_C^{l,\kappa} = F} \Big(& \| P_k F_A \|_{L_t^1(L_x^2)} + \| P_k F_B \|_{X_1^{0,-\frac{1}{2}}} \\ & + \sum_{\pm} \sum_{l>10} \Big(\sum_{\kappa} \inf_{\omega \notin 2\kappa} \operatorname{dist}(\omega,\kappa)^{-2} \| Q^{\pm}_{<k-2l} P_{k,\pm\kappa} F_C^{l,\kappa} \|^2_{L^1_{t_\omega}(L^2_{x_\omega})} \Big)^{\frac{1}{2}} \Big).\end{aligned} \tag{4.28}$$

We will often refer to the last component on the RHS above as *NFA*, and the norm applied to a fixed $Q^{\pm} F_C^{l,\kappa}$ as *NFA*$[\pm\kappa]$.

The full N norm is

$$\|F\|_N^2 = \sum_k \|P_k F\|^2_{N_k}.$$

All of these spaces have versions which are restricted to time intervals I, denoted, e.g., by $S[I]$, respectively $N[I]$. Since the interval truncation does not commute with the time Fourier transform, some minor technical issues arise in the process. These are skipped here.

4.3.1 Frequency envelopes

In many places of the subsequent analysis involving the S spaces it pays to keep a more careful track of how much of the S norm of wave maps is concentrated at various frequencies. This is conveniently expressed in the language of frequency envelopes.

A sequence c_k is called a frequency envelope for ϕ in S if the following three requirements are satisfied:

- Norm control:

$$\|\phi_k\|_S \le c_k.$$

- Norm equivalence:

$$\sum c_k^2 \approx \sum \|\phi_k\|_S^2.$$

- Slowly varying:
$$|c_j/c_k| \lesssim 2^{\delta|j-k|}$$
for a fixed small universal constant δ.

A similar terminology is used with respect to all of the other norms in the paper, e.g., the initial data space $\dot{H}^1 \times L^2$, the space N, etc.

4.3.2 Linear analysis in the S and N spaces

The linear component of our estimates has the form

Proposition 4.2. *The following estimate holds for functions which are localized at frequency 2^k:*
$$\|\phi_k\|_{S_k} \lesssim \|\phi_k[0]\|_{\dot{H}^1\times L^2} + \|\Box\phi_k\|_{N_k}. \tag{4.29}$$

Outline of the proof. The proof is relatively straightforward when interpreted in terms of the U^2 norms. Set $F = \Box\phi_k$. With notations as in the above definition of the N_k norm, consider first the case when $F = F_A + F_B$. By Strichartz type embeddings and the dual to (4.6) it is fairly easy to see that $F \in DU^2_\Box L^2$, therefore the corresponding solution ϕ_k belongs to $U^2_\Box \dot{H}^1$, so it remains to show that $U^2_\Box \dot{H}^1 \subset S$. The first and third components of the S norm are easy to estimate via bounds for free waves and then for atoms. The third component of the S norm is bounded by (4.6). It remains to consider the $S[k,j]$ norms. The U^2 space is well behaved with respect to frequency and modulation localizations:

$$\sum_{\kappa\in K_l} \| Q^{\pm}_{<k-2l} P_{k,\pm\kappa}\phi_k \|^2_{U^2_\Box \dot{H}^1} \lesssim \|\phi_k\|^2_{U^2_\Box \dot{H}^1},$$

so it remains to estimate the $S[k,\kappa]$ norm for each localized piece. But this is easily done again by starting with the known bounds for free solutions, which are then transferred to U^2H^1 atoms.

Lastly, consider the case when $F = \sum_{l,\kappa} F_C^{l,\kappa}$. On one hand, we can place $F_C^{l,\kappa}$ in $DV^2_\Box L^2$, which follows by duality from the embedding of $U^2_\Box \dot{H}^1$ into the NFA^* component of the S_k space. On the other hand, we can place $F_C^{l,\kappa}$ into $DU^2_\omega L^2$, which is the U^2 space corresponding to the wave evolution in the null direction associated to ω.

Thus, denoting by $\phi = \sum \phi^{l,\kappa}$, where $\phi^{l,\kappa}$ is the solution to

$$\Box\phi^{l,\kappa} = F_C^{l,\kappa}, \quad \phi^{l,\kappa}[0] = 0,$$

we can first bound $\phi^{l,\kappa}$ in $V^2_\Box \dot{H}^1$. By frequency orthogonality, this leads to a $V^2_\Box \dot{H}^1$ for ϕ, and this suffices for the first three components of the S_k norm.

Secondly, we can bound $\phi^{l,\kappa}$ in $U^2_\omega \dot{H}^1$ with $\omega = \omega(\kappa)$. Now to estimate the $S[k,j]$ norm of ϕ we consider two cases:

(i) $l > l' = \frac{j-k}{2}$. Then each $\phi^{l,\kappa}$ is measured with respect to a collection of $S[k,\kappa']$ norms with $\kappa' \in K_{l'}$. We can argue separately for each ω that $U^2_\omega \dot{H}^1 \subset S[k,j]$ and then use the square summability with respect to κ to sum up the results.

(ii) $l < l' = \frac{j-k}{2}$. Then each corresponding $S[k,\kappa']$ norm applies to a collection of $\phi^{l,\kappa}$. For the first two components of the $S[k,\kappa']$ norm, we estimate them directly for each $\phi^{l,\kappa}$, and then use L^2 orthogonality based on the frequency/modulation localization to add them up. For the last component of the $S[k,\kappa']$ norm, we simply sum up the bounds for each $\phi^{l,\kappa}$; orthogonality does not hold, but it is also not needed. $\square$

4.3.3 Multilinear estimates

For the nonlinear side of our problem we need not only the bilinear null form estimate described earlier, but also additional bounds which account for the role of the $\mathcal{S}(\phi)$ factor. To start with, we have:

Proposition 4.3. *The following bilinear and trilinear estimates hold for the S and N spaces:*

- *Product estimates:*

$$\| \phi^{(1)}_{<k+O(1)} \cdot \phi^{(2)}_k \|_S \lesssim \| \phi^{(1)}_{<k+O(1)} \|_S \cdot \| \phi^{(2)}_k \|_S, \tag{4.30}$$

$$\| P_k(\phi^{(1)}_{k_1} \cdot \phi^{(2)}_{k_2}) \|_S \lesssim 2^{-(\max\{k_i\}-k)} \| \phi^{(1)}_{k_1} \|_S \cdot \| \phi^{(2)}_{k_2} \|_S, \tag{4.31}$$

$$\| P_k(\phi_{<k+O(1)} \cdot F_k) \|_N \lesssim \| \phi \|_S \cdot \| F_k \|_N, \tag{4.32}$$

$$\| P_k(\phi_{k_1} \cdot F_{k_2}) \|_N \lesssim 2^{-\delta(k-k_2)_+} \| \phi_{k_1} \|_S \cdot \| F_{k_2} \|_N. \tag{4.33}$$

- *Bilinear Null Form Estimates:*

$$\|P_k\big(\partial^\alpha \phi^{(1)}_{k_1} \cdot \partial_\alpha \phi^{(2)}_{k_2}\big)\|_{L^2_t(L^2_x)} \lesssim 2^{\frac{1}{2}\min\{k_i\}} 2^{-(\frac{1}{2}+\delta)(\max\{k_i\}-k)} \prod_i \|\phi^{(i)}_{k_i}\|_S \ , \tag{4.34}$$

$$\| P_k(\partial^\alpha \phi^{(1)}_{k_1} \cdot \partial_\alpha \phi^{(2)}_{k_2}) \|_N \lesssim 2^{-\delta(\max\{k_i\}-k)} \prod_i \| \phi^{(i)}_{k_i} \|_S \ . \tag{4.35}$$

- *Trilinear Null Form Estimate:*

$$\| P_k(\phi^{(1)}_{k_1} \cdot \partial^\alpha \phi^{(2)}_{k_2} \cdot \partial_\alpha \phi^{(3)}_{k_3}) \|_N \lesssim 2^{-\delta(\max\{k_i\}-k)} 2^{-\delta(k_1-\min\{k_2,k_3\})_+} \prod_i \| \phi^{(i)}_{k_i} \|_S. \tag{4.36}$$

The bilinear estimates are essentially the dyadic counterparts of the bounds discussed in Section 4.2.2. The last trilinear estimate provides a key improvement over the composition of bilinear bounds, which plays a major role in the renormalization procedure in Section 4.2.4. The proofs follow largely from the principles discussed in Section 4.2.3, and are omitted; instead we refer the reader to [47], [45] and [37]. As a consequence of the above results we have

Proposition 4.4. a) *The space S is an algebra, and the following Moser type estimates hold for any bounded function G with uniformly bounded derivatives:*

$$\| G(\phi) \|_S \lesssim \| \phi \|_S (1 + \| \phi \|_S^3) \ , \tag{4.37}$$

In addition, if c_k is a frequency envelope for ϕ, then

$$\| G(\phi)_k \|_S \ \lesssim \ (1 + \| \phi \|_S^3) c_k.$$

b) *The product estimate $S \times N \to N$ holds.*

Outline of proof. The nontrivial part of the proposition is the Moser estimate. For that, following [48], we use multilinear paradifferential decompositions. For $h \in \mathbb{R}$ we can write

$$\frac{d}{dh} F(\phi_{<h}) = \phi_h F'(\phi_{<h})$$

or in integral form

$$F(\phi) = F(\phi_{<l}) + \int_l^\infty \phi_h F'(\phi_{<h}) dh. \tag{4.38}$$

This suffices for energy estimates, but not for estimates in the S type spaces. Hence we iterate this computation to obtain

$$\begin{aligned} F(\phi) = F(\phi_{<l}) + F'(\phi_{<l}) \int_l^\infty \chi(h) \phi_h dh + F''(\phi_{<l}) \int_{[l,\infty)^2} \chi(h) \phi_{h_0} \phi_{h_1} dh \\ + \int_{[l,\infty)^3} \chi(h) \phi_{h_0} \phi_{h_1} \phi_{h_2} \ F^{'''}(\phi_{<h_2}) dh \end{aligned} \tag{4.39}$$

where by $\chi(h)$ we denote the ordering function

$$\chi(h) = 1_{h_j \leq h_{j-1} \leq \cdots \leq h_0}.$$

This expansion allows us to successively build estimates for $F(\phi_{<l})$ as follows:

(i) First, by direct differentiation, we have

$$\|\nabla F(\phi_{<k})\|_{L^\infty} \lesssim 2^k, \quad \|\nabla F(\phi_{<k})\|_{L^\infty L^2} \lesssim 1.$$

(ii) Next, repeated differentiation followed by Littlewood-Paley projections yields the high frequency decay

$$\|P_j F(\phi_{<k})\|_{L^\infty} \lesssim 2^{N(k-j)}, \quad \|P_j F(\phi_{<k})\|_{L^\infty L^2} \lesssim 2^{-k+N(k-j)}, \quad j > k.$$

(iii) Applying (4.39) to $\phi_{<k}$ and letting $l \to -\infty$ the first term drops, and using the Strichartz estimates[1] for ϕ and the bounds in the previous steps we obtain the better high frequency decay

$$\|P_j F(\phi_{<k})\|_{L^2} \lesssim 2^{-\frac{3k}{2}+N(k-j)}, \quad j > k+10.$$

For all practical purposes this allows us to assume that $F(\phi_{<k})$ is localized at frequency $\lesssim 2^k$; the contributions of higher frequencies are easier to estimate.

(iv) To estimate a component of the S_k norm of $F(\phi)$ which involves a modulation truncation at modulation $2^j < 2^k$, we apply (4.39) to ϕ with $l = j-10$. The factors $F'(\phi_{<l})$ and $F''(\phi_{<l})$ are bounded, so they preserve all mixed L^p norm, without affecting the frequency localization (except for better behaved tails). In the last term in (4.39) we have the higher frequency factor $F'''(\phi_{<h_2})$. However, this is combined with a trilinear expression $\phi_{h_0}\phi_{h_1}\phi_{h_2}$ which by Strichartz and multilinear S estimates has an L^2 structure on the 2^{h_2} frequency scale; hence, it again suffices to use the L^∞ bound for $F'''(\phi_{<h_2})$. □

4.4 Renormalization

The idea behind the renormalization is to consider a linear paradifferential equation of the type

$$(\Box + 2A_i(\phi)_{<k-m}\partial^\alpha \phi^i_{<k-m}\partial_\alpha)\psi_k = F_k$$

with antisymmetric A_i's, and to obtain estimates of the type

$$\|\psi_k\|_S \lesssim \|\psi_k[0]\|_{\dot H^1 \times L^2} + \|F_k\|_N. \tag{4.40}$$

Here m is a large parameter which depends on the S size of ϕ in the coefficients; it is essential in the large data problem, but it plays no role for small data.

The strategy is to use a renormalization matrix $O_{<k-m}$ to perform a change of variable $w_k = O_{<k-m}\psi_k$ so that the equation for w_k is

$$\Box w_k = O_{<k-m}F_k + \mathrm{error}(N).$$

To motivate the choice of O we compute the above error,

$$\begin{aligned}\mathrm{error} = &\left(\Box O_{<k-m} - O_{<k-m}(\Box + 2A_i(\phi)_{<k-m}\partial^\alpha\phi^i_{<k-m}\partial_\alpha)\right)\phi_k\\ &\Box O_{<k-m}\phi_k + 2(\partial^\alpha O_{<k-m} - O_{<k-m}A_i(\phi)_{<k-m}\partial^\alpha\phi^i_{<k-m})\partial_\alpha\phi_k.\end{aligned}$$

The first term in the error is in some[2] sense better behaved because both derivatives apply to the lower frequency factor. In the second term, in view of the trilinear

[1] It takes exactly three Strichartz estimates to place a product in L^2.

[2] This still has to be proved once $O_{<k}$ is constructed, and it it not entirely straightforward.

estimate (4.36), we can neglect the terms where the frequency of A_i is comparable to or larger than the frequency of ϕ^i. Hence, defining

$$B_k = A_i(\phi)_{<k-C}\phi^i_k,$$

a reasonable choice would be to select O_k so that

$$O_k = O_{<k}B_k. \tag{4.41}$$

Since O_k is continuously interpreted as $\partial_k O_{<k}$, it follows that $O_{<k}$ is defined as the solution to the ode

$$\frac{d}{dk}O_{<k} = O_{<k}B_k, \qquad \lim_{k\to-\infty} O_{<k} = I_m. \tag{4.42}$$

Defining O_k as such has one key advantage, namely that the antisymmetry of B_k insures that $O_{<k}$ remains an orthogonal matrix, and provides good L^p type bounds for its derivatives. There is also a significant disadvantage, namely that the frequency localization is lost; fortunately, the frequency tails turn out to decrease rapidly. The bounds for $O_{<k}$ are summarized as follows:

Proposition 4.5. *Let ϕ be a wave-map with energy E, S norm F, and S frequency envelope $\{c_k\}$. Then the orthogonal matrix $O_{<k}$ defined above and its k derivative O_k have the following properties:*

- (S_k *bounds for* O_k) *Each O_k obeys the bounds:*

$$\| P_{k'}O_k \|_S \lesssim_F 2^{-\delta|k-k'|}2^{-C(k'-k)_+}c_k, \tag{4.43}$$

$$\|P_{k'}\nabla^J_{t,x}O_k\|_{L^1_t(L^1_x)} \lesssim_F 2^{(|J|-3)k}2^{-C(k'-k)}c_k, \quad k' > k+10, \quad |J| \leqslant 2, \tag{4.44}$$

$$\| P_{k'}\big(O_{<k-20}\cdot G_k\big) \|_N \lesssim_F 2^{-|k'-k|}\| G_k \|_N, \tag{4.45}$$

$$\| P_k\big(\Box O_{k_1}\cdot \psi_{k_2}\big) \|_N \lesssim_F 2^{-|k-k_2|}2^{-\delta(k_2-k_1)}c_{k_1}\| \psi_{k_2} \|_S, \quad k_1 < k_2 - 10. \tag{4.46}$$

- (*The matrix O approximately renormalizes $A_\alpha = \nabla_\alpha B$*) *We have the formula:*

$$O^\dagger_{<k}\nabla_\alpha O_{<k} = \nabla_\alpha B_{<k} - \int_{-\infty}^{k} \big[B_{k'}, O^\dagger_{<k'}\nabla_\alpha O_{<k'}\big]\, dk'. \tag{4.47}$$

Proof. The main difficulty in the proof is that, since B_k are not small, it is not possible to directly bootstrap the estimates for O_k. Instead, the proof is by direct arguments, iterating separately the various components of the S norm, in the following order:

- L^∞ and $L^\infty L^2$ bounds,

- Strichartz bounds,
- High modulation bounds (i.e., L^2 bounds for $\Box U_k$),
- High frequency bounds (i.e., the estimate (4.44)),
- S norm bounds.

Here each step is carried out based on the previous steps, without bootstrapping. The most difficult part, i.e., the S bound, is obtained by using iterated expansions akin to the proof of the Moser estimates. For further details we refer the reader to [37]. $\square$

The main use of the renormalization matrix $O_{<k}$ is in the proof of the $N \to S$ estimates for the paradifferential equation:

Proposition 4.6 (Gauge Covariant S Estimate)**.** *Let $\psi_k = P_k\psi$ be a solution to the linear problem:*

$$\Box \psi_k = -2A^{\alpha}_{<k-m}\partial_\alpha \psi_k + G_k, \tag{4.48}$$

where $\mathcal{A}^{\alpha}_{<k-m}$ is the $\mathfrak{so}(n)$ matrix

$$(\mathcal{A}^{\alpha}_{<k-m})^a_b = \left(\mathcal{S}^a_{bc}(\phi) - \mathcal{S}^b_{ac}(\phi)\right)_{<k-m}\partial^\alpha \phi^c_{<k-m}. \tag{4.49}$$

Assume that ϕ is a smooth wave map on I with the bounds:

$$\| \phi \|_{\underline{E}[I]} + \| \phi \|_{\underline{X}[I]} + \| \phi \|_{S[I]} \leqslant F. \tag{4.50}$$

Furthermore, assume that $m \geqslant m(F) > 20$, for a certain function $m(F) \sim \ln(F)$. Then we have the estimate:

$$\| \psi_k \|_S \lesssim_F \| \psi_k[0] \|_{\dot{H}^1 \times L^2} + \| G_k \|_N. \tag{4.51}$$

We remark on the role of the parameter m. If ϕ is small (i.e., F is small in the theorem) then any $m \geq 10$ suffices. However, if ϕ is large, then we need an alternate source for smallness.

Outline of the proof. The proof of this result comes in two flavors:

a) Small ϕ. Set $m = 10$. A direct use of the renormalization matrix $O_{<k-m}$, as shown in the previous section, reduces the problem to an equation for $w_k = O_{<k-m}\psi_k$, namely

$$\Box w_k = O_{<k-m}F_k + \mathrm{Err}\,\psi_k,$$

where the terms on the right are estimated directly using the bilinear and trilinear estimates in Proposition 4.3:

$$\|O_{<k-m}F_k\|_N \lesssim \|F_k\|_N, \quad \|Err\ \psi_k\|_N \lesssim F\|\psi_k\|_S.$$

The smallness of F yields the smallness of the error term, therefore one can conclude using the $N \to S$ estimate (4.29) for the $\Box$ equation.

b) Large ϕ. In this case the previous argument no longer works because the errors are no longer small. This is where the large parameter m plays a key role, and provides a more subtle form of smallness which replaces the smallness coming from ϕ.

In the first step we consider energy estimates. Precisely, our paradifferential equation is essentially a covariant wave equation, therefore energy estimates can be established directly using integration by parts, with an error which is small for large m. In addition, characteristic energy estimates, i.e., bounds for the $\underline{E}$ norm, are just as easy to obtain. Precisely, we have

$$\|\psi_k\|_{\underline{E}} \lesssim_F \|\psi_k[0]\|_{\dot{H}^1 \times L^2} + 2^{-\delta m}\|\psi_k\|_S.$$

In a second step we apply the renormalization procedure; however, instead of directly applying the bilinear and trilinear estimates in Proposition 4.3, we refine them so that the bulk of the error is estimated using the characteristic energy estimates, and only a small part, corresponding to small angle interactions, is done using the full S norm of ψ_k,

$$\|\mathrm{Err}\, w_k\|_N \lesssim_F \epsilon^{-N}\|\psi_k\|_{\underline{E}} + \epsilon\|\psi_k\|_S, \quad \epsilon \ll 1.$$

Combining the last two estimates, we obtain (with a new $\delta > 0$)

$$\|\mathrm{Err}\, w_k\|_N \lesssim_F \|\psi[0]\|_{\dot{H}^1 \times L^2} + 2^{-\delta m}\|\psi_k\|_S,$$

and now we can use again (4.29) to close the argument provided that m is large enough (depending on F).

We note that all implicit constants are polynomial in F, which leads to a logarithmic dependence of $m(F)$ on F. $\square$

4.5 The small data result

Here we outline the proof of the small data result in Theorem 3.6. This is achieved in several steps:

4.5.1 The a priori estimate

The aim here is to start with a smooth wave map on a time interval I, which is a priori assumed to satisfy the bound

$$\|\phi\|_S \leq \epsilon \tag{4.52}$$

for some sufficiently small ϵ. Then we establish the following two estimates:

$$\|\phi\|_S \lesssim \|\phi[0]\|_{\dot{H}^1 \times L^2}, \tag{4.53}$$

$$\|\phi\|_{S^N} \lesssim \|\phi[0]\|_{\dot{H}^N \times \dot{H}^{N-1}}, \tag{4.54}$$

where S^N stands for functions with $N-1$ spatial derivatives in S.

In effect, it is very convenient to provide a more precise version of this, expressed in the language of frequency envelopes. Precisely, one starts with a frequency envelope c_k for the initial data $\phi[0]$, i.e.,

$$\|\phi_k[0]\|_{\dot{H}^1 \times L^2} \leq c_k.$$

Then the estimate to prove is

$$\|\phi_k\|_S \lesssim c_k. \tag{4.55}$$

A similar analysis can be carried out at the level of the S^N norms.

To achieve this we begin with the full equation

$$\Box \phi^i = -\mathcal{S}^i_{jl}(\phi)\partial^\alpha \phi^j \partial_\alpha \phi^l,$$

apply Littlewood–Paley projections, and rewrite it in the paradifferential form

$$(\Box + 2A_j(\phi)_{<k}\partial^\alpha \phi^j_{<k}\partial_\alpha)\phi_k = F_k.$$

The functions F_k contain all the interactions not included in the left, and can be estimated directly using the bilinear and trilinear estimates in Proposition 4.3:

$$\|F_k\|_{N_k} \lesssim d_k \|\phi\|_S F(\|\phi\|_S),$$

where d_k is a frequency envelope for ϕ in S, for now unrelated to c_k.

Applying Proposition 4.6 we obtain the bound

$$\|\phi_k\|_S \lesssim \|\phi_k[0]\|_{\dot{H}^1 \times L^2} + \|F_k\|_N,$$

which leads to

$$d_k \lesssim c_k + \epsilon d_k.$$

Given our assumption on the smallness of ϵ, we obtain $d_k \lesssim c_k$, and the desired conclusion follows.

4.5.2 Global existence and regularity

Consider a smooth initial data set $\phi[0]$ with small energy, say $\ll \epsilon$. Then for a short time there is a smooth solution ϕ, which can be easily shown to be small in S. We consider the set of times T for which a smooth solution satisfying $\|\phi\|_{S[-T,T]} \leq \frac{1}{2}\epsilon$ exists in $[-T,T]$. The family of rescaled functions $\phi(t/T, x/T)$ depends smoothly on T, so it will have an S norm depending continuously on T. By Step 1 it follows that the threshold $\frac{1}{2}\epsilon$ is never reached. By an open/close argument this shows that the solution exists for all t, and satisfies the bound $\|\phi\|_{S[-T,T]} \leq \frac{1}{2}\epsilon$.

4.5.3 Weak Lipschitz dependence on the initial data

Here we consider the linearized wave map equation, which has the form

$$\Box \psi^l = -(\partial_m \mathcal{S}^l_{ij})(\phi)\psi^m \partial^\alpha \phi^i \partial_\alpha \phi^j - 2\mathcal{S}^l_{ij}(\phi)\partial^\alpha \phi^i \partial_\alpha \psi^j. \tag{4.56}$$

The function ψ must satisfy the compatibility condition

$$\psi(t,x) \in T_{\phi(t,x)}M. \tag{4.57}$$

Understanding the behavior of these equations is the key to comparing different solutions of the wave maps equation.

The goal here is to show that under the assumption $\|\phi\|_S \leq \epsilon$, we have a bound of the form

$$\|\psi\|_{S^{-\delta}} \lesssim \|\psi[0]\|_{\dot{H}^{1-\delta}\times \dot{H}^{-\delta}} \tag{4.58}$$

for some small δ.

The proof of this bound is similar to the proof of the main estimate (4.53). We write the equations for ψ_k, which evolve along the same paradifferential flow as the equations for ϕ_k, and then show that the errors are small and use Proposition 4.6.

A consequence of the above bound is an estimate for the difference of solutions,

$$\|\phi_1 - \phi_2\|_{S^{-\delta}} \lesssim \|\phi_1[0] - \phi_2[0]\|_{\dot{H}^{1-\delta}\times \dot{H}^{-\delta}}. \tag{4.59}$$

4.5.4 Rough solutions and continuous dependence on the initial data

Given any small energy datum $\phi[0]$, we approximate it with a sequence of regularized data

$$\phi_n[0] \to \phi[0] \quad \text{in } (\dot{H}^1 \times L^2) \cap (\dot{H}^{1-\delta} \times \dot{H}^{-\delta}).$$

It is not difficult to show that $\phi_n[0]$ can be chosen to inherit the frequency envelope from $\phi[0]$. Then we have a corresponding sequence $\phi^{(n)}$ of smooth solutions, which by the previous step is Cauchy in $S^{-\delta}$. It also has a common frequency envelope in S. Together these two facts show that $\phi^{(n)}$ is actually Cauchy in S; thus we obtain a unique limit ϕ which is small in S.

The same argument yields continuous dependence on the initial data in the $(\dot{H}^1 \times L^2) \cap (\dot{H}^{1-\delta} \times \dot{H}^{-\delta})$ topology. Due to the finite speed of propagation, a localized form of this result is also available; it asserts continuous dependence on the initial data in the $H^1_{\text{loc}} \times L^2_{\text{loc}}$ topology.

4.6 Energy dispersion

Here we discuss the first step toward the study of the large data problem. The idea is that there should be some dichotomy between concentration of wave maps and the global existence of large data solutions. In other words, it would be reasonable to expect that if no concentration occurs then solutions persist globally. This was the viewpoint adopted by Sterbenz-Tataru in [37].

The interesting question though is what is the meaning of "*concentration*". To address that, in [37] was introduced the notion of energy dispersion. For a time interval I we set

$$\|\phi\|_{ED[I]} = \sup_k \| P_k\phi \|_{L^\infty_{t,x}[I\times\mathbb{R}^2]}. \tag{4.60}$$

Then the main result asserts that energy dispersed solutions are good:

Theorem 4.7 (Energy Dispersed Regularity Theorem [37])**.** *There exist two functions,*

$$1 \ll F(E), \quad 0 < \epsilon(E) \ll 1,$$

of the energy such that the following statement is true. If ϕ is a finite energy solution to (4.1) *on the open interval I with energy E and:*

$$\sup_k \| \phi \|_{ED[I]} \leqslant \epsilon(E), \tag{4.61}$$

then one also has

$$\| \phi \|_{S[I]} \leqslant F(E). \tag{4.62}$$

Finally, such a solution $\phi(t)$ extends in a regular way to a neighborhood of the closure of the interval I.

In the remainder of this section we provide an outline of the proof of Theorem 4.7.

In order to construct the functions $F(E)$ and $\epsilon(E)$ such that (4.61) and (4.62) hold we use the induction on energy method. Precisely, we will show that there exists a strictly positive nonincreasing function defined for *all* values of E, $c_0 = c_0(E) \ll 1$, so that if the conclusion of the Theorem holds up to energy E, then it also holds up to energy $E + c_0$. It is important here that c_0 depends *only* on E and not on the size of $F(E)$ or $\epsilon(E)$, as otherwise we would only be able to conclude the usual first step in an induction on energy proof, which is establishing that the set of regular energies is open.

According to Theorem 3.6 we know that $\epsilon(E)$ and $F(E)$ can be constructed up to some $E_0 \ll 1$. We now assume that E_0 is fixed by induction, and to increase its range we consider a solution ϕ defined on an interval I with energy $E[\phi] = E_0 + c$, $c \leqslant c_0(E_0)$, and with energy dispersion $\leqslant \epsilon$ (at first this is a free parameter which we may take as small as we like). We will compare ϕ with a wave map $\tilde{\phi}$ with energy E_0. To construct $\tilde{\phi}$ we reduce the initial datum energy of $\phi[0]$ by truncation in frequency. We define the *cut frequency* $k_* \in \mathbb{R}$ according to (this can be done by adjusting the definition of the $P_{<k}$ continuously if necessary)

$$E[\Pi P_{\leqslant k_*}\phi[0]] = E_0.$$

Here we work in the extrinsic setting, and the small energy dispersion insures that the low frequency projections $P_{\leq k}\phi_0$ stay close to the manifold. Then one can use any reasonable projection operator Π to return back to the manifold.

We consider the wave-map $\tilde{\phi}$ with this initial data $\tilde{\phi}[0] = \Pi P_{\leqslant k_*}\phi[0]$. This wave-map exists classically for at least a short amount of time according to Cauchy stability, and where it exists we have:

$$E[\tilde{\phi}(t)] = E_0. \tag{4.63}$$

Since ϕ has energy dispersion $\leqslant \epsilon$, by (4.71) it follows that $\tilde{\phi}$ has energy dispersion $\lesssim_{E_0} \epsilon^{\frac{1}{4}}$ at time $t = 0$. Again by the usual Cauchy stability theory, if ϵ is chosen small enough in comparison to the inductively defined parameter $\epsilon(E_0)$ it follows that there exists a non-empty interval J_0 where $\tilde{\phi}$ satisfies

$$\sup_k \| P_k\tilde{\phi} \|_{L_t^\infty(L_x^\infty)[J_0]} \leqslant \epsilon(E_0). \tag{4.64}$$

Then our induction hypothesis guarantees that we have the dispersive bounds:

$$\| \tilde{\phi} \|_{S[J_0]} \leqslant F(E_0). \tag{4.65}$$

The plan is now very simple. On one hand, we try to pass the space-time control (i.e. the S bound) from $\tilde{\phi}$ to ϕ via linearization around $\tilde{\phi}$ to control the low frequencies, and conservation of energy and perturbation theory to control the high frequencies. On the other hand, we need to pass the good energy dispersion bounds from ϕ back down to $\tilde{\phi}$ in order to increase the size of $J \subseteq I$ on which (4.64) holds, until it eventually fills up all of I.

To summarize, we have the two wave maps $\tilde{\phi}$ and ϕ on an interval I with energies E, respectively $E + c$, so that

$$\|\tilde{\phi}\|_S \leq \tilde{F} = F(E_0), \quad \|\phi\|_{ED} \leq \epsilon, \tag{4.66}$$

and we want to prove that

$$\|\tilde{\phi}\|_{ED} \leq \tilde{\epsilon} = \epsilon(E_0), \quad \|\phi\|_S \leq F. \tag{4.67}$$

In doing this, we can freely make the bootstrap assumption

$$\|\tilde{\phi}\|_{ED} \leq 2\tilde{\epsilon}, \quad \|\phi\|_S \leq 2F. \tag{4.68}$$

We are also free to independently choose F sufficiently large and ϵ sufficiently small. But the delicate part is that c can only depend on E. The analysis is carried out in several steps:

4.6.1 Energy dispersion and multilinear estimates

Ideally, one would like to know that having small energy dispersion improves the multilinear bounds in Proposition 4.3. To understand this better, let us first discuss the null form estimate (4.8) in the easiest case when both inputs are free waves. As discussed earlier, there is an angular gain for small angle interactions, so one only needs to consider large angles, i.e. bilinear estimates for transversal waves. In that case the null form does not help, so we just treat this as a bilinear product estimate.

On one hand, using Strichartz estimates for one factor and the energy dispersion for the other we obtain an improved L^6 product estimate. On the other hand, the large angle bilinear estimate of Wolff [50] and Tao [44] shows that one also has an $L^{\frac{5}{3}}$ bound (the exact exponent does not matter, only that it is less than 2). Interpolating, one obtains an improved L^2 bound. That suffices, because the output of transversal free waves is at high modulation.

One downside of the above reasoning is that in the case of unbalanced frequency interactions one ends up with the wrong balance of the powers of the two frequencies, namely with an estimate of the type

$$\|\partial^\alpha \phi_j \partial_\alpha \phi_k\|_N \lesssim 2^{c|j-k|} \|\phi_j[0]\|_{\dot H^1\times L^2} \|\phi_k[0]\|^{1-\delta}_{\dot H^1\times L^2} \|\phi_k\|^{\delta}_{L^\infty}, \quad c,\delta > 0.$$

Hence this energy dispersion gain is effective only in the case of balanced factors.

Ideally one would like to have the same estimate for S inputs. While this is not out of question, we were unable to prove that. Instead, we only have weaker estimates of the form

$$\begin{aligned}&\|\partial^\alpha \phi_j \partial_\alpha \phi_k\|_N\\ &\lesssim 2^{c|j-k|} (\|\phi_j\|_{\dot H^1\times L^2} + \|\Box \phi_j\|_N)(\|\phi_k[0]\|_{\dot H^1\times L^2} + \|\Box\phi_k\|_N)^{1-\delta} \|\phi_k\|^{\delta}_{L^\infty}, \quad c,\delta > 0.\end{aligned}$$

The dyadic portions of our wave maps do not have this regularity. However, they do have it after renormalization. This is the reason why we introduce the following definition:

Definition 4.8 (Renormalizable Functions)**.** *We define a non-linear functional $\mathcal{W}_k$ on S as follows:*

$$\begin{aligned}\| \phi \|_{\mathcal{W}_k} := \inf_{U\in SO(d)} \Big[&\big(\| U \|_{S\cap \underline{X}} + \sup_{j\geqslant k} 2^{C(j-k)} \| P_j U \|_{S\cap\underline{X}}\big)\\ &\cdot \sup_{k'} 2^{|k'-k|} \big(\| P_{k'}(U\phi_k)[0] \|_{\dot H^1\times L^2} + \| P_{k'}\Box(U\phi_k) \|_N\big)\Big].\end{aligned} \tag{4.69}$$

Using this notation, the above bound is improved to

$$\|\partial^\alpha \phi_j \partial_\alpha \phi_k\|_N \lesssim 2^{c|j-k|} \|\phi_j\|_{\mathcal{W}_j} \|\phi_k\|^{1-\delta}_{\mathcal{W}_k} \|\phi_k\|^{\delta}_{L^\infty}, \quad c,\delta > 0. \tag{4.70}$$

Similar improvements apply to the other multilinear estimates in Proposition 4.3, provided that at least two of the interacting frequencies are balanced.

These improved estimates are crucial in order to gain the large gap m that is needed in Proposition 4.6.

4.6.2 Compare the initial data of ϕ and $\tilde{\phi}$

At the linearized level we have $\tilde{\phi}[0] = P_{<k_*}\phi[0]$. This is not an identity, but the errors are higher order, and they will be small due to the energy dispersion:

$$\| P_k\big(P_{<k_*}\phi[0] - \tilde{\phi}[0]\big) \|_{\dot{H}^1\times L^2} \lesssim_E \epsilon^{\frac{1}{4}} 2^{-\frac{1}{2}|k-k_*|}. \tag{4.71}$$

4.6.3 Compare the low frequencies of ϕ and $\tilde{\phi}$.

The previous step shows that the low frequencies of the data for ϕ and $\tilde{\phi}$ are very close. Here we aim to show that a similar bound holds for the difference of the solutions,

$$\| P_k\big(P_{<k_*}\phi - \tilde{\phi}\big) \|_S \lesssim_F 2^{-\delta_0|k-k_*|}\epsilon^{\delta_0}. \tag{4.72}$$

This yields the small energy dispersion for $\tilde{\phi}$, provided that ϵ is small enough. To prove (4.72) we consider the equation for the difference $\psi = P_{<k_*}\phi - \tilde{\phi}$. This has the form

$$\begin{aligned}
\Box\psi &= -\mathcal{S}(\tilde{\phi})\partial^\alpha\tilde{\phi}\partial_\alpha\tilde{\phi} + P_{<k_*}(\mathcal{S}(\phi)\partial^\alpha\phi\partial_\alpha\phi)\\
&= -\mathcal{S}(\tilde{\phi})\partial^\alpha\tilde{\phi}\partial_\alpha\tilde{\phi} + \mathcal{S}(\tilde{\phi}+\psi)\partial^\alpha(\tilde{\phi}+\psi)\partial_\alpha(\tilde{\phi}+\psi) + R(\phi),
\end{aligned}$$

where

$$R(\phi) = P_{<k_*}(\mathcal{S}(\phi)\partial^\alpha\phi\partial_\alpha\phi) - \mathcal{S}(P_{<k^*}\phi)\partial^\alpha P_{<k^*}\phi\partial_\alpha P_{<k^*}\phi.$$

We rewrite the above equation in the paradifferential form

$$\Box\psi_k = -2\mathcal{A}^\alpha_{<k-m}(\tilde{\phi})\partial_\alpha\psi_k + \mathrm{Err}_k(\psi) + P_kR(\phi).$$

Provided ϵ is small enough, the remaining part evolves essentially along the linearized flow along $\tilde{\phi}$, and can be solved perturbatively using the linear covariant estimates in Proposition 4.6, with respect to a norm defined as in (4.72). It remains to establish good estimates for the last two terms on the right.

The term $R(\phi)$ is estimated in N using the S norm for ϕ and its energy dispersion:

$$\|P_kR(\phi)\|_N \lesssim_F 2^{-\delta_0|k-k_*|}\epsilon^{\delta_0}. \tag{4.73}$$

The term $\mathrm{Err}_k(\psi)$ is at least quadratic in ψ. It is estimated directly, using Proposition 4.3 for unbalanced frequency interactions, and its energy dispersed improvement for the balanced ones. We remark that here we use the energy dispersion of $\tilde{\phi}$, but that it still can be assumed to be small enough to defeat the S norm of ψ.

4.6.4 Compare the high frequencies

Here we estimate directly the difference $\psi = \phi - \tilde{\phi}$,

$$\|\phi - \tilde{\phi}\|_S \lesssim_{\tilde{F}} 1. \tag{4.74}$$

This yields the S bound for ϕ in (4.67). The tricky bit is to do this with a constant c which depends only on E and not on $\tilde{F}$.

The function ψ has initial datum of size c, and solves the equation

$$\Box\psi = -\mathcal{S}(\tilde{\phi})\partial^\alpha\tilde{\phi}\partial_\alpha\tilde{\phi} + \mathcal{S}(\tilde{\phi}+\psi)\partial^\alpha(\tilde{\phi}+\psi)\partial_\alpha(\tilde{\phi}+\psi).$$

We need to estimate only its high frequencies, i.e., larger than k_*. The idea is to reduce the problem again to a perturbation of the gauge covariant equation (4.48), but this time with coefficients depending on ϕ rather than $\tilde{\phi}$. The difficulty is that the size of $\tilde{\phi}$ is large, and this would force the needed smallness of c to depend on $\tilde{F}$ rather than on E. To remedy this, we need several intermediate steps:

(i) Establish uniform energy bounds for ψ in the energy norm, which do not depend on F. This is done using the energy estimates for both ϕ and $\tilde{\phi}$, combined with the bound (4.72), which guarantees their almost orthogonality.

(ii) Prove a partial divisibility result for the S norm of $\tilde{\phi}$, as follows:

Lemma 4.9. *Let $\tilde{\phi}$ be a wave map with energy E and S norm $\tilde{F}$. Then there exists a collection of subintervals $I = \bigcup_{i=1}^K I_i$, such that $K = K(\tilde{F})$ depends only on $\tilde{F}$, and such that the following bound holds on each I_i:*

$$\| \tilde{\phi} \|_{S[I_i]} \lesssim_E 1 . \tag{4.75}$$

(iii) Use the perturbative argument to estimate the S norm of ψ in each interval I_k. In each interval we do have the small energy dispersion for ϕ, but all other constants depend only on E; hence the smallness condition on c will also depend only on E, and so will the S bound on ψ on I_k. On the other hand, the number of intervals and thus the global S bound for ψ will depend on $\tilde{F}$.

4.7 Energy and Morawetz estimates

The study of the large data problem for wave maps relies on the finite speed of propagation property of the wave equation. Because of this and of the small data result, the following conclusion follows:

If a wave map blows up at a point, then its energy must concentrate toward the tip of the light cone originating at that point. Similarly, if scattering fails, then it fails inside a light cone.

Thus, in order to study both blow up and scattering, it suffices to consider finite energy wave maps inside a light cone. In one case we are interested in what happens at the tip of the light cone, in the other we are interested in what happens inside the cone but toward infinity. We will see that the two problems are virtually identical. The main tools in the study of the energy distribution inside the cone are the energy and the Morawetz estimates. These are described in the sequel.

4.7.1 Notations

We consider the forward light cone

$$C = \{0 \le t < \infty,\ r \le t\}$$

and its subsets

$$C_{[t_0,t_1]} = \{t_0 \le t \le t_1,\ r \le t\}\ .$$

The lateral boundary of $C_{[t_0,t_1]}$ is denoted by $\partial C_{[t_0,t_1]}$. The time sections of the cone are denoted by

$$S_{t_0} = \{t = t_0,\ |x| \le t\}\ .$$

We also use the translated cones

$$C^\delta = \{\delta \le t < \infty,\ r \le t - \delta\},$$

as well as the corresponding notations $C^\delta_{[t_0,t_1]}$, $\partial C^\delta_{[t_0,t_1]}$ and $S^\delta_{t_0}$ for $t_0 > \delta$.

For some of the computations below it is convenient to use the null frame

$$L = \partial_t + \partial_r, \quad \bar{L} = \partial_t - \partial_r, \quad \not\partial = r^{-1}\partial_\theta.$$

4.7.2 The energy-momentum tensor

A systematic way to derive both the energy and the Morawetz estimates is by using the energy-momentum tensor:

$$T_{\alpha\beta}[\Phi] = g_{ij}(\Phi)\big[\partial_\alpha\phi^i\partial_\beta\phi^j - \frac{1}{2}m_{\alpha\beta}\,\partial^\gamma\phi^i\partial_\gamma\phi^j\big], \tag{4.76}$$

with a well chosen vector field. Here $\Phi = (\phi^1, \ldots, \phi^n)$ is a set of local coordinates on the target manifold $(\mathbb{M}, g)$ and $(m_{\alpha\beta})$ stands for the Minkowski metric. The main two properties of $T_{\alpha\beta}[\Phi]$ are:

- it is divergence free, $\nabla^\alpha T_{\alpha\beta} = 0$;
- it obeys the positive energy condition $T(X,Y) \geqslant 0$ whenever both $m(X,X) \leqslant 0$ and $m(Y,Y) \leqslant 0$.

Our estimates are obtained by contracting the energy-momentum tensor with a well chosen vector field. The above properties imply that contracting $T_{\alpha\beta}[\Phi]$ with timelike/null vector fields will result in good energy estimates on characteristic and space-like hypersurfaces.

If X is some vector field, we can form its associated momentum density (i.e., its Noether current)

$${}^{(X)}P_\alpha = T_{\alpha\beta}[\Phi]X^\beta.$$

This one-form obeys the divergence rule

$$\nabla^{\alpha\ (X)}P_\alpha = \frac{1}{2}T_{\alpha\beta}[\Phi]^{(X)}\pi^{\alpha\beta}, \tag{4.77}$$

where ${}^{(X)}\pi_{\alpha\beta}$ is the deformation tensor of X,

$$ {}^{(X)}\pi_{\alpha\beta} = \nabla_\alpha X_\beta + \nabla_\beta X_\alpha . $$

A simple computation shows that one can also express

$$ {}^{(X)}\pi = \mathcal{L}_X g. $$

This latter formulation is very convenient when dealing with coordinate derivatives. Recall that in general one has

$$ (\mathcal{L}_X g)_{\alpha\beta} = X(g_{\alpha\beta}) + \partial_\alpha(X^\gamma) g_{\gamma\beta} + \partial_\beta(X^\gamma) g_{\alpha\gamma}. $$

The energy estimates are obtained by integrating the relation (4.77) over cones $C^\delta_{[t_1,t_2]}$. Then from (4.77) we obtain, for $\delta \le t_1 \le t_2$:

$$ \int_{S^\delta_{t_2}} {}^{(X)}P_0 \, dx + \frac{1}{2} \int_{C^\delta_{[t_1,t_2]}} T_{\alpha\beta}[\Phi] {}^{(X)}\pi^{\alpha\beta} \, dxdt = \int_{S^\delta_{t_1}} {}^{(X)}P_0 \, dx + \int_{\partial C^\delta_{[t_1,t_2]}} {}^{(X)}P_L \, dA, \quad (4.78) $$

where dA is an appropriately normalized (Euclidean) surface area element on the lateral boundary of the cone $r = t - \delta$.

4.7.3 Energy estimates

The standard energy estimates come from contracting $T_{\alpha\beta}[\Phi]$ with $Y = \partial_t$. Then we have

$$ {}^{(Y)}\pi = 0, \quad {}^{(Y)}P_0 = \frac{1}{2}(|\partial_t\Phi|^2 + |\nabla_x\Phi|^2), \quad {}^{(Y)}P_L = \frac{1}{4}|L\Phi|^2 + \frac{1}{2}|\partial\!\!\!/\Phi|^2. $$

Applying (4.78) over $C_{[t_1,t_2]}$ we obtain the energy-flux relation

$$ E_{S_{t_1}}[\Phi] = E_{S_{t_0}}[\Phi] + F_{[t_0,t_1]}[\Phi], \quad (4.79) $$

where E_{S_t} represents the energy of Φ on time sections,

$$ E_{S_t}[\Phi] = \frac{1}{2} \int_{S_t} (|\partial_t\Phi|^2 + |\nabla_x\Phi|^2) dx, $$

and $F_{[t_0,t_1]}[\Phi]$ represents the lateral flux of Φ between t_0 and t_1:

$$ F_{[t_0,t_1]}[\Phi] = \int_{\partial C_{[t_0,t_1]}} \big(\frac{1}{4}|L\Phi|^2 + \frac{1}{2}|\partial\!\!\!/\Phi|^2 \big) \, dA. $$

The energy relation (4.79) shows that $E_{S_t}[\Phi]$ is a nondecreasing function of t. It also shows that for the blow-up problem we have

$$ \lim_{t_1,t_2\to 0} F_{[t_1,t_2]}[\Phi] = 0 $$

and for the scattering problem we have

$$\lim_{t_1, t_2 \to \infty} F_{[t_1,t_2]}[\Phi] = 0.$$

This is the main decay estimate arising as a consequence of the energy relation. Later we will want to turn the flux decay on the boundary of the cone into an integrated decay inside the cone. This is accomplished using Morawetz type estimates.

Finally, we remark that applying (4.78) over $C^{\delta}_{[\delta,1]}$ yields

$$\int_{\partial C^{\delta}_{[\delta,1]}} \frac{1}{4}|L\Phi|^2 + \frac{1}{2}|\not{\partial}\Phi|^2 \, dA \leq E_1[\Phi]. \tag{4.80}$$

This will be used later on.

4.7.4 The energy of self-similar maps

A map $\Phi : C \to (M, g)$ is said to be self-similar if

$$\Phi(\lambda t, \lambda x) = \Phi(t, x), \quad (t, x) \in C, \quad \lambda > 0.$$

Such a map, if it had finite energy, would be a natural obstruction to global existence of wave maps. Later we will argue that finite energy self-similar wave maps do not exist. Here we carry out a preliminary step, which is to compute the energy $E[\Phi]$ (which is independent of time) in hyperbolic coordinates.

Hyperbolic coordinates (ρ, y, Θ) are introduced inside C via

$$t = \rho \cosh(y), \quad r = \rho \sinh(y), \quad \theta = \Theta, \tag{4.81}$$

and self-similar maps Φ can be viewed as functions $\Phi = \Phi(y, \Theta)$ on $\mathbb{H}^2$.

In this system of coordinates, the Minkowski metric becomes

$$-dt^2 + dr^2 + r^2 d\theta^2 = -d\rho^2 + \rho^2\big(dy^2 + \sinh^2(y) d\Theta^2\big). \tag{4.82}$$

A quick calculation shows that the contraction on line (4.84) becomes the one-form

$${}^{(Y)}P^{\alpha} dV_{\alpha} = T(\partial_\rho, \partial_t)\rho^2 dA_{\mathbb{H}^2}, \quad dA_{\mathbb{H}^2} = \sinh(y) dy d\Theta. \tag{4.83}$$

The area element $dA_{\mathbb{H}^2}$ is that of the hyperbolic plane $\mathbb{H}^2$. To continue, we note that

$$\partial_t = \frac{t}{\rho}\partial_\rho - \frac{r}{\rho^2}\partial_y,$$

so in particular

$$T(\partial_\rho, \partial_t) = \frac{\cosh(y)}{2}|\partial_\rho \Phi|^2 - \frac{\sinh(y)}{\rho}\partial_\rho \Phi \cdot \partial_y \Phi + \frac{\cosh(y)}{2\rho^2}\Big(|\partial_y \Phi|^2 + \frac{1}{\sinh^2(y)}|\partial_\Theta \Phi|^2\Big).$$

This computation allows us to obtain a version of the usual energy estimate adapted to the hyperboloids $\sqrt{t^2 - r^2} = 1$. Integrating the divergence of the ${}^{(Y)}P_\alpha$ momentum density over regions of the form $\mathcal{R} = \{\rho \geqslant \rho_0, t \leqslant t_0\}$ we have:

$$\int_{\{\rho=1\}\cap\{t\leqslant t_0\}} {}^{(Y)}P^\alpha dV_\alpha = \int_{\{\rho>1\}\cap\{t=t_0\}} {}^{(Y)}P_0 dx, \tag{4.84}$$

where the integrand on the LHS denotes the interior product of ${}^{(Y)}P$ with the Minkowski volume element.

Letting $t_0 \to \infty$ in (4.84) we obtain a useful consequence of this, namely a weighted hyperbolic space estimate for special solutions to the wave-map equations, which will be used in the sequel to rule out the existence of non-trivial finite energy self-similar solutions:

Lemma 4.10. *Let Φ be a self-similar finite energy smooth wave map in the interior of the cone C Then one has:*

$$\mathcal{E}[\Phi] = \frac{1}{2}\int_{\mathbb{H}^2} |\nabla_{\mathbb{H}^2}\Phi|^2 \cosh(y) dA_{\mathbb{H}^2}. \tag{4.85}$$

Here

$$|\nabla_{\mathbb{H}^2}\Phi|^2 = |\partial_y \Phi|^2 + \frac{1}{\sinh^2(y)}|\partial_\Theta \Phi|^2$$

is the covariant energy density for the hyperbolic metric.

4.7.5 Morawetz estimates

Our goal here is to obtain decay estimates for time-like components of the energy density. For this we use the energy momentum estimate (4.78) with respect to the timelike/null vector field

$$X_\epsilon = \frac{1}{\rho_\epsilon}\left((t+\epsilon)\partial_t + r\partial_r\right), \quad \rho_\epsilon = \sqrt{(t+\epsilon)^2 - r^2}. \tag{4.86}$$

In order to gain some intuition, we first consider the case of X_0. This is most readily expressed in the system of hyperbolic coordinates (4.81). One easily checks that the coordinate derivatives turn out to be

$$\partial_\rho = X_0, \quad \partial_y = r\partial_t + t\partial_r.$$

In particular, X_0 is uniformly timelike with $m(X_0, X_0) = -1$, and one should expect it to generate good energy estimate on time slices $t = \text{const}$. In the system of coordinates (4.81) one also has that

$$\mathcal{L}_{X_0} m = 2\rho\big(dy^2 + \sinh^2(y) d\Theta^2\big).$$

Raising indices, one then computes

$$^{(X_0)}\pi^{\alpha\beta} = \frac{2}{\rho^3}\big(\partial_y \otimes \partial_y + \sinh^{-2}(y)\partial_\Theta \otimes \partial_\Theta\big).$$

Therefore, we have the contraction identity

$$\frac{1}{2}T_{\alpha\beta}[\Phi]^{(X_0)}\pi^{\alpha\beta} = \frac{1}{\rho}|X_0\Phi|^2.$$

To compute the components of $^{(X_0)}P_0$ and $^{(X_0)}P_L$ we use the associated optical functions

$$u = t - r, \quad v = t + r, \quad uv = \rho^2.$$

Then we have

$$X_0 = \frac{1}{\rho}\left(\frac{1}{2}vL + \frac{1}{2}u\bar{L}\right), \quad \partial_t = \frac{1}{2}L + \frac{1}{2}\bar{L}. \tag{4.87}$$

Finally, we record here the components of $T_{\alpha\beta}[\Phi]$ in the null frame:

$$T(L,L) = |L\Phi|^2, \quad T(\bar{L},\bar{L}) = |\bar{L}\Phi|^2, \quad T(L,\bar{L}) = |\not{\partial}\Phi|^2.$$

By combining the above calculations, we see that we may compute

$$^{(X_0)}P_0 = T(\partial_t, X_0) = \frac{1}{4}\left(\frac{v}{u}\right)^{\frac{1}{2}}|L\Phi|^2 + \frac{1}{4}\left[\left(\frac{v}{u}\right)^{\frac{1}{2}} + \left(\frac{u}{v}\right)^{\frac{1}{2}}\right]|\not{\partial}\Phi|^2 + \frac{1}{4}\left(\frac{u}{v}\right)^{\frac{1}{2}}|\bar{L}\Phi|^2,$$
$$^{(X_0)}P_L = T(L, X_0) = \frac{1}{2}\left(\frac{v}{u}\right)^{\frac{1}{2}}|L\Phi|^2 + \frac{1}{2}\left(\frac{u}{v}\right)^{\frac{1}{2}}|\not{\partial}\Phi|^2.$$

These are essentially the same as the components of the usual energy currents $^{(\partial_t)}P_0$ and $^{(\partial_t)}P_L$ modulo ratios of the optical functions u and v.

One would expect to get nice space-time estimates for $X_0\Phi$ by integrating (4.77) over the interior cone $r \leqslant t \leqslant 1$. The problem is that the boundary terms degenerate when $\rho \to 0$. To avoid this difficulty we simply redo everything with the shifted version X_ϵ from line (4.86). The above formulas remain valid with u, v replaced by their time shifted versions

$$u_\epsilon = (t+\epsilon) - r, \quad v_\epsilon = (t+\epsilon) + r.$$

Furthermore, notice that for small t one has the bounds

$$\left(\frac{v_\epsilon}{u_\epsilon}\right)^{\frac{1}{2}} \approx 1, \quad \left(\frac{u_\epsilon}{v_\epsilon}\right)^{\frac{1}{2}} \approx 1, \quad 0 < t \leq \epsilon$$

within the cone C. Thus,

$$^{(X_\epsilon)}P_0 \approx {}^{(\partial_t)}P_0, \quad 0 < t \leq \epsilon.$$

In what follows we work with a wave map Φ in $C_{[\epsilon,1]}$. We denote its total energy and flux by

$$E = E_{S_1}[\Phi], \quad F = F_{[\epsilon,1]}[\Phi].$$

In the limiting case $F = 0$, $\epsilon = 0$ one could apply (4.78) to obtain

$$\int_{S^0_{t_2}} {}^{(X_\epsilon)}P_0 \, dx + \int_{C^0_{[t_1,t_2]}} \frac{1}{\rho_\epsilon} |X_\epsilon \Phi|^2 \, dxdt = \int_{S^0_{t_1}} {}^{(X_\epsilon)}P_0 \, dx.$$

By (4.88), letting $t_1 \to 0$ followed by $\epsilon \to 0$ and taking supremum over $0 < t_2 \leqslant 1$ we would get the model estimate

$$\sup_{t\in(0,1]} \int_{S^0_t} {}^{(X_0)}P_0 \, dx + \int_{C^0_{[0,1]}} \frac{1}{\rho} |X_0 \Phi|^2 \, dxdt \leq E.$$

However, here we need to deal with a small nonzero flux. Observing that

$${}^{(X_\epsilon)}P_L \lesssim \epsilon^{-\frac{1}{2}} {}^{(\partial_t)}P_L,$$

from (4.78) we obtain the weaker bound

$$\int_{S^0_{t_2}} {}^{(X_\epsilon)}P_0 \, dx + \int_{C^0_{[t_1,t_2]}} \frac{1}{\rho_\epsilon} |X_\epsilon \Phi|^2 \, dxdt \lesssim \int_{S^0_{t_1}} {}^{(X_\epsilon)}P_0 dx + \epsilon^{-\frac{1}{2}} F.$$

Letting $t_1 = \epsilon$ and taking the supremum over $\epsilon \leqslant t_2 \leqslant 1$ we obtain

$$\sup_{t\in(\epsilon,1]} \int_{S^0_t} {}^{(X_\epsilon)}P_0 \, dx + \int_{C^0_{[0,1]}} \frac{1}{\rho_\epsilon} |X_\epsilon \Phi|^2 \, dxdt \lesssim E + \epsilon^{-\frac{1}{2}} F. \tag{4.88}$$

A consequence of this is the following, which will be used to rule out the case of asymptotically null pockets of energy:

Lemma 4.11. *Let Φ be a smooth wave map in the cone $C_{(\epsilon,1]}$ which satisfies the flux-energy relation $F \lesssim \epsilon^{\frac{1}{2}} E$. Then*

$$\int_{S^0_1} {}^{(X_\epsilon)}P_0 \, dx \lesssim E. \tag{4.89}$$

Next, we show that we can replace X_ϵ by X_0 in (4.88) if we restrict the integrals on the left to $r < t - \epsilon$. In this region we have

$${}^{(X_\epsilon)}P_0 \approx {}^{(X_0)}P_0, \quad \rho_\epsilon \approx \rho.$$

In addition, a direct computation shows that in $r < t - \epsilon$

$$\frac{1}{\rho} |X_0 \Phi|^2 \lesssim \frac{1}{\rho_\epsilon} |X_\epsilon \Phi|^2 + \frac{\epsilon^2}{\rho^3} |\partial_t \Phi|^2,$$

and also

$$\int_{C^\epsilon_{(\epsilon,1]}} \frac{\epsilon^2}{\rho^3}|\partial_t\Phi|^2 dxdt \le \int_{C^\epsilon_{(\epsilon,1]}} \frac{\epsilon^{\frac12}}{t^{\frac32}}|\partial_t\Phi|^2 dxdt \lesssim E.$$

Thus, using the last three relations in (4.88) we have proved the following estimate which will be used to conclude that rescaling of Φ is asymptotically stationary, and also used to help trap uniformly time-like pockets of energy:

Lemma 4.12. *Let Φ be a smooth wave-map in the cone $C_{(\epsilon,1]}$ which satisfies the flux-energy relation $F \lesssim \epsilon^{\frac12}E$. Then we have*

$$\sup_{t\in(\epsilon,1]} \int_{S^\epsilon_t} {}^{(X_0)}P_0\ dx + \int_{C^\epsilon_{[\epsilon,1]}} \frac{1}{\rho}|X_0\Phi|^2 dxdt \lesssim E. \tag{4.90}$$

Finally, we use the last lemma to propagate pockets of energy forward away from the boundary of the cone. By (4.78) for X_0 we have

$$\int_{S^\delta_1} {}^{(X_0)}P_0\ dx \le \int_{S^\delta_{t_0}} {}^{(X_0)}P_0\ dx + \int_{\partial C^\delta_{[t_0,1]}} {}^{(X_0)}P_L\ dA\ , \qquad \epsilon \le \delta < t_0 < 1.$$

We consider the two components of ${}^{(X_0)}P_L$ separately. For the angular component, by (4.80) we have the bound

$$\int_{\partial C^\delta_{[t_0,1]}} \left(\frac{u}{v}\right)^{\frac12} |\partial\!\!\!/\Phi|^2\, dA \lesssim \left(\frac{\delta}{t_0}\right)^{\frac12} \int_{\partial C^\delta_{[t_0,1]}} |\partial\!\!\!/\Phi|^2\, dA \lesssim \left(\frac{\delta}{t_0}\right)^{\frac12} E.$$

For the L component a direct computation shows that

$$|L\Phi| \lesssim \left(\frac{u}{v}\right)^{\frac12} |X_0\Phi| + \left(\frac{u}{v}\right) |\bar{L}\Phi|.$$

Thus we obtain

$$\int_{S^\delta_1} {}^{(X_0)}P_0\ dx \lesssim \int_{S^\delta_{t_0}} {}^{(X_0)}P_0\, dx + \left(\frac{\delta}{t_0}\right)^{\frac12} E + \int_{\partial C^\delta_{[t_0,1]}} \left(\left(\frac{u}{v}\right)^{\frac12} |X_0\phi|^2 + \left(\frac{u}{v}\right)^{\frac32} |\bar{L}\Phi|^2\right) dA.$$

For the last term we optimize with respect to $\delta \in [\delta_0, \delta_1]$ to obtain:

Lemma 4.13. *Let Φ be a smooth wave-map in the cone $C_{(\epsilon,1]}$ which satisfies the flux-energy relation $F \lesssim \epsilon^{\frac12}E$. Suppose that $\epsilon \le \delta_0 \ll \delta_1 \le t_0$. Then*

$$\int_{S^{\delta_1}_1} {}^{(X_0)}P_0\ dx \lesssim \int_{S^{\delta_0}_{t_0}} {}^{(X_0)}P_0\ dx + \left(\left(\frac{\delta_1}{t_0}\right)^{\frac12} + (\ln(\delta_1/\delta_0))^{-1}\right) E. \tag{4.91}$$

To prove this lemma, it suffices to choose $\delta \in [\delta_0, \delta_1]$ so that

$$\int_{\partial C^{\delta}_{[t_0,1]}} \left[\left(\frac{u}{v}\right)^{\frac{1}{2}} |X_0\phi|^2 + \left(\frac{u}{v}\right)^{\frac{3}{2}} |\bar{L}\Phi|^2 \right] dA \lesssim |\ln(\delta_1/\delta_0)|^{-1} E.$$

This follows by pigeonholing the estimate

$$\int_{C^{\delta_0}_{[t_0,1]} \setminus C^{\delta_1}_{[t_0,1]}} \frac{1}{u} \left[\left(\frac{u}{v}\right)^{\frac{1}{2}} |X_0\phi|^2 + \left(\frac{u}{v}\right)^{\frac{3}{2}} |\bar{L}\Phi|^2 \right] dxdt \lesssim E.$$

The first term is estimated directly by (4.90). For the second we simply use energy bounds, since in the domain of integration we have the relation

$$\frac{1}{u}\left(\frac{u}{v}\right)^{\frac{3}{2}} \leq \frac{\delta_1^{\frac{1}{2}}}{t^{\frac{3}{2}}}.$$

4.8 The threshold theorem

Using the energy dispersed result in Theorem 4.7 and the energy/Morawetz estimates in the previous section we can now approach the large data problem. For the blow-up question we prove the following:

Theorem 4.14 ([38]). *Let $\Phi : C_{(0,1]} \to \mathbb{M}$ be a C^∞ wave map. Then exactly one of the following possibilities must hold:*

(1) *There exists a sequence of points $(t_n, x_n) \in C_{[0,1]}$ and scales r_n with*

$$(t_n, x_n) \to (0,0), \quad \limsup \frac{|x_n|}{t_n} < 1, \quad \lim \frac{r_n}{t_n} = 0$$

such that the rescaled sequence of wave-maps

$$\Phi^{(n)}(t,x) = \Phi\big(t_n + r_n t, x_n + r_n x\big) \tag{4.92}$$

converges strongly in H^1_{loc} to a Lorentz transform of an entire harmonic map of nontrivial energy:

$$\Phi^{(\infty)} : \mathbb{R}^2 \to \mathbb{M}, \quad 0 < \| \Phi^{(\infty)} \|_{\dot{H}^1(\mathbb{R}^2)} \leqslant \lim_{t\to 0} E_{S_t}[\Phi] \ .$$

(2) *For each $\epsilon > 0$ there exists $0 < t_0 \leqslant 1$ and a wave map extension*

$$\Phi : \mathbb{R}^2 \times (0, t_0] \to \mathbb{M}$$

with bounded energy

$$E[\Phi] \leq (1 + \epsilon^8) \lim_{t\to 0} E_{S_t}[\Phi] \tag{4.93}$$

and energy dispersion

$$\sup_{t\in(0,t_0]} \sup_{k\in\mathbb{Z}} \big(\|P_k\Phi(t)\|_{L^\infty_x} + 2^{-k}\|P_k\partial_t\Phi(t)\|_{L^\infty_x} \big) \leq \epsilon \ . \tag{4.94}$$

The analogue result for the scattering problem also holds:

Theorem 4.15. *Let $\Phi : C_{[1,\infty)} \to \mathbb{M}$ be a C^∞ wave map with finite energy. Then exactly one of the following possibilities must hold:*

(1) *There exists a sequence of points $(t_n, x_n) \in C_{[1,\infty)}$ and scales r_n with*

$$t_n \to \infty, \quad \limsup \frac{|x_n|}{t_n} < 1, \quad \lim \frac{r_n}{t_n} = 0,$$

so that the rescaled sequence of wave maps

$$\Phi^{(n)}(t,x) = \Phi\big(t_n + r_n t, x_n + r_n x\big) \tag{4.95}$$

converges strongly in H^1_{loc} to a Lorentz transform of an entire harmonic map of nontrivial energy:

$$\Phi^{(\infty)} : \mathbb{R}^2 \to \mathbb{M}, \quad 0 < \| \Phi^{(\infty)} \|_{\dot H^1(\mathbb{R}^2)} \leqslant \lim_{t\to\infty} E_{S_t}[\Phi].$$

(2) *For each $\epsilon > 0$ there exist $t_0 > 1$ and a wave map extension*

$$\Phi : \mathbb{R}^2 \times [t_0, \infty) \to \mathbb{M}$$

with bounded energy

$$E[\Phi] \le (1+\epsilon^8) \lim_{t\to\infty} E_{S_t}[\Phi] \tag{4.96}$$

and energy dispersion,

$$\sup_{t\in[t_0,\infty)} \sup_{k\in\mathbb{Z}} \big(\|P_k\Phi(t)\|_{L^\infty_x} + 2^{-k}\|P_k\partial_t\Phi(t)\|_{L^\infty_x} \big) \le \epsilon. \tag{4.97}$$

We recall that a nontrivial harmonic map $\Phi^{(\infty)} : \mathbb{R}^2 \to \mathbb{M}$ cannot have an arbitrarily small energy. Precisely, there are two possibilities. Either there are no such harmonic maps (for instance, in the case when $\mathbb{M}$ is negatively curved, see [24]) or there exists a lowest energy nontrivial harmonic map, which we have denoted by $E_{\mathrm{crit}} > 0$. Furthermore, a simple computation shows that the energy of any harmonic map will increase if we apply a Lorentz transformation. Hence, combining the results of Theorem 4.14 and Theorem 4.60 we obtain the following:

Corollary 4.16 (Global regularity for wave maps)**.** *The following statements hold:*

(1) *Assume that $\mathbb{M}$ is a compact Riemannian manifold such that there are no nontrivial finite energy harmonic maps $\Phi^{(\infty)} : \mathbb{R}^2 \to \mathbb{M}$. Then for any finite energy datum $\Phi[0] : \mathbb{R}^2 \times \mathbb{R}^2 \to \mathbb{M} \times T\mathbb{M}$ for the wave map equation there exists a global solution $\Phi \in S$. In addition, this global solution retains any additional regularity of the initial data.*

(2) *Let* $\pi : \tilde{\mathbb{M}} \to \mathbb{M}$ *be a Riemannian covering, with* $\mathbb{M}$ *compact, and such that there are no nontrivial finite energy harmonic maps* $\Phi^{(\infty)} : \mathbb{R}^2 \to \mathbb{M}$*. If* $\Phi[0] : \mathbb{R}^2 \times \mathbb{R}^2 \to \tilde{\mathbb{M}} \times T\tilde{\mathbb{M}}$ *is* C^∞*, then there is a global* C^∞ *solution to* $\tilde{\mathbb{M}}$ *with this datum.*

(3) *Suppose that there exists a lowest energy nontrivial harmonic map into* $\mathbb{M}$ *with energy* E_{crit}*. Then for any datum* $\Phi[0] : \mathbb{R}^2 \times \mathbb{R}^2 \to \mathbb{M} \times T\mathbb{M}$ *for the wave map equation with energy below* E_{crit}*, there exists a global solution* $\Phi \in S$*.*

We remark that the statement in part (2) is a simple consequence of (1) and restricting the projection $\pi \circ \Phi$ to a sufficiently small section S_t of a cone where one expects blowup of the original map into $\tilde{\mathbb{M}}$. In particular, since this projection is regular by part A), its image lies in a simply connected set for sufficiently small t. Thus, this projection can be inverted to yield regularity of the original map close to the suspected blow-up point. Because of this trivial reduction, we work exclusively with compact $\mathbb{M}$ in the sequel. It should be remarked however, that as a (very) special case of this result one obtains global regularity for smooth wave maps into all hyperbolic spaces $\mathbb{H}^n$, which has been a long-standing and important conjecture in geometric wave equations due to its relation with problems in general relativity (see Chapter 16 of [11]).

The statement of Corollary 4.16 in its full generality was known as the *Threshold Conjecture*. Similar results were previously established for the wave map problem via symmetry reductions in the works [12], [35], [41], and [40].

The proof of Theorems 4.14,4.15 are similar, and are outlined in what follows.

Step 1: Extension. Here one constructs an extension for small t in the blow-up problem, respectively for large t in the scattering problem, so that the energy is increased very little, as in (4.93), respectively (4.96).

This argument uses the flux decay in an essential way; this allows us to initiate the extension at a time t_0 where $\not\nabla\Phi$ is very small on the boundary ∂S_{t_0} of the cone, thus guaranteeing the smallness of the energy outside the cone.

By energy estimates, this guarantees that the energy remains small outside the cone up to time zero for the blow-up problem, respectively up to time infinity for the scattering problem. By the small data result, this suffices in order to insure that our extended solution remains regular outside the cone.

Step 2: Energy dispersion and scaling. Here we work with the extensions constructed above. Either they have small energy dispersion, in which case we are done by the energy dispersion result in Theorem 4.7, or not, in which case we have a sequence of points (t_n, x_n) and frequencies k_n with either $t_n \to 0$ or $t_n \to \infty$, so that

$$|P_{k_n}\phi(t_n, x_n)| + 2^{-k_n}|P_{k_n}\partial_t\phi(t_n, x_n)| \geq \epsilon.$$

Using also the flux decay in (4.79) and rescaling t_n to 1, we arrive at a setting where we have the sequence of wave maps

$$\Phi^{(n)}(t,x) = \Phi(t_n t, t_n x)$$

in the increasing regions $C_{[\epsilon_n,1]}$, with $\epsilon_n \to 0$, so that

$$F_{[\epsilon_n,1]}[\Phi^{(n)}] \leq \epsilon_n^{\frac{1}{2}} E[\Phi] \ ,$$

and also points $x_n \in \mathbb{R}^2$ and frequencies $k_n \in \mathbb{Z}$ so that

$$|P_{k_n}\Phi^{(n)}(1,x_n)| + 2^{-k_n}|P_{k_n}\partial_t\Phi^{(n)}(1,x_n)| > \epsilon \ . \tag{4.98}$$

From this point on, the proofs of Theorems 4.14 and 4.15 are identical.

Step 3: Elimination of null concentration scenario. Using the fixed time portion of the X_0 energy bounds in (4.89) we eliminate the case of null concentration

$$|x_n| \to 1, \quad k_n \to \infty$$

in estimate (4.98), and show that the sequence of maps $\Phi^{(n)}$ at time $t=1$ must either have low frequency concentration in the range

$$m(\epsilon,E) < k_n < M(\epsilon,E), \quad |x_n| < R(\epsilon,E)$$

or high frequency concentration strictly inside the cone:

$$k_n \geq M(\epsilon,E), \quad |x_n| < \gamma(\epsilon,E) < 1.$$

Step 4: Time-like energy concentration. In both remaining cases above we show that a nontrivial portion of the energy of $\Phi^{(n)}$ at time 1 must be located inside a smaller cone:

$$\frac{1}{2}\int_{t=1,|x|<\gamma_1} \left(|\partial_t\Phi^{(n)}|^2 + |\nabla_x\Phi^{(n)}|^2\right)dx \geq E_1,$$

where $E_1 = E_1(\epsilon,E)$ and $\gamma_1 = \gamma_1(\epsilon,E) < 1$.

Step 5: Uniform propagation of nontrivial time-like energy. Using again the X_0 energy bounds as in Lemma 4.13 we propagate the above time-like energy concentration for $\Phi^{(n)}$ from time 1 to smaller times $t \in [\epsilon_n^{\frac{1}{2}}, \epsilon_n^{\frac{1}{4}}]$,

$$\frac{1}{2}\int_{|x|<\gamma_2(\epsilon,E)t} \left(|\partial_t\Phi^{(n)}|^2 + |\nabla_x\Phi^{(n)}|^2\right)dx \geq E_0(\epsilon,E), \quad t \in [\epsilon_n^{\frac{1}{2}}, \epsilon_n^{\frac{1}{4}}].$$

At the same time, we obtain bounds for $X_0\Phi^{(n)}$ outside smaller and smaller neighborhoods of the cone, namely

$$\int_{C^{\epsilon_n}_{[\epsilon_n^{\frac{1}{2}},\epsilon_n^{\frac{1}{4}}]}} \rho^{-1}|X_0\Phi^{(n)}|^2 dxdt \lesssim 1.$$

Step 6: Final rescaling. By a pigeonhole argument and rescaling we end up producing another sequence of maps, still denoted by $\Phi^{(n)}$, which are sections the original wave map Φ and are defined in increasing regions $C_{[1,T_n]}$, $T_n = e^{|\ln \epsilon_n|^{\frac{1}{2}}}$, and satisfy the following three properties:

$$
\begin{aligned}
E_{S_t}[\Phi^{(n)}] \approx E, \quad t \in [1, T_n] \qquad & \text{(Bounded Energy)},\\
E_{S_t^{(1-\gamma_2)t}}[\Phi^{(n)}] \geq E_2, \quad t \in [1, T_n] \qquad & \text{(Nontrivial Time-like Energy)},\\
\int_{C^{\epsilon_n^{\frac{1}{2}}}_{[1,T_n]}} \frac{1}{\rho} |X_0 \Phi^{(n)}|^2 dxdt \lesssim |\log \epsilon_n|^{-\frac{1}{2}} \qquad & \text{(Decay to Self-similar Mode)}.
\end{aligned}
$$

Step 7: Isolating the concentration scales. Using several additional pigeonholing arguments we show that one of the following two scenarios must occur:

(1) (Energy Concentration) On a subsequence there exist $(t_n, x_n) \to (t_0, x_0)$, with (t_0, x_0) inside $C^{\frac{1}{2}}_{[\frac{1}{2},\infty)}$, and scales $r_n \to 0$ so that we have

$$
\begin{aligned}
E_{B(x_n,r_n)}[\Phi^{(n)}](t_n) &= \frac{1}{10} E_0,\\
E_{B(x,r_n)}[\Phi^{(n)}](t_n) &\leq \frac{1}{10} E_0, \quad x \in B(x_0, r),\\
r_n^{-1} \int_{t_n - r_n/2}^{t_n + r_n/2} \int_{B(x_0,r)} &|X_0 \Phi^{(n)}|^2 dxdt \to 0.
\end{aligned}
$$

(2) (Non-concentration) For each $j \in \mathbb{N}$ there exists an $r_j > 0$ such that for every (t, x) inside $C_j = C^1_{[1,\infty)} \cap \{2^j < t < 2^{j+1}\}$ one has

$$
\begin{aligned}
& E_{B(x,r_j)}[\Phi^{(n)}](t) \leq \frac{1}{10} E_0 \ , \qquad \forall (t,x) \in C_j \ ,\\
& E_{S_t^{(1-\gamma_2)t}}[\Phi^{(n)}](t) \geq E_2 \ ,\\
& \int_{C_j} |X_0 \Phi^{(n)}|^2 dxdt \to 0 \ .
\end{aligned}
$$

uniformly in n.

Here E_0 represents the threshold in the small data result.

Step 8: The compactness argument. In case i) above we consider the rescaled wave maps

$$
\Psi^{(n)}(t, x) = \Phi^{(n)}(t_n + r_n t, x_n + r_n x)
$$

and show that on a subsequence they converge locally in the energy norm to a finite energy nontrivial wave map Ψ in $\mathbb{R}^2 \times [-\frac{1}{2}, \frac{1}{2}]$ which satisfies $X(t_0, x_0)\Psi = 0$. Thus Ψ must be a Lorentz transform of a nontrivial harmonic map.

In case ii) above we show directly that the sequence $\Phi^{(n)}$ converges locally on a subsequence in the energy norm to a finite energy nontrivial wave map Ψ, defined in the interior of a translated cone $C^2_{[2,\infty)}$, which satisfies $X_0\Psi = 0$. Consequently, in hyperbolic coordinates we may interpret Ψ as a nontrivial harmonic map

$$\Psi : \mathbb{H}^2 \to M \ .$$

Compactifying this and using conformal invariance, we obtain a nontrivial finite energy harmonic map

$$\Psi : \mathbb{D}^2 \to M$$

from the unit disk $\mathbb{D}^2$, which according to the estimates of Section 4.7 obeys the additional weighted energy bound

$$\int_{\mathbb{D}^2} |\nabla_x \Phi|^2 \frac{dx}{1-r} < \infty.$$

But such maps do not exist due to a combination of a theorem of Qing [29] and a theorem of Lemaire [24].

4.9 Further developments

We begin with some comments concerning the higher dimensional case. First of all, we remark that, while not explicitly proved in [37], the result in Theorem 4.7 extends to higher dimensions with no change other that the role of the energy is played by the critical Sobolev norm of the initial data. However, the analogue of Theorem 3.7 is not true as stated. Instead we have the following

Open Problem 4.17. *Consider wave maps in dimension $n \geq 3$ with uniformly bounded critical Sobolev norms.*

a) *Identify all possible concentration scenarios (at the very least, this must include solitons and self-similar solutions).*

b) *Establish a dichotomy, as in Theorem 3.7, between energy dispersion and concentration scenarios.*

Next we return to the two dimensional case. In the results above we have considered solutions below the ground state energy. But what happens if we take data with size slightly above the ground state energy? For simplicity we will discuss the special case of maps from $\mathbb{R}^{2+1}$ into $(M, g) = \mathbb{S}^2$. There we have the harmonic maps Q_k which are the unique energy minimizers in their homotopy class modulo symmetries[3]. Recall that the ground state $Q = Q_1$ is the stereographic projection.

Consider the wave map equation with data which are close in the energy norm to Q_k. Such data must be in the same homotopy class as Q_k, and the

[3]Namely, isometries of $\mathbb{R}^2$, rotations of $\mathbb{S}^2$, and scaling.

corresponding solution stays there as long as no blow up occurs. Then, due to energy conservation, we conclude that the ground states are orbitally stable, i.e., the solution must stay close to Q_k modulo symmetries. However, this does not lead to a global result since the group of symmetries is noncompact. Precisely, it is the scaling that generates the noncompactness[4] and may lead to blow up.

A natural simplification is to look at equivariant solutions. Then all other components of the symmetry group are eliminated, and we are left only with scaling. Thus we are looking at solutions of the form

$$\phi = Q_k(\lambda r) + o(1), \tag{4.99}$$

where λ is some function of t. Blow up at time t_0 would correspond to $\lambda(t) \to \infty$ as $t \to t_0$. Blow-up solutions have been proved to exist:

Theorem 4.18 (Krieger, Schlag, and Tataru [21])**.** *Let $k = 1$. Then there exist equivariant blow-up solutions with the concentration rate*

$$\lambda(t) = t^{-\nu-1}, \quad \nu > 1. \tag{4.100}$$

Theorem 4.19 (Rodnianski and Sterbenz [34], Raphael and Rodnianski [32])**.** *Let $k \geq 1$. Then there exist equivariant blow-up solutions with the concentration rate*

$$\lambda(t) = t^{-1} |\log t|^{-\frac{1}{2k-2}}, \quad k \geq 2, \tag{4.101}$$

$$\lambda(t) = e^{c\sqrt{|\log t|}}, \quad k = 1. \tag{4.102}$$

We expect the first result to be true for all $\nu > 0$. The second result seems to be in some sense an extreme case. The proof of these results is strongly related to the linearized wave map flow around the ground states Q_k. There is a fundamental difference between the case $k = 1$ and $k \geq 2$. In the latter case, the linearized elliptic operator has a zero eigenvalue, which is the source of instability. In the former case, we have instead a zero resonance, which still leads to instability, but in a more subtle way. A natural follow-up problem would be

Open Problem 4.20. *Classify all possible blow-up rates in the equivariant case, and study their stability.*

Is blow up a generic phenomenon or an atypical one ? The knowledge that we have so far seems to indicate that the following may be plausible at least for $k = 1$:

Conjecture 4.21. *Consider the equivariant wave map equation with data near the soliton Q_1. Then there exists a codimension one stable manifold of data separating the data set into two components, as follows:*

---

[4]The spatial translations are another source of noncompactness, but cannot lead to blow up because of the finite speed of propagation.

a) *Data in one component leads to a shrinking soliton and to finite time blow up.*

b) *Data in the other component leads to an expanding soliton.*

This picture may require small adjustments as more data becomes available. As in initial step, in recent work [5] we are able to construct a codimension two stable manifold.

It would also be very interesting to consider non-equivariant data:

Open Problem 4.22. *Classify all possible blow ups for wave maps $\phi : \mathbb{R}^{2+1} \to \mathbb{S}^2$ with data near the ground state Q, in terms of a description akin to* (4.99)*, but with scaling replaced by all the symmetries, and with good asymptotics for the symmetry group parameters as functions of t near the blow-up time.*

Chapter 5

Schrödinger maps

Here we consider Schrödinger maps $\phi : \mathbb{R} \times \mathbb{R}^n \to \mathbb{S}^2$, $n \geq 2$, and prove the small data result in Theorem 3.8. We recall that in n space dimensions the initial data belongs to the space $\dot{H}^{\frac{n}{2}}$. To keep the notations simple we will confine the discussion to the energy critical case $n = 2$; this is also the most difficult case. Beside the finite energy condition, it is technically convenient to assume that for some $Q \in \mathbb{S}^2$ we have

$$M(\phi) = \int_{\mathbb{R}^2} |\phi - Q|^2 \, dx < \infty$$

This is a conserved quantity. The size of M plays no role in any of the estimates. Its only purpose is to insure convergence to the constant state Q along the harmonic heat flow; this in turn is used in the construction of the caloric gauge. The use of this condition can be bypassed, but that is not pursued here.

5.1 Frames and gauges

The formulation we adopt for this problem uses the frame method. At each point (t, x) we consider an orthonormal frame (v, w) in $T_{\phi(t,x)}\mathbb{S}^2$, and use the complex representation of tangent vectors $X \to \langle X, v\rangle + i\langle X, w\rangle$. In particular we can express $\partial_m \phi$ in the (v, w) frame as

$$\psi_m = v \cdot \partial_m \phi + i \; w \cdot \partial_m \phi. \tag{5.1}$$

Here $m = 1, \dots, n, n+1$ and ψ_{n+1} corresponds to the time variable.

Given the frame coefficients

$$A_m = w \cdot \partial_m v, \tag{5.2}$$

we define the covariant differentiation operators

$$\mathbf{D}_m = \partial_m + iA_m.$$

H. Koch et al., *Dispersive Equations and Nonlinear Waves: Generalized Korteweg–de Vries, Nonlinear Schrödinger, Wave and Schrödinger Maps*, Oberwolfach Seminars 45,
DOI 10.1007/978-3-0348-0736-4_14,

The differentiated variables ψ_k are subject to the compatibility conditions

$$\mathbf{D}_m \psi_k = \mathbf{D}_k \psi_m, \tag{5.3}$$

while the connection A_k satisfies the curvature conditions

$$\partial_l A_m - \partial_m A_l = \Im(\psi_l \overline{\psi_m}) = q_{lm}. \tag{5.4}$$

A direct computation shows that the Schrödinger map equation expressed in terms of the differentiated fields reads

$$\psi_{n+1} = i \sum_{l=1}^{n} \mathbf{D}_l \psi_l. \tag{5.5}$$

Using (5.3) and (5.4), it follows that for $m = 1, \dots, n$ we have

$$\mathbf{D}_t \psi_m = i \sum_{l=1}^{n} \mathbf{D}_l \mathbf{D}_l \psi_m + \sum_{l=1}^{d} q_{lm} \psi_l, \tag{5.6}$$

which is equivalent to

$$(i\partial_t + \Delta_x)\psi_m = -2i \sum_{l=1}^{n} A_l \partial_l \psi_m + \Big(A_{d+1} + \sum_{l=1}^{n} (A_l^2 - i\partial_l A_l)\Big)\psi_m - i \sum_{l=1}^{n} \psi_l \Im(\overline{\psi}_l \psi_m). \tag{5.7}$$

To view this as a self-contained system, we need to make a gauge choice, which would uniquely determine the A_j's in terms of the ψ_j's. Ideally, we would like to have a gauge that would make the right-hand side of the above equation perturbative. The analogy we have in mind here is with the cubic NLS problem. Indeed, in view of the relations (5.4), it is reasonable to assume that the A_j's are quadratic and higher order in ψ, therefore the right-hand side above will only contain terms which are cubic and higher order.

The main difficulty primarily originates with the term

$$A_l \partial_l \psi_m,$$

which has an unfavorable balance of derivatives. Consider for instance the simplest gauge, namely the Coulomb gauge, which yields an expression of the form

$$D^{-1}(\psi\bar{\psi})D\psi$$

This causes some difficulties in the case of high $\times$ high $\to$ low interactions in the first factor; these can be resolved in high dimension ($n \geq 4$, see [3]), but the singularity at frequency zero is too strong in two and three dimensions.

This is what causes us to look for a different choice of gauge which avoids the above difficulty. A reason to hope that such a gauge might exist is given by the exact form of the right-hand side in the equations (5.4). Precisely, the functions ψ_l and ψ_m there are not independent, instead they are connected via (5.3). This indicates that to the leading order, the expression $\psi_l\bar{\psi}_m$ is real when the two factors have equal frequencies. Such a cancellation is not at all captured by the Coulomb gauge. As it turns out, there is indeed a more favorable gauge, namely the caloric gauge. This was proposed in [42] in the context of the wave map equation, and then as a possible gauge for Schrödinger maps.

Precisely, at each time t we solve the harmonic heat flow equation with $\phi(t)$ as the initial data:

$$\begin{cases} \partial_s\tilde{\phi} = \Delta_x\tilde{\phi} + \tilde{\phi}\cdot\sum_{m=1}^n |\partial_m\tilde{\phi}|^2 & \text{on } [0,\infty)\times\mathbb{R}^d, \\ \tilde{\phi}(0,t,x) = \phi(t,x). \end{cases} \tag{5.8}$$

We note that the Schrödinger map and the harmonic heat flow do not commute. Thus, the Schrödinger map equation is only valid at $s=0$, and not for larger s.

We heuristically remark that as the heat time s approaches infinity, the solution $\phi(s)$ approaches the equilibrium state Q. This is related to our assumption that the "mass" of ϕ_0 is finite, and would not necessarily be true otherwise. This allows us to arbitrarily pick (v_∞, w_∞) at $s=\infty$ as an orthonormal basis in $T_Q\mathbb{S}^2$, independently of t and x. To define the orthonormal frame (v,w) for all $s\geq 0$ we pull back (v_∞, w_∞) along the backward heat flow using parallel transport. This translates into the relation

$$w\cdot\partial_s v = 0. \tag{5.9}$$

Setting $\partial_0 = \partial_s$, we define the functions ψ_m and A_m for all $s\in[0,\infty)$ and $m=0,\dots,d+1$ by

$$\begin{cases} \psi_m = v\cdot\partial_m\tilde{\phi} + i\, w\cdot\partial_m\tilde{\phi}, \\ A_m = w\cdot\partial_m v. \end{cases} \tag{5.10}$$

In addition, the parallel transport relation $w\cdot\partial_s v = 0$ yields the main gauge condition

$$A_0 = 0. \tag{5.11}$$

As in the case of the Schrödinger equation, a direct computation using the heat equation (5.8) and (5.3), (5.4) shows that

$$\psi_0 = \sum_{l=1}^d \mathbf{D}_l\psi_l. \tag{5.12}$$

Thus, using again (5.4), for any $m=1,\dots,d+1$,

$$\begin{aligned} \partial_0\psi_m = \mathbf{D}_m\psi_0 &= \sum_{l=1}^d \mathbf{D}_m\mathbf{D}_l\psi_l = \sum_{l=1}^d \mathbf{D}_l\mathbf{D}_m\psi_l + i\sum_{l=1}^d q_{ml}\psi_l \\ &= \sum_{l=1}^d \mathbf{D}_l\mathbf{D}_l\psi_m + i\sum_{l=1}^d \Im(\psi_m\overline{\psi_l})\psi_l, \end{aligned}$$

which is equivalent to

$$(\partial_s - \Delta_x)\psi_m = 2i\sum_{l=1}^{d} A_l\partial_l\psi_m - \sum_{l=1}^{d}(A_l^2 - i\partial_l A_l)\psi_m + i\sum_{l=1}^{d}\Im(\psi_m\overline{\psi_l})\psi_l. \quad (5.13)$$

On the other hand, from (5.4) we obtain

$$\partial_s A_m = \Im(\psi_0\overline{\psi_m}).$$

Then we can integrate back from $s=\infty$ to obtain

$$A_m(s) = -\int_s^\infty \Im(\psi_0\overline{\psi_m})(r)\,dr = -\sum_{l=1}^{n}\int_s^\infty \Im\big(\overline{\psi_m}(\partial_l\psi_l + iA_l\psi_l)\big)(r)\,dr, \quad (5.14)$$

for any $m = 1,\dots,d+1$ and $s\in[0,\infty)$. Thus $A_m|_{s=0}$ represents our choice of the gauge for the Schrödinger map equation. The reason we prefer the caloric gauge to the Coulomb gauge is the way the high-high frequency interactions are handled. Indeed, while in the Coulomb gauge the connection coefficients can be conceptually written in the form

$$A \approx \sum_{j<k} 2^{-k}P_j\psi P_k\psi + \sum_{j\le k} 2^{-j}P_j(P_k\psi P_k\psi),$$

substituting the first approximation $\psi(s)\approx e^{s\Delta}\psi(0)$ in (5.14) yields the relation

$$A \approx \sum_{j<k} 2^{-k}P_j\psi P_k\psi + \sum_{j\le k} 2^{-k}P_j(P_k\psi P_k\psi). \quad (5.15)$$

This has a better frequency factor in the high $\times$ high $\to$ low frequency interactions.

5.2 Function spaces

To motivate our choice of spaces, recall the Schrödinger nonlinearities, see (5.7)

$$L_m = -2i\sum_{l=1}^{d} A_l\partial_l\psi_m + \Big(A_{d+1} + \sum_{l=1}^{d}(A_l^2 - i\partial_l A_l)\Big)\psi_m - i\sum_{l=1}^{d}\psi_l\Im(\overline{\psi_l}\psi_m). \quad (5.16)$$

We would like to analyze these nonlinearities perturbatively in suitable spaces. The main difficulty is caused by the magnetic terms $-2i\sum_{l=1}^{d} A_l\partial_l\psi_m$. Using (5.15) (for simplicity consider only the terms corresponding to $k=j$) they can be written schematically in the form

$$\sum_{k,k'\in\mathbb{Z}} 2^{-k}P_k\psi P_k\psi\cdot 2^{k'}P_{k'}\psi. \quad (5.17)$$

If $k > k'$, then this is a Strichartz type term, but if $k < k'$, then we need to recover a full derivative at frequency k'. The way to do that is by using the lateral energy spaces $L_e^{\infty,2}$ associated to Schrödinger waves with a suitable angular localization in a lateral frame with direction e. These, and more generally the $L_e^{p,q}$ spaces, are defined as

$$L_e^{p,q} = L_{x_e}^p L_{t,x_e'}^q,$$

where $(x_e = x \cdot e, x_e')$ is the orthogonal frame associated to the direction e.

Then the above expression (5.17) needs to be estimated in a dual space $L_e^{1,2}$. For this to work it would appear that we need to bound $P_k\psi$ in $L_e^{2,\infty}$. This estimate is valid in dimensions three and higher. However, in two space dimensions this is precisely the forbidden endpoint of the (lateral) Strichartz estimates.

Nevertheless, the corresponding L^2 bilinear estimate for free Schrödinger waves is valid:

$$\|\psi_k \psi_{k'}\|_{L^2} \lesssim 2^{\frac{k-k'}{2}} \|\psi_k(0)\|_{L^2} \|\psi_{k'}(0)\|_{L^2}, \quad k < k'.$$

This suggests that there might be a way to still close by more subtle adjustments to the function spaces. The key observation which allows us to fix the above argument in two space dimensions is that in the lateral energy spaces $L_e^{\infty,2}$ used at frequency k' we are free to add Galilean transformations T_v as long as $|v| \ll 2^{k'}$. Here

$$T_v \phi(x,t) = e^{-i(\frac{1}{2}xv + \frac{1}{4}|v|^2 t)} \phi(x + vt, t).$$

In other words, we can set

$$\|\phi\|_{L_{e,v}^{p,q}} = \|T_v \phi\|_{L_e^{p,q}}$$

and work with the smaller space

$$\bigcap_{|v| \ll 2^{k'}} L_{e,v}^{\infty,2}.$$

This would allow us to relax the bound for $P_k\psi$ to the space

$$\sum_{|v| \approx 2^k} L_{e,v}^{2,\infty}.$$

This strategy actually works. Furthermore, we do not need to use all such v, it suffices to restrict our attention to those which are parallel to e. In addition, by restricting time to a large but finite interval, we can discretize the above continuous set of v's. Precisely, for large $\mathcal{K}$ we restrict time to $t \in [0, 2^{2\mathcal{K}}]$, and then define the set of indices

$$W_k = W_k(\mathcal{K}) = \{\lambda \in [-2^k, 2^k] : 2^{k+2\mathcal{K}}\lambda \in \mathbb{Z}\}.$$

and the associated space

$$L^{2,\infty}_{e,W_k} = \sum_{v \in eW_k} L^{2,\infty}_{e,v}.$$

To use these spaces we need the projectors $P_{k,e}$ which select the region $|\xi \cdot e| \approx 2^k$. Then we have the following

Lemma 5.1. *Let $d = 2$. For any $f \in L^2$, $k \in \mathbb{Z}$, and $e \in \mathbb{S}^1$ we have*

$$\|e^{it\Delta} P_{k,e} f\|_{L^{\infty,2}_{e,v}} \lesssim 2^{-k/2} \|f\|_{L^2}, \qquad |v| \ll 2^k. \tag{5.18}$$

In addition, if $T \in (0, 2^{2\mathcal{K}}]$, then

$$\|1_{[-T,T]}(t) e^{it\Delta} P_k f\|_{L^{2,\infty}_{e,W_{k+40}}} \lesssim 2^{k/2} \|f\|_{L^2}. \tag{5.19}$$

Proof. We begin with (5.18). After a Galilean transformation the problem reduces to the case $v = 0$, where by translation invariance it suffices to estimate

$$\|u\|_{L^2_{t,x'}} \lesssim 2^{-\frac{k}{2}} \|f\|_{L^2}, \quad u = e^{it\Delta} P_{k,e} f(t, 0, x'_e).$$

Without any restriction in generality we assume that $P_{k,e}$ is confined to the positive side $\xi \cdot e \approx 2^k$ (and not -2^k). Then a direct computation shows that

$$\hat{u}(\tau, \xi'_e) = \frac{1}{2\xi_e} p_{k,e}(\xi) \hat{f}(\xi), \quad \tau = \xi^2, \quad \xi_e > 0.$$

Hence (5.18) follows by a simple change of coordinates in the integral defining the L^2 norm.

Next we prove (5.19). For that we define two more classes of spaces. Given a finite subset $W \subseteq \mathbb{R}$ and $r \in [1, \infty]$, we define the spaces $\sum^r L^{p,q}_{e,W}$ and $\bigcap^r L^{p,q}_{e,W}$ using the norms

$$\|\phi\|^r_{\sum^r L^{p,q}_{e,W}} = |W|^{r-1} \inf_{\phi = \sum_{\lambda \in W} \phi_\lambda} \sum_{\lambda \in W} \|\phi_\lambda\|^r_{L^{p,q}_{e,\lambda}} \tag{5.20}$$

and

$$\|\phi\|^r_{\bigcap^r L^{p,q}_{e,W}} = |W|^{-1} \sum_{\lambda \in W} \|\phi\|^r_{L^{p,q}_{e,\lambda}}. \tag{5.21}$$

Clearly, $\sum^1 L^{p,q}_{e,W} = L^{p,q}_{e,W}$ and

$$\|\phi\|_{\sum^r L^{p,q}_{e,W}} \leq \|\phi\|_{\sum^{r'} L^{p,q}_{e,W}} \quad \text{if } r \leq r'. \tag{5.22}$$

We fix $e \in \mathbb{S}^1$. By rescaling we can assume that $\mathcal{K} = 0$. We may also assume that $k \geq 1$, since for $k \leq 0$ one has the stronger bound

$$\|\mathbf{1}_{[-1,1]}(t) e^{it\Delta} P_k f\|_{L^2_x L^\infty_t} \lesssim \|f\|_{L^2}.$$

We need to show that

$$\|1_{[-1,1]}(t)e^{it\Delta}P_k f\|_{\sum^2 L^{2,\infty}_{e,W_{k+5}}} \lesssim 2^{k/2}\|f\|_{L^2}. \tag{5.23}$$

Due to the duality relation[1]

$$\left(\bigcap^2 L^{2,1}_{e,W_{k+5}}\right)' = \sum^2 L^{2,\infty}_{e,W_{k+5}}.$$

it suffices to show that if $\|g\|_{\bigcap^2 L^{2,1}_{e,W_{k+5}}} \leq 1$, then

$$\Big|\int_{\mathbb{R}^2\times\mathbb{R}} \overline{g(x,t)}\mathbf{1}_{[-1,1]}(t)(e^{it\Delta}P_k f)(x,t)\,dxdt\Big| \lesssim 2^{k/2}\|f\|_{L^2}. \tag{5.24}$$

This can be rewritten as

$$\Big|\int_{\mathbb{R}^2\times\mathbb{R}} \overline{(e^{-it\Delta}P_k g(t))(x)}\mathbf{1}_{[-1,1]}(t)f(x)\,dtdx\Big| \lesssim 2^{k/2}\|f\|_{L^2},$$

or, equivalently,

$$\Big\|\int_{-1}^{1} e^{-it\Delta}P_k g(t)\Big\|_{L^2} \lesssim 2^{k/2}.$$

Hence by a TT^* argument it suffices to show that

$$\Big|\int_{\mathbb{R}^2\times\mathbb{R}}\int_{\mathbb{R}^2\times\mathbb{R}} g(x,t)\mathbf{1}_{[-1,1]}(t)\overline{g}(y,s)\mathbf{1}_{[-1,1]}(s)K_k(x-y,t-s)\,dxdtdyds\Big| \lesssim 2^k, \tag{5.25}$$

where

$$K_k(x,t) = \int_{\mathbb{R}^2} e^{ix\cdot\xi}e^{-it|\xi|^2}\chi_k(|\xi|)^2\,d\xi. \tag{5.26}$$

By stationary phase

$$|K_k(t,x)| \lesssim \begin{cases} 2^{2k}(1+2^{2k}|t|)^{-1}, & \text{if } |x| \leq 2^{k+4}|t|, \\ 2^{2k}(1+2^k|x|)^{-N}, & \text{if } |x| \geq 2^{k+4}|t|. \end{cases}$$

The key idea is to foliate K_k in the e direction with respect to (thickened) rays with speed less than 2^{k+5}. We observe that for $t\in[-2,2]$

$$|K_k(t,x)| \lesssim \sum_{\lambda\in W_{k+5}} K_{k,\lambda}(t,x), \quad K_{k,\lambda}(t,x) = (1+2^k|x\cdot e-\lambda t|)^{-N}.$$

[1] This is not entirely straightforward.

Hence the left hand side of (5.25) can be bounded by

$$\sum_{\lambda\in W_{k+5}} \int_{-1}^{1}\int_{-1}^{1} K_{k,\lambda}(t-s,x-y)|g(y,s)||g(x,t)|dxdydsdt$$
$$\lesssim \sum_{\lambda\in W_{k+5}} \|K_{k,\lambda}\|_{L^{1,\infty}_{e,\lambda}} \|g\|_{L^{2,1}_{e,\lambda}} \|g\|_{L^{2,1}_{e,\lambda}} \lesssim 2^{-k} \sum_{\lambda\in W_{k+5}} \|g\|^2_{L^{2,1}_{e,\lambda}} \lesssim 2^{k}\|g\|^2_{\bigcap^2 L^{2,1}_{e,W_{k+5}}},$$

where we used the fact that $|W_{k+5}| \approx 2^{2k}$. Thus (5.24) follows. □

We are now ready to define the dyadic function spaces where we want to study the equation (5.7). We will denote by G_k the spaces for the solutions ψ_m and by N_k the spaces for the right hand side L_m. Heuristically the G_k norms should contain Strichartz type norms, plus the above $\bigcap_{|v|\ll 2^k} L^{\infty,2}_{e,v}$ and the sum space $L^{\infty,2}_{e,W_k}$.

One difficulty we encounter is that the norms of nearby G_k's are not equivalent, and that makes it difficult to propagate them along the harmonic heat flow. For this reason we introduce a third space F_k with a weaker topology than G_k, $G_k \subset F_k$, but which does vary nicely with respect to k.

For comparison purposes, we also provide the corresponding definitions in dimensions three and higher.

Definition 5.2. *Assume $n \geq 3$ and $k \in \mathbb{Z}$. Then $F_k(T)$, $G_k(T)$, and $N_k(T)$ are the Banach spaces of functions localized at frequency 2^k for which the corresponding norms are finite:*

$$\|\phi\|_{F_k(T)} = \|\phi\|_{L^\infty_t L^2_x} + \|\phi\|_{L^{p_d}} + 2^{-kd/(d+2)}\|\phi\|_{L^{p_d}_x L^\infty_t} + 2^{-k(d-1)/2} \sup_{e\in\mathbb{S}^{d-1}} \|\phi\|_{L^{2,\infty}_e}, \tag{5.27}$$

$$\|\phi\|_{G_k(T)} = \|\phi\|_{F_k} + 2^{k/2} \sup_{|j-k|\leq 20} \sup_{e\in\mathbb{S}^{d-1}} \|P_{j,e}\phi\|_{L^{\infty,2}_e}, \tag{5.28}$$

and

$$\|f\|_{N_k(T)} = \inf_{f=f_1+f_2} \left(\|f_1\|_{L^{p'_d}} + 2^{-k/2} \sup_{e\in\mathbb{S}^{d-1}} \|f_2\|_{L^{1,2}_e} \right). \tag{5.29}$$

Definition 5.3. *Assume that $n = 2$, $k \in \mathbb{Z}$, $\mathcal{K} \in \mathbb{Z}_+$, and $T \in (0, 2^{2\mathcal{K}}]$. For functions ϕ at frequency 2^k let*

$$\|\phi\|_{F^0_k(T)} = \|\phi\|_{L^\infty_t L^2_x} + \|\phi\|_{L^4} + 2^{-k/2}\|\phi\|_{L^4_x L^\infty_t} + 2^{-k/2} \sup_{e\in\mathbb{S}^1} \|\phi\|_{L^{2,\infty}_{e,W_{k+40}}}. \tag{5.30}$$

We define $F_k(T)$, $G_k(T)$, and $N_k(T)$ as the spaces of functions for which the corresponding norms are finite:

$$\|\phi\|_{F_k(T)} = \inf_{J,m_1,\dots,m_J\in\mathbb{Z}_+} \inf_{f=f_{m_1}+\dots+f_{m_J}} \sum_{j=1}^{J} 2^{m_j}\|f_{m_j}\|_{F^0_{k+m_j}}, \tag{5.31}$$

$$\begin{aligned}\|\phi\|_{G_k(T)} = \|\phi\|_{F_k^0} + 2^{-k/6}\sup_{e\in\mathbb{S}^1}\|\phi\|_{L_e^{3,6}} + 2^{k/6}\sup_{|j-k|\le 20}\sup_{e\in\mathbb{S}^1}\|P_{j,e}\phi\|_{L_e^{6,3}}\\ + 2^{k/2}\sup_{|j-k|\le 20}\sup_{e\in\mathbb{S}^1}\sup_{|\lambda|<2^{k-40}}\|P_{j,e}\phi\|_{L_{e,\lambda}^{\infty,2}},\end{aligned}\tag{5.32}$$

and

$$\begin{aligned}&\|f\|_{N_k(T)}\\ &= \inf_{f=f_1+f_2+f_3+f_4}\left(\|f_1\|_{L^{\frac{4}{3}}} + 2^{\frac{k}{6}}\|f_2\|_{L_{e_1}^{\frac{3}{2},\frac{6}{5}}} + 2^{\frac{k}{6}}\|f_3\|_{L_{e_2}^{\frac{3}{2},\frac{6}{5}}} + 2^{-\frac{k}{2}}\sup_{e\in\mathbb{S}^1}\|f_4\|_{L^{1,2}_{e,W_{k-40}}}\right),\end{aligned}\tag{5.33}$$

where (e_1, e_2) is the canonical basis in $\mathbb{R}^2$.

In all dimensions $d \ge 2$ the spaces $N_k(T)$ and $G_k(T)$ are related by the following linear estimate:

Proposition 5.4 (Main linear estimate)**.** *Assume $\mathcal{K} \in \mathbb{Z}_+$, $T \in (0, 2^{2\mathcal{K}}]$, and $k \in \mathbb{Z}$. Then for each $u_0 \in L^2$ which is localized at frequency 2^k and any $h \in N_k(T)$, the solution u to*

$$(i\partial_t + \Delta_x)u = h, \quad u(0) = u_0$$

satisfies

$$\|u\|_{G_k(T)} \lesssim \|u(0)\|_{L_x^2} + \|h\|_{N_k(T)}.$$

To bound products of functions in $F_k(T)$ we often use a more relaxed criterion. Precisely, since for $e \in \mathbb{S}^1$ and f localized at frequency 2^k we have

$$\|f\|_{L^{2,\infty}_{e,W_{k+m_j}}} \le \|f\|_{L_e^{2,\infty}} \lesssim 2^{k(d-1)/2}\|f\|_{L_x^2 L_t^\infty},$$

it follows that, in all dimensions $d \ge 2$,

$$\|f\|_{F_k(T)} \lesssim \|f\|_{L_x^2 L_t^\infty} + \|f\|_{L^{p_d}}. \tag{5.34}$$

This criterion is often used to estimate bilinear expressions, by exploiting the $L_x^{p_d} L_t^\infty$ norms in the spaces $F_k(T)$.

We also need to evolve $F_k(T)$ functions along the heat flow. Since the $F_k(T)$ norm is translation invariant, it immediately follows that if $h \in F_k(T)$ then

$$\|e^{s\Delta_x}h\|_{F_k(T)} \lesssim (1 + s2^{2k})^{-20}\|h\|_{F_k(T)}, \quad s \ge 0. \tag{5.35}$$

To prove useful bounds on the connection coefficients A_m, $m = 1, \ldots, d$, for $k \in \mathbb{Z}$ and $\omega \in [0, 1/2]$ we define the normed spaces $S_k^\omega(T)$ of functions in $L_k^2(T)$ for which

$$\|f\|_{S_k^\omega(T)} = 2^{k\omega}(\|f\|_{L_t^\infty L_x^{2\omega}} + \|f\|_{L_t^{p_d} L_x^{p_{d,\omega}}} + 2^{-kd/(d+2)}\|f\|_{L_x^{p_{d,\omega}} L_t^\infty}) < \infty, \tag{5.36}$$

where the exponents 2_ω and $p_{d,\omega}$ are such that

$$\frac{1}{2_\omega} - \frac{1}{2} = \frac{1}{p_{d,\omega}} - \frac{1}{p_d} = \frac{\omega}{d}.$$

The spaces $S_k^\omega(T)$ are at the same scale as the spaces $F_k(T)$ and $F_k(T) \hookrightarrow S_k^0(T)$. By Sobolev embeddings we have

$$\|f\|_{S_k^{\omega'}(T)} \lesssim \|f\|_{S_k^\omega(T)} \quad \text{if } \omega' \leq \omega. \tag{5.37}$$

Thus the spaces $S_k^\omega(T)$ can be interpreted as refinements of the Strichartz part of the spaces $F_k(T)$ (which corresponds to $S_k^0(T)$). It is important to be able to prove bounds on the coefficients A_m, $m = 1, \dots, d$, in both spaces $F_k(T)$ and $S_k^{1/2}(T)$. These bounds quantify an essential gain of smoothness of the coefficients A_m compared to the fields ψ_m.

5.3 The small data result

Here we outline the main steps in the proof of the small data result for Schrödinger maps in Theorem 3.8.

5.3.1 Bounds for the harmonic heat flow

We begin with the L^2 bounds for the harmonic heat flow. Below we state them for small data, but by the work of Smith [36] similar results hold up to the critical energy E_{crit}. For the next result we fix the Schrödinger time:

Proposition 5.5 (Construction of the caloric gauge)**.** *Let $\phi : \mathbb{R}^n \to \mathbb{S}^2$ with $\phi - Q \in L^2$ which satisfies the smallness condition*

$$\|\phi\|_{\dot{H}^{\frac{n}{2}}} = \gamma^2 \ll 1. \tag{5.38}$$

Let c_k be a frequency envelope for ϕ. Then there is a unique smooth solution $\tilde{\phi} \in C^\infty((0,\infty) \times \mathbb{R}^n)$ of the covariant heat equation

$$\begin{cases} \partial_s \tilde{\phi} = \Delta_x \tilde{\phi} + \tilde{\phi} \cdot \sum_{m=1}^d |\partial_m \tilde{\phi}|^2 & \text{on } [0,\infty) \times \mathbb{R}^d, \\ \tilde{\phi}(0,x,t) = \phi(x,t). \end{cases} \tag{5.39}$$

In addition, there are smooth functions $v, w : [0,\infty) \times \mathbb{R}^d \to \mathbb{S}^2$ with the properties

$$v \cdot \tilde{\phi} = w \cdot \tilde{\phi} = v \cdot w = w \cdot \partial_s v = 0 \text{ on } [0,\infty) \times \mathbb{R}^d \times (-T,T), \tag{5.40}$$

and for any $F \in \{\tilde{\phi}, v, w\}$ we have the bounds

$$\|P_k F(s)\|_{L_x^2} \lesssim c_k (1 + s2^{2k})^{-20} 2^{-\frac{n}{2}k}. \tag{5.41}$$

The key caloric gauge condition is the last identity in (5.40), namely ${}^t w \cdot \partial_s v \equiv 0$, which leads to the identity $A_0 \equiv 0$. It is also important that the functions $\tilde{\phi}, v, w$ become trivial as $s \to \infty$.

The L^2 bounds are far from sufficient for our analysis. Instead we need additional F_k bounds for the harmonic heat flow. This happens at the level of space-time estimates, so we add a Schrödinger time variable back into the picture. Again it is convenient to add the frequency envelopes to this picture. This is done with respect to the F_k norm. Thus, let c_k be an F_k frequency envelope for the ψ_m's. To this envelope we associate the sequence

$$c_{>k} = \Big(\sum_{j \geq k} c_j^2 \Big)^{1/2}.$$

Proposition 5.6 (Heat flow bootstrap estimates). *For $T \in (0, \infty)$ and ϕ small in $L^\infty \dot{H}^1(T)$ we consider $\tilde{\phi}, v, w$ as in Proposition 5.5, and ψ_m and A_m the associated fields and connection coefficients.*

(a) *Suppose that the functions $\{\psi_m\}_{m=1,d}$ satisfy*

$$\|P_k \psi_m(0)\|_{F_k(T)} \leq 2^{-k(d-2)/2} c_k, \quad \epsilon := \|c\|_{l^2} \ll 1, \tag{5.42}$$

as well as the bootstrap condition

$$\|P_k \psi_m(s)\|_{F_k(T)} \leq \epsilon^{-1/2} c_k 2^{-k(d-2)/2} (1 + s 2^{2k})^{-4}. \tag{5.43}$$

Then we have

$$\|P_k \psi_m(s)\|_{F_k(T)} \lesssim c_k 2^{-k(d-2)/2} (1 + s 2^{2k})^{-4}. \tag{5.44}$$

Also, for $l, m = 1, \dots, n$ we have the $F_k(T)$ bounds

$$\|P_k (A_m(s) \psi_l(s))\|_{F_k(T)} \lesssim c_k 2^{-k(d-4)/2} (2^{2k} s)^{-\frac{3}{8}} (1 + s 2^{2k})^{-2}, \tag{5.45}$$

as well as the L^{p_d} estimate at $s = 0$

$$\|P_k A_m(0)\|_{L^{p_d}} \lesssim c_k 2^{-k(d-2)/2}. \tag{5.46}$$

(b) *Assume in addition that*

$$\|P_k \psi_{d+1}(0)\|_{L^{p_d}} \lesssim c_k 2^{-k(d-4)/2} 2^k. \tag{5.47}$$

Then we have

$$\|P_k \psi_{d+1}(s)\|_{L^{p_d}} \lesssim c_k 2^{-k(d-4)/2} (1 + 2^{2k} s)^{-2}, \tag{5.48}$$

and the connection coefficient A_{d+1} satisfies the L^2 estimate at $s = 0$

$$\|P_k A_{d+1}(0)\|_{L^2} \lesssim c_k 2^{-k(d-2)/2}, \quad n \geq 3, \tag{5.49}$$

respectively

$$\|A_{d+1}(0)\|_{L^2} \lesssim \epsilon^2, \quad \|P_k A_{d+1}(0)\|_{L^2} \lesssim c_{>k}^2 \quad d = 2. \tag{5.50}$$

The bootstrap assumption (5.43) can be then eliminated.

5.3.2 Bounds for the Schrödinger map flow

Since the connection coefficients A_m are defined via the harmonic heat flow, we cannot use a direct fixed point argument in order to solve the Schrödinger map equation. Instead, we use a bootstrap argument. Our main Schrödinger bootstrap result is the following.

Proposition 5.7 (Schrödinger bootstrap estimates). *Assume that $T \in (0, 2^{2\mathcal{K}}]$ and $Q \in \mathbb{S}^2$. Let $\{c_k\}_{k\in\mathbb{Z}}$ be an ϵ-frequency envelope with $\epsilon \ll 1$. Let ϕ be a smooth Schrödinger map in $[0,T]$ whose initial datum satisfies*

$$\|P_k \nabla \phi_0\|_{L^2_x} \le c_k 2^{-k(d-2)/2}. \tag{5.51}$$

Assume that ϕ satisfies the bootstrap condition

$$\|P_k \nabla \phi\|_{L^\infty_t L^2_x} \le \epsilon^{-1/2} c_k 2^{-k(d-2)/2} \tag{5.52}$$

and let (ϕ, v, w) be the caloric extension of ϕ given by Proposition 5.5, with the corresponding fields ψ_m, A_m. Suppose also that at the initial parabolic time $s=0$ the functions $\{\psi_m\}_{m=1,d}$ satisfy the additional bootstrap condition

$$\|P_k \psi_m(0)\|_{G_k(T)} \le \epsilon^{-1/2} 2^{-(d-2)k/2} c_k. \tag{5.53}$$

Then we have

$$\|P_k \psi_m(0)\|_{G_k(T)} \lesssim 2^{-(d-2)k/2} c_k. \tag{5.54}$$

The above proposition is proved by applying the linear result in Proposition 5.4 to the equation (5.7). The right-hand side in (5.7) is estimated in the $N_k(T)$ spaces using the bounds in Proposition 5.6 for the differentiated fields ψ_m and the connection coefficients A_m.

We note that the bootstrap assumption (5.53) is eliminated via a continuity argument. The additional bootstrap condition (5.52) can also be improved to

$$\|P_k \nabla \phi\|_{L^\infty_t L^2_x} \lesssim c_k \tag{5.55}$$

and then eliminated, by first transferring it to v and w using Proposition 5.5, and then by recovering $\nabla\phi$ via the relations (5.1).

5.3.3 Rough solutions and continuous dependence.

To define rough solutions and study the dependence of solutions on the initial data, we consider the linearized Schrödinger map equation. Expressed in the frame, this has the form

$$\begin{aligned}(i\partial_t + \Delta_x)\psi_{\text{lin}} \\ = -2i\sum_{l=1}^d A_l \partial_l \psi_{\text{lin}} + \Big(A_{d+1} + \sum_{l=1}^d (A_l^2 - i\partial_l A_l)\Big)\psi_{\text{lin}} - i\sum_{l=1}^d \psi_l \Im(\overline{\psi_l}\psi_{\text{lin}}).\end{aligned} \tag{5.56}$$

This can be derived by direct computations as before. Heuristically, one can also think of a one parameter family of solutions $\phi(h)$ for the Schrödinger map equation so that $\phi(0) = \phi$ and $\psi_{\rm lin}$ is the expression in the frame of $\partial_h \phi|_{h=0}$, and extend the frame (v, w) as h varies. For this we will prove that it is well-posed in $\dot{H}^{(d-2)/2}$.

Proposition 5.8. *Let ϕ be a Schrödinger map as above. Then for each initial datum $\psi_{lin}(0) \in H^\infty$ there exists an unique solution $\psi_{lin} \in C(\mathbb{R}, H^\infty)$ for (5.56), which satisfies the bounds*

$$\sum_k 2^{(d-2)k} \|P_k \psi_{lin}\|^2_{G_k(T)} \lesssim \|\psi_{lin}(0)\|^2_{\dot{H}^{\frac{d-2}{2}}}. \tag{5.57}$$

The proof of this result is identical to the proof of Proposition 5.6. As a consequence of this we obtain the Lipschitz dependence of solutions in terms of the initial data in a weaker topology:

Proposition 5.9. *Consider two initial data ϕ_0^0 and ϕ_0^1 in H_Q^∞ which satisfy the smallness condition $\|\phi_0^h\|_{\dot{H}^{\frac{d}{2}}} \ll 1$, $h = 0, 1$, and let ϕ^0 and ϕ^1 be the corresponding global solutions for (5.56). Then*

$$\sum_k 2^{(d-2)k} \|P_k(\phi^0 - \phi^1)\|^2_{L^\infty \dot{H}^{\frac{d-2}{2}}} \lesssim \|\phi_0^0 - \phi_0^1\|^2_{\dot{H}^{\frac{d-2}{2}}}. \tag{5.58}$$

To prove this, one needs to show that any two initial data ϕ_0^0 and ϕ_0^1 which are small in $\dot{H}^1$ can be joined by a one parameter family $\{\phi_0^h\}_{h\in[0,1]} \in C^\infty([0,1]; H^\infty)$ of initial data so that:

$$\int_0^1 \|\partial_h \phi_0^h\|_{\dot{H}^{\frac{d-2}{2}}} \approx \|\phi_0^0 - \phi_0^1\|_{\dot{H}^{\frac{d-2}{2}}}. \tag{5.59}$$

This was proved in [47].

The above proposition allows us to conclude the proof of the strong continuous dependence on the initial data. Precisely, we show that the datum to solution map S_Q admits an unique continuous extension

$$S_Q : \dot{H}^{\frac{d}{2}} \cap \dot{H}_Q^{\frac{d-2}{2}} \to C(\mathbb{R}; \dot{H}^{\frac{d}{2}} \cap \dot{H}_Q^{\frac{d-2}{2}}).$$

It suffices to consider a sequence of smooth initial data $\phi_0^n \in H_Q^\infty$ which satisfy uniformly the smallness condition $\|\phi_0^n\|_{\dot{H}^{\frac{d}{2}}} \ll 1$ and such that $\phi_0^n \to \phi_0$ in $\dot{H}^{\frac{d}{2}} \cap \dot{H}_Q^{\frac{d-2}{2}}$, and show that the corresponding sequence of global solutions is Cauchy in the space in $C(\mathbb{R}; \dot{H}^{\frac{d}{2}} \cap \dot{H}_Q^{\frac{d-2}{2}})$. By Proposition 5.9, it follows that the sequence ϕ^n is Cauchy in $C(\mathbb{R}; \dot{H}_Q^{\frac{d-2}{2}})$:

$$\lim_{n,m\to\infty} \|\phi^n - \phi^m\|_{C(\mathbb{R}; \dot{H}^{\frac{d-2}{2}})} = 0. \tag{5.60}$$

Consider frequency envelopes $\{c_k^n\}$ associated to ϕ_0^n. Since ϕ_0^n is convergent in $\dot{H}^{\frac{d}{2}}$ we can choose the corresponding envelopes $\{c_k^n\}$ to converge in l^2. Then we have the uniform summability property

$$\lim_{k_0\to\infty} \sup_n \sum_{k>k_0} (c_k^n)^2 = 0. \tag{5.61}$$

Now we estimate

$$\begin{aligned}
&\|\phi^n - \phi^m\|^2_{C(\mathbb{R};\dot{H}^{\frac{d}{2}})} \\
&\leq \|P_{\leq k_0}(\phi^n - \phi^m)\|^2_{C(\mathbb{R};\dot{H}^{\frac{d}{2}})} + \|P_{>k_0}\phi^n\|^2_{C(\mathbb{R};\dot{H}^{\frac{d}{2}})} + \|P_{>k_0}\phi^m\|^2_{C(\mathbb{R};\dot{H}^{\frac{d}{2}})} \\
&\lesssim 2^{k_0}\|P_{\leq k_0}(\phi^n - \phi^m)\|^2_{C(\mathbb{R};\dot{H}^{\frac{d-2}{2}})} + \sum_{k>k_0} (c_k^n)^2 + (c_k^m)^2.
\end{aligned}$$

Hence using (5.60) we have

$$\limsup_{n,m\to\infty} \|\phi^n - \phi^m\|^2_{C(\mathbb{R};\dot{H}^{\frac{d}{2}})} \lesssim \sup_n \sum_{k>k_0} (c_k^n)^2.$$

Letting $k_0 \to \infty$, by (5.61) we obtain

$$\limsup_{n,m\to\infty} \|\phi^n - \phi^m\|_{C(\mathbb{R};\dot{H}^{\frac{d}{2}})} = 0,$$

and the argument is concluded.

The continuity of the solution operator S_Q in higher Sobolev spaces

$$S_Q : \dot{H}^\sigma \cap \dot{H}_Q^{\frac{d-2}{2}} \to C(\mathbb{R}; \dot{H}^\sigma \cap \dot{H}_Q^{\frac{d-2}{2}}), \qquad \frac{d}{2} < \sigma \leq \sigma_1$$

can be obtained in the same manner.

5.4 Further developments

5.4.1 Other targets

The frame method works well in the case of the $\mathbb{S}^2$ or $\mathbb{H}^2$ targets, but arbitrary Kähler targets are a different story. There the frame method would not yield a self-contained system for the differentiated fields ψ_m.

Open Problem 5.10. *Prove small data well-posedness for the Schrödinger map equation with values into an arbitrary (say compact) Kähler manifold.*

5.4.2 Large data

For the purpose of this subsection we assume that we are in two space dimensions, i.e., the energy critical case. The reason for this is that in this case the energy is a meaningful invariant object which can be used in the description of the global behavior of solutions.

We begin with the case of the $\mathbb{H}^2$ target, where there are no finite energy harmonic maps, and no other known obstructions to global well-posedness. This is the geometric version of a defocusing problem. Then we have

Conjecture 5.11 (Defocusing Conjecture)**.** *Consider the Schrödinger map problem in two space dimensions, with values in $\mathbb{H}^2$. Then global well-posedness and scattering hold for all finite energy data.*

In the case of the $\mathbb{S}^2$ target, the harmonic maps provide an obvious obstruction to a large data result. In addition, scattering can only occur for solutions in the zero homotopy class. The smallest nontrivial soliton, on the other hand, is the stereographic projection, Q_1, which belongs to the homotopy one class. In order to emulate such a soliton in the zero homotopy class, one needs to wrap the sphere and then unwrap it; this requires twice the soliton energy. Thus the natural conjecture is:

Conjecture 5.12 (Strong Threshold Conjecture)**.** *Consider the Schrödinger map problem in two space dimensions, with values in $\mathbb{S}^2$. Then global well-posedness and scattering holds for all zero homotopy data which satisfy $E(\phi) < 2E(Q_1)$.*

These conjectures parallel recently proved results for wave maps. Both conjectures are still open for Schrödinger maps. However, the equivariant case has recently been studied.

Theorem 5.13 (Bejenaru, Kenig, Ionescu, and Tataru [2])**.** *Consider the Schrödinger map problem in two space dimensions, with values in $\mathbb{H}^2$. For this problem, global well-posedness and scattering hold in the* 1*-equivariant class for all finite energy data.*

Theorem 5.14 (Bejenaru, Kenig, Ionescu, and Tataru [1])**.** *Consider the Schrödinger map problem in two space dimensions, with values in $\mathbb{S}^2$. For this problem, global well-posedness and scattering hold in the* 1*-equivariant class for all zero homotopy data which satisfies $E(\phi) < E(Q_1)$.*

The proof uses the Kenig–Merle method, which involves

- proving that if the result does not hold, then minimal energy blow-up solutions exist and
- eliminating the minimal energy blow-up solutions via mass and momentum Morawetz type estimates.

The key difficulties in the proof are as follows:

- Gauge formulation of the problem: via the Coulomb gauge one obtains two coupled NLS type equations, and the coupling needs to survive in the concentration compactness argument.
- Morawetz (momentum) estimates are harder, and only yield local energy decay in a restricted regime; in particular, we cannot reach the conjectured $2E(Q_1)$ threshold for $\mathbb{S}^2$ targets.

5.4.3 Near soliton behavior

In this section we consider the behavior of solutions with energy above the ground state threshold. For clarity we discuss only the simplest such problem, which is still wide open. Thus, we consider the case of the $\mathbb{S}^2$ target and solutions in the homotopy one class, which have energy just above the soliton energy:

$$E(Q_1) \leq E(\phi) < E(Q_1) + \epsilon. \tag{5.62}$$

We note that if $E(Q_1) = E(\phi)$, then ϕ must belong to the class $\mathcal{Q}_1$ of ground states obtained from Q_1 via symmetries. We also remark that energy considerations show that any such state ϕ must satisfy

$$\operatorname{dist}(\phi, \mathcal{Q}_1) \lesssim \epsilon.$$

Thus the family $\mathcal{Q}_1$ is orbitally stable. Unfortunately, this does not say as much as one might want since the group of symmetries is noncompact. Thus we have the following

Open Problem 5.15. *For Schrödinger maps from $\mathbb{R}^{2+1}$ to $\mathbb{S}^2$ which have homotopy one and satisfy* (5.62)*, understand the possible global dynamics for the flow.*

The key element in this is understanding the motion of solutions along the Q_1 family. Possible issues to consider are

- Can finite time blow up occur? If so, what are the possible rates?
- For global solutions, what is the asymptotic behavior at infinity (if any)?
- Can solutions drift away to spatial infinity in finite time? In infinite time?
- Are there any breather type solutions in this class?

While in such generality the above problem seems out of reach for now, some partial results have been obtained for equivariant solutions. An advantage of working in the equivariant class is that the dimension of the symmetry group is reduced to two, namely scaling and horizontal rotations. The first is noncompact, but the second is compact. Thus we can parametrize the ground states as

$$\mathcal{Q}_1^{eq} = \{Q_{\alpha,\lambda};\ \lambda \in \mathbb{R}^+, \alpha \in \mathbb{S}^1\}$$

The equivariant solutions are represented as

$$\phi(t) = Q_{\alpha(t),\lambda(t)} + O_{\dot{H}^1}(\epsilon)$$

and the question is to understand the behavior of the functions $\alpha(t)$ and $\lambda(t)$.

In chronological order, the results we have so far are as follows:

Theorem 5.16 (Gustafson, Nakanishi, and Tsai [17])**.** *$\mathcal{Q}_k$ ground states are stable in the k equivariant class for $k \geq 3$.*

We remark that this result is very different from the wave-map picture. Also, it seems somewhat unlikely that the result will survive outside the equivariant class.

Theorem 5.17 (Bejenaru and Tataru ($k = 1$, [6]) ($k = 2$, in progress))**.**

a) *$\mathcal{Q}_1$ ground states are unstable in the energy norm $\dot{H}^1$.*

b) *$\mathcal{Q}_1$ ground states are stable in the one equivariant class with respect to a stronger topology X satisfying*

$$H^1 \subset X \subset \dot{H}^1.$$

A key role in this analysis is played by the linearized equation near Q_1 expressed in a suitable gauge. This is a linear Schrödinger equation governed by an explicit operator

$$H = -\Delta + V, \qquad V(r) = \frac{1}{r^2} - \frac{8}{(1+r^2)^2}.$$

A key difficulty is that H has a zero resonance

$$\phi_0 = r\partial_r Q_1 = \frac{2r}{1+r^2},$$

which corresponds to motion along the soliton family.

This is unlike what happens in higher equivariance classes $k \geq 3$, where the analogue of ϕ_0 is not only an eigenvalue, but also belongs to H^{-1}. This allows one to define a corresponding orthogonal projection for functions in $\dot{H}^1$ and opens the door to a more standard perturbation theory.

The proof of the above result requires developing a complete spectral resolution for the operator H. In addition, the parameter $\lambda(t)$ is the main nonperturbative parameter in this analysis, so one in effect needs to work with a linear evolution of the form

$$(i\partial_t + H_{\lambda(t)})\psi = f$$

with a nontrivial dependence of λ on t.

Finally, the last and most recent results in this direction that we mention are

Theorem 5.18 (Merle, Raphael, and Rodnianski [26], Perelman [28])**.** *Finite time blow-up equivariant solutions exist near* $\mathcal{Q}_1$*.*

The first result [26] adapts to the Schrödinger map setting the techniques in the similar work for wave maps in [34], [32]. The second [28] is the Schrödinger map counterpart of the wave map results in [21].

Bibliography

[1] I. Bejenaru, A. Ionescu, C. E. Kenig, and D. Tataru. Equivariant Schrödinger Maps in two spatial dimensions. *ArXiv e-prints*, December 2011.

[2] I. Bejenaru, A. Ionescu, C. E. Kenig, and D. Tataru. Equivariant Schrödinger Maps in two spatial dimensions: the $\H^2$ target. *ArXiv e-prints*, December 2012.

[3] I. Bejenaru, A. D. Ionescu, and C. E. Kenig. Global existence and uniqueness of Schrödinger maps in dimensions $d \geq 4$. *Adv. Math.*, 215(1):263–291, 2007.

[4] I. Bejenaru, A. D. Ionescu, C. E. Kenig, and D. Tataru. Global Schrödinger maps in dimensions $d \geq 2$: small data in the critical Sobolev spaces. *Ann. of Math. (2)*, 173(3):1443–1506, 2011.

[5] I. Bejenaru, J. Krieger, and D. Tataru. A codimension two stable manifold of near soliton equivariant wave maps. *ArXiv e-prints*, September 2011.

[6] I. Bejenaru and D. Tataru. Near soliton evolution for equivariant Schrödinger Maps in two spatial dimensions. *ArXiv e-prints*, September 2010.

[7] Fabrice Bethuel. On the singular set of stationary harmonic maps. *Manuscripta Math.*, 78(4):417–443, 1993.

[8] J. Bourgain. Fourier transform restriction phenomena for certain lattice subsets and applications to nonlinear evolution equations. II. The KdV-equation. *Geom. Funct. Anal.*, 3(3):209–262, 1993.

[9] Yun Mei Chen and Wei Yue Ding. Blow-up and global existence for heat flows of harmonic maps. *Invent. Math.*, 99(3):567–578, 1990.

[10] Yun Mei Chen and Michael Struwe. Existence and partial regularity results for the heat flow for harmonic maps. *Math. Z.*, 201(1):83–103, 1989.

[11] Yvonne Choquet-Bruhat. *General relativity and the Einstein equations*. Oxford Mathematical Monographs. Oxford University Press, Oxford, 2009.

[12] Demetrios Christodoulou and A. Shadi Tahvildar-Zadeh. On the regularity of spherically symmetric wave maps. *Comm. Pure Appl. Math.*, 46(7):1041–1091, 1993.

H. Koch et al., *Dispersive Equations and Nonlinear Waves: Generalized Korteweg–de Vries, Nonlinear Schrödinger, Wave and Schrödinger Maps*, Oberwolfach Seminars 45,
DOI 10.1007/978-3-0348-0736-4,

[13] Raphael Cote, Carlos E. Kenig, and Frank Merle. Scattering below critical energy for the radial 4d Yang-Mills equation and for the 2d corotational wave map system. arXiv:0709.3222.

[14] Lawrence C. Evans. Partial regularity for stationary harmonic maps into spheres. *Arch. Rational Mech. Anal.*, 116(2):101–113, 1991.

[15] Manoussos G. Grillakis and Vagelis Stefanopoulos. Lagrangian formulation, energy estimates, and the Schrödinger map problem. *Comm. Partial Differential Equations*, 27(9-10):1845–1877, 2002.

[16] S. Gustafson, K. Nakanishi, and T.-P. Tsai. Asymptotic Stability, Concentration, and Oscillation in Harmonic Map Heat-Flow, Landau-Lifshitz, and Schrödinger Maps on {mathbb Rˆ2}. *Communications in Mathematical Physics*, 300:205–242, November 2010.

[17] Stephen Gustafson, Kenji Nakanishi, and Tai-Peng Tsai. Asymptotic stability, concentration, and oscillation in harmonic map heat-flow, Landau-Lifshitz, and Schrödinger maps on $\mathbb{R}^2$. *Comm. Math. Phys.*, 300(1):205–242, 2010.

[18] Martin Hadac, Sebastian Herr, and Herbert Koch. Well-posedness and scattering for the KP-II equation in a critical space. *Ann. Inst. H. Poincaré Anal. Non Linéaire*, 26(3):917–941, 2009.

[19] Frédéric Hélein. *Harmonic maps, conservation laws and moving frames*, volume 150 of *Cambridge Tracts in Mathematics*. Cambridge University Press, Cambridge, second edition, 2002.

[20] S. Klainerman and M. Machedon. Space-time estimates for null forms and the local existence theorem. *Comm. Pure Appl. Math.*, 46(9):1221–1268, 1993.

[21] J. Krieger, W. Schlag, and D. Tataru. Renormalization and blow up for charge one equivariant critical wave maps. *Invent. Math.*, 171(3):543–615, 2008.

[22] Joachim Krieger. Global regularity of wave maps from $\mathbb{R}^{2+1}$ to H^2. Small energy. *Comm. Math. Phys.*, 250(3):507–580, 2004.

[23] Joachim Krieger and Wilhelm Schlag. *Concentration compactness for critical wave maps*. EMS Monographs in Mathematics. European Mathematical Society (EMS), Zürich, 2012.

[24] Luc Lemaire. Applications harmoniques de surfaces riemanniennes. *J. Differential Geom.*, 13(1):51–78, 1978.

[25] Fanghua Lin and Changyou Wang. *The analysis of harmonic maps and their heat flows*. World Scientific Publishing Co. Pte. Ltd., Hackensack, NJ, 2008.

[26] F. Merle, P. Raphaël, and I. Rodnianski. Blow up dynamics for smooth equivariant solutions to the energy critical Schrödinger map. *ArXiv e-prints*, June 2011.

[27] Petru Mironescu. Sobolev maps on manifolds: degree, approximation, lifting. In *Perspectives in nonlinear partial differential equations*, volume 446 of *Contemp. Math.*, pages 413–436. Amer. Math. Soc., Providence, RI, 2007.

[28] G. Perelman. Blow up dynamics for equivariant critical Schrödinger maps. *ArXiv e-prints*, December 2012.

[29] Jie Qing. Boundary regularity of weakly harmonic maps from surfaces. *J. Funct. Anal.*, 114(2):458–466, 1993.

[30] Jie Qing and Gang Tian. Bubbling of the heat flows for harmonic maps from surfaces. *Comm. Pure Appl. Math.*, 50(4):295–310, 1997.

[31] P. Raphael and R. Schweyer. Stable blow up dynamics for the 1-corotational energy critical harmonic heat flow. *ArXiv e-prints*, June 2011.

[32] Pierre Raphaël and Igor Rodnianski. Stable blow up dynamics for the critical co-rotational wave maps and equivariant Yang-Mills problems. *Publ. Math. Inst. Hautes Études Sci.*, pages 1–122, 2012.

[33] Tristan Rivière. Conservation laws for conformally invariant variational problems. *Invent. Math.*, 168(1):1–22, 2007.

[34] Igor Rodnianski and Jacob Sterbenz. On the formation of singularities in the critical O(3) σ-model. *Ann. of Math. (2)*, 172(1):187–242, 2010.

[35] Jalal Shatah and A. Shadi Tahvildar-Zadeh. On the Cauchy problem for equivariant wave maps. *Comm. Pure Appl. Math.*, 47(5):719–754, 1994.

[36] P. Smith. Geometric renormalization below the ground state. *ArXiv e-prints*, September 2010.

[37] Jacob Sterbenz and Daniel Tataru. Energy dispersed large data wave maps in $2+1$ dimensions. *Comm. Math. Phys.*, 298(1):139–230, 2010.

[38] Jacob Sterbenz and Daniel Tataru. Regularity of wave-maps in dimension $2+1$. *Comm. Math. Phys.*, 298(1):231–264, 2010.

[39] Michael Struwe. On the evolution of harmonic mappings of Riemannian surfaces. *Comment. Math. Helv.*, 60(4):558–581, 1985.

[40] Michael Struwe. Equivariant wave maps in two space dimensions. *Comm. Pure Appl. Math.*, 56(7):815–823, 2003. Dedicated to the memory of Jürgen K. Moser.

[41] Michael Struwe. Radially symmetric wave maps from $(1+2)$-dimensional Minkowski space to general targets. *Calc. Var. Partial Differential Equations*, 16(4):431–437, 2003.

[42] Terence Tao. Gauges for the Schrödinger map. (unpublished), http://www.math.ucla.edu/ tao/preprints/Expository.

[43] Terence Tao. Global regularity of wave maps VI. Minimal energy blowup solutions. arXiv:0906.2833.

[44] Terence Tao. Endpoint bilinear restriction theorems for the cone, and some sharp null form estimates. *Math. Z.*, 238(2):215–268, 2001.

[45] Terence Tao. Global regularity of wave maps. II. Small energy in two dimensions. *Comm. Math. Phys.*, 224(2):443–544, 2001.

[46] Terence Tao. Multilinear weighted convolution of L^2-functions, and applications to nonlinear dispersive equations. *Amer. J. Math.*, 123(5):839–908, 2001.

[47] Daniel Tataru. On global existence and scattering for the wave maps equation. *Amer. J. Math.*, 123(1):37–77, 2001.

[48] Daniel Tataru. Rough solutions for the wave maps equation. *Amer. J. Math.*, 127(2):293–377, 2005.

[49] Peter Topping. Repulsion and quantization in almost-harmonic maps, and asymptotics of the harmonic map flow. *Ann. of Math. (2)*, 159(2):465–534, 2004.

[50] Thomas Wolff. A sharp bilinear cone restriction estimate. *Ann. of Math. (2)*, 153(3):661–698, 2001.

Dispersive Equations

Monica Vişan

Acknowledgements

While the principal result presented here is from the paper [40], the exposition owes much to a long series of joint works with my collaborators Rowan Killip and Xiaoyi Zhang. We also reproduce here some of the innovations and simplifications introduced in [22]; those notes also contain a far more detailed discussion of the history. Finally, I am grateful to Casey Jao and Jason Murphy for their careful reading of these notes.

Chapter 1

Notation

Throughout this text, we will be regularly referring to the space-time norms

$$\|u\|_{L_t^q L_x^r(\mathbb{R}\times\mathbb{R}^d)} := \left(\int_{\mathbb{R}} \left[\int_{\mathbb{R}^d} |u(t,x)|^r \, dx \right]^{\frac{q}{r}} dt \right)^{\frac{1}{q}}, \tag{1.1}$$

with obvious changes if q or r are infinity. We will often use the abbreviation

$$\|f\|_r := \|f\|_{L_x^r} \quad \text{and} \quad \|u\|_{q,r} := \|u\|_{L_t^q L_x^r}.$$

We write $X \lesssim Y$ to indicate that $X \leq CY$ for some constant C, which is permitted to depend on the ambient spatial dimension, d, without further comment. Other dependencies of C will be indicated with subscripts, for example, $X \lesssim_u Y$. We will write $X \sim Y$ to indicate that $X \lesssim Y \lesssim X$.

We use the 'Japanese bracket' convention: $\langle x \rangle := (1 + |x|^2)^{1/2}$ as well as $\langle \nabla \rangle := (1 - \Delta)^{1/2}$. Similarly, $|\nabla|^s$ denotes the Fourier multiplier with symbol $|\xi|^s$. These are used to define the Sobolev norms

$$\|f\|_{H^{s,r}} := \|\langle \nabla \rangle^s f\|_{L_x^r} \quad \text{and} \quad \|f\|_{\dot{H}^{s,r}} := \||\nabla|^s f\|_{L_x^r}.$$

When $r = 2$ we abbreviate $H^s = H^{s,2}$ and $\dot{H}^s = \dot{H}^{s,2}$.

Our convention for the Fourier transform is

$$\hat{f}(\xi) = (2\pi)^{-\frac{d}{2}} \int_{\mathbb{R}^d} e^{-ix\cdot\xi} f(x) \, dx,$$

so that

$$f(x) = (2\pi)^{-\frac{d}{2}} \int_{\mathbb{R}^d} e^{ix\cdot\xi} \hat{f}(\xi) \, d\xi \quad \text{and} \quad \int_{\mathbb{R}^d} |\hat{f}(\xi)|^2 \, d\xi = \int_{\mathbb{R}^d} |f(x)|^2 \, dx.$$

Notations associated to Littlewood–Paley projections are discussed in the Appendix (Chapter 11).

H. Koch et al., *Dispersive Equations and Nonlinear Waves: Generalized Korteweg–de Vries, Nonlinear Schrödinger, Wave and Schrödinger Maps*, Oberwolfach Seminars 45,
DOI 10.1007/978-3-0348-0736-4_15,

Chapter 2

Dispersive and Strichartz estimates

What all linear dispersive-type equations have in common is a dispersive-type estimate, which expresses the fact that wave packets spread out as time goes by. An expression of this on the Fourier side is the fact that different frequencies move with different speeds and/or in different directions. Below we will discuss several instances of this phenomenon.

2.1 The linear Schrödinger equation

The initial-value problem for the linear Schrödinger equation reads

$$i\partial_t u = -\Delta u \quad \text{with} \quad u(0,x) = u_0(x). \tag{2.1}$$

Here u denotes a complex-valued function on the space-time $\mathbb{R}_t \times \mathbb{R}_x^d$ with spatial dimension $d \geq 1$. By taking Fourier transforms, we observe that

$$\hat{u}(t,\xi) = e^{-it|\xi|^2}\hat{u}_0(\xi). \tag{2.2}$$

In particular, solutions with Schwartz initial data are Schwartz for all $t \in \mathbb{R}$.

Using (2.2) and Plancherel, it is easy to see that solutions to (2.1) conserve *mass*, that is,

$$\|e^{it\Delta}u_0\|_{L_x^2}^2 = \|u_0\|_{L_x^2}^2, \tag{2.3}$$

and *kinetic energy*, that is,

$$\|\nabla e^{it\Delta}u_0\|_{L_x^2}^2 = \|\nabla u_0\|_{L_x^2}^2.$$

To derive an explicit formula for solutions to (2.1), we will first study the particular case of modulated Gaussian initial data, namely,

$$u_0(x) = \exp\Big\{-\tfrac{|x|^2}{4\sigma^2} + ix\xi_0\Big\} \quad \text{with} \quad \sigma > 0 \quad \text{and} \quad \xi_0 \in \mathbb{R}^d.$$

H. Koch et al., *Dispersive Equations and Nonlinear Waves: Generalized Korteweg–de Vries, Nonlinear Schrödinger, Wave and Schrödinger Maps*, Oberwolfach Seminars 45,
DOI 10.1007/978-3-0348-0736-4_16, © Springer Basel 2014

This initial datum is a Gaussian that lives at scale σ and has wave vector ξ_0, that is, it has wave length $\frac{2\pi}{|\xi_0|}$ and the wave fronts are perpendicular to ξ_0. A straightforward computation yields that the solution u to (2.1) with this initial datum is given by

$$\begin{aligned}
[e^{it\Delta}u_0](x) &= (2\pi)^{-\frac{d}{2}} \int_{\mathbb{R}^d} e^{ix\xi - it|\xi|^2} \hat{u}_0(\xi)\, d\xi \\
&= (2\pi)^{-d} \int_{\mathbb{R}^d}\int_{\mathbb{R}^d} e^{ix\xi - it|\xi|^2} e^{-iy\xi} e^{-\frac{|y|^2}{4\sigma^2} + iy\xi_0}\, dy\, d\xi \\
&= (2\pi)^{-d} \int_{\mathbb{R}^d}\int_{\mathbb{R}^d} e^{ix\xi - it|\xi|^2} e^{-\sigma^2|\xi-\xi_0|^2} e^{-|\frac{y}{2\sigma} + i\sigma(\xi-\xi_0)|^2}\, dy\, d\xi \\
&= (2\pi)^{-d}(4\pi\sigma^2)^{\frac{d}{2}} e^{-it|\xi_0|^2 + ix\xi_0 - \frac{|x-2t\xi_0|^2}{4(\sigma^2+it)}} \int_{\mathbb{R}^d} e^{-(\sigma^2+it)\left|\xi - \frac{ix+2\sigma^2\xi_0}{2(\sigma^2+it)}\right|^2}\, d\xi \\
&= \Big(\frac{\sigma^2}{\sigma^2+it}\Big)^{\frac{d}{2}} \exp\Big\{-it|\xi_0|^2 + ix\xi_0 - \tfrac{|x-2t\xi_0|^2}{4(\sigma^2+it)}\Big\}.
\end{aligned} \tag{2.4}$$

In the formulas above, $|v|^2 := \sum_{j=1}^d v_j^2$ for all vectors $v \in \mathbb{C}^d$.

Exercise 2.1. Justify all steps in the derivation of (2.4).

Remark. From the exact formula (2.4), we read the following:
- the wave packet travels at speed $2\xi_0$ (called the *group velocity*)
- the wave vector is still ξ_0 (called the *phase velocity*)
- while the amplitude of the wave packet decreases with time, the wave packet also spreads out: $\operatorname{Re} \frac{1}{4(\sigma^2+it)} < \frac{1}{4\sigma^2}$. This is consistent with the conservation of mass.

We are now ready to derive an exact formula for solutions to (2.1), at least for Schwartz initial data $u_0 \in \mathcal{S}(\mathbb{R}^d)$. Using the linearity of the propagator $e^{it\Delta}$ and (2.4), we get

$$e^{it\Delta}\Big[(4\pi\sigma^2)^{-\frac{d}{2}} \int_{\mathbb{R}^d} e^{-\frac{|x-y|^2}{4\sigma^2}} u_0(y)\, dy\Big] = [4\pi(\sigma^2+it)]^{-\frac{d}{2}} \int_{\mathbb{R}^d} e^{-\frac{|x-y|^2}{4(\sigma^2+it)}} u_0(y)\, dy.$$

To continue, the key observation is that for $u_0 \in \mathcal{S}(\mathbb{R}^d)$,

$$\lim_{\sigma\to 0} (4\pi\sigma^2)^{-\frac{d}{2}} \int_{\mathbb{R}^d} e^{-\frac{|x-y|^2}{4\sigma^2}} u_0(y)\, dy = u_0(x), \tag{2.5}$$

both pointwise in x and in the L_x^2 topology. Using also that the propagator $e^{it\Delta}$ is continuous in the L_x^2 topology (on Schwartz space), we get the exact formula

$$[e^{it\Delta}u_0](x) = (4\pi it)^{-\frac{d}{2}} \int_{\mathbb{R}^d} e^{\frac{i|x-y|^2}{4t}} u_0(y)\, dy \quad \text{for} \quad t \neq 0, \tag{2.6}$$

for all $u_0 \in \mathcal{S}(\mathbb{R}^d)$, where the equality is meant in the L_x^2 sense.

This leads directly to the dispersive inequality for the linear Schrödinger propagator:

$$\|e^{it\Delta}u_0\|_{L^\infty_x} \lesssim |t|^{-\frac{d}{2}}\|u_0\|_{L^1_x} \quad \text{for} \quad t \neq 0. \tag{2.7}$$

Interpolating with (2.3), we obtain the full range of dispersive estimates for the linear Schrödinger propagator:

$$\|e^{it\Delta}u_0\|_{L^p_x} \lesssim |t|^{\frac{d}{2}(\frac{1}{p}-\frac{1}{p'})}\|u_0\|_{L^{p'}_x} \quad \text{for} \quad t \neq 0, \tag{2.8}$$

for all $2 \leq p \leq \infty$, where p' denotes the exponent conjugate to p, that is, $\frac{1}{p}+\frac{1}{p'} = 1$.

Exercise 2.2. Prove that for all $u_0 \in L^2_x$, the equality (2.5) holds both a.e. in x and in the L^2_x topology.

2.2 The Airy equation

The initial-value problem for the Airy equation reads

$$\partial_t u = -\partial_x^3 u \quad \text{with} \quad u(0,x) = u_0(x). \tag{2.9}$$

Here u denotes a real-valued function on the space-time $\mathbb{R}_t \times \mathbb{R}_x$. Note that complex-valued solutions to (2.9) have the property that their real and imaginary parts individually solve (2.9).

Using the Fourier transform, we arrive at

$$[e^{-t\partial_x^3}u_0](x) = (3t)^{-1/3}\int_{\mathbb{R}} \mathrm{Ai}\big(\tfrac{x-y}{(3t)^{1/3}}\big)u_0(y)\,dy \quad \text{for} \quad t \neq 0, \tag{2.10}$$

where

$$\mathrm{Ai}(x) := \pi^{-1}\int_0^\infty \cos\big(\tfrac{1}{3}\xi^3 + x\xi\big)\,d\xi$$

denotes the Airy function of the first kind.

Exercise 2.3. Prove that the Airy function is uniformly bounded. Indeed, show that $\mathrm{Ai}(x) \to 0 \quad \text{as} \quad x \to \pm\infty$.
Hint: Use non-stationary phase for $x \geq 1$; van der Corput for $|x| \leq 1$; van der Corput for $x \leq -1$ on $|\xi| \sim |x|^{1/2}$ and the complementary region, separately.

As a consequence of this exercise and (2.10), we obtain the dispersive estimate for the Airy equation:

$$\|e^{-t\partial_x^3}u_0\|_{L^\infty_x} \lesssim |t|^{-\frac{1}{3}}\|u_0\|_{L^1_x} \quad \text{for} \quad t \neq 0. \tag{2.11}$$

Interpolating with the conservation law

$$\|e^{-t\partial_x^3}u_0\|_{L^2_x} = \|u_0\|_{L^2_x},$$

we obtain the full range of dispersive estimates, namely,

$$\|e^{-t\partial_x^3}u_0\|_{L_x^p} \lesssim |t|^{\frac{1}{3}(\frac{1}{p}-\frac{1}{p'})}\|u_0\|_{L_x^{p'}} \quad \text{for} \quad t \neq 0, \tag{2.12}$$

for all $2 \leq p \leq \infty$, where p' denotes the exponent conjugate to p, $\frac{1}{p}+\frac{1}{p'}=1$.

We may strengthen the dispersive estimate (2.11) by localizing in frequency:

Exercise 2.4 (Frequency-localized dispersive estimate for the Airy propagator)**.** Let $f \in \mathcal{S}(\mathbb{R})$. Prove that

$$\|e^{-t\partial_x^3}P_N u_0\|_{L_x^\infty} \lesssim \min\{|t|^{-\frac{1}{3}}, (N|t|)^{-\frac{1}{2}}\}\|P_N u_0\|_{L_x^1}$$

uniformly for $N \in 2^{\mathbb{Z}}$ and $t \neq 0$. Here P_N denotes the Littlewood–Paley projection to frequencies $|\xi| \sim N$; see Appendix (11) for definitions and basic properties.

2.3 The linear wave equation

The initial-value problem for the linear wave equation reads

$$\partial_t^2 u = \Delta u \quad \text{with} \quad u(0,x) = u_0(x) \quad \text{and} \quad \partial_t u(0,x) = u_1(x). \tag{2.13}$$

Here u denotes a real-valued function on the space-time $\mathbb{R}_t \times \mathbb{R}_x^d$ with spatial dimension $d \geq 1$.

Using the Fourier transform, we find

$$\begin{pmatrix} u \\ u_t \end{pmatrix}(t) = \begin{pmatrix} \cos(t|\nabla|) & |\nabla|^{-1}\sin(t|\nabla|) \\ -|\nabla|\sin(t|\nabla|) & \cos(t|\nabla|) \end{pmatrix} \begin{pmatrix} u_0 \\ u_1 \end{pmatrix}.$$

One can derive an explicit formula for the wave propagator in spatial variables; see, for example, [31]. One advantage of this expression is that it immediately yields Huygens' principle. This exact formula can also be used to derive the dispersive estimate we give below; however, we prefer to take a Fourier analytic approach that generalizes to more equations.

Lemma 2.1 (Frequency-localized dispersive estimate for the half-wave propagator)**.** *For any $d \geq 1$ and any frequency $N \in 2^{\mathbb{Z}}$, we have*

$$\|e^{\pm it|\nabla|}P_N f\|_{L_x^\infty} \lesssim \big(1+|t|N\big)^{-\frac{d-1}{2}}N^d\|P_N f\|_{L_x^1}. \tag{2.14}$$

In particular, interpolating with $\|e^{it|\nabla|}P_N f\|_{L_x^2} = \|P_N f\|_{L_x^2}$ we get

$$\|e^{\pm it|\nabla|}P_N f\|_{L_x^p} \lesssim \big(1+|t|N\big)^{-\frac{(d-1)(p-2)}{2p}}N^{\frac{d(p-2)}{p}}\|P_N f\|_{L_x^{p'}}, \tag{2.15}$$

for all $2 \leq p \leq \infty$, where p' denotes the exponent conjugate to p, that is, $\frac{1}{p}+\frac{1}{p'}=1$.

Proof. By symmetry, it suffices to prove the dispersive estimate for the propagator $e^{it|\nabla|}$. If $d = 1$ or $d \geq 2$ and $|t| \lesssim N^{-1}$, the claim (2.14) follows easily from the Bernstein inequality:

$$\|e^{it|\nabla|}P_N f\|_{L^\infty_x} \lesssim N^{\frac{d}{2}}\|e^{it|\nabla|}P_N f\|_{L^2_x} \lesssim N^{\frac{d}{2}}\|P_N f\|_{L^2_x} \lesssim N^d\|P_N f\|_{L^1_x}.$$

It thus remains to prove the claim for $d \geq 2$ and $|t| \gg N^{-1}$, to which we now turn. We write

$$e^{it|\nabla|}P_N f = e^{it|\nabla|}\tilde{P}_N f_N = \big[e^{it|\xi|}\tilde{\psi}\big(\tfrac{\xi}{N}\big)\widehat{f_N}(\xi)\big]^\vee = \big[e^{it|\xi|}\tilde{\psi}\big(\tfrac{\xi}{N}\big)\big]^\vee * f_N,$$

where $\tilde{P}_N = P_{N/2} + P_N + P_{2N}$ denotes the fattened Littlewood–Paley projection, ψ denotes the Fourier multiplier associated with P_1, and $\tilde{\psi}$ denotes the Fourier multiplier associated with $\tilde{P}_1$. To establish (2.14), it thus suffices to show

$$\Big|\int_{\mathbb{R}^d} e^{ix\xi+it|\xi|}\tilde{\psi}\big(\tfrac{\xi}{N}\big)\,d\xi\Big| \lesssim N^{\frac{d+1}{2}}|t|^{-\frac{d-1}{2}} \tag{2.16}$$

for all $d \geq 2$ and $|t| \gg N^{-1}$.

Using a change of variables and switching to polar coordinates, we write

$$\int_{\mathbb{R}^d} e^{ix\xi+it|\xi|}\tilde{\psi}\big(\tfrac{\xi}{N}\big)\,d\xi = N^d\int_0^\infty\int_{S^{d-1}} e^{ixNr\omega+itNr}\tilde{\psi}(r)\,d\sigma(\omega)r^{d-1}\,dr \tag{2.17}$$

$$= N^d\int_0^\infty e^{itNr}\tilde{\psi}(r)\check{\sigma}(Nr|x|)r^{d-1}\,dr, \tag{2.18}$$

where $d\sigma$ denotes the surface measure on the sphere $S^{d-1} \subset \mathbb{R}^d$.

If $|x| \ll |t|$, we note that the phase $\phi(r) := Nrx\omega + Nrt$ has no critical points; indeed, $|\phi'(r)| \gtrsim N|t|$. Thus, writing $e^{i\phi(r)} = \frac{1}{i\phi'(r)}\partial_r e^{i\phi(r)}$ and integrating by parts k times in (2.17), we get the bound

$$\Big|\int_{\mathbb{R}^d} e^{ix\xi+it|\xi|}\tilde{\psi}\big(\tfrac{\xi}{N}\big)\,d\xi\Big| \lesssim_k N^d(N|t|)^{-k} \lesssim N^{\frac{d+1}{2}}|t|^{-\frac{d-1}{2}}.$$

To obtain the last inequality, we take $k = \frac{d-1}{2}$ if the dimension d is odd, or $k = \frac{d}{2}$ if the dimension d is even (recalling that $|t| \gg N^{-1}$).

It remains to consider the case $|x| \gtrsim |t|$. In this case we use (2.18) together with the following lemma:

Lemma 2.2. *Let $d \geq 2$ and let $d\sigma$ denote the surface measure on the sphere $S^{d-1} \subset \mathbb{R}^d$. Then*

$$|\check{\sigma}(x)| \lesssim \langle x\rangle^{-\frac{d-1}{2}}.$$

Proof. Exercise! *Hint:* Using the fact that $d\sigma$ is rotationally invariant, we may write

$$\check{\sigma}(x) = (2\pi)^{-\frac{d}{2}}\int_{S^{d-1}} e^{i|x|\xi_d}\,d\sigma(\xi) \sim \int_0^\pi e^{i|x|\cos\theta}(\sin\theta)^{d-2}\,d\theta,$$

where θ is the angle x makes with e_d. Now use stationary phase and van der Corput. $\square$

Returning to the proof of Lemma 2.1, for $|x| \gtrsim |t| \gg N^{-1}$ we use (2.18) and Lemma 2.2 to estimate

$$\left| \int_{\mathbb{R}^d} e^{ix\xi + it|\xi|} \tilde{\psi}\big(\tfrac{\xi}{N}\big)\, d\xi \right| \lesssim N^d \int_{\mathbb{R}^d} |\tilde{\psi}(r)| (Nr|x|)^{-\frac{d-1}{2}} r^{d-1}\, dr \lesssim N^{\frac{d+1}{2}} |t|^{-\frac{d-1}{2}},$$

which gives (2.16) in this case. This completes the proof of (2.16), and so the proof of Lemma 2.14. $\square$

2.4 From dispersive to Strichartz estimates

In this subsection, we will only present details for the derivation of Strichartz estimates for the wave equation. Strichartz estimates for Schrödinger and Airy are left as exercises for the reader.

Definition 2.3. We say that (q, r) is *wave admissible* if

$$\frac{1}{q} + \frac{d-1}{2r} \le \frac{d-1}{4}, \quad q, r, d \ge 2, \quad \text{and} \quad (q, r, d) \ne (2, \infty, 3).$$

Proposition 2.4 (Frequency-localized Strichartz estimates for the half-wave propagator)**.** *Let $d \ge 2$ and (q, r) be wave admissible such that $\frac{1}{q} + \frac{d}{r} = \frac{d}{2} - \gamma$ for some $\gamma > 0$. Then*

$$\|e^{\pm it|\nabla|} P_N f\|_{L^q_t L^r_x} \lesssim N^\gamma \|P_N f\|_{L^2_x}, \tag{2.19}$$

$$\left\| \int_{\mathbb{R}} e^{\mp it|\nabla|} P_N F(t)\, dt \right\|_{L^2_x} \lesssim N^\gamma \|P_N F\|_{L^{q'}_t L^{r'}_x}. \tag{2.20}$$

Moreover, if $(\tilde{q}, \tilde{r})$ is also a wave admissible pair, then we have the retarded estimate

$$\left\| \int_{s<t} e^{\pm i(t-s)|\nabla|} P_N F(s)\, ds \right\|_{L^q_t L^r_x} \lesssim N^{d - \frac{1}{q} - \frac{1}{\tilde{q}} - \frac{d}{r} - \frac{d}{\tilde{r}}} \|P_N F\|_{L^{\tilde{q}'}_t L^{\tilde{r}'}_x}. \tag{2.21}$$

Proof. We will only prove the proposition in the non-endpoint cases, that is, omitting the pair $(2, \frac{2(d-1)}{d-3})$ for $d > 3$. For the endpoint case, see [17].

By the TT^* argument, (2.19) is equivalent to (2.20) and they are both equivalent to

$$\left\| \int_{\mathbb{R}} e^{\pm i(t-s)|\nabla|} P_N F(s)\, ds \right\|_{L^q_t L^r_x} \lesssim N^{2\gamma} \|F\|_{L^{q'}_t L^{r'}_x}. \tag{2.22}$$

When $\frac{1}{q} + \frac{d-1}{2r} < \frac{d-1}{4}$ we use (2.15) and Young's inequality to estimate

$$\text{LHS(2.22)} \lesssim \left\| \int_{\mathbb{R}} \big(1 + |t-s|N\big)^{-\frac{(d-1)(r-2)}{2r}} N^{\frac{d(r-2)}{r}} \|P_N F(s)\|_{L^{r'}_x}\, ds \right\|_{L^q_t}$$

$$\lesssim N^{\frac{d(r-2)}{r}} \|P_N F\|_{L_t^{q'} L_x^{r'}} \big\| \big(1+|t|N\big)^{-\frac{(d-1)(r-2)}{2r}} \big\|_{L_t^{q/2}}$$
$$\lesssim N^{\frac{d(r-2)}{r}} N^{-\frac{2}{q}} \|P_N F\|_{L_t^{q'} L_x^{r'}},$$

which gives (2.22) in this case. When $\frac{1}{q} + \frac{d-1}{2r} = \frac{d-1}{4}$ we use instead the Hardy–Littlewood–Sobolev inequality to obtain

$$\text{LHS(2.22)} \lesssim \Big\| \int_{\mathbb{R}} |t-s|^{-\frac{(d-1)(r-2)}{2r}} N^{\frac{(d+1)(r-2)}{2r}} \|P_N F(s)\|_{L_x^{r'}}\, ds \Big\|_{L_t^q}$$
$$\lesssim N^{\frac{(d+1)(r-2)}{2r}} \|P_N F\|_{L_t^{q'} L_x^{r'}},$$

which gives (2.22) in this case. Note that the application of Hardy–Littlewood–Sobolev requires $r < \frac{2(d-1)}{d-3}$. This completes the proof of (2.22), and so the proof of (2.19) and (2.20).

We now turn to (2.21). First we note that, by Bernstein's inequality, it suffices to prove the claim for those admissible pairs that are *sharp* admissible in the sense that $\frac{1}{q} + \frac{d-1}{2r} = \frac{d-1}{4} = \frac{1}{\tilde q} + \frac{d-1}{2\tilde r}$. Next, we remark that the proof of (2.22) gives (2.21) for $(\tilde q, \tilde r) = (q, r)$. Finally, to obtain the full range of sharp admissible pairs, one interpolates between this and the following two estimates, which are simple consequences of duality and (2.19) and (2.20):

$$\Big\| \int_{\mathbb{R}} e^{\pm i(t-s)|\nabla|} P_N F(s)\, ds \Big\|_{L_t^\infty L_x^2} \lesssim N^{\frac{d}{2} - \frac{1}{\tilde q} - \frac{d}{\tilde r}} \|P_N F\|_{L_t^{\tilde q'} L_x^{\tilde r'}},$$
$$\Big\| \int_{\mathbb{R}} e^{\pm i(t-s)|\nabla|} P_N F(s)\, ds \Big\|_{L_t^q L_x^r} \lesssim N^{\frac{d}{2} - \frac{1}{q} - \frac{d}{r}} \|P_N F\|_{L_t^1 L_x^2}.$$

This completes the proof of the lemma. $\square$

Corollary 2.5 (Strichartz estimates for the half-wave propagator)**.** *Let $d \geq 2$ and (q, r) be wave admissible such that $r \neq \infty$ and $\frac{1}{q} + \frac{d}{r} = \frac{d}{2} - \gamma$ for some $\gamma > 0$. Then*

$$\|e^{\pm it|\nabla|} f\|_{L_t^q L_x^r} \lesssim \||\nabla|^\gamma f\|_{L_x^2},$$
$$\Big\| \int_{\mathbb{R}} e^{\mp it|\nabla|} F(t)\, dt \Big\|_{L_x^2} \lesssim \||\nabla|^\gamma F\|_{L_t^{q'} L_x^{r'}}.$$

Moreover, if $(\tilde q, \tilde r)$ is also a wave admissible pair with $\tilde r \neq \infty$, then

$$\Big\| \int_{s<t} e^{\pm i(t-s)|\nabla|} F(s)\, ds \Big\|_{L_t^q L_x^r} \lesssim \||\nabla|^{d - \frac{1}{q} - \frac{1}{\tilde q} - \frac{d}{r} - \frac{d}{\tilde r}} F\|_{L_t^{\tilde q'} L_x^{\tilde r'}}.$$

Proof. In view of Proposition 2.4, it suffices to prove that

$$\|F\|_{L_t^q L_x^r} \lesssim \Big\{ \sum_{N \in 2^{\mathbb{Z}}} \|P_N F\|_{L_t^q L_x^r}^2 \Big\}^{1/2} \quad \text{for all } 2 \leq q \leq \infty \text{ and } 2 \leq r < \infty, \tag{2.23}$$

which by duality is equivalent to

$$\Big\{\sum_{N\in 2^{\mathbb{Z}}}\|P_N F\|^2_{L_t^{q'}L_x^{r'}}\Big\}^{1/2}\lesssim \|F\|_{L_t^{q'}L_x^{r'}}\quad \text{for all } 2\le q\le\infty \text{ and } 2\le r<\infty. \tag{2.24}$$

To see that (2.23) and (2.24) are equivalent, consider the operator $T : L_t^{q'}L_x^{r'} \to l^2(L_t^{q'}L_x^{r'})$ given by $T(F) = \{P_N F\}_{N\in 2^{\mathbb{Z}}}$. The operator T being bounded is equivalent to (2.24). It is easy to check that the adjoint of T is $T^* : l^2(L_t^q L_x^r) \to L_t^q L_x^r$ given by $T^*(\{G_N\}_{N\in 2^{\mathbb{Z}}}) = \sum_{N\in 2^{\mathbb{Z}}} P_N G_N$. Boundedness of T^* implies

$$\Big\|\sum_{N\in 2^{\mathbb{Z}}} P_N G_N\Big\|_{L_t^q L_x^r}\lesssim \Big\{\sum_{N\in 2^{\mathbb{Z}}}\|G_N\|^2_{L_t^q L_x^r}\Big\}^{1/2}. \tag{2.25}$$

Writing $F = \sum_{N\in 2^{\mathbb{Z}}} P_N F = \sum_{N\in 2^{\mathbb{Z}}} P_N \tilde{P}_N F$ and applying (2.25) with $G_N = \tilde{P}_N F$, we obtain (2.23). Thus (2.24) implies (2.23). To see that (2.23) implies (2.25) and so (2.24), we estimate

$$\begin{aligned}\Big\|\sum_{N\in 2^{\mathbb{Z}}} P_N G_N\Big\|_{L_t^q L_x^r} &\lesssim \Big\{\sum_{N\in 2^{\mathbb{Z}}}\Big\|P_N\sum_{M\in 2^{\mathbb{Z}}} P_M G_M\Big\|^2_{L_t^q L_x^r}\Big\}^{1/2}\\ &\lesssim \Big\{\sum_{N\in 2^{\mathbb{Z}}}\Big\|\sum_{M\sim N} G_M\Big\|^2_{L_t^q L_x^r}\Big\}^{1/2}\\ &\lesssim \Big\{\sum_{N\in 2^{\mathbb{Z}}}\|G_N\|^2_{L_t^q L_x^r}\Big\}^{1/2}.\end{aligned}$$

It thus remains to prove (2.23); for this it suffices to show that

$$\|f\|_{L_x^r}\lesssim \Big\{\sum_{N\in 2^{\mathbb{Z}}}\|P_N f\|^2_{L_x^r}\Big\}^{1/2}\quad \text{for all}\quad 2\le r<\infty, \tag{2.26}$$

since then, for $q\ge 2$, we obtain

$$\begin{aligned}\|F\|_{L_t^q L_x^r} &\lesssim \Big\|\Big\{\sum_{N\in 2^{\mathbb{Z}}}\|P_N F(t)\|^2_{L_x^r}\Big\}^{1/2}\Big\|_{L_t^q} = \Big\|\sum_{N\in 2^{\mathbb{Z}}}\|P_N F(t)\|^2_{L_x^r}\Big\|^{1/2}_{L_t^{\frac{q}{2}}}\\ &\lesssim \Big\{\sum_{N\in 2^{\mathbb{Z}}}\Big\|\|P_N F(t)\|^2_{L_x^r}\Big\|_{L_t^{\frac{q}{2}}}\Big\}^{1/2} = \Big\{\sum_{N\in 2^{\mathbb{Z}}}\|P_N F\|^2_{L_t^q L_x^r}\Big\}^{1/2}.\end{aligned}$$

Finally, to prove (2.26) we use the square function estimate and the same argument as above:

$$\|f\|_{L_x^r}\sim \Big\|\Big\{\sum_{N\in 2^{\mathbb{Z}}}|P_N f|^2\Big\}^{1/2}\Big\|_{L_x^r}\lesssim \Big\{\sum_{N\in 2^{\mathbb{Z}}}\|P_N f\|^2_{L_x^r}\Big\}^{1/2}\quad \text{for all}\quad 2\le r<\infty.$$

This completes the proof of the corollary. $\square$

Corollary 2.6 (Strichartz estimates for the wave equation). *Let $d \geq 2$ and let (q,r) and $(\tilde{q},\tilde{r})$ be wave admissible pairs such that $r, \tilde{r} < \infty$ and $\frac{1}{q}+\frac{d}{r} = \frac{d}{2}-\gamma = \frac{1}{\tilde{q}'}+\frac{d}{\tilde{r}'}-2$ for some $\gamma > 0$. If u solves*

$$\partial_t^2 u = \Delta u + F \quad \text{with} \quad u(0) = u_0 \quad \text{and} \quad \partial_t u(0) = u_1$$

on $I \times \mathbb{R}^d$ for some time interval $I \ni 0$, then

$$\|u\|_{L_t^\infty \dot{H}_x^\gamma} + \|\partial_t u\|_{L_t^\infty \dot{H}_x^{\gamma-1}} + \|u\|_{L_t^q L_x^r} \lesssim \|u_0\|_{\dot{H}_x^\gamma} + \|u_1\|_{\dot{H}_x^{\gamma-1}} + \|F\|_{L_t^{\tilde{q}'} L_x^{\tilde{r}'}},$$

where all space-time norms are over $I \times \mathbb{R}^d$.

Proof. Exercise! □

For the Schrödinger equation we have the following Strichartz estimates:

Lemma 2.7 (Strichartz estimates for the Schrödinger equation). *Let $d \geq 1$ and let (q,r) and $(\tilde{q},\tilde{r})$ be such that $2 \leq q, r, \tilde{q}, \tilde{r} \leq \infty$, $\frac{2}{q}+\frac{d}{r} = \frac{d}{2} = \frac{2}{\tilde{q}}+\frac{d}{\tilde{r}}$, and $(q,r,d) \neq (2,\infty,2)$ and $(\tilde{q},\tilde{r},d) \neq (2,\infty,2)$. If u solves*

$$i\partial_t u = -\Delta u + F \quad \text{with} \quad u(0) = u_0$$

on $I \times \mathbb{R}^d$ for some time interval $I \ni 0$, then

$$\|u\|_{L_t^\infty L_x^2(I\times\mathbb{R}^d)} + \|u\|_{L_t^q L_x^r(I\times\mathbb{R}^d)} \lesssim \|u_0\|_{L_x^2} + \|F\|_{L_t^{\tilde{q}'} L_x^{\tilde{r}'}(I\times\mathbb{R}^d)}.$$

Proof. Using as a model the proof of Proposition 2.4, prove the lemma for all pairs of exponents except the endpoints, that is, whenever $r \neq \frac{2d}{d-2}$ and $\tilde{r} \neq \frac{2d}{d-2}$ for $d \geq 3$. For a proof in the endpoint case, see [17]. □

Finally, we record the Strichartz estimates for the Airy equation:

Lemma 2.8 (Strichartz estimates for the Airy equation). *Let (q,r) and $(\tilde{q},\tilde{r})$ be such that $2 \leq q, r, \tilde{q}, \tilde{r} \leq \infty$, $\frac{1}{q}+\frac{1}{3r} = \frac{1}{6} = \frac{1}{\tilde{q}}+\frac{1}{3\tilde{r}}$. If u solves*

$$\partial_t u = -\partial_x^3 u + F \quad \text{with} \quad u(0) = u_0$$

on $I \times \mathbb{R}$ for some time interval $I \ni 0$, then

$$\|u\|_{L_t^\infty L_x^2(I\times\mathbb{R})} + \|u\|_{L_t^q L_x^r(I\times\mathbb{R})} + \left\||\nabla|^{1/6} u\right\|_{L_{t,x}^6(I\times\mathbb{R})} \lesssim \|u_0\|_{L_x^2} + \|F\|_{L_t^{\tilde{q}'} L_x^{\tilde{r}'}(I\times\mathbb{R})}.$$

Proof. Exercise! □

2.5 Bilinear Strichartz and local smoothing estimates

In this subsection, we restrict attention to the Schrödinger propagator.

Theorem 2.9 (Bilinear Strichartz I, [3, 13, 28]). *Fix $d \geq 1$ and $M \leq N$. Then*

$$\big\|[e^{it\Delta}P_M f][e^{it\Delta}P_N g]\big\|_{L^2_{t,x}(\mathbb{R}\times\mathbb{R}^d)} \lesssim M^{\frac{d-1}{2}} N^{-\frac{1}{2}} \|f\|_{L^2_x(\mathbb{R}^d)} \|g\|_{L^2_x(\mathbb{R}^d)}. \tag{2.27}$$

When $d = 1$ we require $M \leq \frac{1}{4}N$, so that $P_N P_M = 0$.

Proof. For $M \sim N$ and $d \neq 1$, the result follows from the $L^2_x \to L^4_t L^{\frac{2d}{d-1}}_x$ Strichartz inequality and Bernstein.

Turning to the case $M \leq \frac{1}{4}N$, we note that by duality and the Parseval identity, it suffices to show

$$\begin{aligned}\Big|\iint_{\mathbb{R}^d\times\mathbb{R}^d} F(|\xi|^2+|\eta|^2, \xi+\eta)\widehat{f_M}(\xi)\widehat{g_N}(\eta)\,d\xi\,d\eta\Big| \\ \lesssim M^{\frac{d-1}{2}} N^{-\frac{1}{2}} \|F\|_{L^2_{\omega,\xi}(\mathbb{R}^{1+d})} \|\hat f\|_{L^2_\xi(\mathbb{R}^d)} \|\hat g\|_{L^2_\xi(\mathbb{R}^d)}.\end{aligned} \tag{2.28}$$

By decomposing the region of integration into several pieces (and rotating the coordinate system appropriately), we can restrict the region of integration to a set where $\eta_1 - \xi_1 \gtrsim N$. Next, we make the change of variables $\zeta = \xi+\eta$, $\omega = |\xi|^2+|\eta|^2$, and $\beta = (\xi_2, \ldots, \xi_d)$. Note that $|\beta| \lesssim M$ while the Jacobian is $J \sim N^{-1}$. Using this information together with Cauchy–Schwarz, we get

$$\begin{aligned}\text{LHS(2.28)} &= \Big|\iiint F(\omega,\zeta)\widehat{f_M}(\xi)\widehat{g_N}(\eta) J\,d\omega\,d\zeta\,d\beta\Big| \\ &\leq \|F\|_{L^2_{\omega,\xi}(\mathbb{R}^{1+d})} \int \Big[\iint |\widehat{f_M}(\xi)|^2 |\widehat{g_N}(\eta)|^2 J^2\,d\omega\,d\zeta\Big]^{\frac{1}{2}} d\beta \\ &\lesssim \|F\|_{L^2_{\omega,\xi}(\mathbb{R}^{1+d})} M^{\frac{d-1}{2}} \Big(\iiint |\widehat{f_M}(\xi)|^2 |\widehat{g_N}(\eta)|^2 J^2\,d\omega\,d\zeta\,d\beta\Big)^{\frac{1}{2}} \\ &\lesssim \|F\|_{L^2_{\omega,\xi}(\mathbb{R}^{1+d})} M^{\frac{d-1}{2}} \Big(\iint |\widehat{f_M}(\xi)|^2 |\widehat{g_N}(\eta)|^2 N^{-1}\,d\xi\,d\eta\Big)^{\frac{1}{2}},\end{aligned}$$

which implies (2.27). □

Corollary 2.10 (Bilinear Strichartz II). *Let M, N, and d be as above. Given any space-time slab $I \times \mathbb{R}^d$ and any functions u, v defined on $I \times \mathbb{R}^d$,*

$$\big\|u_{\leq M} v_{\geq N}\big\|_{L^2_{t,x}(I\times\mathbb{R}^d)} \lesssim M^{\frac{d-3}{2}} N^{-\frac{1}{2}} \|\nabla u_{\leq M}\|_{S^*_0(I)} \|v_{\geq N}\|_{S^*_0(I)},$$

where we use the notation

$$\|u\|_{S^*_0(I)} := \|u\|_{L^\infty_t L^2_x(I\times\mathbb{R}^d)} + \|(i\partial_t+\Delta)u\|_{L^{\frac{2(d+2)}{d+4}}_{t,x}(I\times\mathbb{R}^d)}.$$

Proof. See [39, Lemma 2.5], which builds on earlier versions in [4, 13]. $\square$

Lemma 2.11 (Local smoothing, [14, 32, 38]). *For all $f \in L^2_x$ we have*

$$\int_{\mathbb{R}} \int_{\mathbb{R}^d} \big| \big[|\nabla|^{1/2} e^{it\Delta} f \big](x) \big|^2 e^{-|x|^2} \, dx \, dt \lesssim \|f\|^2_{L^2_x}.$$

In particular, by scaling, for all $R > 0$ we have

$$\big\| |\nabla|^{1/2} e^{it\Delta} f \big\|_{L^2_{t,x}(\mathbb{R}\times B(0,R))} \lesssim R^{1/2} \|f\|_{L^2_x}.$$

Proof. Given $a : \mathbb{R}^d \to [0, \infty)$, we have

$$\begin{aligned}
&\int_{\mathbb{R}} \int_{\mathbb{R}^d} \big| \big[|\nabla|^{1/2} e^{it\Delta} f \big](x) \big|^2 a(x) \, dx \, dt \\
&= (2\pi)^{-d} \int_{\mathbb{R}} \int_{\mathbb{R}^d} \int_{\mathbb{R}^d} \int_{\mathbb{R}^d} e^{ix\xi - it|\xi|^2} |\xi|^{1/2} \hat{f}(\xi) e^{-ix\eta + it|\eta|^2} |\eta|^{1/2} \overline{\hat{f}(\eta)} a(x) \, d\xi \, d\eta \, dx \, dt \\
&= \int_{\mathbb{R}^d} \int_{\mathbb{R}^d} \hat{a}(\eta - \xi) \delta(|\eta|^2 - |\xi|^2) |\xi|^{1/2} |\eta|^{1/2} \hat{f}(\xi) \overline{\hat{f}(\eta)} \, d\xi \, d\eta \\
&= \int_{\mathbb{R}^d} \int_{\mathbb{R}^d} \hat{a}(\eta - \xi) \delta(|\eta| - |\xi|) \frac{|\xi|^{1/2}|\eta|^{1/2}}{|\xi| + |\eta|} \hat{f}(\xi) \overline{\hat{f}(\eta)} \, d\xi \, d\eta.
\end{aligned}$$

By Schur's test it thus suffices to show that

$$\int_{\mathbb{R}^d} \hat{a}(\eta - \xi) \delta(|\eta| - |\xi|) \frac{|\xi|^{1/2}|\eta|^{1/2}}{|\xi| + |\eta|} \, d\xi \lesssim 1 \quad \text{uniformly in} \quad \eta \in \mathbb{R}^d. \tag{2.29}$$

Recalling that in our case $a(x) = e^{-|x|^2}$ and passing to polar coordinates, we obtain

$$\begin{aligned}
&\int_{\mathbb{R}^d} \hat{a}(\eta - \xi) \delta(|\eta| - |\xi|) \frac{|\xi|^{1/2}|\eta|^{1/2}}{|\xi| + |\eta|} \, d\xi \\
&\qquad \lesssim \int_{S^{d-1}} \int_0^\infty e^{-|r\omega - \eta|^2} \delta(|\eta| - r) \frac{r^{1/2}|\eta|^{1/2}}{r + |\eta|} r^{d-1} \, dr \, d\sigma(\omega) \\
&\qquad \lesssim \int_{S^{d-1}} \int_0^\infty e^{-|\eta|^2 \left|\omega - \frac{\eta}{|\eta|}\right|^2} |\eta|^{d-1} \, d\sigma(\omega) \\
&\qquad \lesssim \int_0^\pi e^{-2|\eta|^2(1-\cos\theta)} |\eta|^{d-1} (\sin\theta)^{d-2} \, d\theta \\
&\qquad \lesssim \int_0^{\frac{\pi}{2}} e^{-\frac{|\eta|^2\theta^2}{100}} |\eta|^{d-1} \theta^{d-2} \, d\theta \lesssim \int_0^\infty e^{-\frac{\tau^2}{100}} \tau^{d-2} \, d\tau \lesssim 1.
\end{aligned}$$

In the computation above, θ denotes the angle ω makes with $\frac{\eta}{|\eta|}$. This proves (2.29) and so completes the proof of the lemma. $\square$

The next result is a consequence of local smoothing; see Lemma 3.7 in [18]. The proof we present here is the one from [23]; see also [22].

Lemma 2.12. *Given $\phi \in \dot{H}^1(\mathbb{R}^d)$,*

$$\|\nabla e^{it\Delta}\phi\|^3_{L^2_{t,x}([-T,T]\times\{|x|\le R\})} \lesssim T^{\frac{2}{d+2}} R^{\frac{3d+2}{d+2}} \|e^{it\Delta}\phi\|_{L_{t,x}^{\frac{2(d+2)}{d-2}}} \|\nabla\phi\|^2_{L^2_x}.$$

Proof. Given $N > 0$, Hölder's and Bernstein's inequalities imply

$$\begin{aligned}\|\nabla e^{it\Delta}\phi_{<N}\|_{L^2_{t,x}([-T,T]\times\{|x|\le R\})} &\lesssim T^{\frac{2}{d+2}} R^{\frac{2d}{d+2}} \|e^{it\Delta}\nabla\phi_{<N}\|_{L_{t,x}^{\frac{2(d+2)}{d-2}}} \\ &\lesssim T^{\frac{2}{d+2}} R^{\frac{2d}{d+2}} N \|e^{it\Delta}\phi\|_{L_{t,x}^{\frac{2(d+2)}{d-2}}}.\end{aligned}$$

On the other hand, the high frequencies can be estimated using local smoothing:

$$\begin{aligned}\|\nabla e^{it\Delta}\phi_{\ge N}\|_{L^2_{t,x}([-T,T]\times\{|x|\le R\})} &\lesssim R^{1/2}\||\nabla|^{1/2}\phi_{\ge N}\|_{L^2_x} \\ &\lesssim N^{-1/2}R^{1/2}\|\nabla\phi\|_{L^2_x}.\end{aligned}$$

The lemma now follows by optimizing the choice of N. □

Chapter 3

An inverse Strichartz inequality

In this section, we develop tools that we will employ to prove a linear profile decomposition for the Schrödinger propagator for bounded sequences in $\dot{H}^1(\mathbb{R}^d)$ with $d \geq 3$. Such a linear profile decomposition was first obtained by Keraani [18], relying on an improved Sobolev inequality proved by Gérard, Meyer, and Oru [16]. We should also note the influential precursor [1], which treated the wave equation. In these notes we present a different proof of the result in [18], which relies instead on an inverse Strichartz inequality.

A linear profile decomposition for the Schrödinger propagator for bounded sequences in $L^2(\mathbb{R}^d)$ was proved by Merle and Vega [26] for $d = 2$, Carles and Keraani [7] for $d = 1$, and Bégout and Vargas [2] for $d \geq 3$. For a different approach to these results, which is similar in spirit to what we present in these notes, see [22].

We start by noting that combining the Strichartz inequality for the Schrödinger propagator from Lemma 2.7 and Sobolev embedding, we obtain

$$\|e^{it\Delta} f\|_{L_{t,x}^{\frac{2(d+2)}{d-2}}(\mathbb{R}\times\mathbb{R}^d)} \lesssim \|e^{it\Delta}\nabla f\|_{L_t^{\frac{2(d+2)}{d-2}} L_x^{\frac{2d(d+2)}{d^2+8}}(\mathbb{R}\times\mathbb{R}^d)} \lesssim \|f\|_{\dot{H}_x^1} \tag{3.1}$$

for all $d \geq 3$.

Our next result is a refinement of (3.1), which says that if the linear evolution of f is large in $L_{t,x}^{\frac{2(d+2)}{d-2}}$, then the linear evolution of a single Littlewood–Paley piece of f is, at least partially, responsible.

Lemma 3.1 (Refined Strichartz estimate)**.** *Let $d \geq 3$ and $f \in \dot{H}^1(\mathbb{R}^d)$. Then*

$$\|e^{it\Delta} f\|_{L_{t,x}^{\frac{2(d+2)}{d-2}}(\mathbb{R}\times\mathbb{R}^d)} \lesssim \|f\|_{\dot{H}_x^1}^{\frac{d-2}{d+2}} \sup_{N\in 2^{\mathbb{Z}}} \|e^{it\Delta} f_N\|_{L_{t,x}^{\frac{2(d+2)}{d-2}}(\mathbb{R}\times\mathbb{R}^d)}^{\frac{4}{d+2}}.$$

Proof. We will present the proof in dimensions $d \geq 6$. The proof in dimensions $d = 3, 4$ is easier, as $\frac{2(d+2)}{d-2}$ is an even integer in those cases. The proof in dimension $d = 5$ is a small modification of the argument below. We leave the cases $d = 3, 4, 5$ as an exercise for the conscientious reader.

H. Koch et al., *Dispersive Equations and Nonlinear Waves: Generalized Korteweg–de Vries, Nonlinear Schrödinger, Wave and Schrödinger Maps*, Oberwolfach Seminars 45,
DOI 10.1007/978-3-0348-0736-4_17, © Springer Basel 2014

Fix $d \geq 6$. From the square function estimate, the subaditivity of fractional powers (using the fact that $\frac{d+2}{2(d-2)} \leq 1$ in dimensions $d \geq 6$), and the Bernstein and Strichartz inequalities,

$$
\begin{aligned}
&\|e^{it\Delta} f\|_{L_{t,x}^{\frac{2(d+2)}{d-2}}}^{\frac{2(d+2)}{d-2}} \\
&\quad\lesssim \iint_{\mathbb{R}\times\mathbb{R}^d} \Big(\sum_{N\in 2^{\mathbb{Z}}} |e^{it\Delta} f_N|^2 \Big)^{\frac{d+2}{d-2}} \, dx\, dt \\
&\quad\lesssim \sum_{M\leq N} \iint_{\mathbb{R}\times\mathbb{R}^d} |e^{it\Delta} f_M|^{\frac{d+2}{d-2}} |e^{it\Delta} f_N|^{\frac{d+2}{d-2}} \, dx\, dt \\
&\quad\lesssim \sum_{M\leq N} \|e^{it\Delta} f_M\|_{L_{t,x}^{\frac{2(d+2)}{d-4}}} \|e^{it\Delta} f_M\|_{L_{t,x}^{\frac{2(d+2)}{d-2}}}^{\frac{4}{d-2}} \|e^{it\Delta} f_N\|_{L_{t,x}^{\frac{2(d+2)}{d-2}}}^{\frac{4}{d-2}} \|e^{it\Delta} f_N\|_{L_{t,x}^{\frac{2(d+2)}{d}}} \\
&\quad\lesssim \sup_{N\in 2^{\mathbb{Z}}} \|e^{it\Delta} f_N\|_{L_{t,x}^{\frac{2(d+2)}{d-2}}}^{\frac{8}{d-2}} \sum_{M\leq N} M^2 \|e^{it\Delta} f_M\|_{L_t^{\frac{2(d+2)}{d-4}} L_x^{\frac{2d(d+2)}{d^2+8}}} \|f_N\|_{L_x^2} \\
&\quad\lesssim \sup_{N\in 2^{\mathbb{Z}}} \|e^{it\Delta} f_N\|_{L_{t,x}^{\frac{2(d+2)}{d-2}}}^{\frac{8}{d-2}} \sum_{M\leq N} M^2 \|f_M\|_{L_x^2} \|f_N\|_{L_x^2} \\
&\quad\lesssim \sup_{N\in 2^{\mathbb{Z}}} \|e^{it\Delta} f_N\|_{L_{t,x}^{\frac{2(d+2)}{d-2}}}^{\frac{8}{d-2}} \sum_{M\leq N} \tfrac{M}{N} \|\nabla f_M\|_{L_x^2} \|\nabla f_N\|_{L_x^2} \\
&\quad\lesssim \sup_{N\in 2^{\mathbb{Z}}} \|e^{it\Delta} f_N\|_{L_{t,x}^{\frac{2(d+2)}{d-2}}}^{\frac{8}{d-2}} \|f\|_{\dot H_x^1}^2 .
\end{aligned}
$$

This completes the proof of the lemma in dimensions $d \geq 6$. □

The refined Strichartz inequality shows that linear solutions with non-trivial space-time norm must concentrate on at least one frequency annulus. The next proposition goes one step further and shows that they contain a bubble of concentration around some point in space-time.

Proposition 3.2 (Inverse Strichartz inequality)**.** *Let $d \geq 3$ and let $\{f_n\} \subset \dot H^1(\mathbb{R}^d)$. Suppose that*

$$\lim_{n\to\infty} \|f_n\|_{\dot H_x^1} = A < \infty \quad \textit{and} \quad \lim_{n\to\infty} \|e^{it\Delta} f_n\|_{L_{t,x}^{\frac{2(d+2)}{d-2}}(\mathbb{R}\times\mathbb{R}^d)} = \varepsilon > 0.$$

Then there exist a subsequence in n, $\phi \in \dot H_x^1$, $\{\lambda_n\} \subset (0,\infty)$, and $\{(t_n, x_n)\} \subset \mathbb{R}\times\mathbb{R}^d$ such that

$$\lambda_n^{\frac{d-2}{2}} [e^{it_n\Delta} f_n](\lambda_n x + x_n) \rightharpoonup \phi(x) \quad \textit{weakly in} \quad \dot H_x^1, \tag{3.2}$$

$$\liminf_{n\to\infty} \Big\{ \|f_n\|_{\dot H_x^1}^2 - \|f_n - \phi_n\|_{\dot H_x^1}^2 \Big\} = \|\phi\|_{\dot H_x^1}^2 \gtrsim A^2 (\tfrac{\varepsilon}{A})^{\frac{d(d+2)}{4}}, \tag{3.3}$$

$$\liminf_{n\to\infty} \Big\{ \|e^{it\Delta} f_n\|_{L_{t,x}^{\frac{2(d+2)}{d-2}}}^{\frac{2(d+2)}{d-2}} - \|e^{it\Delta}(f_n - \phi_n)\|_{L_{t,x}^{\frac{2(d+2)}{d-2}}}^{\frac{2(d+2)}{d-2}} \Big\} \gtrsim \varepsilon^{\frac{2(d+2)}{d-2}} (\tfrac{\varepsilon}{A})^{\frac{(d+2)(d+4)}{4}}, \tag{3.4}$$

where

$$\phi_n(x) := \lambda_n^{-\frac{d-2}{2}} [e^{-i\lambda_n^{-2} t_n \Delta} \phi]\big(\tfrac{x-x_n}{\lambda_n}\big). \tag{3.5}$$

Proof. Passing to a subsequence, we may assume that

$$\lim_{n\to\infty} \|f_n\|_{\dot H^1_x} \le 2A \quad \text{and} \quad \lim_{n\to\infty} \|e^{it\Delta} f_n\|_{L_{t,x}^{\frac{2(d+2)}{d-2}}(\mathbb{R}\times\mathbb{R}^d)} \ge \tfrac{\varepsilon}{2}.$$

Thus, using Lemma 3.1 we see that for each n there exists $N_n \in 2^{\mathbb{Z}}$ such that

$$\|e^{it\Delta} P_{N_n} f_n\|_{L_{t,x}^{\frac{2(d+2)}{d-2}}} \gtrsim \varepsilon^{\frac{d+2}{4}} A^{-\frac{d-2}{4}}.$$

On the other hand, from the Strichartz and Bernstein inequalities we get

$$\|e^{it\Delta} P_{N_n} f_n\|_{L_{t,x}^{\frac{2(d+2)}{d}}} \lesssim \|P_{N_n} f_n\|_{L^2_x} \lesssim N_n^{-1} A.$$

By Hölder's inequality, these imply

$$\begin{aligned} \varepsilon^{\frac{d+2}{4}} A^{-\frac{d-2}{4}} &\lesssim \|e^{it\Delta} P_{N_n} f_n\|_{L_{t,x}^{\frac{2(d+2)}{d-2}}} \\ &\lesssim \|e^{it\Delta} P_{N_n} f_n\|_{L_{t,x}^{\frac{2(d+2)}{d}}}^{\frac{d-2}{d}} \|e^{it\Delta} P_{N_n} f_n\|_{L^\infty_{t,x}}^{\frac{2}{d}} \\ &\lesssim N_n^{-\frac{d-2}{d}} A^{\frac{d-2}{d}} \|e^{it\Delta} P_{N_n} f_n\|_{L^\infty_{t,x}}^{\frac{2}{d}}, \end{aligned}$$

and so

$$N_n^{-\frac{d-2}{2}} \|e^{it\Delta} P_{N_n} f_n\|_{L^\infty_{t,x}} \gtrsim A\big(\tfrac{\varepsilon}{A}\big)^{\frac{d(d+2)}{8}}.$$

Thus there exist $(t_n, x_n) \in \mathbb{R}\times\mathbb{R}^d$ such that

$$N_n^{-\frac{d-2}{2}} \big|[e^{it_n\Delta} P_{N_n} f_n](x_n)\big| \gtrsim A\big(\tfrac{\varepsilon}{A}\big)^{\frac{d(d+2)}{8}}. \tag{3.6}$$

We define the spatial scales $\lambda_n := N_n^{-1}$.

It remains to find the profile ϕ and to prove it satisfies (3.2) through (3.4). To this end, we set

$$g_n(x) := \lambda_n^{\frac{d-2}{2}} [e^{it_n\Delta} f_n](\lambda_n x + x_n).$$

A simple change of variables gives

$$\|g_n\|_{\dot H^1_x} = \|f_n\|_{\dot H^1_x} \lesssim A$$

and so, passing to a subsequence, we can choose ϕ so that $g_n \rightharpoonup \phi$ weakly in $\dot H^1_x$. This proves (3.2).

We now turn to (3.3). The asymptotic decoupling statement is immediate since $\dot H^1_x$ is a Hilbert space. We are left to prove the lower bound in (3.3). Toward this end, let $\check\psi := P_1\delta_0$ denote the convolution kernel associated with P_1. Using a change of variables and (3.6), we get

$$\begin{aligned}\big|\langle\phi,\check\psi\rangle_{L^2_x}\big| = \big|\lim_{n\to\infty}\langle g_n,\check\psi\rangle_{L^2_x}\big| &= \big|\lim_{n\to\infty}\big\langle e^{it_n\Delta}f_n, \lambda_n^{-\frac{d+2}{2}}\check\psi\big(\tfrac{x-x_n}{\lambda_n}\big)\big\rangle_{L^2_x}\big|\\ &= N_n^{-\frac{d-2}{2}}\big|[e^{it_n\Delta}P_{N_n}f_n](x_n)\big| \gtrsim A\big(\tfrac{\varepsilon}{A}\big)^{\frac{d(d+2)}{8}}. \end{aligned}\tag{3.7}$$

On the other hand, by Hölder's inequality and Sobolev embedding,

$$\big|\langle\phi,\check\psi\rangle_{L^2_x}\big| \lesssim \|\phi\|_{L^6_x}\|\check\psi\|_{L^{6/5}_x} \lesssim \|\phi\|_{\dot H^1_x}.$$

Putting the two inequalities together, we derive the lower bound in (3.3).

It remains to prove (3.4). We start by proving decoupling for the $L^{\frac{2(d+2)}{d-2}}_{t,x}$ norm. Note that

$$(i\partial_t)^{\frac12}e^{it\Delta} = (-\Delta)^{\frac12}e^{it\Delta},$$

as can be checked by testing against Schwartz functions in $\mathbb{R}\times\mathbb{R}^d$. Thus, by Hölder's inequality, on any compact set K in $\mathbb{R}\times\mathbb{R}^d$ we have

$$\|e^{it\Delta}g_n\|_{H^{\frac12}_{t,x}(K)} \lesssim \|\langle -\Delta\rangle^{\frac12}e^{it\Delta}g_n\|_{L^2_{t,x}(K)} \lesssim_K A.$$

Using this together with Rellich–Kondrashov and passing to a subsequence, we get

$$e^{it\Delta}g_n \to e^{it\Delta}\phi \quad \text{strongly in} \quad L^2_{t,x}(K).$$

(In order to identify the limit in the display above, we note that $g_n \rightharpoonup \phi$ weakly in $\dot H^1_x$ implies that $e^{it\Delta}g_n$ converges to $e^{it\Delta}\phi$ as distributions on $\mathbb{R}\times\mathbb{R}^d$.) Passing to a further subsequence, we deduce that $e^{it\Delta}g_n \to e^{it\Delta}\phi$ a.e. on K. Finally, using a diagonal argument and passing again to a subsequence if necessary, we obtain

$$e^{it\Delta}g_n \to e^{it\Delta}\phi \quad \text{a.e. in } \mathbb{R}\times\mathbb{R}^d.$$

To continue, we use this convergence together with the refined Fatou lemma (see Lemma 11.3) due to Brezis and Lieb and a change of variables; we obtain

$$\lim_{n\to\infty}\Big\{\|e^{it\Delta}f_n\|^{\frac{2(d+2)}{d-2}}_{L^{\frac{2(d+2)}{d-2}}_{t,x}} - \|e^{it\Delta}(f_n-\phi_n)\|^{\frac{2(d+2)}{d-2}}_{L^{\frac{2(d+2)}{d-2}}_{t,x}}\Big\} = \|e^{it\Delta}\phi\|^{\frac{2(d+2)}{d-2}}_{L^{\frac{2(d+2)}{d-2}}_{t,x}},$$

from which (3.4) will follow once we prove

$$\|e^{it\Delta}\phi\|_{L^{\frac{2(d+2)}{d-2}}_{t,x}} \gtrsim \varepsilon\big(\tfrac{\varepsilon}{A}\big)^{\frac{d^2+2d-8}{8}}. \tag{3.8}$$

To see this, we use (3.7), the Mikhlin multiplier theorem, and Bernstein to estimate

$$A(\tfrac{\varepsilon}{A})^{\frac{d(d+2)}{8}} \lesssim \big|\langle \phi, \check{\psi}\rangle_{L^2_x}\big| = \big|\langle e^{it\Delta}\phi, e^{it\Delta}\check{\psi}\rangle_{L^2_x}\big| \lesssim \|e^{it\Delta}\phi\|_{L_x^{\frac{2(d+2)}{d-2}}} \|e^{it\Delta}\check{\psi}\|_{L_x^{\frac{2(d+2)}{d+6}}}$$
$$\lesssim \|e^{it\Delta}\phi\|_{L_x^{\frac{2(d+2)}{d-2}}},$$

uniformly in $|t| \leq 1$. Integrating in t leads to (3.8). □

Exercise 3.1. Under the hypotheses of Proposition 3.2 and passing to a further subsequence if necessary, prove decoupling of the potential energy, namely,

$$\liminf_{n\to\infty}\Big\{\|f_n\|_{\frac{2d}{d-2}}^{\frac{2d}{d-2}} - \|f_n - \phi_n\|_{\frac{2d}{d-2}}^{\frac{2d}{d-2}} - \|e^{-i\lambda_n^{-2}t_n\Delta}\phi\|_{\frac{2d}{d-2}}^{\frac{2d}{d-2}}\Big\} = 0.$$

Hint: Passing to a subsequence, we may assume that $\lambda_n^{-2}t_n \to t_0 \in [-\infty, \infty]$. If $t_0 = \pm\infty$, then approximate ϕ in $\dot{H}^1_x$ by Schwartz functions and use the fact that, by the dispersive estimate for the Schrödinger propagator,

$$\|e^{-i\lambda_n^{-2}t_n\Delta}\psi\|_{\frac{2d}{d-2}} \to 0 \quad \text{as} \quad n \to \infty$$

for any $\psi \in \mathcal{S}(\mathbb{R}^d)$. If instead $t_0 \in (-\infty, \infty)$, then (3.2) can be upgraded to $\lambda_n^{\frac{d-2}{2}} f_n(\lambda_n x + x_n) \rightharpoonup e^{-it_0\Delta}\phi(x)$ weakly in $\dot{H}^1_x$. Now use Rellich–Kondrashov and refined Fatou as in the proof of (3.4).

Chapter 4

A linear profile decomposition

In this section, we use the inverse Strichartz inequality Proposition 3.2 to derive a linear profile decomposition for the Schrödinger propagator.

One can view the linear profile decomposition as a tool for measuring the defects of compactness in the Strichartz inequality (3.1). More precisely, given a bounded sequence of functions $\{f_n\}_{n\geq 1} \subset \dot{H}^1(\mathbb{R}^d)$ we would like to be able to say that, after possibly passing to a subsequence, $\{e^{it\Delta} f_n\}_{n\geq 1}$ converges in $L_{t,x}^{\frac{2(d+2)}{d-2}}$. Unfortunately, every non-compact symmetry of the inequality (3.1) is a reason why we would fail to extract a convergent subsequence.

The non-compact symmetries of (3.1) are space and time translations and $\dot{H}^1_x$-preserving scaling. To see how these work against us, consider the simple scenario where $f_n(x) = f(x + x_n)$ with $f \in \dot{H}^1_x$ and $\{x_n\}_{n\geq 1} \subset \mathbb{R}^d$ is a sequence that diverges to infinity; in this case, $\{e^{it\Delta} f_n\}_{n\geq 1}$ converges weakly to zero. We leave it to the reader to use time translations and $\dot{H}^1_x$-preserving scaling to construct bounded sequences of functions $\{f_n\}_{n\geq 1} \subset \dot{H}^1(\mathbb{R}^d)$ for which $\{e^{it\Delta} f_n\}_{n\geq 1}$ converges weakly to zero.

At this point we might imagine that if suitably translate and rescale our sequence, then we might be able to extract a convergent subsequence. Proposition 3.2 gives us hope, since it exhibits a bubble of concentration living inside each $e^{it\Delta} f_n$, which captures a nontrivial portion of the $L_{t,x}^{\frac{2(d+2)}{d-2}}$ norm of $e^{it\Delta} f_n$. However, even this modified goal is naive and doomed to fail, as one can see by considering the following scenario: $f_n(x) = f(x) + f(x + x_n)$ with $f \in \dot{H}^1_x$ and $\{x_n\}_{n\geq 1} \subset \mathbb{R}^d$ is a sequence that diverges to infinity; in this case, the evolutions $e^{it\Delta} f_n$ contain two diverging bubbles of concentration and translating our sequence would still fail to exhibit a convergent subsequence.

Nevertheless, this suggests that if we take out enough bubbles of concentration living inside $e^{it\Delta} f_n$, then we might be able to say that the remainders do indeed converge to zero in $L_{t,x}^{\frac{2(d+2)}{d-2}}$. This is precisely the content of the following theorem.

H. Koch et al., *Dispersive Equations and Nonlinear Waves: Generalized Korteweg–de Vries, Nonlinear Schrödinger, Wave and Schrödinger Maps*, Oberwolfach Seminars 45,
DOI 10.1007/978-3-0348-0736-4_18,

Theorem 4.1 ($\dot H^1_x$ linear profile decomposition for the Schrödinger propagator)**.** *Fix $d \geq 3$ and let $\{f_n\}_{n\geq 1}$ be a sequence of functions bounded in $\dot H^1(\mathbb{R}^d)$. Passing to a subsequence if necessary, there exist $J^* \in \{0,1,\dots\}\cup\{\infty\}$, functions $\{\phi^j\}_{j=1}^{J^*} \subset \dot H^1(\mathbb{R}^d)$, $\{\lambda_n^j\} \subset (0,\infty)$, and $\{t_n^j, x_n^j\} \subset \mathbb{R}\times\mathbb{R}^d$ such that for each finite $0 \leq J \leq J^*$, we have the decomposition*

$$f_n = \sum_{j=1}^{J} (\lambda_n^j)^{-\frac{d-2}{2}} [e^{it_n^j\Delta}\phi^j]\big(\tfrac{x-x_n^j}{\lambda_n^j}\big) + w_n^J, \tag{4.1}$$

with the following properties:

$$\lim_{J\to J^*} \limsup_{n\to\infty} \big\| e^{it\Delta} w_n^J \big\|_{L_{t,x}^{\frac{2(d+2)}{d-2}}(\mathbb{R}\times\mathbb{R}^d)} = 0, \tag{4.2}$$

$$\lim_{n\to\infty} \Big[\|\nabla f_n\|_2^2 - \sum_{j=1}^{J} \|\nabla \phi^j\|_2^2 - \|\nabla w_n^J\|_2^2 \Big] = 0, \tag{4.3}$$

$$\lim_{n\to\infty} \Big[\|f_n\|_{\frac{2d}{d-2}}^{\frac{2d}{d-2}} - \sum_{j=1}^{J} \|e^{it_n^j\Delta}\phi^j\|_{\frac{2d}{d-2}}^{\frac{2d}{d-2}} - \|w_n^J\|_{\frac{2d}{d-2}}^{\frac{2d}{d-2}} \Big] = 0, \tag{4.4}$$

$$e^{-it_n^J\Delta}\big[(\lambda_n^J)^{\frac{d-2}{2}} w_n^J(\lambda_n^J x + x_n^J)\big] \rightharpoonup 0 \quad \textit{weakly in } \dot H^1(\mathbb{R}^d). \tag{4.5}$$

Moreover, for each $j \neq k$ we have the following asymptotic decoupling of parameters:

$$\frac{\lambda_n^j}{\lambda_n^k} + \frac{\lambda_n^k}{\lambda_n^j} + \frac{|x_n^j - x_n^k|^2}{\lambda_n^j\lambda_n^k} + \frac{\big|t_n^j(\lambda_n^j)^2 - t_n^k(\lambda_n^k)^2\big|}{\lambda_n^j\lambda_n^k} \to \infty \quad \textit{as} \quad n \to \infty. \tag{4.6}$$

Lastly, we may additionally assume that for each j either $t_n^j \equiv 0$ or $t_n^j \to \pm\infty$.

Proof. To keep formulas within margins, we will use the notation

$$(g_n^j f)(x) := (\lambda_n^j)^{-\frac{d-2}{2}} f\big(\tfrac{x-x_n^j}{\lambda_n^j}\big) \quad \text{and} \quad [(g_n^j)^{-1} f](x) := (\lambda_n^j)^{\frac{d-2}{2}} f\big(\lambda_n^j x + x_n^j\big).$$

Note that $\|g_n^j f\|_{\dot H^1_x} = \|f\|_{\dot H^1_x} = \|(g_n^j)^{-1} f\|_{\dot H^1_x}$ and

$$\langle g_n^j f_1, f_2\rangle_{\dot H^1_x} = \langle f_1, (g_n^j)^{-1} f_2\rangle_{\dot H^1_x} \quad \text{for all} \quad f_1, f_2 \in \dot H^1_x.$$

We will also use the notation

$$\phi_n^j(x) := (\lambda_n^j)^{-\frac{d-2}{2}} [e^{it_n^j\Delta}\phi^j]\big(\tfrac{x-x_n^j}{\lambda_n^j}\big) = [g_n^j e^{it_n^j\Delta}\phi^j](x).$$

To prove the theorem we will proceed inductively, extracting one bubble at a time. To start, we set $w_n^0 := f_n$. Now suppose we have a decomposition up to level $J \geq 0$ obeying (4.3) through (4.5). (Conditions (4.2) and (4.6) will be verified at the end.) Passing to a subsequence if necessary, we set

$$A_J := \lim_{n\to\infty} \|w_n^J\|_{\dot H^1_x} \quad \text{and} \quad \varepsilon_J := \lim_{n\to\infty} \|e^{it\Delta} w_n^J\|_{L_{t,x}^{\frac{2(d+2)}{d-2}}}.$$

If $\varepsilon_J = 0$, we stop and set $J^* = J$. If not, we apply Proposition 3.2 to w_n^J. Thus, passing to a subsequence in n, we find $\phi^{J+1} \in \dot{H}_x^1$, $\{\lambda_n^{J+1}\} \subset (0, \infty)$, and $\{(t_n^{J+1}, x_n^{J+1})\} \subset \mathbb{R} \times \mathbb{R}^d$, where we renamed the time parameters given by Proposition 3.2 as follows: $t_n^{J+1} = -\lambda_n^{-2} t_n$.

According to Proposition 3.2, the profile ϕ^{J+1} is defined as a weak limit, namely,

$$\phi^{J+1} = \operatorname*{w-lim}_{n\to\infty} (g_n^{J+1})^{-1} \big[e^{-it_n^{J+1}(\lambda_n^{J+1})^2 \Delta} w_n^J \big] = \operatorname*{w-lim}_{n\to\infty} e^{-it_n^{J+1}\Delta} [(g_n^{J+1})^{-1} w_n^J].$$

We let $\phi_n^{J+1} := g_n^{J+1} e^{it_n^{J+1}\Delta} \phi^{J+1}$.

Now define $w_n^{J+1} := w_n^J - \phi_n^{J+1}$. By the definition of ϕ^{J+1},

$$e^{-it_n^{J+1}\Delta} (g_n^{J+1})^{-1} w_n^{J+1} \rightharpoonup 0 \quad \text{weakly in } \dot{H}_x^1.$$

This proves (4.5) at the level $J+1$. Moreover, from Proposition 3.2 we also have

$$\lim_{n\to\infty} \Big\{ \|w_n^J\|_{\dot{H}_x^1}^2 - \|w_n^{J+1}\|_{\dot{H}_x^1}^2 - \|\phi^{J+1}\|_{\dot{H}_x^1}^2 \Big\} = 0.$$

Combining this with the inductive hypothesis gives (4.3) at the level $J+1$. A similar argument using Exercise 3.1 establishes (4.4) at the same level.

Passing to a further subsequence and using Proposition 3.2, we obtain

$$\begin{aligned} A_{J+1}^2 &= \lim_{n\to\infty} \|w_n^{J+1}\|_{\dot{H}_x^1}^2 \le A_J^2 \Big[1 - C\big(\tfrac{\varepsilon_J}{A_J}\big)^{\frac{d(d+2)}{4}} \Big] \le A_J^2, \\ \varepsilon_{J+1}^{\frac{2(d+2)}{d-2}} &= \lim_{n\to\infty} \|e^{it\Delta} w_n^{J+1}\|_{L_{t,x}^{\frac{2(d+2)}{d-2}}}^{\frac{2(d+2)}{d-2}} \le \varepsilon_J^{\frac{2(d+2)}{d-2}} \Big[1 - C\big(\tfrac{\varepsilon_J}{A_J}\big)^{\frac{(d+2)(d+4)}{4}} \Big]. \end{aligned} \tag{4.7}$$

If $\varepsilon_{J+1} = 0$, we stop and set $J^* = J+1$; in this case, (4.2) is automatic. If $\varepsilon_{J+1} > 0$, we continue the induction. If the algorithm does not terminate in finitely many steps, we set $J^* = \infty$; in this case, (4.7) implies $\varepsilon_J \to 0$ as $J \to \infty$ and so (4.2) follows.

Next we verify the asymptotic orthogonality condition (4.6). We argue by contradiction. Assume (4.6) fails to be true for some pair (j, k). Without loss of generality, we may assume that this is the first pair for which (4.6) fails, that is, $j < k$ and (4.6) holds for all pairs (j, l) with $j < l < k$. Passing to a subsequence, we may assume that

$$\frac{\lambda_n^j}{\lambda_n^k} \to \lambda_0 \in (0, \infty), \quad \frac{x_n^j - x_n^k}{\sqrt{\lambda_n^j \lambda_n^k}} \to x_0, \quad \text{and} \quad \frac{t_n^j (\lambda_n^j)^2 - t_n^k (\lambda_n^k)^2}{\lambda_n^j \lambda_n^k} \to t_0. \tag{4.8}$$

From the inductive relation

$$w_n^{k-1} = w_n^j - \sum_{l=j+1}^{k-1} \phi_n^l$$

and the definition of ϕ^k, we obtain

$$\begin{aligned}\phi^k &= \operatorname*{w-lim}_{n\to\infty} e^{-it_n^k\Delta}[(g_n^k)^{-1}w_n^{k-1}] \\ &= \operatorname*{w-lim}_{n\to\infty} e^{-it_n^k\Delta}[(g_n^k)^{-1}w_n^{j}] - \sum_{l=j+1}^{k-1} \operatorname*{w-lim}_{n\to\infty} e^{-it_n^k\Delta}[(g_n^k)^{-1}\phi_n^{l}]. \end{aligned} \tag{4.9}$$

We will prove that these weak limits are all zero and so obtain a contradiction to the nontriviality of ϕ^k.

We write

$$\begin{aligned} e^{-it_n^k\Delta}[(g_n^k)^{-1}w_n^j] &= e^{-it_n^k\Delta}(g_n^k)^{-1}g_n^j e^{it_n^j\Delta}[e^{-it_n^j\Delta}(g_n^j)^{-1}w_n^j] \\ &= (g_n^k)^{-1}g_n^j e^{i\left(t_n^j - t_n^k\frac{(\lambda_n^k)^2}{(\lambda_n^j)^2}\right)\Delta}[e^{-it_n^j\Delta}(g_n^j)^{-1}w_n^j]. \end{aligned}$$

Note that, by (4.8),

$$t_n^j - t_n^k\frac{(\lambda_n^k)^2}{(\lambda_n^j)^2} = \frac{t_n^j(\lambda_n^j)^2 - t_n^k(\lambda_n^k)^2}{\lambda_n^j\lambda_n^k}\cdot\frac{\lambda_n^k}{\lambda_n^j} \to \frac{t_0}{\lambda_0}.$$

Using this together with (4.5), Exercise 4.2, and the fact that the adjoints of the unitary operators $(g_n^k)^{-1}g_n^j$ converge strongly, we obtain that the first term on RHS(4.9) is zero.

To complete the proof of (4.6), it remains to show that the second term on RHS(4.9) is zero. For all $j<l<k$ we write

$$e^{-it_n^k\Delta}(g_n^k)^{-1}\phi_n^l = (g_n^k)^{-1}g_n^j e^{i\left(t_n^j - t_n^k\frac{(\lambda_n^k)^2}{(\lambda_n^j)^2}\right)\Delta}[e^{-it_n^j\Delta}(g_n^j)^{-1}\phi_n^l].$$

Arguing as for the first term on RHS(4.9), it thus suffices to show that

$$e^{-it_n^j\Delta}(g_n^j)^{-1}\phi_n^l = e^{-it_n^j\Delta}(g_n^j)^{-1}g_n^l e^{it_n^l\Delta}\phi^l \rightharpoonup 0 \quad \text{weakly in } \dot H_x^1.$$

Using a density argument, this reduces to

$$I_n := e^{-it_n^j\Delta}(g_n^j)^{-1}g_n^l e^{it_n^l\Delta}\phi \rightharpoonup 0 \quad \text{weakly in } \dot H_x^1, \tag{4.10}$$

for all $\phi\in C_c^\infty(\mathbb{R}^d)$. Note that we can rewrite I_n as follows:

$$I_n = \left(\frac{\lambda_n^j}{\lambda_n^l}\right)^{\frac{d-2}{2}}\left[e^{i\left(t_n^l - t_n^j\left(\frac{\lambda_n^j}{\lambda_n^l}\right)^2\right)\Delta}\phi\right]\left(\frac{\lambda_n^j x + x_n^j - x_n^l}{\lambda_n^l}\right).$$

Recalling that (4.6) holds for the pair (j,l), we first prove (4.10) when the scaling parameters are not comparable, that is,

$$\lim_{n\to\infty}\frac{\lambda_n^j}{\lambda_n^l}+\frac{\lambda_n^l}{\lambda_n^j}=\infty. \tag{4.11}$$

By Cauchy–Schwarz,

$$\begin{aligned}\big|\langle I_n, \psi\rangle_{\dot H^1_x}\big| &\lesssim \min\Big\{\|\Delta I_n\|_{L^2_x}\|\psi\|_{L^2_x}, \|I_n\|_{L^2_x}\|\Delta\psi\|_{L^2_x}\Big\}\\ &\lesssim \min\bigg\{\frac{\lambda_n^j}{\lambda_n^l}\|\Delta\phi\|_{L^2_x}\|\psi\|_{L^2_x}, \frac{\lambda_n^l}{\lambda_n^j}\|\phi\|_{L^2_x}\|\Delta\psi\|_{L^2_x}\bigg\},\end{aligned}$$

which converges to zero as $n\to\infty$, for all $\psi\in C_c^\infty(\mathbb{R}^d)$. This establishes (4.10) when (4.11) holds.

Henceforth we may assume

$$\lim_{n\to\infty}\frac{\lambda_n^j}{\lambda_n^l}=\lambda_1\in(0,\infty).$$

We now suppose the time parameters diverge, that is,

$$\lim_{n\to\infty}\frac{|t_n^j(\lambda_n^j)^2-t_n^l(\lambda_n^l)^2|}{\lambda_n^j\lambda_n^l}=\infty;$$

then we also have

$$\left|t_n^l-t_n^j\left(\frac{\lambda_n^j}{\lambda_n^l}\right)^2\right|=\frac{|t_n^l(\lambda_n^l)^2-t_n^j(\lambda_n^j)^2|}{\lambda_n^l\lambda_n^j}\cdot\frac{\lambda_n^j}{\lambda_n^l}\to\infty\quad\text{as}\quad n\to\infty.$$

Under this condition, (4.10) follows from

$$\lambda_1^{\frac{d-2}{2}}\left[e^{i\left(t_n^l-t_n^j\left(\frac{\lambda_n^j}{\lambda_n^l}\right)^2\right)\Delta}\phi\right]\left(\lambda_1 x+\frac{x_n^j-x_n^l}{\lambda_n^l}\right)\rightharpoonup 0\quad\text{weakly in }\dot H^1_x,$$

which is an immediate consequence of Exercise 4.3.

Finally, we deal with the situation when

$$\frac{\lambda_n^j}{\lambda_n^l}\to\lambda_1\in(0,\infty),\qquad\frac{t_n^l(\lambda_n^l)^2-t_n^j(\lambda_n^j)^2}{\lambda_n^j\lambda_n^l}\to t_1,\quad\text{but}\quad\frac{|x_n^j-x_n^l|^2}{\lambda_n^j\lambda_n^l}\to\infty.\tag{4.12}$$

Then we also have $t_n^l-t_n^j(\lambda_n^j)^2/(\lambda_n^l)^2\to\lambda_1 t_1$. Thus, it suffices to show that

$$\lambda_1^{\frac{d-2}{2}}e^{it_1\lambda_1\Delta}\phi(\lambda_1 x+y_n)\rightharpoonup 0\quad\text{weakly in }\dot H^1_x,\tag{4.13}$$

where

$$y_n:=\frac{x_n^j-x_n^l}{\lambda_n^l}=\frac{x_n^j-x_n^l}{\sqrt{\lambda_n^l\lambda_n^j}}\sqrt{\frac{\lambda_n^j}{\lambda_n^l}}\to\infty\quad\text{as}\quad n\to\infty.$$

The desired weak convergence (4.13) follows again from Exercise 4.3.

Finally, we prove the last assertion in the theorem regarding the behaviour of t_n^j. For each j, by passing to a subsequence we may assume $t_n^j \to t^j \in [-\infty, \infty]$. Using a standard diagonal argument, we may assume that the limit exists for all $j \geq 1$.

Fix $j \geq 1$. If $t^j = \pm\infty$, there is nothing more to be proved. If $t^j \in (-\infty, \infty)$, we claim that we may redefine $t_n^j \equiv 0$, provided we replace the original profile ϕ^j by $e^{it^j\Delta}\phi^j$. Indeed, we merely need to prove that the errors introduced by these changes can be incorporated into w_n^J, namely,

$$\lim_{n\to\infty} \big\| g_n^j e^{it_n^j\Delta}\phi^j - g_n^j e^{it^j\Delta}\phi^j \big\|_{\dot H_x^1} = 0.$$

But this follows easily from the strong convergence of the linear propagator.

This completes the proof of Theorem 4.1. □

Exercise 4.1. Under the hypotheses of Proposition 3.2, prove that

$$e^{-it_n^j\Delta}\big[(\lambda_n^j)^{\frac{d-2}{2}} w_n^J(\lambda_n^j x + x_n^j)\big] \rightharpoonup 0 \quad \text{weakly in } \dot H^1(\mathbb{R}^d) \text{ for all } j \leq J.$$

Exercise 4.2. Let $f_n \in \dot H^1(\mathbb{R}^d)$ be such that $f_n \rightharpoonup 0$ weakly in $\dot H^1(\mathbb{R}^d)$ and let $t_n \to t_\infty \in \mathbb{R}$. Then

$$e^{it_n\Delta} f_n \rightharpoonup 0 \quad \text{weakly in } \dot H_x^1 \text{ as } n \to \infty.$$

Exercise 4.3. Let $f \in C_c^\infty(\mathbb{R}^d)$ and let $\{(t_n, x_n)\}_{n\geq 1} \subset \mathbb{R} \times \mathbb{R}^d$. Then

$$e^{it_n\Delta} f(x + x_n) \rightharpoonup 0 \quad \text{weakly in } \dot H_x^1 \text{ as } n \to \infty$$

whenever $|t_n| \to \infty$ or $|x_n| \to \infty$.

Chapter 5

Stability theory for the energy-critical NLS

In this section we develop a stability theory for the energy-critical NLS

$$i\partial_t u = -\Delta u \pm |u|^{\frac{4}{d-2}} u \quad \text{with} \quad u(0) = u_0 \in \dot{H}^1_x. \tag{5.1}$$

Definition 5.1 (Solution)**.** A function $u : I \times \mathbb{R}^d \to \mathbb{C}$ on a non-empty time interval $0 \in I \subset \mathbb{R}$ is a *solution* (more precisely, a strong $\dot{H}^1_x$ solution) to (5.1) if it lies in the class $C^0_t \dot{H}^1_x(K \times \mathbb{R}^d) \cap L^{\frac{2(d+2)}{d-2}}_{t,x}(K \times \mathbb{R}^d)$ for all compact $K \subset I$, and satisfies the Duhamel formula

$$u(t) = e^{it\Delta} u(0) \mp i \int_0^t e^{i(t-s)\Delta} |u(s)|^{\frac{4}{d-2}} u(s)\, ds \tag{5.2}$$

for all $t \in I$. We refer to the interval I as the *lifespan* of u. We say that u is a *maximal-lifespan solution* if the solution cannot be extended to any strictly larger interval. We say that u is a *global solution* if $I = \mathbb{R}$.

Solutions to (5.1) conserve the energy

$$E(u(t)) = \int_{\mathbb{R}^d} \tfrac{1}{2} |\nabla u(t,x)|^2 \pm \tfrac{d-2}{2d} |u(t,x)|^{\frac{2d}{d-2}}\, dx.$$

Note that taking data in $\dot{H}^1_x$ renders the energy finite. Indeed, Sobolev embedding shows that $\dot{H}^1_x$ is precisely the energy space.

The equation is called *energy-critical* because the scaling associated with this equation, namely,

$$u(t,x) \mapsto \lambda^{\frac{d-2}{2}} u\big(\lambda^2 t, \lambda x\big) \quad \text{for} \quad \lambda > 0,$$

leaves invariant not only the class of solutions to (5.1), but also the energy.

Throughout the section, we use S^0 to denote the intersection of any finite number of Strichartz spaces $L^q_t L^r_x$ with (q,r) obeying the conditions of Lemma 2.7,

H. Koch et al., *Dispersive Equations and Nonlinear Waves: Generalized Korteweg–de Vries, Nonlinear Schrödinger, Wave and Schrödinger Maps*, Oberwolfach Seminars 45,
DOI 10.1007/978-3-0348-0736-4_19,

and N^0 to denote the sum of any finite number of dual Strichartz spaces $L_t^{q'} L_x^{r'}$. For an interval $I \subset \mathbb{R}$ we define the norms

$$\|u\|_{S^0(I)} := \|u\|_{S^0(I \times \mathbb{R}^d)} \quad \text{and} \quad \|F\|_{N^0(I)} := \|F\|_{N^0(I \times \mathbb{R}^d)}.$$

We start by reviewing the standard local well-posedness statement for (6.1).

Theorem 5.2 (Standard local well-posedness, [8, 9, 10]). *Let $d \geq 3$ and $u_0 \in H^1(\mathbb{R}^d)$. There exists $\eta_0 = \eta_0(d) > 0$ such that if $0 < \eta \leq \eta_0$ and I is a compact interval containing zero such that*

$$\big\|\nabla e^{it\Delta} u_0\big\|_{L_t^{\frac{2(d+2)}{d-2}} L_x^{\frac{2d(d+2)}{d^2+4}} (I \times \mathbb{R}^d)} \leq \eta, \tag{5.3}$$

then there exists a unique solution u to (5.1) on $I \times \mathbb{R}^d$. Moreover, we have the bounds

$$\begin{aligned}
\big\|\nabla u\big\|_{L_t^{\frac{2(d+2)}{d-2}} L_x^{\frac{2d(d+2)}{d^2+4}} (I \times \mathbb{R}^d)} &\leq 2\eta, \\
\big\|\nabla u\big\|_{S^0(I \times \mathbb{R}^d)} &\lesssim \big\|\nabla u_0\big\|_{L_x^2} + \eta^{1+p}, \\
\|u\|_{S^0(I \times \mathbb{R}^d)} &\lesssim \|u_0\|_{L_x^2}.
\end{aligned}$$

Proof. Exercise! *Hint:* use contraction mapping with the distance given by an S^0 norm. □

Remarks. 1. By the Strichartz inequality,

$$\big\|\nabla e^{it\Delta} u_0\big\|_{L_t^{\frac{2(d+2)}{d-2}} L_x^{\frac{2d(d+2)}{d^2+4}} (I \times \mathbb{R}^d)} \lesssim \big\|\nabla u_0\big\|_{L_x^2}.$$

Thus, (5.3) holds with $I = \mathbb{R}$ for initial data with sufficiently small $\dot{H}_x^1$ norm. In particular, we obtain global well-posedness for initial data in H_x^1 that are small in $\dot{H}_x^1$.

2. By the monotone convergence theorem, given an arbitrary $u_0 \in \dot{H}_x^1$ we can choose a sufficiently small interval I to ensure that (5.3) holds. Note however that the length of I will depend upon u_0 and not merely on its norm.

This standard local well-posedness result suffers from the fact that the initial data belong to the inhomogeneous Sobolev space H_x^1, rather than the energy space $\dot{H}_x^1$; the stronger requirement $u_0 \in H_x^1$ is needed in the proof of Theorem 5.2 in order to prove that the solution map is a contraction. To remove this restriction, we need the following stability result:

Theorem 5.3 (Energy-critical stability result, [22, 34]). *Let I a compact time interval and let $\tilde{u}$ be an approximate solution to (5.1) on $I \times \mathbb{R}^d$ in the sense that*

$$i\tilde{u}_t = -\Delta\tilde{u} \pm |\tilde{u}|^{\frac{4}{d-2}}\tilde{u} + e$$

for some function e. Assume that

$$\|\tilde{u}\|_{L_t^\infty \dot{H}_x^1(I\times\mathbb{R}^d)} \leq E, \tag{5.4}$$

$$\|\tilde{u}\|_{L_{t,x}^{\frac{2(d+2)}{d-2}}(I\times\mathbb{R}^d)} \leq L, \tag{5.5}$$

for some positive constants E and L. Let $t_0 \in I$ and $u_0 \in \dot{H}_x^1$ and assume the smallness conditions

$$\|u_0 - \tilde{u}(t_0)\|_{\dot{H}_x^1} \leq \varepsilon \tag{5.6}$$

$$\|\nabla e\|_{N^0(I)} \leq \varepsilon \tag{5.7}$$

for some $0 < \varepsilon < \varepsilon_1 = \varepsilon_1(E, L)$. Then there exists a unique strong solution $u : I \times \mathbb{R}^d \to \mathbb{C}$ to (5.1) with initial datum u_0 at time $t = t_0$ satisfying

$$\|u - \tilde{u}\|_{L_{t,x}^{\frac{2(d+2)}{d-2}}(I\times\mathbb{R}^d)} \leq C(E, L)\varepsilon^c, \tag{5.8}$$

$$\|\nabla(u - \tilde{u})\|_{S^0(I)} \leq C(E, L), \tag{5.9}$$

$$\|\nabla u\|_{S^0(I)} \leq C(E, L), \tag{5.10}$$

where $c = c(d) > 0$.

This stability result was first proved for $d = 3$ in the work of Colliander, Keel, Staffilani, Takaoka, and Tao [13] on the defocusing energy-critical NLS. For $d = 4$, it can be found in [30]. The same proof extends easily to dimensions $d = 5, 6$. To prove Theorem 5.3 in dimensions $d \geq 7$, new ideas are needed. To see why, let us consider the question of continuous dependence of the solution upon the initial data, which corresponds to taking $e = 0$ in Theorem 5.3. To make things as simple as possible, we choose initial data $u_0, \tilde{u}_0 \in H_x^1$ which are small in the sense that

$$\|u_0\|_{\dot{H}_x^1} + \|\tilde{u}_0\|_{\dot{H}_x^1} \leq \eta_0.$$

By the first remark above, if η_0 is sufficiently small there exist unique global solutions u and $\tilde{u}$ to (5.1) with initial data u_0 and $\tilde{u}_0$, respectively; moreover, they satisfy

$$\|\nabla u\|_{S^0(\mathbb{R})} + \|\nabla \tilde{u}\|_{S^0(\mathbb{R})} \lesssim \eta_0.$$

We would like to see that if u_0 and $\tilde{u}_0$ are *close* in $\dot{H}_x^1$, say $\|\nabla(u_0 - \tilde{u}_0)\|_2 \leq \varepsilon \ll \eta_0$, then u and $\tilde{u}$ remain ε-close in *energy-critical* norms. An application of the Strichartz inequality combined with the bounds above yields

$$\|\nabla(u-\tilde{u})\|_{S^0(\mathbb{R})} \lesssim \|\nabla(u_0 - \tilde{u}_0)\|_{L_x^2} + \eta_0^{\frac{4}{d-2}} \|\nabla(u-\tilde{u})\|_{S^0(\mathbb{R})} + \eta_0 \|\nabla(u-\tilde{u})\|_{S^0(\mathbb{R})}^{\frac{4}{d-2}}.$$

If $4/(d-2) \geq 1$, a simple bootstrap argument implies continuous dependence of the solution on the initial data. However, this will not work if $4/(d-2) < 1$, that is, if $d \geq 7$. The last term in the inequality above causes the bootstrap argument to break down in high dimensions; indeed, tiny numbers become much larger when raised to a fractional power.

To prove Theorem 5.3 in dimensions $d \geq 7$, the authors of [34] work in spaces with fractional derivatives (rather than a full derivative), while still maintaining criticality with respect to the scaling. A similar technique was employed by Nakanishi [27] for the energy-critical Klein–Gordon equation in high dimensions.

The result in [34] assumes the less stringent smallness condition

$$\Big(\sum_{N \in 2^{\mathbb{Z}}} \big\| \nabla P_N e^{i(t-t_0)\Delta} \big(u_0 - \tilde{u}(t_0)\big) \big\|^2_{L_t^{\frac{2(d+2)}{d-2}} L_x^{\frac{2d(d+2)}{d^2+4}} (I \times \mathbb{R}^d)} \Big)^{1/2} \leq \varepsilon$$

in place of (5.6). There is also an improvement over the result in [34], in which the smallness condition above is replaced by

$$\big\| e^{i(t-t_0)\Delta} \big(u_0 - \tilde{u}(t_0)\big) \big\|_{L_{t,x}^{\frac{2(d+2)}{d-2}} (I \times \mathbb{R}^d)} \leq \varepsilon.$$

To prove Theorem 5.3 with this particular hypothesis (which was helpful in early treatments of the energy-critical NLS), it becomes necessary to work in spaces with fractional derivatives even in small dimensions; see [22] for the proof.

In what follows, we will present the proof of Theorem 5.3 in dimensions $3 \leq d \leq 6$. For higher dimensions, see [22, 34].

Proof of Theorem 5.3 *for* $3 \leq d \leq 6$. We will prove the result under the additional assumption that $u_0 \in L_x^2$ (and so $u_0 \in H_x^1$). This allows us to invoke Theorem 5.2 and so guarantee that u exists. Thus, it suffices to prove (5.8) through (5.10) as *a priori* estimates, that is, we assume that u exists and satisfies $\nabla u \in S^0(I)$. Once we have proved (5.8) through (5.10), we may remove the additional assumption $u_0 \in L_x^2$ by the usual limiting argument: Approximate $u_0 \in \dot{H}_x^1$ by $\{f_n\}_{n \geq 1} \subset H_x^1$ and let u_n be the solution to (5.1) with initial data $u_n(t_0) = f_n$. Applying Theorem 5.3 with $\tilde{u} := u_m$, $u := u_n$, and $e = 0$, we deduce that the sequence of solutions $\{u_n\}_{n \geq 1}$ is Cauchy in energy-critical norms. Therefore, u_n converges to a solution u with data $u(t_0) = u_0$ which satisfies $\nabla u \in S^0(I)$.

We first prove the theorem under the hypothesis

$$\|\nabla \tilde{u}\|_{L_t^{\frac{2(d+2)}{d-2}} L_x^{\frac{2d(d+2)}{d^2+4}} (I \times \mathbb{R}^d)} \leq \delta \tag{5.11}$$

for some $\delta > 0$ sufficiently small depending on E. Without loss of generality, we may assume $t_0 = \inf I$.

To continue, let $v := u - \tilde{u}$ and for $t \in I$ define

$$A(t) := \big\| \nabla \big[(i\partial_t + \Delta) v + e \big] \big\|_{L_t^2 L_x^{\frac{2d}{d+2}} ([t_0, t] \times \mathbb{R}^d)}.$$

By Sobolev embedding, Strichartz, (5.6), and (5.7), we get

$$\begin{aligned}\|v\|_{L_{t,x}^{\frac{2(d+2)}{d-2}}([t_0,t]\times\mathbb{R}^d)} &\lesssim \|\nabla v\|_{L_t^{\frac{2(d+2)}{d-2}} L_x^{\frac{2d(d+2)}{d^2+4}}([t_0,t]\times\mathbb{R}^d)}\\ &\lesssim \|v(t_0)\|_{\dot H_x^1} + A(t) + \|\nabla e\|_{L_t^2 L_x^{\frac{2d}{d+2}}([t_0,t]\times\mathbb{R}^d)}\\ &\lesssim A(t)+\varepsilon. \end{aligned} \tag{5.12}$$

On the other hand, by Hölder, (5.11), (5.12), and Sobolev embedding, we get

$$\begin{aligned} A(t) \lesssim\ & \|\nabla\tilde u\|_{L_t^{\frac{2(d+2)}{d-2}} L_x^{\frac{2d(d+2)}{d^2+4}}} \|v\|_{L_{t,x}^{\frac{2(d+2)}{d-2}}} \Big[\|v\|_{L_{t,x}^{\frac{2(d+2)}{d-2}}} + \|\tilde u\|_{L_{t,x}^{\frac{2(d+2)}{d-2}}}\Big]^{\frac{6-d}{d-2}}\\ &+ \|\nabla v\|_{L_t^{\frac{2(d+2)}{d-2}} L_x^{\frac{2d(d+2)}{d^2+4}}} \Big[\|v\|_{L_{t,x}^{\frac{2(d+2)}{d-2}}} + \|\tilde u\|_{L_{t,x}^{\frac{2(d+2)}{d-2}}}\Big]^{\frac{4}{d-2}}\\ \lesssim\ & \delta[A(t)+\varepsilon][A(t)+\varepsilon+\delta]^{\frac{6-d}{d-2}} + [A(t)+\varepsilon][A(t)+\varepsilon+\delta]^{\frac{4}{d-2}}, \end{aligned}$$

where all space-time norms are over $[t_0,t]\times\mathbb{R}^d$.

Taking δ,ε sufficiently small (depending only on the ambient dimension so far), a standard continuity argument gives

$$A(t)\lesssim\varepsilon \quad \text{for all} \quad t\in I, \tag{5.13}$$

with $c=c(d)=1$. Together with (5.12), this gives (5.8). To obtain (5.9), we use the Strichartz inequality, (5.6), (5.7), and (5.13), as follows:

$$\begin{aligned} \|\nabla(u-\tilde u)\|_{S^0(I)} \lesssim\ & \|u_0-\tilde u(t_0)\|_{\dot H_x^1} + \big\|\nabla\big[(i\partial_t+\Delta)v+e\big]\big\|_{L_t^2 L_x^{\frac{2d}{d+2}}(I\times\mathbb{R}^d)}\\ &+ \|\nabla e\|_{L_t^2 L_x^{\frac{2d}{d+2}}(I\times\mathbb{R}^d)}\\ \lesssim\ & \varepsilon. \end{aligned}$$

To obtain (5.10), we first note that, by (5.11) and (5.12),

$$\begin{aligned} &\|\nabla u\|_{L_t^{\frac{2(d+2)}{d-2}} L_x^{\frac{2d(d+2)}{d^2+4}}(I\times\mathbb{R}^d)}\\ &\le \|\nabla v\|_{L_t^{\frac{2(d+2)}{d-2}} L_x^{\frac{2d(d+2)}{d^2+4}}(I\times\mathbb{R}^d)} + \|\nabla\tilde u\|_{L_t^{\frac{2(d+2)}{d-2}} L_x^{\frac{2d(d+2)}{d^2+4}}(I\times\mathbb{R}^d)}\\ &\lesssim \varepsilon+\delta. \end{aligned}$$

Using this together with the Strichartz inequality, Sobolev embedding, and (5.4),

$$\begin{aligned} \|\nabla u\|_{S^0(I)} &\lesssim \|\tilde u(t_0)\|_{\dot H_x^1} + \|u_0-\tilde u(t_0)\|_{\dot H_x^1} + \|\nabla u\|^{\frac{d+2}{d-2}}_{L_t^{\frac{2(d+2)}{d-2}} L_x^{\frac{2d(d+2)}{d^2+4}}(I\times\mathbb{R}^d)}\\ &\lesssim E+\varepsilon+[\varepsilon+\delta]^{\frac{d+2}{d-2}} \lesssim E, \end{aligned}$$

provided $\delta, \varepsilon \le \varepsilon_0 = \varepsilon_0(E)$.

To complete the proof of Theorem 5.3 in small dimensions, it remains to restore the hypothesis (5.5) in place of (5.11). We first note that (5.5) implies $\nabla u \in S^0(I)$. Indeed, subdividing I into $N_0 \sim (1+\frac{L}{\eta})^{\frac{2(d+2)}{d-2}}$ subintervals J_k such that on each J_k we have

$$\|\tilde{u}\|_{L_{t,x}^{\frac{2(d+2)}{d-2}}(J_k\times\mathbb{R}^d)} \le \eta,$$

and using the Strichartz inequality, Sobolev embedding, and (5.4), we estimate

$$\begin{aligned}\|\nabla\tilde{u}\|_{S^0(J_k)} &\lesssim \|\tilde{u}\|_{L_t^\infty \dot{H}_x^1(I\times\mathbb{R}^d)} + \|\nabla\tilde{u}\|_{S^0(J_k)}\|\tilde{u}\|^{\frac{4}{d-2}}_{L_{t,x}^{\frac{2(d+2)}{d-2}}(J_k\times\mathbb{R}^d)} + \|\nabla e\|_{L_t^2L_x^{\frac{2d}{d+2}}(I\times\mathbb{R}^d)}\\ &\lesssim E + \eta^{\frac{4}{d-2}}\|\nabla\tilde{u}\|_{S^0(J_k)} + \varepsilon.\end{aligned}$$

Thus for η sufficiently small depending on d,

$$\|\nabla\tilde{u}\|_{S^0(J_k)} \lesssim E + \varepsilon.$$

Summing these bounds over all the intervals J_k we obtain

$$\|\nabla\tilde{u}\|_{S^0(I)} \le C(E, L).$$

We can now subdivide I into $N_1 = N_1(E, L)$ subintervals $I_j = [t_j, t_{j+1}]$ such that on each I_j we have

$$\|\nabla\tilde{u}\|_{L_t^{\frac{2(d+2)}{d-2}} L_x^{\frac{2d(d+2)}{d^2+4}}(I_j\times\mathbb{R}^d)} \le \delta,$$

where δ is as in (5.11). Choosing ε_1 sufficiently small depending on ε_0 and N_1, the argument above implies that for each j and all $0 < \varepsilon < \varepsilon_1$,

$$\begin{aligned}\|u-\tilde{u}\|_{L_{t,x}^{\frac{2(d+2)}{d-2}}(I_j\times\mathbb{R}^d)} &\le C(j)\varepsilon,\\ \|\nabla(u-\tilde{u})\|_{S^0(I_j)} &\le C(j)\varepsilon,\\ \|\nabla u\|_{S^0(I_j)} &\le C(j)E,\\ \big\|\nabla\big[(i\partial_t+\Delta)(u-\tilde{u})+e\big]\big\|_{L_t^2L_x^{\frac{2d}{d+2}}(I_j\times\mathbb{R}^d)} &\le C(j)\varepsilon,\end{aligned}$$

provided we can show that (5.6) holds when t_0 is replaced by t_j. We check this using an inductive argument. By the Strichartz inequality,

$$\begin{aligned}\|u(t_{j+1})-\tilde{u}(t_{j+1})\|_{\dot{H}_x^1} &\lesssim \|u_0-\tilde{u}(t_0)\|_{\dot{H}_x^1} + \|\nabla e\|_{L_t^2L_x^{\frac{2d}{d+2}}(I\times\mathbb{R}^d)}\\ &\qquad + \big\|\nabla\big[(i\partial_t+\Delta)(u-\tilde{u})+e\big]\big\|_{L_t^2L_x^{\frac{2d}{d+2}}([t_0,t_{j+1}]\times\mathbb{R}^d)}\\ &\lesssim \varepsilon + \sum_{k=0}^{j} C(k)\varepsilon.\end{aligned}$$

Choosing ε_1 sufficiently small depending on ε_0 and E, we can continue the inductive argument.

This completes the proof of Theorem 5.3 in dimensions $3 \leq d \leq 6$. $\square$

Chapter 6

A large data critical problem

Throughout the remainder of these notes we restrict attention to the defocusing energy-critical NLS

$$i\partial_t u + \Delta u = |u|^{\frac{4}{d-2}} u \quad \text{with} \quad u(0) = u_0 \in \dot{H}^1_x. \tag{6.1}$$

For arguments and further references in the focusing case, see [22]. For equation (6.1) we have the following large data global result:

Theorem 6.1 (Global well-posedness and scattering)**.** *Let $d \geq 3$ and $u_0 \in \dot{H}^1_x$. Then there exists a unique global solution u to (6.1) and it satisfies*

$$\int_{\mathbb{R}} \int_{\mathbb{R}^d} |u(t,x)|^{\frac{2(d+2)}{d-2}} \, dx \, dt \leq C(E(u_0)).$$

In particular, the solution scatters, that is, there exist asymptotic states $u_\pm \in \dot{H}^1_x$ such that

$$\|u(t) - e^{it\Delta} u_\pm\|_{\dot{H}^1_x} \to 0 \quad \text{as} \quad t \to \pm\infty.$$

The proof of this theorem sparked the recent progress on dispersive equations at the critical regularity. It was first proved for spherically symmetric initial data in dimensions $d = 3, 4$ by Bourgain [5]. In this work, Bourgain introduced the induction on energy paradigm as a means for breaking the scaling symmetry; this allowed him to use non-critical monotonicity formulas like the Morawetz inequality (which scales like $\dot{H}^{1/2}_x$). Building on Bourgain's argument, Tao [33] proved the theorem in dimensions $d \geq 5$ for spherically symmetric data.

The radial assumption was removed in dimension $d = 3$ by Colliander, Keel, Staffilani, Takaoka, and Tao [13]. This work further advanced the induction on energy argument, introducing important new ideas that informed subsequent developments. To deal with arbitrary data, the authors employed a frequency-localized interaction Morawetz inequality, which is even further away from scaling (it scales like $\dot{H}^{1/4}_x$). The work [13] was extended to four dimensions in [30]. Finally, for dimensions $d \geq 5$, Theorem 6.1 was proved in [39]; for a different proof reflecting new advances see [23], which also treats the focusing problem.

H. Koch et al., *Dispersive Equations and Nonlinear Waves: Generalized Korteweg–de Vries, Nonlinear Schrödinger, Wave and Schrödinger Maps*, Oberwolfach Seminars 45,
DOI 10.1007/978-3-0348-0736-4_20,

In these notes, we will present the proof of Theorem 6.1 in dimension $d = 4$. The proof below is taken from [40], which revisits the argument in [30] in light of the recent advances made by Dodson [15] on the mass-critical NLS. For a proof of the three-dimensional case treated in [13] that also incorporates these advances see [21].

We note that parts of the argument we will present in these notes work in all dimensions $d \geq 3$; in particular, we will demonstrate the existence of a minimal counterexample to Theorem 6.1 in all dimensions $d \geq 3$.

To start, for any $0 \leq E < \infty$, we define

$$L(E) := \sup\{S_I(u) : \, u : I \times \mathbb{R}^d \to \mathbb{C} \text{ such that } E(u) \leq E\},$$

where the supremum is taken over all solutions $u : I \times \mathbb{R}^d \to \mathbb{C}$ to (6.1). Here, we use the notation

$$S_I(u) := \int_I \int_{\mathbb{R}^d} |u(t,x)|^{\frac{2(d+2)}{d-2}} \, dx \, dt$$

for the scattering size of u on an interval I.

Note that $L : [0, \infty) \to [0, \infty]$ is a non-decreasing function. Moreover, from the small data theory,

$$L(E) \lesssim E^{\frac{d+2}{d-2}} \quad \text{for} \quad E \leq \eta_0,$$

where $\eta_0 = \eta_0(d)$ is the small data threshold.

Exercise 6.1. Prove that the set $\{E > 0 : \, L(E) < \infty\}$ is open.
Hint: Use Theorem 5.3.

Therefore, there must exist a unique *critical energy* $0 < E_c \leq \infty$ such that

$$L(E) < \infty \quad \text{for} \quad E < E_c \quad \text{and} \quad L(E) = \infty \quad \text{for} \quad E \geq E_c.$$

This plays the role of the inductive hypothesis because it says that Theorem 6.1 holds for energies $E < E_c$. The argument is called induction on energy, because this inductive hypothesis will be used to prove that $L(E_c) < \infty$, thus providing the desired contradiction.

Chapter 7

A Palais–Smale type condition

In this section we prove a Palais–Smale condition for minimizing sequences of blowup solutions to the defocusing energy-critical NLS. It was already observed in [5, 13] that such minimizing sequences have good tightness and equicontinuity properties. This was taken to its ultimate conclusion by Keraani [19], who showed the existence and almost periodicity of minimal blowup solutions in the context of the mass-critical NLS. The proof of the Palais–Smale condition is the crux of this argument.

We first define operators T_n^j on general functions of space-time. These act on linear solutions in a manner corresponding to the action of $g_n^j e^{it_n^j \Delta}$ on initial data:

$$(T_n^j u)(t,x) := (\lambda_n^j)^{-\frac{d-2}{2}} u\left(\frac{t}{(\lambda_n^j)^2} + t_n^j, \frac{x - x_n^j}{\lambda_n^j}\right).$$

Here, the parameters $\lambda_n^j, t_n^j, x_n^j$ are as defined in Theorem 4.1. Using the asymptotic orthogonality condition (4.6), it is not hard to prove the following

Lemma 7.1 (Asymptotic decoupling)**.** *Suppose that the parameters associated to j, k are orthogonal in the sense of (4.6). Then for any $\psi^j, \psi^k \in C_c^\infty(\mathbb{R} \times \mathbb{R}^d)$,*

$$\|T_n^j \psi^j T_n^k \psi^k\|_{L_{t,x}^{\frac{d+2}{d-2}}} + \|T_n^j \psi^j \nabla(T_n^k \psi^k)\|_{L_{t,x}^{\frac{d+2}{d-1}}} + \|\nabla(T_n^j \psi^j)\nabla(T_n^k \psi^k)\|_{L_{t,x}^{\frac{d+2}{d}}}$$

converges to zero as $n \to \infty$.

Proof. Using a change of variables, we get

$$\begin{aligned} &\|T_n^j \psi^j T_n^k \psi^k\|_{L_{t,x}^{\frac{d+2}{d-2}}} + \|T_n^j \psi^j \nabla(T_n^k \psi^k)\|_{L_{t,x}^{\frac{d+2}{d-1}}} + \|\nabla(T_n^j \psi^j)\nabla(T_n^k \psi^k)\|_{L_{t,x}^{\frac{d+2}{d}}} \\ &\quad = \|\psi^j (T_n^j)^{-1} T_n^k \psi^k\|_{L_{t,x}^{\frac{d+2}{d-2}}} + \|\psi^j \nabla (T_n^j)^{-1} T_n^k \psi^k\|_{L_{t,x}^{\frac{d+2}{d-1}}} \\ &\qquad + \|\nabla \psi^j \nabla (T_n^j)^{-1} T_n^k \psi^k\|_{L_{t,x}^{\frac{d+2}{d}}}, \end{aligned}$$

H. Koch et al., *Dispersive Equations and Nonlinear Waves: Generalized Korteweg–de Vries, Nonlinear Schrödinger, Wave and Schrödinger Maps*, Oberwolfach Seminars 45,
DOI 10.1007/978-3-0348-0736-4_21, © Springer Basel 2014

where all space-time norms are over $\mathbb{R}\times\mathbb{R}^d$. Note that

$$[(T_n^j)^{-1}T_n^k\psi^k](t,x) = \big(\tfrac{\lambda_n^j}{\lambda_n^k}\big)^{\frac{d-2}{2}}\psi^k\Big(\big(\tfrac{\lambda_n^j}{\lambda_n^k}\big)^2\big(t-\tfrac{t_n^j(\lambda_n^j)^2-t_n^k(\lambda_n^k)^2}{(\lambda_n^j)^2}\big), \tfrac{\lambda_n^j}{\lambda_n^k}\big(x-\tfrac{x_n^k-x_n^j}{\lambda_n^j}\big)\Big).$$

We will only present the details for decoupling in the $L_{t,x}^{\frac{d+2}{d-2}}$ norm; the argument for decoupling in the other norms is very similar.

We first assume that $\frac{\lambda_n^j}{\lambda_n^k}+\frac{\lambda_n^k}{\lambda_n^j}\to\infty$. Using Hölder's inequality and a change of variables, we estimate

$$\begin{aligned}
&\|\psi^j(T_n^j)^{-1}T_n^k\psi^k\|_{L_{t,x}^{\frac{d+2}{d-2}}}\\
&\qquad\le \min\Big\{\|\psi^j\|_{L^\infty_{t,x}}\|(T_n^j)^{-1}T_n^k\psi^k\|_{L_{t,x}^{\frac{d+2}{d-2}}}+\|\psi^j\|_{L_{t,x}^{\frac{d+2}{d-2}}}\|(T_n^j)^{-1}T_n^k\psi^k\|_{L^\infty_{t,x}}\Big\}\\
&\qquad\lesssim \min\Big\{\big(\tfrac{\lambda_n^j}{\lambda_n^k}\big)^{-\frac{d-2}{2}}, \big(\tfrac{\lambda_n^j}{\lambda_n^k}\big)^{\frac{d-2}{2}}\Big\}\to 0 \quad\text{as}\quad n\to\infty.
\end{aligned}$$

Henceforth, we may assume $\frac{\lambda_n^j}{\lambda_n^k}\to\lambda_0\in(0,\infty)$.

If $\frac{|t_n^j(\lambda_n^j)^2-t_n^k(\lambda_n^k)^2|}{\lambda_n^k\lambda_n^j}\to\infty$, it is easy to see that the temporal supports of ψ^j and $(T_n^j)^{-1}T_n^k\psi^k$ become disjoint for n sufficiently large. Hence

$$\lim_{n\to\infty}\|\psi^j(T_n^j)^{-1}T_n^k\psi^k\|_{L_{t,x}^{\frac{d+2}{d-2}}}=0.$$

If instead

$$\frac{\lambda_n^j}{\lambda_n^k}\to\lambda_0, \quad \frac{t_n^j(\lambda_n^j)^2-t_n^k(\lambda_n^k)^2}{\lambda_n^k\lambda_n^j}\to t_0, \quad\text{and}\quad \frac{|x_n^j-x_n^k|}{\sqrt{\lambda_n^j\lambda_n^k}}\to\infty,$$

then the spatial supports of ψ^j and $(T_n^j)^{-1}T_n^k\psi^k$ become disjoint for n sufficiently large. Indeed, in this case we have

$$\frac{|x_n^j-x_n^k|}{\lambda_n^j}=\frac{|x_n^j-x_n^k|}{\sqrt{\lambda_n^j\lambda_n^k}}\sqrt{\frac{\lambda_n^k}{\lambda_n^j}}\to\infty \quad\text{as}\quad n\to\infty.$$

This completes the proof of the lemma. □

Recall that failure of Theorem 6.1 implies the existence of a critical energy $0<E_c<\infty$ so that

$$L(E)<\infty \quad\text{for}\quad E<E_c \quad\text{and}\quad L(E)=\infty \quad\text{for}\quad E\ge E_c, \tag{7.1}$$

where $L(E)$ denotes the supremum of $S_I(u)$ over all solutions $u: I\times\mathbb{R}^d\to\mathbb{C}$ with $E(u)\le E$.

The positivity of E_c is a consequence of the small data global well-posedness. Indeed, the argument proves the stronger statement

$$\|u\|_{\dot X^1(\mathbb{R}\times\mathbb{R}^d)}\lesssim E(u_0)^{\frac12} \quad\text{for all data with } E(u_0)\le\eta_0, \tag{7.2}$$

where η_0 denotes the small data threshold. Here,

$$\dot{X}^1 := L_{t,x}^{\frac{2(d+2)}{d-2}} \cap L_t^{\frac{2(d+2)}{d}} \dot{H}_x^{1,\frac{2(d+2)}{d}}.$$

Using the induction on energy argument together with (7.1) and the stability result Theorem 5.3, we now prove a compactness result for optimizing sequences of blowup solutions.

Proposition 7.2 (Palais–Smale condition)**.** *Let $u_n : I_n \times \mathbb{R}^d \to \mathbb{C}$ be a sequence of solutions to the defocusing energy-critical NLS with $E(u_n) \to E_c$, for which there is a sequence of times $t_n \in I_n$ so that*

$$\lim_{n\to\infty} S_{\geq t_n}(u_n) = \lim_{n\to\infty} S_{\leq t_n}(u_n) = \infty.$$

Then the sequence $u_n(t_n)$ has a subsequence that converges in $\dot{H}_x^1$ modulo scaling and spatial translations.

Proof. Using time translation symmetry, we may take $t_n \equiv 0$ for all n; thus,

$$\lim_{n\to\infty} S_{\geq 0}(u_n) = \lim_{n\to\infty} S_{\leq 0}(u_n) = \infty. \tag{7.3}$$

Applying Theorem 4.1 to the bounded sequence $\{u_n(0)\}_{n\geq 1} \subset \dot{H}_x^1$ and passing to a subsequence if necessary, we decompose

$$u_n(0) = \sum_{j=1}^{J} g_n^j e^{it_n^j\Delta}\phi^j + w_n^J, \tag{7.4}$$

with the properties stated in that theorem. In particular, for any finite $0 \leq J \leq J^*$ we have the energy decoupling condition

$$\lim_{n\to\infty} \Big\{ E(u_n) - \sum_{j=1}^{J} E(e^{it_n^j\Delta}\phi^j) - E(w_n^J) \Big\} = 0. \tag{7.5}$$

To prove the proposition, we need to show that $J^* = 1$, that $w_n^1 \to 0$ in $\dot{H}_x^1$, and that $t_n^1 \equiv 0$. All other possibilities will be shown to contradict (7.3). We discuss two scenarios:

Scenario I: $\sup_j \limsup_{n\to\infty} E(e^{it_n^j\Delta}\phi^j) = E_c$.

From the non-triviality of the profiles, we have $\liminf_{n\to\infty} E(e^{it_n^j\Delta}\phi^j) > 0$ for every finite $1 \leq j \leq J^*$. Thus, using (7.5) together with the hypothesis $E(u_n) \to E_c$ (and passing to a subsequence if necessary), we deduce that there is a single profile in the decomposition (7.4) (that is, $J^* = 1$) and we can write

$$u_n(0) = g_n e^{it_n\Delta}\phi + w_n \quad \text{with} \quad \lim_{n\to\infty} \|w_n\|_{\dot{H}_x^1} = 0 \tag{7.6}$$

and $t_n \equiv 0$ or $t_n \to \pm\infty$. If $t_n \equiv 0$, then we obtain the desired compactness. Thus, we only need to preclude the scenario when $t_n \to \pm\infty$.

Let us suppose $t_n \to \infty$; the case $t_n \to -\infty$ can be treated symmetrically. In this case, the Strichartz inequality and the monotone convergence theorem yield

$$S_{\geq 0}(e^{it\Delta} u_n(0)) \lesssim S_{\geq t_n}(e^{it\Delta}\phi) + S(e^{it\Delta} w_n) \to 0 \quad \text{as} \quad n \to \infty.$$

By Theorem 5.3, this implies that $S_{\geq 0}(u_n) \to 0$, which contradicts (7.3).

Scenario 2: $\sup_j \limsup_{n\to\infty} E(e^{it_n^j\Delta}\phi^j) \leq E_c - 2\delta$ for some $\delta > 0$.

We first observe that in this case, for each finite $J \leq J^*$ we have $E(e^{it_n^j\Delta}\phi^j) \leq E_c - \delta$ for all $1 \leq j \leq J$ and n sufficiently large.

Next we define nonlinear profiles corresponding to each bubble in the decomposition of $u_n(0)$. If $t_n^j \equiv 0$, we define $v^j : I^j \times \mathbb{R}^d \to \mathbb{C}$ to be the maximal-lifespan solution to the defocusing energy-critical NLS with initial data $v^j(0) = \phi^j$. If instead $t_n^j \to \pm\infty$, we define $v^j : I^j \times \mathbb{R}^d \to \mathbb{C}$ to be the maximal-lifespan solution to the defocusing energy-critical NLS which scatters to $e^{it\Delta}\phi^j$ as $t \to \pm\infty$. Now define $v_n^j := T_n^j v^j$. Then v_n^j is also a solution to the defocusing energy-critical NLS on the time interval $I_n^j := (\lambda_n^j)^2(I^j - \{t_n^j\})$. In particular, for n sufficiently large we have $0 \in I_n^j$ and

$$\lim_{n\to\infty} \|v_n^j(0) - g_n^j e^{it_n^j\Delta}\phi^j\|_{\dot{H}_x^1} = 0. \tag{7.7}$$

Combining this with $E(e^{it_n^j\Delta}\phi^j) \leq E_c - \delta < E_c$ and the inductive hypothesis (7.1), we deduce that for n sufficiently large, v_n^j (and so also v^j) are global solutions that satisfy

$$S_{\mathbb{R}}(v^j) = S_{\mathbb{R}}(v_n^j) \leq L(E_c - \delta) < \infty.$$

(Note in particular that this implies v_n^j are global for *all* $n \geq 1$ and they admit a common space-time bound.)

Combining this with the Strichartz inequality shows that all Strichartz norms of v^j and v_n^j are finite; in particular,

$$\|v^j\|_{\dot{X}^1(\mathbb{R}\times\mathbb{R}^d)} = \|v_n^j\|_{\dot{X}^1(\mathbb{R}\times\mathbb{R}^d)} \leq_{E_c,\delta} 1.$$

This allows us to approximate v_n^j in $\dot{X}^1(\mathbb{R} \times \mathbb{R}^d)$ by $C_c^\infty(\mathbb{R} \times \mathbb{R}^d)$ functions. More precisely, for any $\varepsilon > 0$ there exist $\psi_\varepsilon^j \in C_c^\infty(\mathbb{R} \times \mathbb{R}^d)$ so that

$$\|v_n^j - T_n^j \psi_\varepsilon^j\|_{\dot{X}^1(\mathbb{R}\times\mathbb{R}^d)} < \varepsilon. \tag{7.8}$$

Moreover, we may use (7.2) together with our bounds on the space-time norms of v_n^j and the finiteness of E_c to deduce that

$$\|v_n^j\|_{\dot{X}^1(\mathbb{R}\times\mathbb{R}^d)} \lesssim_{E_c,\delta} E(e^{it_n^j\Delta}\phi^j)^{\frac{1}{2}} \lesssim_{E_c,\delta} 1. \tag{7.9}$$

Combining this with (7.5) we deduce that

$$\limsup_{n\to\infty} \sum_{j=1}^{J} \|v_n^j\|^2_{\dot X^1(\mathbb{R}\times\mathbb{R}^d)} \lesssim_{E_c,\delta} \limsup_{n\to\infty} \sum_{j=1}^{J} E(e^{it_n^j\Delta}\phi_n^j) \lesssim_{E_c,\delta} 1, \tag{7.10}$$

uniformly for finite $J \le J^*$.

The asymptotic orthogonality condition (4.6) gives rise to asymptotic decoupling of the nonlinear profiles.

Lemma 7.3 (Decoupling of nonlinear profiles)**.** *For $j \neq k$ we have*

$$\lim_{n\to\infty} \|v_n^j v_n^k\|_{L_{t,x}^{\frac{d+2}{d-2}}(\mathbb{R}\times\mathbb{R}^d)} + \|v_n^j \nabla v_n^k\|_{L_{t,x}^{\frac{d+2}{d-1}}(\mathbb{R}\times\mathbb{R}^d)} + \|\nabla v_n^j \nabla v_n^k\|_{L_{t,x}^{\frac{d+2}{d}}(\mathbb{R}\times\mathbb{R}^d)} = 0.$$

Proof. We only present the argument for decoupling in the $L_{t,x}^{\frac{d+2}{d-2}}$ norm; the argument for decoupling in the other norms is similar. Recall that for any $\varepsilon > 0$ there exist $\psi_\varepsilon^j, \psi_\varepsilon^k \in C_c^\infty(\mathbb{R}\times\mathbb{R}^d)$ so that

$$\|v_n^j - T_n^j\psi_\varepsilon^j\|_{\dot X^1(\mathbb{R}\times\mathbb{R}^d)} + \|v_n^k - T_n^k\psi_\varepsilon^k\|_{\dot X^1(\mathbb{R}\times\mathbb{R}^d)} < \varepsilon.$$

Thus, using (7.9) and Lemma 7.1 we get

$$\begin{aligned}
&\|v_n^j v_n^k\|_{L_{t,x}^{\frac{d+2}{d-2}}} \\
&\quad \le \|v_n^j(v_n^k - T_n^k\psi_\varepsilon^k)\|_{L_{t,x}^{\frac{d+2}{d-2}}} + \|(v_n^j - T_n^j\psi_\varepsilon^j)T_n^k\psi_\varepsilon^k\|_{L_{t,x}^{\frac{d+2}{d-2}}} + \|T_n^j\psi_\varepsilon^j\, T_n^k\psi_\varepsilon^k\|_{L_{t,x}^{\frac{d+2}{d-2}}} \\
&\quad \lesssim \|v_n^j\|_{\dot X^1}\|v_n^k - T_n^k\psi_\varepsilon^k\|_{\dot X^1} + \|v_n^j - T_n^j\psi_\varepsilon^j\|_{\dot X^1}\|\psi_\varepsilon^k\|_{\dot X^1} + \|T_n^j\psi_\varepsilon^j\, T_n^k\psi_\varepsilon^k\|_{L_{t,x}^{\frac{d+2}{d-2}}} \\
&\quad \lesssim_{E_c,\delta} \varepsilon + o(1) \quad \text{as} \quad n\to\infty.
\end{aligned}$$

As $\varepsilon > 0$ was arbitrary, this proves the asymptotic decoupling statement. □

As a consequence of this decoupling we can bound the sum of the nonlinear profiles in $\dot X^1$, as follows:

$$\limsup_{n\to\infty} \Big\| \sum_{j=1}^{J} v_n^j \Big\|_{\dot X^1(\mathbb{R}\times\mathbb{R}^d)} \lesssim_{E_c,\delta} 1 \quad \text{uniformly for finite } J \le J^*. \tag{7.11}$$

Indeed, by Young's inequality, (7.9), (7.10), and Lemma 7.3,

$$S_{\mathbb{R}}\Big(\sum_{j=1}^{J} v_n^j\Big) \lesssim \sum_{j=1}^{J} S_{\mathbb{R}}(v_n^j) + C_J \sum_{j\neq k} \|v_n^j v_n^k\|_{L_{t,x}^{\frac{d+2}{d-2}}}^{\frac{d+2}{d-2}} \lesssim_{E_c,\delta} 1 + C_J o(1) \quad \text{as} \quad n\to\infty.$$

Similarly,

$$\begin{aligned}\Big\|\sum_{j=1}^J \nabla v_n^j\Big\|^2_{L_{t,x}^{\frac{2(d+2)}{d}}} &= \Big\|\Big(\sum_{j=1}^J \nabla v_n^j\Big)^2\Big\|_{L_{t,x}^{\frac{d+2}{d}}} \\ &\lesssim \sum_{j=1}^J \|\nabla v_n^j\|^2_{L_{t,x}^{\frac{2(d+2)}{d}}} + \sum_{j\neq k} \|\nabla v_n^j \nabla v_n^k\|_{L_{t,x}^{\frac{d+2}{d}}} \\ &\lesssim_{E_c,\delta} 1 + o(1) \quad \text{as} \quad n\to\infty.\end{aligned}$$

This completes the proof of (7.11). The same argument combined with (7.5) shows that given $\eta > 0$, there exists $J' = J'(\eta)$ such that

$$\limsup_{n\to\infty}\Big\|\sum_{j=J'}^J v_n^j\Big\|_{\dot X^1(\mathbb{R}\times\mathbb{R}^d)} \le \eta \quad \text{uniformly in } J \ge J'. \tag{7.12}$$

Now we are ready to construct an approximate solution to the defocusing energy-critical NLS. For each n and J, we define

$$u_n^J := \sum_{j=1}^J v_n^j + e^{it\Delta} w_n^J.$$

Obviously u_n^J is defined globally in time. In order to apply the stability result, it suffices to verify the following three claims for u_n^J:

Claim 1: $\|u_n^J(0) - u_n(0)\|_{\dot H_x^1} \to 0$ as $n \to \infty$ for any J.

Claim 2: $\limsup_{n\to\infty} \|u_n^J\|_{\dot X^1(\mathbb{R}\times\mathbb{R}^d)} \lesssim_{E_c,\delta} 1$ uniformly in J.

Claim 3: $\lim_{J\to J^*} \limsup_{n\to\infty} \big\|\nabla\big[(i\partial_t + \Delta)u_n^J - |u_n^J|^{\frac{4}{d-2}} u_n^J\big]\big\|_{N^0(\mathbb{R})} = 0$.

The three claims imply that for sufficiently large n and J, u_n^J is an approximate solution to the defocusing energy-critical NLS with finite scattering size, which asymptotically matches $u_n(0)$ at time $t = 0$. Using the stability result we see that for n, J sufficiently large, the solution u_n inherits the space-time bounds of u_n^J, thus contradicting (7.3). Therefore, to complete the treatment of the second scenario, it suffices to verify the three claims above.

The first claim follows trivially from (7.4) and (7.7). To derive the second claim, we use (7.11) and the Strichartz inequality, as follows:

$$\limsup_{n\to\infty} \|u_n^J\|_{\dot X^1(\mathbb{R}\times\mathbb{R}^d)} \lesssim \limsup_{n\to\infty}\Big\|\sum_{j=1}^J v_n^j\Big\|_{\dot X^1(\mathbb{R}\times\mathbb{R}^d)} + \limsup_{n\to\infty} \|w_n^J\|_{\dot H_x^1} \lesssim_{E_c,\delta} 1.$$

It remains to verify the third claim. Adopting the notation $F(z) = |z|^{\frac{4}{d-2}} z$, we write

$$
\begin{aligned}
(i\partial_t + \Delta)u_n^J - F(u_n^J) &= \sum_{j=1}^{J} F(v_n^j) - F(u_n^J) \\
&= \sum_{j=1}^{J} F(v_n^j) - F\Big(\sum_{j=1}^{J} v_n^j\Big) + F\big(u_n^J - e^{it\Delta}w_n^J\big) - F(u_n^J).
\end{aligned}
\tag{7.13}
$$

Taking the derivative, in dimensions $d \geq 6$ we estimate

$$
\bigg|\nabla\bigg\{\sum_{j=1}^{J} F(v_n^j) - F\Big(\sum_{j=1}^{J} v_n^j\Big)\bigg\}\bigg| \lesssim_J \sum_{j\neq k} |\nabla v_n^j||v_n^k|^{\frac{4}{d-2}}.
$$

In dimensions $d = 3, 4, 5$ there is an additional term on the right-hand side of the inequality above, namely, $\sum_{j\neq k} |\nabla v_n^j||v_n^k||v_n^j|^{\frac{6-d}{d-2}}$. Using (7.9) and Lemma 7.3, in dimensions $d \geq 6$ we estimate

$$
\begin{aligned}
\bigg\|\nabla\bigg[\sum_{j=1}^{J} F(v_n^j) - F\Big(\sum_{j=1}^{J} v_n^j\Big)\bigg]\bigg\|_{N^0(\mathbb{R})} &\lesssim_J \sum_{j\neq k} \big\||\nabla v_n^j||v_n^k|^{\frac{4}{d-2}}\big\|_{L_{t,x}^{\frac{2(d+2)}{d+4}}} \\
&\lesssim_J \sum_{j\neq k} \|\nabla v_n^j v_n^k\|_{L_{t,x}^{\frac{d+2}{d-1}}}^{\frac{4}{d-2}} \|\nabla v_n^k\|_{L_{t,x}^{\frac{2(d+2)}{d}}}^{\frac{d-6}{d-2}} \\
&\lesssim_{J,E_c,\delta} o(1) \quad \text{as} \quad n \to \infty.
\end{aligned}
$$

The additional term in dimensions $d = 3, 4, 5$ can be treated analogously. Thus,

$$
\lim_{J\to J^*} \limsup_{n\to\infty} \bigg\|\nabla\bigg[\sum_{j=1}^{J} F(v_n^j) - F\Big(\sum_{j=1}^{J} v_n^j\Big)\bigg]\bigg\|_{N^0(\mathbb{R})} = 0. \tag{7.14}
$$

We now turn to estimating the second difference in (7.13). We will show that

$$
\lim_{J\to J^*} \limsup_{n\to\infty} \big\|\nabla\big[F(u_n^J - e^{it\Delta}w_n^J) - F(u_n^J)\big]\big\|_{N^0(\mathbb{R})} = 0. \tag{7.15}
$$

In dimensions $d \geq 6$,

$$
\begin{aligned}
\big\|\nabla\big[F\big(u_n^J - e^{it\Delta}w_n^J\big) - F(u_n^J)\big]\big\|_{L_{t,x}^{\frac{2(d+2)}{d+4}}} &\lesssim \|\nabla e^{it\Delta}w_n^J\|_{L_{t,x}^{\frac{2(d+2)}{d}}} \|e^{it\Delta}w_n^J\|_{L_{t,x}^{\frac{2(d+2)}{d-2}}}^{\frac{4}{d-2}} \\
&\quad + \|\nabla u_n^J\|_{L_{t,x}^{\frac{2(d+2)}{d}}} \|e^{it\Delta}w_n^J\|_{L_{t,x}^{\frac{2(d+2)}{d-2}}}^{\frac{4}{d-2}} \\
&\quad + \big\||u_n^J|^{\frac{4}{d-2}} \nabla e^{it\Delta}w_n^J\big\|_{L_{t,x}^{\frac{2(d+2)}{d+4}}}.
\end{aligned}
$$

In dimensions $d = 3, 4, 5$, one must add the term

$$
\|\nabla u_n^J\|_{L_{t,x}^{\frac{2(d+2)}{d}}} \|e^{it\Delta}w_n^J\|_{L_{t,x}^{\frac{2(d+2)}{d-2}}} \|u_n^J\|_{L_{t,x}^{\frac{2(d+2)}{d-2}}}^{\frac{6-d}{d-2}}
$$

to the right-hand side above. Using the second claim together with (4.2), and the Strichartz inequality combined with the fact that w_n^J is bounded in $\dot H_x^1$, we see that (7.15) will follow once we establish

$$\lim_{J\to J^*}\limsup_{n\to\infty}\big\| |u_n^J|^{\frac{4}{d-2}}\nabla e^{it\Delta}w_n^J\big\|_{L_{t,x}^{\frac{2(d+2)}{d+4}}(\mathbb{R}\times\mathbb{R}^d)} = 0. \tag{7.16}$$

We will only prove (7.16) in dimensions $d\geq 6$. We leave the remaining low dimensions as an exercise for the conscientious reader. Using Hölder's inequality, the second claim, and the Strichartz inequality, we get

$$\begin{aligned}
\big\||u_n^J|^{\frac{4}{d-2}}\nabla e^{it\Delta}w_n^J\big\|_{L_{t,x}^{\frac{2(d+2)}{d+4}}} &\lesssim \|u_n^J\nabla e^{it\Delta}w_n^J\|_{L_{t,x}^{\frac{d+2}{d-1}}}^{\frac{4}{d-2}}\|\nabla e^{it\Delta}w_n^J\|_{L_{t,x}^{\frac{2(d+2)}{d}}}^{\frac{d-6}{d-2}}\\
&\lesssim_{E_c,\delta} \|e^{it\Delta}w_n^J\nabla e^{it\Delta}w_n^J\|_{L_{t,x}^{\frac{d+2}{d-1}}}^{\frac{4}{d-2}} + \Big\|\sum_{j=1}^J v_n^j\nabla e^{it\Delta}w_n^J\Big\|_{L_{t,x}^{\frac{d+2}{d-1}}}^{\frac{4}{d-2}}\\
&\lesssim_{E_c,\delta} \|e^{it\Delta}w_n^J\|_{L_{t,x}^{\frac{2(d+2)}{d-2}}}^{\frac{4}{d-2}}\|\nabla e^{it\Delta}w_n^J\|_{L_{t,x}^{\frac{2(d+2)}{d}}}^{\frac{4}{d-2}}\\
&\qquad + \Big\|\sum_{j=1}^J v_n^j\nabla e^{it\Delta}w_n^J\Big\|_{L_{t,x}^{\frac{d+2}{d-1}}}^{\frac{4}{d-2}}\\
&\lesssim_{E_c,\delta} \|e^{it\Delta}w_n^J\|_{L_{t,x}^{\frac{2(d+2)}{d-2}}}^{\frac{4}{d-2}} + \Big\|\sum_{j=1}^J v_n^j\nabla e^{it\Delta}w_n^J\Big\|_{L_{t,x}^{\frac{d+2}{d-1}}}^{\frac{4}{d-2}}.
\end{aligned}$$

By (4.2), the contribution of the first term to (7.16) is acceptable. We now turn to the second term.

By (7.12),

$$\begin{aligned}
\limsup_{n\to\infty}\Big\|\Big(\sum_{j=J'}^J v_n^j\Big)\nabla e^{it\Delta}w_n^J\Big\|_{L_{t,x}^{\frac{d+2}{d-1}}} &\lesssim \limsup_{n\to\infty}\Big\|\sum_{j=J'}^J v_n^j\Big\|_{\dot X^1}\|\nabla e^{it\Delta}w_n^J\|_{L_{t,x}^{\frac{2(d+2)}{d}}}\\
&\lesssim_{E_c,\delta} \eta,
\end{aligned}$$

where $\eta>0$ is arbitrary and $J'=J'(\eta)$ is as in (7.12). Thus, proving (7.16) reduces to showing

$$\lim_{J\to J^*}\limsup_{n\to\infty}\|v_n^j\nabla e^{it\Delta}w_n^J\|_{L_{t,x}^{\frac{d+2}{d-1}}} = 0 \quad\text{for each}\quad 1\leq j<J'. \tag{7.17}$$

Fix $1\leq j<J'$. By a change of variables,

$$\|v_n^j\nabla e^{it\Delta}w_n^J\|_{L_{t,x}^{\frac{d+2}{d-1}}} = \big\|v^j\nabla\tilde w_n^J\big\|_{L_{t,x}^{\frac{d+2}{d-1}}},$$

where $\tilde w_n^J := (T_n^j)^{-1}\big(e^{it\Delta}w_n^J\big)$. Note that

$$\|\tilde w_n^J\|_{\dot X^1(\mathbb{R}\times\mathbb{R}^d)} = \|e^{it\Delta}w_n^J\|_{\dot X^1(\mathbb{R}\times\mathbb{R}^d)}. \tag{7.18}$$

By density, we may assume $v^j \in C_c^\infty(\mathbb{R}\times\mathbb{R}^d)$. Invoking Hölder's inequality, it thus suffices to show

$$\lim_{J\to J^*} \limsup_{n\to\infty} \|\nabla \tilde{w}_n^J\|_{L^2_{t,x}(K)} = 0$$

for any compact $K \subset \mathbb{R}\times\mathbb{R}^d$. This however follows immediately from Lemma 2.12, (4.2), and (7.18), thus completing the proof of (7.17).

This proves (7.16) and so (7.15). Combining (7.14) and (7.15) yields the third claim. This completes the treatment of the second scenario and so the proof of the proposition. □

Chapter 8

Existence of minimal blowup solutions and their properties

In this section we prove the existence of minimal counterexamples to Theorem 6.1 and we study some of their properties.

Theorem 8.1 (Existence of minimal counterexamples). *Suppose Theorem 6.1 fails to be true. Then there exist a critical energy $0 < E_c < \infty$ and a maximal-lifespan solution $u : I \times \mathbb{R}^d \to \mathbb{C}$ to the defocusing energy-critical NLS with $E(u) = E_c$, which blows up in both time directions in the sense that*

$$S_{\geq 0}(u) = S_{\leq 0}(u) = \infty,$$

and whose orbit $\{u(t) : t \in \mathbb{R}\}$ is precompact in $\dot{H}^1_x$ modulo scaling and spatial translations.

Proof. If Theorem 6.1 fails to be true, then there must exist a critical energy $0 < E_c < \infty$ and a sequence of solutions $u_n : I_n \times \mathbb{R}^d \to \mathbb{C}$ such that $E(u_n) \to E_c$ and $S_{I_n}(u_n) \to \infty$. Let $t_n \in I_n$ be such that $S_{\geq t_n}(u_n) = S_{\leq t_n}(u_n) = \frac{1}{2} S_{I_n}(u_n)$; then

$$\lim_{n\to\infty} S_{\geq t_n}(u_n) = \lim_{n\to\infty} S_{\leq t_n}(u_n) = \infty. \tag{8.1}$$

Applying Proposition 7.2 and passing to a subsequence, we find $\phi \in \dot{H}^1_x$ such that $u_n(t_n)$ converge to ϕ in $\dot{H}^1_x$ modulo scaling and spatial translations. Using the scaling and space-translation invariance of the equation and modifying $u_n(t_n)$ appropriately, we may assume $u_n(t_n) \to \phi$ in $\dot{H}^1_x$. In particular, $E(\phi) = E_c$.

Let $u : I \times \mathbb{R}^d \to \mathbb{C}$ be the maximal-lifespan solution to the defocusing energy-critical NLS with initial data $u(0) = \phi$. From the stability result Theorem 5.3 and (8.1), we get

$$S_{\geq 0}(u) = S_{\leq 0}(u) = \infty. \tag{8.2}$$

H. Koch et al., *Dispersive Equations and Nonlinear Waves: Generalized Korteweg–de Vries, Nonlinear Schrödinger, Wave and Schrödinger Maps*, Oberwolfach Seminars 45,
DOI 10.1007/978-3-0348-0736-4_22,

Finally, we prove that the orbit of u is precompact in $\dot{H}_x^1$ modulo scaling and space translations. For any sequence $\{t_n'\} \subset I$, (8.2) implies $S_{\geq t_n'}(u) = S_{\leq t_n'}(u) = \infty$. Thus, by Proposition 7.2, we see that $u(t_n')$ admits a subsequence that converges in $\dot{H}_x^1$ modulo scaling and space translations. This completes the proof of the theorem. $\square$

By Corollary 11.2, the maximal-lifespan solution found in Theorem 8.1 is *almost periodic modulo symmetries*, that is, there exist (possibly discontinuous) functions $N : I \to \mathbb{R}^+$, $x : I \to \mathbb{R}^d$, and $C : \mathbb{R}^+ \to \mathbb{R}^+$ such that

$$\int_{|x-x(t)|\geq C(\eta)/N(t)} |\nabla u(t,x)|^2\, dx + \int_{|\xi|\geq C(\eta)N(t)} |\xi \hat{u}(t,\xi)|^2\, d\xi \leq \eta$$

for all $t \in I$ and $\eta > 0$. We refer to the function N as the *frequency scale function*, x is the *spatial center function*, and C is the *compactness modulus function*.

Another consequence of the precompactness in $\dot{H}_x^1$ modulo symmetries of the orbit of the solution found in Theorem 8.1 is that for every $\eta > 0$ there exists $c(\eta) > 0$ such that

$$\int_{|x-x(t)|\leq c(\eta)/N(t)} |\nabla u(t,x)|^2\, dx + \int_{|\xi|\leq c(\eta)N(t)} |\xi \hat{u}(t,\xi)|^2\, d\xi \leq \eta,$$

uniformly for all $t \in I$.

In what follows, we record some basic properties of almost periodic (modulo symmetries) solutions. We start with the following definition:

Definition 8.2 (Normalized solution)**.** Let $u : I \times \mathbb{R}^d \to C$ be a solution to (6.1), which is almost periodic modulo symmetries with parameters $N(t)$ and $x(t)$. We say that u is *normalized* if the lifespan I contains zero and

$$N(0) = 1 \quad \text{and} \quad x(0) = 0.$$

More generally, we can define the *normalization* of a solution u at a time $t_0 \in I$ by

$$u^{[t_0]}(s,x) := N(t_0)^{-\frac{d-2}{2}} u\big(t_0 + N(t_0)^{-2}s, x(t_0) + N(t_0)^{-1}x\big). \tag{8.3}$$

Note that $u^{[t_0]}$ is a normalized solution which is almost periodic modulo symmetries with lifespan $I^{[t_0]} := \{s \in \mathbb{R} : t_0 + N(t_0)^{-2}s \in I\}$. The parameters of $u^{[t_0]}$ satisfy

$$N^{[t_0]}(s) := \frac{N\big(t_0 + sN(t_0)^{-2}\big)}{N(t_0)} \quad \text{and} \quad x^{[t_0]}(s) := N(t_0)\big[x\big(t_0 + sN(t_0)^{-2}\big) - x(t_0)\big]$$

and $u^{[t_0]}$ has the same compactness modulus function as u. Furthermore, if u is a maximal-lifespan solution, then so is $u^{[t_0]}$.

Lemma 8.3 (Local constancy of $N(t)$ and $x(t)$, [20, 22])**.** *Let $u : I \times \mathbb{R}^d \to \mathbb{C}$ be a non-zero almost periodic modulo symmetries solution to (6.1) with parameters $N(t)$ and $x(t)$. Then there exists a small number δ, depending on u, such that for every $t_0 \in I$ we have*

$$\left[t_0 - \delta N(t_0)^{-2}, t_0 + \delta N(t_0)^{-2}\right] \subset I \tag{8.4}$$

and

$$N(t) \sim_u N(t_0) \quad \text{and} \quad |x(t) - x(t_0)| \lesssim_u N(t_0)^{-1} \tag{8.5}$$

whenever $|t - t_0| \leq \delta N(t_0)^{-2}$.

Proof. We first prove (8.4). Arguing by contradiction, we assume (8.4) fails. Thus, there exist sequences $t_n \in I$ and $\delta_n \to 0$ such that $t_n + \delta_n N(t_n)^{-2} \notin I$ for all n. Then $u^{[t_n]}$ given by (8.3) are normalized solutions whose lifespans $I^{[t_n]}$ contain 0 but not δ_n. Invoking almost periodicity and passing to a subsequence, we conclude that $u^{[t_n]}(0)$ converge to some $v_0 \in \dot{H}^1_x$. Let $v : J \times \mathbb{R}^d \to \mathbb{C}$ be the maximal-lifespan solution with data $v(0) = v_0$. By the local well-posedness theory, J is an open interval and so contains δ_n for all sufficiently large n. By the stability result Theorem 5.3, for n sufficiently large we must have that $\delta_n \in I^{[t_n]}$. This contradicts the hypothesis and so gives (8.4).

We now turn to (8.5). Again, we argue by contradiction, taking δ even smaller if necessary. Suppose one of the two claims in (8.5) failed no matter how small one chose δ. Then one can find sequences $t_n, t_n' \in I$ so that $s_n := (t_n' - t_n)N(t_n)^2 \to 0$ but $N(t_n')/N(t_n)$ converge to either zero or infinity (if the first claim failed) or $|x(t_n') - x(t_n)|N(t_n) \to \infty$ (if the second claim failed). Therefore, $N^{[t_n]}(s_n)$ converge to either zero or infinity or $x^{[t_n]}(s_n) \to \infty$. By almost periodicity, this implies that $u^{[t_n]}(s_n)$ must converge weakly to zero.

On the other hand, using almost periodicity and passing to a subsequence we find that $u^{[t_n]}(0)$ converge to some $v_0 \in \dot{H}^1_x$. As $s_n \to 0$, we conclude that $u^{[t_n]}(s_n)$ converge to v_0 in $\dot{H}^1_x$. Thus $v_0 = 0$ and $E(u) = E(u^{[t_n]}) \to E(v_0) = 0$. This means $u \equiv 0$, a contradiction. This completes the proof of (8.5). □

An immediate consequence of Lemma 8.3 is the following corollary, which describes the behaviour of the frequency scale function.

Corollary 8.4 ($N(t)$ at blowup, [20, 22])**.** *Let $u : I \times \mathbb{R}^d \to \mathbb{C}$ be a non-zero maximal-lifespan solution to (6.1) that is almost periodic modulo symmetries with frequency scale function $N : I \to \mathbb{R}^+$. If T is any finite endpoint of the lifespan I, then $N(t) \gtrsim_u |T - t|^{-1/2}$; in particular, $\lim_{t\to T} N(t) = \infty$. If I is infinite or semi-infinite, then for any $t_0 \in I$ we have $N(t) \gtrsim_u \min\{N(t_0), |t - t_0|^{-1/2}\}$.*

Proof. Exercise! □

Our next result shows how energy-critical norms of an almost periodic solution can be computed in terms of its frequency scale function; see [20] for the mass-critical analogue.

Lemma 8.5 (Strichartz norms via $N(t)$, [22]). *Let $u : I \times \mathbb{R}^d \to \mathbb{C}$ be a non-zero almost periodic modulo symmetries solution to (6.1) with frequency scale function $N : I \to \mathbb{R}^+$. Then*

$$\int_I N(t)^2\,dt \lesssim_u \int_I \int_{\mathbb{R}^d} |u(t,x)|^{\frac{2(d+2)}{d-2}}\,dx\,dt \lesssim_u 1 + \int_I N(t)^2\,dt.$$

Proof. We first prove

$$\int_I \int_{\mathbb{R}^d} |u(t,x)|^{\frac{2(d+2)}{d-2}}\,dx\,dt \lesssim_u 1 + \int_I N(t)^2\,dt. \tag{8.6}$$

Let $0 < \eta < 1$ be a small parameter to be chosen shortly and partition I into subintervals I_j so that

$$\int_{I_j} N(t)^2\,dt \le \eta; \tag{8.7}$$

this requires at most $\eta^{-1} \times \mathrm{RHS}(8.6)$ many intervals.

For each j, we may choose $t_j \in I_j$ so that

$$N(t_j)^2 |I_j| \le 2\eta. \tag{8.8}$$

By Sobolev embedding, Strichartz, Hölder, and Bernstein, we obtain

$$\begin{aligned}
\|u\|_{L_{t,x}^{\frac{2(d+2)}{d-2}}} &\lesssim \|\nabla u\|_{L_t^{\frac{2(d+2)}{d-2}} L_x^{\frac{2d(d+2)}{d^2+4}}} \\
&\lesssim \|e^{i(t-t_j)\Delta} \nabla u(t_j)\|_{L_t^{\frac{2(d+2)}{d-2}} L_x^{\frac{2d(d+2)}{d^2+4}}} + \|\nabla u\|_{L_t^{\frac{2(d+2)}{d-2}} L_x^{\frac{2d(d+2)}{d^2+4}}}^{\frac{d+2}{d-2}} \\
&\lesssim \|u_{\ge N_0}(t_j)\|_{\dot H_x^1} + |I_j|^{\frac{d-2}{2(d+2)}} N_0^{\frac{d-2}{d+2}} \|u(t_j)\|_{\dot H_x^1} + \|\nabla u\|_{L_t^{\frac{2(d+2)}{d-2}} L_x^{\frac{2d(d+2)}{d^2+4}}}^{\frac{d+2}{d-2}},
\end{aligned}$$

where all space-time norms are over $I_j \times \mathbb{R}^d$. Choosing N_0 as a large multiple of $N(t_j)$ and using almost periodicity modulo symmetries, we can make the first term as small as we wish. Subsequently, choosing η sufficiently small depending on $E(u)$ and invoking (8.8), we may also render the second term arbitrarily small. Thus, by the usual bootstrap argument we obtain

$$\|u\|_{L_{t,x}^{\frac{2(d+2)}{d-2}}(I_j \times \mathbb{R}^d)} \lesssim \|\nabla u\|_{L_t^{\frac{2(d+2)}{d-2}} L_x^{\frac{2d(d+2)}{d^2+4}}(I_j \times \mathbb{R}^d)} \le 1.$$

Using the bound on the number of intervals I_j, this leads to (8.6).

Next we prove

$$\int_I \int_{\mathbb{R}^d} |u(t,x)|^{\frac{2(d+2)}{d-2}}\,dx\,dt \gtrsim_u \int_I N(t)^2\,dt. \tag{8.9}$$

Using almost periodicity and Sobolev embedding, we can guarantee that

$$\int_{|x-x(t)|\leq C(u)N(t)^{-1}} |u(t,x)|^{\frac{2d}{d-2}}\,dx \gtrsim_u 1 \tag{8.10}$$

uniformly for $t \in I$. On the other hand, by Hölder,

$$\int_{\mathbb{R}^d} |u(t,x)|^{\frac{2(d+2)}{d-2}}\,dx \gtrsim_u \Big(\int_{|x-x(t)|\leq C(u)N(t)^{-1}} |u(t,x)|^{\frac{2d}{d-2}}\,dx\Big)^{\frac{d+2}{d}} N(t)^2.$$

Using (8.10) and integrating over I we obtain (8.9). □

Corollary 8.6. *Let $u : I\times\mathbb{R}^d \to \mathbb{C}$ be a non-zero almost periodic modulo symmetries solution to (6.1) with frequency scale function $N : I \to \mathbb{R}^+$. Then*

$$\|\nabla u\|^2_{L^2_t L_x^{\frac{2d}{d-2}}(I\times\mathbb{R}^d)} \lesssim_u 1 + \int_I N(t)^2\,dt.$$

Proof. Exercise! □

The next proposition tells us that for a minimal blowup solution $u : I\times\mathbb{R}^d \to \mathbb{C}$, the free evolution coming for the endpoints of the maximal-lifespan I converges weakly to zero in $\dot{H}^1_x$. Intuitively, we expect this to be the case since the free evolution is nothing but radiation and radiation does not directly contribute to blowup. However, a minimal blowup solution needs all its norm in order to blow up and so cannot waste any norm on a radiation term.

Proposition 8.7 (Reduced Duhamel formulas, [22, 36])**.** *Let $u : I \times \mathbb{R}^d \to \mathbb{C}$ be a maximal-lifespan almost periodic modulo symmetries solution to (6.1). Then $e^{-it\Delta}u(t)$ converges weakly to zero in $\dot{H}^1_x$ as $t \to \sup I$ or $t \to \inf I$. In particular, we have the 'reduced' Duhamel formulas*

$$\begin{aligned} u(t) &= i \lim_{T\to\sup I} \int_t^T e^{i(t-s)\Delta}|u(s)|^{\frac{4}{d-2}}u(s)\,ds \\ &= -i \lim_{T\to\inf I} \int_T^t e^{i(t-s)\Delta}|u(s)|^{\frac{4}{d-2}}u(s)\,ds, \end{aligned} \tag{8.11}$$

where the limits are to be understood in the weak $\dot{H}^1_x$ topology.

Proof. We prove the claim for the case $t \to \sup I$; the proof in the reverse time direction is similar.

Assume first that $\sup I < \infty$. Then by Corollary 8.4,

$$\lim_{t\to\sup I} N(t) = \infty.$$

By almost periodicity, this implies that $u(t)$ converges weakly to zero as $t \to \sup I$. As $\sup I < \infty$ and the map $t \mapsto e^{it\Delta}$ is continuous in the strong operator topology on $\dot{H}^1_x$, we see that $e^{-it\Delta}u(t)$ converges weakly to zero, as desired.

Now suppose that $\sup I = \infty$. We need to prove that

$$\lim_{t\to\infty} \langle u(t), e^{it\Delta}\phi\rangle_{\dot H^1_x} = 0$$

for all test functions $\phi \in C_c^\infty(\mathbb{R}^d)$. Let $\eta > 0$ be a small parameter. By Cauchy–Schwarz and almost periodicity,

$$\begin{aligned}
\Big|\langle u(t), e^{it\Delta}\phi\rangle_{\dot H^1_x}\Big|^2 &\lesssim \Big|\int_{|x-x(t)|\leq C(\eta)/N(t)} \nabla u(t,x)\overline{e^{it\Delta}\nabla\phi(x)}\,dx\Big|^2 \\
&\qquad + \Big|\int_{|x-x(t)|\geq C(\eta)/N(t)} \nabla u(t,x)\overline{e^{it\Delta}\nabla\phi(x)}\,dx\Big|^2 \\
&\lesssim \|u(t)\|_{\dot H^1_x}^2 \int_{|x-x(t)|\leq C(\eta)/N(t)} |e^{it\Delta}\nabla\phi(x)|^2\,dx + \eta\|\phi\|_{\dot H^1_x}^2.
\end{aligned}$$

Therefore, to obtain the claim we merely need to show that

$$\int_{|x-x(t)|\leq C(\eta)/N(t)} |e^{it\Delta}\nabla\phi(x)|^2\,dx \to 0 \quad \text{as} \quad \eta\to 0.$$

This follows from Lemma 8.8 below, Corollary 8.4, and a change of variables. □

Lemma 8.8 (Fraunhofer formula)**.** *For $\psi \in L^2(\mathbb{R}^d)$ and $t\to\pm\infty$,*

$$\big\|[e^{it\Delta}\psi](x) - (2it)^{-\frac{d}{2}} e^{i|x|^2/4t}\hat\psi\big(\tfrac{x}{2t}\big)\big\|_{L^2_x} \to 0. \tag{8.12}$$

Proof. This asymptotic is most easily understood in terms of stationary phase. However, our proof will be based on the exact formula for the Schrödinger propagator, which we derived in Section 2. We have the identity

$$\begin{aligned}
\text{LHS}(8.12) &= \Big\|(4\pi it)^{-\frac{d}{2}}\int_{\mathbb{R}^d} e^{i|x-y|^2/4t}[1-e^{-i|y|^2/4t}]\psi(y)\,dy\Big\|_{L^2_x} \\
&= \big\|e^{it\Delta}\big[(1-e^{-i|\cdot|^2/4t})\psi\big]\big\|_{L^2_x} \\
&= \big\|(1-e^{-i|\cdot|^2/4t})\psi\big\|_{L^2_x}.
\end{aligned}$$

The result now follows from the dominated convergence theorem. □

So far we have proved that if Theorem 6.1 fails, then there exists a minimal witness to its failure. This is a maximal-lifespan almost periodic solution $u : I\times\mathbb{R}^d \to \mathbb{C}$ which blows up in both time directions; see Theorem 8.1. Moreover, we have recorded some basic properties satisfied by the modulation parameters $N(t)$ and $x(t)$. Thus, to prove Theorem 6.1 we have to rule out the existence of these minimal counterexamples. In order to achieve this, we need more quantitative information regarding $N(t)$ and $x(t)$. The first modest step in this direction is the following theorem, which asserts that we may assume $N(t)$ is bounded from below;

the price we pay for this information is that we can no longer guarantee that u blows up in *both* time directions.

For an argument that is upside down relative to the one we present below, see Theorem 3.3 in [35]. This reference treats the mass-critical NLS and restricts attention to almost periodic solutions with $N(t) \le 1$.

Theorem 8.9. *Suppose Theorem 6.1 fails to be true. Then there exists an almost periodic modulo symmetries solution $u : I \times \mathbb{R}^d \to \mathbb{C}$ such that $S_I(u) = \infty$ and*

$$N(t) \ge 1 \quad \text{for all} \quad t \in I. \tag{8.13}$$

Proof. By Theorem 8.1, there exists a maximal-lifespan solution $v : J \times \mathbb{R}^d \to \mathbb{C}$ to the defocusing energy-critical NLS which is almost periodic modulo symmetries and which blows up in both time directions in the sense that $S_{\ge 0}(v) = S_{\le 0}(v) = \infty$. Let $N_v(t)$ denote the frequency scale function associated to v. We will obtain the desired u satisfying (8.13) from v, by rescaling appropriately.

Write J as a nested union of compact intervals $J_1 \subset J_2 \subset \cdots \subset J$. On each compact interval J_n, we have $v \in C_t \dot{H}^1_x(J_n \times \mathbb{R}^d)$, which easily implies that $N_v(t)$ is bounded above and below on J_n. Thus, we may find $t_n \in J_n$ such that

$$N_v(t_n) \le 2N_v(t) \quad \text{for all} \quad t \in J_n. \tag{8.14}$$

Now consider the normalizations $v^{[t_n]} : I_n \times \mathbb{R}^d \to \mathbb{C}$ with $I_n := \{t \in \mathbb{R} : t_n + N_v(t_n)^{-2} t \in J_n\}$. Using almost periodicity and passing to a subsequence, we get that $v^{[t_n]}(0)$ converge in $\dot{H}^1_x$ to some u_0. From the conservation of energy, we see that u_0 is not identically zero. Let $u : (-T_-, T_+) \times \mathbb{R}^d \to \mathbb{C}$ be the maximal-lifespan solution with data $u(0) = u_0$.

Now let $v_n : \tilde{I}_n \times \mathbb{R}^d \to \mathbb{C}$ be the maximal-lifespan solution which agrees with $v^{[t_n]}$ on I_n. If K is any compact subinterval of $(-T_-, T_+)$ containing 0, then $S_K(u) < \infty$. From the stability result Theorem 5.3, for sufficiently large n we must have $K \subseteq \tilde{I}_n$ and $S_K(v_n) < \infty$ uniformly in n. As $S_{J_n}(v) = S_{I_n}(v_n) \to \infty$ as $n \to \infty$, we must have $I_n \nsubseteq K$ for n large. Passing to subsequence if necessary, this leaves only two possibilities:

- For every $0 < t < T_+$, $[0, t] \subseteq I_n$ for all sufficiently large n.
- For every $-T_- < t < 0$, $[t, 0] \subseteq I_n$ for all sufficiently large n.

By time reversal symmetry, it suffices to consider the former possibility. Let $I := [0, T_+)$. We will prove that $u : I \times \mathbb{R}^d \to \mathbb{C}$ satisfies the conclusions of Theorem 8.9.

We first note that $u : I \times \mathbb{R}^d \to \mathbb{C}$ is almost periodic modulo symmetries. Indeed, for any $0 \le t < T_+$, $u(t)$ can be approximated to arbitrary accuracy in $\dot{H}^1_x$ by $v^{[t_n]}(t)$, which is a rescaled version of a function in the orbit $\{v(t) : t \in J\}$. As the orbit of v is precompact in $\dot{H}^1_x$ modulo symmetries, then so is $\{u(t) : 0 \le t < T_+\}$.

Next we prove that $S_I(u) = \infty$. Otherwise we would have $T_+ = \infty$ and $[0, \infty) \subseteq I_n$ for n large. Moreover, by the stability theory, for n large we get

$S_{\geq 0}(v^{[t_n]}) = S_{\geq t_n}(v) < \infty$, which contradicts the fact that v blows up forward in time.

Finally, we prove (8.13). Let $\eta > 0$ to be chosen later. Fix $t \in I$. By the stability result, for n large we have $t \in I_n$ and

$$\|v^{[t_n]}(t) - u(t)\|_{\dot{H}^1_x} \to 0 \quad \text{as} \quad n \to \infty.$$

Combining this with (8.14) and almost periodicity, we find that there exists $c(\eta) > 0$ such that

$$\begin{aligned} \eta &\geq \int_{|\xi| \leq c(\eta) N_v(t)} |\xi \hat{v}(t,\xi)|^2 \, d\xi = \int_{|\xi| \leq c(\eta) \frac{N_v(t_n + N_v(t_n)^{-2} t)}{N_v(t_n)}} |\xi \widehat{v^{[t_n]}}(t,\xi)|^2 \, d\xi \\ &\geq \int_{|\xi| \leq \frac{1}{2} c(\eta)} |\xi \widehat{v^{[t_n]}}(t,\xi)|^2 \, d\xi \to \int_{|\xi| \leq \frac{1}{2} c(\eta)} |\xi \hat{u}(t,\xi)|^2 \, d\xi. \end{aligned}$$

Recalling the definition of almost periodicity, we derive (8.13). This completes the proof of the theorem. □

Putting together the results of this section we can restrict attention to the following very specific enemy to Theorem 6.1:

Theorem 8.10. *Suppose Theorem 6.1 fails to be true. Then there exists an almost periodic solution $u : [0, T_{\max}) \times \mathbb{R}^d \to \mathbb{C}$ such that*

$$S_{[0,T_{\max})}(u) = \int_0^{T_{\max}} \int_{\mathbb{R}^d} |u(t,x)|^{\frac{2(d+2)}{d-2}} \, dx \, dt = +\infty.$$

Moreover, we may write $[0, T_{\max}) = \bigcup_k J_k$, with J_k being intervals of local constancy and

$$N(t) \equiv N_k \geq 1 \quad \textit{for each} \quad t \in J_k.$$

In the following two sections we will see how to preclude the existence of the almost periodic solution described in Theorem 8.10 for the defocusing energy-critical NLS in four spatial dimensions:

$$i\partial_t u = -\Delta u + |u|^2 u \quad \text{with} \quad u(0) = u_0 \in \dot{H}^1_x(\mathbb{R}^4). \tag{8.15}$$

Some of the arguments that follow work also in higher dimensions, as well as for the focusing equation; however, in these notes we are not aiming for the greatest generality, but rather we try to demonstrate how these techniques can be used to settle Theorem 6.1 in the particular case $d = 4$.

Before we launch into the involved argument that will preclude the existence of the enemy described in Theorem 8.10, let us first pause and collect the rewards of this section. In particular, we will see that our enemy must be global forward in time; strictly speaking, this step is not necessary for the argument that follows, but it is always good to realize how far we have come and how much further there is to go.

Theorem 8.11. *Let $u : [0, T_{\max}) \times \mathbb{R}^4 \to \mathbb{C}$ be an almost periodic solution to (8.15) with $S_{[0,T_{\max})}(u) = \infty$. Then $T_{\max} = \infty$.*

Proof. We argue by contradiction. Assume that $T_{\max} < \infty$. Using Proposition 8.7, the Strichartz inequality, Hölder's inequality, and the conservation of energy, we estimate

$$\begin{aligned}
\|u_{\geq N}(t)\|_{L_x^2} &\lesssim \|P_{\geq N}(|u|^2 u)\|_{L_t^2 L_x^{4/3}([t,T_{\max})\times\mathbb{R}^4)} \\
&\lesssim (T_{\max} - t)^{1/2} \|u\|^3_{L_t^\infty L_x^4([t,T_{\max})\times\mathbb{R}^4)} \\
&\lesssim_u (T_{\max} - t)^{1/2},
\end{aligned}$$

uniformly in $N \in 2^{\mathbb{Z}}$. Letting $N \to 0$ we deduce that u has finite mass; letting $t \to T_{\max}$ and invoking the conservation of mass, we deduce that

$$M(u(t)) = \int_{\mathbb{R}^4} |u(t,x)|^2 \, dx = 0 \quad \text{for all} \quad t \in [0, T_{\max}).$$

In particular, $u \equiv 0$, which contradicts the fact that $S_{[0,T_{\max})}(u) = \infty$.

This completes the proof of the theorem. $\square$

Chapter 9

Long-time Strichartz estimates and applications

In this section, we prove a long-time Strichartz inequality for solutions to (8.15) as described in Theorem 8.10. This will then be used to rule out rapid frequency cascade solutions, namely, solutions which also satisfy

$$\int_0^{T_{\max}} N(t)^{-1}\,dt < \infty.$$

9.1 A long-time Strichartz inequality

Long-time Strichartz inequalities originate in the work of Dodson [15] on the mass-critical NLS. The main result of this section is a long-time Strichartz estimate for solutions to (8.15). This was proved in [40]; we review the proof below.

Theorem 9.1 (Long-time Strichartz estimates)**.** *Let $u : [0, T_{\max}) \times \mathbb{R}^4 \to \mathbb{C}$ be an almost periodic solution to (8.15) with $N(t) \equiv N_k \geq 1$ on each characteristic interval $J_k \subset [0, T_{\max})$. Then, on any compact time interval $I \subset [0, T_{\max})$, which is a union of contiguous intervals J_k, and for any frequency $M > 0$,*

$$\|\nabla u_{\leq M}\|_{L_t^2 L_x^4(I\times\mathbb{R}^4)} \lesssim_u 1 + M^{3/2}K^{1/2}, \tag{9.1}$$

where $K := \int_I N(t)^{-1}\,dt$. Moreover, for any $\eta > 0$ there exists $M_0 = M_0(\eta) > 0$ such that for all $M \leq M_0$,

$$\|\nabla u_{\leq M}\|_{L_t^2 L_x^4(I\times\mathbb{R}^4)} \lesssim_u \eta\big(1 + M^{3/2}K^{1/2}\big). \tag{9.2}$$

Importantly, the constant M_0 and the implicit constants in (9.1) and (9.2) are independent of the interval I.

Proof. Fix a compact time interval $I \subset [0, T_{\max})$, which is a union of contiguous intervals J_k. Throughout the proof all space-time norms will be on $I \times \mathbb{R}^4$, unless

H. Koch et al., *Dispersive Equations and Nonlinear Waves: Generalized Korteweg–de Vries, Nonlinear Schrödinger, Wave and Schrödinger Maps*, Oberwolfach Seminars 45,
DOI 10.1007/978-3-0348-0736-4_23, © Springer Basel 2014

we specify otherwise. Let $\eta_0 > 0$ be a small parameter to be chosen later. By almost periodicity, there exists $c_0 = c_0(\eta_0)$ such that

$$\|\nabla u_{\leq c_0 N(t)}\|_{L_t^\infty L_x^2} \leq \eta_0. \tag{9.3}$$

For $M > 0$ we define

$$A(M) := \|\nabla u_{\leq M}\|_{L_t^2 L_x^4(I\times\mathbb{R}^4)}.$$

Note that Corollary 8.5 implies

$$A(M) \lesssim_u 1 + M^{3/2}K^{1/2} \quad \text{whenever} \quad M \geq \left(\frac{\int_I N(t)^2\,dt}{\int_I N(t)^{-1}\,dt}\right)^{1/3}, \tag{9.4}$$

and, in particular, whenever $M \geq N_{max} := \sup_{t\in I} N(t)$. We will obtain the result for arbitrary frequencies $M > 0$ by induction. Our first step is to obtain a recurrence relation for $A(M)$. We start with an application of the Strichartz inequality:

$$A(M) \lesssim \inf_{t\in I} \|\nabla u_{\leq M}(t)\|_{L_x^2} + \big\|\nabla P_{\leq M} F(u)\big\|_{L_t^2 L_x^{4/3}}. \tag{9.5}$$

To continue, we decompose $u = u_{\leq M/\eta_0} + u_{>M/\eta_0}$ and then further decompose $u(t) = u_{\leq c_0 N(t)}(t) + u_{>c_0 N(t)}(t)$. Thus we may write

$$F(u) = \emptyset\big(u_{>M/\eta_0}u^2\big) + \emptyset\big(\big(P_{\leq c_0 N(t)}u_{\leq M/\eta_0}\big)^3\big) + \emptyset\big(u^2_{\leq M/\eta_0}u_{>c_0 N(t)}\big), \tag{9.6}$$

where we use the notation $\emptyset(X)$ to denote a quantity that resembles X, that is, a finite linear combination of terms that look like those in X, but possibly with some factors replaced by their complex conjugates and/or restricted to various frequencies. Next, we will estimate the contributions of each of these terms to (9.5).

To estimate the contribution of the first term on the right-hand side of (9.6), we use the Bernstein inequality followed by Lemma 11.9, Lemma 11.6, Hölder, and Sobolev embedding:

$$\begin{aligned}
&\big\|\nabla P_{\leq M}\emptyset\big(u_{>M/\eta_0}u^2\big)\big\|_{L_t^2 L_x^{4/3}} && (9.7)\\
&\lesssim M^{5/3}\big\||\nabla|^{-2/3}\emptyset\big(u_{>M/\eta_0}u^2\big)\big\|_{L_t^2 L_x^{4/3}}\\
&\lesssim M^{5/3}\big\||\nabla|^{-2/3}u_{>M/\eta_0}\big\|_{L_t^2 L_x^4}\big\||\nabla|^{2/3}\emptyset\big(u^2\big)\big\|_{L_t^\infty L_x^{3/2}}\\
&\lesssim M^{5/3}\big\||\nabla|^{-2/3}u_{>M/\eta_0}\big\|_{L_t^2 L_x^4}\big\||\nabla|^{2/3}u\big\|_{L_t^\infty L_x^{12/5}}\|u\|_{L_t^\infty L_x^4}\\
&\lesssim M^{5/3}\big\||\nabla|^{-2/3}u_{>M/\eta_0}\big\|_{L_t^2 L_x^4}\|u\|^2_{L_t^\infty \dot H_x^1}\\
&\lesssim_u \sum_{L>M/\eta_0}\left(\frac{M}{L}\right)^{5/3}A(L). && (9.8)
\end{aligned}$$

We turn now to the contribution of the second term on the right-hand side of (9.6). Employing Hölder and (9.3), we obtain

$$\begin{aligned}\big\|\nabla P_{\leq M}\emptyset\big(\big(P_{\leq c_0 N(t)}u_{\leq M/\eta_0}\big)^3\big)\big\|_{L_t^2 L_x^{4/3}} &\lesssim \|\nabla u_{\leq M/\eta_0}\|_{L_t^2 L_x^4}\|u_{\leq c_0 N(t)}\|_{L_t^\infty L_x^4}^2\\ &\lesssim_u \eta_0^2 A(M/\eta_0). \end{aligned}\tag{9.9}$$

Finally, we consider the contribution of the third term on the right-hand side of (9.6). By Bernstein and then Hölder,

$$\begin{aligned}\big\|\nabla P_{\leq M}\emptyset\big(u_{\leq M/\eta_0}^2 u_{>c_0 N(t)}\big)\big\|_{L_t^2 L_x^{4/3}} &\lesssim M\|u_{\leq M/\eta_0}\|_{L_t^\infty L_x^4}\|u_{\leq M/\eta_0}u_{>c_0 N(t)}\|_{L_{t,x}^2}\\ &\lesssim_u M\|u_{\leq M/\eta_0}u_{>c_0 N(t)}\|_{L_{t,x}^2}.\end{aligned}$$

To continue, we decompose the time interval I into intervals of local constancy J_k and apply the bilinear Strichartz estimate Corollary 2.10 on each J_k. Note that by Lemma 8.5, Corollary 8.6, and Hölder's inequality, on each J_k we have

$$\|\nabla u\|_{L_t^2 L_x^4(J_k\times\mathbb{R}^4)} + \|\nabla F(u)\|_{L_{t,x}^{3/2}(J_k\times\mathbb{R}^4)} \lesssim_u 1 \quad \text{and hence} \quad \|\nabla u\|_{S_0^*(J_k)} \lesssim_u 1.$$

Thus, using also Bernstein's inequality,

$$\begin{aligned}\big\|u_{\leq M/\eta_0}u_{>c_0 N(t)}\big\|_{L_{t,x}^2(J_k\times\mathbb{R}^4)} &\lesssim \frac{(M/\eta_0)^{1/2}}{(c_0 N_k)^{1/2}}\|\nabla u_{\leq M/\eta_0}\|_{S_0^*(J_k)}\|u_{>c_0 N_k}\|_{S_0^*(J_k)}\\ &\lesssim_u \frac{M^{1/2}}{\eta_0^{1/2}c_0^{3/2}N_k^{3/2}}\|\nabla u_{\leq M/\eta_0}\|_{S_0^*(J_k)}.\end{aligned}$$

The term $\|\nabla u_{\leq M/\eta_0}\|_{S_0^*(J_k)}$ will be essential in obtaining the small parameter η in claim (9.2) and this is why we choose to keep it in the display above rather than discarding it. Summing the estimates above over the intervals J_k and invoking again the local constancy property Lemma 8.3, we find

$$\begin{aligned}\big\|u_{\leq M/\eta_0}u_{>c_0 N(t)}\big\|_{L_{t,x}^2(I\times\mathbb{R}^4)} &\lesssim_u \frac{M^{1/2}}{\eta_0^{1/2}c_0^{3/2}}\Big(\sum_{J_k\subset I}\frac{1}{N_k^3}\Big)^{1/2}\sup_{J_k\subset I}\|\nabla u_{\leq M/\eta_0}\|_{S_0^*(J_k)}\\ &\lesssim_u \frac{M^{1/2}K^{1/2}}{\eta_0^{1/2}c_0^{3/2}}\sup_{J_k\subset I}\|\nabla u_{\leq M/\eta_0}\|_{S_0^*(J_k)}.\end{aligned}$$

Thus, the contribution of the third term on the right-hand side of (9.6) can be bounded as follows:

$$\big\|\nabla P_{\leq M}\emptyset\big(u_{\leq M/\eta_0}^2 u_{>c_0 N(t)}\big)\big\|_{L_t^2 L_x^{4/3}} \lesssim_u \frac{M^{3/2}K^{1/2}}{\eta_0^{1/2}c_0^{3/2}}\sup_{J_k\subset I}\|\nabla u_{\leq M/\eta_0}\|_{S_0^*(J_k)}. \tag{9.10}$$

Collecting (9.5) through (9.10), we obtain

$$A(M) \lesssim_u \inf_{t\in I} \|\nabla u_{\le M}(t)\|_{L_x^2} + \frac{M^{3/2}K^{1/2}}{\eta_0^{1/2}c_0^{3/2}} \sup_{J_k\subset I} \|\nabla u_{\le M/\eta_0}\|_{S_0^*(J_k)} + \sum_{L\ge \frac{M}{\eta_0}} \Big(\frac{M}{L}\Big)^{5/3} A(L). \tag{9.11}$$

The inductive step in the proof of claims (9.1) and (9.2) will rely on this recurrence relation.

Let us first address (9.1). Recall that by (9.4), the claim holds for $M \ge N_{max}$, that is,

$$A(M) \le C(u)\big[1 + M^{3/2}K^{1/2}\big], \tag{9.12}$$

for some constant $C(u) > 0$ and all $M \ge N_{max}$. Now using the fact that (9.11) implies

$$A(M) \le \tilde C(u)\Big\{1 + \frac{M^{3/2}K^{1/2}}{\eta_0^{1/2}c_0^{3/2}} + \sum_{L\ge\frac{M}{\eta_0}} \Big(\frac{M}{L}\Big)^{5/3} A(L)\Big\}, \tag{9.13}$$

we can inductively prove the claim by halving the frequency M at each step. For example, assuming that (9.12) holds for frequencies larger or equal to M, an application of (9.13) (with $\eta_0 \le 1/2$) yields

$$\begin{aligned} A\big(M/2\big) &\le \tilde C(u)\Big\{1 + \frac{(M/2)^{3/2}K^{1/2}}{\eta_0^{1/2}c_0^{3/2}} + C(u)\sum_{L\ge\frac{M}{2\eta_0}} \Big(\frac{M}{2L}\Big)^{5/3}\big[1 + L^{3/2}K^{1/2}\big]\Big\} \\ &\le \tilde C(u)\Big\{1 + \frac{(M/2)^{3/2}K^{1/2}}{\eta_0^{1/2}c_0^{3/2}} + 2\eta_0^{5/3}C(u) + 2\eta_0^{1/6}C(u)(M/2)^{3/2}K^{1/2}\Big\}. \end{aligned}$$

Choosing $\eta_0 = \eta_0(u)$ small enough so that $\eta_0^{1/6}\tilde C(u) \le 1/4$, we thus obtain

$$A\big(M/2\big) \le \frac12 C(u)\Big\{1 + (M/2)^{3/2}K^{1/2}\Big\} + \tilde C(u)\Big\{1 + \frac{(M/2)^{3/2}K^{1/2}}{\eta_0^{1/2}c_0^{3/2}}\Big\}.$$

The claim now follows by setting $C(u) \ge 2\tilde C(u)\eta_0^{-1/2}c_0^{-3/2}$.

Next we turn to (9.2). To exhibit the small constant η, we will need the following

Lemma 9.2 (Vanishing of the small frequencies)**.** *Under the assumptions of Theorem 9.1, we have*

$$f(M) := \|\nabla u_{\le M}\|_{L_t^\infty L_x^2([0,T_{\max})\times\mathbb{R}^4)} + \sup_{J_k\subset[0,T_{\max})} \|\nabla u_{\le M}\|_{S_0^*(J_k)} \to 0 \quad \textit{as } M \to 0.$$

Proof. As by hypothesis $\inf_{t\in[0,T_{\max})} N(t) \geq 1$, almost periodicity yields

$$\lim_{M\to 0} \|\nabla u_{\leq M}\|_{L_t^\infty L_x^2([0,T_{\max})\times\mathbb{R}^4)} = 0. \tag{9.14}$$

Now fix a characteristic interval $J_k \subset [0, T_{\max})$ and recall that all Strichartz norms of u are bounded on J_k. In particular, we have

$$\|\nabla u\|_{L_t^2 L_x^4(J_k\times\mathbb{R}^4)} + \|u\|_{L_t^3 L_x^{12}(J_k\times\mathbb{R}^4)} + \|u\|_{L_{t,x}^6(J_k\times\mathbb{R}^4)} \lesssim_u 1.$$

Using this followed by the decomposition $u = u_{\leq M^{1/2}} + u_{>M^{1/2}}$, Hölder, and Bernstein, for any frequency $M > 0$ we estimate

$$\begin{aligned}
&\|\nabla u_{\leq M}\|_{S_0^*(J_k)} \\
&= \|\nabla u_{\leq M}\|_{L_t^\infty L_x^2} + \|\nabla P_{\leq M} F(u)\|_{L_{t,x}^{3/2}} \\
&\lesssim \|\nabla u_{\leq M}\|_{L_t^\infty L_x^2} + \|\nabla P_{\leq M} F(u_{>M^{1/2}})\|_{L_{t,x}^{3/2}} + \|\nabla u_{>M^{1/2}} u_{\leq M^{1/2}} u\|_{L_{t,x}^{3/2}} \\
&\quad + \|\nabla u_{\leq M^{1/2}} u^2\|_{L_{t,x}^{3/2}} \\
&\lesssim \|\nabla u_{\leq M}\|_{L_t^\infty L_x^2} + M \|u_{>M^{1/2}}\|_{L_t^2 L_x^4} \|u_{>M^{1/2}}\|_{L_{t,x}^6} \|u_{>M^{1/2}}\|_{L_t^\infty L_x^4} \\
&\quad + \|\nabla u_{>M^{1/2}}\|_{L_t^2 L_x^4} \|u_{\leq M^{1/2}}\|_{L_t^\infty L_x^4} \|u\|_{L_{t,x}^6} + \|\nabla u_{\leq M^{1/2}}\|_{L_t^\infty L_x^2} \|u\|^2_{L_t^3 L_x^{12}} \\
&\lesssim_u \|\nabla u_{\leq M}\|_{L_t^\infty L_x^2} + M^{1/2} + \|\nabla u_{\leq M^{1/2}}\|_{L_t^\infty L_x^2}.
\end{aligned}$$

All space-time norms in the estimates above are on $J_k \times \mathbb{R}^4$. As $J_k \subset [0, T_{\max})$ was arbitrary, we find

$$\begin{aligned}
\sup_{J_k\subset[0,T_{\max})} \|\nabla u_{\leq M}\|_{S_0^*(J_k)} \lesssim_u M^{1/2} &+ \|\nabla u_{\leq M}\|_{L_t^\infty L_x^2([0,T_{\max})\times\mathbb{R}^4)} \\
&+ \|\nabla u_{\leq M^{1/2}}\|_{L_t^\infty L_x^2([0,T_{\max})\times\mathbb{R}^4)}.
\end{aligned}$$

The claim now follows by combining this with (9.14). □

We are now ready to prove (9.2). Using (9.1) and Lemma 9.2, the estimate (9.11) implies

$$\begin{aligned}
A(M) &\lesssim_u f(M) + \frac{M^{3/2}K^{1/2}}{\eta_0^{1/2} c_0^{3/2}} f(M) + \sum_{L\geq \frac{M}{\eta_0}} \Big(\frac{M}{L}\Big)^{5/3} A(L) \\
&\lesssim_u f(M) + \eta_0^{5/3} + \left\{ \frac{f(M)}{\eta_0^{1/2} c_0^{3/2}} + \eta_0^{1/6} \right\} M^{3/2} K^{1/2}.
\end{aligned}$$

Thus, for any $\eta > 0$, choosing first $\eta_0 = \eta_0(\eta)$ such that $\eta_0^{1/6} \leq \eta$ and then $M_0 = M_0(\eta)$ such that $\frac{f(M_0)}{\eta_0^{1/2} c_0^{3/2}} \leq \eta$, we obtain

$$A(M) \lesssim_u \eta\big(1 + M^{3/2}K^{1/2}\big) \quad \text{for all} \quad M \leq M_0.$$

This completes the proof of Theorem 9.1. □

Next, we record a consequence of Theorem 9.1, which will be useful in the derivation of a frequency-localized interaction Morawetz inequality.

Corollary 9.3 (Low and high frequencies control). *Let $u : [0, T_{\max}) \times \mathbb{R}^4 \to \mathbb{C}$ be an almost periodic solution to (8.15) with $N(t) \equiv N_k \geq 1$ on each characteristic interval $J_k \subset [0, T_{\max})$. Then, on any compact time interval $I \subset [0, T_{\max})$, which is a union of contiguous intervals J_k, and for any frequency $M > 0$,*

$$\|u_{\geq M}\|_{L_t^q L_x^r(I\times\mathbb{R}^4)} \lesssim_u M^{-1}(1 + M^3 K)^{\frac{1}{q}} \quad \textit{for all } \tfrac{1}{q} + \tfrac{2}{r} = 1 \textit{ with } 3 < q \leq \infty. \tag{9.15}$$

Moreover, for any $\eta > 0$ there exists $M_0 = M_0(\eta)$ such that for all $M \leq M_0$ we have

$$\|\nabla u_{\leq M}\|_{L_t^q L_x^r(I\times\mathbb{R}^4)} \lesssim_u \eta(1 + M^3 K)^{\frac{1}{q}} \quad \textit{for all } \tfrac{1}{q} + \tfrac{2}{r} = 1 \textit{ with } 2 \leq q \leq \infty. \tag{9.16}$$

The constant M_0 and the implicit constants in (9.15) and (9.16) are independent of the interval I.

Proof. We first address (9.15). By (9.1) and Bernstein's inequality, for any $\varepsilon > 0$ and any frequency $M > 0$ we have

$$\begin{aligned}
\big\||\nabla|^{-1/2-\varepsilon} u_{\geq M}\big\|_{L_t^2 L_x^4(I\times\mathbb{R}^4)} &\lesssim \sum_{L \geq M} L^{-3/2-\varepsilon} \|\nabla u_L\|_{L_t^2 L_x^4(I\times\mathbb{R}^4)} \\
&\lesssim_u \sum_{L\geq M} L^{-3/2-\varepsilon}(1 + L^{3/2} K^{1/2}) \\
&\lesssim_u M^{-3/2-\varepsilon}(1 + M^3 K)^{1/2}.
\end{aligned}$$

The claim now follows by interpolating with the energy bound:

$$\begin{aligned}
\|u_{\geq M}\|_{L_t^q L_x^r(I\times\mathbb{R}^4)} &\lesssim \big\||\nabla|^{-\frac{1}{2}-\frac{q-3}{2}} u_{\geq M}\big\|^{2/q}_{L_t^2 L_x^4(I\times\mathbb{R}^4)} \|\nabla u_{\geq M}\|^{1-2/q}_{L_t^\infty L_x^2(I\times\mathbb{R}^4)} \\
&\lesssim_u M^{-1}(1 + M^3 K)^{1/q},
\end{aligned}$$

whenever $\frac{1}{q} + \frac{2}{r} = 1$ and $3 < q \leq \infty$.

We turn now to (9.16). As $\inf_{t\in I} N(t) \geq 1$, using almost periodicity, for any $\eta > 0$ we can find $M_0(\eta)$ such that for all $M \leq M_0$,

$$\|\nabla u_{\leq M}\|_{L_t^\infty L_x^2(I\times\mathbb{R}^4)} \leq \eta.$$

The claim follows by interpolating with (9.2). □

9.2 The rapid frequency cascade scenario

In this subsection, we preclude the existence of almost periodic solutions as in Theorem 8.10 for which $\int_0^{T_{\max}} N(t)^{-1}\,dt < \infty$. We will show their existence is inconsistent with the conservation of mass.

Theorem 9.4 (No rapid frequency cascades)**.** *There are no almost periodic solutions $u : [0, T_{\max}) \times \mathbb{R}^4 \to \mathbb{C}$ to (8.15) with $N(t) \equiv N_k \geq 1$ on each characteristic interval $J_k \subset [0, T_{\max})$ such that $\|u\|_{L^6_{t,x}([0,T_{\max})\times\mathbb{R}^4)} = +\infty$ and*

$$\int_0^{T_{\max}} N(t)^{-1}\,dt < \infty. \tag{9.17}$$

Proof. We argue by contradiction. Let u be such a solution. Then, by Corollary 8.4,

$$\lim_{t\to T_{\max}} N(t) = \infty, \tag{9.18}$$

whether $T_{\max}$ is finite or infinite. Thus, by almost periodicity we have

$$\lim_{t\to T_{\max}} \|\nabla u_{\leq M}(t)\|_{L^2_x} = 0 \quad \text{for any} \quad M > 0. \tag{9.19}$$

Now let I_n be a nested sequence of compact subintervals of $[0, T_{\max})$ that are unions of contiguous characteristic intervals J_k. On each I_n we may now apply Theorem 9.1. Specifically, using (9.11) together with the hypothesis (9.17), we get

$$\begin{aligned} A_n(M) &:= \|\nabla u_{\leq M}\|_{L^2_t L^4_x(I_n\times\mathbb{R}^4)} \\ &\lesssim_u \inf_{t\in I_n} \|\nabla u_{\leq M}(t)\|_{L^2_x} + \frac{M^{3/2}}{\eta_0^{1/2} c_0^{3/2}} \Big[\int_0^{T_{\max}} N(t)^{-1}\,dt\Big]^{1/2} + \sum_{L\geq \frac{M}{\eta_0}} \Big(\frac{M}{L}\Big)^{5/3} A_n(L) \\ &\lesssim_u \inf_{t\in I_n} \|\nabla u_{\leq M}(t)\|_{L^2_x} + \frac{M^{3/2}}{\eta_0^{1/2} c_0^{3/2}} + \sum_{L\geq \frac{M}{\eta_0}} \Big(\frac{M}{L}\Big)^{5/3} A_n(L) \end{aligned}$$

for all frequencies $M > 0$. Arguing as for (9.1), we find

$$\|\nabla u_{\leq M}\|_{L^2_t L^4_x(I_n\times\mathbb{R}^4)} \lesssim_u \inf_{t\in I_n} \|\nabla u_{\leq M}(t)\|_{L^2_x} + M^{3/2} \quad \text{for all} \quad M > 0.$$

Letting n tend to infinity and invoking (9.19), we obtain

$$\|\nabla u_{\leq M}\|_{L^2_t L^4_x([0,T_{\max})\times\mathbb{R}^4)} \lesssim_u M^{3/2} \quad \text{for all} \quad M > 0. \tag{9.20}$$

Our next claim is that (9.20) implies

$$\|\nabla u_{\leq M}\|_{L^\infty_t L^2_x([0,T_{\max})\times\mathbb{R}^4)} \lesssim_u M^{3/2} \quad \text{for all} \quad M > 0. \tag{9.21}$$

Fix $M > 0$. Using the Duhamel formula from Proposition 8.7 together with the Strichartz inequality, the decomposition $u = u_{\leq M} + u_{>M}$, Lemma 11.9, Lemma 11.6, (9.20), Bernstein, Hölder, and Sobolev embedding, we find that

$$\begin{aligned}
\|\nabla u_{\leq M}\|_{L_t^\infty L_x^2} &\lesssim \|\nabla P_{\leq M} F(u)\|_{L_t^2 L_x^{4/3}} \\
&\lesssim \|\nabla P_{\leq M} F(u_{\leq M})\|_{L_t^2 L_x^{4/3}} + \|\nabla P_{\leq M} \text{Ø}(u_{>M} u^2)\|_{L_t^2 L_x^{4/3}} \\
&\lesssim \|\nabla u_{\leq M}\|_{L_t^2 L_x^4} \|u_{\leq M}\|_{L_t^\infty L_x^4}^2 + M^{5/3} \big\||\nabla|^{-2/3} \text{Ø}(u_{>M} u^2)\big\|_{L_t^2 L_x^{4/3}} \\
&\lesssim_u M^{3/2} + M^{5/3} \big\||\nabla|^{-2/3} u_{>M}\big\|_{L_t^2 L_x^4} \big\||\nabla|^{2/3} u\big\|_{L_t^\infty L_x^{12/5}} \|u\|_{L_t^\infty L_x^4} \\
&\lesssim_u M^{3/2} + M^{5/3} \sum_{L>M} L^{-5/3} \|\nabla u_L\|_{L_t^2 L_x^4} \\
&\lesssim_u M^{3/2} + M^{5/3} \sum_{L>M} L^{-1/6} \\
&\lesssim_u M^{3/2}.
\end{aligned}$$

All space-time norms in the estimates above are on $[0, T_{\max}) \times \mathbb{R}^4$.

With (9.21) in place, we are now ready to finish the proof of Theorem 9.4. First note that by Bernstein's inequality and (9.21), $u \in L_t^\infty \dot{H}_x^{-1/4}([0, T_{\max}) \times \mathbb{R}^4)$; indeed,

$$\begin{aligned}
\big\||\nabla|^{-1/4} u\|_{L_t^\infty L_x^2} &\lesssim \big\||\nabla|^{-1/4} u_{>1}\|_{L_t^\infty L_x^2} + \big\||\nabla|^{-1/4} u_{\leq 1}\|_{L_t^\infty L_x^2} \\
&\lesssim_u \sum_{M>1} M^{-5/4} + \sum_{M\leq 1} M^{1/4} \lesssim_u 1.
\end{aligned}$$

Now fix $t \in [0, T_{\max})$ and let $\eta > 0$ be a small constant. By almost periodicity, there exists $c(\eta) > 0$ such that

$$\int_{|\xi| \leq c(\eta) N(t)} |\xi|^2 |\hat{u}(t,\xi)|^2 \, d\xi \leq \eta.$$

Interpolating with $u \in L_t^\infty \dot{H}_x^{-1/4}$, we find

$$\int_{|\xi| \leq c(\eta) N(t)} |\hat{u}(t,\xi)|^2 \, d\xi \lesssim_u \eta^{1/5}. \tag{9.22}$$

Meanwhile, by elementary considerations,

$$\int_{|\xi| \geq c(\eta) N(t)} |\hat{u}(t,\xi)|^2 \, d\xi \leq [c(\eta) N(t)]^{-2} \int_{\mathbb{R}^4} |\xi|^2 |\hat{u}(t,\xi)|^2 \, d\xi \lesssim_u [c(\eta) N(t)]^{-2}. \tag{9.23}$$

Collecting (9.22) and (9.23) and using Plancherel's theorem, we obtain

$$0 \leq M(u(t)) := \int_{\mathbb{R}^4} |u(t,x)|^2 \, dx \lesssim_u \eta^{1/5} + c(\eta)^{-2} N(t)^{-2}$$

for all $t \in [0, T_{\max})$. Letting η tend to zero and invoking (9.18) and the conservation of mass, we conclude $M(u) = 0$ and hence u is identically zero. This contradicts $\|u\|_{L^6_{t,x}([0,T_{\max})\times\mathbb{R}^4)} = \infty$, thus settling Theorem 9.4. $\square$

Chapter 10

Frequency-localized interaction Morawetz inequalities and applications

Our goal in this section is to prove a frequency-localized interaction Morawetz inequality. This will then be used to preclude the existence of almost periodic solutions as in Theorem 8.10 for which $\int_0^{T_{\max}} N(t)^{-1}\,dt = \infty$. These results appear in [40]; we review the proof below.

Before we delve into the gory details, let us pause to assess where we are. In view of Theorems 8.11 and 9.4, the only enemy we are left to face is an almost periodic solution $u : [0,\infty) \times \mathbb{R}^4 \to \mathbb{C}$ to (8.15) with $N(t) \equiv N_k \geq 1$ on each characteristic interval $J_k \subset [0,\infty)$ such that $\|u\|_{L^6_{t,x}([0,\infty)\times\mathbb{R}^4)} = +\infty$ and

$$\int_0^\infty N(t)^{-1}\,dt < \infty.$$

In order to rule out this quasi-soliton solution, we need tools that express the defocusing nature of the equation. These are the various versions of the Morawetz inequality.

The Morawetz inequality originates in classical mechanics: in the presence of a repulsive potential, the quantity $\mathbf{p}(t) \cdot \frac{\mathbf{x}(t)}{|\mathbf{x}(t)|}$ is monotone. Here $\mathbf{p}$ denotes the momentum of the particle and $\mathbf{x}$ denotes its position. The natural quantum mechanical analogue of the quantity $\mathbf{p}(t) \cdot \frac{\mathbf{x}(t)}{|\mathbf{x}(t)|}$ is the Morawetz action

$$M(t) := 2\,\mathrm{Im} \int_{\mathbb{R}^4} \overline{u(t,x)} \nabla u(t,x) \cdot \frac{x}{|x|}\,dx,$$

where u is a solution to (8.15). A direct computation shows that

$$\partial_t M(t) \geq 2 \int_{\mathbb{R}^4} \frac{|u(t,x)|^2}{|x|^3}\,dx + 3 \int_{\mathbb{R}^4} \frac{|u(t,x)|^4}{|x|}\,dx.$$

Integrating with respect to time and using Cauchy–Schwarz we derive the Lin–

H. Koch et al., *Dispersive Equations and Nonlinear Waves: Generalized Korteweg–de Vries, Nonlinear Schrödinger, Wave and Schrödinger Maps*, Oberwolfach Seminars 45,
DOI 10.1007/978-3-0348-0736-4_24, © Springer Basel 2014

Strauss Morawetz inequality, [25]:

$$\int_I \int_{\mathbb{R}^4} \frac{|u(t,x)|^4}{|x|}\, dx\, dt \lesssim \|u\|_{L_t^\infty L_x^2(I\times\mathbb{R}^4)} \|u\|_{L_t^\infty \dot{H}_x^1(I\times\mathbb{R}^4)}. \tag{10.1}$$

There are two obvious drawbacks when attempting to use this formula to preclude our final enemy. The first one is that it favours the origin: it basically says that if the solution is in $L_t^\infty H_x^1$, then it cannot spend a lot of time near the spatial origin. Secondly, in order to exploit inequality (10.1), we need the solution to lie in $L_t^\infty H_x^1$. However, even if we only cared about Schwartz solutions, when we apply the concentration compactness argument to exhibit a minimal counterexample to Theorem 6.1, we lose all information about the solution that is not left invariant by the symmetries of the equation; in particular, we are left with a solution that is merely in $L_t^\infty \dot{H}_x^1$.

Bourgain [5] showed us how to resolve the second issue above. His solution was to truncate in space; this is equivalent to throwing away the low frequencies of the solution. (Incidentally, truncating an $L_t^\infty \dot{H}_x^1$ solution to high frequencies places it in $L_t^\infty H_x^1$, although the truncation will no longer be a solution.) In this way, Bourgain obtained the following Morawetz inequality:

$$\int_I \int_{|x|\le A|I|^{1/2}} \frac{|u(t,x)|^4}{|x|}\, dx\, dt \lesssim A|I|^{1/2} \|u\|^2_{L_t^\infty \dot{H}_x^1(I\times\mathbb{R}^4)}. \tag{10.2}$$

Compared with (10.1), it still favours the spatial origin, but at least now we can control the right-hand side.

Let us quickly see how to use (10.2) to complete the proof of Theorem 6.1 for *radial* initial data in dimension $d=4$:

Step 1: We note that by rotation invariance and uniqueness of solutions to (8.15), solutions with radial initial data are radial for all time.

Step 2: Radial almost periodic solutions must concentrate near the spatial origin. Indeed, if $|x(t)| \gg N(t)^{-1}$, then by spherical symmetry there exist a very large number of disjoint balls on which $u(t)$ concentrates a nontrivial portion of its energy. This however contradicts the conservation of energy. Thus we must have $|x(t)| \lesssim N(t)^{-1}$. At this point we may set $x(t)\equiv 0$ by modifying the compactness modulus function accordingly.

Step 3: By Sobolev embedding and almost periodicity, we can find $C(u)>0$ such that

$$\int_{|x|\le C(u)/N(t)} |u(t,x)|^4\, dx \gtrsim_u 1 \quad \text{uniformly for } t\in[0,\infty).$$

Step 4: Using (10.2) and Step 3 above, for any time interval $I\subset[0,\infty)$ which is a contiguous union of intervals of local constancy J_k we obtain

$$\begin{aligned}
|I|^{1/2} &\gtrsim_u \int_I \int_{|x|\leq C(u)|I|^{1/2}} \frac{|u(t,x)|^4}{|x|}\,dx\,dt \\
&\gtrsim_u \sum_{J_k\subset I} \int_{J_k} \int_{|x|\leq C(u)|J_k|^{1/2}} \frac{|u(t,x)|^4}{|x|}\,dx\,dt \\
&\gtrsim_u \sum_{J_k\subset I} \int_{J_k} \int_{|x|\leq C(u)/N(t)} N(t)|u(t,x)|^4\,dx\,dt \\
&\gtrsim_u \sum_{J_k\subset I} \int_{J_k} N(t)\,dt \\
&\gtrsim_u \int_I N(t)\,dt.
\end{aligned}$$

Recalling that $\inf_{t\in[0,\infty)} N(t)\geq 1$, we derive a contradiction by taking the interval $I\subset[0,\infty)$ sufficiently large.

This completes the proof of Theorem 6.1 for radial initial data in dimension $d=4$.

To handle nonradial initial data, Colliander, Keel, Staffilani, Takaoka, and Tao [13] made use of an interaction Morawetz inequality, which they introduced in [12]. (Strictly speaking they treated the case $d=3$. In what follows we consider the $d=4$ analogue; see also [30].) Their idea was to center the Morawetz action not at the origin, but rather where the solution actually lives:

$$M_{\text{interact}}(t) := 2\,\mathrm{Im}\int_{\mathbb{R}^4}\int_{\mathbb{R}^4} \overline{u(t,x)}\nabla u(t,x)\cdot\frac{x-y}{|x-y|}|u(t,y)|^2\,dx\,dy.$$

A computation gives

$$\partial_t M_{\text{interact}}(t) \gtrsim \int_{\mathbb{R}^4}\int_{\mathbb{R}^4} \frac{|u(t,x)|^2|u(t,y)|^2}{|x-y|^3} + \frac{|u(t,x)|^4|u(t,y)|^2}{|x-y|}\,dx\,dy.$$

Thus, by the fundamental theorem of calculus and Cauchy–Schwarz,

$$\begin{aligned}
\int_I\int_{\mathbb{R}^4}\int_{\mathbb{R}^4} &\frac{|u(t,x)|^2|u(t,y)|^2}{|x-y|^3} + \frac{|u(t,x)|^4|u(t,y)|^2}{|x-y|}\,dx\,dy\,dt \\
&\lesssim \|u\|^3_{L_t^\infty L_x^2(I\times\mathbb{R}^4)}\|u\|_{L_t^\infty \dot H_x^1(I\times\mathbb{R}^4)}. \qquad (10.3)
\end{aligned}$$

This interaction Morawetz inequality has an obvious drawback, namely, in order to exploit it we need the solution to belong to $L_t^\infty H_x^1$. However, as noted before, our last enemy belongs merely to $L_t^\infty \dot H_x^1$. Therefore, in order to employ this new monotonicity formula, Colliander, Keel, Staffilani, Takaoka, and Tao truncated the solution to frequencies greater than some frequency $N\in 2^{\mathbb{Z}}$, which is chosen small enough so that the truncation captures most of the norm of the solution. By almost periodicity, it is possible to chose N independent of time, since our enemy satisfies $\inf_{t\in[0,\infty)} N(t) \geq 1$. Of course, since $u_{\geq N}$ no longer solves (8.15), there are additional errors introduced on the right-hand side of (10.3). Schematically, we obtain something of the form

$$\begin{aligned}\int_I \int_{\mathbb{R}^4} \int_{\mathbb{R}^4} \frac{|u_{\geq N}(t,x)|^2 |u_{\geq N}(t,y)|^2}{|x-y|^3}\, dx\, dy\, dt \\ &\lesssim \|u_{\geq N}\|^3_{L_t^\infty L_x^2(I\times\mathbb{R}^4)} \|u_{\geq N}\|_{L_t^\infty \dot H_x^1(I\times\mathbb{R}^4)} + \text{errors} \\ &\lesssim_u N^{-3} + \text{errors}. \end{aligned} \tag{10.4}$$

If these errors were magically zero, then it would be a relatively easy task to use (10.4) to rule out our last enemy; see Theorem 10.3 below. However, these errors are not zero and controlling them is highly nontrivial.

Nowadays, there are two ways of handling the error terms on the right-hand side of (10.4). Colliander, Keel, Staffilani, Takaoka, and Tao estimate these errors using solely the left-hand side in (10.4). The smallness needed to close the resulting bootstrap comes from the fact that $u_{\geq N}$ captures most of the norm of the solution and so $\|u_{\leq N}\|_{L_t^\infty \dot H_x^1} \ll 1$. A second approach, inspired by Dodson's work on the mass-critical NLS, is to first obtain additional a priori control in the form of the long-time Strichartz inequality we derived in Section 9; this is then used to control error terms in (10.4). It is this second approach that we will discuss here following [40]. This approach has also been adapted to the three dimensional problem originally treated by Colliander, Keel, Staffilani, Takaoka, and Tao [13] in [21].

10.1 A frequency-localized interaction Morawetz inequality

In this subsection we derive a frequency-localized interaction Morawetz inequality, using the Dodson approach to control the error terms. We start by recalling the interaction Morawetz inequality in four spatial dimensions in slightly more generality; for details, see [30]. For a solution $\phi : I \times \mathbb{R}^4 \to \mathbb{C}$ to the equation $i\phi_t + \Delta\phi = \mathcal{N}$, we define the interaction Morawetz action by

$$M_{\text{interact}}(t) := 2\,\mathrm{Im} \int_{\mathbb{R}^4}\int_{\mathbb{R}^4} |\phi(t,y)|^2 \frac{x-y}{|x-y|} \nabla\phi(t,x)\overline{\phi(t,x)}\, dx\, dy.$$

Standard computations show

$$\begin{aligned}\partial_t M_{\text{interact}}(t) \geq 3 \int_{\mathbb{R}^4}\int_{\mathbb{R}^4} \frac{|\phi(t,x)|^2|\phi(t,y)|^2}{|x-y|^3}\, dx\, dy \\ + 4\,\mathrm{Im} \int_{\mathbb{R}^4}\int_{\mathbb{R}^4} \{\mathcal{N},\phi\}_m(t,y) \frac{x-y}{|x-y|} \nabla\phi(t,x)\overline{\phi(t,x)}\, dx\, dy\end{aligned}$$

$$+ 2 \int_{\mathbb{R}^4} \int_{\mathbb{R}^4} |\phi(t,y)|^2 \frac{x-y}{|x-y|} \{\mathcal{N}, \phi\}_p(t,x)\, dx\, dy,$$

where the mass bracket is given by $\{\mathcal{N}, \phi\}_m := \operatorname{Im}(\mathcal{N}\bar{\phi})$ and the momentum bracket is given by $\{\mathcal{N}, \phi\}_p := \operatorname{Re}(\mathcal{N}\nabla\overline{\phi} - \phi\nabla\overline{\mathcal{N}})$. Thus, integrating with respect to time, we obtain

Proposition 10.1 (Interaction Morawetz inequality)**.**

$$\begin{aligned} 3 \int_I \int_{\mathbb{R}^4} \int_{\mathbb{R}^4} & \frac{|\phi(t,x)|^2 |\phi(t,y)|^2}{|x-y|^3}\, dx\, dy\, dt \\ &+ 2 \int_I \int_{\mathbb{R}^4} \int_{\mathbb{R}^4} |\phi(t,y)|^2 \frac{x-y}{|x-y|} \{\mathcal{N}, \phi\}_p(t,x)\, dx\, dy\, dt \\ &\leq 2\|\phi\|^3_{L^\infty_t L^2_x} \|\phi\|_{L^\infty_t \dot{H}^1_x} + 4\|\phi\|_{L^\infty_t L^2_x} \|\phi\|_{L^\infty_t \dot{H}^1_x} \|\{\mathcal{N}, \phi\}_m\|_{L^1_{t,x}}, \end{aligned}$$

where all space-time norms are over $I \times \mathbb{R}^4$.

We will apply Proposition 10.1 with $\phi = u_{\geq M}$ and $\mathcal{N} = P_{\geq M}(|u|^2 u)$ for M small enough so that the Littlewood–Paley projection captures most of the solution. More precisely, we will prove

Proposition 10.2 (Frequency-localized interaction Morawetz estimate, [40])**.** *Let $u : [0, T_{\max}) \times \mathbb{R}^4 \to \mathbb{C}$ be an almost periodic solution to (8.15) such that $N(t) \equiv N_k \geq 1$ on each characteristic interval $J_k \subset [0, T_{\max})$. Then for any $\eta > 0$ there exists $M_0 = M_0(\eta)$ such that for $M \leq M_0$ and any compact time interval $I \subset [0, T_{\max})$, which is a union of contiguous intervals J_k, we have*

$$\int_I \int_{\mathbb{R}^4} \int_{\mathbb{R}^4} \frac{|u_{\geq M}(t,x)|^2 |u_{\geq M}(t,y)|^2}{|x-y|^3}\, dx\, dy\, dt \lesssim_u \eta \Big[M^{-3} + \int_I N(t)^{-1}\, dt \Big].$$

The implicit constant does not depend on the interval I.

Proof. Fix a compact interval $I \subset [0, T_{\max})$, which is a union of contiguous intervals J_k, and let $K := \int_I N(t)^{-1}\, dt$. Throughout the proof, all space-time norms will be on $I \times \mathbb{R}^4$.

Fix $\eta > 0$ and let $M_0 = M_0(\eta)$ be small enough so that claim (9.16) of Corollary 9.3 holds; more precisely, for all $M \leq M_0$,

$$\|\nabla u_{\leq M}\|_{L^q_t L^r_x} \lesssim_u \eta (1 + M^3 K)^{1/q} \quad \text{for all} \quad \tfrac{1}{q} + \tfrac{2}{r} = 1 \quad \text{with} \quad 2 \leq q \leq \infty. \tag{10.5}$$

Choosing M_0 even smaller if necessary, we can also guarantee that

$$\|u_{\geq M}\|_{L^\infty_t L^2_x} \lesssim_u \eta^6 M^{-1} \quad \text{for all} \quad M \leq M_0. \tag{10.6}$$

Now fix $M \leq M_0$ and write $u_{\text{lo}} := u_{\leq M}$ and $u_{\text{hi}} := u_{>M}$. With this notation, (10.5) becomes

$$\|\nabla u_{\text{lo}}\|_{L^q_t L^r_x} \lesssim_u \eta (1 + M^3 K)^{1/q} \quad \text{for all} \quad \tfrac{1}{q} + \tfrac{2}{r} = 1 \quad \text{with} \quad 2 \leq q \leq \infty. \tag{10.7}$$

We will also need claim (9.15) of Corollary 9.3, which reads

$$\|u_{\mathrm{hi}}\|_{L_t^q L_x^r} \lesssim_u M^{-1}(1+M^3K)^{1/q} \quad \text{for all} \quad \tfrac{1}{q}+\tfrac{2}{r}=1 \quad \text{with} \quad 3<q\leq\infty. \tag{10.8}$$

Note that by (10.6), the endpoint $q=\infty$ of the inequality above is strengthened to

$$\|u_{\mathrm{hi}}\|_{L_t^\infty L_x^2} \lesssim_u \eta^6 M^{-1}. \tag{10.9}$$

To continue, we apply Proposition 10.1 with $\phi = u_{\mathrm{hi}}$ and $\mathcal{N} = P_{\mathrm{hi}}F(u)$ and use (10.9); we obtain

$$\begin{aligned} &\int_I \int_{\mathbb{R}^4}\int_{\mathbb{R}^4} \frac{|u_{\mathrm{hi}}(t,x)|^2|u_{\mathrm{hi}}(t,y)|^2}{|x-y|^3}\,dx\,dy\,dt \\ &\qquad + \int_I \int_{\mathbb{R}^4}\int_{\mathbb{R}^4} |u_{\mathrm{hi}}(t,y)|^2 \frac{x-y}{|x-y|}\{P_{\mathrm{hi}}F(u),u_{\mathrm{hi}}\}_p(t,x)\,dx\,dy\,dt \\ &\qquad\qquad \lesssim_u \eta^{18}M^{-3} + \eta^6 M^{-1}\|\{P_{\mathrm{hi}}F(u),u_{\mathrm{hi}}\}_m\|_{L^1_{t,x}(I\times\mathbb{R}^4)}. \end{aligned} \tag{10.10}$$

We first consider the contribution of the momentum bracket term. We write

$$\begin{aligned} &\{P_{\mathrm{hi}}F(u),u_{\mathrm{hi}}\}_p \\ &\quad = \{F(u),u\}_p - \{F(u_{\mathrm{lo}}),u_{\mathrm{lo}}\}_p - \{F(u)-F(u_{\mathrm{lo}}),u_{\mathrm{lo}}\}_p - \{P_{\mathrm{lo}}F(u),u_{\mathrm{hi}}\}_p \\ &\quad = -\tfrac{1}{2}\nabla[|u|^4 - |u_{\mathrm{lo}}|^4] - \{F(u)-F(u_{\mathrm{lo}}),u_{\mathrm{lo}}\}_p - \{P_{\mathrm{lo}}F(u),u_{\mathrm{hi}}\}_p \\ &\quad =: I + II + III. \end{aligned}$$

After an integration by parts, the term I contributes to the left-hand side of (10.10) a multiple of

$$\begin{aligned} &\int_I \int_{\mathbb{R}^4}\int_{\mathbb{R}^4} \frac{|u_{hi}(t,y)|^2|u_{hi}(t,x)|^4}{|x-y|}\,dx\,dy\,dt \\ &\qquad + \sum_{j=1}^3 \int_I \int_{\mathbb{R}^4}\int_{\mathbb{R}^4} \frac{|u_{\mathrm{hi}}(t,y)|^2|u_{\mathrm{hi}}(t,x)|^j|u_{\mathrm{lo}}(t,x)|^{4-j}}{|x-y|}\,dx\,dy\,dt. \end{aligned}$$

In order to estimate the contribution of II to (10.10), we use $\{f,g\}_p = \nabla\emptyset(fg) + \emptyset(f\nabla g)$ to write

$$\{F(u)-F(u_{\mathrm{lo}}),u_{\mathrm{lo}}\}_p = \sum_{j=1}^3 \nabla\emptyset(u_{hi}^j u_{lo}^{4-j}) + \sum_{j=1}^3 \emptyset(u_{hi}^j u_{lo}^{3-j}\nabla u_{\mathrm{lo}}).$$

Integrating by parts for the first term and bringing absolute values inside the integrals for the second term, we find that II contributes to the right-hand side of (10.10) a multiple of

$$\sum_{j=1}^{3}\int_I\int_{\mathbb{R}^4}\int_{\mathbb{R}^4}\frac{|u_{\mathrm{hi}}(t,y)|^2|u_{\mathrm{hi}}(t,x)|^j|u_{\mathrm{lo}}(t,x)|^{4-j}}{|x-y|}\,dx\,dy\,dt$$
$$+\sum_{j=1}^{3}\int_I\int_{\mathbb{R}^4}\int_{\mathbb{R}^4}|u_{\mathrm{hi}}(t,y)|^2|u_{\mathrm{hi}}(t,x)|^j|\nabla u_{\mathrm{lo}}(t,x)||u_{\mathrm{lo}}(t,x)|^{3-j}\,dx\,dy\,dt.$$

Finally, integrating by parts when the derivative (from the definition of the momentum bracket) falls on u_{hi}, we estimate the contribution of III to the right-hand side of (10.10) by a multiple of

$$\int_I\int_{\mathbb{R}^4}\int_{\mathbb{R}^4}\frac{|u_{\mathrm{hi}}(t,y)|^2|u_{\mathrm{hi}}(t,x)||P_{\mathrm{lo}}F(u(t,x))|}{|x-y|}\,dx\,dy\,dt$$
$$+\int_I\int_{\mathbb{R}^4}\int_{\mathbb{R}^4}|u_{\mathrm{hi}}(t,y)|^2|u_{\mathrm{hi}}(t,x)||\nabla P_{\mathrm{lo}}F(u(t,x))|\,dx\,dy\,dt.$$

Consider now the mass bracket appearing in (10.10). Exploiting cancellation, we write

$$\begin{aligned}&\{P_{\mathrm{hi}}F(u),u_{\mathrm{hi}}\}_m\\ &\quad=\{P_{\mathrm{hi}}F(u)-F(u_{\mathrm{hi}}),u_{\mathrm{hi}}\}_m\\ &\quad=\{P_{\mathrm{hi}}\big[F(u)-F(u_{\mathrm{hi}})-F(u_{\mathrm{lo}})\big],u_{\mathrm{hi}}\}_m+\{P_{\mathrm{hi}}F(u_{\mathrm{lo}}),u_{\mathrm{hi}}\}_m-\{P_{\mathrm{lo}}F(u_{\mathrm{hi}}),u_{\mathrm{hi}}\}_m\\ &\quad=\emptyset(u_{hi}^3u_{\mathrm{lo}})+\emptyset(u_{hi}^2u_{lo}^2)+\{P_{\mathrm{hi}}F(u_{\mathrm{lo}}),u_{\mathrm{hi}}\}_m-\{P_{\mathrm{lo}}F(u_{\mathrm{hi}}),u_{\mathrm{hi}}\}_m.\end{aligned}$$

Putting everything together and using (10.9), (10.10) becomes

$$\int_I\int_{\mathbb{R}^4}\int_{\mathbb{R}^4}\frac{|u_{\mathrm{hi}}(t,x)|^2|u_{\mathrm{hi}}(t,y)|^2}{|x-y|^3}\,dx\,dy\,dt+\int_I\int_{\mathbb{R}^4}\int_{\mathbb{R}^4}\frac{|u_{\mathrm{hi}}(t,x)|^2|u_{\mathrm{hi}}(t,y)|^4}{|x-y|}\,dx\,dy\,dt \tag{10.11}$$

$$\lesssim_u \eta^{18}M^{-3} \tag{10.12}$$
$$+\eta^6M^{-1}\{\|u_{hi}^3u_{\mathrm{lo}}\|_{L^1_{t,x}}+\|u_{hi}^2u_{lo}^2\|_{L^1_{t,x}} + \|u_{\mathrm{hi}}P_{\mathrm{hi}}F(u_{\mathrm{lo}})\|_{L^1_{t,x}}+\|u_{\mathrm{hi}}P_{\mathrm{lo}}F(u_{\mathrm{hi}})\|_{L^1_{t,x}}\} \tag{10.13}$$
$$+\eta^{12}M^{-2}\sum_{j=1}^{3}\|u_{hi}^ju_{lo}^{3-j}\nabla u_{\mathrm{lo}}\|_{L^1_{t,x}}+\eta^{12}M^{-2}\|u_{\mathrm{hi}}\nabla P_{\mathrm{lo}}F(u)\|_{L^1_{t,x}} \tag{10.14}$$
$$+\sum_{j=1}^{3}\int_I\int_{\mathbb{R}^4}\int_{\mathbb{R}^4}\frac{|u_{\mathrm{hi}}(t,y)|^2|u_{\mathrm{hi}}(t,x)|^j|u_{\mathrm{lo}}(t,x)|^{4-j}}{|x-y|}\,dx\,dy\,dt \tag{10.15}$$
$$+\int_I\int_{\mathbb{R}^4}\int_{\mathbb{R}^4}\frac{|u_{\mathrm{hi}}(t,y)|^2|u_{\mathrm{hi}}(t,x)||P_{\mathrm{lo}}F(u(t,x))|}{|x-y|}\,dx\,dy\,dt. \tag{10.16}$$

Thus, to complete the proof of Proposition 10.2 we have to show that the error terms (10.13) through (10.16) are acceptable; clearly, (10.12) is acceptable.

Consider now error term (10.13). Using (10.7), (10.8), and Sobolev embedding, we estimate

$$\begin{aligned}\|u_{hi}^3 u_{\rm lo}\|_{L^1_{t,x}} &\lesssim \|u_{\rm hi}\|_{L_t^\infty L_x^4}\|u_{\rm hi}\|^2_{L_t^{7/2}L_x^{14/5}}\|u_{\rm lo}\|_{L_t^{7/3}L_x^{28}} \lesssim_u \eta M^{-2}(1+M^3K)\\ \|u_{hi}^2 u_{lo}^2\|_{L^1_{t,x}} &\lesssim \|u_{\rm hi}\|^2_{L_t^4 L_x^{8/3}}\|u_{\rm lo}\|^2_{L_t^4 L_x^8} \lesssim_u \eta^2 M^{-2}(1+M^3K).\end{aligned}$$

Using Bernstein's inequality as well, we estimate

$$\begin{aligned}\|u_{\rm hi}P_{\rm hi}F(u_{\rm lo})\|_{L^1_{t,x}} &\lesssim \|u_{\rm hi}\|_{L_t^4 L_x^{8/3}} M^{-1}\|\nabla F(u_{\rm lo})\|_{L_t^{4/3}L_x^{8/5}}\\ &\lesssim_u M^{-2}(1+M^3K)^{1/4}\|\nabla u_{\rm lo}\|_{L_t^2 L_x^4}\|u_{\rm lo}\|^2_{L_t^8 L_x^{16/3}}\\ &\lesssim_u \eta^3 M^{-2}(1+M^3K).\end{aligned}$$

Finally, by Hölder, Bernstein, Sobolev embedding, (10.7) and (10.8),

$$\begin{aligned}\|u_{\rm hi}P_{\rm lo}F(u_{\rm hi})\|_{L^1_{t,x}} &\lesssim \|u_{\rm hi}\|_{L_t^{10/3}L_x^{20/7}} M^{7/5}\|F(u_{\rm hi})\|_{L_t^{10/7}L_x^1}\\ &\lesssim_u M^{2/5}(1+M^3K)^{3/10}\|u_{\rm hi}\|^{7/3}_{L_t^{10/3}L_x^{20/7}}\|u_{\rm hi}\|^{2/3}_{L_t^\infty L_x^{40/11}}\\ &\lesssim_u M^{2/5-7/3}(1+M^3K)\||\nabla|^{9/10}u_{\rm hi}\|^{2/3}_{L_t^\infty L_x^2}\\ &\lesssim_u M^{-2}(1+M^3K).\end{aligned}$$

Collecting the estimates above we find

$$(10.13) \lesssim_u \eta^6 M^{-3}(1+M^3K) \lesssim_u \eta(M^{-3}+K),$$

and thus this error term is acceptable.

Consider next error term (10.14). By (10.7), (10.8), (10.9), Sobolev embedding, and Bernstein,

$$\begin{aligned}\|u_{hi}u_{lo}^2\nabla u_{\rm lo}\|_{L^1_{t,x}} &\lesssim \|\nabla u_{\rm lo}\|_{L_t^2 L_x^4}\|u_{\rm hi}\|_{L_t^\infty L_x^2}\|u_{\rm lo}\|^2_{L_t^4 L_x^8} \lesssim_u \eta^9 M^{-1}(1+M^3K)\\ \|u_{hi}^2 u_{lo}\nabla u_{\rm lo}\|_{L^1_{t,x}} &\lesssim \|\nabla u_{\rm lo}\|_{L_t^2 L_x^4}\|u_{\rm hi}\|^2_{L_t^4 L_x^{8/3}}\|u_{\rm lo}\|_{L^\infty_{t,x}} \lesssim_u \eta^2 M^{-1}(1+M^3K)\\ \|u_{hi}^3\nabla u_{\rm lo}\|_{L^1_{t,x}} &\lesssim \|\nabla u_{\rm lo}\|_{L_t^{7/3}L_x^{28}}\|u_{\rm hi}\|^2_{L_t^{7/2}L_x^{14/5}}\|u_{\rm hi}\|_{L_t^\infty L_x^4} \lesssim_u \eta M^{-1}(1+M^3K).\end{aligned}$$

To estimate the second term in (10.14), we write $F(u) = F(u_{\rm lo}) + \emptyset(u_{\rm hi}u_{lo}^2 + u_{hi}^2 u_{\rm lo} + u_{hi}^3)$. Arguing as above, we obtain

$$\begin{aligned}\|u_{\rm hi}\nabla P_{\rm lo}F(u_{\rm lo})\|_{L^1_{t,x}} &\lesssim \|u_{\rm hi}\|_{L_t^\infty L_x^2}\|\nabla u_{\rm lo}\|_{L_t^2 L_x^4}\|u_{\rm lo}\|^2_{L_t^4 L_x^8} \lesssim_u \eta^9 M^{-1}(1+M^3K)\\ \|u_{\rm hi}\nabla P_{\rm lo}\emptyset(u_{\rm hi}u_{lo}^2)\|_{L^1_{t,x}} &\lesssim M\|u_{\rm hi}\|^2_{L_t^4 L_x^{8/3}}\|u_{\rm lo}\|^2_{L_t^4 L_x^8} \lesssim_u \eta^2 M^{-1}(1+M^3K)\\ \|u_{\rm hi}\nabla P_{\rm lo}\emptyset(u_{hi}^2 u_{\rm lo})\|_{L^1_{t,x}} &\lesssim M\|u_{\rm hi}\|_{L_t^\infty L_x^4}\|u_{\rm hi}\|^2_{L_t^{7/2}L_x^{14/5}}\|u_{\rm lo}\|_{L_t^{7/3}L_x^{28}}\\ &\lesssim_u \eta M^{-1}(1+M^3K)\end{aligned}$$

$$\begin{aligned}\|u_{\rm hi}\nabla P_{\rm lo}\emptyset(u_{hi}^3)\|_{L^1_{t,x}} &\lesssim \|u_{\rm hi}\|_{L_t^{10/3}L_x^{20/7}} M^{12/5}\|u_{hi}^3\|_{L_t^{10/7}L_x^1}\\ &\lesssim M^{12/5}\|u_{\rm hi}\|^{10/3}_{L_t^{10/3}L_x^{20/7}}\|u_{\rm hi}\|^{2/3}_{L_t^\infty L_x^{40/11}}\\ &\lesssim_u M^{-1}(1+M^3K).\end{aligned}$$

Putting everything together, we find

$$(10.14) \lesssim_u \eta^{12}M^{-3}(1+M^3K)\lesssim_u \eta(M^{-3}+K),$$

and thus this error term is also acceptable.

We now turn to error term (10.15). By easy considerations, we only have to consider the cases $j=1$ and $j=3$. We start with the case $j=1$; using Hölder together with the Hardy–Littlewood–Sobolev inequality, Sobolev embedding, (10.7), (10.8), and (10.9), we estimate

$$\begin{aligned}\int_I\int_{\mathbb{R}^4}\int_{\mathbb{R}^4} &\frac{|u_{\rm hi}(t,y)|^2|u_{\rm hi}(t,x)||u_{\rm lo}(t,x)|^3}{|x-y|}\,dx\,dy\,dt\\ &\lesssim \|u_{\rm hi}\|^2_{L_t^{12}L_x^{24/11}}\Big\|\tfrac{1}{|x|}*\big(|u_{\rm hi}||u_{\rm lo}|^3\big)\Big\|_{L_t^{6/5}L_x^{12}}\\ &\lesssim_u M^{-2}(1+M^3K)^{1/6}\|u_{\rm hi}u_{lo}^3\|_{L^{6/5}_{t,x}}\\ &\lesssim_u M^{-2}(1+M^3K)^{1/6}\|u_{\rm hi}\|_{L_t^\infty L_x^2}\|u_{\rm lo}\|^3_{L_t^{18/5}L_x^9}\\ &\lesssim_u \eta^9M^{-3}(1+M^3K).\end{aligned}$$

Finally, to estimate the error term corresponding to $j=3$, we consider two scenarios: If $|u_{\rm lo}|\le\delta|u_{\rm hi}|$ for some small $\delta>0$, we absorb this contribution into the term

$$\int_I\int_{\mathbb{R}^4}\int_{\mathbb{R}^4}\frac{|u_{\rm hi}(t,x)|^2|u_{\rm hi}(t,y)|^4}{|x-y|}\,dx\,dy\,dt,$$

which appears in (10.11). If instead $|u_{\rm hi}|\le\delta^{-1}|u_{\rm lo}|$, we may estimate the contribution of this term by that of the error term corresponding to $j=1$. Thus,

$$(10.15)\lesssim_u \eta(M^{-3}+K)+\delta\int_I\int_{\mathbb{R}^4}\int_{\mathbb{R}^4}\frac{|u_{\rm hi}(t,x)|^2|u_{\rm hi}(t,y)|^4}{|x-y|}\,dx\,dy\,dt,$$

where $0<\delta<1$ is a constant small enough so that the second term on the right-hand side above can be absorbed by (10.11). Thus, the error term (10.15) is acceptable.

We are left to consider error term (10.16). Arguing as for the case $j=1$ of the error term (10.15), we derive

$$\begin{aligned}\int_I\int_{\mathbb{R}^4}\int_{\mathbb{R}^4} &\frac{|u_{\rm hi}(t,y)|^2|u_{\rm hi}(t,x)||P_{\rm lo}F(u(t,x))|}{|x-y|}\,dx\,dy\,dt\\ &\lesssim \|u_{\rm hi}\|^2_{L_t^{12}L_x^{24/11}}\Big\|\tfrac{1}{|x|}*\big(|u_{\rm hi}||P_{\rm lo}F(u)|\big)\Big\|_{L_t^{6/5}L_x^{12}}\end{aligned}$$

$$
\begin{aligned}
&\lesssim_u M^{-2}(1+M^3K)^{1/6}\|u_{\text{hi}}P_{\text{lo}}F(u)\|_{L^{6/5}_{t,x}}\\
&\lesssim_u M^{-2}(1+M^3K)^{1/6}\|u_{\text{hi}}\|_{L^4_tL^{8/3}_x}\|P_{\text{lo}}F(u)\|_{L^{12/7}_tL^{24/11}_x}\\
&\lesssim_u M^{-3}(1+M^3K)^{5/12}\|P_{\text{lo}}F(u)\|_{L^{12/7}_tL^{24/11}_x}.
\end{aligned}
$$

We now write $F(u)=F(u_{\text{hi}})+\emptyset(u_{lo}^3+u_{lo}^2u_{\text{hi}}+u_{\text{lo}}u_{hi}^2)$. Using Hölder, Bernstein, Sobolev embedding, (10.7), (10.8), and (10.9), we estimate

$$
\begin{aligned}
\|P_{\text{lo}}\emptyset(u_{lo}^3)\|_{L^{12/7}_tL^{24/11}_x} &\lesssim \|u_{\text{lo}}\|_{L^{12}_tL^{24/5}_x}\|u_{\text{lo}}\|^2_{L^4_tL^8_x}\lesssim_u \eta^3(1+M^3K)^{7/12}\\
\|P_{\text{lo}}\emptyset(u_{lo}^2u_{\text{hi}})\|_{L^{12/7}_tL^{24/11}_x} &\lesssim M\|u_{lo}^2u_{\text{hi}}\|_{L^{12/7}_tL^{24/17}_x}\lesssim M\|u_{\text{lo}}\|^2_{L^4_tL^8_x}\|u_{\text{hi}}\|_{L^{12}_tL^{24/11}_x}\\
&\lesssim_u \eta^2(1+M^3K)^{\frac{7}{12}}\\
\|P_{\text{lo}}\emptyset(u_{\text{lo}}u_{hi}^2)\|_{L^{12/7}_tL^{24/11}_x} &\lesssim M\|u_{\text{lo}}u_{hi}^2\|_{L^{12/7}_tL^{24/17}_x}\\
&\lesssim M\|u_{\text{lo}}\|_{L^3_tL^{12}_x}\|u_{\text{hi}}\|_{L^4_tL^{8/3}_x}\|u_{\text{hi}}\|_{L^\infty_tL^4_x}\\
&\lesssim_u \eta(1+M^3K)^{7/12}.
\end{aligned}
$$

Finally, using Bernstein, Hölder, interpolation, (10.7), (10.8), and (10.9), we get

$$
\begin{aligned}
\|P_{\text{lo}}F(u_{\text{hi}})\|_{L^{12/7}_tL^{24/11}_x} &\lesssim M^{13/6}\|F(u_{\text{hi}})\|_{L^{12/7}_tL^1_x}\\
&\lesssim M^{13/6}\|u_{\text{hi}}\|^2_{L^{24/7}_tL^{48/17}_x}\|u_{\text{hi}}\|_{L^\infty_tL^{24/7}_x}\\
&\lesssim_u M^{1/6}(1+M^3K)^{7/12}\||\nabla|^{5/6}u_{\text{hi}}\|_{L^\infty_tL^2_x}\\
&\lesssim_u \eta(1+M^3K)^{7/12}.
\end{aligned}
$$

Collecting these estimates, we find

$$
(10.16)\lesssim_u \eta M^{-3}(1+M^3K)\lesssim_u \eta(M^{-3}+K),
$$

and thus this last error term is also acceptable.

This completes the proof of Proposition 10.2. □

10.2 The quasi-soliton scenario

With Proposition 10.2 in place, we are now ready to preclude our last enemy, namely, solutions as in Theorem 8.10 for which $\int_0^{T_{\max}} N(t)^{-1}\,dt=\infty$.

Theorem 10.3 (No quasi-solitons). *There exist no almost periodic solutions $u:[0,T_{\max})\times\mathbb{R}^4\to\mathbb{C}$ to (8.15) with $N(t)\equiv N_k\geq 1$ on each characteristic interval $J_k\subset[0,T_{\max})$ which satisfy $\|u\|_{L^6_{t,x}([0,T_{\max})\times\mathbb{R}^4)}=+\infty$ and*

$$
\int_0^{T_{\max}} N(t)^{-1}\,dt=\infty. \tag{10.17}
$$

Proof. We argue by contradiction. Assume there exists such a solution u.

Let $\eta > 0$ be a small parameter to be chosen later. By Proposition 10.2, there exists $M_0 = M_0(\eta)$ such that for all $M \leq M_0$ and any compact time interval $I \subset [0, T_{\max})$, which is a union of contiguous intervals J_k, we have

$$\int_I \int_{\mathbb{R}^4} \int_{\mathbb{R}^4} \frac{|u_{\geq M}(t,x)|^2 |u_{\geq M}(t,y)|^2}{|x-y|^3} \, dx \, dy \, dt \lesssim_u \eta \Big[M^{-3} + \int_I N(t)^{-1} \, dt \Big]. \quad (10.18)$$

As $\inf_{t \in [0, T_{\max})} N(t) \geq 1$, choosing M_0 even smaller if necessary (depending on η) we can also ensure that

$$\|u_{\leq M}\|_{L_t^\infty L_x^4([0,T_{\max})\times\mathbb{R}^4)} + \|u_{\leq M}\|_{L_t^\infty \dot{H}_x^1([0,T_{\max})\times\mathbb{R}^4)} \leq \eta \quad \text{for all} \quad M \leq M_0. \quad (10.19)$$

Exercise 10.1. Use almost periodicity to prove that there exists $C(u) > 0$ such that

$$N(t)^2 \int_{|x-x(t)| \leq C(u)/N(t)} |u(t,x)|^2 \, dx \gtrsim_u 1/C(u) \quad (10.20)$$

uniformly for $t \in [0, T_{\max})$.

Using Hölder's inequality and (10.19), we find that

$$\begin{aligned} \int_{|x-x(t)| \leq C(u)/N(t)} |u_{\leq M}(t,x)|^2 \, dx &\lesssim \Big\{ \frac{C(u)}{N(t)} \|u_{\leq M}\|_{L_t^\infty L_x^4([0,T_{\max})\times\mathbb{R}^4)} \Big\}^2 \\ &\lesssim_u \eta^2 C(u)^2 N(t)^{-2} \end{aligned}$$

for all $t \in [0, T_{\max})$ and all $M \leq M_0$. Combining this with (10.20) and choosing η sufficiently small depending on u, we find that

$$\inf_{t \in [0,T_{\max})} N(t)^2 \int_{|x-x(t)| \leq C(u)/N(t)} |u_{\geq M}(t,x)|^2 \, dx \gtrsim_u 1 \quad \text{for all} \quad M \leq M_0.$$

Thus, on any compact time interval $I \subset [0, T_{\max})$ and for any $M \leq M_0$ we have

$$\begin{aligned} &\int_I \int_{\mathbb{R}^4} \int_{\mathbb{R}^4} \frac{|u_{\geq M}(t,x)|^2 |u_{\geq M}(t,y)|^2}{|x-y|^3} \, dx \, dy \, dt \\ &\geq \int_I \iint_{|x-y| \leq \frac{2C(u)}{N(t)}} \Big[\frac{N(t)}{2C(u)}\Big]^3 |u_{\geq M}(t,x)|^2 |u_{\geq M}(t,y)|^2 \, dx \, dy \, dt \\ &\geq \int_I \Big[\frac{N(t)}{2C(u)}\Big]^3 \int_{|x-x(t)| \leq \frac{C(u)}{N(t)}} |u_{\geq M}(t,x)|^2 \, dx \int_{|y-x(t)| \leq \frac{C(u)}{N(t)}} |u_{\geq M}(t,y)|^2 \, dy \, dt \\ &\gtrsim_u \int_I N(t)^{-1} \, dt. \end{aligned}$$

Invoking (10.18) and choosing η small depending on u, we find

$$\int_I N(t)^{-1}\,dt \lesssim_u M^{-3} \quad \text{for all} \quad M \leq M_0$$

and all intervals $I \subset [0, T_{\max})$, which are unions of contiguous intervals J_k. Recalling the hypothesis (10.17), we derive a contradiction by choosing the interval $I \subset [0, T_{\max})$ sufficiently large.

This completes the proof of the theorem. □

Chapter 11

Appendix A: Background material

11.1 Compactness in L^p

Recall that by the Arzelà–Ascoli theorem, a family of continuous functions on a compact set $K \subset \mathbb{R}^d$ is precompact in $C^0(K)$ if and only if it is uniformly bounded and equicontinuous. The natural generalization to L^p spaces is due to M. Riesz [29] and reads as follows:

Proposition 11.1. *Fix $1 \le p < \infty$. A family of functions $\mathcal{F} \subset L^p(\mathbb{R}^d)$ is precompact in this topology if and only if it obeys the following three conditions:*

(i) *There exists $A > 0$ so that $\|f\|_p \le A$ for all $f \in \mathcal{F}$.*

(ii) *For any $\varepsilon > 0$ there exists $\delta > 0$ so that $\int_{\mathbb{R}^d} |f(x) - f(x+y)|^p \, dx < \varepsilon$ for all $f \in \mathcal{F}$ and all $|y| < \delta$.*

(iii) *For any $\varepsilon > 0$ there exists R so that $\int_{|x| \ge R} |f|^p \, dx < \varepsilon$ for all $f \in \mathcal{F}$.*

Remark. By analogy to the case of continuous functions (or of measures), it is natural to refer to the three conditions as uniform boundedness, equicontinuity, and tightness, respectively.

Proof. If $\mathcal{F}$ is precompact, it may be covered by balls of radius $\frac{1}{2}\varepsilon$ around a finite collection of functions $\{f_j\}$. As any single function obeys (i)–(iii), these properties can be extended to the whole family by approximation by an f_j.

We now turn to sufficiency. Given $\varepsilon > 0$, our job is to show that there are finitely many functions $\{f_j\}$ such that the ε-balls centered at these points cover $\mathcal{F}$. We will find these points via the usual Arzelà–Ascoli theorem, which requires us to approximate $\mathcal{F}$ by a family of continuous functions of compact support. Let $\phi : \mathbb{R}^d \to [0, \infty)$ be a smooth function supported by $\{|x| \le 1\}$ with $\phi(x) = 1$ in a neighbourhood of $x = 0$ and $\int_{\mathbb{R}^d} \phi(x)\, dx = 1$. Given $R > 0$ we define

$$f_R(x) := \phi\big(\tfrac{x}{R}\big) \int_{\mathbb{R}^d} R^d \phi\big(R(x-y)\big) f(y)\, dy$$

H. Koch et al., *Dispersive Equations and Nonlinear Waves: Generalized Korteweg–de Vries, Nonlinear Schrödinger, Wave and Schrödinger Maps*, Oberwolfach Seminars 45,
DOI 10.1007/978-3-0348-0736-4_25,

and write $\mathcal{F}_R := \{f_R : f \in \mathcal{F}\}$. Employing the three conditions, we see that it is possible to choose R so large that $\|f - f_R\|_p < \frac{1}{2}\varepsilon$ for all $f \in \mathcal{F}$. We also see that $\mathcal{F}_R$ is a uniformly bounded family of equicontinuous functions on the compact set $\{|x| \leq R\}$. Thus, $\mathcal{F}_R$ is precompact and we can find a finite family $\{f_j\} \subseteq C^0(\{|x| \leq R\})$ such that $\mathcal{F}_R$ is covered by the L^p-balls of radius $\frac{1}{2}\varepsilon$ around these points. By construction, the ε-balls around these points cover $\mathcal{F}$. □

In the L^2 case it is natural to replace (ii) by a condition on the Fourier transform:

Corollary 11.2. *A family of functions is precompact in $L^2(\mathbb{R}^d)$ if and only if it obeys the following two conditions:*

(i) *There exists $A > 0$ so that $\|f\| \leq A$ for all $f \in \mathcal{F}$.*

(ii) *For all $\varepsilon > 0$ there exists $R > 0$ so that $\int_{|x|\geq R} |f(x)|^2\, dx + \int_{|\xi|\geq R} |\hat{f}(\xi)|^2\, d\xi < \varepsilon$ for all $f \in \mathcal{F}$.*

Proof. Necessity follows as before. Regarding the sufficiency of these conditions, we note that

$$\int_{\mathbb{R}^d} |f(x+y) - f(x)|^2\, dx \sim \int_{\mathbb{R}^d} |e^{i\xi y} - 1|^2 |\hat{f}(\xi)|^2\, d\xi,$$

which allows us to rely on the preceding proposition. □

In our applications, regularity allows us to upgrade weak-$*$ convergence to almost everywhere convergence. The lower semicontinuity of the norm under this notion of convergence is essentially Fatou's lemma. The following quantitative version of this is due to Brezis and Lieb [6] (see also [24, Theorem 1.9]):

Lemma 11.3 (Refined Fatou)**.** *Suppose $\{f_n\} \subseteq L^p_x(\mathbb{R}^d)$ with $\limsup \|f_n\|_p < \infty$. If $f_n \to f$ almost everywhere, then*

$$\int_{\mathbb{R}^d} \Big| |f_n|^p - |f_n - f|^p - |f|^p \Big|\, dx \to 0.$$

In particular, $\|f_n\|_p^p - \|f_n - f\|_p^p \to \|f\|_p^p$.

11.2 Littlewood–Paley theory

Let $\varphi(\xi)$ be a radial bump function supported in the ball $\{\xi \in \mathbb{R}^d : |\xi| \leq \frac{11}{10}\}$ and equal to 1 on the ball $\{\xi \in \mathbb{R}^d : |\xi| \leq 1\}$. For each number $N > 0$, we define the Fourier multipliers

$$\begin{aligned}
\widehat{P_{\leq N} f}(\xi) &:= \varphi(\xi/N)\hat{f}(\xi),\\
\widehat{P_{> N} f}(\xi) &:= (1 - \varphi(\xi/N))\hat{f}(\xi),
\end{aligned}$$

$$\widehat{P_N f}(\xi) := (\varphi(\xi/N) - \varphi(2\xi/N))\hat{f}(\xi),$$

and similarly $P_{<N}$ and $P_{\geq N}$. We also define

$$P_{M<\cdot\leq N} := P_{\leq N} - P_{\leq M} = \sum_{M<N'\leq N} P_{N'}$$

whenever $M < N$. We will usually use these multipliers when M and N are *dyadic numbers* (that is, of the form 2^n for some integer n); in particular, all summations over N or M are understood to be over dyadic numbers. Nevertheless, it will occasionally be convenient to allow M and N to not be a power of 2.

Like all Fourier multipliers, the Littlewood–Paley operators commute with the propagator $e^{it\Delta}$, as well as with differential operators such as $i\partial_t + \Delta$. We will use basic properties of these operators many times, including

Lemma 11.4 (Bernstein estimates). *For* $1 \leq p \leq q \leq \infty$,

$$\begin{aligned}
\big\||\nabla|^{\pm s} P_N f\big\|_{L^p(\mathbb{R}^d)} &\sim N^{\pm s}\|P_N f\|_{L^p(\mathbb{R}^d)},\\
\|P_{\leq N} f\|_{L^q(\mathbb{R}^d)} &\lesssim N^{\frac{d}{p}-\frac{d}{q}}\|P_{\leq N} f\|_{L^p(\mathbb{R}^d)},\\
\|P_N f\|_{L^q(\mathbb{R}^d)} &\lesssim N^{\frac{d}{p}-\frac{d}{q}}\|P_N f\|_{L^p(\mathbb{R}^d)}.
\end{aligned}$$

Lemma 11.5 (Square function estimates). *Given a Schwartz function f, let*

$$S(f)(x) := \Big(\sum_N \big|P_N f(x)\big|^2\Big)^{1/2}$$

denote the Littlewood–Paley square function. For $1 < p < \infty$,

$$\|S(f)\|_{L^p(\mathbb{R}^d)} \sim \|f\|_{L^p(\mathbb{R}^d)}.$$

More generally,

$$\Big\|\Big(\sum_N N^{2s}\big|P_N f(x)\big|^2\Big)^{1/2}\Big\|_{L^p(\mathbb{R}^d)} \sim \big\||\nabla|^s f\big\|_{L^p(\mathbb{R}^d)} \tag{11.1}$$

for all $s > -d$ and $1 < p < \infty$.

11.3 Fractional calculus

We first record the fractional product rule from [11]:

Lemma 11.6 (Fractional product rule, [11]). *Let $s \in (0,1]$ and $1 < r, p_1, p_2, q_1, q_2 < \infty$ such that $\frac{1}{r} = \frac{1}{p_i} + \frac{1}{q_i}$ for $i = 1, 2$. Then,*

$$\big\||\nabla|^s(fg)\big\|_{L^r(\mathbb{R}^d)} \lesssim \|f\|_{L^{p_1}(\mathbb{R}^d)}\big\||\nabla|^s g\big\|_{L^{q_1}(\mathbb{R}^d)} + \big\||\nabla|^s f\big\|_{L^{p_2}(\mathbb{R}^d)}\|g\|_{L^{q_2}(\mathbb{R}^d)}.$$

We will also need the following fractional chain rule from [11]. For a textbook treatment, see [37, §2.4].

Lemma 11.7 (Fractional chain rule, [11]). *Suppose $G \in C^1(\mathbb{C})$, $s \in (0,1]$, and $1 < p, p_1, p_2 < \infty$ are such that $\frac{1}{p} = \frac{1}{p_1} + \frac{1}{p_2}$. Then,*

$$\||\nabla|^s G(u)\|_{L^p(\mathbb{R}^d)} \lesssim \|G'(u)\|_{L^{p_1}(\mathbb{R}^d)} \||\nabla|^s u\|_{L^{p_2}(\mathbb{R}^d)}.$$

Although we will not need it in our applications here, for completeness we record the following fractional chain rule for when the function G is no longer C^1, but merely Hölder continuous:

Lemma 11.8 (Fractional chain rule for a Hölder continuous function, [39]). *Let G be a Hölder continuous function of order $0 < \alpha < 1$. Then, for every $0 < s < \alpha$, $1 < p < \infty$, and $\frac{s}{\alpha} < \sigma < 1$ we have*

$$\big\||\nabla|^s G(u)\big\|_{L^p(\mathbb{R}^d)} \lesssim \big\||u|^{\alpha - \frac{s}{\sigma}}\big\|_{L^{p_1}(\mathbb{R}^d)} \big\||\nabla|^\sigma u\big\|^{\frac{s}{\sigma}}_{L^{\frac{s}{\sigma}p_2}(\mathbb{R}^d)}, \tag{11.2}$$

provided $\frac{1}{p} = \frac{1}{p_1} + \frac{1}{p_2}$ and $(1 - \frac{s}{\alpha\sigma})p_1 > 1$.

11.4 A paraproduct estimate

In Section 9, we made use of a paraproduct estimate from [40]. The proof we present here is different from the one in [40]; however, it only requires basic knowledge of harmonic analysis and so it is better suited to these lecture notes.

Lemma 11.9 (Paraproduct estimate, [40]). *We have*

$$\big\||\nabla|^{-2/3}(fg)\big\|_{L^{4/3}(\mathbb{R}^4)} \lesssim \big\||\nabla|^{-2/3} f\big\|_{L^p(\mathbb{R}^4)} \big\||\nabla|^{2/3} g\big\|_{L^q(\mathbb{R}^4)},$$

for any $\frac{4}{3} < p < \infty$ and $1 < q < \infty$ such that $\frac{1}{p} + \frac{1}{q} = \frac{11}{12}$.

Proof. The claim is equivalent to the following estimate:

$$\big\||\nabla|^{-\frac{2}{3}}\{(|\nabla|^{\frac{2}{3}} f)(|\nabla|^{-\frac{2}{3}} g)\}\big\|_{L^{4/3}(\mathbb{R}^4)} \lesssim \|f\|_{L^p(\mathbb{R}^4)} \|g\|_{L^q(\mathbb{R}^4)}, \tag{11.3}$$

for $\frac{4}{3} < p < \infty$, $1 < q < \infty$ such that $\frac{1}{p} + \frac{1}{q} = \frac{11}{12}$. To prove this, we start by performing the following decomposition:

$$\begin{aligned} |\nabla|^{-\frac{2}{3}}\{(|\nabla|^{\frac{2}{3}} f)(|\nabla|^{-\frac{2}{3}} g)\} = |\nabla|^{-\frac{2}{3}}\Big\{ &\sum_{\frac{1}{8} \le \frac{N_1}{N_2} \le 8} P_{N_1}\big(|\nabla|^{\frac{2}{3}} f\big) P_{N_2}\big(|\nabla|^{-\frac{2}{3}} g\big) \\ &+ \sum_{N_1} P_{N_1}\big(|\nabla|^{\frac{2}{3}} f\big) P_{>8N_1}\big(|\nabla|^{-\frac{2}{3}} g\big) \\ &+ \sum_{N_1} P_{N_1}\big(|\nabla|^{\frac{2}{3}} f\big) P_{<\frac{1}{8}N_1}\big(|\nabla|^{-\frac{2}{3}} g\big)\Big\}. \end{aligned} \tag{11.4}$$

Next, we will show how to control the contribution of each of the terms on the right-hand side of (11.4) to (11.3).

Using Sobolev embedding, Cauchy–Schwarz, and the square function estimate (11.1), we estimate the contribution of the first term on the right-hand side of (11.4) as follows:

$$\begin{aligned}
&\Big\| |\nabla|^{-\frac{2}{3}} \sum_{\frac{1}{8}\leq \frac{N_1}{N_2}\leq 8} P_{N_1}\big(|\nabla|^{\frac{2}{3}} f\big) P_{N_2}\big(|\nabla|^{-\frac{2}{3}} g\big)\Big\|_{L^{4/3}} \\
&\lesssim \Big\| \sum_{\frac{1}{8}\leq \frac{N_1}{N_2}\leq 8} N_1^{-\frac{2}{3}} N_2^{\frac{2}{3}} \big|P_{N_1}\big(|\nabla|^{\frac{2}{3}} f\big)\big| \big|P_{N_2}\big(|\nabla|^{-\frac{2}{3}} g\big)\big|\Big\|_{L^{12/11}} \\
&\lesssim \Big\| \Big(\sum_{\frac{1}{8}\leq \frac{N_1}{N_2}\leq 8} \big|N_1^{-\frac{2}{3}} P_{N_1}\big(|\nabla|^{\frac{2}{3}} f\big)\big|^2\Big)^{\frac{1}{2}} \Big(\sum_{\frac{1}{8}\leq \frac{N_1}{N_2}\leq 8} \big|N_2^{\frac{2}{3}} P_{N_2}\big(|\nabla|^{-\frac{2}{3}} g\big)\big|^2\Big)^{\frac{1}{2}}\Big\|_{L^{12/11}} \\
&\lesssim \Big\| \Big(\sum_{\frac{1}{8}\leq \frac{N_1}{N_2}\leq 8} \big|N_1^{-\frac{2}{3}} P_{N_1}\big(|\nabla|^{\frac{2}{3}} f\big)\big|^2\Big)^{\frac{1}{2}}\Big\|_{L^p} \Big\| \Big(\sum_{\frac{1}{8}\leq \frac{N_1}{N_2}\leq 8} \big|N_2^{\frac{2}{3}} P_{N_2}\big(|\nabla|^{-\frac{2}{3}} g\big)\big|^2\Big)^{\frac{1}{2}}\Big\|_{L^q} \\
&\lesssim \|f\|_{L^p}\|g\|_{L^q}.
\end{aligned}$$

Arguing similarly, we estimate the contribution of the second term on the right-hand side of (11.4) as follows:

$$\begin{aligned}
&\Big\| |\nabla|^{-\frac{2}{3}} \sum_{N_1} P_{N_1}\big(|\nabla|^{\frac{2}{3}} f\big) P_{>8N_1}\big(|\nabla|^{-\frac{2}{3}} g\big)\Big\|_{L^{4/3}} \\
&\qquad\lesssim \Big\| \sum_{N_1} N_1^{-\frac{2}{3}} \big|P_{N_1}\big(|\nabla|^{\frac{2}{3}} f\big)\big| N_1^{\frac{2}{3}} \big|P_{>8N_1}\big(|\nabla|^{-\frac{2}{3}} g\big)\big|\Big\|_{L^{12/11}} \\
&\qquad\lesssim \Big\| \Big(\sum_{N_1} \big|N_1^{-\frac{2}{3}} P_{N_1}\big(|\nabla|^{\frac{2}{3}} f\big)\big|^2\Big)^{\frac{1}{2}} \Big(\sum_{N_1} \big|N_1^{\frac{2}{3}} P_{>8N_1}\big(|\nabla|^{-\frac{2}{3}} g\big)\big|^2\Big)^{\frac{1}{2}}\Big\|_{L^{12/11}} \\
&\qquad\lesssim \|f\|_{L^p}\|g\|_{L^q},
\end{aligned}$$

where we also used the following consequence of (11.1):

$$\Big\| \Big(\sum_N N^{2s} \big|P_{\geq N} h\big|^2\Big)^{\frac{1}{2}}\Big\|_{L^p} \sim \big\| |\nabla|^s h\big\|_{L^p} \quad \text{for all} \quad s>0 \quad \text{and} \quad 1<p<\infty.$$

It remains to estimate the contribution of the third term on the right-hand side of (11.4). To do this, we use Lemma 11.5, the easy estimates $|P_N h| \lesssim M(h)$ and $|P_{\leq N} h| \lesssim M(h)$, and the vector maximal inequality:

$$\Big\| |\nabla|^{-\frac{2}{3}} \sum_{N_1} P_{N_1}\big(|\nabla|^{\frac{2}{3}} f\big) P_{<\frac{1}{8}N_1}\big(|\nabla|^{-\frac{2}{3}} g\big)\Big\|_{L^{4/3}}$$

$$
\begin{aligned}
&\lesssim \Big\| \Big(\sum_N \Big| N^{-\frac{2}{3}} P_N \Big[\sum_{N_1} P_{N_1}\big(|\nabla|^{\frac{2}{3}} f\big) P_{<\frac{1}{8}N_1}\big(|\nabla|^{-\frac{2}{3}} g\big)\Big]\Big|^2 \Big)^{\frac{1}{2}} \Big\|_{L^{4/3}} \\
&\lesssim \Big\| \Big(\sum_N \Big| N^{-\frac{2}{3}} M \Big[\sum_{N_1 \sim N} P_{N_1}\big(|\nabla|^{\frac{2}{3}} f\big)\Big]\Big|^2 \Big)^{\frac{1}{2}} M\big(|\nabla|^{-\frac{2}{3}} g\big) \Big\|_{L^{4/3}} \\
&\lesssim \Big\| \Big(\sum_N \sum_{N_1 \sim N} \big| N^{-\frac{2}{3}} P_{N_1}\big(|\nabla|^{\frac{2}{3}} f\big)\big|^2 \Big)^{\frac{1}{2}} \Big\|_{L^p} \big\| M\big(|\nabla|^{-\frac{2}{3}} g\big)\big\|_{L^r} \\
&\lesssim \|f\|_{L^p} \big\| |\nabla|^{-\frac{2}{3}} g \big\|_{L^r},
\end{aligned}
$$

where r is such that $\frac{1}{p} + \frac{1}{r} = \frac{3}{4}$. (Note that this is the source of the restriction $p > \frac{4}{3}$.) The claim now follows by applying Sobolev embedding to the second factor on the right-hand side of the inequality above. $\square$

Bibliography

[1] Hajer Bahouri and Patrick Gérard, *High frequency approximation of solutions to ccritical nonlinear wave equations*, Amer. J. Math. **121** (1999), no. 1, 131–175. MR 1705001 (2000i:35123)

[2] P. Bégout and A. Vargas, *Mass concentration phenomena for the L^2-critical nonlinear Schrödinger equation*, Trans. Amer. Math. Soc. **359** (2007), no. 11, 5257–5282. MR 2327030 (2008g:35190)

[3] J. Bourgain, *Refinements of Strichartz' inequality and applications to 2D-NLS with critical nonlinearity*, Internat. Math. Res. Notices (1998), no. 5, 253–283. MR 1616917 (99f:35184)

[4] ———, *Global solutions of nonlinear Schrödinger equations*, American Mathematical Society Colloquium Publications, vol. 46, American Mathematical Society, Providence, RI, 1999. MR 1691575 (2000h:35147)

[5] ———, *Global wellposedness of defocusing critical nonlinear Schrödinger equation in the radial case*, J. Amer. Math. Soc. **12** (1999), no. 1, 145–171. MR 1626257 (99e:35208)

[6] H. Brézis and E. Lieb, *A relation between pointwise convergence of functions and convergence of functionals*, Proc. Amer. Math. Soc. **88** (1983), no. 3, 486–490. MR 699419 (84e:28003)

[7] R. Carles and S. Keraani, *On the role of quadratic oscillations in nonlinear Schrödinger equations. II. The L^2-critical case*, Trans. Amer. Math. Soc. **359** (2007), no. 1, 33–62 (electronic). MR 2247881 (2008a:35260)

[8] T. Cazenave, *Semilinear Schrödinger equations*, Courant Lecture Notes in Mathematics, vol. 10, New York University Courant Institute of Mathematical Sciences, New York, 2003. MR 2002047 (2004j:35266)

[9] T. Cazenave and F. B. Weissler, *Some remarks on the nonlinear Schrödinger equation in the critical case*, Nonlinear semigroups, partial differential equations and attractors (Washington, DC, 1987), Lecture Notes in Math., vol. 1394, Springer, Berlin, 1989, pp. 18–29. MR 1021011 (91a:35149)

[10] ———, *The Cauchy problem for the critical nonlinear Schrödinger equation in H^s*, Nonlinear Anal. **14** (1990), no. 10, 807–836. MR 1055532 (91j:35252)

H. Koch et al., *Dispersive Equations and Nonlinear Waves: Generalized Korteweg–de Vries, Nonlinear Schrödinger, Wave and Schrödinger Maps*, Oberwolfach Seminars 45,
DOI 10.1007/978-3-0348-0736-4,

[11] F. M. Christ and M. I. Weinstein, *Dispersion of small amplitude solutions of the generalized Korteweg-de Vries equation*, J. Funct. Anal. **100** (1991), no. 1, 87–109. MR 1124294 (92h:35203)

[12] J. Colliander, M. Keel, G. Staffilani, H. Takaoka, and T. Tao, *Global existence and scattering for rough solutions of a nonlinear Schrödinger equation on* $\mathbb{R}^3$, Comm. Pure Appl. Math. **57** (2004), no. 8, 987–1014. MR 2053757 (2005b:35257)

[13] ———, *Global well-posedness and scattering for the energy-critical nonlinear Schrödinger equation in* $\mathbb{R}^3$, Ann. of Math. (2) **167** (2008), no. 3, 767–865. MR 2415387 (2009f:35315)

[14] P. Constantin and J.-C. Saut, *Local smoothing properties of dispersive equations*, J. Amer. Math. Soc. **1** (1988), no. 2, 413–439. MR 928265 (89d:35150)

[15] B. Dodson, *Global well-posedness and scattering for the defocusing, L^2-critical, nonlinear Schrödinger equation when $d \geq 3$*, Preprint `arXiv:0912.2467`.

[16] P. Gérard, Y. Meyer, and F. Oru, *Inégalités de Sobolev précisées*, Séminaire sur les Équations aux Dérivées Partielles, 1996–1997, École Polytech., Palaiseau, 1997, pp. Exp. No. IV, 11. MR 1482810

[17] M. Keel and T. Tao, *Endpoint Strichartz estimates*, Amer. J. Math. **120** (1998), no. 5, 955–980. MR 1646048 (2000d:35018)

[18] S. Keraani, *On the defect of compactness for the Strichartz estimates of the Schrödinger equations*, J. Differential Equations **175** (2001), no. 2, 353–392. MR 1855973 (2002j:35281)

[19] ———, *On the blow up phenomenon of the critical nonlinear Schrödinger equation*, J. Funct. Anal. **235** (2006), no. 1, 171–192. MR 2216444 (2007e:35260)

[20] R. Killip, T. Tao, and M. Vişan, *The cubic nonlinear Schrödinger equation in two dimensions with radial data*, J. Eur. Math. Soc. (JEMS) **11** (2009), no. 6, 1203–1258. MR 2557134 (2010m:35487)

[21] R. Killip and M. Vişan, *Global well-posedness and scattering for the defocusing quintic nls in three dimensions*, To appear in Analysis and PDE.

[22] ———, *Nonlinear Schrödinger equations at critical regularity*, To appear in Proceedings of the 2008 Clay summer school, "Evolution Equations" held at the ETH, Zürich.

[23] ———, *The focusing energy-critical nonlinear Schrödinger equation in dimensions five and higher*, Amer. J. Math. **132** (2010), no. 2, 361–424. MR 2654778 (2011e:35357)

[24] E. H. Lieb and M. Loss, *Analysis*, second ed., Graduate Studies in Mathematics, vol. 14, American Mathematical Society, Providence, RI, 2001. MR 1817225 (2001i:00001)

[25] J. E. Lin and W. A. Strauss, *Decay and scattering of solutions of a nonlinear Schrödinger equation*, J. Funct. Anal. **30** (1978), no. 2, 245–263. MR 515228 (80k:35056)

[26] F. Merle and L. Vega, *Compactness at blow-up time for L^2 solutions of the critical nonlinear Schrödinger equation in 2D*, Internat. Math. Res. Notices (1998), no. 8, 399–425. MR 1628235 (99d:35156)

[27] K. Nakanishi, *Scattering theory for the nonlinear Klein-Gordon equation with Sobolev critical power*, Internat. Math. Res. Notices (1999), no. 1, 31–60. MR 1666973 (2000a:35174)

[28] T. Ozawa and Y. Tsutsumi, *Space-time estimates for null gauge forms and nonlinear Schrödinger equations*, Differential Integral Equations **11** (1998), no. 2, 201–222. MR 1741843 (2000m:35167)

[29] M. Riesz, *Sur les ensembles compacts de fonctions sommable*, Acta Sci. Math. (Szeged) **6** (1933), 136–142.

[30] E. Ryckman and M. Vişan, *Global well-posedness and scattering for the defocusing energy-critical nonlinear Schrödinger equation in $\mathbb{R}^{1+4}$*, Amer. J. Math. **129** (2007), no. 1, 1–60. MR 2288737 (2007k:35474)

[31] J. Shatah and M. Struwe, *Geometric wave equations*, Courant Lecture Notes in Mathematics, vol. 2, New York University Courant Institute of Mathematical Sciences, New York, 1998. MR 1674843 (2000i:35135)

[32] P. Sjölin, *Regularity of solutions to the Schrödinger equation*, Duke Math. J. **55** (1987), no. 3, 699–715. MR 904948 (88j:35026)

[33] T. Tao, *Global well-posedness and scattering for the higher-dimensional energy-critical nonlinear Schrödinger equation for radial data*, New York J. Math. **11** (2005), 57–80. MR 2154347 (2006e:35308)

[34] T. Tao and M. Vişan, *Stability of energy-critical nonlinear Schrödinger equations in high dimensions*, Electron. J. Differential Equations (2005), No. 118, 28. MR 2174550 (2006e:35307)

[35] T. Tao, M. Vişan, and X. Zhang, *Global well-posedness and scattering for the defocusing mass-critical nonlinear Schrödinger equation for radial data in high dimensions*, Duke Math. J. **140** (2007), no. 1, 165–202. MR 2355070 (2010a:35249)

[36] ———, *Minimal-mass blowup solutions of the mass-critical NLS*, Forum Math. **20** (2008), no. 5, 881–919. MR 2445122 (2009m:35495)

[37] M. E. Taylor, *Tools for PDE*, Mathematical Surveys and Monographs, vol. 81, American Mathematical Society, Providence, RI, 2000, Pseudodifferential operators, paradifferential operators, and layer potentials. MR 1766415 (2001g:35004)

[38] L. Vega, *Schrödinger equations: pointwise convergence to the initial data*, Proc. Amer. Math. Soc. **102** (1988), no. 4, 874–878. MR 934859 (89d:35046)

[39] M. Vişan, *The defocusing energy-critical nonlinear Schrödinger equation in higher dimensions*, Duke Math. J. **138** (2007), no. 2, 281–374. MR 2318286 (2008f:35387)

[40] ______, *Global well-posedness and scattering for the defocusing cubic nonlinear Schrödinger equation in four dimensions*, Int. Math. Res. Not. IMRN (2012), no. 5, 1037–1067. MR 2899959