VARIATIONS ON A THEME: ON THE DISPERSION OF WAVES

SUNG-JIN OH

Abstract. The linear wave equation

$$(-\partial_t^2 + \sum_{j=1}^{d} \partial_j^2)\phi = 0$$

underlies description of many fundamental wave phenomena in physics; examples include vibration of the string, elasticity, acoustics, optics, electromagnetism and gravity (general relativity), to mention just a few. A key property of a solution to the linear wave equation is *dispersion*, i.e., the decay of the amplitude of the solution while the total energy is conserved. Not only is it of obvious physical relevance, dispersion is often the central mechanism for stability and regularity in the mathematical investigation of nonlinear wave equations.

In this lecture series, I will describe not one, nor two, but *three* distinct proofs of dispersion for the wave equation, using 1) Fourier analysis and oscillatory integrals; 2) Klainerman's vector field method; and 3) decomposition into wave packets. Each proof has varying strengths and weaknesses, which I will demonstrate by discussing different nonlinear applications, respectively.

Introduction

The topic of this lecture series is the (linear) *wave equation*:

$$(-\partial_t^2 + \partial_{x^1}^2 + \cdots + \partial_{x^d}^2)\phi = 0, \quad (t, x) \in \mathbb{R}^{1+d}.$$

This PDE underlies many physical phenomena of basic importance:

- Motion of an elastic string (d'Alembert), membrane or body (elasticity);
- Propagation of sound (compressible Euler equation; gas dynamics);
- Propagation of light (Maxwell equation; electrodynamics, optics);
- Gravity (Einstein equation; general relativity).

Most of these interesting PDEs from physics are *nonlinear wave equations*. A suitable understanding of the simple linear wave equation is often the first step for studying such nonlinear equations.

We always consider the *initial value problem* for the wave equation. Introducing the shorthand

$$\Delta = \partial_{x^1}^2 + \cdots + \partial_{x^d}^2, \quad \Box = -\partial_t^2 + \partial_{x^1}^2 + \cdots + \partial_{x^d}^2$$

These notes were prepared for a sequence of three one-hour lectures at "2018 IBS–CGP Mathematics Festival"; the author thanks the participants of the Festival for their attention and enthusiasm. The author learned of the content of Lecture III below ("wave packet approach") through a seminar organized by D. Tataru at UC Berkeley in Spring 2014. The author would like to acknowledge contribution from all the participants of the seminar: M. Beceanu, B. Dodson, C. Gavrus, B. Haberman, B. Harrop-Griffiths, M. Ifrim, A. Lawrie, G. Liu and P. Smith.

The initial value problem for the wave equation consists of the following: Given a pair of functions ϕ_0, ϕ_1 on $\mathbb{R}^d$ and a function f on $I \times \mathbb{R}^d$, find the unique solution ϕ that satisfies

$$\begin{cases} \Box\phi = f \\ (\phi, \partial_t \phi)|_{t=0} = (\phi_0, \phi_1). \end{cases} \tag{0.1}$$

We will discuss the well-posedness (existence and uniqueness) of this problem in the course of this lecture.

We emphasize two absolutely fundamental features of this PDE. First, there is a naturally associated notion of *energy* that is conserved.

Theorem 0.1 (Energy conservation)**.** *Let ϕ be a "nice" solution to $\Box\phi = 0$ defined on $I \times \mathbb{R}^d$, where I is any interval in $\mathbb{R}$. Then the conserved energy of ϕ at time t, defined by*

$$E[\phi](t) = \int \frac{1}{2}\left(|\partial_t \phi|^2 + \sum_{i=1}^{d} |\partial_i \phi|^2 \right)(t)\, \mathrm{d}x$$

is constant in time.

Here, "nice" means that ϕ is smooth and decays to 0 sufficiently fast as $x \to \infty$.

Remark 0.2. From the energy conservation, uniqueness of a "nice" solution to the IVP with $f = 0$ follows.

Proof. We differentiate $E[\phi](t)$ in time, and use the equation.

$$\begin{aligned} \frac{\mathrm{d}}{\mathrm{d}t} E[\phi](t) &= \int \left(\partial_t \phi \partial_t^2 \phi + \sum_i \partial_i \phi \partial_i \partial_t \phi \right) \mathrm{d}x \\ &= \int \left(\partial_t \phi \partial_t^2 \phi - \sum_i \partial_t \phi \partial_i^2 \phi \right) \mathrm{d}x = 0. \end{aligned}$$

□

Remark 0.3. It may seem mysterious where $E[\phi](t)$ comes from. In specific applications, this quantity is associated with the physical notion of total energy of the system described by the solution $u(t)$. More generally, $E[u](t)$ arises as the conserved quantity associated with the time translation symmetry, via the so-called *Nöther principle.*

The energy $E[\phi](t)$ is a certain measure of the size of ϕ; when $\phi(t, x) \to 0$ as $x \to \infty$, then $E[\phi](t) = 0$ implies $\phi(t, \cdot) \equiv 0$. Therefore, Theorem 0.1 (conservation of energy) tells us that something stays the same. Nonetheless, when $d \geqslant 2$ it turns out that some other measures of the size of ϕ, namely the maximum amplitude of ϕ ($\|\phi\|_{L^\infty}$) and its differential $\partial\phi$ ($\|\partial\phi\|_{L^\infty}$), go to zero as $t \to \pm\infty$, by a mechanism called *dispersion*:

Theorem 0.4 (Dispersion)**.** *Let ϕ be a "nice" solution to $\Box\phi = 0$. Then*

$$\sup_{x \in \mathbb{R}^d} (|\phi(t, x)| + |\partial\phi(t, x)|) \lesssim_{\phi_0, \phi_1} (1 + |t|)^{-\frac{d-1}{2}}.$$

A rough description of the mechanism is as follows: Consider a solution ϕ to $\Box\phi = 0$, whose energy density is compactly supported initially. Although the energy is conserved, the solution u "disperses" in time, and its the support of the energy density spreads to larger and larger volumes. Since the total integral, which is the conserved energy, has to remain the same in time, $\|\partial\phi\|_{L^\infty}$ has to decay.

Remark 0.5. As we will see below, the situation is different in $d = 1$; the solution does not disperse, but is only transported.

The goal of this lecture series is to give not one, nor two, but *three* distinct *proofs* of this important fact.

(1) Proof via a representation formula;
(2) Proof via the vector field method;
(3) Proof via a wave packet decomposition.

Each proof has its own strengths and weaknesses, leading to different applications in the nonlinear case.

1. Lecture I: The representation formula approach

1.1. **d'Alembert and Kirchoff's formulae.** We start with some classical representation formulae in dimensions $d = 1, 3$ (and $d = 2$). For a reference, see [Evans, *Partial Differential Equations*, §2.4].

d'Alembert's formula in $d = 1$. We consider

$$(-\partial_t^2 + \partial_x^2)\phi = 0.$$

Note the obvious factorization

$$(-\partial_t^2 + \partial_x^2) = -(\partial_t - \partial_x)(\partial_t + \partial_x) = -(\partial_t + \partial_x)(\partial_t - \partial_x).$$

Thus

$$\phi(t, x) = \phi_{left}(x + t) + \phi_{right}(x - t).$$

To specify ϕ_{left} and ϕ_{right}, note that

$$\begin{aligned} \phi_{left}(x) + \phi_{right}(x) =& \phi(0, x) = \phi_0(x), \\ \partial_x\phi_{left}(x) - \partial_x\phi_{right}(x) =& \partial_t\phi(0, x) = \phi_1(x), \end{aligned}$$

which can be immediately solved as follows:

$$\partial_x\phi_{left}(x) = \frac{1}{2}\left(\partial_x\phi_0 + \phi_1\right)(x), \quad \partial_x\phi_{right}(x) = \frac{1}{2}\left(\partial_x\phi_0 - \phi_1\right)(x).$$

In conclusion,

$$\phi(t, x) = \frac{1}{2}\phi_0(x + t) + \frac{1}{2}\phi_0(x - t) + \frac{1}{2}\int_{x-t}^{x+t} \phi_1(y)\,\mathrm{d}y.$$

Kirchoff's formula in $d = 3$. In the polar coordinates $(t, r, \omega) = (t, r, \theta, \varphi)$,

$$\left(-\partial_{t^2} + \partial_r^2 + \frac{2}{r}\partial_r + \frac{1}{r^2}\Delta_\omega\right)\phi = 0.$$

where $\Delta_\omega = \partial_\theta^2 + \frac{\cos\theta}{\sin\theta}\partial_\theta + \frac{1}{\sin^2\theta}\partial_\varphi^2$ is the Laplacian on $\mathbb{S}^2$. We introduce the spherical mean

$$U(t, r, \omega) = \int_{\mathbb{S}^2} \phi(t, r, \omega)\,\frac{\mathrm{d}\omega}{4\pi},$$

where $\mathrm{d}\omega = \sin\theta\mathrm{d}\theta\mathrm{d}\varphi$. Then

$$\left(-\partial_t^2 + \partial_r^2 + \frac{2}{r}\partial_r\right)U(t, r) = 0.$$

A simple algebra gives, for $t, r > 0$

$$\left(-\partial_t^2 + \partial_r^2\right)(rU)(t,r) = 0$$

By extending U past $r = 0$ by an even reflection $U(r) = U(-r)$ for $r < 0$, the preceding equation holds for all $r \in \mathbb{R}$. Then using d'Alembert's formula, for $0 < r \leqslant t$ we have

$$rU(t,r) = \frac{1}{2}(r-t)U(0,t-r) + \frac{1}{2}(r+t)U(0,r+t) + \frac{1}{2}\int_{t-r}^{r+t} s\partial_t U(0,s)\,\mathrm{d}s$$

Since u is regular, we must have $\phi(t,0) = \lim_{r\to 0} U(t,r)$. Therefore,

$$\begin{aligned}\phi(t,0) =& \lim_{r\to 0+}\left(\frac{1}{2}(U(0,t-r)+U(0,t+r)) + \frac{t}{2r}(U(t+r)-U(t-r)) + \frac{1}{2r}\int_{t-r}^{t+r} s\partial_t U(0,s)\,\mathrm{d}s\right)\\ =& U(0,t) + tU'(t) + t\partial_t U(0,t)\\ =& \int_{\mathbb{S}^2} \phi_0(t,\omega) + t\partial_r\phi_0(t,\omega) + t\phi_1(t,\omega)\,\frac{\mathrm{d}\omega}{4\pi}\\ =& \frac{1}{4\pi t^2}\int_{S_t(0)} \left(\phi_0 + t\partial_r\phi_0 + t\phi_1\right)\mathrm{d}A(y).\end{aligned}$$

Translating any point x to the origin, we obtain

$$\phi(t,x) = \frac{1}{4\pi t^2}\int_{S_t(x)} \phi_0\,\mathrm{d}A(y) + \frac{1}{4\pi t}\int_{S_t(x)}\left(\frac{y-x}{|y-x|}\cdot\partial_y\phi_0 + \phi_1\right)\mathrm{d}A(y). \tag{1.1}$$

This formula makes clear the dispersion effect for a smooth, compactly supported initial data set (ϕ_0, ϕ_1).

Exercise 1.1 (Poisson's formula in $d = 2$)**.** By *the method of descent* from Kirchoff's formula, a representation formula in $d = 2$ can be derived:

$$\phi(t,x) = \frac{1}{2\pi t}\int_{B_t(x)} \frac{\phi_0}{(t^2-|y-x|^2)^{\frac{1}{2}}}\,\mathrm{d}A(y) + \frac{1}{2\pi}\int_{B_t(x)} \frac{\frac{y-x}{|y-x|}\cdot\partial_y\phi_0 + \phi_1}{(t^2-|y-x|^2)^{1/2}}\,\mathrm{d}A(y).$$

In this case, although it is trickier to see, it can be shown that ϕ has a uniform pointwise decay rate of $t^{-\frac{1}{2}}$ for a smooth, compactly supported initial data set (ϕ_0, ϕ_1).

1.2. **Notation.** To continue, we introduce a few notation and conventions that will be used throughout the lectures.

- *L^p norms.* For any $1 \leqslant p < \infty$ and any (nice) function f on $\mathbb{R}^d$, define

$$\|f\|_{L^p} = \left(\int |f|^p\,\mathrm{d}x\right)^{\frac{1}{p}}$$

  In case $p = \infty$, we let $\|f\|_{L^\infty} = \sup_{x\in\mathbb{R}^d}|f(x)|$.
- *Asymptotic notation.* $A \lesssim B$ means that there exists a constant $C > 0$ such that $A \leqslant CB$. We specify the parameters that C depends on by subscripts: For instance, $A \lesssim_a B$ means that $A \leqslant C(a)B$. We will often suppress the dependence of the constants on the dimension d. We use $A \simeq B$ to mean $A \lesssim B$ and $B \lesssim A$.

1.3. **A review of the Fourier transform.** To derive a representation formula which is convenient for all dimensions, we will use the Fourier transform. See [Evans, *Partial Differential Equations*, Ch. 4.3.1].

We introduce the *Schwartz class* $\mathcal{S}(\mathbb{R}^d)$ of $\mathbb{C}$-valued functions on $\mathbb{R}^d$:

$$\mathcal{S}(\mathbb{R}^d) = \{u \in C^\infty(\mathbb{R}^d) : \sup_x |x|^k |\partial_x^{(\ell)} u| < \infty \text{ for all } k, \ell \in \mathbb{N}_0\}.$$

Given a Schwartz function $u \in \mathcal{S}(\mathbb{R}^d)$, its Fourier transform $\hat{u} = \mathcal{F}(u)$ is defined by

$$\hat{u}(\xi) = \int_{\mathbb{R}^d} u(x) e^{-ix\cdot\xi} \, \mathrm{d}x =: \mathcal{F}(u)(\xi).$$

It is not difficult to check that $\hat{u} \in \mathcal{S}(\mathbb{R}^d)$ as well (see the properties below).

Two fundamental properties are as follows:

- *Fourier inversion formula:*

$$u(x) = \int_{\mathbb{R}^d} \hat{u}(\xi) e^{ix\cdot\xi} \, \frac{\mathrm{d}\xi}{(2\pi)^d} =: \mathcal{F}^{-1}(\hat{u})(x)$$

  and vice versa. In short, $\mathcal{F}^{-1}\mathcal{F} = \mathcal{F}\mathcal{F}^{-1} = \mathrm{Id}$.
- *Plancherel's identity:*

$$\int_{\mathbb{R}^d} |u(x)|^2 \, \mathrm{d}x = \int_{\mathbb{R}^d} |\hat{u}(\xi)|^2 \, \frac{\mathrm{d}\xi}{(2\pi)^d}.$$

We take the above two facts as granted. Other important properties of the Fourier transform follow rather easily from the Fourier inversion formula. For instance,

- *Diagonalization of partial differentiation:*

$$\widehat{\partial_j u}(\xi) = i\xi_j \hat{u}(\xi).$$

  To see this, we compute

$$\begin{aligned} \partial_j u(x) &= \int_{\mathbb{R}^d} \hat{u}(\xi) \partial_j e^{ix\cdot\xi} \, \frac{\mathrm{d}\xi}{(2\pi)^d} \\ &= \int_{\mathbb{R}^d} i\xi_j \hat{u}(\xi) e^{ix\cdot\xi} \, \frac{\mathrm{d}\xi}{(2\pi)^d}. \end{aligned}$$

- *Scaling property:* For any $\lambda > 0$, let $u_\lambda(x) := u(x/\lambda)$.

$$\widehat{u_\lambda}(\xi) = \lambda^d \hat{u}(\lambda\xi).$$

  We compute

$$\begin{aligned} u_\lambda(x) &= \int_{\mathbb{R}^d} \hat{u}(\xi) e^{i\lambda^{-1}x\cdot\xi} \, \frac{\mathrm{d}\xi}{(2\pi)^d} \\ &= \int_{\mathbb{R}^d} \hat{u}(\xi) e^{ix\cdot\lambda^{-1}\xi} \, \frac{\mathrm{d}\xi}{(2\pi)^d} \\ &= \int_{\mathbb{R}^d} \lambda^d \hat{u}(\lambda\eta) e^{ix\cdot\eta} \, \frac{\mathrm{d}\eta}{(2\pi)^d}. \end{aligned}$$

  (In fact, the effect of any linear change of variables in x can be easily computed.)

Finally, we study the behavior of products under the Fourier transform:

$$\begin{aligned} uv(x) &= \int_{\mathbb{R}^d} \hat{u}(\eta)\partial_j e^{ix\cdot\eta} \frac{\mathrm{d}\eta}{(2\pi)^d} \int_{\mathbb{R}^d} \hat{v}(\zeta)\partial_j e^{ix\cdot\zeta} \frac{\mathrm{d}\zeta}{(2\pi)^d} \\ &= \int_{\mathbb{R}^d}\int_{\mathbb{R}^d} \hat{u}(\eta)\hat{v}(\zeta) e^{ix\cdot(\eta+\zeta)} \frac{\mathrm{d}\eta}{(2\pi)^d}\frac{\mathrm{d}\zeta}{(2\pi)^d} \\ &= \int_{\mathbb{R}^d} \left(\int_{\mathbb{R}^d} \hat{u}(\eta)\hat{v}(\xi-\eta) \frac{\mathrm{d}\eta}{(2\pi)^d} \right) e^{ix\cdot\xi} \frac{\mathrm{d}\xi}{(2\pi)^d}, \end{aligned}$$

where on the last line, we made a change of variables $(\eta, \zeta) \mapsto (\eta, \xi = \eta + \zeta)$. Motivated by this computation, we introduce the *convolution* operation:

$$u * v(x) = \int u(y) v(x-y) \, \mathrm{d}y.$$

Then:

- *Product to convolution:*

$$\widehat{uv}(\xi) = \frac{1}{(2\pi)^d} \hat{u} * \hat{v}(\xi).$$

  Conversely,

$$\widehat{u * v}(\xi) = \hat{u}(\xi)\hat{v}(\xi).$$

1.4. **Representation formula by Fourier transform.** Let ϕ be a "nice" solution to the wave equation $\Box\phi = 0$ in $\mathbb{R}^{1+d}$. Taking the Fourier transform in space, we arrive at

$$\partial_t^2 \hat{\phi}(t,\xi) = -|\xi|^2 \hat{\phi}(t,\xi),$$

which is a second order ODE in t for each fixed value of $\xi \in \mathbb{R}^d$. Solving this ODE, we see that $\hat{\phi}$ is of the form:

$$\hat{\phi}(t,\xi) = \hat{\phi}_+(\xi) e^{it|\xi|} + \hat{\phi}_-(\xi) e^{-it|\xi|}$$

The coefficients $\hat{\phi}_\pm$ are determined by the initial data via the following linear relation

$$\begin{cases} \hat{\phi}_+(\xi) + \hat{\phi}_-(\xi) = \hat{\phi}_0(\xi), \\ i|\xi|(\hat{\phi}_+(\xi) - \hat{\phi}_-(\xi)) = \hat{\phi}_1(\xi), \end{cases}$$

Solving this linear relation, we arrive at

$$\hat{\phi}_\pm(\xi) = \frac{1}{2}\left(\hat{\phi}_0(\xi) \pm \frac{1}{i|\xi|}\hat{\phi}_1(\xi) \right). \tag{1.2}$$

By the inverse Fourier transform, we have the representation formula

$$\phi(t,x) = \int_{\mathbb{R}^d} \phi_+(\xi) e^{i(t|\xi| + x\cdot\xi)} \frac{\mathrm{d}\xi}{(2\pi)^d} + \int_{\mathbb{R}^d} \phi_-(\xi) e^{i(-t|\xi| + x\cdot\xi)} \frac{\mathrm{d}\xi}{(2\pi)^d}. \tag{1.3}$$

1.5. **Simple model problem: Oscillatory integrals.** Before we continue, we consider the simpler problem of estimating a model 1-dimensional oscillatory

$$I = \int_{\mathbb{R}} e^{i\lambda\Phi(\xi)} a(\xi)\, \mathrm{d}\xi.$$

We assume that the (unweighted) phase function $\Phi(\xi)$ is real-valued, and that the amplitude $a(\xi)$ has compact support in $(-1, 1)$. On this set, we assume that Φ, a are uniformly smooth, in the sense that

$$|\Phi^{(k)}(\xi)| + |a^{(k)}(\xi)| \lesssim_k 1.$$

We are interested in the size of I as $\lambda \to \infty$.

The basic idea is to play the rapid oscillation of $e^{i\lambda\Phi(\xi)}$ against the slowly varying amplitude $a(\xi)$, by using the formula

$$e^{i\lambda\Phi(\xi)} = \frac{1}{i\lambda\, \partial_\xi \Phi(\xi)} \partial_\xi e^{i\lambda\Phi(\xi)}$$

and integrating by parts in ξ. Consider the following two basic examples:

Example 1: No critical points. Consider

$$I = \int e^{i\lambda\Phi(\xi)} a(\xi)\, \mathrm{d}\xi$$

with $\Phi'(\xi) > 1$ on $(-1, 1)$. Then repeated integration by parts yields

$$I = O_N(\lambda^{-N}).$$

Example 2: One nondegenerate critical point. Consider $\Phi(\xi) = \xi^n$, i.e.,

$$I = \int e^{i\lambda\xi^n} a(\xi)\, \mathrm{d}\xi.$$

Note that we would see no oscillation in the interval $|\lambda\xi^n| \ll 1$; it is reasonable to expect that the contribution of this interval gives the main term. Indeed, this idea turns out to be true: To see this, we make the change of variables

$$\eta = \lambda\xi^n, \quad \mathrm{d}\xi = \lambda^{-\frac{1}{n}} \eta^{-\frac{n-1}{n}} \mathrm{d}\eta.$$

Then

$$I = \lambda^{-\frac{1}{n}} \int e^{i\eta} a(\lambda^{-\frac{1}{n}} \eta^{\frac{1}{n}}) \eta^{-\frac{n-1}{n}}\, \mathrm{d}\eta.$$

It is not difficult to verify that

$$I = \lambda^{-\frac{1}{n}} \int e^{i\eta} a(0) \eta^{-\frac{n-1}{n}}\, \mathrm{d}\eta + O(\lambda^{-\frac{2}{n}}).$$

Exercise 1.2. Prove the preceding formula.

Interlude: Dyadic decomposition. In anticipation of what to come, however, we present a different proof, which does not involve change of variables, but rather a "dyadic decomposition" of the interval, adapted to the phase $\lambda\xi^n$.

We begin by defining the notion of a smooth dyadic decomposition. Let $\chi_{<1}$ be a smooth function on $\mathbb{R}^d$ supported on $\{|x| < 2\}$ and is equal to 1 on $\{|x| < 1\}$. For any $\mu > 0$, let

$$\chi_{<\mu}(x) = \chi_{<1}(x/\mu), \quad \chi_\mu(x) = \chi_{<\mu}(x) - \chi_{<\mu/2}(x).$$

Note that $\chi_{<\mu}$ is supported in $\{|x| < 2\mu\}$, and χ_μ is supported in $\{\frac{\mu}{2} < |x| < 2\mu\}$. Positive numbers of the form 2^k with $k \in \mathbb{Z}$ are called *dyadic numbers*, and the set of all dyadic numbers is denoted by $2^{\mathbb{Z}}$. Note that for any $\mu, \nu \in 2^{\mathbb{Z}}$ with $\nu < \mu$,

$$\chi_{<\mu} = \sum_{\mu' \in 2^{\mathbb{Z}} : \mu' \leqslant \mu} \chi_{\mu'} = \chi_{<\nu} + \sum_{\mu' \in 2^{\mathbb{Z}} : \nu < \mu' \leqslant \mu} \chi_{\mu'}.$$

Dyadic decomposition is an effective way to reduce the continuous problem to more manageable discrete problems, because our phase $\lambda\xi^n$ is a *polynomial power*, so that its values are roughly equivalent on each dyadic interval $\{\frac{1}{2}\mu < \xi < 2\mu\}$.

Example 2': One degenerate critical point, dyadic decomposition. Motivated by the above consideration, we use

$$\chi_{<1} = \chi_{<\lambda^{-1/n}} + \sum_{\substack{\alpha \in 2^{\mathbb{Z}} \\ 1 < \alpha \leqslant \lambda^{1/n}}} \chi_{\alpha\lambda^{-1/n}},$$

to split I (we will be loose about the endpoints of the dyadic sums!). In the first region, there is no oscillation to exploit. Thus,

$$I_0 := \int \chi_{<\lambda^{-1/n}}(\xi) a(\xi) e^{i\lambda\xi^n} \, \mathrm{d}\xi = O\left(\frac{1}{\lambda^{1/n}}\right).$$

On the other hand,

$$\begin{aligned} I_\alpha &:= \int \chi_{\alpha\lambda^{-1/n}}(\xi) a(\xi) e^{i\lambda\xi^n} \, \mathrm{d}\xi \\ &= \int \chi_{\alpha\lambda^{-1/n}}(\xi) a(\xi) \frac{1}{ni\lambda\xi^{n-1}} \partial_\xi e^{i\lambda\xi^n} \, \mathrm{d}\xi \\ &= -\int \partial_\xi \left(\chi_{\alpha\lambda^{-1/n}}(\xi) a(\xi) \frac{1}{ni\lambda\xi^{n-1}} \right) e^{i\lambda\xi^n} \, \mathrm{d}\xi. \end{aligned}$$

Note that $\partial_\xi \left(\chi_{\alpha\lambda^{-1/n}}(\xi) a(\xi) \frac{1}{ni\lambda\xi^{n-1}} \right) = O(\frac{1}{\lambda\xi^n})$, and it is still supported in $\{\frac{1}{2}\alpha\lambda^{-1/n} < |\xi| < 2\alpha\lambda^{-1/n}\}$. Therefore, integration yields

$$|I_\alpha| \lesssim \frac{1}{\alpha^{n-1}\lambda^{\frac{1}{n}}}.$$

which may be summed up for $\alpha > 1$.

Heuristic principles. The preceding simple computations generalize to the following heuristic principles:

(1) *Localization (or nonstationary phase) principle:* If Φ is nonstationary (i.e., $\nabla\Phi \neq 0$) on the support of a, then I can be shown to be rapidly decaying like $O(\lambda^{-N})$ for any $N \in \mathbb{N}_0$. Therefore, the problem of estimating I can be *localized* to the regions near stationary points $\{\xi : \nabla\Phi(\xi) = 0\}$.

(2) *Scaling (or stationary phase) principle:* The size of an oscillatory integral I whose amplitude is localized near a stationary point ξ_0 is determined by the order of vanishing of $\nabla\Phi$. The main contribution comes from the ξ-region where $|\lambda\Phi(\xi) - \lambda\Phi(\xi_0)| \lesssim 1$.

1.6. **Proof of the dispersive inequality.** The precise statement of the dispersive inequality that we will prove is as follows.

Theorem 1.3 (Dispersive inequality)**.**

$$\|\phi(t)\|_{L^\infty} \lesssim |t|^{-\frac{d-1}{2}} \sum_{\mu\in 2^{\mathbb{Z}}} \mu^{\frac{d+1}{2}} \|(P_\mu \phi_0, \mu^{-1} P_\mu \phi_1)\|_{L^1}.$$

Here, $P_\mu u = \mathcal{F}^{-1}(\chi_\mu \hat{u})$. The interest in this inequality lies in the range $|t| > 1$.

Proof. We proceed in several steps.

Step 1: Reduction via symmetries. We begin by reducing the problem using the symmetries of the problem. First, by the symmetry under reflection $(t, x) \mapsto (-t, -x)$, we may assume that $t > 0$. Next, we decompose

$$1 = \sum_{\mu\in 2^{\mathbb{Z}}} \chi_\mu(\xi),$$

and also introduce $\tilde{\chi}_\mu(\xi) = (\chi_{\mu/2} + \chi_\mu + \chi_{2\mu})(\xi)$ so that $\tilde{\chi}_\mu \chi_\mu = \chi_\mu$. Then we may decompose

$$\hat{\phi}(t, \xi) = \sum_{\pm} \sum_{\mu\in\mathbb{Z}} \tilde{\chi}_\mu(\xi) \hat{\phi}_{0,\pm}(\xi) \chi_\mu(\xi) e^{\pm it|\xi|}.$$

Write

$$I_{\mu,\pm}(t, x) = \int \chi_\mu(\xi) e^{i(\pm t|\xi| + x\cdot\xi)} \frac{\mathrm{d}\xi}{(2\pi)^d}.$$

Then

$$\phi(t, x) = \sum_{\mu} (I_{\mu,\pm}(t) * \psi_{\mu,\pm})(x).$$

where $\hat{\psi}_{\mu,\pm} = \tilde{\chi}_\mu \hat{\phi}_{0,\pm}$. By Hölder's inequality,

$$|\phi(t, x)| \lesssim \sum_{\mu} \|I_\mu(t)\|_{L^\infty} \|\psi_{\mu,\pm}\|_{L^1}.$$

It is not difficult to show that

$$\|\psi_{\mu,\pm}\|_{L^1} = \|\mathcal{F}^{-1}(\tilde{\chi}_\mu \hat{\phi}_{0,\pm})\|_{L^1} \lesssim \sum_{\mu'\in\{\mu/2,\mu,2\mu\}} \|(P_{\mu'}\phi_0, \mu^{-1} P_{\mu'}\phi_1)\|_{L^1}.$$

Therefore, it is left to verify

$$\|I_{\mu,\pm}(t, x)\|_{L^\infty_x} \lesssim t^{-\frac{d-1}{2}} \mu^{\frac{d+1}{2}}.$$

By the time reversal and scaling symmetries, we may assume that $\pm = +$ and $\mu = 1$.

Step 2: Oscillatory integral estimate. By the previous step, the problem is reduced to analyzing the size of the oscillatory integral

$$I_1 = \int_{\mathbb{R}^d} \chi_1(\xi) e^{i\Phi(t,x,\xi)} \,\mathrm{d}\xi, \tag{1.4}$$

where the phase Φ is given by

$$\Phi(t, x, \xi) = t|\xi| + x \cdot \xi$$

and the amplitude $\chi_1(\xi)$ is a smooth function supported in the annulus $\{\frac{1}{2} < |\xi| < 2\}$.

Step 2.1: Basic observations. By a suitable rotation, we may assume that the point x lies on the x^1-axis, i.e., $x = (x^1, 0, \dots, 0)$. The phase function thus becomes

$$\Phi(t, x, \xi) = t|\xi| + x^1 \xi_1.$$

As before, the basic idea is to use the formula

$$e^{i\Phi(t,x,\xi)} = \frac{1}{i\partial_{\xi_j} \Phi(t, x, \xi)} \partial_{\xi_j} e^{i\Phi(t,x,\xi)}$$

and to integrate by parts in ξ. Note that

$$\partial_{\xi_i} \Phi(t, x, \xi) = t\frac{\xi_i}{|\xi|} + x^1 \delta_{1i}$$

and

$$\partial_{\xi_i} \partial_{\xi_j} \Phi(t, x, \xi) = \frac{t}{|\xi|} \left(\delta_{ij} - \frac{\xi_i \xi_j}{|\xi|^2} \right).$$

We also note the easy higher derivative bounds (for $k \geqslant 3$)

$$|\partial_\xi^{(k)} \Phi(t, x, \xi)| \lesssim t \quad \text{for } \frac{1}{2} < |\xi| < 2.$$

Stationary phase (i.e., critical point of Φ) when $t > 0$ occurs when

$$\frac{\xi_1}{|\xi|} = \frac{x^1}{t}, \quad \xi_2 = \cdots = \xi_d = 0,$$

which forces $\xi_1 = |\xi|$ and $t = x^1$.

Step 2.2: Region with no stationary phase. We first treat the region with no stationary phase. Let η be a smooth function on $\mathbb{R}$ which equals 1 on $(-\infty, -\frac{1}{4})$ and 0 on $(0, \infty)$. Consider

$$I_{nonstat} = \int \chi_1(\xi) \eta(\xi_1 - \tfrac{1}{2}) e^{i\Phi(t,x,\xi)} \, \mathrm{d}\xi.$$

It is not difficult to verify that $|\nabla \Phi(t, x, \xi)| \gtrsim t$ in the set $\{\frac{1}{2} < |\xi| < 2,\ \xi_1 < \frac{1}{4}\}$. Repeated integration by parts then gives

$$|I_{nonstat}| \lesssim_N O(t^{-N}).$$

Step 2.3: Region with stationary phase. We may now focus on the set $\{\frac{1}{4} < \xi_1 < 2\}$, where there are possibly stationary phases. We introduce the notation

$$\tilde{\chi}_1(\xi) = \chi_1(\xi)(1 - \eta)(\xi_1 - \tfrac{1}{2}),$$

and write

$$I_{stat} = I - I_{nonstat} = \int \tilde{\chi}_1(\xi) e^{i\Phi(t,x,\xi)} \, \mathrm{d}\xi.$$

The Hessian of Φ is

$$\nabla_\xi^2 \Phi = \frac{t}{|\xi|} \begin{pmatrix} 0 & 0 \\ 0 & \delta_{ij} \end{pmatrix}.$$

This means that near the the critical points of $\nabla\Phi$,

$$\Phi(t, x, \xi) = \frac{t}{\xi_1} (\xi_2^2 + \cdots \xi_d^2) + \cdots ,$$

where $\xi_1 \simeq 1$. We expect the dominant contribution to be the ξ-volume of the region where $\frac{t}{\xi_1}(\xi_2^2 + \cdots \xi_d^2) \lesssim 1$, which is $O(t^{-\frac{d-1}{2}})$.

To make this idea precise, we note that

$$\left| \iint \chi_{<t^{-1/2}}(\xi') \tilde{\chi}_1(\xi) e^{i\Phi(t,x,\xi)} \, \mathrm{d}\xi \right| \lesssim t^{-\frac{d-1}{2}},$$

where $\xi' = (\xi_2, \ldots, \xi_d)$. On the other hand, in the region

$$\{\tfrac{1}{4}\alpha^2 \leqslant t(\xi_2^2 + \cdots + \xi_d^2) \leqslant 4\alpha^2\} \cap \{\tfrac{1}{2} < |\xi| < 2,\ \tfrac{1}{8} < \xi_1\},$$

we have the bound

$$\left| \partial_{\xi'}^{(k)} \left(\frac{1}{\partial_{\xi_j} \Phi(t,x,\xi)} \right) \right| \lesssim_k \frac{1}{t} (\alpha t^{-1/2})^{-1-k}.$$

so that repeated integration by parts (k-times) yields

$$\left| \iint \chi_{\alpha t^{-1/2}}(\xi') \tilde{\chi}_1(\xi) e^{i\Phi(t,x,\xi)} \, \mathrm{d}\xi \right| \lesssim_k \frac{1}{t^k} (\alpha t^{-1/2})^{-2k} \alpha^{d-1} t^{-\frac{d-1}{2}} \lesssim \alpha^{(d-1)-2k} t^{-\frac{d-1}{2}}$$

which may be summed up for $\alpha \gtrsim 1$ if $k \gtrsim d$. □

Exercise 1.4. For $\alpha > 0$, consider the *fractional Schrödinger equation*

$$i\partial_t \phi + |D|^\alpha \phi = 0.$$

where $|D|^\alpha \phi = \mathcal{F}^{-1}(|\xi|^\alpha \phi)$. The case $\alpha = 2$ corresponds to the usual Schrödinger equation, and $\alpha = 1$ is the (half-)wave equation, that we just considered.

Formulate and prove the dispersive inequality in the case $\alpha > 1$. What is the uniform decay rate in t? Also, what happens when $0 < \alpha < 1$?

2. Lecture II: The vector field method

In this lecture, we follow an idea of Klainerman to prove the dispersive property of the wave equation using only physical space methods. The reference is [Klainerman, *Uniform decay estimates and the Lorentz invariance of the classical wave equation*, Comm. Pure. Appl. Math., 1985].

2.1. **More on the energy method.** Let ϕ be a (real-valued) solution to

$$\Box \phi = f.$$

Multiplying the equation by $\partial_t \phi$, we compute

$$-\partial_t \left(\frac{1}{2} |\partial_t \phi|^2 + \frac{1}{2} \sum_{j=1}^{d} |\partial_j \phi|^2 \right) + \partial_j \left(\partial_j \phi \partial_t \phi \right) = f \partial_t \phi. \tag{2.1}$$

Integrating this identity on $(t_1, t_2) \times \mathbb{R}^d$, and integrating by parts (or, in fancy terms, use the divergence theorem):

$$E[\phi](t_2) = E[\phi](t_1) + \int_{t_1}^{t_2} \int_{\mathbb{R}^d} f \partial_t \phi \, \mathrm{d}x \mathrm{d}t.$$

For instance, by the Cauchy–Schwarz inequality, we have a useful basic inequality

$$\|\nabla_{t,x} \phi(t_2)\|_{L^2} \leqslant \|\nabla_{t,x} \phi(t_1)\|_{L^2} + C \int_{t_1}^{t_2} \|f(t')\|_{L^2} \, \mathrm{d}t'.$$

Exercise 2.1. Integrating the above identity over $C(t_0, x_0) \cap \{t_1 < t < t_2\}$, where

$$C(t_0, x_0) = \{(t, x) : |x - x_0| < t_0 - t,\, t > 0\}.$$

it follows that

$$\int_{B_{t_0-t_2}(x_0)} \frac{1}{2}|\partial_t \phi|^2 + \frac{1}{2}\sum_{j=1}^{d} |\partial_j \phi|^2 \, \mathrm{d}x \leqslant \int_{B_{t_0-t_1}(x_0)} \frac{1}{2}|\partial_t \phi|^2 + \frac{1}{2}\sum_{j=1}^{d} |\partial_j \phi|^2 \mathrm{d}x + \int_{C(t_0,x_0)\cap\{t_1<t<t_2\}} f \partial_t \phi \, \mathrm{d}t\mathrm{d}x.$$

This estimate proves *finite speed of propagation*: If $(\phi_0, \phi_1) = 0$ on $B_R(x_0)$, then the solution $\Box\phi = 0$ is zero on $C_R(x_0)$ (called the future *domain of dependence* of $B_R(x_0)$). This property implies a nice uniqueness statement for $\Box\phi = f$!

2.2. **Pointwise estimate via the energy method.** Conservation of energy allows us to have an L^2-type control on the solution. How do we convert this to a pointwise control on, say, $\partial\phi$?

Step 1: Commute $\Box\phi = f$ with ∂_μ, so that $\Box\partial_\mu\phi = \partial_\mu f$.
Step 2: Apply the energy inequality to control $\|\partial^{(k+1)}\phi\|_{L^2}$
Step 3: Apply the following *Sobolev inequality*:

Lemma 2.2 (Sobolev inequality)**.** *Let ϕ be a smooth compactly supported function on $\mathbb{R}^d$. Then we have*

$$|\phi(x)| \lesssim \sum_{j=0}^{\lfloor d/2 \rfloor+1} \|\partial^{(k)}\phi\|_{L^2(\mathbb{R}^d)}.$$

For a nice exposition of Sobolev spaces, see [Evans, *Partial Differential Equations*, Ch. 5]. Here, we present a simple proof on $\mathbb{R}^d$ using the Fourier transform.

Proof. Write $N = \lfloor d/2 \rfloor + 1$. We use the Fourier transform. By the inverse Fourier transform,

$$|\phi(x)| = \left| \int \hat{\phi}(\xi) \frac{\mathrm{d}\xi}{(2\pi)^d} \right| \leqslant \int |\hat{\phi}| \, \mathrm{d}\xi$$

We split the last integral and bound each term as follows:

$$\int_{\{|\xi|<1\}} |\hat{\phi}| \, \mathrm{d}\xi + \int_{\{|\xi|\}\geqslant 1} |\hat{\phi}| \, \mathrm{d}\xi \leqslant \|\hat{\phi}\|_{L^2} + \||\xi|^N \hat{\phi}\|_{L^2},$$

where we used that $|\xi|^{-2N}$ is integrable on $\{|\xi| \geqslant 1\}$. Using the properties of the Fourier transform, the desired statement follows. ☐

As a result, we have:

$$\sup_{(t,x)\in[0,T]\times\mathbb{R}^d} |\phi(t, x)| \lesssim \sum_{j=0}^{\lfloor d/2 \rfloor+1} \left(\|\partial^{(k)}\partial\phi(0)\|_{L^2} + \int_0^T \|\partial^{(k)} f(t)\|_{L^2} \, \mathrm{d}t \right).$$

2.3. **Overview of the strategy for dispersion.** We pursue the same strategy to prove pointwise *decay*, with some extra ingredients: First, instead of just controlling higher derivatives of ϕ, we attempt to control *weighted* higher derivatives of ϕ. Second, rather than the usual Sobolev inequality, we use a tailored version (often called the *Klainerman–Sobolev inequality*) which exploits the control of the weighted derivatives of ϕ.

2.4. **Symmetries of the $\Box$.** Just as the Laplacian Δ is intimately related to the Euclidean space $(\mathbb{R}^d, \delta)$, the d'Alembertian $\Box$ is associated with the scalar product $g(v,w)$ of the form

$$g(v,w) = -v^0 w^0 + v^1 w^1 + \cdots v^d w^d,$$

This scalar product is called the *Minkowski metric*, and the pair $(\mathbb{R}^{d+1}, g)$ is referred to as the *Minkowski spacetime*. Introducing the matrix notation

$$g_{\mu\nu} = \mathrm{diag}(-1, +1, \dots, +1),$$

we may write

$$\Box = (g^{-1})^{\mu\nu} \partial_\mu \partial_\nu$$

where we implicitly sum over repeated indices. From this expression, it is clear that $\Box$ is invariant under the Lorentz transformations, i.e., affine transformations of $\mathbb{R}^{d+1}$ that preserve g.

The Lorentz transformations consist of the following:

- **Translations.** $x^\mu \mapsto x^\mu + s v^\mu$.
- **Rotations.** Rotation in the (x^1, x^2) plane is given by the matrix

$$\begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & \cos\theta & -\sin\theta & 0 \\ 0 & \sin\theta & \cos\theta & 0 \\ 0 & 0 & 0 & \mathrm{Id} \end{pmatrix}.$$

- **Lorentz boosts.** Lorentz boost in the $(t = x^0, x^1)$ plane is given by the matrix

$$\begin{pmatrix} \frac{1}{\sqrt{1-v^2}} & \frac{-v}{\sqrt{1-v^2}} & 0 \\ \frac{-v}{\sqrt{1-v^2}} & \frac{1}{\sqrt{1-v^2}} & 0 \\ 0 & 0 & \mathrm{Id} \end{pmatrix}.$$

The infinitesimal generators of these symmetries are:

- **Translations.** $T_\mu = \partial_\mu$
- **Rotations.** $\Omega_{jk} = x_j \partial_k - x_k \partial_j$
- **Lorentz boosts.** $H_j = \Omega_{j0} = x_j \partial_t + t \partial_j$

It can be verified that all these vector fields commute with $\Box$:

$$[\Box, T_\mu] = 0, \quad [\Box, \Omega_{\mu\nu}] = 0.$$

Although it is not a symmetry of the Minkowski spacetime, the $\Box$ is also invariant under the scaling transformation $x^\mu \mapsto \lambda x^\mu$, whose infinitesimal generator is:

- **Scaling.** $S = x^\mu \partial_\mu$

In fact, we have

$$[\Box, S] = 2\Box.$$

The commutator relations for these vector fields are as follows.

$$\begin{aligned} [T_\mu, T_\nu] =& 0, \\ [\Omega_{\alpha\beta}, T_\mu] =& g_{\alpha\mu} T_\beta - g_{\beta\mu} T_\alpha, \\ [S, T_\mu] =& -T_\mu, \\ [\Omega_{\alpha\beta}, \Omega_{\mu\nu}] =& g_{\alpha\mu}\Omega_{\beta\nu} - g_{\beta\mu}\Omega_{\alpha\nu} + g_{\beta\nu}\Omega_{\alpha\mu} - g_{\alpha\nu}\Omega_{\beta\mu}, \\ [\Omega_{\alpha\beta}, S] =& 0, \end{aligned}$$

We use the notation Γ_j to refer to vector fields with different homogeneities. More precisely,

$$\Gamma_0 \in \{T_\mu\}, \quad \Gamma_1 \in \{\Omega_{jk}, H_j, S\}.$$

Schematically, one may summarize the commutation relations between these vector fields as follows.

$$[\Gamma_0, \Gamma_0] = 0, \quad [\Gamma_0, \Gamma_1] = \Gamma_0, \quad [\Gamma_1, \Gamma_1] = \Gamma_1.$$

We leave the derivation of these commutator formulae as an exercise.

Let $\mathcal{Z} := \{T_\mu, \Omega_{jk}, H_j, S\}$. For $\Gamma \in \mathcal{Z}$, note that

$$\Box\phi = 0 \Rightarrow \Box(\Gamma\phi) = 0.$$

Finally, we introduce the following schematic notation: We write

$$|\partial^{(k)}\phi|^2 = \sum_{\alpha_1, \dots, \alpha_k \in \{0, \dots, d\}} |\partial_{\alpha_1} \cdots \partial_{\alpha_k}\phi|^2, \quad |\partial^{(\leqslant k)}\phi|^2 = \sum_{j=0}^{d} |\partial^{(j)}\phi|^2,$$

$$|\Gamma^{(k)}\phi|^2 = \sum_{\Gamma_1, \dots, \Gamma_k \in \mathcal{Z}} |\Gamma_1 \cdots \Gamma_k\phi|^2, \quad |\Gamma^{(\leqslant k)}\phi|^2 = \sum_{j=0}^{k} |\Gamma^{(j)}\phi|^2.$$

Accordingly, we also write

$$|\partial\Gamma^{(k)}\phi|^2 = \sum_{\Gamma_1, \dots, \Gamma_k \in \mathcal{Z}} |\partial\Gamma_1 \cdots \Gamma_k\phi|^2, \quad |\partial\Gamma^{(\leqslant k)}\phi|^2 = \sum_{j=0}^{k} |\partial\Gamma^{(j)}\phi|^2.$$

2.5. **Weights from $\mathcal{Z}$.** We now study the weights obtained by commuting with the vector fields $\mathcal{Z}$. In what follows, we restrict to the case $t > 0$.

A general principle is that the control of vector field commuters in $\mathcal{Z}$ gives rise to control of $u\partial$, where $u = t - r$. More precisely, we have the relation

$$\partial_\mu = (-t^2 + |x|^2)^{-1}(x^\nu \Omega_{\mu\nu} + x_\mu S). \tag{2.2}$$

Note that, schematically,

$$[-t^2 + |x|^2, \Gamma_1] = 0, \quad [x, \Gamma_1] = x.$$

We therefore arrive at the following lemma.

Lemma 2.3. *We have*

$$u^k |\partial^{(k)}\phi| \lesssim |\Gamma_1^{(\leqslant k)}\phi|$$

where $\Gamma_1 \in \{\Omega, H, S\}$.

Remark 2.4. Lemma 2.3 and the trivial observation that the rotation vector fields Ω are invariant under scaling (which is closely related to the fact that Ω's are essentially the weight r times the normalized angular derivatives) suffice for the proof of the Klainerman–Sobolev inequality below.

2.6. **Klainerman–Sobolev inequality.**

Theorem 2.5. *Let ϕ be a smooth function on $\mathbb{R}^{1+d}$. Then the following inequality holds for $t \geqslant 0$:*

$$(1+|v|)^{\frac{d-1}{2}}(1+|u|)^{\frac{1}{2}}|\phi(t,x)| \lesssim \sum_{k=0}^{\lfloor d/2 \rfloor+1} \|\Gamma^{(k)}\phi(t)\|_{L^2(\mathbb{R}^d)}. \tag{2.3}$$

To prove this theorem, we need the following two ingredients:

Lemma 2.6 (Localized Sobolev inequality)**.** *For any smooth function ψ on $\mathbb{R}^d$ and $R > 0$, the following inequality holds.*

$$R^d|\psi(x)|^2 \lesssim_d \sum_{0\leqslant k\leqslant \lfloor d/2 \rfloor+1} R^{2k}\int_{B_R(x)} |\partial_y^k \psi|^2 \,\mathrm{d}V, \tag{2.4}$$

Proof. Without loss of generality, we may set $x = 0$. In the case $R = 1$, this lemma follows from the usual Sobolev inequality (Lemma 2.2) after a smooth cutoff. The general case $R > 0$ then follows by scaling. □

Lemma 2.7 (Localized Sobolev inequality in polar coordinates)**.** *Let ψ be a smooth function on $\mathbb{R}^d$ ($d \geqslant 2$). Then for any $x \neq 0$ and λ such that $0 < \lambda \leqslant r/2$ (where $r = |x|$), the following inequality holds.*

$$\lambda r^{d-1}|\psi(x)|^2 \lesssim_d \sum_{0\leqslant k+\ell\leqslant \lfloor d/2 \rfloor+1} \lambda^{2k}\int_{A_\lambda(r)} |\partial_r^k \Omega_x^{(\ell)}\psi(y)|^2 \,\mathrm{d}V \tag{2.5}$$

where $A_\lambda(r)$ is the annulus $\{y \in \mathbb{R}^d : ||y| - r| < \lambda\}$.

Proof. By scaling, it suffices to consider the case $r = 1$, in which case $0 < \lambda \leqslant \frac{1}{2}$. We can moreover restrict our attention to the angular sector $\{y : y^1/|y| \geqslant 1/10\}$, as we can cover the whole annulus $A_\lambda(1)$ by a finite number (depending on d) of its rotated copies (**Exercise:** Prove that the RHS remains equivalent under rotations).

The idea now is to flatten-out the angular directions. One concrete way to do it is simply to take $(r, y^2, \dots, y^d)$ as the coordinates. Because of the localization

$$r \in (1-\lambda, 1+\lambda) \subseteq (\tfrac{1}{2}, \tfrac{3}{2}), \quad \frac{y^1}{|y|} \geqslant \frac{1}{10}$$

we can check, with concrete computation, that:

$$\mathrm{d}V = J\mathrm{d}r \wedge \mathrm{d}y^2 \wedge \cdots \wedge \mathrm{d}y^d, \quad J \simeq 1,$$
$$|\partial_r^{(k)}\partial_{y'}^{(\ell)}\psi| \lesssim_{k,\ell} |\partial_r^k \Omega^{(\ell)}\psi|.$$

Then the proof of (2.5) is reduced to

$$\lambda|\psi(x)|^2 \lesssim_d \sum_{0\leqslant k+\ell\leqslant \lfloor d/2 \rfloor+1} \lambda^{2k}\int_{|y^1-1|<\lambda, |y'|\leqslant 1} |\partial_1^k \partial_{y'}^\ell \psi(y)|^2 \,\mathrm{d}y,$$

This follows from the usual Sobolev inequality by localizing to $(1/2, 3/2) \times \{|y'| \leqslant 1\}$, and scaling the first variable around 1. □

Proof of Theorem 2.5. We divide into two cases.

Case 1: $r \leqslant \frac{t}{2}$**.** By Lemma 2.3, we have

$$u^k |\partial^{(k)} \phi| \lesssim |\Gamma_1^{(\leqslant k)} \phi|.$$

For the region where $u \leqslant 1$, we also have the trivial schematic relation

$$|\partial^{(k)} \phi| \leqslant |\Gamma_0^{(\leqslant k)} \phi|.$$

Note furthermore that $u \simeq v \simeq t$ in this region. Applying Lemma 2.6 to balls $B_{t/2}(0)$, we obtain

$$(1 + t^{\frac{d}{2}}) |\phi(t, x)| \lesssim \sum_{k=0}^{\lfloor d/2 \rfloor + 1} \|\Gamma^{(k)} \phi(t)\|_{L^2(\mathbb{R}^d)},$$

which is sufficient.

Case 2: $r \geqslant \frac{t}{4}$**.** By Lemma 2.3, we have the schematic relation

$$u^k |\partial_r^k \Omega^{(\ell)} \phi| \lesssim |\Gamma_1^{(\leqslant k + \ell)} \phi|.$$

On the other hand, we also have the trivial schematic relation

$$|\partial_r^k \Omega^{(\ell)} \phi| \lesssim |\Gamma_0^{(\leqslant k)} \Gamma_1^{(\ell)} \phi|.$$

Performing a dyadic decomposition in u and applying Lemma 2.7, we obtain the desired statement. □

2.7. **Uniform decay of the derivative of solutions.**

Theorem 2.8. *Let ϕ be a "nice" solution to $\Box \phi = 0$ on $\mathbb{R}^{1+d}$. Then for $t \geqslant 0$, we have*

$$(1 + t + |x|)^{\frac{d-1}{2}} (1 + |u|)^{\frac{1}{2}} |\partial \phi(t, x)| \leqslant C \sum_{k=0}^{\lfloor d/2 \rfloor + 1} \|\partial \Gamma^{(k)} \phi(0, x)\|_{L^2_x}$$

Sketch of the proof. We follow the following strategy:

Step 1: Commute $\Box \phi = 0$ with the vector fields Γ; note that $\Box \Gamma \phi = 0$ as well.
Step 2: Apply the energy inequality to control $\|\partial \Gamma^{(k)} \phi\|_{L^2}$. In this process, we need:

Lemma 2.9. *We have*

$$|\partial \Gamma^{(k)} \phi| \leqslant C |\Gamma^{(\leqslant k)} (\partial \phi)|, \quad |\Gamma^{(k)} \partial \phi| = C |\partial \Gamma^{(\leqslant k)} \phi|$$

The proof is a straightforward application of the commutator identities.
Step 3: Apply the Klainerman–Sobolev inequality.

We leave the details to as an exercise. □

Remark 2.10. As discussed above, the following decay estimate for the wave equation with a forcing term $\Box \phi = f$ can be easily formulated and proved by the same strategy:

$$\sup_{(t,x) \in [0,T] \times \mathbb{R}^d} |\partial \phi(t, x)| \lesssim (1 + |t| + |x|)^{-\frac{d-1}{2}} (1 + |u|)^{-\frac{1}{2}} \sum_{j=0}^{\lfloor d/2 \rfloor + 1} \left(\|\Gamma^{(k)} \partial \phi(0)\|_{L^2} + \int_0^T \|\Gamma^{(k)} f(t)\|_{L^2} \, \mathrm{d}t \right).$$

Exercise 2.11. Consider the Schrödinger equation

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

Using the commuting operators

$$L_k = x_k + i2t\partial_k,$$

prove uniform t-decay of a "nice" solution ϕ with the sharp rate $t^{-\frac{d}{2}}$. [Hint: To prove an analogue of the Klainerman–Sobolev inequality, use the identity $e^{-i\frac{|x|^2}{4t}} L_k e^{i\frac{|x|^2}{4t}} = 2it\partial_k$.]

3. Lecture III: Wave packet approach

3.1. **Phase space decomposition and the uncertainty principle.** The two approaches so far for the proof of the dispersive property proceeded either in the Fourier space (Lecture I) or in the physical space (Lecture II). In this third and final lecture, we present the so-called *wave packet* or *phase space* approach, which is a powerful philosophy that "bridges" the Fourier and physical space approaches.

Before discussing the wave equation, let us discuss the more basic issue of expressing a given function f on $\mathbb{R}^d$ in various ways. One way to express f is by its pointwise values in the physical (or x-) space: $\mathbb{R}^d \ni x \mapsto f(x)$. Alternatively, we can take the Fourier transform and express f by its pointwise values in the Fourier (or ξ-) space: $\mathbb{R}^d \ni \xi \mapsto \hat{f}(\xi)$. Informally, we may think of these two viewpoints as decomposing f into "basis elements" consisting of, respectively, "delta distributions"[1] $\{\delta_{x_0}\}_{x_0 \in \mathbb{R}^d}$, which are sharply localized in the physical space, or plane waves $\{e^{i\xi_0 \cdot x}\}_{\xi_0 \in \mathbb{R}^d}$, which are sharply localized in the Fourier space. Each viewpoint has its own strength; the operation of differentiation is best understood with the Fourier-localized basis (since each plane wave $e^{i\xi_0 \cdot x}$ diagonalize all partial differentiation operators), but the operation of multiplication by another function is easier to understand with the physical-space-localized basis.

The *phase space viewpoint* is an idea that attempts to take the best of both worlds, by decomposing f into "basis elements" that are well-localized in both the physical and Fourier spaces. However, a fundamental complication of this viewpoint is that there does not exist a way to decompose functions into basis elements with arbitrarily good localization in both the physical and Fourier spaces. The following celebrated result embodies this property:

Theorem 3.1 (Uncertainty principle). *For any $f \in \mathcal{S}(\mathbb{R})$ and $x_0, \xi_0 \in \mathbb{R}^d$, we have*

$$\left(\int |x - x_0|^2 |f|^2 \, \mathrm{d}x \right) \left(\int |\xi - \xi_0|^2 |\hat{f}|^2 \, \frac{\mathrm{d}\xi}{2\pi} \right) \geqslant \frac{1}{4} \|f\|_{L^2}^4,$$

Proof. By translation, modulation and normalization, we may assume that $x_0 = \xi_0 = 0$ and $\|f\|_{L^2} = 1$. By Cauchy–Schwarz and Plancherel,

$$\left| \int \mathrm{Re}(x f \overline{\partial_x f}) \, \mathrm{d}x \right| \leqslant \left(\int |x|^2 |f|^2 \, \mathrm{d}x \right)^{1/2} \left(\int |\xi|^2 |\hat{f}|^2 \, \frac{\mathrm{d}\xi}{2\pi} \right)^{1/2}.$$

On the other hand,

$$\int \mathrm{Re}(x f \overline{\partial_x f}) \, \mathrm{d}x = \frac{1}{2} \int \mathrm{Re}([x, \partial_x] f \overline{f}) \, \mathrm{d}x = -\frac{1}{2} \int \mathrm{Re} f \overline{f} \, \mathrm{d}x = -\frac{1}{2}. \qquad \square$$

[1] At this informal level, the reader may view the delta distribution δ_{x_0} as a "generalized function" that is only supported at the point $\{x_0\}$, but is somehow nontrivial (in particular, has "integral" 1). This discussion may be made more precise with the help of measure theory, viewing δ_{x_0} as a measure with mass 1 supported only at $\{x_0\}$.

Remark 3.2. Following the proof, it can be verified that the extremizers of the uncertainty principle are the Gaussians.

Applying the uncertainty principle to each ξ_i-axis, we have the informal formulation:

$$\Delta x^i \cdot \Delta \xi_i \gtrsim 1.$$

Informally, functions which saturate the uncertainty principle are called *wave packets*, or *coherent waves*.

As in Remark 3.2, the Gaussians are the precise extremizers of the uncertainty principle, as formulated in Theorem 3.1. However, in applications it helps to take a more general view, and consider any function which obey appropriate decay (or localization) properties in both the physical and Fourier spaces. An archetypical example is given by fixing a Schwarz function χ, and considering the functions given by rescalings, translations and modulations (i.e., translation in the Fourier space) of χ to be wave packets. For a more precise formulation, see Section 3.2 below.

Remark 3.3. It would be convenient if we can take χ be compactly supported in both the physical and the Fourier space. However, there does not exist any nontrivial such function. This is due to the *Paley–Wiener theorem.* For simplicity we sketch the case of $\mathbb{R}$. If $\operatorname{supp} \hat{f} \subset (-A, A)$, then

$$f(z) = \int \hat{f}(\xi) e^{iz\cdot\xi} \frac{\mathrm{d}\xi}{2\pi}$$

is, in fact, an *entire (analytic)* function of $z \in \mathbb{C}$; such a function cannot be zero in an interval.

3.2. **Notation and conventions.** We start by introducing some notation.

- *Fourier multiplier.* For any function $f(\xi)$ on $\mathbb{R}^d$, we define the corresponding *Fourier multiplier* $f(D)$ to be the operator

$$\mathcal{F}(f(D)\phi)(\xi) = f(\xi)\hat{\phi}(\xi).$$

  Correspondingly, we use the notation $D_j = \frac{1}{i}\partial_j$, which is the Fourier multiplier corresponding to ξ_j. Fourier multipliers are flexible generalizations of constant coefficient differential operators.
- *Localization scales & orientation.* We denote by Δx (resp. $\Delta \xi$) a rectangular box of dimension $\Delta x^1 \times \cdots \times \Delta x^d$ (resp. $\Delta \xi_1 \times \cdots \times \Delta \xi_d$) in an orthonormal frame $(e_1, \ldots, e_d)$ in $\mathbb{R}^d_x$ (resp. $\theta^1 \times \cdots \times \theta^d$ in $\mathbb{R}^d_\xi = (\mathbb{R}^d_x)^*$). The numbers $(\Delta x^1, \ldots, \Delta x^d)$ (resp. $(\Delta \xi_1, \ldots, \Delta \xi_d)$) are called *localization scales*, and the frame $(e_1, \ldots, e_d)$ (resp. $(\theta^1, \ldots, \theta^d)$) is called the *orientation* of the rectangular box.

  We say that Δx and $\Delta \xi$ are *dual* if $(e_1, \ldots, e_n)$ and $(\theta^1, \ldots, \theta^d)$ are dual to each other and $\Delta x^i \Delta \xi_i = 1$.

  Oftentimes, we will rotate the axes and work with Δx, $\Delta \xi$ whose orientations coincide with the usual coordinate axes.
- Let χ be a Schwartz function on $\mathbb{R}^d$. Without loss of generality, we set:

  Δx and $\Delta \xi$ are dual localization scales whose orientation coincides with the usual coordinate axes.

A *normalized wave packet* based on χ centered at $(x_0, \xi_0) \in \mathbb{R}^d_x \times \mathbb{R}^d_\xi$ with localization scales $(\Delta x, \Delta\xi)$ is given by

$$\phi^{\Delta x,\Delta\xi}_{x_0,\xi_0}(x) = \frac{1}{(\Delta x^1 \cdots \Delta x^d)^{\frac{1}{2}}} e^{ix\cdot\xi_0} \chi\left(\frac{x^1 - x^1_0}{\Delta x^1}, \dots, \frac{x^d - x^d_0}{\Delta x^d}\right)$$

Note that

$$\hat{\phi}^{\Delta x,\Delta\xi}_{x_0,\xi_0} = \frac{1}{(\Delta\xi_1 \cdots \Delta\xi_d)^{\frac{1}{2}}} e^{-i(\xi-\xi_0)\cdot x_0} \hat{\chi}\left(\frac{\xi_1 - (\xi_0)_1}{\Delta\xi_1}, \dots, \frac{\xi_d - (\xi_0)_d}{\Delta\xi_d}\right).$$

and that $\|\phi^{\Delta x,\Delta\xi}_{x_0,\xi_0}\|_{L^2} = \|\chi\|_{L^2}$.

3.3. **Evolution of a single wave packet.** We would like to understand the evolution of a single wave packet $\phi^{\Delta x,\Delta\xi}_{x_0,\xi_0}$ under the evolution

$$\begin{cases} i\partial_t \phi \pm |D|\phi = 0, \\ \phi(0) = \phi^{\Delta x,\Delta\xi}_{x_0,\xi_0}. \end{cases}$$

To simplify the notation, we write $\phi_{x_0,\xi_0} = \phi^{\Delta x,\Delta\xi}_{x_0,\xi_0}$. By scaling and rotational symmetries, we may assume that $\frac{1}{2} < |\xi_0| < 2$ and $\xi_0 = ((\xi_0)_1, 0, \dots, 0)$. Without loss of generality, we may also assume that $t > 0$ and $\pm = +$.

Working in the Fourier space, we expand the symbol $|\xi|$ around $\xi = \xi_0$:

$$\begin{aligned} |\xi| &= \sqrt{|\xi_0 + (\xi - \xi_0)|^2} \\ &= \frac{\xi_0}{|\xi_0|} \cdot \xi + r_{\xi_0}(\xi - \xi_0). \end{aligned}$$

where $r_{\xi_0}(\xi - \xi_0)$ consist of quadratic of higher terms in $\xi - \xi_0$. Thus, back in the physical space,

$$\partial_t \phi + \frac{\xi_0}{|\xi_0|} \cdot \partial_x \phi = i r_{\xi_0}(D - \xi_0)\phi.$$

Since $\hat{\phi}(t, \cdot)$ is expected to be localized near ξ_0, we expect the RHS to be small. The linear operator on the LHS is nothing but the transport operator with constant velocity $\frac{\xi_0}{|\xi_0|}$; thus we expect

$$\phi(t, x) = \phi_{x_0(t),\xi_0}(x) + \text{error}$$

where

$$x_0(t) = x_0 + t\frac{\xi_0}{|\xi_0|}.$$

To quickly read off the time scale Δt on which such an approximation is valid, which will be related with $\Delta\xi$, we make one iteration and consider

$$\begin{cases} \left(\partial_t + \frac{\xi_0}{|\xi_0|} \cdot \partial_x\right) e_1 = i r_{\xi_0}(D - \xi_0)\phi_{x_0(t),\xi_0}, \\ e_1(t = 0) = 0 \end{cases}$$

and solve this equation for $0 \leqslant t \leqslant \Delta t$. By the energy method (i.e., multiplying by e_1 and integrating by parts)

$$\|e_1(t)\|_{L^2} \leqslant \int_0^t \|r_{\xi_0}(D - \xi_0)\phi_{x_0(t'),\xi_0}(t')\|_{L^2} \, \mathrm{d}t'.$$

Taking the Fourier transform, observe that $\hat{\phi}_{x_0(t'),\xi_0}(t') = e^{-i\frac{\xi_0}{|\xi_0|}t}\hat{\phi}_{x_0,\xi_0}(0)$. By the Plancherel identity, we may estimate

$$\begin{aligned}\int_0^t \|r_{\xi_0}(D-\xi_0)\phi_{x_0(t'),\xi_0}(t')\|_{L^2_x}\,\mathrm{d}t' \lesssim& \int_0^t \|r_{\xi_0}(\xi-\xi_0)\hat{\phi}_{x_0,\xi_0}\|_{L^2_\xi}\,\mathrm{d}t' \\ \lesssim& \Delta t \|r_{\xi_0}(\xi-\xi_0)\hat{\phi}_{x_0,\xi_0}\|_{L^2_\xi}\end{aligned}$$

Since

$$r_{\xi_0}(\xi-\xi_0) = \mathrm{Hess}_{\xi_0}|\xi|(\Delta\xi,\Delta\xi),$$

and the main term ϕ_{x_0,ξ_0} is normalized in L^2, we see that the error is small as long as

$$\mathrm{Hess}_{\xi_0}|\xi|(\Delta\xi,\Delta\xi)\Delta t \ll 1. \tag{3.1}$$

Since

$$\mathrm{Hess}_{\xi_0}|\xi| = \frac{1}{|\xi_0|}\begin{pmatrix} 0 & 0 \\ 0 & \mathrm{Id}_{(d-1)\times(d-1)}\end{pmatrix},$$

we see that the optimal choice is

$$\Delta\xi = (1, (\Delta t)^{-\frac{1}{2}}, \dots (\Delta t)^{-\frac{1}{2}}).$$

Remark 3.4. The velocity

$$v = \frac{\xi_0}{|\xi_0|} = \partial_\xi|\xi|(\xi_0)$$

is called the *group velocity* corresponding to the dispersion relation $\tau = |\xi|$ (here, τ is the temporal frequency). As we have seen, it is the velocity of the wave packet centered at ξ_0. Note that $\mathrm{Hess}_{\xi_0}|\xi|\Delta\xi$ can be interpreted as the *group velocity spread* Δv. Note that the relation (3.1) can be rewritten as

$$\Delta v \Delta t = \Delta x,$$

which means that Δt is not only the coherent time, but also the time when the nearby wave packets, which may be initially overlapping, become essentially disjoint.

The heuristics concerning the coherence time Δt can be made precise as follows.

Proposition 3.5 (Coherence)**.** *Let $\frac{1}{2} < |\xi_0| < 2$ and $\xi_0 = ((\xi_0)_1, 0, \dots, 0)$. Given $\Delta t > 0$, take*

$$\Delta\xi = (1, (\Delta t)^{-\frac{1}{2}}, \dots, (\Delta t)^{-\frac{1}{2}})$$

oriented with the usual coordinate axes, and let Δx be the dual localization scale.

Let ϕ be a solution to

$$(i\partial_t + |D|)\phi = 0.$$

Define

$$x_0(t) = x_0 + t\frac{\xi_0}{|\xi_0|},$$

and

$$\chi(t,x) = (\Delta x^1 \cdots \Delta x^d)^{\frac{1}{2}} e^{-i\sum_j \Delta x^j x^j (\xi_0)_j}\phi(t, x_0^1(t) + \Delta x^1 x^1, \dots, x_0^d(t) + \Delta x^d x^d).$$

If $\chi(0,x)$ obeys the Schwartz bounds

$$\sup_{x\in\mathbb{R}^d} ||x|^n \partial^{(m)}\chi(0,x)| \leqslant C_{n,m},$$

then there exist positive constants $\{\tilde{C}_{n,m}\}$ depending on $\{C_{n,m}\}$ such that

$$\sup_{(t,x)\in[0,\Delta t]\times\mathbb{R}^d} ||x|^n \partial^{(m)}\chi(t,x)| \leqslant \tilde{C}_{n,m} \quad \textit{for } 0 \leqslant t \leqslant \Delta t.$$

We defer the proof of Proposition 3.5 until Section 3.6.

Remark 3.6 (Connection with Knapp counterexample). In the constant-coefficient case, like the wave equation we have been considering, there is a Fourier-analytic way to interpret and construct a wave packet solution. In fact, the process is nothing but that of finding a *Knapp counterexample*, which is well-known. Namely, recall the Fourier representation formula for the (positive) half-wave equation:

$$\phi(t,x) = \int e^{i(t|\xi|+x\cdot\xi)}\hat{\phi}(0,\xi)\,\mathrm{d}\xi,$$

Recall further that a smooth cutoff $\hat{\phi}(0,\xi)$ to a suitable parallelepiped R in ξ corresponds to the Fourier transform of a wave packet initial data (with physical space center $x_0 = 0$; we may translate this around by multiplying $\hat{\phi}(0,\xi)$ by $e^{i\xi\cdot x_0}$). Then looking for a solution coherent for time Δt boils down to looking for a parallelepiped R in ξ such that the phase $e^{i(t|\xi|+x\cdot\xi)}$ is not oscillatory in the physical-space region $[0,\Delta t]\times R'$, where R' is the dual parallelepiped to R in the physical space (i.e., where $\phi(0,x)$ is localized); outside this region, the oscillation would take over and we would see decay. The choice of localization scale for $\hat{\phi}(0,\xi)$ as in Proposition 3.5 leads to $\Delta t \approx 1$; this is the standard Knapp counterexample for the wave equation.

3.4. **Wave packet decomposition.** We now wish to understand the evolution of more general initial data by decomposition of wave packets. As in Section 1.6, it suffices to understand

$$I_1 = \int \chi_1(\xi) e^{i(t|\xi|+x\cdot\xi)}\,\mathrm{d}\xi, \tag{3.2}$$

or equivalently, the solution to

$$\begin{cases} i\partial_t\phi + |D|\phi = 0, \\ \phi(0) = \mathcal{F}^{-1}(\chi_1)(x). \end{cases} \tag{3.3}$$

Given $\Delta t > 0$, $\Delta\xi$ is determined by (3.1); note that $\Delta\xi$ depends on the center ξ_0, in the sense that it is the rectangle of dimension $1\times(\Delta t)^{-\frac{1}{2}}\times\cdots\times(\Delta t)^{-\frac{1}{2}}$ oriented towards ξ_0. Consider uniformly separated covering of the annulus $\{\frac{1}{2} < |\xi| < 2\}$ by such rectangles $R^{\Delta\xi}_{\xi_0}$ (of which there are $O((\Delta t)^{\frac{d-1}{2}})$ many). Accordingly, in the Fourier space, we decompose

$$\chi_1(\xi) = \sum \chi^{\Delta\xi}_{\xi_0}(\xi)$$

where each $\chi^{\Delta\xi}_{\xi_0}$ is a smooth bump function essentially supported on the rectangle $R^{\Delta\xi}_{\xi_0}$. Back in the physical space, note that $\mathcal{F}^{-1}\chi^{\Delta\xi}_{\xi_0}$ is essentially supported in the dual localization scale Δx centered at $x_0 = 0$. Thus, we write

$$\chi^{\Delta\xi}_{\xi_0}(\xi) = \phi_{0,\xi_0}.$$

3.5. **Wave packet proof of the dispersive inequality.** Here, we give an alternative proof of Theorem 1.3 using the wave packet method. As in the previous proof, it suffices to consider the solution (3.3) and prove

$$\sup_{x\in\mathbb{R}^d} |\phi(t,x)| \lesssim t^{-\frac{d-1}{2}} \qquad t > 0.$$

Fix $t > 0$. We apply the wave packet decomposition as in Section 3.4 with $\Delta t = t$. In the physical space, the wave packets are all centered at $x_0 = 0$. Geometrically, we may see that at time t, the overlap among the wave packets is $O(1)$ (see Remark 3.4). Therefore, the maximum amplitude of ϕ at t is comparable to the amplitude of one wave packet, i.e.,

$$\sup_x |\phi(t,x)| \lesssim \sup_{\xi_0} \sup_x |\phi_{0,\xi_0}(t,x)| \lesssim \sup_{\xi_0} \sup_x |\phi_{0,\xi_0}(x)|.$$

According to the wave packet decomposition in Section 3.4, χ_0 is split into $(\Delta t)^{\frac{d-1}{2}}$ many pieces, corresponding to decomposition of the angular variables into caps of radius $(\Delta t)^{-\frac{1}{2}}$. Each piece has an (essentially) equal L^2-norm N, supported on a Fourier-space region with (essentially) equal volume V; these numbers are determined by

$$(\Delta t)^{\frac{d-1}{2}} N^2 \simeq 1, \quad (\Delta t)^{\frac{d-1}{2}} V \simeq 1,$$

or equivalently, $\|\phi_{0,\xi_0}\|_{L^2} \simeq (\Delta t)^{-\frac{d-1}{4}}$ and $|\mathrm{supp}\, \hat{\phi}_{0,\xi_0}| \simeq (\Delta t)^{-\frac{d-1}{2}}$. It follows that

$$|\phi_{0,\xi_0}| \leqslant \int |\hat{\phi}_{0,\xi_0}| \frac{\mathrm{d}\xi}{(2\pi)^d} \leqslant |\Delta\xi_1 \cdots \Delta\xi_d|^{\frac{1}{2}} \|\hat{\phi}_{0,\xi_0}\|_{L^2_\xi} \lesssim t^{-\frac{d-1}{2}},$$

as desired.

Remark 3.7. The strategy presented here is robust; it can be applied to the study of

$$\frac{1}{i}\partial_t \phi + A\phi = f$$

for a general partial differential (or pseudo-differential) operator A with variable coefficients. See [H. Koch and D. Tataru, *Dispersive estimates for principally normal pseudodifferential operator*, Comm. Pure. Appl. Math.].

3.6. **Proof of coherence.** Finally, we prove Proposition 3.5.

Step 1. We first treat the case $\Delta t = 1$, $\Delta \xi = (1, \dots, 1)$, $\Delta x = (1, \dots, 1)$, which is very simple.

We need to understand the evolution under the equation

$$\frac{1}{i}\partial_t \phi + |D|\phi = 0$$

of the initial data

$$\phi(0) = \phi_{x_0,\xi_0}^{\Delta x, \Delta \xi}.$$

As we have seen, for a wave packet localized near $\xi = \xi_0$, the leading order approximate equation is

$$\frac{1}{i}\left(\partial_t + \frac{\xi_0}{|\xi_0|}\cdot \partial_x\right)\phi = \cdots$$

Consider the solution operator $S_{\xi_0}[t]$ for the approximate equation with zero RHS:

$$S_{\xi_0}[t]\psi(x) = \psi\left(x - \frac{\xi_0}{|\xi_0|}t\right).$$

We write

$$\phi = S_{\xi_0}[t]\psi(t)$$

Without difficulty, we may translate in space and rotate the axes so that

$$x_0 = 0, \qquad \xi_0 = ((\xi_0)_1, 0, \dots, 0), \qquad \frac{1}{2} < |\xi_0| < 2.$$

The goal is to show that ψ remains well-localized at scales $\Delta x, \Delta\xi = (1, \dots, 1)$ near $(0, \xi_0)$ up until time $\Delta t = 1$.

Note that $\hat{\psi}$ obeys the equation:

$$\frac{1}{i}\partial_t \hat{\psi} + r_{\xi_0}(\xi - \xi_0)\hat{\psi} = 0, \qquad r_{\xi_0}(\eta) := |\xi_0 + \eta| - |\xi_0| - \frac{\xi_0}{|\xi_0|} \cdot \eta.$$

The ξ-localization of ψ therefore remains invariant. To determine the x-localization, we need to commute with ∂_{ξ_j}. However, since the symbol $r_{\xi_0}(\eta)$ in the range $\frac{1}{2} < |\xi_0| < 2$ and $|\eta| \lesssim 1$ clearly obey the bound

$$|\partial_\eta^{(n)} r_{\xi_0}(\eta)| \lesssim_n 1$$

it is not difficult to prove, by the energy method and an induction on the number of derivatives, that

$$\sum_{0 \leqslant j \leqslant n} \|\partial_\xi^{(j)} \hat{\psi}(t)\|_{L^2_\xi} \lesssim_n \sum_{0 \leqslant j \leqslant n} \|\partial_\xi^{(j)} \hat{\psi}(0)\|_{L^2_\xi}.$$

for $0 \leqslant t \leqslant 1$. The desired statement then follows.

Step 2. Next, we upgrade the special case in Step 1 to the general case using the Lorentz transformation and scaling.

Let ϕ, x_0, ξ_0 etc. be as in the statement of Proposition 3.5. Without loss of generality, assume that ξ_0 lies on the ξ_1-axis, i.e., $\xi_0 = ((\xi_0)_1, 0, \dots, 0)$. We apply the Lorentz transformation L_v in the (t, x^1)-plane (where $0 \leqslant v \leqslant 1$ will be determined below) and make the change of variables

$$x = L_v \tilde{x} = \left(\frac{\tilde{t} - v\tilde{x}^1}{\sqrt{1 - v^2}}, \frac{\tilde{x}^1 - v\tilde{t}}{\sqrt{1 - v^2}}, x^2, \dots, x^d \right),$$

so that the time interval $0 \leqslant x^0 \leqslant \Delta t$ is mapped to $0 \leqslant \tilde{x}^0 \leqslant \sqrt{1 - v^2}\Delta t$, and the initial localization scale Δx is mapped to $(\frac{1}{\sqrt{1-v^2}}\Delta x^1, \Delta x^2, \dots, \Delta x^d)$. Then we apply scaling

$$\tilde{x} = \mu^{-1} y$$

so that the time interval $0 \leqslant x^0 \leqslant \Delta t$ is mapped to $0 \leqslant y^0 \leqslant \sqrt{1 - v^2}\mu\Delta t$, and the initial localization scale Δx is mapped to $(\frac{1}{\sqrt{1-v^2}}\mu\Delta x^1, \mu\Delta x^2, \dots, \mu\Delta x^d)$. Choosing

$$\frac{1}{\sqrt{1 - v^2}}\mu = 1, \quad \mu = (\Delta t)^{-1/2} \quad \Rightarrow \sqrt{1 - v^2}\mu = (\Delta t)^{-1},$$

the situation is reduced to that treated in Step 1.

3.7. **Optional: An alternative method for the proof of coherence.** It is possible to avoid the use of Lorentz boosts.

When $\xi = (\xi_1, 0, \dots, 0)$, we know that the optimal localization scale is

$$\Delta\xi = (1, (\Delta t)^{-\frac{1}{2}}, \dots, (\Delta t)^{-\frac{1}{2}}),$$

and Δx is dual to $\Delta\xi$. However, due to the degeneracy in the radial (or ξ_1-) direction, the physical space localization is a bit tricky to propagate.

The problem is simplified if we instead work with the localization scale

$$\Delta\xi = ((\Delta t)^{-\frac{1}{2}}, (\Delta t)^{-\frac{1}{2}}, \dots, (\Delta t)^{-\frac{1}{2}}).$$

If we work with wave packets with such a localization scale, then we get an overlap of $O(t^{1/2})$ wave packets at time t in the proof of the dispersive inequality, but that is exactly compensated by the fact that each wave packet is smaller by a factor of $O(t^{-1/2})$.

Exercise 3.8. Give an alternative proof of the dispersive inequality for the fractional Schrödinger equation with $\alpha > 1$ (see Exercise 1.4) using the wave packet approach.

UC Berkeley, Berkeley, CA, USA 94720
Email address: sjoh@math.berkeley.edu