Topics in Nonlinear Dispersive PDE

Introduction (January 19, 2016)

This course (and consequently these notes) will focus on Nonlinear Dispersive PDE. The prototypical example we will consider is the Nonlinear Schrödinger equation (NLS):

$$i\partial_t u + \Delta u = |u|^{p-1}u, \quad p \geq 2, \tag{1}$$

where $u : \mathbb{R}^d \times \mathbb{R} \to \mathbb{C}$.

The Schrödinger equation models the state of a quantum system over time, in general. Recall the Linear Schrödinger equation,

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

where we have already deviated from a strict physical derivation which would yield constants such as

$$i\partial_t u + \frac{\hbar}{2m}\Delta u = 0,$$

which describes a freely moving quantum particle of mass m. Since our aim is in demonstrating various analytic estimates, abstracting this constant away (by, for instance, choosing appropriate units) poses no concern.

Our main goal will be to understand the dynimics of u as time evolves. In order to do that we will first develop a local in time theory for the NLS and then a global in time theory, as well. Much of our discussion will use techniques from Harmonic Analysis. As motivation (and a reminder) for why this is useful, recall that taking the Fourier Transform of the Linear Schrödinger equation yields

$$i\partial_t u + \Delta u = 0 \rightsquigarrow i\partial_t \widehat{u} - |\xi|^2 \widehat{u} = 0$$

which we can solve as an ODE to find $\widehat{u} = Ce^{i|\xi|^2 t}$. Provided with an appropriate initial condition, this is effectively solved. This is typical of the general case with linear PDEs - taking the Fourier Transform converts the PDE to an ODE (or even an algebraic equation) which we can solve and convert back.

Informal definition of dispersion

Dispersion reffers to the fact that solutions u, corresponding to different frequencies ξ, will tend to propagate at different velocities, thus dispersing the solution over time. This behavior is in contrast to, for example, the Transport equation where solutions of all frequencies travel with the same velocity, or dissipateive equations (eg. Heat equation) where solutions of any frequency do not propagate but instead decay in time toward zero.

Dispersion relation

A plane wave described by $e^{i(\xi x + \omega t)}$ propagates with velocity $-\omega/\xi$ (phase velocity). In order for the plane wave to be a solution of the differential equation at hand, the angular frequency ω has to be a function of the (spatial) frequency ξ, i.e. $\omega = h(\xi)$, which depends, each time, on the differential equation.

The function $h(\xi)$ is called the *dispersion relation* [1] of the equation, and its negative gradient defines the *group velocity* $v_g = -\nabla h(\xi)$. A full discussion of all the above is forthcoming, but for now we record a few facts which make our loose description of dispersion a little more tangible.

[1] The discussion here is intentionally vague at this point, and in the examples that follow we are not claiming to give a proof or explanation of the various dispersion relations as stated. A good reference for more information is the Dispersion Wiki: `http://wiki.math.toronto.edu/DispersiveWiki/index.php/Dispersion_relation`

Examples of Dispersion

Example 1. (Transport Equation)

$$\begin{cases} \partial_t u + \vec{a} \cdot \nabla u = 0 \\ u(x,0) = u_0(x) \end{cases}$$

where $u : \mathbb{R}^d \times (0,\infty) \to \mathbb{R}$, $u_0 : \mathbb{R}^d \to \mathbb{R}$ and $\vec{a} \in \mathbb{R}^d$. Recall the solution (found, for instance, by the method of characteristics) is $u(x,t) = u_0(x - \vec{a}t)$ and therefore the initial data is transported with velocity $\vec{a}$. The dispersion relation is $h(\xi) = -\vec{a}\xi$, and the group velocity $-\nabla h(\xi) = \vec{a}$ is constant, independent of ξ. Thus, the Transport Equation is not dispersive.

Example 2. (Linear Schrödinger Equation)

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

We will show that the dispersion relation is $h(\xi) = -|\xi|^2$, and the group velocity is $-\nabla h(\xi) = 2\xi$.

Example 3. (Airy Equation)

$$u_t + u_{xxx} = 0, \quad \text{where } u : \mathbb{R} \times \mathbb{R} \to \mathbb{R}$$

The dispersion relation is $h(\xi) = \xi^3$, and the group velocity is $-\nabla h(\xi) = -3\xi^2$.This is fully dispersive (that is, as dispersive as Shrödinger). The Airy Equation is also (along with NLS and the Wave equation) Hamiltonian.

The Airy equation is the linear part of the Korteweg-de Vries (KdV) equation,

$$u_t + u_{xxx} = 6uu_x,$$

which is also dispersive.

Example 4. (Wave Equation)

$$u_{tt} - \Delta u = 0, \quad \text{where } u : \mathbb{R}^d \times \mathbb{R} \to \mathbb{R}$$

The presence of the u_{tt} term makes this a little more complicated, but one can express this as a first order system and find that the dispersion relation is $h(\xi) = \pm|\xi|$, yielding the group velocity $-\nabla h(\xi) = \mp\frac{\xi}{|\xi|}$. This case is what we call *partially* dispersive, because the frequency determines the direction of the propagation but not the speed (all frequencies travel with speed 1).

Plan

Our focus will be on Scrödinger in $\mathbb{R}^d$. The first step is to understand the analytic properties of the solutions to the linear problem. To do so, we need a few tools:

Strichartz Estimates will allow us to measure dispersion over time. We will develop these through a bootstrapping proceedure:

1. Estimates for the linear homogeneous problem.
2. Estimates for the linear nonhomogeneous problem.
3. Combine the above two results to get Strichartz estimates in general for any solution to NLS.

Duhamel's Principle loosely states that the solution to a general nonhomogeneous problem is a superposition of solutions of the homogeneous problem with nonhomogeneous boundary condition.

The Banach Fixed Point Argument (or contraction mapping theorem) plays a similar role to Picard iteration from ODEs.

Duhamel's Principle

For L a linear partial differential operator not depending on t, the solution of

$$\begin{cases} u_t + Lu = f & \text{in } \mathbb{R}^d \times (0,\infty), \\ u(x,0) = g(x) & \text{on } \mathbb{R}^d \times \{t=0\}, \end{cases} \tag{2}$$

A good example to keep in mind is $L = \Delta$. An even simpler example which will prove quite useful is $L = c$, so that our PDEs become linear first order ODEs. This will motivate the notation to follow.

is a sum of solutions to

$$\begin{cases} v_t + Lv = 0 & \text{in } \mathbb{R}^d \times (0,\infty), \\ v(x,0) = g(x) & \text{on } \mathbb{R}^d \times \{t=0\}, \end{cases} \tag{3}$$

and

$$\begin{cases} w_t + Lw = f & \text{in } \mathbb{R}^d \times (0,\infty), \\ w(x,0) = 0 & \text{on } \mathbb{R}^d \times \{t=0\}. \end{cases} \tag{4}$$

In the notation of semigroups, we write the solution of (3) as

$$v(x,t) = e^{-tL} g(x),$$

where e^{-tL} is the solution operator.

Example 5. If $L = \Delta$, then (3) is the heat equation, and so

$$e^{-t\Delta} g = \frac{1}{(4\pi t)^{n/2}} \int_{\mathbb{R}^d} e^{-\frac{|x-y|^2}{4t}} g(y)\, dy.$$

Duhamel's principle states that (4) is solved by a *continuous superposition* of solutions of

$$\begin{cases} v_t + Lv = 0 & \text{in } \mathbb{R}^d \times (s, \infty), \\ v(x,s) = f(x,s) & \text{on } \mathbb{R}^d \times \{t = s\}, \end{cases} \tag{5}$$

which, after translation in time, is of exactly the same form as (3).

That is, the solution of (4) has the form

$$w(x,t) = \int_0^t e^{(s-t)L} f(x,s)\, ds. \tag{6}$$

Note that we could also have interpreted this as convolution in time with the solution operator e^{-tl}.

Hence the solution of our original PDE (2) is

$$u(x,t) = e^{-tL} g(x) + \int_0^t e^{(s-t)L} f(x,s)\, ds.$$

Compare this to the situation where $L = c$, and our PDE (2) was really a linear first-order ODE. It's unfortunate that the similarity doesn't stretch to even the mild case of $L = F(t)$, but c'est la vie.

Fourier Transform (January 21, 2016)

We begin with some background on the Fourier Transform, which is most naturally and easily developed on the space of Schwartz functions, where questions of integral convergence are completely nullified because the functions decay so rapidly.

Intuitively, a Schwartz function is a function whose derivatives decrease more rapidly than any polynomial.

Definition 1. The *Schwartz space* (on $\mathbb{R}^d$) is the space of rapidly decreasing functions

$$\mathcal{S}(\mathbb{R}^d) := \{f \in C^\infty(\mathbb{R}^d \to \mathbb{C}) : \|f\|_{\alpha,\beta} < \infty \forall \alpha, \beta\}$$

where α, β are multiindices[2] , and

$$\|f\|_{\alpha,\beta} = \sup_{x\in\mathbb{R}^d} |x^\alpha D^\beta f(x)|.$$

[2] A multiindex is a vector of the form

$$\alpha = (\alpha_1, \ldots, \alpha_d) \in \mathbb{N}^d.$$

The order of the multiindex is given by the ℓ_1 norm,

$$|\alpha| = \alpha_1 + \cdots + \alpha_d.$$

We define

$$x^\alpha := x_1^{\alpha_1} \cdots x_d^{\alpha_d}$$

and

$$D^\alpha f(x) := \frac{\partial^{|\alpha|} f(x)}{\partial x_1^{\alpha_1} \cdots \partial x_d^{\alpha_d}}.$$

Schwartz functions are dense in any L^p space, and also in any Sobolev space. The prototypical example of a Schwartz function is $e^{-|x|^2}$. More generally, $x^\alpha e^{-b|x|^2}$ is Schwartz for any multiindex α and any positive real number b. Note that the space of compactly supported smooth functions $C_c^\infty(\mathbb{R}^d)$ is contained in $S(\mathbb{R}^d)$.

Definition 2. Given $f \in \mathcal{S}(\mathbb{R}^d)$, the Fourier Transform of f is a function $\widehat{f} : \mathbb{R}^d \to \mathbb{C}$ defined as

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

There are many different ways of adding constants to the definition of the Fourier Transform, but they all share many of the same properties and for our purposes this is the most useful.

We recall the following properties of the Fourier Transform.

Theorem 1. *(Properties of Fourier Transform)*[3]

Given $f, g \in \mathcal{S}(\mathbb{R}^d)$, $x_o, \xi_0 \in \mathbb{R}^d$, $\lambda > 0$, and α a multiindex, we have

[3] See, for instance, *Classical Fourier Analysis* by Grafakos.

	Function	*Fourier Transform*
(a)	$f(x - x_o)$	$\widehat{f}(\xi) e^{-ix_0\cdot\xi}$
(b)	$f(x) e^{ix\cdot\xi_0}$	$\widehat{f}(\xi - \xi_0)$
(c)	$\overline{f}(x)$	$\overline{\widehat{f}}(-\xi)$
(d)	$f(\lambda x)$	$\frac{1}{\lambda^d}\widehat{f}(\frac{\xi}{\lambda})$
(e)	$f * g(x)$	$\widehat{f}(\xi)\widehat{g}(\xi)$
(f)	$f(x)g(x)$	$\widehat{f} * \widehat{g}(\xi)$
(g)	$D^\alpha f(x)$	$(i\xi)^\alpha \widehat{f}(\xi)$
(h)	$\nabla f(x)$	$i\xi\widehat{f}(\xi)$

Proof. The proofs of (a) - (d) are straightforward from the definition via change of variables.

For (e) and (f), recall the definition of convolution:

$$f * g(x) := \int_{\mathbb{R}^d} f(x-y)g(y)\,dy = \int_{\mathbb{R}^d} f(y)g(x-y)\,dy.$$

Note that $f * g$ is a global operation (it requires knowledge of the value of f and g on all of $\mathbb{R}^d$), while pointwise multiplication is local. The global nature of convolution is encapsulated in the Fourier Transform (itself a global operation).

The fact that $\widehat{f * g}(\xi) = \widehat{f} * \widehat{g}(\xi)$ in (d) is then a straightforward verification from the definition, and (e) follows similarly.

Finally, (g) and (h) follow from integration by parts. □

Definition 3. For $1 \leq p < \infty$, the L^p norm of a function is

$$\|f\|_{L^p(\mathbb{R}^d)} := \left(\int_{\mathbb{R}^d} |f(x)|^p\,dx\right)^{\frac{1}{p}},$$

and the L^∞ norm is

$$\|f\|_{L^\infty} := \operatorname{ess\,sup} |f(x)| = \inf\{\alpha \mid \{|f(x)| > \alpha\} \text{has measure zero}\}.$$

We think of ess sup as the supremum, modulo null sets.

Note that $L^1(\mathbb{R}^d)$ and $L^\infty(\mathbb{R}^d)$ are special, in the sense that they often are excluded or handled separately from theorems involving L^p spaces. The next theorem is an example of this. Recall that the dual space (denoted with a $*$) is the Banach space of all continuous linear functionals.

Theorem 2. *(Duality) $L^p(\mathbb{R}^d)^* \cong L^{p'}(\mathbb{R}^d)$ for $1 < p < \infty$, where p' is Hölder dual to p, i.e. $\frac{1}{p} + \frac{1}{p'} = 1$.*

Proof. The isomorphism[4] is given explicitly for $g \in \mathcal{L}^{p'}$ as

$$g \mapsto \left[f \mapsto \int f(x)g(x)\,dx\right].$$

[4] It can actually be shown that this is an isometry.

The fact that this satisfies the requirements follows from Hölder's inequality and the Radon-Nikodym theorem.[5] □

[5] See `https://en.wikipedia.org/wiki/Lp_space#Dual_spaces`

This duality theorem does not extend completely to 1 and ∞, in particular $L^1(\mathbb{R}^d)^* = L^\infty(\mathbb{R}^d)$, but $L^\infty(\mathbb{R}^d)^* \neq L^1(\mathbb{R}^d)$.[6] As another example, the Hardy-Littlewood maximal function

[6] For a counterexample, and much more, see Folland, *Real Analysis*, §6.2.

$$Mf(x) = \sup_{r>0} \frac{1}{|B(x,r)|} \int_{B(x,r)} |f(y)|\,dy$$

takes L^p into itself for $p > 1$, but takes L^1 into $L^{1,\infty}$ (that is, L^1-weak).

We note that the Fourier Transform $\mathcal{F}$ is an operator on $\mathcal{S}$ and also on $\mathcal{S}' := \mathcal{S}^*$, the space of tempered distributions. In fact the Fourier

Transform of an L^1 function is a uniformly continuous function. We recall the following standard results regarding the Fourier Transform without proof.[7]

[7] References include *Fourier Analysis* by Stein & Shakarchi, *Real Analysis* by Folland, and *Classical Fourier Analysis* by Grafakos.

Recall that $L^2(\mathbb{R}^d)$ is a Hilbert space, with inner product

$$\langle f, g \rangle = \int_{\mathbb{R}^d} f(x)\overline{g(x)}\, dx.$$

Theorem 3. *(Plancherel) If $f, g \in L^2(\mathbb{R}^d)$,*

$$\langle f, g \rangle = \int_{\mathbb{R}^d} f(x)\overline{g(x)}\, dx = \int_{\mathbb{R}^d} \widehat{f}(\xi)\overline{\widehat{g}(\xi)}\, d\xi = \langle \widehat{f}, \widehat{g} \rangle.$$

Corollary 1. *The Fourier Transform $\mathcal{F}$ is an operator on L^2. In fact,*

$$\|f\|_{L^2} = \|\widehat{f}\|_{L^2} \qquad \text{(Parseval's Identity)}$$

Theorem 4. *(Inversion) The Fourier Transform $\mathcal{F} : \mathcal{S} \to \mathcal{S}$ is invertible, and*

$$f(x) = \int_{\mathbb{R}^d} \widehat{f}(\xi) e^{ix\cdot\xi}\, d\xi.$$

We will denote the inverse Fourier Transform as

$$\mathcal{F}^{-1}(f(\xi))(x) = \check{f}(x).$$

Littlewood-Paley Theory

Definition 4. A *Fourier multiplier* is a function $m : \mathbb{R}^d \to \mathbb{C}$ which is locally integrable and has at most polynomial growth which allows us to define an operator T_m on $\mathcal{S}(\mathbb{R}^d)$ via

Multipliers give rise to operators. Here it is natural to establish the definition on the Fourier side first, and pull back to the spatial side. This flexibility will become natural.

$$\widehat{T_m f}(\xi) = m(\xi)\widehat{f}(\xi).$$

That is,

$$T_m f(x) = \widecheck{m\widehat{f}}(x),$$

therefore

$$T_m f(x) = \int_{\mathbb{R}^d} m(\xi)\widehat{f}(\xi) e^{ix\cdot\xi}\, d\xi.$$

Theorem 5. *If $m \in L^\infty$, $T_m : L^2 \to L^2$.*

It is easiest to work with the operator on the Fourier side. This gives us a nice representation to play with. (Operators do not, in general, have this property.) A special class of multipliers which arise in Littlewood-Paley theorey are appropriately known as Littlewood-Paley (LP) multipliers. The idea is to find a multiplier m which is a bump function, and so that

$m(\xi) = 1$ on $B(0,r)$ for some $r \in (0,1)$,

$m(\xi) = 0$ on $\mathbb{R}^d \setminus B(0,2)$,

$\check{m} = \varphi \in L^1(\mathbb{R}^d) \cap \mathcal{S}(\mathbb{R}^d)$,

$\int_{\mathbb{R}^d} \varphi(x)\,dx = 1$,

and m is compactly supported.

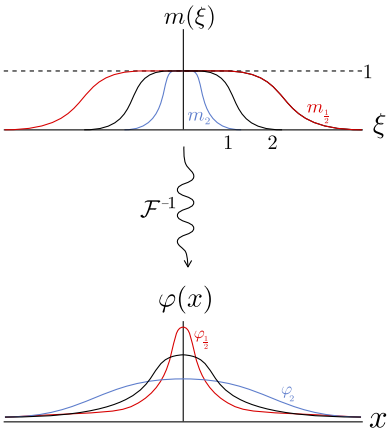

Figure 1: An example LP multiplier

Such a function will also have the following dilation properties:

As $\lambda \to 0$, $m(\lambda\xi) \to 1$.

As $\lambda \to \infty$, $m(\lambda\xi) \to \chi_{\{0\}}$.

So, an LP multiplier is a radially symmetric bump function $m(\xi)$, adapted to the ball $B(0,2)$, such that $m(\xi) = 1$ on a smaller ball $B(0,r)$, where, for instance $r < 1$. The exact choice of m is not relevant, but we can always choose m so that $\check{m} = \varphi \geq 0$, $\varphi \in L^1$, and $\int_{\mathbb{R}^d} \varphi(x)\,dx = 1$.

If $m_\lambda(\xi) := m(\lambda\xi)$, we have that

$$\check{m}_\lambda(x) = \frac{1}{\lambda^d}\varphi\left(\frac{x}{\lambda}\right) =: \varphi_\lambda(x)$$

so $\|\varphi_\lambda\|_{L^1} = \|\varphi\|_{L^1}$.

The φ function we are building is known as an approximation to the identity. The prototypical example is the Poisson kernel:

$$P(x) = \frac{c_d}{(1+|x|^2)^{\frac{d+1}{2}}},$$

where c_d is a constant such that $\int_{\mathbb{R}^d} P(x)\,dx = 1$. If $d = 1$, we find $c_1 = 1$.

With this LP multiplier in hand, we can proceed to describe LP theory. Note that if $\widehat{Rf}(\xi) := m(\xi)\widehat{f}(\xi)$, then $Rf(x) = \varphi * f(x)$, which is essentially the average of f on a ball. (See Figures 2 and 3, which demonstrate how convolving with an approximation to the identity essentially blurs the function.) If $\widehat{R_\lambda f}(\xi) = m_\lambda(\xi)\widehat{f}(\xi)$, then $R_\lambda f(x) = \varphi_\lambda * f(x)$.

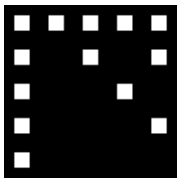

Figure 2: Simple Function

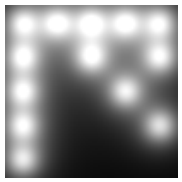

Figure 3: Convolution with Poisson Kernel

We will choose dyadic λ, i.e. $\lambda \sim 2^j$ for $j \in \mathbb{Z}$. Instead of λ, we'll call these dyadic scaling parameters N, so $N = 2^j$ for some $j \in \mathbb{Z}$. For each such N, define the following operators:

$$\widehat{P_{\leq N} f}(\xi) := m\left(\frac{\xi}{N}\right)\widehat{f}(\xi) \tag{7}$$

$$\widehat{P_N f}(\xi) := \left(m\left(\frac{\xi}{N}\right) - m\left(\frac{\xi}{N/2}\right)\right)\widehat{f}(\xi) \tag{8}$$

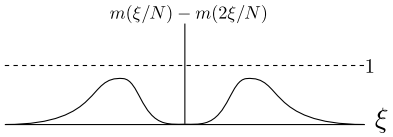

Figure 4: LP multiplier subtraction, $\widehat{\psi}_N$

Defining $\widehat{\psi}_N := m\left(\frac{\xi}{N}\right) - m\left(\frac{2\xi}{N}\right)$, we have the important result that

$$P_N f = \psi_N * f, \quad \text{where } \psi_N \in L^1 \text{ and } \int_{\mathbb{R}^d} \psi_N(x)\,dx = 0,$$

since $\widehat{\psi}_N(0) = 0$. Furthermore, we have that in the L^2-sense, $f = \sum_N P_N f$. If $f \in \mathcal{S}(\mathbb{R}^d)$, then this equality even holds pointwise.

Here we've used a slightly less familiar property of the Fourier Transform, $\int_{\mathbb{R}^d} \psi(x)\,dx = \int_{\mathbb{R}^d} \psi(x)e^{ix\cdot 0}\,dx = \widehat{\psi}(0)$.

We think of P_N as the projection operator that looks at what happens to f at frequencies of order $|\xi| \sim N$.