An Introduction to Dispersive Equations

Lecture Notes given at Michigan State University

Willie Wai-Yeung Wong

Revision based on commit c81ba63 (HEAD -> master, origin/master, origin/HEAD) as of January 10, 2019

Preface

This document is currently being prepared for a second-year graduate level PDE course offered at Michigan State University. The goal is to offer a one semester, conceptual overview of some of the issues and techniques in the study of dispersive equations. The intended audience for these notes are advanced undergraduate students and beginning graduate students. The only real assumptions made in terms of prerequisite preparation are a firm grasp of multivariable calculus and basic real analysis. However, the firmer the students' background is in analysis the more benefit they can derive from these notes. So while not strictly necessary, some familiarity with Fourier theory, measure theory, and Hilbert space theory (as, for example, expounded in volumes one and three of the textbooks by Eli Stein and Rami Shakarchi) will be immensely helpful. Indeed, seeing as the mainstay of the classical techniques for analyzing linear dispersive equations is that of oscillatory integrals, large portions of these notes can be aptly subtitled "applications of Fourier analysis". In terms of scope, this document is truly at the level of an introduction. In the presentation explicit computations are preferred over far-reaching general theorems. And while some nonlinear theory and applications will be introduced, the selected material represents only one corner of the recent developments in the field. The hope is that some of the discussions in these notes will pique the interests of students, and lead them to further study in this direction.

Stein and Shakarchi, Fourier analysis; *Stein and Shakarchi*, Real analysis

The subject of these notes are *dispersive equations*. A first question one may ask is: "What is dispersion?" The dictionary definition says that to disperse is to spread out over a wide area. In physics, and, by extension, in the study of partial differential equations, dispersion refers to the specific phenomenon that a collection of "particles" travelling at different speeds will tend to spread out. The word "particles" in the previous sentence was left in quotes because the origins of this terminology lie in the study of optics from the 18th and 19th century (to refer to, for example, the emittance of a rainbow of colors when white light shines through a prism), and so from the very get-go the phenomenon was understood with "wave packets" playing the role of "particles". With the great hindsight available to us from quantum theory, however, we will begin these notes by tackling the problem of the dynamics of collections of particles, and later on moving toward the equations governing continuous fields. The initial focus on the particle picture is not accidental: the modern development of the study of dispersive equations relies a lot on analyses performed on "phase space". The particle picture where the phase space is *commutative* forms a good starting point before tackling the "quantum" theory where the phase space is not commutative (Heisenberg's uncertainty principle). With in mind these connections both physical and mathematical, the beginning chapter of this notes makes some short detours into the realms of basic theoretical physics, and throughout appeals will be made toward "physical intuition" whenever available.

Of course, this description is still anachronistic. Newton's original theory of optics, with which he gave an explanation of refraction (the mechanism behind the splitting of white light by triangular prisms), is corpuscular. *While Huygens had advocated, around the same time, his wave theory, the nature of light as waves was not fully appreciated until Young's interference experiments a century later.*

The second and third chapters develop the main Fourier analytic tools used for the classical study of dispersive equations. The importance of Fourier analysis can be seen through the quantization process moving from classical to quantum mechanics; and from this point of view the properties of solutions to dispersive equations is tied to the geometry of the corresponding Fresnel surface. As is well-known the principle curvatures of this surface play important roles in determining both the long-time decay behavior of solutions, as well as their short-time integrability. These are usually studied through oscillatory integral techniques which will be described in chapter 3 of these notes. Some of the dispersive decay estimates can also be recovered through purely physical space arguments based on considering the inherent symmetries of the partial differential equations. We develop these "vector field" type methods in chapter 5. The application of this method to the wave and Klein-Gordon equations are fairly well known and originates in the works of Sergiu Klainerman in the 1980s. The subsequent development to other models are more recent, with contributions by David Fajman, Jérémie Joudioux, Jacques Smulevici, and (separately) the author all in the mid 2010s.

Fresnel surface: also called the wave surface, it is the hypersurface in the space-time frequency space defined by the dispersion relation of the physical model.

One of the main results in the study of linear dispersive equations is that of *Strichartz estimates*. These estimates can be regarded as dual estimates to the Fourier restriction estimates and manifests in the control of space-time integrals of the solutions based on their initial data. Modern proofs of Strichartz estimates use the TT^* method and a healthy dose of interpolation theory. The relevant material are developed in chapter 4.

In chapter 6, the material developed thus far are applied to solve some basic problems in local and global wellposedness of the initial value problems for nonlinear dispersive equations. The results here are far from the cutting edge, and are presented mainly as a tool to introduce the notion of wellposedness and expose the students to some of the basic results in the field. Students interested in further study of nonlinear dispersive equations should consult the lecture notes of Terence Tao for development up to the early 2000s. The Oberwolfach Seminars of Herbert Koch, Daniel Tataru, and Monica Vişan contain some further results through 2012. More recent results are typically only found in the original research articles.

Tao, Nonlinear dispersive equations: local and global analysis

Koch et al., Dispersive Equations and Nonlinear Waves

If you have any suggestions for improvements, questions on the material, or general comments, please feel free to e-mail me at `wongwwy@member.ams.org`. I would also love to hear from you if you've found these notes useful in anyway, whether for your own studies, for research, or for teaching.

Design philosophy and TEXnical details

The present document is prepared and typeset using LaTeX 2_ε. The document class used is a custom class called `wwwnotes3` built over the standard `report` class; you can find the source code at the Git repository `https://gitlab.msu.edu/wongwil2/www-textools`. The font used is from the Johannes Kepler project, accessible as the package `kpfonts` on CTAN.

The layout of the pages is heavily inspired by the works of Edward Tufte. In particular are the use of a wide margin and side notes instead of footnotes, the limit to *two* levels of topic headings (Chapter and Section in this document), and the citable "thought units" (in this document paragraphs delimited by a number at the start and the symbol ¶ at the end). The following paragraphs contain some tips on how to effectively use this document.

0.1 (Cross referencing) Every item that can be cross-referenced (such as this one) is labelled `chapter#.item#`. The item number increases monotonically throughout the chapter. At the top of every page you can find, similar to what appears in dictionaries, a numeric range; this is designed to help you locate antecedents of references. The notation "(Prev. ref. #)" indicates that there are no new items defined on the current page, and the # shows the most recently defined item number.

Yes, I know this is slightly buggy at the moment with some "off-by-one" errors. But it mostly works as indicated.

The items that can be cross-referenced are: equations, theorem-like assertions, conjecture-like queries, and thoughts (such as this one). Occasionally the references are found together with a note reminding you what the antecedent is. For example, Thought 0.1 refers to this very item. And in the margin you will see a short description of the reference.

Ref. 0.1: "Concerning cross references"

The items that can be cross-referenced all have fixed scope; that is, they have a beginning and an ending. The equations are easy to identify. For the other three types, they all begin with the item number and end with a 'mark'. For "assertions" the mark is '■'. For "queries" the mark is '◇'. And for "thoughts" the mark is '¶'. Since this is a mathematical text, there are also "proofs", whose ends are denoted by '□'. ¶

"Assertions" refer to definite statements such as Theorems or Definitions which may or may not be justified. "queries" refer to tentative statements such as Conjectures, and "thoughts" refer to a block of text, potentially consisting of several paragraphs, following a coherent idea.

0.2 (Margin notes) This document makes extensive use of margin notes. The guiding principle of margin notes is that they provide helpful, but optional annotation to the running text. In particular, their removal should not be detrimental to the understanding of the text proper. Therefore, there will not be any footnote in the traditional sense (where the flow of reading is interrupted by the appearance of a superscript, with the reader compelled to certain distraction by tangential remarks); if something is important enough to grab the attention of the reader, it should appear in the text proper. Marginalia will largely comprise historical and tangential remarks, as well as suggestions for further reading. (In addition to the aide to deciphering cross references described in Thought 0.1.) The reader should feel free to ignore them all. Additionally, the author hopes that active readers of mathematical treatises may find beneficial the ample margins. ¶

Ref. 0.1: "Concerning cross references"

0.3 (Citations) Since these are lecture notes, in-line citations will not be given. Suggestions for further reading, as well as discussion of the history of progression of a particular result, can all be found in the margin notes insofar as they appear. The list of further readings are also collected at the end of this document. ¶

Acknowledgements

I thank the students of my MTH943 class in Spring of 2017 for their patience with the development of these notes: these include from Michigan State, Leo Abbrescia, Anqi Chen, Michael Gundlach, Seonghyeon Jeong, Rodrigo Matos, Woong Bae Park, Wenchuan Tian, and Xin Yang; and from Xi'an Jiaotong University, Rulin Liu, Zeyu Liu, Zhixin Wang, and Yizhen Zhao. For the development of the LaTeX 2_ε code doing a lot of the heavy-lifting of the formatting of this manuscript, I give my thanks to the help

I received and the tricks I learned from the community at `http://tex.stackexchange.com`.

I would also like to thank Pin Yu and the Yau Mathematical Sciences Center at Tsinghua University (Beijing) for their hospitality. The current presentation of chapter 5 is based on a mini-course I gave at the YMSC in the Summer of 2017.

Contents

List of exercises

Chapter 1

Dispersion: A First Look

Dispersion is most easily understood in terms of collections of particles. In this chapter we will start from this point of view and derive some basic consequences thereof. We will then switch gears and discuss *linear* dispersive equations, especially the physical intuitions for the phenomena exhibited by wave-like theories of continuous fields. We will introduce our main examples in this chapter, while leaving their analysis to subsequent ones.

The story of N particles

Consider N particles moving on $\mathbb{R}^d$; for simplicity we will assume that the particles don't interact with each other. Then each particle is described by its position $x_i : \mathbb{R} \to \mathbb{R}^d$ (with $i \in \{1, \ldots, N\}$) as a function of time. By Newton's first law the particles travel on straight lines, so we can write

$$x_i(t) = x_i(0) + v_i t \tag{1.1}$$

where v_i is the (constant-in-time) velocity of the ith particle. We have very simple upper and lower bounds for the distances. Using that

$$d_{ij}(t) \stackrel{\text{def}}{=} x_i(t) - x_j(t) = [x_i(0) - x_j(0)] + (v_i - v_j)t,$$

we can bound

$$t \cdot \left|v_i - v_j\right| - \left|d_{ij}(0)\right| \leq \left|d_{ij}(t)\right| \leq t \cdot \left|v_i - v_j\right| + \left|d_{ij}(0)\right|. \tag{1.2}$$

In particular, for *generic* velocities, the distance between any pair of particles is approximately a linear function of time.

The equation (1.2) captures the fundamental idea of behind *dispersion*. From it we can derive some immediate consequences.

1.3 Exercise
What are the limits

$$\lim_{t\to\pm\infty} \frac{\max_{i,j}\left|d_{ij}(t)\right|}{t}, \qquad \lim_{t\to\pm\infty} \frac{\min_{i\neq j}\left|d_{i,j}(t)\right|}{t}$$

in terms of the velocities $\{v_i\}$? ■

1.4 Exercise
Consider the function

$$\rho(R,t) = \sup_{y\in\mathbb{R}^d} \#\Big\{i \in \{1,\dots,N\} \;\Big|\; |x_i(t) - y| < R\Big\},$$

which can be described as the maximum number density *at scale* R and time t.

1. Show that, as long as $N > 0$, for any $R > 0$ and $t \in \mathbb{R}$ it holds that $\rho(R,t) \geq 1$.
2. Show that, for *generic* initial data, for any $R \in (0,\infty)$ the density converges to 1 as $t \to \pm\infty$. More precisely, show that whenever R, t satisfies
$$|t| > \frac{\max_{i,j}\left|x_i(0) - x_j(0)\right| + 2R}{\min_{i\neq j}\left|v_i - v_j\right|}$$
we have $\rho(R,t) = 1$. ■

The estimate in the second part of the previous exercise can be refined. First, given $t_1, t_2 \in \mathbb{R}$ and $\Omega_1, \Omega_2 \subset \mathbb{R}^d$, we can let

$$X(t_1,\Omega_1,t_2,\Omega_2) = \Big\{i \in \{1,\dots,N\} \;\Big|\; x_i(t_1) \in \Omega_1 \text{ and } x_i(t_2) \in \Omega_2\Big\}$$

be the set of particles that traverses from Ω_1 at time t_1 to Ω_2 at time t_2. Then clearly $i \in X$ requires $v_i \in \frac{1}{t_2 - t_1}(\Omega_2 - \Omega_1)$ (with element-wise subtraction). Now fix $t_1 = 0$ and Ω_1 any set that contains in the initial $\{x_i(0)\}$. Running Ω_2 over all balls of radius R at some late time t_2, we see that the density of the distribution of particles can be directly controlled by the *velocity density* of the initial distribution. This naturally leads us to the formulation in terms of *kinetic theory*.

Kinetic theory

Kinetic theory is an intermediate description of molecular motion between the Newtonian particle picture and the continuum fluidic approximation. In kinetic theory, instead of considering individual particles, we consider the *distribution* of particles in phase space and its associated evolution. Supposing the physical space has dimension d, Newton's laws of motion tell us that the trajectories of particles are determined by their instantaneous position and velocity (being a second order differential equation). So we can take as our phase space for the particles $\mathbb{R}^d \times \mathbb{R}^d$. A distribution of particles, then, is a function

$$\rho : \mathbb{R} \times \mathbb{R}^d \times \mathbb{R}^d \to \overline{\mathbb{R}_+}, \tag{1.5}$$

representing the density (over phase space) of the particles at a particular position x with a particular velocity v at time t. It takes non-negative values as we don't allow negative number of particles.

The study of kinetic theory proper captures the Newtonian interaction of particles by collision by some multilinear integral operator acting on ρ. We, however, will again take the (unphysical) simplifying assumption that *individual particles don't interact.* This reduces the equation of motion to Newton's first law:

See the vast literature on the Boltzmann and Landau equations.

$$\partial_t \rho(t,x,v) + v \cdot \nabla^{(x)} \rho(t,x,v) = 0. \tag{1.6}$$

The symbol $\nabla^{(x)}$ is the *gradient* with respect to the spatial coordinates x; that $\nabla^{(x)}\rho$ is a $\mathbb{R}^d$ valued function whose ith component is $\frac{\partial}{\partial x^i}\rho$. The equation (1.6) is in conservative form, and states simply that along the straight-line trajectories $x + vt$, the particle density is constant (if a particle starts at position x and moves with velocity v, then it will remain always on the line $x + vt$). This equation is sometimes called the *Vlasov equation* or simply the *linear transport equation*.

The particle interpretation allows us to write down explicitly the solution to this first order scalar partial differential equation. Since the density remains constant on the line $x + vt$, we have that

This is in fact nothing more than the "method of characteristics".

$$\rho(t,x,v) = \rho(0, x - tv, v). \tag{1.7}$$

The explicit solution (1.7) allows us to write down our first "dispersive estimate". First, thinking back to the N-particle picture, what we are interested in is the physical space density given by

$$\overline{\rho}(t,x) \stackrel{\text{def}}{=} \int_{\mathbb{R}^d} \rho(t,x,v)\,\mathrm{d}v \tag{1.8}$$

where we summed up all the particles (with different velocities) located at the point x. From the N-particle picture, we expect this physical space density to decay.

1.9 Remark
The fact that we have integrated in v is important here. The Vlasov equation (1.6) acts by linear transport, and one can easily check that for any $p \in [1,\infty]$, the norm $\left\|\rho(t,\bullet,\bullet)\right\|_{L^p(\mathbb{R}^d\times\mathbb{R}^d)}$ is constant in time. What we assert in the previous paragraph, however, is that the norm $\|\overline{\rho}(t,\bullet)\|_{L^p(\mathbb{R}^d)}$ for $p > 1$ decays as a function of time. ■

1.10 Theorem (Dispersive estimate for Vlasov equation, version 1)

See the next chapter for a discussion of Schwartz space.

Let $\rho_0 \in \mathscr{S}(\mathbb{R}^{2d})$. The solution to (1.6) with initial data $\rho(0,x,v) = \rho_0(x,v)$ satisfies

$$\sup_{x\in\mathbb{R}^d} \overline{\rho}(t,x) \lesssim \langle t\rangle^{-d}. \tag{1.11}$$

■

1.12 Convention ($\lesssim$ and the Japanese bracket)
In the statement of Theorem 1.10, several notational conventions were introduced. We will throughout adopt the convention, which is by now standard in the literature, that

$$A \lesssim B$$

means

$$\exists C \in \mathbb{R}_+ \text{ s.t. } A \leq CB.$$

So the statement in the previous theorem should be read as

> ...with initial data $\rho(0,x,v) = \rho_0(x,v)$ satisfies, for some $C > 0$, the estimate $\sup_{x\in\mathbb{R}^d} \overline{\rho}(t,x) \leq C\langle t\rangle^{-d}$.

The question of "on what does the constant C depend" is generally clear from context. For example, in Theorem 1.10 it is understood that the constant C may depend on ρ_0, but not on t. Sometimes it pays to emphasize the dependence, in which case the notation $\lesssim_{E,F,G}$ will be used where E, F, and G (for example) are quantities which influence the value of C.

Another notation introduced is the *Japanese bracket*, which is defined as

$$\langle t \rangle \stackrel{\text{def}}{=} \sqrt{1+t^2} \tag{1.13}$$

for any real-valued t. ■

Proof (Theorem 1.10) Using (1.7) we can write

$$\int_{\mathbb{R}^d} \rho(t,x,v)\,\mathrm{d}v = \int_{\mathbb{R}^d} \rho_0(x - tv, v)\,\mathrm{d}v.$$

Doing a change of variable with $w = \langle t \rangle v$ we have that

$$= \frac{1}{\langle t \rangle^d} \int_{\mathbb{R}^d} \rho_0\Big(x - \frac{t}{\langle t \rangle} w, \frac{1}{\langle t \rangle} w\Big)\,\mathrm{d}w.$$

The integral is now over the plane

$$\left\{ \Big(x - \frac{t}{\langle t \rangle} w, \frac{1}{\langle t \rangle} w\Big) \in \mathbb{R}^{2d} \,\middle|\, w \in \mathbb{R}^d \right\}$$

with the *induced surface measure*. So, writing Π for the set of all d dimensional affine subspaces of $\mathbb{R}^{2d}$, we have

$$\int_{\mathbb{R}^d} \rho(t,x,v)\,\mathrm{d}v \leq \underbrace{\sup_{P \in \Pi} \int_P \rho_0 \,\mathrm{d}\sigma}_{C}\ \langle t \rangle^{-d},$$

with the constant C clearly independent of t and x. So the estimate is proved.

Note that the norm

$$\rho_0 \mapsto \sup_{P \in \Pi} \int_P \rho_0 \,\mathrm{d}\sigma$$

is the L^1 *trace norm* for restricting a function on $\mathbb{R}^{2d}$ to its d dimensional affine subspaces. By Gagliardo's Sobolev trace theorem, this means that the constant C can also be bounded by the $W^{d,1}(\mathbb{R}^{2d})$ norm of the initial data.□

For more on Sobolev spaces and embedding theorems, a standard reference is Adams and Fournier, Sobolev spaces.

1.14 Exercise
In the proof of Theorem 1.10, we see that the constant C can be determined by the $W^{d,1}$ norm of the initial data. (For completeness, let us recall the definition

$$\left\|\phi\right\|_{W^{k,1}(\mathbb{R}^d\times\mathbb{R}^d)} \stackrel{\text{def}}{=} \sum_{|\alpha|+|\beta|\leq k} \int_{\mathbb{R}^d\times\mathbb{R}^d} \left|\partial_x^\alpha \partial_v^\beta \phi(x,v)\right| \mathrm{d}v\ \mathrm{d}x$$

where α, β are d-dimensional multi-indices.) Show that this estimate is *sharp*. More precisely, show that if $k < d$ is a non-negative integer, then there exists a sequence of initial data $\{\rho_0^{(j)}\}_{j\in\mathbb{N}}$ such that

- $\left\|\rho_0^{(j)}\right\|_{W^{k,1}} \leq 1$ for every j;
- writing $\rho^{(j)}$ for the corresponding solutions to (1.6), there exists some $\epsilon > 0$ fixed such that for every t, there exists some $j \in \mathbb{N}$ satisfying $\sup_x \overline{\rho}^{(j)}(t,x) \geq \epsilon$.

Hint: Think back at the particle picture. What would ρ_0 be for a point particle? ■

1.15 (Some characteristics of dispersive estimates) As we have seen in the preceding discussion, our dispersive estimates controls the decay in time of the L^∞ norm (for the present discussion, ignore the v variable) by the initial data measured with d derivatives in the L^1 norm. This exhibits two characteristics that are common-place for dispersive estimates: first is that L^∞ decay is controlled by measuring the data in L^1, and the second is that there is a definite loss of smoothness (see previous exercise). Both of these have physical explanations.

That the point-wise decay in time is controlled by some integral norm of the initial data arises from the intuition that dispersion occurs because the physical extent of the particles spreads out while the total mass is conserved. With the same mass divided among a greater volume, the spatial density (which is what $\int \rho \ \mathrm{d}v$ measures) must decay. But for this argument to be sensible, the total mass had better be finite: otherwise the particles leaving a spatial domain can be replenished by particles arriving from arbitrarily far away. This would be the case, for example, for a solution of the Vlasov equation that is *spatially homogeneous*.

The same idea of "conserved mass divided among ever-growing volume" also serves to explain the dependence on smoothness. For the intuition to hold, that the pointwise density decays, we need that the conserved mass is approximately evenly distributed among the available volume. If not, one can easily imagine a lopsided distribution where almost all the mass concentrate at one point and decay is not evident. Smoothness of the distribution clearly provides a measure of equidistribution of the mass density. Another way to see this is to recall our discussion in the N particle picture, where we have seen that the rate of spreading out of the particles is proportional to how disparate the particle velocities are. (In fact, the worst case scenario for our dispersive estimate is precisely the particle picture, where all the mass concentrate at discrete points!) For two phase-space distributions, the one that is more concentrated and particle-like will tend to be rougher. And so, as seen in the previous exercise, the loss of smoothness in the estimates are necessary, and reflect our need to distinguish between particle-like initial distributions, and those distributions that are more like a fluid. ¶

1.16 (Methods of proof) The proof we have given above for Theorem 1.10 is based on the fundamental solution (1.7) to the linear Vlasov equation (1.6). This method of proof is relatively direct, and can be easily adapted to deal with *semilinear* type problems (imagine (1.6) but with the right hand side, instead of being zero, being some nonlinear function of ρ) using Duhamel's principle. On the other hand, this method is not-so-stable when the equations are perturbed in a way that changes the principal part (i.e. the part with the highest order derivatives) of the equation. In the remainder of this section we develop an alternative dispersive estimate using the *vector field method*. This method was originally developed in the context of wave equations and has been shown to be more stable for applications in both linear and nonlinear applications where the principal part of the equation differs from the standard linear expressions. The main idea is a careful exploitation of conservation laws to obtain weighted integral estimates. ¶

1.17 (Conservation laws) Since the Vlasov equation (1.6) is in divergence form (by definition $\nabla^{(x)}$ commutes with multiplication by v), we have that for every fixed v the integral

$$\int_{\mathbb{R}^d} \rho(t,x,v)\,\mathrm{d}x = \int_{\mathbb{R}^d} \rho(0,x,v)\,\mathrm{d}x$$

is conserved, provided that it makes sense at $t = 0$. This then implies, in fact, a conservation law for every velocity multiplier: let $g : \mathbb{R}^d \to \mathbb{R}$ be any measurable function, then

$$E[g](t) \stackrel{\text{def}}{=} \int_{\mathbb{R}^{2d}} \rho(t,x,v)g(v)\,\mathrm{d}x\,\mathrm{d}v$$

is constant in time. Recalling that ρ represents the particle density, so the quantity $E[1]$ is the conserved total mass. Next, let $e \in \mathbb{R}^d$ be any *unit vector*, and let $g(v) = e \cdot v$. The quantity $E[g]$ represents the conserved total *linear momentum* in the direction e. If we let $g(v) = \frac{1}{2} v \cdot v$, then $E[g]$ analogously gives the conserved total *energy*. ¶

Another way to interpret the conservation laws above is to regard them as the statement that the functions $g = g(v)$ commutes with the differential operator $\partial_t + v \cdot \nabla^{(x)}$. In fact, for any differential operator K that commutes with $\partial_t + v \cdot \nabla^{(x)}$, we can define $E[K]$ its associated conserved quantity using the same argument as before.

1.18 Exercise (More conservation laws)

1. For $i, j \in \{1, \ldots, d\}$, $i \neq j$, set $K = x^i v^j - x^j v^i$. Verify that the multiplier K generates a conservation law. What is its physical meaning?

2. For $i \in \{1, \cdots, d\}$, set $K = v^i t - x^i$.

   (a) Verify that K generates a conservation law.

   (b) The conservation law generated by K can be re-written in the form

   $$\partial_t \int_{\mathbb{R}^{2d}} \rho(t,x,v)x^i\,\mathrm{d}x\,\mathrm{d}v = \partial_t \int_{\mathbb{R}^{2d}} \rho(t,x,v)v^i t\,\mathrm{d}x\,\mathrm{d}v.$$

   Give a physical interpretation of this formula. ■

1.19 (Even more conservation laws) The view of conservation laws as arising from operators commuting with the evolution equation is in fact that of *Noether's theorem* in Hamiltonian mechanics. The Lagrangian counterpart of Noether's theorem ties these conservation laws to symmetries of the equations. So we are lead to considering the natural symmetry of the Vlasov equation (1.6), namely that of the Galilean symmetry. If $\rho(t,x,v)$ is a solution, then so are

- Space-time translations: for any $\tau \in \mathbb{R}$ and $\xi \in \mathbb{R}^d$, $\rho(t+\tau, x+\xi, v)$.

- Rotational symmetry: for any orthogonal matrix O, $\rho(t, Ox, Ov)$.
- Galilean boost: for any $w \in \mathbb{R}^d$, $\rho(t, x - tw, v - w)$.
- Spatial scaling: for any $\lambda \in \mathbb{R}_+$, $\rho(t, \lambda x, \lambda v)$.
- Temporal scaling: for any $\lambda \in \mathbb{R}_+$, $\rho(\lambda^{-1} t, x, \lambda v)$.

Note, in particular, that these symmetries are continually parameterized: the space-time translation by $(\tau, \xi) \in \mathbb{R} \times \mathbb{R}^d$, the rotations by $O \in \mathrm{SO}(d, \mathbb{R})$, boosts by $w \in \mathbb{R}^d$, and scalings by $\lambda \in \mathbb{R}_+$; furthermore, the symmetry operations acts by the identity when, in the translation case $(\tau, \xi) = (0, 0)$, in the rotation case $O = \mathrm{Id}$, in the boost case $w = 0$, and in the scaling cases $\lambda = 1$. Now, let ρ be a solution, and let ρ_γ be a one parameter family of solutions, obtained, for each γ, from the action of the above symmetries. Assume that this family is "differentiable in γ". Then by the linearity of (1.6) the derivative $\frac{d}{d\gamma}\rho|_{\gamma=0}$ is also a solution of the Vlasov equations.

What are the various $\frac{d}{d\gamma}\rho$? It turns out the possible ones corresponding to the classes of the symmetries above can all be obtained as linear combinations of the following linear operators on ρ:

- Corresponding to space-time translations: $\partial_t \rho$ and $\partial_{x^i} \rho$.
- Corresponding to rotations: $(x^i \partial_{x^j} - x^j \partial_{x^i} + v^i \partial_{v^j} - v^j \partial_{v^i})\rho$.
- Corresponding to boosts: $t\partial_{x^i}\rho + \partial_{v^i}\rho$.
- Corresponding to spatial scaling: $x \cdot \nabla^{(x)}\rho + v \cdot \nabla^{(v)}\rho$.
- Corresponding to temporal scaling: $-t\partial_t \rho + v \cdot \nabla^{(v)}\rho$.

In particular, all of these operators can be written as acting on ρ by some vector field K. ¶

Now, if one were to compute the conserved quantities associated to most of the symmetries described in the previous paragraph, one would find that the integral $E[K]$ evaluates to identically zero. Part of the problem is that while ρ is by definition non-negative, $K\rho$, for some arbitrary differential operator K, is generally unsigned. And in fact in our situation we have often perfect cancellations. This can however by remedied by the following observation: our linear equation (1.6) really says that the directional derivative of ρ, as a function on $\mathbb{R}^{2d+1}$, in the direction of a certain fixed vector field $\partial_t + v \cdot \nabla^{(x)}$ is zero. So in particular, if ρ solves the Vlasov equation

Without using the words, what is really going on is that the continuous family of symmetries above are actually via actions by Lie groups. The discussion here is getting at the concept of infinitesimal generators of symmetries, i.e. elements of the corresponding Lie algebra.

(without necessarily satisfying the constraint that $\rho \geq 0$ everywhere), and if $f : \mathbb{R} \to \mathbb{R}$ is any function, then $f(\rho)$ solves (1.6) also, in the sense that the particular directional derivative of $f(\rho)$ exists and is equal to zero, *even when f is not necessarily differentiable*.

In the case where f is differentiable, this conclusion is simply drawn from the chain rule

$$(\partial_t + v \cdot \nabla^{(x)}) f(\rho) = f'(\rho)(\partial_t + v \cdot \nabla^{(x)})\rho. \tag{1.20}$$

1.21 Theorem (Dispersive estimate for Vlasov equation, version 2)
Let $\rho_0 \in \mathscr{S}(\mathbb{R}^{2d})$. The solution to the Vlasov equation (1.6) with initial data $\rho(0,x,v) = \rho_0(x,v)$ satisfies

$$\sup_{x\in\mathbb{R}^d} \overline{\rho}(t,x) \leq |t|^{-d} \left\|\rho_0\right\|_{L^1_x \mathring{W}^{d,1}_v}$$

where the norm

$$\|f\|_{L^1_x \mathring{W}^{k,1}_v} = \sum_{|\alpha|=k} \iint_{\mathbb{R}^{2d}} |\partial_v^\alpha f| \,\mathrm{d}v\,\mathrm{d}x.$$

■

1.22 Remark
Compare the conclusion of this theorem to that of Theorem 1.10, we see that in this case if we sacrifice the estimate on *short time* behavior (namely, when $t \approx 0$; this is the difference between $|t|$ and $\langle t \rangle$), we can sharpen the constant so that instead of using derivatives over the full phase-space $\mathbb{R}^d$, we only need derivatives in the velocity (v) variables. This makes precise the idea which was mentioned before, that the dispersion depends on the regularity of the velocity distribution of the initial data. ■

This proof is an adaptation of the argument from Smulevici, "Small data solutions of the Vlasov-Poisson system and the vector field method".

Proof (Theorem 1.21) Observe that we can translate between pointwise and integral estimates by integration. More precisely, for $x \in \mathbb{R}^d$, let R_x denote the rectangle

$$R_x \stackrel{\text{def}}{=} \left\{ y \in \mathbb{R}^d \;\middle|\; y^i \leq x^i \quad \forall i \in \{1,\ldots,d\} \right\}.$$

Then by fundamental theorem of calculus we have

$$\begin{aligned}\overline{\rho}(t,x) &= \int_{R_x} \partial_{x^1}\partial_{x^2}\cdots\partial_{x^d}\overline{\rho}(t,y)\,\mathrm{d}y \\ &= \int_{R_x}\int_{\mathbb{R}^d} \partial_{x^1}\partial_{x^2}\cdots\partial_{x^d}\rho(t,y,v)\,\mathrm{d}v\,\mathrm{d}y. \end{aligned}\tag{1.23}$$

Next, observe that by rapid decay we have

$$\int_{\mathbb{R}^d} \partial_{v^i} \rho(t,x,v)\, \mathrm{d}v = 0$$

for any $i \in \{1,\ldots,d\}$. So we can rewrite (1.23) using the Galilean boosts:

$$\overline{\rho}(t,x) = t^{-d} \int_{R_x} \int_{\mathbb{R}^d} G_1 \cdots G_d \rho(t,y,v)\, \mathrm{d}v\, \mathrm{d}y. \tag{1.24}$$

where $G_i = t\partial_{x^i} + \partial_{v^i}$. From here we estimate

$$\overline{\rho}(t,x) \leq |t|^{-d} \iint_{\mathbb{R}^{2d}} \left| G_1 \cdots G_d \rho(t,y,v) \right| \mathrm{d}v\, \mathrm{d}y.$$

From our discussion above, we observe that $\left| G_1 \cdots G_d \rho(t,y,v) \right|$ solves Vlasov's equation, and so the L^1 conservation of mass holds. Which means

$$\overline{\rho}(t,x) \leq |t|^{-d} \iint_{\mathbb{R}^{2d}} \left| G_1 \cdots G_d \rho(0,y,v) \right| \mathrm{d}v\, \mathrm{d}y.$$

Now, when $t = 0$, the vector field G_i simplifies to ∂_{v^i}, and so we have, finally,

$$\overline{\rho}(t,x) \leq |t|^{-d} \iint_{\mathbb{R}^{2d}} \left| \partial_{v^1} \cdots \partial_{v^d} \rho(t,y,v) \right| \mathrm{d}v\, \mathrm{d}y \leq |t|^{-d} \left\| \rho_0 \right\|_{L^1_x \mathring{W}^{d,1}_v}$$

as desired. □

1.25 Exercise
As discussed in Remark 1.22, the estimate proved in Theorem 1.21 degenerates as $|t| \to 0$. Prove, using a modification of the arguments immediately above, that

$$\sup_{x \in \mathbb{R}^d} \overline{\rho}(t,x) \leq \langle t \rangle^{-d} \left\| \rho_0 \right\|_{W^{d,1}(\mathbb{R}^d \times \mathbb{R}^d)}.$$

(*Hint:* we want to make use of a conservation law, so we want to do this by applying a set of d vector fields. The final result tells you that, at the initial data level, these vector fields involve differentiations also in the x coordinates.) ■

The quantum phase space

When one sees "dispersive partial differential equations" in the literature, generally one refers not to the *(semi-)classical* equations of kinetic theory described in the previous section, but to the partial differential equations governing their *quantum mechanical* counterparts. In the quantum mechanical picture, the position distribution representing the particles cannot be specified independently from the momentum (velocity) distribution, thanks to the celebrated Heisenberg Uncertainty Principle.

There is a vast literature on formalizing "quantization", the mathematical process of starting from a classical mechanical system (described by, for example, a symplectic manifold) and forming a quantum mechanical system. Our discussion here is much more rudimentary; see the first part of Cohen-Tannoudji et al., Quantum Mechanics *(among many other standard textbooks) for more details.*

1.26 (A crash course in quantum mechanics) The fundamental idea here is that of the *de Broglie hypothesis*. In his 1924 PhD dissertation, de Broglie proposed the idea of wave-particle duality. Earlier works of Planck and Einstein have postulated that electromagnetic waves can be considered as *particles* with energy $E = \hbar\omega$ and momentum $p = \hbar k$, with ω being the temporal angular frequency and k being the wave vector (a monochromatic plane wave with frequency ω and wave vector k is described by the amplitude function $\phi(t,x) = \exp(i\omega t + ik\cdot x)$; observe that $E/p = \omega/k$ is the wave velocity as expected). De Broglie took the idea one step further and proposed that all particles can be considered as waves, with frequency and wave vector obtained from what is now known as the *de Broglie relations* (which is really just rewriting the Planck-Einstein equations),

$$\omega = \frac{E}{\hbar}, \qquad k = \frac{p}{\hbar},$$

where E is the kinetic energy and p the momentum (which is proportional to particle velocity by its mass) of the particle.

Now, a pure particle of energy E and momentum p is then associated to the monochromatic plane wave

$$\phi(t,x) = \exp(i\omega t + ik\cdot x)$$

where ω, k satisfy the de Broglie relations given above. Notice that for monochromatic plane waves

$$k = -i\bar{\phi}(t,x)\nabla^{(x)}\phi(t,x)$$

where $\bar{\phi}$ denotes complex conjugation, we see the idea that, in quantum mechanics, the spatial gradient is (up to a scaling constant) the "momentum operator". (Similarly, the time derivative is the "energy operator".)

Assuming a linear physical theory, waves obey the superposition principle. So at any space-time position we should expect the field strength ϕ

for our particles to be the sum of the corresponding monochromatic plane waves of a bunch of different particles with different energies and momenta. That is to say, we expect that the "amplitude" of our "matter wave" to be in fact

$$\phi(t,x) = \int_{\mathbb{R}\times\mathbb{R}^d} \widetilde{\phi}(\omega,k)\, \exp(i\omega t + ik\cdot x)\, \mathrm{d}\omega\, \mathrm{d}k. \tag{1.27}$$

The amplitude here should be compared with the spatial density $\overline{\rho}(t,x) = \int_{\mathbb{R}^d} \rho(t,x,v)\, \mathrm{d}v$ in the kinetic theory picture: it is the sum of all particles of different energy-momentum (velocity in the Vlasov case) that impinge on a space-time point.

Equation (1.27) should remind us of the (inverse) Fourier transform, where ω and t are conjugate variables, and also k and x. Whereas a classical matter distribution is a function over the classical phase space $\mathbb{R}^d \times \mathbb{R}^d$, a quantum matter distribution is interpreted as *either* a physical space(-time) distribution $\phi(t,x)$, *or alternatively* a conjugate space(-time) distribution $\widetilde{\phi}(\omega,k)$. In the classical case specifying either or both of the physical space or velocity space traces (in other words specifying $\int_{\mathbb{R}^d} \rho(t,x,v)\, \mathrm{d}v$ or $\int_{\mathbb{R}^d} \rho(t,x,v)\, \mathrm{d}x$) is not enough to constraint the distribution itself; in the quantum case specify either ϕ or $\widetilde{\phi}$ gives the other by the Fourier transform. ¶

1.28 **(Equations of motion)** In this picture, the equations of motion are captured in the relation between energy and momentum; in other words, between ω and k. Here we give some examples.

1. *Schrödinger's equation*. For classical Newtonian physics, the kinetic energy is proportional to the square of the momentum. This leads us to postulate $\omega = |k|^2$ as the relation. This in turn means that the corresponding monochromatic plane waves $\exp(i\omega t + ik\cdot x)$ solve the equation
$$i\partial_t \phi - \triangle\phi = 0, \tag{1.29}$$
where the Laplace operator is $\triangle \stackrel{\text{def}}{=} \sum_{i=1}^d \partial^2_{x^i x^i}$. This equation can also be "derived" from the Vlasov equation (1.6) by replacing, in the term $v\cdot\nabla^{(x)}$, the velocity v by the momentum operator $i\nabla^{(x)}$.

2. *Airy equation*. Let P be any polynomial of d variables $k^1,\dots,k^d$ with real coefficients; more precisely let
$$P(k) = \sum_{|\alpha|\leq N} A_\alpha k^\alpha$$

where α ranges over multi-indices, A_α are real numbers, and for $\alpha = (\alpha_1, \alpha_2, \ldots, \alpha_d)$ the monomial

$$k^\alpha = (k^1)^{\alpha_1}(k^2)^{\alpha_2}\cdots(k^d)^{\alpha_d}.$$

The postulate $\omega = P(k)$ leads to the corresponding plane waves being solutions to the equation

$$i\partial_t\phi + P(i\nabla^{(x)})\phi = i\partial_t\phi + \sum_{|\alpha|\leq N} A_\alpha \partial^\alpha \phi = 0.$$

In the case where $d = 1$ and $P(k) = k^3$, this is known as the Airy equation, or the linear Korteweg-de Vries equation

$$\partial_t\phi - \partial^3_{xxx}\phi = 0. \tag{1.30}$$

3. *Relativistic dynamics*. For the special-relativistic dynamics of particles, the relation between energy and momentum are such that the difference between the square of the energy and the square of the momentum is a constant depending on the rest mass of the particle. Therefore for massive particles we can postulate the relation $\omega^2 - |k|^2 = 1$ and for massless particles the relation $\omega^2 - |k|^2 = 0$. The former leads to the *Klein-Gordon equation*:

$$\partial^2_{tt}\phi - \triangle\phi + \phi = 0; \tag{1.31}$$

the latter leads to the linear *wave equation*:

$$\partial^2_{tt}\phi - \triangle\phi = 0. \tag{1.32}$$

Equations (1.29), (1.30), (1.31) and (1.32) are the fundamental examples that we will discuss in this set of notes. Their solutions, going back to the description above, can be identified as the (inverse) Fourier transform (1.27) of distributions supported on the set $\{\omega = |k|^2\}$, $\{\omega = k^3\}$, $\{\omega^2 - |k|^2 = 1\}$, and $\{\omega^2 = |k|^2\}$ respectively. Note that these sets are smooth hypersurfaces, except for the isolated singularity at the origin of the last example. In fact, in the first three examples, the sets are *graphs* over the hyperplane $\{(0,k) \in \mathbb{R}\times\mathbb{R}^d\}$. In view of this, we will frequently decouple the time parameter t from the spatial position, and treat our field ϕ as a time-dependent distribution that has either a physical space representation $\phi(t,x)$ or a momentum space representation $\widehat{\phi}(t,k)$. ¶

The connection of the quantum phase-space with the Fourier transform is both a blessing and a curse. On the plus side, that Heisenberg's Uncertainty Principle implies that the physical and momentum space representations can not be both concentrated near a point implies that there are some limits to how much the "particle" picture can place obstructions on the decay. Returning to Exercise 1.4, we see that in the N particle case, we have an obstruction to the decay of the number density at scale R. Since particles are indivisible, when there are at least one particle, we have $\rho(R, t) \geq 1$ always. In the case of Vlasov equation, on the other hand, the spatial density $\overline{\rho}(t, x)$ decays *provided the initial distribution is smooth enough*; and as we saw in Exercise 1.14, the "particle-like" situations provide genuine obstructions to dispersive decay. In the quantum set-up, by removing the pure particle situation from consideration, we expect that this obstruction to be ameliorated. On the minus side, in the classical Vlasov situation, the position and momentum are independent; but in the quantum picture, the they are two-sides of the same coin. So while in the discussion above for kinetic theory we can act on the velocity coordinate (see, for example, the arguments in Thought 1.17) with zero effect on the spatial coordinate and vice versa, for quantum systems the fact that the position and momentum operators do not commute will introduce additional complications and artifacts into our computations. These ideas will be explored in more detail in the following chapters.

Incidentally this is another way of stating Heisenberg's Uncertainty Principle.

To give an illustration of the effects of quantization, where we identify the velocity variable v with $i\nabla^{(x)}$, let us return to the discussion of conservation laws given in Thought 1.17, Exercise 1.18, and Thought 1.19. For the classical Vlasov equation we've seen that the quantity

$$\int_{\mathbb{R}^{2d}} \rho(t, x, v) v \cdot v \, \mathrm{d}x \, \mathrm{d}v$$

is conserved in time, and we associated this to the conservation of kinetic energy. On the other hand, we also saw that if ρ solves the Vlasov equation, so does $\nabla^{(x)} \cdot \nabla^{(x)} \rho$ due to the spatial translation symmetry (applied twice). This implies that

$$\int_{\mathbb{R}^{2d}} \nabla^{(x)} \cdot \nabla^{(x)} \rho(t, x, v) \, \mathrm{d}x \, \mathrm{d}v$$

(which evaluates to zero) is a conserved quantity. In the quantum situation of the Schrödinger equation, these two conservation laws should be "identified" by our discussion above. Similarly, we have that the *multiplier* $e \cdot v$

corresponding to linear momentum can be identified with the *differential operator* $e \cdot \nabla^{(x)}$ that is a symmetry of the Vlasov equation, and that the two multipliers given in Exercise 1.18 can be identified with the differential operators for rotations and Galilean boosts respectively. We will return to conservation laws and commuting vector fields for the Schrödinger equation in a later chapter.

Chapter 2

Fourier Theory: the Basics

At the end of the previous chapter we described briefly the quantum phase space, where if we accept the de Broglie hypothesis of wave-particle duality, then the principle of superposition leads us directly to (1.27) which relates the physical space distribution of the particle density to the momentum space distribution of the particles. The relation is reminiscent of the Fourier transform. Therefore it shouldn't come as a surprise that much of the modern understanding of dispersive equations are built upon the foundation of Fourier theory. In this chapter we will review the more pertinent of its basic aspects. Many results will be stated without proof. For more background, the first chapter of Stein and Weiss, *Introduction to Fourier analysis on Euclidean spaces* is a good reference, as is Stein and Shakarchi, *Fourier analysis*.

Techniques based on Fourier theory that are more specific to the analysis of dispersive equations will be presented in the course of the following chapters with full details. Readers familiar with the basics of the Fourier transform can skip or skim this chapter.

The Fourier transform

Let $\phi : \mathbb{R}^d \to \mathbb{C}$. The basics of our discussion is the formula

$$\widehat{\phi}(\xi) \stackrel{\text{def}}{=} \mathscr{F}[\phi](\xi) \stackrel{\text{def}}{=} \frac{1}{(2\pi)^{\frac{d}{2}}} \int_{\mathbb{R}^d} \phi(x) e^{-i\xi\cdot x} \, \mathrm{d}x. \tag{2.1}$$

We interpret $\mathscr{F}$ as an operator sending functions on $\mathbb{R}^d$ to functions on $\mathbb{R}^d$, whenever the above expression is *well-defined*. The linearity of integration in particular implies that $\mathscr{F}$ is a linear operator. Noting that $|\exp(-i\xi \cdot x)| = 1$, we have that a *sufficient* condition for $\mathscr{F}[\phi]$ to be well-defined is that $\phi \in L^1(\mathbb{R}^d)$; that is, ϕ is absolutely integrable. In this case, we have the estimate

$$\left|\mathscr{F}[\phi](\xi)\right| \leq \frac{1}{(2\pi)^{\frac{d}{2}}} \int\limits_{\mathbb{R}^d} \left|\phi(x)e^{-i\xi\cdot x}\right| \, \mathrm{d}x = \frac{1}{(2\pi)^{\frac{d}{2}}} \left\|\phi\right\|_{L^1(\mathbb{R}^d)}$$

and so we conclude that $\mathscr{F} : L^1(\mathbb{R}^d) \to L^\infty(\mathbb{R}^d)$ is a bounded linear operator. In fact, we have more than just boundedness.

2.2 Proposition (Uniform continuity)
If $\phi \in L^1(\mathbb{R}^d)$, then $\widehat{\phi}$ is uniformly continuous. ■

2.3 Exercise
Prove the above proposition following the outline below:

1. First prove that
$$\lim_{R\to\infty} \int\limits_{|x|>R} \left|\phi\right| \, \mathrm{d}x = 0.$$
2. Show that for every $R, \epsilon > 0$, there exists $\delta > 0$ such that whenever $|x| \leq R$ and $|\xi - \xi'| < \delta$,
$$\left|e^{-i\xi\cdot x} - e^{-i\xi'\cdot x}\right| < \epsilon.$$
3. Prove uniform continuity, by splitting the integral
$$\int\limits_{\mathbb{R}^d} \phi(x)\left[e^{-i\xi\cdot x} - e^{-i\xi'\cdot x}\right] \, \mathrm{d}x$$
into an integral over a compact region and an integral near infinity. ■

2.4 Remark (Normalizations of the Fourier transform)
In the literature, the Fourier transform is defined usually as one of the following three operations.

$$\mathscr{F}[\phi](\xi) = \begin{cases} \int_{\mathbb{R}^d} \phi(x)e^{-i\xi\cdot x} \, \mathrm{d}x \\ \frac{1}{(2\pi)^{d/2}} \int_{\mathbb{R}^d} \phi(x)e^{-i\xi\cdot x} \, \mathrm{d}x \\ \int_{\mathbb{R}^d} \phi(x)e^{-2\pi i\xi\cdot x} \, \mathrm{d}x \end{cases}$$

The choice of convention is largely cosmetic. The first or second are often preferred by specialists in partial differential equations as in those cases the momentum operator is $i\nabla^{(x)}$ instead of $2\pi i\nabla^{(x)}$. The second and (more frequently) the third are often preferred by harmonic analysts and functional analysts as in those definitions the Fourier transform extends to a unitary mapping of $L^2(\mathbb{R}^d)$ to itself. The purpose of this remark is just to warn the readers that, when comparing formulae from different sources, make sure to double check the normalization of the Fourier transform, as several conventions are in use in the literature. ■

The following two Propositions can be proved by direction computation and we omit their proofs here.

2.5 Proposition (Basic properties 1: conjugation, reflection)
Let $\phi \in L^1(\mathbb{R}^d)$. Denote by $R : L^1(\mathbb{R}^d) \to L^1(\mathbb{R}^d)$ the mapping $\phi(x) \mapsto \phi(-x)$. Then

$$\mathscr{F}[R\phi] = R\mathscr{F}[\phi] \tag{2.6a}$$

$$\mathscr{F}[\overline{\phi}] = R\overline{\mathscr{F}[\phi]} \tag{2.6b}$$

■

2.7 Proposition (Basic properties 2: scaling, translation, modulation)
Let $\phi \in L^1(\mathbb{R}^d)$. For $\lambda > 0$, $y \in \mathbb{R}^d$, and $\zeta \in \mathbb{R}^d$, denote by $S_\lambda, \tau_y, \mu_\zeta$ the mappings from $L^1(\mathbb{R}^d)$ to itself given by

$$S_\lambda\phi(x) = \lambda^{d/2}\phi(\lambda x), \quad \tau_y\phi(x) = \phi(x+y), \quad \mu_\zeta\phi(x) = \phi(x)\cdot e^{i\zeta\cdot x}.$$

Then

$$\mathscr{F}[S_\lambda\phi] = S_{\lambda^{-1}}\mathscr{F}[\phi] \tag{2.8a}$$

$$\mathscr{F}[\tau_y\phi] = \mu_y\mathscr{F}[\phi] \tag{2.8b}$$

$$\mathscr{F}[\mu_\zeta\phi] = \tau_{-\zeta}\mathscr{F}[\phi] \tag{2.8c}$$

■

The next Proposition is also standard, and captures the interchange of the momentum parameter and its physical space representation as a differential operator.

2.9 Proposition (Basic properties 3: differentiation)
Let $\phi \in W^{1,1}(\mathbb{R}^d)$. Then for any $j \in \{1,\ldots,d\}$,

$$\mathscr{F}[\partial_{x^j}\phi](\xi) = i\xi^j\widehat{\phi}(\xi).$$

Conversely, if ϕ is such that $\langle\bullet\rangle\phi \in L^1(\mathbb{R}^d)$, then $\widehat{\phi}$ is differentiable and

$$\mathscr{F}[x^j\phi](\xi) = i\partial_{\xi^j}\widehat{\phi}(\xi).$$

■

2.10 Exercise (Riemann-Lebesgue Lemma)
Prove that whenever $\phi \in L^1(\mathbb{R}^d)$, for every $\epsilon > 0$ there exists $R > 0$ such that

$$\sup_{|\xi|>R}\left|\widehat{\phi}(\xi)\right| < \epsilon.$$

(Hint: approximate ϕ by a function in $W^{1,1}$ and apply the previous Proposition.) ■

2.11 Exercise (Baby Paley-Wiener)
In this exercise you will prove that if $\phi \in L^1(\mathbb{R}^d)$ is such that there exists $R > 0$ where $\phi|_{|x|>R} \equiv 0$, then $\widehat{\phi}$ is *real analytic*. This is a primitive version of the Paley-Wiener theorem.

1. First show that the assumption implies $\widehat{\phi}$ is differentiable using Proposition 2.9. By induction show that $\widehat{\phi}$ is infinitely differentiable.
2. Using the assumption show the following uniform bound: for every multi-index α
$$\left|\partial^\alpha\widehat{\phi}(\xi)\right| \leq R^{|\alpha|}\|\phi\|_{L^1}.$$
3. Use this to conclude that the Taylor series of $\widehat{\phi}$ at every point has infinite radius of convergence.
4. Show further (using Taylor's remainder theorem) that there exists $r > 0$ such that for every ξ, ξ' satisfying $|\xi - \xi'| < r$, the Taylor series of $\widehat{\phi}$, centered at ξ, converges to $\widehat{\phi}(\xi')$ pointwise. From this conclude that $\widehat{\phi}$ is real analytic. ■

The space $\mathscr{S}$; Fourier inversion

We have already seen that the Fourier transform is a bounded linear map from L^1 to L^∞: the domain and co-domain are not equal. Even though the Riemann-Lebesgue Lemma (Exercise 2.10) guarantees that the Fourier transforms of absolutely integrable functions decay at infinity, the decay does not come at a precise rate and in general the Fourier transforms are *not* absolutely integrable. This is also seen in Proposition 2.7. Using the same notation as given there, we observe that

$$\|\phi\|_{L^1} = \|\lambda^{d/2} S_\lambda \phi\|_{L^1}.$$

And therefore simply from scaling considerations we see that $\mathscr{F}$ *cannot* be a bounded linear map from L^1 to itself.

2.12 Exercise
Make the above argument rigorous; that is, using the scaling property in Proposition 2.7, show the Fourier transform is not a bounded linear map from L^1 to itself. ■

If we want to think of the Fourier transform as a mapping from some function space to itself, a constraint is given by Proposition 2.9. We saw there that differentiability of a function translates to decay of the Fourier transform, and that decay of the function translates to the differentiability of the Fourier transform. Therefore in constructing a function space that is preserved under the Fourier transform, differentiability and decay should go hand-in-hand. One option is, of course, to forego both differentiability and decay; we will address some of those function spaces later on. For much of our purpose (proving *a priori* estimates for partial differential equations), it is convenient for the computations to assuming that we have differentiability available.

2.13 Definition (Schwartz space)
An infinitely differentiable function ϕ is said to be in the *Schwartz space* $\mathscr{S}(\mathbb{R}^d)$ if

$$\sup_{x \in \mathbb{R}^d} \left| x^\alpha \partial^\beta \phi \right| < \infty$$

for every pair of multiindices α and β. ■

One can equip $\mathscr{S}$ with the topology of a Fréchet space, by using this defining condition, indexed by the multiindices α and β, as its countable family of seminorms.

2.14 Definition (Convolution)
When the formula below is well-defined, we say that the convolution of two functions f and g is

$$f * g(x) \stackrel{\text{def}}{=} \int_{\mathbb{R}^d} f(x-y) g(y) \, \mathrm{d}y.$$

Notice that by a direct change of variables we also have

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

that is to say, $f * g = g * f$. ■

2.15 Lemma
If $\phi, \psi \in \mathscr{S}$, then so are

- $\partial^\alpha \phi$ for any multi-index α;
- $P \cdot \phi$ for any polynomial P;
- the pointwise sum $\phi + \psi$;
- the pointwise product $\phi \cdot \psi$;
- the convolution $\phi * \psi$. ■

2.16 Proposition
If $\phi \in \mathscr{S}(\mathbb{R}^d)$, then $\widehat{\phi} \in \mathscr{S}(\mathbb{R}^d)$. ■

Ref. 2.9: "Fourier transform properties: differentiation"

Proof Follows from Proposition 2.9 directly. □

2.17 Example (Gaussian)
Consider the family of functions, parametrized by $s \in (0, \infty)$,

$$g_s(x) = e^{-\frac{s}{2}|x|^2}$$

defined on $\mathbb{R}^d$. We claim that $g_s \in \mathscr{S}$. First, using the well-known fact that exponential grow faster than any power, it is easy to check that $\lim_{|x|\to\infty} P(x) g_s(x) = 0$ for any polynomial P. Next, observe that for any multi-index α, the derivative can be written as

$$\partial^\alpha g_s = Q^{(\alpha)} g_s$$

for some polynomial $Q^{(\alpha)}$. Therefore $g_s \in \mathscr{S}$ for every $s \in (0, \infty)$.

The L^1 norm of g_s can be computed with the following well-known trick. First, observe the formula

$$g_s(x) = g_s(x^1) g_s(x^2) g_s(x^3) \cdots g_s(x^d).$$

And hence if we compute $\int_{\mathbb{R}} \exp(-s|x|^2/2)\, \mathrm{d}x$ we can get $\int_{\mathbb{R}^d} g_s\, \mathrm{d}x$ by raising it to the dth power by Fubini's theorem. The case $d = 2$, however, can be computed directly, using that in polar coordinates

$$\int_{\mathbb{R}^2} \exp(-\frac{s}{2}(x^2 + y^2))\, \mathrm{d}x\, \mathrm{d}y = \int_0^\infty \int_0^{2\pi} \exp(-\frac{s}{2} r^2) r\, \mathrm{d}r\, \mathrm{d}\theta = \frac{2\pi}{s}.$$

So that

$$\int_{\mathbb{R}^d} g_s(x)\, \mathrm{d}x = \left(\frac{2\pi}{s}\right)^{\frac{d}{2}}.$$

With this we can also compute the convolution of two such functions:

$$\begin{aligned} g_s * g_t(x) &= \int \exp\left(-\frac{s}{2}\left|x-y\right|^2 - \frac{t}{2}\left|y\right|^2\right) \mathrm{d}y \\ &= \int \exp\left(-\frac{s+t}{2}\left|y\right|^2 + sx\cdot y - \frac{s}{2}|x|^2\right) \mathrm{d}y \\ &= \int \exp\left(-\frac{s+t}{2}\left|y - \frac{s}{s+t}x\right|^2\right)\cdot \exp\left(-\frac{st}{2(s+t)}|x|^2\right) \mathrm{d}y \\ &= \left(\frac{2\pi}{s+t}\right)^{\frac{d}{2}} g_{\frac{st}{s+t}}(x). \end{aligned}$$

The Fourier transform of g_s can also be computed. Noting that $g_s = s^{-d/4} S_{\sqrt{s}} g_1$ (in the notations of Proposition 2.7), it suffices to compute $\widehat{g_1}$. For this we will repeatedly use Proposition 2.9. First, since $g_1 \in \mathscr{S}$, we have that $\widehat{g_1}$ is differentiable, and

Ref. 2.7: "Fourier transform properties: scaling, translation, modulation"

Ref. 2.9: "Fourier transform properties: differentiation"

$$\partial_{\xi^j} \widehat{g_1} = -i\mathscr{F}[x^j g_1].$$

Next, observe that

$$\partial_{x^j} g_1 = -x^j g_1$$

by the chain-rule of differentiation, and hence we have that

$$\partial_{\xi^j} \widehat{g_1} = i\mathscr{F}[\partial_{x^j} g_1] = -\xi^j \widehat{g_1}$$

with the second equality again following from Proposition 2.9. Using integrating factors we see that this implies

$$\nabla^{(\xi)}\left[e^{\frac{1}{2}|\xi|^2}\widehat{g_1}(\xi)\right] = 0,$$

and hence

$$\widehat{g_1}(\xi) = C g_1(\xi)$$

for some constant C. To find out the value of this constant, we evaluate using (2.1)

$$\widehat{g_1}(0) = \frac{1}{(2\pi)^{d/2}} \int_{\mathbb{R}^d} g_1(x)\, \mathrm{d}x = 1,$$

and conclude that g_1 *is its own Fourier transform*. ∎

2.18 Lemma (Parseval's formula)
If $\phi, \psi \in \mathscr{S}$, then

$$\int \phi(x)\widehat{\psi}(x)\,\mathrm{d}x = \int \widehat{\phi}(x)\psi(x)\,\mathrm{d}x. \qquad \blacksquare$$

Proof (Sketch) This follows by direct computation using (2.1). □

2.19 Theorem (Fourier inversion)
If $\phi \in \mathscr{S}$, and $\widehat{\phi}$ is defined as in (2.1), then

$$\phi(x) = \frac{1}{(2\pi)^{\frac{d}{2}}} \int_{\mathbb{R}^d} \widehat{\phi}(\xi) e^{ix\cdot\xi}\,\mathrm{d}\xi = \mathscr{F}[\widehat{\phi}](-x). \qquad \blacksquare$$

Proof Using that $g_s = s^{-d/4} S_{\sqrt{s}} g_1$, we have that $\widehat{g_s} = s^{-d/2} g_{1/s}$, and $\int \widehat{g_s}\,\mathrm{d}x = (2\pi)^{d/2}$. Therefore $(2\pi)^{-d/2}\widehat{g_s}$, as $s \to 0$, is an approximation of identity. In particular, we have that

$$\phi(y) = \lim_{s\to 0} \frac{1}{(2\pi)^{d/2}} \int \phi(y+x)\widehat{g_s}(x)\,\mathrm{d}x.$$

Applying Parseval's formula we get

$$\phi(y) = \lim_{s\to 0} \frac{1}{(2\pi)^{d/2}} \int \mathscr{F}[\tau_y\phi](x) g_s(x)\,\mathrm{d}x.$$

As $s \to 0$ clearly $g_s(x) \to 1$ pointwise, and hence we get

$$\phi(y) = \frac{1}{(2\pi)^{d/2}} \int \mathscr{F}[\tau_y\phi](x)\,\mathrm{d}x.$$

Ref. 2.7: "Fourier transform properties: scaling, translation, modulation"

Now using Proposition 2.7 we get

$$\phi(y) = \frac{1}{(2\pi)^{d/2}} \int \mu_y \widehat{\phi}(x)\,\mathrm{d}x$$

which is exactly as claimed. □

2.20 Example (Eigenspaces of the Fourier transform)
In Example 2.17, we've seen that $g_1(x) = \exp -\frac{1}{2}|x|^2$ is its own Fourier transform. Now, observing that $\mathscr{F} : \mathscr{S} \to \mathscr{S}$, we can ask whether there are

other functions that are also its own Fourier transform. It turns out that there are many such functions.

By the Fourier inversion formula, we have that $\mathscr{F}\mathscr{F}[\phi](x) = \phi(-x)$. This means that $\mathscr{F}^{(4)}[\phi] = \phi$; that is to say, $\mathscr{F}$ is a linear operator whose fourth power is the identity. This means that its eigenvalues are the fourth roots of unity: ± 1 and $\pm i$. To generate eigenfunctions of the Fourier transform, we can observe that, given any $\phi \in \mathscr{S}$, $\mathscr{F}$ acts by cyclic permutation on the ordered quadruple $(\phi, \mathscr{F}[\phi], \mathscr{F}^{(2)}[\phi], \mathscr{F}^{(3)}[\phi])$, and so the subspace spanned by those four functions is an *invariant subspace* of $\mathscr{F}$ acting on $\mathscr{S}$. Using those four functions as the basis vectors of this invariant subspace, $\mathscr{F}$ acts as the permutation matrix

$$\begin{pmatrix} & & & 1 \\ 1 & & & \\ & 1 & & \\ & & 1 & \end{pmatrix}$$

and one checks easily that the eigensubspaces E_λ corresponding to eigenvalues λ are

$$E_1 = \operatorname{span}(1,1,1,1), \qquad E_{-1} = \operatorname{span}(1,-1,1,-1);$$
$$E_i = \operatorname{span}(1,-i,-1,i), \qquad E_{-i} = \operatorname{span}(1,i,-1,-i).$$

This in particular implies that every ϕ can be decomposed as $\phi = \phi_1 + \phi_{-1} + \phi_i + \phi_{-i}$ with each factor living in the corresponding eigenspace of $\mathscr{F}$, and furthermore, this allows us to define *fractional* Fourier transforms. ■

First made rigorous by Norbert Wiener.

The relation between the Fourier transform an its inverse can be used to derive the following propositions. The latter, Plancherel's Theorem, shows that the Fourier transform is an isometry for the L^2 inner product.

2.21 Proposition (Basic properties 4: convolution and product) If $\phi, \psi \in \mathscr{S}$, then

$$\mathscr{F}[\phi * \psi] = (2\pi)^{d/2} \widehat{\phi} \cdot \widehat{\psi} \tag{2.22a}$$

$$\mathscr{F}[\phi \cdot \psi] = \frac{1}{(2\pi)^{d/2}} \widehat{\phi} * \widehat{\psi}. \tag{2.22b}$$

■

Proof (Sketch) The first equality follows by direct computation and switching the order of integration, using Proposition 2.7 in the process. The second equality follows by applying the Fourier inversion formula to the first one. □

Ref. 2.7: "Fourier transform properties: scaling, translation, modulation"

2.23 Proposition (Plancherel)
If $\phi, \psi \in \mathscr{S}$, then

$$\int \phi\overline{\psi}\, \mathrm{d}x = \int \widehat{\phi}\overline{\widehat{\psi}}\, \mathrm{d}\xi.$$

In particular,

$$\int |\phi|^2\, \mathrm{d}x = \int \left|\widehat{\phi}\right|^2\, \mathrm{d}\xi.$$

■

Proof (Sketch) Follows from applying Lemma 2.18 to ϕ and $\mathscr{F}^{-1}[\overline{\psi}]$, and using Proposition 2.5. □

Ref. 2.5: "Fourier transform properties: conjugation, reflection"

2.24 Exercise (Riemann-Lebesgue Lemma, part 2)
Prove that the decay given by Riemann-Lebesgue Lemma (Exercise 2.10) has no rate. More precisely:

Let $\{\epsilon_n\}_{n\in\mathbb{N}}$ be any sequence of positive real numbers tending to zero as $n \to \infty$. Show that there exists a function $\phi \in L^1(\mathbb{R})$ such that the inequality

$$\left|\widehat{\phi}(n)\right| > \epsilon_n$$

holds for infinitely many $n \in \mathbb{N}$.

(Hint: take some subset $S \subset \mathbb{N}$ and let ϕ be [using the Fourier inversion formula] a function whose Fourier transform consists of many little bumps centered around points in S. It suffices to choose S in a way that guarantees that ϕ is absolutely integrable.) ■

Fourier transform on L^2; the space $\mathscr{S}'$

In the previous sections, we have initially defined the Fourier transform (2.1) as a bounded linear mapping $L^1 \to L^\infty$; we saw further that restricting the domain to $\mathscr{S}$ the image of the Fourier transform also sits in $\mathscr{S}$. One particular aspect of this is Proposition 2.23, which asserts that for functions on $\mathscr{S}$, the Fourier transform preserves the $L^2(\mathbb{R}^d, \mathbb{C})$ inner product

$$\langle \phi, \psi \rangle = \int_{\mathbb{R}^d} \phi(x)\overline{\psi(x)}\, \mathrm{d}x.$$

It turns out that the Fourier transform can be extended uniquely to a Hilbert space isometry of $L^2(\mathbb{R}^d, \mathbb{C})$ to itself. We outline the procedure here:

1. First, we note that $\mathscr{S}$ is dense in L^2, in that given any $\phi \in L^2$ we can find a sequence ϕ_n of functions in $\mathscr{S}$ such that

$$\left\|\phi - \phi_n\right\|_{L^2} \to 0.$$

2. By virtue of convergence, we have that this sequence is Cauchy. Since $\mathscr{F}$ preserves the L^2 norm for Schwartz functions, we have that $\widehat{\phi_n}$ is a sequence of functions in $\mathscr{S}$, bounded in L^2 norm, and *Cauchy with respect to the L^2 distance*.

3. Therefore $\widehat{\phi_n}$ converges to an unique element in L^2: we define this element to be $\widehat{\phi}$.

This procedure is also how we will understand *a priori* estimates for solutions of partial differential equations. Abstractly, in many situations we can construct a solution operator $U(t)$ mapping from data in some space X to some solution space Y; for many of the equations that we will consider, $X = Y = \mathscr{S}$. For initial data in X we can prove an estimate based on norms W and Z of the form

$$\|U(t)f\|_Z \leq C(t)\|f\|_W.$$

Then, provided that X is dense in W and Y is dense in Z, the exact same procedure above allows us to extend the solution operator $U(t)$ to a mapping $U(t) : W \to Z$ with the same bounds. And therefore, for the remainder of these Notes, we generally will not be too concerned with regularity and integrability of functions that are being treated, since for the most part they can be assumed to lie in $\mathscr{S}$ or some related function space.

Aside from $\mathscr{S}$ and L^2, the Fourier transform can be extended also to a self-map on the space $\mathscr{S}'$ of *tempered distributions*.

2.25 Definition (Tempered distributions $\mathscr{S}'$)
The space $\mathscr{S}'(\mathbb{R}^d)$ is the set of all linear mappings $\mathscr{S} \to \mathbb{C}$ satisfying the following continuity condition: If $\Phi \in \mathscr{S}'$, there exists $m, n \in \mathbb{N}$ and a constant C such that for every $\psi \in \mathscr{S}$ the inequality

$$\left|\Phi(\psi)\right| \leq C \sum_{|\alpha| \leq m} \sup_{x \in \mathbb{R}^d} \langle x \rangle^n \left|\partial^\alpha \psi(x)\right|$$

holds (α ranging over multiindices). ■

2.26 Example

For a given $x \in \mathbb{R}^d$, the *evaluation* map $\mathscr{S} \ni \phi \mapsto \phi(x) \in \mathbb{C}$ is a tempered distribution. For historical reasons this distribution is often referred to as the *Dirac δ function centered at x* and the evaluating is written

$$\phi \mapsto \int_{\mathbb{R}^d} \delta_x(y)\phi(y)\, \mathrm{d}y,$$

even though δ_x is in no sense an actually integrable function. Clearly this operation satisfies the definition of the tempered distributions: letting $m = n = 0$, we have trivially

$$\left|\phi(x)\right| \leq \sup_{x\in\mathbb{R}^d} \left|\phi(x)\right|.$$

Similarly, letting Ω be any measurable subset of $\mathbb{R}^d$. The map

$$\phi \mapsto \int_{\Omega} \phi(x)\, \mathrm{d}x$$

satisfies, for $m = 0$ and $n = d + 1$

$$\begin{aligned} \left|\int_\Omega \phi(x)\, \mathrm{d}x\right| &= \left|\int_\Omega \langle x\rangle^{d+1}\phi(x)\langle x\rangle^{-1-d}\, \mathrm{d}x\right| \\ &\leq \int_{\mathbb{R}^d} \langle x\rangle^n \left|\phi(x)\right| \langle x\rangle^{-1-d}\, \mathrm{d}x \\ &\leq \left(\int_{\mathbb{R}^d} \langle x\rangle^{-1-d}\, \mathrm{d}x\right) \cdot \sup_{x\in\mathbb{R}^d} \langle x\rangle^n \left|\phi(x)\right|. \end{aligned}$$

Using that $\langle x\rangle^{-1-d}$ is integrable on $\mathbb{R}^d$, this shows that the above operator $\int_\Omega \bullet\, \mathrm{d}x$ of integrating over Ω is a tempered distribution.

The example in the previous paragraph can be further generalized: let $p \in [1, \infty]$, and let $f \in L^p(\mathbb{R}^d)$. Let P be any polynomial, and finally let α be some multi-index. Then the evaluation

$$\phi \mapsto \int_{\mathbb{R}^d} f(x)P(x)\partial^\alpha \phi(x)\, \mathrm{d}x$$

is another tempered distribution. ■

For applications, the following structure theorem gives a more concrete way of thinking about tempered distributions.

2.27 Theorem (Structure theorem for tempered distributions)
Every tempered distribution can be written as the finite linear combination of operators of the form

See Trèves, Topological Vector Spaces, Distributions and Kernels *for proof and more detailed discussions.*

$$\phi \mapsto \int_{\mathbb{R}^d} p(x) \partial^\alpha \phi(x) \, \mathrm{d}x$$

where $p(x)$ is a continuous function that grows at most polynomially. ■

2.28 Example
Consider the δ_x distribution again. This time let us focus on the case $d = 1$ (the higher dimensional case is analogous). How can we express $\phi(x)$ as an integral in the form given by Theorem 2.27?

The answer can be found by noting that, if δ_x were an integrable function, then by all rights its integral should satisfy

$$\int_{-\infty}^{y} \delta_x(z) \, \mathrm{d}z \text{ “=” } \begin{cases} 0 & y < x \\ 1 & x < y \end{cases}$$

and so morally speaking the Dirac δ is the “derivative” of the Heaviside step function. Inspired by this, we can choose

$$p(y) = \begin{cases} 0 & y \leq 0 \\ y & y > 0 \end{cases}$$

which is clearly continuous. Then we have

$$\int_{-\infty}^{\infty} p(y - x) \phi''(y) \, \mathrm{d}y = \int_{x}^{\infty} (y - x) \phi''(y) \, \mathrm{d}y.$$

Integrating by parts we get

$$= \underbrace{(y - x)\phi'(y)\Big|_x^\infty}_{=0} - \int_x^\infty \phi'(y) \, \mathrm{d}y = \phi(x),$$

giving us the desired representation. ■

Now, the Fourier transform $\mathscr{F} : \mathscr{S} \to \mathscr{S}$, and $\mathscr{S}'$ is the set of continuous linear functionals on $\mathscr{S}$, so from linear algebra the "transpose" of $\mathscr{F}$ should give a map $\mathscr{S}' \to \mathscr{S}'$. In other words,

2.29 Definition
If $\Phi \in \mathscr{S}'$ is a tempered distribution, we define its Fourier transform as the tempered distribution

$$\widehat{\Phi} : \mathscr{S} \ni \phi \mapsto \Phi(\widehat{\phi}).$$ ■

When $\psi \in \mathscr{S}$, we can define a distribution by $\phi \mapsto \int \phi\psi \, \mathrm{d}x$. We see that the above definition is compatible with the notion of Fourier transform for Schwartz functions by virtue of Parseval's identity (Lemma 2.18).

2.30 Example
Returning again to the δ distribution. Consider the distribution δ_0 centered at $x = 0$, and let us compute its Fourier transform. By definition

$$\int \widehat{\delta_0}\phi \, \mathrm{d}x = \int \delta_0 \widehat{\phi} \, \mathrm{d}\xi = \widehat{\phi}(0).$$

On the other hand, by the definition of the Fourier transform (2.1), we have

$$\widehat{\phi}(0) = \frac{1}{(2\pi)^{d/2}} \int \phi \, \mathrm{d}x.$$

So we can say that the Fourier transform of δ_0 is the constant function $(2\pi)^{-d/2}$. ■

2.31 Example
For a more involved example, let's consider the *principal value* distribution for functions on $\mathbb{R}$. Given $\phi \in \mathscr{S}(\mathbb{R})$, we can define

One should think of the principal value distribution as the "multiplicative inverse" of the distribution $\phi \mapsto x\phi$, which, through the Fourier transform, tells us that the principal value distribution is related to the inversion of differentiation. It appears in the definition of the Hilbert transform, and is the simplest example of a singular integral operator.

$$\text{p.v.}\frac{1}{x} : \phi \mapsto \lim_{\epsilon \to 0} \left(\int_{-\infty}^{-\epsilon} \frac{\phi(x)}{x} \, \mathrm{d}x + \int_{\epsilon}^{\infty} \frac{\phi(x)}{x} \, \mathrm{d}x \right).$$

First, let us verify that this is indeed a tempered distribution. We can compute

$$\int_{-\infty}^{-\epsilon} \frac{\phi(x)}{x} \, \mathrm{d}x + \int_{\epsilon}^{\infty} \frac{\phi(x)}{x} \, \mathrm{d}x = \left(\int_{-\infty}^{-1} + \int_{-1}^{-\epsilon} + \int_{\epsilon}^{1} + \int_{1}^{\infty} \right) \frac{\phi(x)}{x} \, \mathrm{d}x.$$

The first and last terms can be bounded by, for example

$$\left|\int_{-\infty}^{-1} \frac{\phi(x)}{x}\,\mathrm{d}x\right| \leq \int_{-\infty}^{-1} \frac{1}{x^2}\cdot\langle x\rangle\big|\phi(x)\big|\,\mathrm{d}x \leq \sup_{x\in\mathbb{R}}\langle x\rangle\big|\phi(x)\big|.$$

For the middle two terms, observe that

$$\int_{-1}^{-\epsilon} \frac{1}{x}\,\mathrm{d}x + \int_{\epsilon}^{1} \frac{1}{x} = 0,$$

we can write

$$\phi(x) = \phi(x) - \phi(0) + \phi(0)$$

to get

$$\int_{-1}^{-\epsilon} \frac{\phi(x)}{x}\,\mathrm{d}x + \int_{\epsilon}^{1} \frac{\phi(x)}{x}\,\mathrm{d}x = \int_{-1}^{-\epsilon} \frac{\phi(x)-\phi(0)}{x}\,\mathrm{d}x + \int_{\epsilon}^{1} \frac{\phi(x)-\phi(0)}{x}\,\mathrm{d}x.$$

Now, since $\phi(x)-\phi(0)$ evaluates to 0 as $x = 0$, and that ϕ is smooth, we have that $\phi(x) - \phi(0) = x\psi(x)$ for some smooth function ψ. Now, $\psi(x) = \frac{\phi(x)-\phi(0)}{x}$ and hence, by the mean value theorem, equals $\phi'(y)$ for some $y \in [0, x]$. This allows us to control uniformly

$$\left|\int_{-1}^{-\epsilon} \frac{\phi(x)}{x}\,\mathrm{d}x + \int_{\epsilon}^{1} \frac{\phi(x)}{x}\,\mathrm{d}x\right| \leq \int_{-1}^{1} \big|\psi(x)\big|\,\mathrm{d}x \leq 2\sup_{x\in\mathbb{R}}\big|\phi'(x)\big|.$$

Putting together the estimates we see that $\mathrm{p.v.}\frac{1}{x}$ is indeed a tempered distribution.

To compute its Fourier transform, first write

$$\left[\widehat{\mathrm{p.v.}\frac{1}{x}}\right](\phi) = \lim_{\epsilon\to 0}\left[\int_{-\infty}^{-\epsilon} \frac{\widehat{\phi}(\xi)}{\xi}\,\mathrm{d}\xi + \int_{\epsilon}^{\infty} \frac{\widehat{\phi}(\xi)}{\xi}\,\mathrm{d}\xi\right].$$

Now

$$\int_{-\infty}^{-\epsilon} \frac{\widehat{\phi}(\xi)}{\xi}\,\mathrm{d}\xi = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{-\epsilon}\int_{-\infty}^{\infty} \phi(x)\frac{e^{-ix\xi}}{\xi}\,\mathrm{d}x\,\mathrm{d}\xi.$$

So

$$\left[\widehat{\mathrm{p.v.}\frac{1}{x}}\right](\phi) = \frac{1}{\sqrt{2\pi}} \lim_{\epsilon\to 0} \int_{\epsilon}^{\infty} \int_{-\infty}^{\infty} \phi(x) \underbrace{\frac{e^{-ix\xi} - e^{ix\xi}}{\xi}}_{=-2i\sin(x\xi)/\xi} \, \mathrm{d}x \, \mathrm{d}\xi.$$

Since for fixed ξ, $\sin(x\xi)/\xi$ is an odd function, we can write

$$\left[\widehat{\mathrm{p.v.}\frac{1}{x}}\right](\phi) = -\frac{i}{\sqrt{2\pi}} \lim_{\epsilon\to 0} \int_{\epsilon}^{\infty} \int_{-\infty}^{\infty} [\phi(x) - \phi(-x)] \frac{\sin(x\xi)}{\xi} \, \mathrm{d}x \, \mathrm{d}\xi.$$

Now, since $\phi(x) - \phi(-x)$ is an odd Schwarz function, we can define $\Phi(x) = \int_{-\infty}^{x} \phi(y) - \phi(-y) \, \mathrm{d}y$. We note that $\Phi(x)$ is obviously even, and smooth, with $\lim_{x\to\pm\infty} \Phi(x) = 0$. The superpolynomial decay of ϕ implies that Φ also has rapid decay, and hence $\Phi \in \mathscr{S}$. We can then integrate by parts in x.

$$\begin{aligned}
\left[\widehat{\mathrm{p.v.}\frac{1}{x}}\right](\phi) &= \frac{i}{\sqrt{2\pi}} \lim_{\epsilon\to 0} \int_{\epsilon}^{\infty} \int_{-\infty}^{\infty} \Phi(x) \cos(x\xi) \, \mathrm{d}x \, \mathrm{d}\xi \\
&= \frac{i}{\sqrt{2\pi}} \int_{0}^{\infty} \int_{-\infty}^{\infty} \Phi(x) e^{ix\xi} \, \mathrm{d}x \, \mathrm{d}\xi
\end{aligned}$$

where we used that Φ is even. Since $\Phi \in \mathscr{S}$ we can replace by its Fourier transform

$$= i \int_{0}^{\infty} \widehat{\Phi}(-\xi) \, \mathrm{d}\xi.$$

By Proposition 2.5, we also have $\widehat{\Phi}$ is even. So

$$\begin{aligned}
&= \frac{i}{2} \int_{-\infty}^{\infty} \widehat{\Phi}(\xi) \, \mathrm{d}\xi \\
&= \frac{i}{2} \sqrt{2\pi} \Phi(0)
\end{aligned}$$

by the Fourier inversion formula. We can expand $\Phi(0) = \int_{-\infty}^{0} \phi(x) - \phi(-x) \, \mathrm{d}x$ to finally get

$$\left[\widehat{\mathrm{p.v.}\frac{1}{x}}\right](\phi) = \int_{-\infty}^{\infty} \left[-\frac{\sqrt{\pi} i}{\sqrt{2}} \operatorname{sgn}(x)\right] \phi(x) \, \mathrm{d}x$$

and identify the term in the brackets with the Fourier transform of the principal value distribution. ■

2.32 Remark

Formally integrating by parts the structure theorem, we have that tempered distributions can built from operations

$$\phi \mapsto \int_{\mathbb{R}^d} (-1)^{|\alpha|} \partial^\alpha p(x) \cdot \phi(x) \, \mathrm{d}x.$$

Therefore it is customary to identify a distribution $\Phi \in \mathscr{S}'$ with the object "$(-1)^{\|\alpha\|} \partial^\alpha p$". This is especially apt as for many reasonable functions f the operation

$$\phi \mapsto \int_{\mathbb{R}^d} f(x)\phi(x) \, \mathrm{d}x$$

is a tempered distribution, so rather than thinking of f and the operation it defines from the above expression as two different objects, it is convenient just to denote the corresponding tempered distribution also by the symbol f. We will adopt this convention throughout: when a function f is given on $\mathbb{R}^d$ and asserted to be a tempered distribution, the linear operation we refer to is the one defined above.

We've already seen this notation used earlier, when looking at the expression $\int_{\mathbb{R}^d} \delta_x(y)\phi(y) \, \mathrm{d}y$ for the evaluation map at x, phrased in terms of the Dirac δ_x. For "integration over a measurable subset Ω", the corresponding function would be the characteristic function $\mathbf{1}_\Omega$. And from the previous example, the Fourier transform of the principal value distribution should be identified with the function $-\frac{\sqrt{\pi} i}{\sqrt{2}} \operatorname{sgn}(x)$. ■

The method with which we defined the Fourier transform can also be extended to other linear operations that send $\mathscr{S}$ to itself. In particular, many of the operations listed in Lemma 2.15 extend to mappings of $\mathscr{S}'$. For example, if Φ is a tempered distribution, then we can define its derivative by the formula

$$\partial_x^\alpha \Phi : \mathscr{S} \ni \phi \mapsto \Phi((-1)^{|\alpha|} \partial_x^\alpha \phi). \tag{2.33}$$

(the factors of -1 come up because we are formally "integrating by parts", see the previous remark) and multiplication by a polynomial can be defined by

$$P \cdot \Phi : \mathscr{S} \ni \phi \mapsto \Phi(P \cdot \phi). \tag{2.34}$$

One can check that when Φ can be represented by integrating against a Schwartz function, the definitions above agree with the interpretation in Remark 2.32.

This method of definition is compatible with *linear* mappings of $\mathscr{S} \to \mathscr{S}$; however, this is not compatible with *bilinear* mappings. And so, products of tempered distributions and convolutions of tempered distributions are not, in general, well-defined. However, we can define, when $\Phi \in \mathscr{S}'$ and $\phi \in \mathscr{S}$ the operations

$$\Phi \cdot \phi : \mathscr{S} \ni \psi \mapsto \Phi(\phi \cdot \psi) \tag{2.35}$$

and

$$\Phi * \phi : \mathscr{S} \ni \psi \mapsto \Phi(\phi * \psi). \tag{2.36}$$

2.37 Remark
Convolution and point-wise multiplication, as mentioned above, cannot be defined as full bilinear mappings on $\mathscr{S}'$. They, however, can be defined as partial functions. For example, if $\Phi \in \mathscr{S}'$ is a distribution *with compact support* (this means that there exists a compact set $K \subset \mathbb{R}^d$ such that whenever $\phi \in \mathscr{S}$ vanishes identically on K, $\Phi(\phi) = 0$), then one can check that $\Phi * \phi$ is in fact in Schwartz class. This means that convolutions of tempered distributions against distributions of compact support are well-defined.

See Hörmander, The analysis of linear partial differential operators. I *for a more complete discussion of such properties of distributions.*

There are similar criteria for the point-wise product of tempered distributions to be well-defined; a thorough analysis of the situation requires the notion of the *wavefront set*, which can be thought of as a method to measure the degree of singularity of a tempered distribution in the full quantum phase space. We will not engage in a discussion here. ■

Uncertainty principle

The colloquial formulation of the uncertainty principle states that

> "You can't know the position and momentum of a particle at the same time."

Some aspects of this we have already seen in our discussion of the Vlasov equation. In Exercise 1.14 we explored the impact of lowered regularity of initial data on decay rates. The $W^{d,1}$ condition in the proof of Theorem 1.10 should be thought of as a "phase space regularity" condition that prevents the initial data to be concentrated around a hyperplane of dimension strictly less than d. This means that the initial phase-space distribution must be such that no more than half of the phase-space variables can be known to infinite precision. (A point particle is represented by a Dirac δ distribution.) So it is underneath this primitive version of uncertainty principle that dispersion holds.

The uncertainty principle is, however, baked into quantum systems. And some of the basic dispersive phenomena that we will described for quantum systems can trace their foundations, at least in part, to this simple concept. And the reason that the uncertainty principle appears really is due to the quantum phase space that underlies the equations: the uncertainty principle is, above all, a statement about the Fourier transform.

A first manifestation of the uncertainty principle comes as a direct consequence of the Paley-Wiener Theorem (see Exercise 2.11), and is commonly stated in the form "a function and its Fourier transform cannot both have compact support."

2.38 Corollary
Suppose $\phi \in L^1(\mathbb{R}^d)$ and has compact support, and $\widehat{\phi}$ also has compact support. Then $\phi \equiv 0$. ■

Proof By the Paley-Wiener Theorem we have that $\widehat{\phi}$ is real analytic from our hypotheses. The hypotheses also implies that $\widehat{\phi}$ vanish on some open set. By analyticity it vanishes everywhere. Parseval's identity implies then for every Schwartz function ψ, $\int_{\mathbb{R}^d} \phi\psi \, \mathrm{d}x = 0$. This implies that $\phi \equiv 0$ by Lusin's theorem. □

Lusin's Theorem states that given a measurable function ϕ on a compact set X, for every ϵ there exists a compact set $E \subset X$ where $X \setminus E$ has measure $< \epsilon$ and $\phi|_E$ is continuous.

2.39 Remark
The final steps of the above proof also gives a proof of the statement that $\mathscr{F} : L^1 \to L^\infty$ is injective. ■

A quantitative version of the uncertainty principle can be derived from the fact that the momentum operator $\nabla^{(x)}$ (which we can think of as being multiplication in frequency space thanks to Proposition 2.9) does not commute with the position operator x.

$$\partial_{x^i} x^i - x^i \partial_{x^i} = 1$$

as differential operators on functions. This manifests in the following computation. Let $f \in \mathscr{S}$. For each $i \in \{1,\ldots,d\}$:

$$\int_{\mathbb{R}^d} |f|^2 \, \mathrm{d}x = \int_{\mathbb{R}^d} (\partial_{x^i} x^i)|f|^2 \, \mathrm{d}x$$
$$= -\int_{\mathbb{R}^d} x^i \left(f \partial_{x^i} \overline{f} + \overline{f} \partial_{x^i} f\right) \, \mathrm{d}x$$

after integrating by parts. Now rewrite as

$$= -\int_{\mathbb{R}^d} x^i f \cdot \partial_{x^i} \overline{f} + x^i \overline{f} \cdot \partial_{x^i} f \, \mathrm{d}x$$

we can apply Cauchy-Schwarz to get

$$\leq 2 \left(\int_{\mathbb{R}^d} (x^i)^2 |f|^2 \, \mathrm{d}x \right)^{\frac{1}{2}} \left(\int_{\mathbb{R}^d} |\partial_{x^i} f|^2 \, \mathrm{d}x \right)^{\frac{1}{2}}.$$

With help with Proposition 2.9 and Proposition 2.23, we can rewrite the inequality as

$$\|f\|_{L^2}^2 \leq 2 \left\| x^i f \right\|_{L^2} \left\| \xi^i \widehat{f} \right\|_{L^2}. \tag{2.40}$$

Inequality (2.40) is usually called *Heisenberg's uncertainty principle*, where the name comes from interpreting $|f|^2$ as a probability density, and hence $(x^i)^2$ is the *variance* of the position operator x^i, interpreted as a random variable. This inequality should be understood as saying that the function f cannot be such that both f and $\hat{f}$ concentrate near the origin (the origin is not special here, due to translation invariance of the L^2 norm; concentration would imply, for example, $\left\| x^i f \right\|_{L^2} \approx 0$), without the function itself to be small.

We conclude this section with *Hardy's uncertainty principle*, which we won't prove here (it relies on complex analytic techniques, specifically the Phragmén-Lindelöf principle). Whereas the Paley-Wiener result mentioned above states that f and $\widehat{f}$ cannot both have compact support, Hardy's principle states that f and $\widehat{f}$ cannot both decay too fast. In fact, the joint decay rates exhibited by the Gaussian functions are optimal, and that the Gaussians are the unique optimizers.

2.41 Theorem (Hardy's uncertainty principle)
Let f be a Schwartz function. Suppose that there exists $C, \hat{C}$ such that the pointwise estimates

$$|f(x)| \leq C e^{-\frac{1}{2}|x|^2}, \qquad \left|\widehat{f}(\xi)\right| \leq \hat{C} e^{-\frac{1}{2}|x|^2}$$

hold for all x, ξ. Then $f(x) = A \exp -\frac{1}{2}|x|^2$ for some constant A. ■

2.42 Exercise
Let f be a Schwartz function. Suppose that there exists positive constants $C, \hat{C}, a, \hat{a}$ such that the pointwise estimates

$$|f(x)| \leq C \exp\left(-\frac{a}{2}|x|^2\right), \qquad \left|\widehat{f}(\xi)\right| \leq \hat{C} \exp\left(-\frac{\hat{a}}{2}|x|^2\right)$$

hold for all x, ξ. Show that

- If $a\hat{a} = 1$, then $f(x) = A \exp -\frac{a}{2}|x|^2$ for some constant A.
- If $a\hat{a} > 1$, then $f(x) \equiv 0$.

(Hint: use Hardy's uncertainty principle plus scaling.) ■

Some applications to dispersive equations

Let us return to our original purpose of understanding (1.27) as a representation formula for the solution to dispersive equations. The discussion above readily shows that, after prescribing Schwartz class initial data for the Schrödinger (1.29), Airy (1.30), and Klein-Gordon (1.31) equations, one can construct solutions that are Schwartz for every instant in time.

Take, for example, Schrödinger's equation. Taking the Fourier transform in the spatial variables formally gives us the equation

$$i\partial_t \widehat{\phi} + |\xi|^2 \widehat{\phi} = 0 \tag{2.43}$$

which is an ordinary differential equation for every fixed frequency ξ. This implies that

$$\widehat{\phi}(t, \xi) = e^{i|\xi|^2 t} \widehat{\phi}(0, \xi). \tag{2.44}$$

Noting that $\xi \mapsto \exp(i|\xi|^2 t)$ is smooth with norm 1 for every $t \in \mathbb{R}$, this implies that as long as $\widehat{\phi}(0, \bullet) \in \mathscr{S}$ (which would be the case if $\phi(0, \bullet) \in \mathscr{S}$), so is $\widehat{\phi}(t, \bullet)$ and hence also $\phi(t, \bullet)$.

Similarly, the Airy equation has Fourier transform

$$\partial_t \widehat{\phi} + i\xi^3 \widehat{\phi} = 0. \tag{2.45}$$

Thus its solutions can be recovered from

$$\widehat{\phi}(t,\xi) = e^{-i\xi^3 t}\widehat{\phi}(0,\xi), \tag{2.46}$$

which implies that Schwartz initial data gives rise to Schwartz (in space) solutions.

The Klein-Gordon equation, being second order, has to be treated slightly differently. The Fourier transform of the equation takes the form

$$\partial^2_{tt}\widehat{\phi} + (|\xi|^2 + 1)\widehat{\phi} = 0. \tag{2.47}$$

The general form of the solution to the ordinary differential equation is

$$\widehat{\phi}(t,\xi) = Ae^{it\langle\xi\rangle} + Be^{-it\langle\xi\rangle}.$$

The two amplitudes A and B corresponds to the fact that, for a second order equation, we need to prescribe both the initial value and the initial velocity. Rewriting we get

$$\widehat{\phi}(t,\xi) = \cos(t\langle\xi\rangle)\widehat{\phi}(0,\xi) + \frac{\sin(t\langle\xi\rangle)}{\langle\xi\rangle}\widehat{\partial_t\phi}(0,\xi). \tag{2.48}$$

Since $\cos(t\langle\xi\rangle)$ and $\sin(t\langle\xi\rangle)/\langle\xi\rangle$ are both smooth, bounded functions, this implies that as long as the initial data $\phi(0,\bullet)$ and $\partial_t\phi(0,\bullet)$ are Schwartz class, so is $\phi(t,\bullet)$ for every time t.

2.49 Exercise
Prove that the functions

$$\mathbb{R}^d \ni \xi \mapsto \cos(|\xi|)$$

and

$$\mathbb{R}^d \ni \xi \mapsto \frac{\sin(|\xi|)}{|\xi|}$$

are smooth and bounded. As a consequence show that given $\phi_0, \phi_1 \in \mathscr{S}(\mathbb{R}^d)$, there exists a solution ϕ of the wave equation (1.32) on $\mathbb{R}\times\mathbb{R}^d$ such that $\phi(0,x) = \phi_0(x)$ and $\partial_t\phi(0,x) = \phi_1(x)$, and that $\phi(t,\bullet) \in \mathscr{S}(\mathbb{R}^d)$. ■

The Fourier representation formulae (2.44), (2.46), and (2.48) involves *multiplying* the Fourier transform of the initial data with a bounded smooth function. Now, recall that bounded smooth functions are in fact tempered distributions. One can easily show that Proposition 2.21 also holds if one takes the product or convolution of $\Phi \in \mathscr{S}'$ with $\phi \in \mathscr{S}$. Therefore, the discussion above proves

2.50 Theorem (Existence of fundamental solution)
For every $t \in \mathbb{R}$, there exist tempered distributions

$$G_t^{(\mathrm{Sch})}, G_t^{(\mathrm{Airy})}, G_t^{(\mathrm{KG})}, G_t^{(\mathrm{wave})} \in \mathscr{S}'(\mathbb{R}^d)$$

such that:

- The function $\phi(t,x) = G_t^{(\mathrm{Sch})} * \phi_0(x)$ solves the Schrödinger equation (1.29) with initial data $\phi(0,x) = \phi_0(x)$.
- The function $\phi(t,x) = G_t^{(\mathrm{Airy})} * \phi_0(x)$ solves the Airy equation (1.30) with data $\phi(0,x) = \phi_0(x)$.
- The function $\phi(t,x) = (\partial_t G_t^{(\mathrm{KG})}) * \phi_0(x) + G_t^{(\mathrm{KG})} * \phi_1(x)$ solves the Klein-Gordon equation (1.31) with data $\phi(0,x) = \phi_0(x)$ and $\partial_t \phi(0,x) = \phi_1(x)$.
- The function $\phi(t,x) = (\partial_t G_t^{(\mathrm{wave})}) * \phi_0(x) + G_t^{(\mathrm{wave})} * \phi_1(x)$ solves the wave equation (1.32) with initial data $\phi(0,x) = \phi_0(x)$ and $\partial_t \phi(0,x) = \phi_1(x)$.

The initial data ϕ_0, ϕ_1 are assumed to be in $\mathscr{S}$. ■

2.51 Remark
We can formally write, for example,

$$G_t^{(\mathrm{Sch})}(x) = \frac{1}{(2\pi)^d} \int_{\mathbb{R}^d} \exp[it|\xi|^2 + ix\cdot\xi] \,\mathrm{d}\xi.$$

This should be compared to (1.27); the integrand is simply the set of *all* monochromatic plane wave solutions to Schrödinger's equation. Similar computations can be performed for the other fundamental solutions.

Note that the power of the factor 2π is d: this is due to $d/2$ coming from the inverse Fourier transform, and $d/2$ coming from Proposition 2.21.

Necessarily, at $t = 0$, we have that $G_0^{(\mathrm{Sch})} = G_0^{(\mathrm{Airy})} = \delta_0$. For the Klein-Gordon and wave equations, it is the terms $\partial_t G_0^{(\mathrm{KG})}$ and $\partial_t G_0^{(\mathrm{wave})}$ that equals the Dirac δ. ■

Ref. 2.21: "Fourier transform properties: convolution and product"

In the subsequent chapters, we will study the qualitative and quantitative properties of the fundamental solutions described above. Here we give one simple example.

Variations of this technique lead to the RAGE theorem (due separately to Ruelle, Amrein and Georgescu, and Enss) of quantum dynamics; see Chapter 5 in Teschl, Mathematical methods in quantum mechanics.

2.52 *Example (Qualitative decay for solutions to Schrödinger's equation)* Consider the formula (2.44), which implies

$$\phi(t,0) = \frac{1}{(2\pi)^{d/2}} \int_{\mathbb{R}^d} e^{i|\xi|^2 t} \widehat{\phi_0}(\xi) \, \mathrm{d}\xi.$$

Express the integral now in polar coordinates $(r,\omega) \in \mathbb{R}_+ \times \mathbb{S}^{d-1}$, we have

$$\phi(t,0) = \frac{1}{(2\pi)^{d/2}} \int_0^\infty \int_{\mathbb{S}^{d-1}} e^{ir^2 t} \widehat{\phi_0}(r\omega) r^{d-1} \, \mathrm{d}\omega \, \mathrm{d}r.$$

Now perform the change of variables $\rho = r^2$, we get

$$\phi(t,0) = \frac{1}{2\cdot(2\pi)^{d/2}} \int_0^\infty \int_{\mathbb{S}^{d-1}} e^{i\rho t} \widehat{\phi_0}(\rho\omega) \rho^{(d-2)/2} \, \mathrm{d}\omega \, \mathrm{d}\rho.$$

Let ψ be the function on $\mathbb{R}$ defined by

$$\psi(\rho) \stackrel{\text{def}}{=} \begin{cases} \rho^{(d-2)/2} \int_{\mathbb{S}^{d-1}} \widehat{\phi_0}(\rho\omega) \, \mathrm{d}\omega, & \rho \geq 0; \\ 0, & \rho < 0. \end{cases}$$

Then

$$\phi(t,0) = \frac{1}{2\cdot(2\pi)^{(d-1)/2}} \widehat{\psi}(-t).$$

Since $\phi_0 \in \mathscr{S}$, we have that $\widehat{\phi_0}$ decays rapidly and is bounded near the origin, and therefore $\psi(\rho)$ is absolutely integrable (as $(d-2)/2 > -d$ for every $d \geq 1$). This means that $\phi(t,0)$ is given by the Fourier transform of an absolutely integrable function, and hence by the Riemann-Lebesgue lemma Exercise 2.10 converges to 0 as $t \to \pm\infty$. Note further that there is nothing special about the origin: we can apply the translation operator to ϕ and run the proof again centered at any point $x \in \mathbb{R}^d$. So as a consequence we have proven that: if ϕ solves Schrödinger's equation (1.29) with Schwartz class initial data, then for any $x \in \mathbb{R}^d$, it holds that $\lim_{t\to\infty} \phi(t,x) = 0$.

The same result holds also for solutions to the Airy, Klein-Gordon, and wave equations, essentially by the same argument. ■

Chapter 3

Oscillatory Integrals: Boundedness and Decay

In the previous chapter we discussed the basics of Fourier theory; in this chapter we will talk about some more advanced applications. The starting point is the Riemann-Lebesgue Lemma, its application in Example 2.52, and Proposition 2.9.

Ref. 2.52: "Decay of Schrödinger via Riemann-Lebesgue"

Ref. 2.9: "Fourier transform properties: differentiation"

Let us recall: by the Riemann-Lebesgue Lemma (Exercise 2.10), the Fourier transform $\widehat{\phi}$ of an absolutely integrable function $\phi \in L^1(\mathbb{R}^d)$ is not only uniformly continuous, but also decays to zero as $|\xi| \to \infty$. On the other hand, this decay carries no guaranteed rate by Exercise 2.24. A rate of decay however is guaranteed when we are willing to sacrifice derivatives. By Proposition 2.9, when ϕ is in the Sobolev space $W^{1,1}(\mathbb{R}^d)$, we can relate the Fourier transforms $\mathscr{F}[\partial_j \phi](\xi) = i\xi^j \widehat{\phi}(\xi)$. Since $\mathscr{F}[\partial_j \phi]$ is a bounded function, this implies that

$$\left|\widehat{\phi}(\xi)\right| \le \left\|\phi\right\|_{W^{1,1}} \langle \xi \rangle^{-1}.$$

More generally, if $\phi \in W^{k,1}(\mathbb{R}^d)$ then $\left|\widehat{\phi}(\xi)\right| \le \left\|\phi\right\|_{W^{k,1}} \langle \xi \rangle^{-k}$, a fact which motivated our definition of the Schwartz space $\mathscr{S}$.

It is natural to ask whether we can do the same thing to get higher rates of time-decay for solutions to Schrödinger's equation, in view of the application of the Riemann-Lebesgue Lemma to derive a "rate-less" decay in Example 2.52. The answer is mixed. Revisiting the proof: we applied the Riemann-Lebesgue Lemma to the function ψ that vanishes on the negative half line and is equal to

$$\psi(\rho) = \rho^{(d-2)/2} \int_{\mathbb{S}^{d-1}} \widehat{\phi_0}(\rho\omega)\, \mathrm{d}\omega$$

when $\rho \geq 0$, where $\widehat{\phi_0} \in \mathscr{S}$ is given. This function ψ is absolutely integrable: in fact, rapidly decreasing near infinity. However, ψ is not, in general, in $\mathscr{S}$. Given that

$$\lim_{\rho \to 0^+} \int_{\mathbb{S}^{d-1}} \widehat{\phi_0}(\rho\omega)\, \mathrm{d}\omega = \left|\mathbb{S}^{d-1}\right| \cdot \widehat{\phi_0}(0),$$

we have that ψ cannot be continuously differentiable more than $\frac{d-2}{2}$ times at the origin, since $\psi(\rho)$ is identically zero to the left and grows like $\rho^{\frac{d-2}{2}}$ to the right. So the integration by parts argument can be applied only $\lfloor\frac{d-2}{2}\rfloor$ times, resulting in the following proposition.

3.1 Proposition
Let ϕ be a solution to Schrödinger's equation (1.29) with initial data $\phi_0 \in \mathscr{S}$ on dimension $d \geq 3$. Then for every x there exists some constant C depending on ϕ_0 and x such that

$$\left|\phi(t,x)\right| \leq C\langle t\rangle^{-\lfloor\frac{d-2}{2}\rfloor}.$$

■

3.2 Remark
The improved decay to the Fourier transform can also be studied for initial data in Hölder classes $C^{k,\sigma}$; in the case of odd dimensions this allows us to bridge the gap between $\frac{d-2}{2}$ and $\lfloor\frac{d-2}{2}\rfloor$. Nonetheless, as we will see later on in this chapter, the sharp decay rate for solutions to Schrödinger's equation is $t^{-d/2}$, so while the method above comes close, it still leaves a gap. Therefore we will not focus too much attention discussing the details of this method; what one should keep in mind from this illustration is that firstly, decay rates generally increase with the dimension d (in full agreement with what we saw with Vlasov's equation), and secondly, the obstacle of decay happens where the dispersion relation $\omega = \omega(\xi) = |\xi|^2$ has a critical point. At this stage the connection between the critical point and the finite rates of decay is a bit harder to see, and explaining it will be the point of departure for this chapter. ■

Stationary phase: obstructions to decay

Let us begin by examining the proof of Proposition 2.9 in more detail. Let $\phi \in \mathscr{S}(\mathbb{R})$. Let us consider the integral

Ref. 2.9: "Fourier transform properties: differentiation"

$$\int \phi(x) e^{-ix\xi} \, \mathrm{d}x.$$

Using that $\partial_x(x\xi) = \xi$, we can write

$$\begin{aligned} \int \phi(x) e^{-ix\xi} \, \mathrm{d}x &= \int \phi(x) \cdot \frac{-i\xi}{-i\xi} \cdot e^{-ix\xi} \, \mathrm{d}x \\ &= \frac{i}{\xi} \int \phi(x) \cdot \partial_x e^{-ix\xi} \, \mathrm{d}x = -\frac{i}{\xi} \int \partial_x \phi(x) e^{-ix\xi} \, \mathrm{d}x. \end{aligned} \tag{3.3}$$

In the last equality, we integrated by parts. The exact argument of (3.3) can be generalized. Recall that the *support* of a (continuous) function is the closed set

$$\operatorname{supp} \phi \stackrel{\text{def}}{=} \overline{\{x \in \mathbb{R}^d \mid \phi(x) \neq 0\}}. \tag{3.4}$$

For convenience we will set the following notation for use in this chapter.

For more about oscillatory integrals, see Stein, Harmonic analysis: real-variable methods, orthogonality, and oscillatory integrals.

3.5 Definition
Let $\eta : \mathbb{R}^d \to \mathbb{R}$ be a smooth function, and let $\phi \in \mathscr{S}(\mathbb{R}^d)$. For every $\lambda \in \mathbb{R}$ we introduce the notation

$$I_{\eta,\phi}(\lambda) \stackrel{\text{def}}{=} \int_{\mathbb{R}^d} \phi(x) e^{i\lambda\eta(x)} \, \mathrm{d}x.$$

The quantity $I_{\eta,\phi}(\lambda)$ is often called an *oscillatory integral of the first kind* in the literature. ■

3.6 Lemma
Let $\eta : \mathbb{R}^d \to \mathbb{R}$ be a smooth function, and let $\phi \in \mathscr{S}(\mathbb{R}^d)$. Suppose there exists $c > 0$ such that $|\nabla \eta(x)| \geq c$ for every $x \in \operatorname{supp}\phi$. Then for every $N \in \mathbb{N}$, there exists a constant C_N which depends on c, N, and on $\sup_{\alpha \leq N+1} \sup_{x \in \operatorname{supp}\phi} |\partial^\alpha \eta(x)|$, such that

$$|I_{\eta,\phi}(\lambda)| \leq \frac{C_N}{\langle \lambda \rangle^N} \|\phi\|_{W^{N,1}}.$$

■

Proof Using that $\nabla \exp(i\lambda\eta) = i\lambda\nabla\eta\exp(i\lambda\eta)$, we have

$$\phi e^{i\lambda\eta} = \frac{\phi}{i\lambda\left|\nabla\eta\right|^2}\nabla\eta\cdot\nabla e^{i\lambda\eta}. \tag{3.7}$$

Integrating by parts this means

$$I_{\eta,\phi}(\lambda) = \frac{i}{\lambda}\int \nabla\cdot\frac{\phi\nabla\eta}{\left|\nabla\eta\right|^2}e^{i\lambda\eta}\,\mathrm{d}x.$$

Writing L the operator

$$L\phi = \nabla\cdot\frac{\phi\nabla\eta}{\left|\nabla\eta\right|^2}$$

we see, by induction

$$\left|I_{\eta,\phi}(\lambda)\right| \leq \lambda^{-N}\int\left|L^{(N)}\phi\right|\,\mathrm{d}x.$$

Point-wise $L^{(N)}\phi$ is bounded by up-to-N derivatives of ϕ multiplied against factors involving up-to-$(N+1)$ derivatives of η, with a linear dependence on the ϕ terms. And hence putting the ϕ terms in L^1 we get the claimed decay rates, after noting also that $\left|I_{\eta,\phi}(\lambda)\right| \leq \left\|\phi\right\|_{L^1}$ by definition. □

Don't worry too much if you don't find this relation obvious. Roughly speaking the singularity means that if you try to rewrite $I_{\eta,\phi}(\lambda)$ in the form of a one-dimensional Fourier transform, whatever change of variables you make there will be a "fold" in the parametrization. The connection can be made a bit clearer if you look through Whitney, "Singularities of Mappings of Euclidean Spaces".

Ref. 2.52: "Decay of Schrödinger via Riemann-Lebesgue"

In the case of the standard Fourier transform, fix $\omega \in \mathbb{S}^{d-1}$ and let $\eta(x) = \omega\cdot x$. This function being linear we have that its gradient has nonzero norm. Writing ξ in polar coordinates as $\lambda\omega$ with $\lambda \in \mathbb{R}_+$, we have that the Fourier transform $\widehat{\phi}(\xi)$ is proportional to $I_{\eta,\phi}(-\lambda)$ for the above-chosen η and λ.

The crucial condition in this lemma, however, is that $\left|\nabla\eta\right| \geq c$. The appearance of $\left|\nabla\eta\right|$ in the denominator of (3.7) is related to the finite differentiability of ψ in the discussion of Example 2.52 at the beginning of this chapter. Returning to the representation formulae (2.44), (2.46), and (2.48), we see that t plays the role of λ, and $\eta(\xi)$ takes the forms $|\xi|^2$, ξ^3, and $\pm\langle\xi\rangle$ respectively. In all three of those cases $\nabla\eta(0) = 0$, and therefore Lemma 3.6 does not apply. The main results of this chapter are extensions which can be applied to study the decay rates of solutions to the Schrödinger, Airy, and Klein-Gordon equations.

Estimates when $d = 1$

As often happens, we understand the situation in the one dimensional case much better. Using the special structures of the one dimensional real line, we can get sharper estimates compared to Lemma 3.6. The first results proven in this section are frequently referred to collectively under the name of "Van der Corput Lemma". Throughout this section we will let (a, b) be an open interval, where the end-points are allowed to be infinite. We take η to be a smooth, real-valued function on (a, b) with continuous extension to its closure, and ϕ a smooth, complex-valued function with continuous extension to the closure of (a, b). We also assume that ϕ and its first derivative ϕ' are absolutely integrable on (a, b).

3.8 Lemma (Van der Corput, part 1)
Suppose that $|\eta'| \geq 1$ on its domain of definition, and suppose further that η' is monotonic. Then

$$\left| \int_a^b \phi e^{i\lambda\eta} \, \mathrm{d}x \right| \leq \frac{1}{\lambda} \left[\|\phi'\|_{L^1} + 3\|\phi\|_{L^\infty} \right].$$

Furthermore, if ϕ has compact support in (a, b), then the estimate can be improved to

$$\left| \int_a^b \phi e^{i\lambda\eta} \, \mathrm{d}x \right| \leq \frac{2}{\lambda} \|\phi'\|_{L^1}.$$

■

Proof We proceed as in the proof of Lemma 3.6. Noting that

$$\phi e^{i\lambda\eta} = \frac{\phi}{i\lambda\eta'} \nabla e^{i\lambda\eta},$$

integrating by parts we get

$$\int_a^b \phi e^{i\lambda\eta} \, \mathrm{d}x = \frac{i}{\lambda} \int_a^b \left(\frac{\phi'}{\eta'} - \frac{\phi\eta''}{(\eta')^2} \right) e^{i\lambda\eta} \, \mathrm{d}x + \frac{\phi}{i\lambda\eta'} e^{i\lambda\eta} \Big|_a^b.$$

Using that $|\eta'| \geq 1$, the first term in the integral we can bound by $\lambda^{-1} \|\phi'\|_{L^1}$.

For the second term we use

$$\left|\int_a^b \frac{\phi\eta''}{(\eta')^2} e^{i\lambda\eta}\, \mathrm{d}x\right| \leq \|\phi\|_{L^\infty} \int_a^b \left|\frac{\eta''}{(\eta')^2}\right| \mathrm{d}x.$$

The assumption that η' is monotonic, however, implies that

$$\int_a^b \left|\frac{\eta''}{(\eta')^2}\right| \mathrm{d}x = \left|\int_a^b \frac{\eta''}{(\eta')^2}\right| = \left|\frac{1}{\eta'(b)} - \frac{1}{\eta'(a)}\right| \leq 1.$$

Putting everything together gives

$$\left|\int_a^b \phi e^{i\lambda\eta}\, \mathrm{d}x\right| \leq \frac{1}{\lambda}\Big[\|\phi'\|_{L^1(a,b)} + \|\phi\|_{L^\infty(a,b)} + |\phi(a)| + |\phi(b)|\Big], \tag{3.9}$$

which immediately implies the desired results. □

Comparing Lemma 3.8 and Lemma 3.6, we see that Van der Corput's lemma has the extra assumption that ϕ' is monotonic. But from this it sharpens the decay result so that the constant C_N, which in Lemma 3.6 would depend on the second derivative of η, now has no dependence on higher derivatives of η, nor on the length of the interval (a,b). This allows us to generalize the result to deal with cases where η has critical points. A first example being:

3.10 Lemma (Van der Corput, part 2)
Suppose $|\eta''| \geq 1$ on its domain of definition. Then

$$\left|\int_a^b \phi e^{i\lambda\eta}\, \mathrm{d}x\right| \leq \frac{1}{\sqrt{|\lambda|}}\Big[\|\phi'\|_{L^1} + 8\|\phi\|_{L^\infty}\Big].$$

■

Proof Observe that if $\eta'(x_0) = 0$, then by the fundamental theorem of calculus

$$\eta'(x_0 + y) = \int_{x_0}^{x_0+y} \eta''(z)\, \mathrm{d}z$$

and so with the hypothesized bound $\left|\eta''\right| \geq 1$ (which implies η'' has constant sign) we get

$$\left|\eta'(x_0 + y)\right| \geq y.$$

This means that for any $\delta > 0$, among (a, b) there exists at most one unique interval I, with length no more than 2δ, on which $\left|\eta'\right| \leq \delta$. The complement of I can thus be written as the union of at most two intervals.

We split the integral

$$\int_a^b \phi e^{i\lambda\eta}\, \mathrm{d}x = \int_I \phi e^{i\lambda\eta}\, \mathrm{d}x + \int_{I^c} \phi e^{i\lambda\eta}\, \mathrm{d}x.$$

Within the region I we estimate

$$\left|\int_I \phi e^{i\lambda\eta}\, \mathrm{d}x\right| \leq |I| \left\|\phi\right\|_{L^\infty} \leq 2\delta \left\|\phi\right\|_{L^\infty}.$$

Let J be a connected component of I^c. The function $\delta^{-1}\eta$ satisfies $\left|(\delta^{-1}\eta)'\right| \geq 1$ on J with $(\delta^{-1}\eta)'$ being monotonic. So we can apply Lemma 3.8 to get

Ref. 3.8: "Van der Corput: non-stationary case"

$$\left|\int_J \phi e^{i\lambda\eta}\, \mathrm{d}x\right| = \left|\int_J \phi e^{i(\delta\lambda)\cdot(\delta^{-1}\eta)}\, \mathrm{d}x\right| \leq \frac{1}{\delta\lambda}\left[\left\|\phi'\right\|_{L^1(J)} + 3\left\|\phi\right\|_{L^\infty}\right].$$

Therefore, since there are at most two pieces, we get

$$\left|\int_{I^c} \phi e^{i\lambda\eta}\, \mathrm{d}x\right| \leq \frac{1}{\delta\lambda}\left[\left\|\phi'\right\|_{L^1(I^c)} + 6\left\|\phi\right\|_{L^\infty}\right].$$

Combining the estimate on I and I^c we get the lemma as claimed. □

The trick going from one derivative to two derivatives can be repeated inductively.

3.11 Lemma (Van der Corput, part 3)
Suppose the kth derivative $\left|\eta^{(k)}\right| \geq 1$ on its domain of definition, for some $k \geq 2$. Then

$$\left|\int_a^b \phi e^{i\lambda\eta}\, \mathrm{d}x\right| \leq \frac{1}{|\lambda|^{\frac{1}{k}}}\left[\left\|\phi'\right\|_{L^1} + (5\cdot 2^{k-1} - 2)\left\|\phi\right\|_{L^\infty}\right].$$

■

3.12 Exercise
Prove Lemma 3.11 by induction on k. ■

3.13 Example (Applications to dispersive equations)
As an immediate application, let us revisit some of the representation formulae discussed before. Fix the dimension $d = 1$. Looking back into the discussion of Example 2.52, for Schrödinger's equation, the solution can be represented by

Ref. 2.52: "Decay of Schrödinger via Riemann-Lebesgue"

$$\phi(t,x) = \frac{1}{\sqrt{2\pi}} \int_{\mathbb{R}} e^{it\xi^2} e^{ix\xi} \widehat{\phi_0}(\xi) \, \mathrm{d}\xi.$$

Setting $\eta(\xi) = \frac{1}{2}\xi^2$ and grouping together $e^{ix\xi}\widehat{\phi_0}(\xi)$ as the amplitude, an application of Lemma 3.10 gives us that

$$\left|\phi(t,x)\right| \leq \frac{1}{\sqrt{4\pi|t|}} \left[\left\|\widehat{\phi_0}'\right\|_{L^1} + |x| \left\|\widehat{\phi_0}\right\|_{L^1} + 8\left\|\widehat{\phi_0}\right\|_{L^\infty} \right]. \qquad (3.14)$$

We see that this already gives us an improvement over both Example 2.52 and Proposition 3.1.

A similar result holds for solutions to Airy's equation, where the decay rate found is now $|t|^{-1/3}$; this uses Lemma 3.11. For solutions to the linear wave equation in $d = 1$, the situations is vastly simpler. The dispersion relation for the wave equation in one dimension implies that, if the initial data (both position and velocity) is in $\mathscr{S}$, then there exists Schwartz functions ϕ_+ and ϕ_- such that

$$\phi(t,x) = \frac{1}{\sqrt{2\pi}} \int_{\mathbb{R}} \phi_+(\xi) e^{it\xi} e^{ix\xi} + \phi_-(\xi) e^{-it\xi} e^{ix\xi} \, \mathrm{d}\xi.$$

And so for fixed x we can conclude that $\phi(t,x)$ decays faster than any polynomial: the same conclusion can be reached using the method of characteristics in physical space.

The Klein-Gordon equation requires a small trick. Observe that the phase function is $\eta(\xi) = \langle\xi\rangle$. We see that $\eta'' = \langle\xi\rangle^{-3}$ is not bounded below, and so Lemma 3.10 cannot be applied directly. To get around this problem, notice that η'' and η' are not small both at once. More precisely, for $|\xi| \geq 1$, we have that $\left|\eta'\right| = \frac{|\xi|}{\langle\xi\rangle} \geq \frac{1}{2}$, while for $|\xi| \leq 1$, we have $\left|\eta''\right| = \langle\xi\rangle^{-3} \geq \frac{1}{8}$. Therefore we can apply Lemma 3.8 to the region $\{|\xi| \geq 1\}$ and Lemma 3.10 to the region $\{|\xi| \leq 1\}$ to conclude that solutions to the Klein-Gordon equation, for fixed $x \in \mathbb{R}$, must decay like $t^{-1/2}$ in time. ■

3.15 Remark (Is the wave equation dispersive?)
Let us re-examine the wave equation (1.32)

$$\partial^2_{tt}\phi - \triangle\phi = 0$$

which, we recall is based on the dispersion relation $\omega^2 = |k|^2$. This we can rewrite as $\omega = \pm\frac{k}{|k|}\cdot k$, and from this we derive that its *classical* analogue equation would be the following modification of Vlasov's equation:

$$\partial_t\rho + \frac{v}{|v|}\cdot\nabla^{(x)}\rho = 0. \tag{3.16}$$

The fact that the velocity dependence is of the form $v/|v|$ means that we are essentially losing one dimension worth of dispersion: particles of "momenta" v and λv for any $\lambda > 0$ will travel with the same speed, and so will not separate spatially.

This is especially noticeable in dimension $d = 1$. The dispersion relation means that all positive momenta particles will be moving with one speed, and all negative momenta ones will be moving with another, and so *in dimension 1, the wave equation behaves not like the kinetic theory picture, but like the N-particle picture*. And in particular, in dimension $d = 1$ the wave equation cannot really be considered as dispersive. Indeed, by the method of characteristics we see that solutions of the linear wave equation in one dimension, like the solutions in the N-particle picture (Exercise 1.4), has a threshold below which the amplitude cannot decay.

In higher dimensions, however, some remnants of dispersion remain available. Due to the fact that particles travel in the direction given by their momentum, *angular* dispersion is still available, and so by dimension counting we expect that wave equation in $d \geq 2$ dimensions decays like Schrödinger equation in $(d-1)$ dimensions. ∎

3.17 Exercise
The equation (3.16) can be rephrased as follows. Fix $d \geq 2$. Let $\rho : \mathbb{R}\times\mathbb{R}^d\times\mathbb{S}^{d-1} \to \overline{\mathbb{R}_+}$ be the distribution function. Assume ρ solves

$$\partial_t\rho(t,x,\omega) + \omega\cdot\nabla^{(x)}\rho(t,x,\omega) = 0,$$

where we naturally identify $\omega \in \mathbb{S}^{d-1}$ with a corresponding unit vector in $\mathbb{R}^d$. Formulate and prove an analogue of Theorem 1.10 for ρ. ∎

3.18 Remark
The discussion of the Klein-Gordon equation in Example 3.13 above brings up an important technique when applying Van der Corput Lemmas. Since

the Lemmas are independent of the length of the interval (a,b), we can freely localize the estimates by chopping up our domain into smaller pieces. For considering asymptotic behavior, one can first restrict to the set where η' is bounded below by some positive constant. Only on the remainder, where η' is small, do we need to consider the higher order Van der Corput estimates. So among the points where η' is small, we use Lemma 3.10 for the subset where η'' is bounded below, and only on the remainder where both η' and η'' are small do we start consider higher order estimates like those in Lemma 3.11. We return to this again below when we discuss the asymptotic behavior of oscillatory integrals. ■

The Van der Corput lemma establishes a decaying upper bound for the oscillatory integral. A natural follow-up question is: can the integral

$$\int_a^b \phi e^{i\lambda\eta}\,\mathrm{d}x$$

be developed as an asymptotic series in inverse powers of λ? The answer to this question is encapsulated in the *method of stationary phase*. We will not give a full account of the method here; rather, we content ourselves with an illustration on how to compute the leading term of the asymptotic series for solutions to the Schrödinger equation.

3.19 Example (Leading order asymptotic for 1 dimensional Schrödinger) Consider the integral

$$\int_{\mathbb{R}} \phi(x) e^{i\lambda x^2}\,\mathrm{d}x,$$

where $\phi \in \mathscr{S}$. Knowing that this integral decays at least as fast as $\lambda^{-1/2}$, we can ask about the limit

$$\lim_{\lambda\to\infty} \lambda^{\frac{1}{2}} \int_{\mathbb{R}} \phi(x) e^{i\lambda x^2}\,\mathrm{d}x = ?$$

A first thing to notice is that if our smooth ϕ is such that $\phi(0) = 0$, then the limit vanishes. This is due to the observation that if ϕ is smooth and $\phi(0) = 0$, then there exists some smooth ψ such that $\phi(x) = x\psi(x)$; we have already seen this argument in action implicitly in Example 2.31, and a stand-alone proof is given in Lemma 3.21 below. Using the observation, we see that when $\phi(0) = 0$, the integral

$$\int_{\mathbb{R}} \phi(x) e^{i\lambda x^2} \,\mathrm{d}x = \int_{\mathbb{R}} \psi(x) x e^{i\lambda x^2} \,\mathrm{d}x = -\frac{i}{2\lambda} \int_{\mathbb{R}} \psi(x) \nabla e^{i\lambda x^2} \,\mathrm{d}x.$$

So after integrating by parts we obtain, as claimed, that

$$\lim_{\lambda\to\infty} \lambda^{\frac{1}{2}} \left| \int_{\mathbb{R}} \phi(x) e^{i\lambda x^2} \,\mathrm{d}x \right| \leq \lim_{\lambda\to\infty} \frac{1}{2\lambda^{\frac{1}{2}}} \left\| \psi' \right\|_{L^1} = 0.$$

By the linearity of the integral in ϕ, we then conclude that

$$\lim_{\lambda\to\infty} \lambda^{\frac{1}{2}} \int_{\mathbb{R}} \phi(x) e^{i\lambda x^2} \,\mathrm{d}x$$

must be proportional to $\phi(0)$. It thus remains to compute what the constant of proportionality is. Our argument above showed that we can choose $\phi(x)$ at our convenience for this computation, so we will choose $\phi(x) = \exp(-\frac{1}{2}x^2)$. We can rewrite

$$-\frac{1}{2}x^2 + i\lambda x^2 = -\frac{1}{2}\langle 2\lambda\rangle x^2 \cdot \frac{1 - 2i\lambda}{\langle 2\lambda\rangle};$$

the final fraction has norm 1, and can be written as $e^{2i\theta_0}$ for some $\theta_0 \in (-\pi/4, \pi/4)$. Now using that the mapping $z \mapsto \exp -z^2$ for $z \in \mathbb{C}$ is holomorphic, and on the sector $z = re^{i\theta}$ with $\theta \in (-\pi/4, \pi/4)$ we have that $\exp(-z^2)$ decays to zero at infinity. So by contour integration, we conclude that

$$\int_{\mathbb{R}} e^{-\frac{1}{2}x^2 + i\lambda x^2} e^{i\theta_0} \,\mathrm{d}x = \frac{1}{\langle 2\lambda\rangle^{\frac{1}{2}}} \int_{\mathbb{R}} e^{-\frac{1}{2}x^2} \,\mathrm{d}x = \left(\frac{2\pi}{\langle 2\lambda\rangle} \right)^{\frac{1}{2}}.$$

Noting that as $\lambda \to \infty$ we have $\theta_0 \to -\pi/4$, we conclude that

$$\lim_{\lambda\to\infty} \lambda^{\frac{1}{2}} \int_{\mathbb{R}} \phi(x) e^{i\lambda x^2} \,\mathrm{d}x = e^{i\pi/4}\sqrt{\pi}\phi(0), \tag{3.20}$$

and thereby giving the first term of the asymptotic expansion of the oscillatory integral.

Now, returning to the actual Schrödinger's equation, we have that in the case $d = 1$

$$\phi(t,x) = \frac{1}{\sqrt{2\pi}} \int_{\mathbb{R}} e^{it\xi^2} e^{ix\xi} \widehat{\phi_0}(\xi)\, \mathrm{d}\xi.$$

And hence asymptotically

$$\lim_{t\to\infty} t^{\frac{1}{2}} \phi(t,x) = \frac{e^{i\pi/4}}{\sqrt{2}} \widehat{\phi_0}(0) = \frac{e^{i\pi/4}}{2\sqrt{\pi}} \int_{\mathbb{R}} \phi_0(y)\, \mathrm{d}y.$$

■

3.21 Lemma (Baby Malgrange Preparation)
If $\phi \in \mathscr{S}(\mathbb{R})$ and $\phi(0) = 0$, then there exists $\psi \in \mathscr{S}(\mathbb{R})$ so that $\phi(x) = x\psi(x)$.■

Proof Since $\phi \in \mathscr{S}$, we can compute $\widehat{\phi} \in \mathscr{S}$. By the Fourier inversion formula applied to $x = 0$ we get

$$0 = \phi(0) = \frac{1}{\sqrt{2\pi}} \int_{\mathbb{R}} \widehat{\phi}(\xi)\, \mathrm{d}\xi.$$

Ref. 2.9: "Fourier transform properties: differentiation"

Consider the function $\widehat{\psi}(x) \stackrel{\text{def}}{=} -i \int_{-\infty}^{x} \widehat{\phi}(\xi)\, \mathrm{d}\xi$. The above identity implies that $\widehat{\psi} \in \mathscr{S}$ also, and therefore is the Fourier transform of some Schwartz function ψ. By Proposition 2.9, this means

$$\mathscr{F}[x\psi](\xi) = i \frac{\mathrm{d}}{\mathrm{d}\xi} \widehat{\psi}(\xi) = \widehat{\phi}(\xi)$$

and so $x\psi = \phi$. □

Estimates in higher dimensions

Ref. 3.11: "Van der Corput: stationary case, $k \geq 3$"

In higher dimensions, we do not have the same η independent bounds that we saw in the Van der Corput Lemmas. The direct analogue of the higher-order Van der Corput Lemma turns out not to be as useful for analyzing dispersive equations; we state it here without proof for completeness. It proof is basically identical to the that of Lemma 3.11, but combined with localizations to small balls using a partition of unity argument.

3.22 Proposition
Suppose $\phi \in \mathscr{S}(\mathbb{R}^d)$ and $\eta : \mathbb{R}^d \to \mathbb{R}$ is smooth and that, for some multi-index α with $k = |\alpha| \geq 1$, we have $\left|\partial^\alpha \eta\right| \geq 1$ on $\operatorname{supp}\phi$. Then there exists some constant C dependent on the dimension d, the number k, and $\sup_{|\beta|\leq k+1} \sup_{x\in\operatorname{supp}\phi}\left|\partial^\beta \eta\right|$, such that

$$\left|I_{\eta,\phi}(\lambda)\right| \leq C\lambda^{-\frac{1}{k}}\left(\left\|\phi\right\|_{L^\infty} + \left\|\nabla\phi\right\|_{L^1}\right). \qquad \blacksquare$$

For our purposes, more interesting are the cases where the critical point is isolated. In the case where the isolated critical points are *non-degenerate*, meaning that the Hessian matrix at the critical point is invertible, one can recover better decay rates. We will not prove the statement for general stationary phase integrals, but will illustrate this approach using the Schrödinger and Klein-Gordon equations.

3.23 Theorem (Decay of solutions to Schrödinger)
Let $\eta(x) = x^2$. Then there exists a constant C depending only on the dimension d such that for every $\phi \in \mathscr{S}(\mathbb{R}^d)$,

$$\left|I_{\eta,\phi}(\lambda)\right| \leq C\langle\lambda\rangle^{-d/2}\left\|\phi\right\|_{W^{d,1}}. \qquad \blacksquare$$

Proof Let χ_0 be a bump function satisfying

- $\chi_0 \in C_0^\infty(\mathbb{R}^d)$ and takes value in $[0,1]$,
- χ_0 is rotationally symmetric: $\chi_0(x) = \chi_0(y)$ when $|x| = \left|y\right|$,
- $\chi_0(x) \equiv 1$ when $|x| \leq 1$ and 0 when $|x| \geq 2$.

Define $\chi_m(x) = \chi_0(2^{-m}x)$; χ_m is supported on the ball of radius 2^{m+1}. For a given m, we can split the integral

$$I_{\eta,\phi}(\lambda) = I_{\eta,\chi_m\phi}(\lambda) + I_{\eta,(1-\chi_m)\phi}(\lambda).$$

The first term is easy to estimate: using the compact support of χ_m we get

$$\left|I_{\eta,\chi_m\phi}(\lambda)\right| \leq 2^{d(m+2)}\left\|\phi\right\|_{L^\infty}.$$

For the second term, we observe that using $r = |x|$ we can write

$$\begin{aligned} I_{\eta,(1-\chi_m)\phi}(\lambda) &= \int_{\mathbb{R}^d} (1-\chi_m(r))\phi(x)e^{i\lambda r^2}\,\mathrm{d}x \\ &= \frac{1}{2i\lambda}\int_{\mathbb{R}^d}(1-\chi_m(r))\phi(x)r^{-1}\partial_r e^{i\lambda r^2}\,\mathrm{d}x; \end{aligned}$$

we used the fact that $(1-\chi_m(r)$ is supported exterior of the ball of radius 2^m. Let L denote the operator $f \mapsto \partial_r \frac{f}{r}$, repeating the integration by parts we get

$$I_{\eta,(1-\chi_m)\phi}(\lambda) = \frac{1}{(2i\lambda)^d} \int_{\mathbb{R}^d} L^{(d)}[(1-\chi_m(r))\phi(x)]e^{i\lambda r^2}\, \mathrm{d}x.$$

The expression $L^{(d)}$ can be written as a sum of terms of the form

$$\frac{1}{r^{d+a_1}} \cdot \partial_r^{a_2}(1-\chi_m)\cdot \partial_r^{a_3}\phi$$

where $a_1 + a_2 + a_3 = d$; this can be shown by induction with the fact that $\partial_r r^{-k} = -kr^{-k-1}$. The coefficients of the sum depends only on d. Therefore, there exists some constant which depends only on the dimension d such that

$$\left|I_{\eta,(1-\chi_m)\phi}(\lambda)\right| \leq \frac{C}{(2\lambda)^d} \sum_{a_1+a_2+a_3=d} \left\|\frac{1}{r^{d+a_1}} \cdot \partial_r^{a_2}(1-\chi_m)\cdot \partial_r^{a_3}\phi\right\|_{L^1}.$$

By Hölder's inequality we can bound

$$\left\|\frac{1}{r^{d+a_1}} \cdot \partial_r^{a_2}(1-\chi_m)\cdot \partial_r^{a_3}\phi\right\|_{L^1} \leq \left\|r^{-d-a_1}\right\|_{L^{d/a_1}} \left\|\partial_r^{a_2}(1-\chi_m)\right\|_{L^{d/a_2}} \left\|\partial_r^{a_3}\phi\right\|_{L^{d/a_3}}$$

where the integration is over $\operatorname{supp}(1-\chi_m)$, which we can possibly enlarge to the exterior of the ball of radius 2^m. Thus a direct computation shows that

$$\left\|r^{-d-a_1}\right\|_{L^{d/a_1}} \lesssim 2^{-md}.$$

Using that χ_m is obtained from χ_0 by scaling we also have

$$\left\|\partial_r^{a_2}(1-\chi_m)\right\|_{L^{d/a_2}} \lesssim 1.$$

And finally by the Gagliardo-Nirenberg-Sobolev inequality we have

$$\left\|\partial_r^{a_3}\phi\right\|_{L^{d/a_3}} \lesssim \left\|\phi\right\|_{\mathring{W}^{d,1}}.$$

The space $\mathring{W}^{k,p}$ is the homogeneous Sobolev space using exactly k derivatives.

Combining all the estimates we get

$$\left|I_{\eta,\phi}(\lambda)\right| \lesssim 2^{dm}\left\|\phi\right\|_{L^\infty} + \lambda^{-d}2^{-dm}\left\|\phi\right\|_{\mathring{W}^{d,1}}.$$

As the inequality holds for all m and λ, we can optimize by choosing $2^{dm} = \lambda^{-d/2}$ to get

$$\left|I_{\eta,\phi}(\lambda)\right| \lesssim \lambda^{-d/2}\left\|\phi\right\|_{\mathring{W}^{d,1}};$$

combining this with the trivial estimate $\left|I_{\eta,\phi}(\lambda)\right| \leq \left\|\phi\right\|_{L^1}$ we get the desired bound. □

3.24 Exercise (Uniform decay in Schrödinger)
In the case of solutions to Schrödinger's equation, the above argument in fact implies a uniform rate of decay. Prove this by following the following outline:

Recall that a solution to Schrödinger's equation has the respresentation formula

$$\phi(t,x) = \frac{1}{(2\pi)^{\frac{d}{2}}} \int_{\mathbb{R}^d} e^{it|\xi|^2 + ix\cdot\xi} \widehat{\phi_0}(\xi)\, \mathrm{d}\xi.$$

For $t > 0$, let $\eta(\xi) = |\xi|^2 + \frac{x}{t}\cdot\xi$. Completing the square and doing a change of variable, apply the estimate from Theorem 3.23 to conclude that

$$\left|\phi(t,x)\right| \leq C\langle t\rangle^{-d/2} \left\|\widehat{\phi_0}\right\|_{W^{d,1}},$$

where it is used that the Sobolev norm $W^{d,1}$ is translation invariant.

This argument uses the specific form of the Schrödinger dispersion relation, and does not easily generalize to other equations. ■

3.25 Theorem (Decay of solutions to Klein-Gordon)
Let $\eta(x) = \langle x\rangle$. Then there exists a constant C depending only on the dimension d such that for every $\phi \in \mathscr{S}(\mathbb{R}^d)$,

$$\left|I_{\eta,\phi}(\lambda)\right| \leq C\langle\lambda\rangle^{-d/2} \left\|\phi\right\|_{W^{d,1}}.$$ ■

Proof (Sketch) The proof is almost identical to the case of Schrödinger equation; the difference is that the operator L should be defined as

$$f \mapsto \partial_r \frac{\langle r\rangle}{r} f.$$

Using that for $k \geq 1$ we have

$$\left|\partial_r^k \frac{\langle r\rangle}{r}\right| \lesssim r^{-k-1}$$

and that $\langle r\rangle / r \leq \max(2, 2/r)$ we can estimate $L^{(d)}[(1-\chi_m)\phi]$ with essentially the same bounds. □

3.26 Remark
The rate $\lambda^{d/2}$ for Schrödinger is sharp. This can be seen by taking ϕ to be the Gaussian and computing the limit

$$\lim_{\lambda\to\infty} \lambda^{d/2} I_{\eta,\phi}(\lambda)$$

explicitly, in the same manner as Example 3.19. ■

3.27 Exercise (Point-wise decay for the wave equation)
Consider the integral

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

Prove that

$$|I(\lambda)| \leq C\langle\lambda\rangle^{-d}\|\phi\|_{W^{d,1}}.$$

Hint: Copy the proof of Theorem 3.23; why can we make m much more negative in this case than before? ■

As seen in the previous exercise, for any fixed x, the solutions to the linear wave equation $\phi(t,x)$ decays like t^{-d} or better. However, unlike the case of Schrödinger's equation, this decay is not spatially uniform. In fact, the best we can prove is that $\sup_{x\in\mathbb{R}^d}|\phi(t,x)| \lesssim t^{-(d-1)/2}$. We return to this in the next section.

Estimates of the fundamental solutions

Another way to approach the decay phenomenon for dispersive equations is to look at the corresponding fundamental solutions $G_t^{(*)}$ defined in Theorem 2.50. There are several advantages to the argument using oscillatory integrals of first kind to study the solutions: for one, it is possible, in principle, to compute the asymptotic series for $I_{\eta,\phi}(\lambda)$ in inverse powers of λ. This gives rather precise information about the asymptotic behavior of the solutions based on the Fourier representation of the initial data. There are, however, also disadvantages. First, the estimates proven are not obviously uniform in the spatial variables x: for the Schrödinger equation it is true by Exercise 3.24; however, as we already saw between Exercise 3.27 and Remark 3.15, the uniformity is not true for the wave equation. Secondly, and more importantly, the estimates proven in this form provide bounds of the solution by norms of the *Fourier transform* of the initial data. For various reasons (for example, iteration arguments for solving nonlinear problems) one would hope to provide bounds by norms of the *physical space representation* of the initial data. In this section we tackle some of these estimates for our archetype equations.

The Schrödinger case depended on the observation that "a boosted parabola is just another parabola".

Let us start with the Schrödinger equation; we start here because it is the simplest case and can be used to illustrate many of the main ideas. We will give two separate proofs of the dispersive decay for the Schrödinger equation: the first, using complex analytic ideas, proceeds by establishing an explicit formula for the kernel $G_t^{(\mathrm{Sch})}$. The requisite estimates can then be simply read off from the formula. This method however is not easily generalized, as the fundamental solutions for many equations (such as the Klein-Gordon equation) do not have known explicit formulas. The second, more real-variables based, method is less explicit, and the constants derived in the estimates are consequently worse. But it has the advantage of easily carrying over to different dispersive equations. This second method will form the basis of our analyses of the other basic equations.

What we are interested in is control of the integral operator

$$\phi_0(x) \mapsto \phi(t,x) \stackrel{\text{def}}{=} \frac{1}{(2\pi)^d} \iint\limits_{\mathbb{R}^d\times\mathbb{R}^d} e^{it|\xi|^2 + i(x-y)\cdot\xi} \phi_0(y) \,\mathrm{d}y\,\mathrm{d}\xi$$

acting on $\phi_0 \in \mathscr{S}(\mathbb{R}^d)$. We formally identify

$$G_t^{(\mathrm{Sch})} = \frac{1}{(2\pi)^d} \int\limits_{\mathbb{R}^d} e^{it|\xi|^2 + ix\cdot\xi}\,\mathrm{d}\xi$$

in the sense of distributions. Now let $\chi_n \in \mathscr{S}$ be any sequence of functions that converges (uniformly on compact sets) to the constant function 1. Then we can check that, for an fixed t, the sequence of tempered distributions

$$e^{it|\xi|^2}\chi_n(\xi)$$

converges to $e^{it|\xi|^2}$. Indeed, given $\phi \in \mathscr{S}$, we can let

$$A = \sup_{\xi\in\mathbb{R}^d} \langle\xi\rangle^{d+1} \left|\phi(\xi)\right|.$$

Then given $\epsilon > 0$ we can choose $R > 0$ such that

$$\int\limits_{|\xi|\geq R} \left|\phi(\xi)\right|\,\mathrm{d}\xi \leq \int\limits_{|\xi|\geq R} \frac{A}{\langle\xi\rangle^{d+1}}\,\mathrm{d}\xi \leq \frac{\epsilon}{2}A.$$

Next, observe that

$$\int\limits_{|\xi|\leq R} |1-\chi_n(\xi)|\left|\phi(\xi)\right|\,\mathrm{d}\xi \leq \sup_{|\xi|\leq r}|1-\chi_n| \cdot \int\limits_{\mathbb{R}^d} \frac{A}{\langle\xi\rangle^{d+1}}\,\mathrm{d}\xi.$$

By the uniform convergence on compact sets, we have that for all sufficiently large n,

$$\int_{|\xi|\leq R} |1-\chi_n(\xi)|\big|\phi(\xi)\big| \,\mathrm{d}\xi \leq \frac{\epsilon}{2}A,$$

and this shows that the difference $e^{it|\xi|^2}(1-\chi_n(\xi))$ tends to zero in the sense of distributions. This approximation procedure is the first step in both of the methods that we will present.

The first method can be summarized in the following theorem.

3.28 Theorem (Explicit formula for $G_t^{(\mathrm{Sch})}$)
For $t \neq 0$, the tempered distribution

$$G_t^{(\mathrm{Sch})} = \frac{e^{i\,\mathrm{sgn}(t)d\pi/4}}{(4\pi|t|)^{d/2}} e^{-\frac{i}{4t}|x|^2}.$$

■

Proof We obtain the claimed expression by studying the limit

$$\lim_{\epsilon\to 0} \frac{1}{(2\pi)^d} \int_{\mathbb{R}^d} e^{it|\xi|^2 + ix\cdot\xi} e^{-\frac{\epsilon}{2}|\xi|^2} \,\mathrm{d}\xi.$$

Noticing that as $\epsilon \to 0$, $e^{-\epsilon|\xi|^2} \to 1$ uniformly on compact sets, by our argument before the limiting distribution given by the expression above is $G_t^{(\mathrm{Sch})}$, using that the Fourier transform sends $\mathscr{S}' \to \mathscr{S}'$. The computation of this limit proceeds largely along the same lines as the argument in Example 3.19. Completing the square

Ref. 3.19: "Leading order asymptotic for 1d Schrödinger"

$$-(\frac{\epsilon}{2} - it)|\xi|^2 + ix\cdot\xi = -\frac{1}{2}(a\xi - \frac{i}{a}x)\cdot(a\xi - \frac{i}{a}x) - \frac{1}{2a^2}|x|^2$$

where $a \in \mathbb{C}$ is given by $re^{i\theta}$ with $\theta \in (-\pi/4, \pi/4)$ and $a^2 = \epsilon - 2it$ (which implies $r = (\epsilon^2 + 4t^2)^{\frac{1}{4}}$). So by contour integration again we get

$$\frac{1}{(2\pi)^d} \int_{\mathbb{R}^d} e^{it|\xi|^2 + ix\cdot\xi} e^{-\frac{\epsilon}{2}|\xi|^2} \,\mathrm{d}\xi = \frac{e^{-id\theta}}{(2\pi)^{d/2}\cdot r^d} e^{-\frac{1}{2(\epsilon-2it)}|x|^2}.$$

Now, as we take the limit $\epsilon \to 0$, we see that $r \to \sqrt{2|t|}$, and $\theta \to -\mathrm{sgn}(t)\cdot\frac{\pi}{4}$. This gives

$$G_t^{(\mathrm{Sch})}(x) = \frac{e^{i\,\mathrm{sgn}(t)d\pi/4}}{(4\pi|t|)^{d/2}} e^{-\frac{i}{4t}|x|^2}$$

as claimed. □

3.29 Corollary (Uniform decay for Schrödinger)
If $\phi(t,x)$ solves Schrödinger's equation with initial data $\phi(0,x) = \phi_0(x)$, then

$$\left|\phi(t,x)\right| \leq \frac{1}{(4\pi|t|)^{d/2}} \left\|\phi_0\right\|_{L^1}. \qquad \blacksquare$$

Proof Using that $\phi(t,x) = G_t^{(\mathrm{Sch})} * \phi_0(x)$ by definition, we see that in the convolution integral

$$\left|G_t^{(\mathrm{Sch})} * \phi_0(x)\right| = \left|\int_{\mathbb{R}^d} G_t^{(\mathrm{Sch})}(x-y)\phi_0(y)\,\mathrm{d}y\right| \leq \left\|G_t^{(\mathrm{Sch})}\right\|_{L^\infty} \left\|\phi_0\right\|_{L^1}$$

by Hölder's inequality. □

This first method exhibits several properties that is typical of complex-analytic arguments. First, the formula obtained is extremely explicit. Second, as a result, the *sharp* constant is found for the uniform decay estimate for Schrödinger equation. The main drawback to this method is that it is not easily generalizable, as it relies on the specific form of the Fourier transform of the Schrödinger kernel as an "imaginary Gaussian". For example, the same argument cannot be directly applied to obtain a closed form representation of $G_t^{(\mathrm{Airy})}$. Before giving the more-generally-applicable second method, let us digress a little and talk about "scaling properties".

Observe that the explicit formula for $G_t^{(\mathrm{Sch})}$ has a scaling homogeneity:

$$G_t^{(\mathrm{Sch})}(x) = \frac{1}{t^{d/2}} G_1^{(\mathrm{Sch})}\left(\frac{x}{\sqrt{t}}\right)$$

for $t > 0$. This is in accordance with the natural scaling of the equation. If $\phi(t,x)$ solve Schrödinger's equation, then so does $\phi_\lambda(t,x) = \phi(\lambda^2 t, \lambda x)$. Scaling properties can also be computed for the Airy and wave equations.

For the Airy equation, we see that if $\phi(t,x)$ is a solution, then so is $\phi(\lambda^3 t, \lambda x)$. Therefore, if one were able to prove that $G_1^{(\mathrm{Airy})}$ is a tempered distribution represented by a uniformly bounded function (which we will do in the sequel), then the scaling will imply immediately that solutions to the Airy equation have uniform $t^{-1/3}$ decay. For the wave equation, we see that if $\phi(t,x)$ is a solution, then so is $\phi(\lambda t, \lambda x)$. Again, if $G_1^{(\mathrm{wave})}$ were bounded, the scaling would imply that solutions to the wave equation would decay like t^{-d}. However, as we have already seen previously, the decay rate for the wave equation should be no better than $t^{-(d-1)/2}$. And indeed, the fundamental solution to the wave equation *cannot* be represented as a bounded function.

The Klein-Gordon equation, on the other hand, does not have good scaling properties, due to the presence of the mass term. This complication will be reflected in the relative difficulty when we try to control its fundamental solution.

Now let us consider the second method of obtaining dispersive estimates for $G_t^{(\mathrm{Sch})}$. As discussed above, by scaling homogeneity it suffices to prove that $G_1^{(\mathrm{Sch})}$ can be represented by a bounded function. We will show directly that

3.30 Theorem
There exists a constant C depending on the dimension d such that solutions of of Schrödinger's equation with initial data ϕ_0 in $\mathscr{S}$ satisfies the uniform estimate

$$\left|\phi(1,x)\right| \leq C\left\|\phi_0\right\|_{L^1}. \qquad \blacksquare$$

Proof By our approximation procedure above, it suffices to show that

$$\iint\limits_{\mathbb{R}^d\times\mathbb{R}^d} e^{i|\xi|^2+i(x-y)\cdot\xi}\phi_0(y)\chi_n(\xi)\,\mathrm{d}\xi\,\mathrm{d}y$$

is bounded with the bound being uniform in x and n, where $\chi_n \to 1$ uniformly on compact sets. We build our χ_n as follows. Fix γ a smooth monotonic function on the real line such that $\gamma(x) \equiv 1$ when $x \leq 1$ and $\gamma(x) \equiv 0$ when $x \geq 2$. Let $\sigma_0(x,y,\xi) = \gamma\left(\left|2\xi-(y-x)\right|\right)$, and for $k \geq 1$ let

$$\sigma_k(x,y,\xi) = \gamma\left(2^{-k}\left|2\xi-(y-x)\right|\right) - \gamma\left(2^{1-k}\left|2\xi-(y-x)\right|\right).$$

Let us write $\eta(x,y,\xi) = |\xi|^2 + (x-y)\cdot\xi$. Observe that $\sum_{k=0}^{\infty}\sigma_k(x,y,\xi) = 1$. We estimate separately the pieces

$$\int\limits_{\mathbb{R}^d} e^{i\eta(x,y,\xi)}\sigma_k(x,y,\xi)\,\mathrm{d}\xi.$$

Using that for x,y fixed, the support in ξ of σ_0 is a ball of radius 1, we have that

$$\left|\int\limits_{\mathbb{R}^d} e^{i\eta(x,y,\xi)}\sigma_0(x,y,\xi)\,\mathrm{d}\xi\right| \lesssim 1.$$

Next, observe that since

$$\nabla^{(\xi)}\eta = 2\xi - (y - x)$$

on the support of σ_k, we have that

$$\left|\nabla^{(\xi)}\eta\right| \geq 2^{k-1}, \quad \left|\nabla^{(\xi)}\nabla^{(\xi)}\eta\right| \leq 2, \quad \nabla^{(\xi)}\nabla^{(\xi)}\nabla^{(\xi)}\eta = 0.$$

This implies that

$$\left(\nabla^{(\xi)}\right)^\ell \frac{\nabla^{(\xi)}\eta}{\left|\nabla^{(\xi)}\eta\right|^2} \lesssim_\ell 2^{-\ell k}. \tag{3.31}$$

Now, let L be the linear operator $Lf = -\nabla^{(\xi)} \cdot \left(f\nabla^{(\xi)}\eta / \left|\nabla^{(\xi)}\eta\right|^2\right)$, we have that

$$\int_{\mathbb{R}^d} e^{i\eta}\sigma_k \,\mathrm{d}\xi = \int_{\mathbb{R}^d} L^{d+1}(\sigma_k)e^{i\eta} \,\mathrm{d}\xi.$$

By the estimate (3.31), we can bound pointwise

$$\left|L^{d+1}(\sigma_k)\right| \lesssim \sum_{\ell=0}^{d+1} \sum_{|\alpha|=d+1-\ell} 2^{-k\ell} \left|\left(\nabla^{(\xi)}\right)^\alpha \sigma_k\right|.$$

From the scaling property of σ_k, we have that pointwise

$$\left|(\nabla^{(\xi)})^\alpha \sigma_k\right| \lesssim 2^{-k|\alpha|},$$

so we conclude that

$$\left|L^{d+1}(\sigma_k)\right| \lesssim 2^{-k(d+1)},$$

with the constant independent of x, y and k. Integrating we obtain

$$\left|\int_{\mathbb{R}^d} e^{i\eta}\sigma_k\right| \mathrm{d}\xi \leq \int_{\mathbb{R}^d} \left|L^{d+1}(\sigma_k)\right| \mathrm{d}\xi \lesssim 2^{-k}.$$

This sequence being absolutely summable we get the desired result. □

3.32 Remark
Very roughly speaking, what we proved in the previous two theorems is that the integral

$$\int_{\mathbb{R}^d} e^{i|\xi|^2 + ix\cdot\xi} \,\mathrm{d}\xi$$

is "uniformly bounded". Of course, the integrand being something of norm 1 at every ξ means that the integral is far from converging, and that the boundedness only really makes sense in a distributional sense. The key factor driving the boundedness here is the fact that the phase function $\eta(\xi) = |\xi|^2 + x\cdot\xi$ not only grows as $\xi \to \infty$, but also has the property that $\left|\nabla\eta\right| \nearrow \infty$ as $|\xi| \nearrow \infty$. One should think of this as one of the key ideas behind oscillatory integrals: high oscillation implies small integrals.

Consider the Fourier transform applied to the function

$$f(x) = \begin{cases} 1 & |x| \leq 1 \\ 0 & |x| > 1 \end{cases}$$

defined on the real line. We can understand the decay of the Fourier coefficients as follows: looking at

$$\int_{\mathbb{R}} e^{-ix\xi} f(x)\,\mathrm{d}x = \int_{-1}^{1} e^{-ix\xi}\,\mathrm{d}x$$

we see that the integrand is an oscillating function on the interval $[-1,1]$. Noticing that integrating over a period, which for $e^{-ix\xi}$ would be for x over an interval $[x_0, x_0 + \frac{2\pi}{\xi}]$, the integral evaluates to zero. And so the "non-zero" contributions to $\int_{-1}^{1} e^{ix\xi}\,\mathrm{d}x$ only comes from the boundary region near the points $\{\pm 1\}$ of thickness no more than π/ξ, where the integral cannot be paired up into a complete period. And from this argument already we see the decay estimate

$$\left|\widehat{f}(\xi)\right| \lesssim \langle\xi\rangle^{-1}.$$

Returning to the case $\int_{\mathbb{R}^d} e^{i|\xi|^2 + ix\cdot\xi}\,\mathrm{d}\xi$, the growth of the phase function as ξ tends to infinity means that for larger and larger ξ, the integral is better and better at oscillatory cancellations, so even though $e^{i|\xi|^2}$ has norm 1, when integrated against a function that has some regularity (say, Schwartz class), the contribution from ξ far away from the origin is relatively small. This is nothing more than just applying Lemma 3.6 after localizing in ξ to regions where we can control the size of $\nabla\eta$. ■

Ref. 3.6: "Arbitrarily fast decay from method of non-stationary phase"

Let us next consider the Airy equation. Using that the function $\eta(\xi) = -\xi^3 + x\xi$ has uniform lower bound $\left|\eta'''(\xi)\right| = 6$ we have that, if χ_n is a cut-off function supported on $[-n-1, n+1]$, with $\chi_n \equiv 1$ on $[-n, n]$, and such that

χ_n is monotonic on $[-n-1,-n]$ and $[n,n+1]$, we have the uniform bound by Lemma 3.11 that

$$\left|\int_{\mathbb{R}} \chi_n e^{i\eta}\, \mathrm{d}\xi\right| \leq \frac{20}{\sqrt[3]{6}}$$

and thus we have that $G_1^{(\mathrm{Airy})}$ can be represented by a bounded function A. Using similar ideas to that used to prove Theorem 3.30, we can get a bit more about the behavior of the function A.

3.33 Theorem
The bounded function $A : \mathbb{R} \to \mathbb{C}$ representing $G_1^{(\mathrm{Airy})}$ satisfies:

- $A(x)$ decays faster than any polynomial of x as $x \to -\infty$.
- $A(x) \lesssim \langle x\rangle^{-1/4}$. ■

Proof Let χ_n be defined as in the paragraph before the statement of the theorem. Let's treat the first case, where $x \leq -1$. There we have that $\eta'(\xi) = -3\xi^2 + x \leq x$. This implies that the derivatives $\left|\nabla^{(k)} \frac{1}{\eta'(\xi)}\right| \lesssim_k \frac{1}{|x|+\xi^2}$ and so are integrable. Writing $L(f) = -\nabla(f/\eta')$ we can integrate by parts N times to get

$$\int_{\mathbb{R}} e^{i\eta} \chi_n \,\mathrm{d}\xi = \int_{\mathbb{R}} L^N(\chi_n) e^{i\eta}\, \mathrm{d}\xi.$$

Now,

$$\left|L^N(\chi_n)\right| \lesssim_N \frac{\|\chi_n\|_{W^{N,\infty}}}{(|x| + \xi^2)^N}$$

by the above computations, which implies

$$\left\|L^N(\chi_n)\right\|_{L^1} \lesssim_N \frac{\|\chi_n\|_{W^{N,\infty}}}{|x|^{N/2}}.$$

Noticing that $\|\chi_n\|_{W^{N,\infty}}$ can be taken to be independent of n, after taking $n \to \infty$ we prove the first assertion.

For the second assertion, it suffices to look at the case $x \geq 1$. Let $\xi_0 = \sqrt{x/6}$. Split the integral

$$\int_{\mathbb{R}} e^{i\eta}\chi_n \,\mathrm{d}\xi = \int_{-\infty}^{-\xi_0} + \int_{-\xi_0}^{\xi_0} + \int_{\xi_0}^{\infty} e^{i\eta}\chi_n \,\mathrm{d}\xi.$$

For the first and third pieces, observe that when $|\xi| \geq \xi_0$ we have

$$\left|\eta''(\xi)\right| = |6\xi| \geq \sqrt{6|x|}.$$

Ref. 3.10: "Van der Corput: stationary case, k = 2"

So by Lemma 3.10 we have that

$$\left|\int_{\xi_0}^{\infty} e^{i\eta} \chi_n \,\mathrm{d}\xi\right| \leq \left(\frac{1}{6|x|}\right)^{\frac{1}{4}} \cdot 9,$$

where we used that $\int_{\xi_0}^{\infty} |\chi_n'| \,\mathrm{d}\xi = 1$ by construction.

For the middle piece, observe that when $|\xi| \leq \xi_0$ we have

$$\left|\eta'(\xi)\right| = \left|-3\xi^2 + x\right| \geq \frac{1}{2}|x|.$$

Ref. 3.8: "Van der Corput: non-stationary case"

So by Lemma 3.8 we have, for all $n > \xi_0$, that

$$\left|\int_{-\xi_0}^{\xi_0} e^{i\eta} \chi_n \,\mathrm{d}\xi\right| \leq \frac{6}{|x|}.$$

Putting these together we get the desired estimate. □

3.34 Remark

As we saw in Theorem 3.28, the norm $\left|G_1^{(\mathrm{Sch})}\right| = \frac{1}{(4\pi)^{d/2}}$ is independent of spatial position. On the other hand, we have just proven that the Airy kernel $G_1^{(\mathrm{Airy})}$ has spatial decay. The spatial decay in this situation arises from the fact that when $x > 1$, due to the cubic growth of the phase function, the critical point of the phase function $\eta(\xi)$ occurs where $\eta'' \approx \sqrt{x}$. Compare this with the Schrödinger case where the critical point occurs where $\eta'' = 2$.

The rapid decay of $G_1^{(\mathrm{Airy})}$ in the positive x axis should be viewed against the classical analogue of the Airy equation. Recalling that the dispersion relation in this case is $\omega(k) = k^3$, the classical "kinetic theory" analogue of the Airy equation is the equation

$$\partial_t \rho + v^2 \partial_x \rho = 0$$

for the distribution function ρ over the classical phase space $\mathbb{R} \times \mathbb{R}^2$. The solution to this classical equation has explicit form

$$\rho(t, x, v) = \rho(0, x - tv^2, v),$$

which indicates that all "particles" move "to the right". The classical analogue of $G_1^{(\text{Airy})}$ would in fact vanish for the entirety of the left half line. The rapid decay of the Airy kernel on the left half line is the typical exponential decay one expects from quantum tunneling into the classically forbidden region. ■

Next, let us look at the wave equation. As mentioned earlier, we do not expect $G_1^{(\text{wave})}$ to be represented by a bounded function, which has as a consequence that convolving against the fundamental solution will not represent a bounded mapping from L^1 to L^∞.

3.35 Proposition (Estimate for truncated wave kernel)
Let $d \geq 2$. Suppose χ_0 is a smooth function with support contained within the annulus with inner radius r_1 and outer radius r_2, then

$$\left| \int_{\mathbb{R}^d} e^{it|\xi| + ix\cdot\xi} \chi_0(\xi) \,\mathrm{d}\xi \right| \lesssim |t|^{-(d-1)/2},$$

where the constant depends on up to $d-1$ derivatives of χ_0, the dimension d, and the numbers r_1, r_2. ■

Proof Let $\omega_0 \in \mathbb{S}^{d-1}$ be the direction of the vector x; when $x = 0$ we choose ω_0 arbitrarily. Let P be the orthogonal hyperplane to ω_0. We write the integral

$$\int_{\mathbb{R}^d} e^{it|\xi| + ix\cdot\xi} \chi_0(\xi) \,\mathrm{d}\xi = \int_{\mathbb{R}} \int_P e^{it|s\omega_0 + \zeta| + is(x\cdot\omega_0)} \chi_0(s\omega_0 + \zeta) \,\mathrm{d}\zeta \,\mathrm{d}s,$$

where $\xi \in \mathbb{R}^d$ is decomposed into a sum of $\zeta \in P$ and $s\omega_0$. Now let ψ_1 be a smooth monotonic function on $\mathbb{R}$ such that $\psi_1(x) \equiv 1$ when $x \leq 1$ and $\psi_1(x) \equiv 0$ when $x \geq 2$. Denote by $\psi_\tau(x) = \psi_1(\sqrt{\tau}x)$. We split the integral into

$$\iint_{\mathbb{R}\times P} e^{it|s\omega_0 + \zeta| + is(x\cdot\omega_0)} \chi_0(s\omega_0 + \zeta) \psi_t(|\zeta|) \,\mathrm{d}\zeta \,\mathrm{d}s$$
$$+ \iint_{\mathbb{R}\times P} e^{it|s\omega_0 + \zeta| + is(x\cdot\omega_0)} \chi_0(s\omega_0 + \zeta) [1 - \psi_t(|\zeta|)] \,\mathrm{d}\zeta \,\mathrm{d}s.$$

Using that χ_0 is supported on the annulus, and that ψ_t is supported in a region of radius $t^{-1/2}$ around the ω_0 axis, we have that the first integral is bounded by $\lesssim t^{-(d-1)/2}$.

For the second integral, let r denote the radial variable on P; then on the support of $1-\psi_t$ we have that the radial derivative ∂_r on P is a smooth vector field. Using that $|s\omega_0 + \zeta| = \sqrt{s^2+r^2}$, we have that the derivative of the phase satisfies

$$\partial_r(t|s\omega_0 + \zeta| + s(x\cdot\omega_0)) = \frac{rt}{\sqrt{s^2+r^2}}.$$

Hence writing the operator $Lf = -\partial_r(f\sqrt{s^2+r^2}/r)$ we have

$$\left| \iint_{\mathbb{R}\times P} e^{it|s\omega_0+\zeta|+is(x\cdot\omega_0)} \chi_0(s\omega_0+\zeta)[1-\psi_t(|\zeta|)] \,\mathrm{d}\zeta\,\mathrm{d}s \right| \leq \frac{1}{t^{d-1}} \iint_{\mathbb{R}\times P} \left| L^{d-1}[\chi_0(1-\psi_t)] \right| \,\mathrm{d}\zeta\,\mathrm{d}s.$$

Now using that $\sqrt{s^2+r^2}$, on $\operatorname{supp}\chi_0$, is a bounded function with bounded derivatives, and using that $\operatorname{supp}\nabla\psi_t$ is contained in a set of volume $\lesssim t^{-(d-1)/2}$, we conclude that

$$\iint_{\mathbb{R}\times P} \left| L^{d-1}[\chi_0(1-\psi_t)] \right| \,\mathrm{d}\zeta\,\mathrm{d}s \lesssim \sup_{\operatorname{supp}(1-\psi_t)} \frac{1}{r^{d-1}},$$

where the constant depends on the norms of up to $d-1$ derivatives of ψ_1 and χ_0, but not at t or x. Using that $r \geq \frac{1}{\sqrt{t}}$ on the support of $1-\psi_t$, we finally arrive at that the second integral is also bounded by $\lesssim t^{-(d-1)/2}$. □

3.36 Remark

In our proof of Proposition 3.35 we used the radial vectorfield orthogonal to x to capture the non-stationary phase behaviour. Observe that this vector field is almost tangent to the unit sphere near the axis given by the direction of x, while the vector field is almost orthogonal to the unit sphere "on the equator". It turns out that only the direction near the axis is crucial. When $t = |x|$, the phase function is constant along a ray in the direction $-x$, and to apply the method of (non)stationary phase we have to use vector fields that are orthogonal to the ray, in order to pick up the correct amount of decay.

(The basic idea behind stationary phase is that the observation that on the real line, if $f'' > 1$ and $f'(0) = 0$, then we have that $|f'(x)| \geq |x|$. This we already saw in our discussion of the Van der Corput Lemma 3.10. For higher dimensional cases, the analogue is that for convex functions, the radial derivatives grow at least linearly from the critical point. Note however that non-radial derivatives can remain small!)

For points that are far from the problematic ray, one easily checks that all derivatives of the phase function have the appropriate lower bounds. This implies that instead of the radial vector field orthogonal to x, another good vector field to consider is one that is tangent to the spheres, and runs between the two poles. If one were to use this vector field, one can upgrade Proposition 3.35 to a statement about integrals of measures supported on spheres. ■

3.37 Exercise (Decay of Fourier transform of spherical measure)
Let $\omega_0 \in \mathbb{S}^d$ be fixed. Prove that

$$\left| \int_{\mathbb{S}^d} e^{ir\omega_0\cdot\omega} \,\mathrm{d}\omega \right| \lesssim r^{-d/2}.$$

(*Hint:* decompose $\mathbb{S}^d$ into three portions, S_+ near the north pole ω_0, S_- near the south pole $-\omega_0$, and S_0 the rest. On $S_\pm$ use volume estimates. On S_0 integrate by parts against the unit-length vector field in the direction of the longitude lines. To implement this it may help to rewrite things using the polar coordinates where $\mathbb{S}^d \cong [0,\pi] \times \mathbb{S}^{d-1} \ni (\theta, \omega')$; from this we get that the desired vector field is simply ∂_θ.) ■

3.38 Exercise (Higher decay of Fourier transform of spherical measure)
Let $\omega_0 \in \mathbb{S}^d$ be fixed. Let $\phi_\pm : \mathbb{S}^d \to \mathbb{C}$ be smooth functions supported on $\{\omega_0 \cdot \omega > -\frac{1}{2}\}$ and $\{\omega_0 \cdot \omega < \frac{1}{2}\}$ respetively. Prove that

$$\left| \partial_r^k \left[e^{\mp ir} \int_{\mathbb{S}^d} e^{ir\omega_0\cdot\omega} \phi_\pm(\omega) \,\mathrm{d}\omega \right] \right| \lesssim r^{-d/2-k}.$$

(*Hint:* $\left|\partial_r e^{-ir+ir\omega_0\cdot\omega}\right| = |\omega_0 \cdot \omega - 1| = 1 - \cos\theta \leq \frac{1}{2}\theta^2$ where θ is defined as in the previous exercise.) ■

3.39 Corollary (Frequency-restricted decay for wave)
Let $0 < r_1 < r_2$, and $d \geq 2$. Then there exists a constant C such that for every $\phi_0 \in \mathscr{S}(\mathbb{R}^d)$ such that $\operatorname{supp}\widehat{\phi_0}$ is contained in the annulus of inner radius

r_1 and other radius r_2, we have the estimates

$$\left|G_t^{(\text{wave})} * \phi_0(x)\right| \le C|t|^{-(d-1)/2}\|\phi_0\|_{L^1};$$
$$\left|\partial_t G_t^{(\text{wave})} * \phi_0(x)\right| \le C|t|^{-(d-1)/2}\|\phi_0\|_{L^1}. \qquad \blacksquare$$

Proof We demonstrate the case for $G_t^{(\text{wave})}$; the case for $\partial_t G_t^{(\text{wave})}$ is analogous. We have by definition

$$G_t^{(\text{wave})} * \phi_0(x) = \frac{1}{(2\pi)^d}\int_{\mathbb{R}^d} \frac{1}{2i|\xi|}\left[\exp(it|\xi| + ix\cdot\xi) - \exp(-it|\xi| + ix\cdot\xi)\right]\widehat{\phi_0}(\xi)\,\mathrm{d}\xi.$$

The integral is well-defined by the support condition on $\widehat{\phi_0}$. Now let $\tilde{\chi}$ be a smooth function supported on the annulus of inner radius $\frac{1}{2}r_1$ and outer radius $2r_2$, and equals 1 identically on the annulus of inner radius r_1 and outer radius r_2. Then we have that $\tilde{\chi}\widehat{\phi_0} = \widehat{\phi_0}$. So using the Fourier inversion formula we can write

$$\int_{\mathbb{R}^d} \frac{e^{it|\xi|+ix\cdot\xi}}{|\xi|}\widehat{\phi_0}(\xi)\,\mathrm{d}\xi = \frac{1}{(2\pi)^{d/2}}\iint_{\mathbb{R}^d\times\mathbb{R}^d} e^{it|\xi|+i(x-y)\cdot\xi}\chi_0(\xi)\phi_0(y)\,\mathrm{d}y\,\mathrm{d}\xi,$$

where $\chi_0(\xi) = \tilde{\chi}/|\xi|$ is smooth and supported in some annulus. Switching the order of integration (since everything converges), we see that the integral in ξ can be estimated uniformly using Proposition 3.35, with the constant depending only on χ_0 (which we chose to be universal for the fixed r_1 and r_2), d, and r_1 and r_2. The Corollary follows. $\square$

Ref. 3.35: "Estimate for wave kernel restricted to an annulus"

We have shown that solutions to the linear wave equation, with initial data having Fourier support bounded away from both 0 and ∞, decay like $|t|^{-(d-1)/2}$ uniformly when the initial data is measured in L^1. However, the constant can yet depend on the radii of the Fourier support. The following captures the radii dependence.

3.40 Corollary

Let $d \ge 2$, then there exists a universal constant C such that for every $\phi_0 \in \mathscr{S}(\mathbb{R}^d)$ such that $\operatorname{supp}\widehat{\phi_0}$ is contained in the annulus of inner radius

2^{k-1} and outer radius 2^{k+1}, we have the estimates

$$\left|G_t^{(\text{wave})} * \phi_0(x)\right| \leq C 2^{k(d-1)/2} |t|^{-(d-1)/2} \left\|\phi_0\right\|_{L^1};$$
$$\left|\partial_t G_t^{(\text{wave})} * \phi_0(x)\right| \leq C 2^{k(d+1)/2} |t|^{-(d-1)/2} \left\|\phi_0\right\|_{L^1}. \qquad \blacksquare$$

Proof The proof is based on the scaling homogeneity of the wave equation. Let $\phi_k(x) = 2^{-kd}\phi_0(2^{-k}x)$, then we have $\left\|\phi_k\right\|_{L^1} = \left\|\phi_0\right\|_{L^1}$. By Proposition 2.7 we have $\widehat{\phi_k}(\xi) = \widehat{\phi_0}(2^k\xi)$ now has support on the annulus of inner radius $\frac{1}{2}$ and outer radius 2. So for some universal constant we have that

Ref. 2.7: "Fourier transform properties: scaling, translation, modulation"

$$\left|G_t^{(\text{wave})} * \phi_k(x)\right| \leq C|t|^{-(d-1)/2} \left\|\phi_0\right\|_{L^1}.$$

On the other hand, we have that

$$\int_{\mathbb{R}^d} \frac{\sin(t|\xi|)}{|\xi|} e^{ix\cdot\xi} \widehat{\phi_k}(\xi) \,\mathrm{d}\xi = \int_{\mathbb{R}^d} \frac{\sin(2^{-k}t\left|2^k\xi\right|)}{2^{-k}\left|2^k\xi\right|} e^{i2^{-k}x\cdot 2^k\xi} \widehat{\phi_0}(2^k\xi) \frac{2^{kd}}{2^{kd}} \,\mathrm{d}\xi.$$

So we have that

$$\frac{1}{2^{k(d-1)}} G_{2^{-k}t}^{(\text{wave})} * \phi_0(2^{-k}x) = G_t^{(\text{wave})} * \phi_k(x)$$

or, as claimed,

$$\left|G_t^{(\text{wave})} * \phi_0(x)\right| \leq C|t|^{-(d-1)/2} 2^{-k(d-1)/2} 2^{k(d-1)} \left\|\phi_0\right\|_{L^1}.$$

The case for $\partial_t G_t^{(\text{wave})}$ is analogous, except that due to its Fourier representation not having the $\frac{1}{|\xi|}$ term, the scaling reveals one extra factor of 2^k in the estimates. □

The above estimates only apply to functions with Fourier support contained on an annulus. To get an estimate that applies to all Schwartz functions, we have to introduce the notion of *Littlewood-Paley projectors* and the notion of *Besov spaces*.

3.41 Definition (Standard Littlewood-Paley projectors)

Let $\psi : \mathbb{R}^d \to [0,1]$ be a smooth function with compact support, such that $\operatorname{supp}\psi$ is contained in the ball of radius 2 around the origin, and that $\psi \equiv 1$ on the ball of radius 1 around the origin. The *Littlewood-Paley*

Projectors based on ψ is the family of operators $\{\Delta_k\}$ indexed by $k \in \mathbb{Z}$, where $\Delta_k : L^2(\mathbb{R}^d) \to L^2(\mathbb{R}^d)$ is given by

$$\Delta_k(f) \stackrel{\text{def}}{=} \mathscr{F}^{-1}[(\psi(2^{-k}\bullet) - \psi(2^{1-k}\bullet))\widehat{f}(\bullet)].$$

We will also have occasion to make use of the partially summed projectors

$$\Delta_{\leq k}(f) \stackrel{\text{def}}{=} \sum_{\ell=-\infty}^{k} \Delta_\ell(f) = \mathscr{F}^{-1}[\psi(2^{-k}\bullet)\widehat{f}(\bullet)].$$

■

3.42 (Basic properties of Littlewood-Paley projectors) Observe that formally the sum $\sum_k \Delta_k$ is telescoping, and therefore we have that for any $f \in L^2$ (and hence also for any $f \in \mathscr{S}$),

$$\sum_{k\in\mathbb{Z}} \Delta_k f = f.$$

The individual functions $\Delta_k f$ has, by construction, Fourier support on the annulus with inner radius 2^{k-1} and outer radius 2^{k+1}, ane hence by the Paley-Wiener theorem (see Exercise 2.11) are real analytic. The projectors are also *almost orthogonal*:

$$\Delta_j \Delta_k f = 0 \text{ whenever } |j-k| > 1,$$

and they are almost idempotent

$$\sum_{j=k-1}^{k+1} \Delta_j \Delta_k = \Delta_k.$$

By Plancherel (Proposition 2.23), and the almost orthogonality, we have

$$\sum_{k\in\mathbb{Z}} \|\Delta_k f\|_{L^2}^2 \leq \|f\|_{L^2}^2 \leq 3 \sum_{k\in\mathbb{Z}} \|\Delta_k f\|_{L^2}^2.$$

And we have the pointwise comparison for derivatives: Notice that if we write $\chi_k(x) = \frac{1}{(2\pi)^{d/2}} \mathscr{F}^{-1}[\psi(2^{-k}\xi) - \psi(2^{1-k}\xi)]$, we have that $\Delta_k f = \chi_k * f$. This implies that $\partial \Delta_k f = (\partial \chi_k) * f$ and hence

$$|\partial \Delta_k f| \leq C 2^k \left| \sum_{j=k-1}^{k+1} \Delta_j f \right|.$$

Further properties of the Littlewood-Paley projectors will be introduced as they are needed. ¶

3.43 Remark
Using the language of the Littlewood-Paley projectors, we can rewrite Corollary 3.40 as the statements

$$\left|\Delta_k G_t^{(\text{wave})}\right| \lesssim 2^{k(d-1)/2}|t|^{-(d-1)/2};$$
$$\left|\Delta_k \partial_t G_t^{(\text{wave})}\right| \lesssim 2^{k(d+1)/2}|t|^{-(d-1)/2}.$$
■

3.44 Definition (Besov spaces)
The homogeneous Besov (semi-)norm $\mathring{B}^{s,p}_q$ on $\mathscr{S}$ is defined by

$$\|f\|_{\mathring{B}^{s,p}_q} \stackrel{\text{def}}{=} \left[\sum_{k\in\mathbb{Z}} 2^{skq}\|\Delta_k f\|^q_{L^p}\right]^{\frac{1}{q}}.$$

The corresponding inhomogeneous norm is

$$\|f\|_{B^{s,p}_q} \stackrel{\text{def}}{=} \left[\|\Delta_{\leq 0} f\|^q_{L^p} + \sum_{k=1}^{\infty} 2^{skq}\|\Delta_k f\|^q_{L^p}\right]^{\frac{1}{q}}.$$
■

3.45 Exercise
Using Plancherel, show that for k an integer, the homogeneous norms $\mathring{B}^{k,2}_2$ and $\mathring{W}^{k,2}$ are equivalent, and that the inhomogeneous norms $B^{k,2}_2$ and $W^{k,2}$ are also equivalent. (Recall that two (semi-)norms A, A' are said to be equivalent if there exists a universal constant C such that

$$C^{-1}\|f\|_A \leq \|f\|_{A'} \leq C\|f\|_A$$

for all f.) ■

With the above definitions, we can summarize the decay estimate for the solutions to wave equations in the following theorem.

3.46 Theorem (Decay estimate for wave)
Let $\phi(t,x)$ solve the linear wave equation, with initial data $\phi(0,x) = \phi_0(x)$ and $\partial_t\phi(0,x) = \phi_1(x)$ both in $\mathscr{S}$. Then there exists a universal constant C depending only on the dimension d such that

$$\left|\phi(t,x)\right| \leq C|t|^{-\frac{d-1}{2}}\left(\left\|\phi_0\right\|_{\mathring{B}_1^{\frac{d+1}{2},1}} + \left\|\phi_1\right\|_{\mathring{B}_1^{\frac{d-1}{2},1}}\right).$$
■

3.47 Remark
One can prove that $\mathring{B}^{k,1}_1 \subset \mathring{W}^{k,1}$: simply write

$$\|f\|_{\mathring{W}^{k,1}} = \left\|\sum \Delta_\ell f\right\|_{\mathring{W}^{k,1}} \leq \sum \|\Delta_\ell f\|_{\mathring{W}^{k,1}} = \sum_{\ell\in\mathbb{Z}} \sum_{|\alpha|=k} \|\partial^\alpha \Delta_k f\|_{L^1}.$$

By the pointwise comparison we have

$$\lesssim \sum_{\ell\in\mathbb{Z}} 2^{k\ell} \|\Delta_k f\|_{L^1} = \|f\|_{\mathring{B}^{k,1}_1}.$$

So the decay estimate morally is asking for not simply L^1 integrabiliy of the initial data, but rather also differentiability up to order $(d-1)/2$. This resonates with our earlier discussion which asserts that the fundamental solution of the wave equation cannot be written in the form of a bounded function. Instead, the fundamental solution $G_t^{(\text{wave})}$ is a bona fide tempered distribution.

Notice, however, that the estimate only requires $(d-1)/2$ derivatives. Compare this to Sobolev embedding results where to get from a space of the form $W^{k,1}$ to L^∞ we need d derivatives overall. The ability to beat the Sobolev embedding result is a mark of dispersion.

Comparing to the Schrödinger and Airy cases, where the estimates do not require any derivatives, we see that the main difference in the wave case is that the gradient of the phase function $\left|\nabla\eta\right|$ remains bounded for large ξ. For Schrödinger and Airy it but grows unboundedly. This directly contributes to needing to assume that $\widehat{\phi_0}$ decays as $\xi \to \infty$ for the wave case, and as weights in frequency space equates to regularity in physical space, this shows that the regularity control is necessary. In the Schrödinger and Airy cases we get additional cancellations for large ξ from the increasingly faster oscillation, and thus additional weights are not needed in frequency space. ■

We will close this chapter with a discussion of the decay estimates for the Klein-Gordon equation. The Klein-Gordon dispersion relation is not homogeneous, unlike the model cases of Schrödinger, Airy, and wave equations, and so we don't have a scaling argument handy. We can catch a few glimpses into the expected behaviors, however, by comparing the dispersion relation $\eta(\xi) = \langle\xi\rangle$ with the other cases that we have already dealt with.

First, notice that $\left|\nabla\eta\right| = |\xi/\langle\xi\rangle| \leq 1$ for all ξ. This means that there are no oscillatory cancellations to use near $\xi \to \infty$, and so any estimate that we prove we will expect to need some suitable number of derivatives, similar to the wave equation case.

On the other hand, computing the Hessian yields

$$\nabla^2_{ij}\eta = \frac{\delta_{ij}}{\langle\xi\rangle} - \frac{\xi_i\xi_j}{\langle\xi\rangle^3} = \frac{1}{\langle\xi\rangle^3}\Big(\delta_{ij}(1+\xi\cdot\xi) - \xi_i\xi_j\Big).$$

By Cauchy-Schwarz, the Hessian matrix is positive definite, and is bounded below by $\langle\xi\rangle^{-3}\delta_{ij}$. This suggests that we don't have the same sort of degeneracy worries that we saw for the wave equation, and that we will be able to prove estimates giving $t^{-d/2}$ decay.

We will approach the estimates similar to how we dealt with the wave equation case: instead of estimating directly $G_t^{(\mathrm{KG})}$, we will estimate the truncated version $\Delta_k G_t^{(\mathrm{KG})}$. Unlike the wave case, however, since our dispersion relation is non-homogeneous, we will only estimate $\Delta_k G_t^{(\mathrm{KG})}$ for $k \geq 1$, and in the low-frequency setting we will do one estimate for $\Delta_{\leq 0} G_t^{(\mathrm{KG})}$. We summarize the results in the following theorem.

3.48 Theorem (Klein-Gordon decay estimate)
The Klein-Gordon kernel has low frequency decay

$$\left|\Delta_{\leq 0} G_t^{(\mathrm{KG})}\right| \lesssim \langle t\rangle^{-d/2}$$

and high frequency decay for $k \geq 1$

$$\left|\Delta_k G_t^{(\mathrm{KG})}\right| \lesssim 2^{\frac{kd}{2}} |t|^{-d/2}.$$

■

Proof We need to provide estimates for

$$\int_{\mathbb{R}^d} \frac{1}{\langle\xi\rangle} e^{it\langle\xi\rangle + x\cdot\xi} \chi(\xi)\, \mathrm{d}\xi$$

where χ is either the low frequency projector ψ or the high frequency projector $\psi(2^{-k}\xi) - \psi(2^{1-k}\xi)$. For convenience we will assume that we have chosen our seed function ψ to be radially symmetric. Without loss of generality we can assume $\omega_0 = x/|x|$ points in the x^1 direction. We will write, as usual, $\eta(\xi) = t\langle\xi\rangle + x\cdot\xi$ the phase function. Observe that for a fixed t, x, the critical point of the phase function occurs when

$$\nabla\eta(\xi) = t\frac{\xi}{\langle\xi\rangle} + x = 0 \iff \frac{\xi}{\langle\xi\rangle} = -\frac{x}{t}. \tag{3.49}$$

Let us deal first with the case where χ is the low frequency projector. The estimate for $|t| \leq 1$ is trivial. We focus our attention to the case of large times. Recall that within the support of χ, we have $|\xi| \leq 2$, which implies that the ratio $|\xi/\langle\xi\rangle| \leq 2/\sqrt{5} < 1$. So in the region $|x| \geq t$, we have that $\left|\nabla\eta(\xi)\right| \geq (1-2/\sqrt{5})t$, and noting that higher derivatives of η are *independent of x* we conclude that by the same argument as Lemma 3.6 we get the *uniform in x* bound by $C_N\langle t\rangle^{-N}$ for any N.

In the region $|x| \leq t$, we have to deal with critical points. Let ξ_0 denote the unique solution to (3.49). Using our bump function ψ we can split the integral into

$$\int e^{i\eta}\frac{\chi(\xi)}{\langle\xi\rangle}\psi(\sqrt{t}(\xi-\xi_0))\,\mathrm{d}\xi + \int e^{i\eta}\frac{\chi(\xi)}{\langle\xi\rangle}\left[1-\psi(\sqrt{t}(\xi-\xi_0)\right]\,\mathrm{d}\xi.$$

The first integral as usual we estimate by the volume $\lesssim t^{-d/2}$. For the second integral, after integrating by parts we see that it suffices to control

$$\left|\int e^{i\eta}L^d\left(\frac{\chi(\xi)}{\langle\xi\rangle}\left[1-\psi(\sqrt{t}(\xi-\xi_0))\right]\right)\,\mathrm{d}\xi\right|$$

where

$$Lf = -\nabla\cdot\left(\frac{\nabla\eta}{\left|\nabla\eta\right|^2}f\right).$$

Where the integrand is supported, using the uniform lower bound of the Hesslan $\left|\nabla^2\eta\right| \gtrsim 1$ on the support of χ, we can conclude that $\left|\nabla\eta\right| \gtrsim \sqrt{t}$. Next, note that we have uniform bounds on higher derivatives of $\nabla\eta/\left|\nabla\eta\right|$. Finally, we have the uniform bound that $\left|\nabla^{(k)}\psi((\xi-\xi_0)\sqrt{t})\right| \lesssim |t|^{k/2}$ but also that $\operatorname{supp}\nabla\psi$ has volume bounded by $|t|^{-d/2}$. So putting everything together we also have uniform estimates for the second integral by $t^{-d/2}$, as needed.

Next we treat the higher frequency case where $|\xi| \approx 2^k$. We can re-write the integral in polar coordinates as

$$\int_0^\infty\int_{\mathbb{S}^{d-1}}\frac{1}{\langle r\rangle}e^{it\langle r\rangle}e^{ir|x|\omega_0\cdot\omega}\chi(r)r^{d-1}\,\mathrm{d}\omega\,\mathrm{d}r$$

where χ is the projection to frequency $\approx 2^k$. Doing a rescaling, we have that, if we take χ to be instead the projection to frequency ≈ 2, equivalently we need to control the integral

$$2^{kd}\int_1^4\int_{\mathbb{S}^{d-1}} \frac{1}{\langle 2^k r\rangle} e^{it\langle 2^k r\rangle} e^{i2^k r|x|\omega_0\cdot\omega}\chi(r) r^{d-1}\,\mathrm{d}\omega\,\mathrm{d}r.$$

In view of Exercise 3.38, we can choose $\phi_\pm : \mathbb{S}^{d-1}\to[0,1]$ supported on $\{\omega_0\cdot\omega > -\frac{1}{2}\}$ and $\{\omega_0\cdot\omega < \frac{1}{2}\}$ respectively, such that $\phi_+ + \phi_- = 1$. It suffices to consider

$$J(x,t) = 2^{kd}\int_1^4\int_{\mathbb{S}^{d-1}} e^{it\langle 2^k r\rangle} e^{i2^k r|x|\omega_0\cdot\omega}\chi(r)\phi_+(\omega)\frac{r^{d-1}}{\langle 2^k r\rangle}\,\mathrm{d}\omega\,\mathrm{d}r. \tag{3.50}$$

Denote by $I(r)$ the spherical integral

$$I(r) \stackrel{\text{def}}{=} e^{-i2^k r|x|}\int_{\mathbb{S}^{d-1}} e^{i2^k r|x|\omega_0\cdot\omega}\phi_+(\omega)\,\mathrm{d}\omega$$

we have that (3.50) can be rewritten as

$$J(x,t) = 2^{kd}\int_1^4 e^{it\langle 2^k r\rangle - i2^k r|x|} I(r)\chi(r)\frac{r^{d-1}}{\langle 2^k r\rangle}\,\mathrm{d}r. \tag{3.51}$$

Note that $I(r)$ satisfies, by Exercise 3.38,

$$\left|\partial_r^\ell I(r)\right| \lesssim |x|^{-(d-1)/2}2^{-k(d-1)/2}.$$

On the other hand, by taking the derivatives directly under the integral sign we also have

$$\left|\partial_r^\ell I(r)\right| \lesssim 2^{k\ell}|x|^\ell.$$

The two together implies that

$$\left|\partial_r^\ell I(r)\right| \lesssim 1.$$

To estimate (3.50), we write $\eta(r) = t\langle 2^k r\rangle - 2^k r|x|$. Observe that $\eta' = 2^k\left(\frac{2^k r}{\langle 2^k r\rangle}t - |x|\right)$; so $\eta'(r) = 0 \implies |x|/t \in [1/\sqrt{2}, 1]$. We consider two different cases. First, suppose $|x|/t \notin [1/2, 2]$. Then $\eta' \gtrsim 2^k(t+|x|)$ for some universal constant. Using that the higher derivatives of η are independent of x, we have the bounds

$$|J(x,t)| \lesssim \frac{2^{kd-k}}{2^{\ell k}(t+|x|)^{\ell}}$$

for all ℓ, we in particular have the desired decay result.

For $|x|/t \in [1/2, 2]$, observe that $\eta''(r) = \frac{2^{2k}}{\langle 2^k r\rangle^3} t \gtrsim 2^{-k} t$. So we can apply Lemma 3.10 and get

$$\begin{aligned} |J(x,t)| &\lesssim 2^{kd} 2^{k/2} t^{-1/2} \left\| r^{d-1} \langle 2^k r\rangle^{-1} \chi(r) I(r) \right\|_{W^{1,\infty}} \\ &\lesssim 2^{kd} 2^{k/2} t^{-1/2} |x|^{-(d-1)/2} 2^{-k(d-1)/2} 2^{-k}. \end{aligned}$$

Simplifying, we have exactly

$$|J(x,t)| \lesssim 2^{kd/2} t^{-d/2}$$

as desired. □

3.52 Remark

Notice that in the final paragraph of the proof above, when considering the case $|x|/t \in [1/2, 2]$, we have additional smoothing in the case $2^k|x| < 1$. In this case another available estimate is simply that

$$\left| e^{i\eta} I(r) \chi(r) \frac{r^{d-1}}{\langle 2^k r\rangle} \right| \lesssim 2^{-k}.$$

Therefore we have

$$|J(x,t)| \lesssim 2^{kd} 2^{-k} \le 2^{k(d/2-1)} t^{-d/2}$$

using that $2^{kd/2} \le |x|^{-d/2} \approx t^{-d/2}$. This agrees with the discussion before the statement of the theorem. In the lower frequency regime $2^k \le |x|^{-1}$, the solution behaves more similar to the Schrödinger case where we have good smoothing effects. This, however, is cancelled out by the higher frequency regime $2^k \ge |x|^{-1}$ where the wave-like effects dominate and we need the full $d/2$ derivatives in the estimate. ■

Chapter 4

Interpolation Theory: a Sampling

In Chapter 2 we saw that the Fourier transform $\mathscr{F}$ functions as a linear operator

$$\mathscr{F} : \begin{cases} L^1(\mathbb{R}^d) \to L^\infty(\mathbb{R}^d) \\ L^2(\mathbb{R}^d) \to L^2(\mathbb{R}^d) \\ \mathscr{S} \to \mathscr{S} \\ \mathscr{S}' \to \mathscr{S}' \end{cases} .$$

It is natural to ask: does the Fourier transform extend to mappings defined on $L^p(\mathbb{R}^d)$ for p between 1 and 2? And if so, what are the corresponding co-domains? One way of addressing this problem is through *Interpolation Theory*, a sample of which we will present in this chapter. The theory is not only applicable to understanding the Fourier transform; it also is an important tool in the modern proofs of Strichartz estimates, which is one of the fundamental dispersive estimates.

For more detailed discussion and a far more complete account, see Bergh and Löfström, Interpolation spaces. An introduction.

What is interpolation theory

Interpolation theory describes the family of results of the following prototypical form.

4.1 Given some normed linear spaces X_0, X_1, Y_0, Y_1. Then one can find some normed spaces X, Y (where X is thought of as "between" X_0 and X_1

and Y is between Y_0 and Y_1) such that whenever T acts as a bounded linear operator $X_0 \to Y_0$ and also $X_1 \to Y_1$, then T also acts as a bounded linear operator $X \to Y$ with the norm $\|T\|_{X\to Y}$ controlled uniformly by its norms $X_0 \to Y_0$ and $X_1 \to Y_1$. ¶

4.2 Remark
Pay attention to the order of quantifiers! The spaces X and Y are independent of the linear operator T: interpolation theory is less about "understanding how an operator behaves on an intermediate space" but more about "how to define intermediate spaces on which operators behave predictably". ■

So what is meant for a space X to be between X_0 and X_1? Suppose that X_0 and X_1 are vector subspaces of some larger vector spaces (for our purposes, the larger vector space can usually be taken to be $\mathscr{S}'$, the space of tempered distributions). Then two natural spaces can be constructed: first is the intersection

$$X_0 \cap X_1;$$

the second is the sum

$$X_0 + X_1 \stackrel{\text{def}}{=} \{x \in \mathscr{S}' \mid \exists x_0 \in X_0, x_1 \in X_1 \text{ s.t. } x = x_0 + x_1\}.$$

Now, suppose that X_0 and X_1 are equipped with norms $\|\bullet\|_{X_0}$ and $\|\bullet\|_{X_1}$ respectively. Then on $X_0 \cap X_1$ we can impose the norm

$$\|x\|_{X_0\cap X_1} \stackrel{\text{def}}{=} \max(\|x\|_{X_0}, \|x\|_{X_1}), \tag{4.3}$$

while on $X_0 + X_1$ we can put the norm

$$\|x\|_{X_0+X_1} \stackrel{\text{def}}{=} \inf_{x=x_0+x_1} \|x_0\|_{X_0} + \|x_1\|_{X_1}, \tag{4.4}$$

where the infimum is taken over all decompositions $x = x_0 + x_1$ where $x_0 \in X_0$ and $x_1 \in X_1$. These two norms have the nice property that for $i \in \{0, 1\}$, the natural injections

$$X_0 \cap X_1 \to X_i \to X_0 + X_1 \tag{4.5}$$

are all continuous.

4.6 Exercise
Check that (4.3) and (4.4) do define norms, and check that the maps in (4.5) are indeed continuous. ■

When we say that a normed space X is an *intermediate space* between X_0 and X_1, what is meant is that X can be identified as a vector subspace of $X_0 + X_1$ satisfying $X_0 \cap X_1 \subset X$, and is equipped with a norm $\|\bullet\|_X$ such that the injections

$$X_0 \cap X_1 \to X \to X_0 + X_1$$

are both continuous.

4.7 Convention
Throughout this chapter, when we speak of the measure space (E, Σ, μ), we always assume that μ is σ-finite. ■

4.8 Example
Let (E, Σ, μ) be any measure space, and let w_0, w_1 be two positive real measurable functions on E. Consider the spaces $X_{0,1}$ the weighted L^1 spaces with norm

$$\|f\|_{X_i} = \int_E |f| w_i \, \mathrm{d}\mu.$$

Then we see that $f \in X_0 \cap X_1$ if and only if both $\|f\|_{X_0}$ and $\|f\|_{X_1}$ are bounded. This implies that

$$\int_E |f| \max(w_0, w_1) \, \mathrm{d}\mu < \infty \tag{4.9}$$

since $\max(w_0, w_1) \leq w_0 + w_1$. Conversely, if f is such that (4.9) holds, then necessarily $\|f\|_{X_0}$ and $\|f\|_{X_1}$ are bounded. The norm defined by (4.9) is equivalent to the norm $\max(\|f\|_{X_0}, \|f\|_{X_1})$, for

$$\begin{aligned} \max(\|f\|_{X_0}, \|f\|_{X_1}) &\leq \int_E |f| \max(w_0, w_1) \, \mathrm{d}\mu \\ &\leq \int_E |f|(w_0 + w_1) \, \mathrm{d}\mu \leq 2 \max(\|f\|_{X_0}, \|f\|_{X_1}). \end{aligned}$$

For f to be in $X_0 + X_1$, we need to be able to write $f = f_0 + f_1$ with $f_i \in X_i$. This implies that

$$\int_E |f| \min(w_0, w_1) \, \mathrm{d}\mu < \infty \tag{4.10}$$

since

$$\int_E |f|\min(w_0,w_1)\,\mathrm{d}\mu \le \int_E (|f_0|+|f_1|)\min(w_0,w_1)\mathrm{D}\mu$$
$$\le \int_E |f_0|w_0\,\mathrm{d}\mu + \int_E |f_1|w_1\,\mathrm{d}\mu \le \|f_0\|_{X_0} + \|f_1\|_{X_1}.$$

Conversely, we also have if (4.10) holds, defining $f_0 = f\cdot\mathbf{1}_{\{w_0\le w_1\}}$ and $f_1 = f\cdot\mathbf{1}_{\{w_1<w_0\}}$ we have $f = f_0+f_1$ and

$$\int_E |f_i|w_i\,\mathrm{d}\mu = \int_E |f_i|\min(w_0,w_1)\,\mathrm{d}\mu \le \int_E |f|\min(w_0,w_1)\,\mathrm{d}\mu \tag{4.11}$$

using the disjoint support of f_0 and f_1. Finally, observe that the norm defined by (4.10) is exactly equal to the norm

$$\inf_{f_0+f_1=f}\|f_0\|_{X_0}+\|f_1\|_{X_1}.$$

For by our computations before we have demonstrated that whenever $f_0+f_1=f$, the inequality

$$\int_E |f|\min(w_0,w_1)\,\mathrm{d}\mu \le \|f_0\|_{X_0}+\|f_1\|_{X_1},$$

which implies this norm is bounded above by $\|f\|_{X_0+X_1}$. On the other hand, the computations surrounding (4.11) shows that there exists some $f_1\in X_1$ and $f_0\in X_0$ with $\|f_1\|_{X_1}+\|f_0\|_{X_0}\le\int_E|f|\min(w_0,w_1)\,\mathrm{d}\mu$, and so (4.11) gives an alternative characterization of $\|\bullet\|_{X_0+X_1}$ in this set-up.

If we let w be any function on E that satisfies the pointwise bound

$$\min(w_0,w_1)\le w\le\max(w_0,w_1),$$

then clearly the space defined by the norm

$$\|f\|_X = \int_E |f|w\,\mathrm{d}\mu$$

satisfies

$$\int_E |f|\min(w_0,w_1)\,\mathrm{d}\mu \le \|f\|_X \le \int_E |f|\max(w_0,w_1)\,\mathrm{d}\mu.$$

This shows, using the discussion above, that

$$X_0 \cap X_1 \to X \to X_0 + X_1$$

are continuous maps. And so X is an intermediate space of X_0 and X_1. ■

4.12 Example
Let again (E, Σ, μ) be a measure space. Consider the usual space of L^p functions on E. We will show that if $p_0 < p < p_1$, then L^p is an intermediate space of L^{p_0} and L^{p_1}. Notice that if p is between p_0 and p_1, there exists a unique $\theta \in (0,1)$ such that

$$\frac{1}{p} = \frac{1-\theta}{p_0} + \frac{\theta}{p_1}.$$

This can be re-written as

$$1 = \left(\frac{p_0}{(1-\theta)p}\right)^{-1} + \left(\frac{p_1}{\theta p}\right)^{-1}. \tag{4.13}$$

Now, let $f \in L^{p_0} \cap L^{p_1}$. Observe that

$$\int_E |f|^p \, \mathrm{d}\mu = \int_E |f|^{p(1-\theta)} |f|^{p\theta} \, \mathrm{d}\mu$$
$$\leq \left\| |f|^{p(1-\theta)} \right\|_{L^{\frac{p_0}{(1-\theta)p}}} \left\| |f|^{p\theta} \right\|_{L^{\frac{p_1}{\theta p}}} = \|f\|_{L^{p_0}}^{(1-\theta)p} \|f\|_{L^{p_1}}^{\theta p}$$

where the inequality is an application of Hölder's inequality in view of (4.13). This result, that $\|f\|_{L^p} \leq \|f\|_{L^{p_0}}^{1-\theta} \|f\|_{L^{p_1}}^{\theta}$, is usually referred to as the *log-convexity of the L^p norms*. By this inequality we have that $L^{p_0} \cap L^{p_1} \to L^p$ continuously, since

$$\|f\|_{L^{p_0}}^{1-\theta} \|f\|_{L^{p_1}}^{\theta} \leq \max(\|f\|_{L^{p_0}}, \|f\|_{L^{p_1}}).$$

Next, let $f \in L^p$ be a nontrivial function. For an arbitrary $\lambda > 0$, we can let $f_0 = f \cdot \mathbf{1}_{\{|f| \geq \lambda\}}$ and $f_1 = f \cdot \mathbf{1}_{\{|f| < \lambda\}}$. Then we have that pointwise, $\left|\frac{f_0}{\lambda}\right|^{p_0} \leq \left|\frac{f}{\lambda}\right|^p$, and $\left|\frac{f_1}{\lambda}\right|^{p_1} \leq \left|\frac{f}{\lambda}\right|^p$. This implies that $f_0 \in L^{p_0}$ and $f_1 \in L^{p_1}$ while $f_0 + f_1 = f$. Therefore $f \in L^{p_0} + L^{p_1}$. We can compute

$$\|f_0\|_{L^{p_0}}^{p_0} = \int_E |f_0|^{p_0} \, \mathrm{d}\mu \leq \lambda^{p_0 - p} \int_E |f|^p \, \mathrm{d}\mu = \lambda^{p_0 - p} \|f\|_{L^p}^p$$

and similarly $\|f_1\|_{L^{p_1}}^{p_1} \leq \lambda^{p_1-p}\|f\|_{L^p}^p$. This implies

$$\|f\|_{L^{p_0}+L^{p_1}} \leq \|f_0\|_{L^{p_0}} + \|f_1\|_{L^{p_1}} \leq \lambda^{1-\frac{p}{p_0}}\|f\|_{L^p}^{\frac{p}{p_0}} + \lambda^{1-\frac{p}{p_1}}\|f\|_{L^p}^{\frac{p}{p_1}}.$$

So if we choose $\lambda = \|f\|_{L^p}$, we conclude that

$$\|f\|_{L^{p_0}+L^{p_1}} \leq 2\|f\|_{L^p}$$

showing that the mapping $L^p \to L^{p_0} + L^{p_1}$ is continuous. ■

4.14 Definition
Given X_0, X_1 and Y_0, Y_1 normed spaces. We say that a pair (X, Y) *interpolates between the pairs* (X_0, Y_0) and (X_1, Y_1) if X is an intermediate space between X_0 and X_1; and Y is an intermediate space between Y_0 and Y_1; such that whenever T acts both as a bounded linear operator from $X_0 \to Y_0$, and also as a bounded linear operator from $X_1 \to Y_1$, then T acts as a bounded linear operator from $X \to Y$. ■

4.15 Remark
Given the pairs (X_0, Y_0) and (X_1, Y_1), then the pairs $(X_0 \cap X_1, Y_0 \cap Y_1)$ and $(X_0 + X_1, Y_0 + Y_1)$ are both interpolants. That they are pairs formed of intermediate spaces are obvious. It remains to check the condition on linear operators. Suppose that the operator norms $\|T\|_{X_0\to Y_0} = M_0$ and $\|T\|_{X_1\to Y_1} = M_1$.

For the intersection spaces, notice that

$$\begin{aligned}\|Tf\|_{Y_0\cap Y_1} &= \max(\|Tf\|_{Y_0}, \|Tf\|_{Y_1}) \leq \max(M_0\|f\|_{X_0}, M_1\|f\|_{X_1}) \\ &\leq \max(M_0, M_1)\max(\|f\|_{X_0}, \|f\|_{X_1}) = \max(M_0, M_1)\|f\|_{X_0\cap X_1}.\end{aligned}$$

So we have that

$$\|T\|_{X_0\cap X_1\to Y_0\cap Y_1} \leq \max(M_0, M_1).$$

For the sum spaces, we have

$$\begin{aligned}\|Tf\|_{Y_0+Y_1} &= \inf_{Tf=g_0+g_1} \|g_0\|_{Y_0} + \|g_1\|_{Y_1} \\ &\leq \inf_{f=f_0+f_1} \|Tf_0\|_{Y_0} + \|Tf_1\|_{Y_1} \leq \inf_{f=f_0+f_1} M_0\|f_0\|_{X_0} + M_1\|f_1\|_{X_1} \\ &\leq \max(M_0, M_1) \inf_{f=f_0+f_1} \|f_0\|_{X_0} + \|f_1\|_{X_1} = \max(M_0, M_1)\|f\|_{X_0+X_1}.\end{aligned}$$

And so

$$\|T\|_{(X_0+X_1)\to(Y_0+Y_1)} \leq \max(M_0, M_1).$$ ■

The idea behind interpolation theory is two-fold:

1. First, given the pairs (X_0, Y_0) and (X_1, Y_1), construct systematically interpolating pairs (X, Y) in between.

2. Second, when the pairs (X_0, Y_0) and (X_1, Y_1) belong to standard scales of spaces (by this we mean that they are L^p spaces, Sobolev spaces, or Besov spaces, for example), identify the interpolant (X, Y) with spaces from the standard scales (up to equivalent norms).

With these types of results available, once one proves that a linear operator T acts continuously both on $X_0 \to Y_0$ and $X_1 \to Y_1$, one can automatically conclude that T acts continuously on any interpolant $X \to Y$.

In this chapter, we will give quick introductions to some of the general theory that allows us to construct the interpolating pairs (X, Y) from given space (X_0, Y_0) and (X_1, Y_1). The general techniques are split up into two flavors: the "complex" and the "real" methods. The former appeals to complex analytic tools in the construction; as we've already seen in some of our discussion of Schrödinger's equation, complex analytic tools tend to give exact formulae and sharp estimates. The same holds true for the complex method of interpolation. The latter appeals to real analytic, divide-and-conquer type tools. We've also already seen this in our discussion of oscillatory integrals: an integral of the form $\int e^{i\lambda\eta}\phi \, \mathrm{d}x$ is controlled by splitting into regions "near" the critical points of η, where one use one technique (volume estimate), and regions "far" from the critical points, where one use another (repeated integration by parts). The real method tends to give weaker bounds, but makes up for it by being more broadly applicable. In the context of interpolation theory, this means that the real method more easily generalizes to nonlinear (especially sublinear or quasilinear) interpolation, and that it can sometimes recover interpolants that cannot be found from the complex method.

As it turns out, these general theory for constructing the interpolating pairs are fairly robust and well-developed, and their main ideas are not that hard to understand. More technically challenging, however, is the identification or computation of the interpolating spaces when the end-points are well-known, standard spaces. We will summarize a collection of such results in the sequel, as well as some of the well-known consequences of interpolation theory, many of which will be useful for studying dispersive equations.

The Complex Method and Riesz-Thorin

A standard presentation of the complex method is Calderón, "Intermediate spaces and interpolation, the complex method".

The complex method of interpolation is based on a particular manifestation of the "maximum modulus principle" in complex analysis, summarized in what is called the *three lines lemma*.

4.16 Lemma (Three Lines Lemma)
Let f be a bounded holomorphic function on the strip $0 < \mathfrak{Re}\, z < 1$ which extends continuously to its closure $0 \leq \mathfrak{Re}\, z \leq 1$. Suppose further that

$$M_0 = \sup_{t\in\mathbb{R}} |f(it)| < \infty, \quad M_1 = \sup_{t\in\mathbb{R}} |f(1+it)| < \infty.$$

Then for any $\theta \in (0,1)$ and any $t \in \mathbb{R}$ we have

$$|f(\theta + it)| \leq M_0^{1-\theta} M_1^{\theta}.$$

■

Proof We may assume $M_0 M_1 > 0$ (else since f is holomorphic, it vanishes identically). Let $\epsilon > 0$ be arbitrary and consider the function

$$F_\epsilon(z) = \frac{e^{-\epsilon z(1-z)} f(z)}{M_0^{1-z} M_1^z}.$$

This function is holomorphic on the strip; furthermore, as $\mathfrak{Re}(z - z^2) = \mathfrak{Re}\, z - (\mathfrak{Re}\, z)^2 + (\mathfrak{Im}\, z)^2$, we see that as $|\mathfrak{Im}\, z| \to \infty$ the exponential weight forces $F_\epsilon(z)$ to decay rapidly, and this decay is uniform in $\mathfrak{Re}\, z$ since by assumption $f(z)$ is bounded. Therefore there exists $\lambda > 0$ such that whenever $|\mathfrak{Im}\, z| \geq \lambda$, then $|F_\epsilon(z)| < 1$. On the other hand, one easily checks that

$$F_\epsilon(it) = \frac{e^{-\epsilon t^2 - i\epsilon t} f(it)}{M_0^{1-it} M_1^{it}}$$

has norm bounded by 1, and similarly for $F_\epsilon(1+it)$. Applying the maximum modulus principle to the rectangle $\{\mathfrak{Re}\, z \in [0,1], \mathfrak{Im}\, z \in [-\lambda, \lambda]\}$ we see that $|F_\epsilon(z)| \leq 1$ on the rectangle. Since outside the rectangle we known from our construction that $|F_\epsilon(z)| < 1$, we conclude that $|F_\epsilon(z)| < 1$ for all z with $\mathfrak{Re}\, z \in [0,1]$. This implies

$$|f(\theta + it)| \leq e^{\epsilon(\theta - \theta^2 + t^2)} M_0^{1-\theta} M_1^{\theta}.$$

Since this holds for all $\epsilon > 0$, taking the limit $\epsilon \to 0$ gives us the desired result. □

4.17 (The Complex Method) For the discussion, we will assume all Banach spaces are defined over the scalars $\mathbb{C}$. The complex method is based on studying Banach-space valued holomorphic functions on the strip $0 < \Re z < 1$; the theory is largely the same (at least of what we need) to the theory of $\mathbb{C}$-valued holomorphic functions. For convenience we will write in this section $\mathcal{D} = \{0 < \Re z < 1\}$ and $\overline{\mathcal{D}} = \{0 \leq \Re z \leq 1\}$.

Let X_0, X_1 be two Banach spaces. Denote by $\mathscr{H}(X_0, X_1)$ the set of all functions $f : \overline{\mathcal{D}} \to X_0 + X_1$, that satisfies the conditions

1. f is bounded and continuous on $\overline{\mathcal{D}}$;
2. f is holomorphic on $\mathcal{D}$;
3. the mapping $\mathbb{R} \ni t \mapsto f(it)$ is continuous into X_0, and converges to 0 as $t \to \pm\infty$;
4. the mapping $\mathbb{R} \ni t \mapsto f(1 + it)$ is continuous into X_1, and converges to 0 as $t \to \pm\infty$.

$\mathscr{H}(X_0, X_1)$ is a vector space over $\mathbb{C}$, we can make it a normed space with the norm

$$\|f\|_{\mathscr{H}(X_0,X_1)} = \max(\sup_{t\in\mathbb{R}} \|f(it)\|_{X_0}, \sup_{t\in\mathbb{R}} \|f(1 + it)\|_{X_1}). \tag{4.18}$$

It turns out that $\mathscr{H}(X_0, X_1)$ is complete with respect to this norm, making it a Banach space; the proof is not difficult but slightly technical, so we omit it here.

We define the following spaces of X_0, X_1, where $\theta \in (0, 1)$ is a parameter.

$$X_{[\theta]} = (X_0, X_1)_{[\theta]} \stackrel{\text{def}}{=} \{x \in X_0 + X_1 \mid \exists f \in \mathscr{H}(X_0, X_1) \text{ s.t. } f(\theta) = x\}. \tag{4.19}$$

This space can be equipped with the norm

$$\|x\|_{X_{[\theta]}} = \inf\{\|f\|_{\mathscr{H}(X_0,X_1)} \mid f \in \mathscr{H}(X_0, X_1) \text{ s.t. } f(\theta) = x\}. \tag{4.20}$$

We claim that $X_{[\theta]}$ is an intermediate space between X_0 and X_1. First, observe that by assumption f is a bounded map into $X_0 + X_1$, so essentially the same argument of Lemma 4.16 shows that

$$\|f(\theta + it)\|_{X_0+X_1} \leq \left(\sup_{t\in\mathbb{R}} \|f(it)\|_{X_0+X_1}\right)^{1-\theta} \left(\sup_{t\in\mathbb{R}} \|f(1 + it)\|_{X_0+X_1}\right)^{\theta}.$$

Using that by definition

$$\|f(it)\|_{X_0+X_1} \leq \|f(it)\|_{X_0}, \qquad \|f(1 + it)\|_{X_0+X_1} \leq \|f(1 + it)\|_{X_1},$$

we conclude that

$$\|f(\theta+it)\|_{X_0+X_1} \leq \max(\sup_{t\in\mathbb{R}}\|f(it)\|_{X_0}, \sup_{t\in\mathbb{R}}\|f(1+it)\|_{X_1}) = \|f\|_{\mathscr{H}(X_0,X_1)}.$$

By the definition of $X_{[\theta]}$, we see this means

$$\|x\|_{X_0+X_1} \leq \|x\|_{X_{[\theta]}}$$

so that the mapping $X_{[\theta]} \to X_0 + X_1$ is continuous.

Next, let $x \in X_0 \cap X_1$. Consider the holomorphic function

$$f(z) = e^{\epsilon(z-\theta)^2} x$$

where $\epsilon > 0$ is any fixed constant. We see that since

$$\mathfrak{Re}(z-\theta)^2 = (\mathfrak{Re} z - \theta)^2 - (\mathfrak{Im} z)^2 \leq 1 - (\mathfrak{Im} z)^2,$$

we have that f is bounded and decays to zero uniformly as $(\mathfrak{Im} z) \to \pm\infty$, and hence $f \in \mathscr{H}(X_0, X_1)$. This test function shows that

$$\begin{aligned}\|x\|_{X_{[\theta]}} \leq \|f\|_{\mathscr{H}(X_0,X_1)} &= \max(\sup_{t\in\mathbb{R}}\|f(it)\|_{X_0}, \sup_{t\in\mathbb{R}}\|f(1+it)\|_{X_0}) \\ &= \max(e^{\epsilon\theta^2}\|x\|_{X_0}, e^{\epsilon(1-\theta)^2}\|x\|_{X_1}) \leq e^{\epsilon}\|x\|_{X_0\cap X_1}.\end{aligned}$$

This shows that $X_0 \cap X_1 \to X_{[\theta]}$ is continuous. And thus $X_{[\theta]}$ is an intermediate space as claimed. ¶

4.21 Theorem (Complex interpolation)
Let X_0, X_1, Y_0, Y_1 be Banach spaces. Let $T : X_0 + X_1 \to Y_0 + Y_1$ be a linear mapping such that T restricts to continuous mappings $X_0 \to Y_0$ and $X_1 \to Y_1$ with norms M_0 and M_1 respectively. Then for any $\theta \in (0,1)$, the mapping $T : X_{[\theta]} \to Y_{[\theta]}$ is bounded with norm $M_0^{1-\theta} M_1^{\theta}$. (In other words, $(X_{[\theta]}, Y_{[\theta]})$ interpolates between (X_0, Y_0) and (X_1, Y_1).) ■

Proof It suffices to prove that, given $x \in X_{[\theta]}$, we have $Tx \in Y_{[\theta]}$ and that $\|Tx\|_{Y_{[\theta]}} \leq M_0^{1-\theta} M_1^{\theta} X_{[\theta]}$. Since $x \in X_0 + X_1$, clearly $Tx \in Y_0 + Y_1$.

Now let $\epsilon > 0$ be arbitrary. By definition there exists $f \in \mathscr{H}(X_0, X_1)$ such that $f(\theta) = x$ and $\|f\|_{\mathscr{H}(X_0,X_1)} \leq \|x\|_{X_{[\theta]}} + \epsilon$. Consider $g = M_0^{z-1} M_1^{-z} (Tf)(z)$. Since f is bounded in $X_1 + X_2$, we have that g is bounded in $Y_1 + Y_2$, using that T extends to a bounded linear operator $X_1 + X_2 \to Y_1 + Y_2$ with norm $\max(M_0, M-1)$ (see Remark 4.15). Since $f(it)$ and $f(1+it)$ are continuous into X_0 and X_1 respectively and vanish as $t \to \infty$, the boundedness of T as linear operators $X_i \to Y_i$ implies that the same can be said of $g(it)$ and $g(1+it)$ as functions into Y_0 and Y_1. Hence we can conclude that $g \in \mathscr{H}(Y_0, Y_1)$. This implies that $g(\theta) = M_0^{\theta-1} M_1^{-\theta} Tx \in Y_{[\theta]}$.

We can compute the norm $\|g\|_{\mathscr{H}(Y_0,Y_1)}$: by definition this is equal to

$$\max(\sup_{t\in\mathbb{R}} \|g(it)\|_{Y_0}, \sup_{t\in\mathbb{R}} \|g(1+it)\|_{Y_1}) \\ \leq \max(\sup_{t\in\mathbb{R}} \|f(it)\|_{X_0}, \sup_{t\in\mathbb{R}} \|f(1+it)\|_{X_1}) = \|f\|_{\mathscr{H}(X_0,X_1)}.$$

So this implies that

$$M_0^{\theta-1} M_1^{-\theta} \|Tx\|_{Y_{[\theta]}} \leq \|g\|_{\mathscr{H}(Y_0,Y_1)} \leq \|f\|_{\mathscr{H}(X_0,X_1)} \leq \|x\|_{X_{[\theta]}} + \epsilon$$

by our initial choice of f. However, since the above inequality holds for every $\epsilon > 0$, we conclude that

$$\|T\|_{X_{[\theta]} \to Y_{[\theta]}} \leq M_0^{1-\theta} M_1^{\theta}$$

as claimed. □

4.22 Remark

Theorem 4.21 highlights an important facet of the construction given in Thought 4.17, namely that the intermediate spaces $X_{[\theta]}$ thus constructed are "universal" as interpolation spaces. In general, however, the fact that (X,Y) interpolates between (X_0,Y_0) and (X_1,Y_1) does not mean that either X and Y functions well for interpolating between other spaces; for example, there is no guarantee that (X,X) in fact interpolates between (X_0,X_0) and (X_1,X_1).

This also highlights one of the main points of interpolation theory, that of finding systematic methods of constructing interpolants. This universality manifest as the fact that the procedure described in Thought 4.17 which is applicable to essentially arbitrary pairs of Banach spaces (provided that $X_0 + X_1$ makes sense) can be captured as a *functor*. ■

Having given a general method of constructing interpolants, it remains to tackle the second goal of interpolation theory, that of identifying $X_{[\theta]}$ when X_0 and X_1 are well-known spaces. The results in this direction are many; we give only the most basic and most classical of the results, and describe some of its consequences.

4.23 Theorem (Riesz-Thorin)
Fix a measure space (E, Σ, μ). Let $p_0, p_1 \in [1, \infty]$, and $\theta \in (0,1)$. Setting p to be given by

$$\frac{1}{p} = \frac{1-\theta}{p_0} + \frac{\theta}{p_1},$$

then $(L^{p_0}, L^{p_1})_{[\theta]} = L^p$ with equal norms. ∎

Proof First we prove that if $f : E \to \mathbb{C}$ is in L^p, then $f \in (L^{p_0}, L^{p_1})_{[\theta]}$ with

$$\|f\|_{(L^{p_0}, L^{p_1})_{[\theta]}} \leq \|f\|_{L^p}.$$

Without loss of generality, we can assume $\|f\|_{L^p} = 1$. Let $\epsilon > 0$ be arbitrary. For $z \in \overline{\mathcal{D}}$ we define the function $F_z : E \to \mathbb{C}$ by

$$F_z(x) = \begin{cases} e^{\epsilon z(z-1) - \epsilon\theta(\theta-1)} |f(x)|^{\frac{p}{p(z)} - 1} f(x), & f(x) \neq 0; \\ 0, & f(x) = 0; \end{cases} \tag{4.24}$$

where $p(z)$ is defined by

$$\frac{1}{p(z)} = \frac{1-z}{p_0} + \frac{z}{p_1}.$$

Notice that $\Re\mathfrak{e}(1/p(z)) = 1/p(\Re\mathfrak{e}\, z)$ and that $p(\theta) = p$, and also that $F_\theta = f$. Furthermore, by construction, we have that

$$\|F_z\|_{L^{p(\Re\mathfrak{e} z)}}^{p(\Re\mathfrak{e} z)} = e^{\epsilon[\Re\mathfrak{e}\, z(z-1) - \theta(\theta-1)]} \underbrace{\|f\|_{L^p}^p}_{=1}. \tag{4.25}$$

which decays uniformly to zero as $\Im\mathfrak{m}(z) \to \pm\infty$. Using that $L^{p(\Re\mathfrak{e} z)}$ is an intermediate space of $L^{p_0} + L^{p_1}$ (see Example 4.12), this implies that the mapping $F : \overline{\mathcal{D}} \ni z \mapsto F_z \in L^{p_0} + L^{p_1}$ is in fact in $\mathscr{H}(L^{p_0}, L^{p_1})$. Therefore we conclude that

$$\|f\|_{(L^{p_0}, L^{p_1})_{[\theta]}} \leq \|F\|_{\mathscr{H}(L^{p_0}, L^{p_1})} = \max(\sup_{t \in \mathbb{R}} \|F_{it}\|_{L^{p_0}}, \sup_{t \in \mathbb{R}} \|F_{1+it}\|_{L^{p_1}}).$$

By (4.25), we have then (using that $p_0, p_1 \geq 1$)

$$\|f\|_{(L^{p_0}, L^{p_1})_{[\theta]}} \leq \exp[\epsilon\theta(1-\theta)].$$

Since $\epsilon > 0$ is arbitrary, this implies $\|f\|_{(L^{p_0}, L^{p_1})_{[\theta]}} \leq 1$ as desired.

To prove the reverse inclusion we will use the dual characterization of L^p spaces, namely that

$$\|f\|_{L^p} = \sup \int_E f g \, \mathrm{d}\mu$$

where the infimum is taken over the set of all $L^{p'}$ functions (where $(p')^{-1} + p^{-1} = 1$; note that by assumption $\theta \in (0,1)$ and hence $p \in (1,\infty)$) with $\|g\|_{L^{p'}} = 1$. Using that functions in $L^{p'}$ can be arbitrarily well approximated by functions which are bounded and have supports that have finite measure, we can in fact take the infimum over this class of functions with $\|g\|_{L^{p'}} = 1$. Note that if g is bounded and $\mu(\operatorname{supp} g) < \infty$ then necessarily $g \in L^1 \cap L^\infty$. Now, for each such g, we can define G_z by first setting $q(z) = ((1-z)/p_0' + z/p_1')^{-1}$, and taking

$$G_z(x) = \begin{cases} |g(x)|^{\frac{p'}{q(z)}-1} g(x), & g(x) \neq 0; \\ 0, & g(x) = 0. \end{cases}$$

Since $p/q(z) \in [0,p]$, we conclude that G_z is a bounded function with finite-measure support, and so is in $L^1 \cap L^\infty$. Since $\|G_z\|_{L^1 \cap L^\infty}$ is bounded, the image of $z \mapsto G_z$ is also bounded in $L^{p_0'} + L^{p_1'}$. We further have that $G_{it} \in L^{p_0'}$ and $G_{1+it} \in L^{p_1'}$ with norms $= 1$ by direct computation.

Now, given $f \in (L^{p_0}, L^{p_1})_{[\theta]}$, by definition for every $\epsilon > 0$ there exists $F \in \mathscr{H}(L^{p_0}, L^{p_1})$ with $F_\theta = f$ and $\|F\|_{\mathscr{H}(L^{p_0},L^{p_1})} \leq (1+\epsilon)\|f\|_{(L^{p_0},L^{p_1})_{[\theta]}}$. Let g and G_z be as in the previous paragraph. Since for each z, $F_z \in L^{p_0} + L^{p_1}$ and $G_z \in L^{p_0'} \cap L^{p_1'}$, we have that $F_z G_z \in L^1$ by Hölder's inequality, so the function

$$h(z) = \int_E F_z G_z \, \mathrm{d}\mu$$

is a holomorphic function on $\mathcal{D}$ that extends boundedly and continuously to $\overline{\mathcal{D}}$. We know that

$$\sup_{t \in \mathbb{R}} |h(it)| \leq \sup_{t \in \mathbb{R}} \|F_{it}\|_{L^{p_0}} \|G_{it}\|_{L^{p_0'}} \leq \|F\|_{\mathscr{H}(L^{p_0},L^{p_1})},$$

and

$$\sup_{t \in \mathbb{R}} |h(1+it)| \leq \sup_{t \in \mathbb{R}} \|F_{1+it}\|_{L^{p_1}} \|G_{it}\|_{L^{p_1'}} \leq \|F\|_{\mathscr{H}(L^{p_0},L^{p_1})}.$$

And so applying Lemma 4.16 to h we get

$$\left|\int_E fg\,\mathrm{d}\mu\right| = |h(\theta)| \leq \|F\|_{\mathscr{H}(L^{p_0},L^{p_1})} \leq (1+\epsilon)\|f\|_{(L^{p_0},L^{p_1})_{[\theta]}}.$$

Taking $\epsilon \to 0$ and the infimum over all g gives us that $\|f\|_{L^p} \leq \|f\|_{(L^{p_0},L^{p_1})_{[\theta]}}$ as claimed. □

4.26 Corollary (Hausdorff-Young)
For any $p \in [1,2]$, the Fourier transform is a bounded linear operator $\mathscr{F} : L^p \to L^{p'}$, where p' is the conjugate exponent defined by

$$1 = \frac{1}{p} + \frac{1}{p'}.$$

■

Ref. 2.23: "Plancherel: Fourier transform is an L^2 isometry"

Proof The Fourier transform is an L^2 isometry by Proposition 2.23; by definition it also satisfies

$$\|\mathscr{F}\|_{L^1(\mathbb{R}^d)\to L^\infty(\mathbb{R}^d)} = \frac{1}{(2\pi)^{\frac{d}{2}}}.$$

So by Theorem 4.21

$$\mathscr{F} : (L^1, L^2)_{[\theta]} \to (L^\infty, L^2)_{[\theta]}$$

with norm $(2\pi)^{-\frac{d}{2}(1-\theta)}$. With $p \in (1,2)$, we can compute the corresponding θ by solving

$$\frac{1}{p} = \frac{(1-\theta)}{1} + \frac{\theta}{2}$$

to get

$$\theta = 2 - \frac{2}{p}.$$

Observe that necessarily

$$\frac{1-\theta}{\infty} + \frac{\theta}{2} = 1 - \frac{1}{p} = \frac{1}{p'}.$$

Incidentally,

$$\|\mathscr{F}\|_{L^p \to L^{p'}} = (2\pi)^{-\frac{(2-p)d}{2p}}.$$

□

4.27 Corollary (Young's inequality)
Suppose $k(x,y)$ is a measurable function on $\mathbb{R}^{d_1}\times\mathbb{R}^{d_2}$, such that there is a constant $C>0$ and some $r\in[1,\infty]$ such that

$$\sup_{x\in\mathbb{R}^{d_1}}\|k(x,\bullet)\|_{L^r(\mathbb{R}^{d_2})}\le C,\qquad \sup_{y\in\mathbb{R}^{d_2}}\left\|k(\bullet,y)\right\|_{L^r(\mathbb{R}^{d}_1)}\le C,$$

then for every $p,q\in[1,\infty]$ satisfying

$$1+\frac{1}{q}=\frac{1}{r}+\frac{1}{p},$$

we have

$$\left\|\int_{\mathbb{R}^{d_2}}k(\bullet,y)f(y)\,\mathrm{d}y\right\|_{L^q(\mathbb{R}^{d_1})}\le C\|f\|_{L^p(\mathbb{R}^{d_2})}$$

for every $f\in L^p$. ■

Proof Write T the mapping $Tf(x)=\int_{\mathbb{R}^{d_2}}k(x,y)f(y)\,\mathrm{d}y$. By Minkowski's inequality we have

$$|Tf(x)|\le\int_{\mathbb{R}^{d_2}}\left|k(x,y)\right|\left|f(y)\right|\,\mathrm{d}y$$

which we can estimate by Hölder's inequality

$$\le\|k(x,\bullet)\|_{L^r(\mathbb{R}^{d_2})}\|f\|_{L^{r'}(\mathbb{R}^{d_2})}\le C\|f\|_{L^{r'}}.$$

On the other hand, Minkowski's inequality also gives

$$\|Tf\|_{L^r(\mathbb{R}^{d_1})}\le\int_{\mathbb{R}^{d_2}}\left\|k(\bullet,y)\right\|_{L^r(\mathbb{R}^{d_1})}\left|f(y)\right|\,\mathrm{d}y.$$

So Hölder's inequality implies

$$\le\sup_{y\in\mathbb{R}^{d_2}}\left\|k(\bullet,y)\right\|_{L^r(\mathbb{R}^{d_1})}\|f\|_{L^1(\mathbb{R}^{d_2})}\le C\|f\|_{L^1}.$$

That is to say, $T:L^1(\mathbb{R}^{d_2})\to L^r(\mathbb{R}^{d_1})$ and $L^{r'}(\mathbb{R}^{d_2})\to L^\infty(\mathbb{R}^{d_1})$ both with norm $\le C$. Applying Theorem 4.21 we get

$$T:(L^1(\mathbb{R}^{d_2}),L^{r'}(\mathbb{R}^{d_2}))_{[\theta]}\to(L^r(\mathbb{R}^{d_1}),L^\infty(\mathbb{R}^{d_1}))_{[\theta]}$$

also with norm θ. A direct computation using Theorem 4.23 gives us the equivalent L^p and L^q norms. □

4.28 Remark
A direct consequence of Young's inequality above is the convolution inequality

$$\|f * g\|_{L^q} \le \|f\|_{L^r} \|g\|_{L^p}$$

when $1 + q^{-1} = r^{-1} + p^{-1}$. ■

4.29 Exercise (L^p decay of Schrödinger)
Let $\phi(t,x)$ solve Schrödinger's equation with initial data $\phi(0,x) = \phi_0(x) \in \mathscr{S}$.

1. Using the Fourier representation (2.44), prove that

$$\left\|\phi(t,\bullet)\right\|_{L^2} = \left\|\phi_0\right\|_{L^2}.$$

2. Combining the above with Corollary 3.29, prove that for every $p \in [2,\infty]$ there exists a constant C dependning on the dimension d and p such that

$$\left\|\phi(t,\bullet)\right\|_{L^p} \le C|t|^{-\frac{d}{2}(1-\frac{2}{p})}\left\|\phi_0\right\|_{L^{p'}}.$$

   Compute also the constant C. ■

4.30 Exercise (Baby Stein-Weiss)
This exercise builds upon Example 4.8. Fix a measure space (E,Σ,μ). Let $p \in [1,\infty)$, and let $w_0, w_1 : E \to [0,\infty)$ two μ-measurable functions. Define the norms, for $i = 0,1$,

$$\|f\|_{X_i} = \left(\int_E |f|^p w_i \,\mathrm{d}\mu\right)^{\frac{1}{p}}.$$

What is the complex interpolation space $(X_0, X_1)_{[\theta]}$? Model your proof after the proof of Theorem 4.23. ■

The Real Method and Marcinkiewicz

The real method is based on examining different *comparable norms* on $X_0 + X_1$. Observe that if we replace the norm $\|f\|_{X_1}$ by $\|f\|_{tX_1} = t\|f\|_{X_1}$, this gives an equivalent norm on X_1. These induce comparable norms of X_0+X_1, as for a fixed t,

$$\|f\|_{X_0+tX_1} = \inf_{f=f_0+f_1} \|f_0\|_{X_0} + t\|f_1\|_{X_1}$$
$$\leq \max(1,t)\cdot \inf_{f=f_0+f_1} \|f_0\|_{X_0} + \|f_1\|_{X_1} = \max(1,t)\|f\|_{X_0+X_1}.$$

But the "optimal splitting" of $f = f_0 + f_1$ can be drastically different for different t. For $f \in X_0 \cap X_1$ for example, one would expect that for $t \gg 1$ the optimum would be to put $f_0 = f$ and $f_1 = 0$; similarly, for $t \ll 1$ one would put $f_0 = 0$ and $f_1 = f$. The real method aims to capture the behavior of how the optimal splitting changes as one changes the weight t, and use that to characterize the intermediate spaces between X_0 and X_1.

4.31 Definition (K-functional)
For $t \in \mathbb{R}_+$ and $x \in X_0 + X_1$, we define

$$K(t,x;X_0,X_1) = \inf_{x=x_0+x_1} \|x_0\|_{X_0} + t\|x_1\|_{X_1}.$$

■

4.32 (Properties of the K-functional) Note that for fixed x_0, x_1, the function $t \mapsto \|x_0\|_{X_0} + t\|x_1\|_{X_1}$ is linear, and hence is both convex and concave. Using that the infimum of any family of concave functions is again concave, we have that for any $x \in X_0 + X_1$, the function $t \mapsto K(t,x;X_0,X_1)$ is a concave function.

We have pointwise control available for $K(t,x;X_0,X_1)$. For $x_0 \in X_0$ and $x_1 \in X_1$, we note first that, since

$$\|x_0\|_{X_0} + s\|x_1\|_{X_1} \leq \|x_0\|_{X_0} + t\|x_1\|_{X_1}, \qquad \text{when } s \leq t,$$

we get that $K(t,x;X_0,X_1)$ is increasing in t. This can be sharpened to

$$\min(1,t/s)\Big[\|x_0\|_{X_0} + s\|x_1\|_{X_1}\Big] \leq \|x_0\|_{X_0}+t\|x_1\|_{X_1} \leq \max(1,t/s)\Big[\|x_0\|_{X_0} + s\|x_1\|_{X_1}\Big].$$

This implies

$$\min(1,t)\|x\|_{X_0+X_1} \leq K(t,x;X_0,X_1) \leq \max(1,t)\|x\|_{X_0+X_1} \tag{4.33}$$

for any $x \in X_0 + X_1$. Furthermore, it is easy to check for $x \in X_0 \cap X_1$ that

$$K(t,x;X_0,X_1) \leq \min(\|x\|_{X_0}, t\|x\|_{X_1}) \leq \min(1,t)\|x\|_{X_0\cap X_1}. \tag{4.34}$$

A further consequence is that

$$\min(1, \frac{t}{s}) K(s,x;X_0,X_1) \leq K(t,x;X_0,X_1) \leq \max(1,\frac{t}{s}) K(s,x;X_0,X_1). \quad (4.35)$$

Lastly, we have the algebraic relation $K(t,x;X_0,X_1) = tK(t^{-1},x;X_1,X_0)$ for interchanging X_0 and X_1. ¶

4.36 Definition

Given X_0, X_1, we define the norms $X_{\theta,q}$ on $X_0 + X_1$. For $q \in [1,\infty)$ and $\theta \in (0,1)$

$$\|x\|_{X_{\theta,q}} \stackrel{\text{def}}{=} \left(\int_0^\infty t^{-\theta q - 1} K(t,x;X_0,X_1)^q \, \mathrm{d}t \right)^{\frac{1}{q}} \quad (4.37)$$

and for $q = \infty$ and $\theta \in [0,1]$

$$\|X\|_{X_{\theta,\infty}} \stackrel{\text{def}}{=} \sup_t t^{-\theta} K(t,x;X_0,X_1). \quad (4.38)$$

In other words, $X_{\theta,q}$ is the L^q norm of the function $t \mapsto t^{-\theta} K(t,x;X_0,X_1)$ on $\mathbb{R}_+$ with respect to the measure $t^{-1}\mathrm{d}t$.

We refer by the spaces $X_{\theta,q}$ the subspaces of $X_0 + X_1$ on which the corresponding norms are finite. ■

4.39 (Basic properties of $X_{\theta,q}$) From the definitions, we can show that

1. $(X_0,X_1)_{\theta,q} = (X_1,X_0)_{1-\theta,q}$.
2. The spaces $(X_0,X_1)_{\theta,q}$ can be shown to be intermediate spaces between X_0 and X_1. Using (4.33) we see that $\min(t^{-\theta}, t^{1-\theta})\|x\|_{X_0+X_1} \leq t^{-\theta}K(t,x;X_0,X_1)$. Let $w_\theta(t) = \min(t^{-\theta}, t^{1-\theta})$, we have that

   $$\|w_\theta\|_{L^q(t^{-1}\mathrm{d}t)} \|x\|_{X_0+X_1} \leq \|x\|_{X_{\theta,q}}.$$

   Using that for each admissible θ, q we have $\|w_\theta\|_{L^q(t^{-1}\mathrm{d}t)}$ is a well-defined real constant, we have that $X_{\theta,q} \to X_0 + X_1$ continuously. Similarly, by (4.34) we have that $t^{-\theta}K(t,x;X_0,X_1) \leq \min(t^{-\theta},t^{1-\theta})\|x\|_{X_0 \cap X_1}$. And so

   $$\|x\|_{X_{\theta,q}} \leq \|w_\theta\|_{L^q(t^{-1}\mathrm{d}t)} \|x\|_{X_0\cap X_1}$$

   showing the continuity of $X_0 \cap X_1 \to X_{\theta,q}$.

3. From (4.35) we get that

$$\min(t^{-\theta}, t^{1-\theta}s^{-1})K(s,x;X_0,X_1) \leq t^{-\theta}K(t,x;X_0,X_1)$$

which we can rewrite as

$$s^{-\theta}w_\theta(t/s)K(s,x;X_0,X_1) \leq t^{-\theta}K(t,x;X_0,X_1).$$

Using that the $L^q(t^{-1}\mathrm{d}t)$ norm is scale invariant, we have

$$s^{-\theta}\|w_\theta\|_{L^q(t^{-1}\mathrm{d}t)}K(s,x;X_0,X_1) \leq \|x\|_{X_{\theta,q}}. \tag{4.40}$$

This implies that $X_{\theta,q} \leq X_{\theta,\infty}$. On the other hand, Hölder's inequality implies

$$\|f\|_{L^r} \leq \|f\|_{L^q}^{q/r}\|f\|_{L^\infty}^{1-q/r}$$

when $r > q$. This implies

$$\|x\|_{X_{\theta,r}} \leq \|x\|_{X_{\theta,q}}^{q/r}\|x\|_{X_{\theta,\infty}}^{1-q/r} \leq \|x\|_{X_{\theta,q}}$$

from the previous step. Therefore we conclude that whenever $r \geq q$ we have that $X_{\theta,q} \subseteq X_{\theta,r}$. ¶

4.41 Theorem (Real interpolation)
Let X_0, X_1, Y_0, Y_1 be Banach spaces. Let $T : X_0 + X_1 \to Y_0 + Y_1$ be a linear mapping such that T restricts continuously to $X_0 \to Y_0$ and $X_1 \to Y_1$ with norms M_0 and M_1 respectively. Then for $(\theta, q) \in (0,1) \times [1,\infty) \cup [0,1] \times \{\infty\}$ we have that $T : X_{\theta,q} \to Y_{\theta,q}$ with norm $\leq M_0^{1-\theta}M_1^\theta$. ■

Proof Observe that for $x = x_0 + x_1$ we have

$$K(t,Tx;Y_0,Y_1) \leq \|Tx_0\|_{Y_0} + t\|Tx_1\|_{Y_1} \leq M_0\|x_0\|_{X_0} + tM_1\|x_1\|_{X_1}.$$

This implies

$$K(t,Tx;Y_0,Y_1) \leq M_0K(t\frac{M_1}{M_0},x;X_0,X_1). \tag{4.42}$$

This we can rewrite, denoting by $\tau = t\frac{M_1}{M_0}$, as

$$t^{-\theta}K(t,Tx;Y_0,Y_1) \leq M_0^{1-\theta}M_1^\theta\tau^{-\theta}K(\tau,x;X_0,Y_0),$$

from which the theorem follows by our definitions. □

4.43 Remark (Sublinear operators)
Looking at proof of the previous theorem, we see that, unlike in the case of the complex method, here the fact that T is linear is not crucial. In fact, it suffices that T satisfies the property:

> For every $x_0 \in X_0$ and $x_1 \in X_1$, there exists $y_0 \in Y_0$ and $Y_1 \in y_1$ such that $T(x_0 + x_1) = y_0 + y_1$ and $\|y_0\|_{Y_0} \leq M_0 \|x_0\|_{X_0}$ and $\|y_1\|_{Y_1} \leq M_1 \|x_1\|_{X_1}$.

Note that by taking one of the $x_i = 0$, the above condition also implies that T as a mapping from $X_i \to Y_i$ is bounded. Such a condition would be sufficient to derive the bound $K(t, Tx; Y_0, Y_1) \leq M_0 K(t\frac{M_1}{M_0}, x; X_0, X_1)$. ■

Notice that in the real method the interpolating space $X_{\theta,q}$ depends on two parameters; this means that compared to the complex method we get a two dimensional (instead of one dimensional) family of spaces. For applications this means that sometimes we can get away with using weaker spaces for the endpoints of our mappings X_0, X_1, Y_0, Y_1. To illustrate this let us compute some of the interpolating spaces for some usual scales of functions. We begin with stating, without proof, a convenient technical lemma, which says that applying real interpolation twice does not give you more spaces.

The proof of the reiteration lemma for the ⊆ inclusion is straightforward and follows from careful application of (4.40). *The reverse inclusion is best done using the dual characterization of $X_{\theta,q}$ via the so-called J-method of real interpolation which we omit in this introduction. Please see Chapter 3 in Bergh and Löfström,* Interpolation spaces. An introduction *for more details.*

4.44 Lemma (Reiteration)
Given X_0, X_1, we have

$$(X_{\theta_0,q_0}, X_{\theta_1,q_1})_{\eta,q} = X_{\theta,q}$$

with equivalent norms whenever $\theta_0, \theta_1 \in (0,1)$ are distinct, $\eta \in (0,1)$, and $\theta = (1-\eta)\theta_0 + \eta\theta_1$. We also have

$$(X_{\theta,q_0}, X_{\theta,q_1})_{\eta,q} = X_{\theta,q}$$

with equivalent norms whenever $\theta, \eta \in (0,1)$, and $q_0, q_1, q \in [1,\infty]$ satisfies $q^{-1} = (1-\eta)q_0^{-1} + \eta q_1^{-1}$. ■

One of the first applications of the real interpolation method is the following definition.

The classical definition of Lorentz spaces present them as quasi-normed spaces with the quasi-norms defined via either the distribution function or via the decreasing rearrangement. Standard presentations then prove that the real interpolation spaces of Lebesgue spaces coincide with the Lorentz spaces. Here we choose to take the reverse route.

4.45 Definition (Lorentz spaces)
Given a measure space (E, Σ, μ), we define the Lorentz space L^p_q with $q \in (1,\infty)$ and $p \in (1,\infty)$ as

$$L^p_q = (L^\infty, L^1)_{\frac{1}{p},q}.$$

For $p \in (1,\infty)$ we also define

$$L^p_\infty = (L^\infty, L^1)_{\frac{1}{p},\infty}. \qquad \blacksquare$$

Then the Reiteration Lemma immediately implies

4.46 Corollary (Interpolation between Lorentz spaces)
Suppose p_0, p_1, q_0, q_1 are such that $L^{p_0}_{q_0}$ and $L^{p_1}_{q_1}$ are meaningfully defined as Lorentz spaces.

1. If $p_0 < p_1$, then for every $q \in (1,\infty]$ and for $p \in (p_0, p_1)$, we have

$$(L^{p_0}_{q_0}, L^{p_1}_{q_1})_{\eta,q} = L^p_q$$

   provided
$$\frac{1}{p} = \frac{1-\eta}{p_0} + \frac{\eta}{p_1}.$$

2. If $p_0 = p_1 = p$ and $q_0 < q_1$, then for every $q \in (q_0, q_1)$ we have

$$(L^p_{q_0}, L^p_{q_1})_{\eta,q} = L^p_q$$

   provided
$$\frac{1}{q} = \frac{1-\eta}{q_0} + \frac{\eta}{q_1}.$$

The cases $p_0 > p_1$ and $q_0 > q_1$ follows analogously. ■

4.47 Remark
Note that by properties of the real interpolation spaces we have $L^p_q \subseteq L^p_r$ whenever $r \geq q$. ■

In practice, to check whether a function is in a Lorentz space, we would like to have a more easily computable criterion. It turns out such a thing is available.

4.48 Definition (Decreasing rearrangement)
Given a measurable function f on a measure space (E, Σ, μ), we define its *decreasing rearrangement* to be the function $f^* : [0,\infty) \mapsto [0,\infty)$ defined by the expression

$$f^*(t) \stackrel{\text{def}}{=} \inf\{s \in [0,\infty) \mid \mu(\{|f| > s\}) \leq t\}. \qquad \blacksquare$$

4.49 (Decreasing rearrangement explained) The notion of decreasing arrangement is based on the definition of the Lebesgue integral. In defining the Lebesgue integral as the "area under the curve", we slice the "area" horizontally (as opposed to Riemann integration where we slice vertically), each of the slices correspond to the superlevel set $\{|f| > s\}$ for some $s \in [0,\infty)$. We then sum over all possible "heights" s the corresponding areas $\mu(\{|f| > s\})$ to obtain the integral. Lebesgue integral is by definition compatible with Cavalieri's principle: horizontally moving the slices maintains the Lebesgue integral of the function. Now imagine our original measure space to be $E = [0,\infty)$, with Σ the Borel σ-algebra and μ the Lebesgue measure. The decreasing rearrangement of a measurable function f then is formed by pushing each of the superlevel sets as far to the left as possible, so they sit right against the vertical axis. This forces the resulting function f^* to be monotonically decreasing, and hence the name.

For general measure spaces, the definition above produces f^* which is equimeasure, in the sense that the Lebesgue measure of the superlevel sets $\big|\{|f^*| > s\}\big| = \mu(\{|f| > s\})$ is equal to the measure of the superlevel sets of the original function. This implies that

$$\|f\|_{L^p} = \|f^*\|_{L^p}$$

when the left is measured with μ and the right is measured with the Lebesgue measure.

How fast is the growth and decay of f^* when $f \in L^p$? This can be answered by *Chebyshev's inequality* which can be written in the form

$$f^*(t) t^{\frac{1}{p}} \leq \|f\|_{L^p} \tag{4.50}$$

for any $f \in L^p$. ¶

4.51 Proposition
If $f \in L^1 + L^\infty$, then $K(\tau, f; L^1, L^\infty) = \int_0^\tau f^*(t)\,\mathrm{d}t$. ■

Proof Observe the following elementary fact: if a,b,c,d are non-negative numbers and $a+b \geq c+d$, then at least one of $a \geq c$ and $b \geq d$ is true. An immediate consequence of this fact is that if $f = f_0 + f_1$ is the sum of two measurable functions, and $s = s_0 + s_1$ is a sum of two non-negative reals, then

$$\mu(\{|f| > s\}) \leq \mu(\{|f_0| > s_0\}) + \mu(\{|f_1| > s_1\}).$$

From this we derive that for any $\theta \in (0,1]$

$$f^*(t) \leq f_0^*((1-\theta)t) + f_1^*(\theta t). \tag{4.52}$$

So

$$\begin{aligned}
\int_0^\tau f^*(t)\,\mathrm{d}t &\leq \int_0^\tau f_0^*((1-\theta)t)\,\mathrm{d}t + \int_0^\tau f_1^*(\theta t)\,\mathrm{d}t \\
&\leq \int_0^\infty f_0^*((1-\theta)t)\,\mathrm{d}t + \int_0^\tau f_1^*(\theta t)\,\mathrm{d}t \\
&\leq \frac{1}{1-\theta}\left\|f_0^*\right\|_{L^1} + \tau\left\|f_1^*\right\|_{L^\infty}.
\end{aligned}$$

The inequality holds for any $\theta \in (0,1]$ and any $f_0 \in L^1$ and $f_1 \in L^\infty$. This means that

$$\int_0^\tau f^*(t)\,\mathrm{d}t \leq K(\tau, f; L^1, L^\infty). \tag{4.53}$$

For the reverse inequality, choose

$$f_1(x) = \begin{cases} f(x) & |f(x)| \leq f^*(\tau) \\ s\cdot\operatorname{sgn}(f) & |f(x)| > f^*(\tau) \end{cases}, \quad f_0(x) = f(x) - f_1(x).$$

We have that $f_0 \in L^1$ and $f_1 \in L^\infty$ with $\|f_1\|_{L^\infty} = f^*(\tau)$. Note that by construction $f^* = f_0^* + f_1^*$, and so

$$\int_0^\tau f^*(t)\,\mathrm{d}t = \int_0^\tau f_0^*(t)\,\mathrm{d}t + \int_0^\tau f_1^*(t)\,\mathrm{d}t.$$

By construction $f_1^*(t) = f^*(\tau)$ for all $t < \tau$, so the second term in the integral is exactly $\tau\|f_1\|_{L^\infty}$. Also by construction $f_0^*(t) = 0$ for all $t > \tau$, so the first integral is exactly $\|f_0\|_{L^1}$. This shows that

$$K(\tau, f; L^1, L^\infty) \leq \|f_0\|_{L^1} + \tau\|f_1\|_{L^\infty} = \int_0^\tau f^*(t)\,\mathrm{d}t$$

and the proposition is proved. □

The following corollary is immediate using the definitions (4.37) and (4.38) of real interpolation spaces.

4.54 Corollary
For $q < \infty$, a function $f \in L^p_q$ if and only if

$$\int_0^\infty \left[t^{\frac{1}{p}-1} \int_0^t f^*(s)\, \mathrm{d}s \right]^q \frac{\mathrm{d}t}{t} < \infty.$$

Furthermore, a function $f \in L^p_\infty$ if and only if

$$\sup_t t^{\frac{1}{p}-1} \int_0^t f^*(s)\, \mathrm{d}s < \infty.$$

■

4.55 Lemma (Standard characterization of Lorentz spaces)
$f \in L^p_q$ if and only if the function $t \mapsto t^{\frac{1}{p}} f^*(t)$ is in $L^q([0,\infty), t^{-1}\mathrm{d}t)$. ■

Proof The lemma boils down to showing that $t^{\frac{1}{p}} f^*(t)$ is in $L^q(t^{-1}\mathrm{d}t)$ if and only if the function $g(t) = t^{\frac{1}{p}-1} \int_0^t f^*(s)\, \mathrm{d}s$ is in $L^q(t^{-1}\mathrm{d}t)$. First note that since $f^*(t)$ is decreasing, $\frac{1}{t}\int_0^t f^*(s)\, \mathrm{d}s \geq f^*(t)$, and hence the implication $\Leftarrow$ is immediate.

For the implication $\Rightarrow$, we first perform a change of variables to write

$$g(t) = t^{\frac{1}{p}} \int_0^1 f^*(ts)\, \mathrm{d}s.$$

Minkowski's inequality then implies

$$\|g(t)\|_{L^q(t^{-1}\mathrm{d}t)} \leq \int_0^1 \left[\int_0^\infty t^{\frac{q}{p}} [f^*(ts)]^q \frac{\mathrm{d}t}{t} \right]^{\frac{1}{q}} \mathrm{d}s.$$

A second change of variables implies

$$\|g(t)\|_{L^q(t^{-1}\mathrm{d}t)} \leq \int_0^1 s^{-\frac{1}{p}} \left[\int_0^\infty t^{\frac{q}{p}} [f^*(t)]^q \frac{\mathrm{d}t}{t} \right]^{\frac{1}{q}} \mathrm{d}s = \frac{p}{p-1} \left\| t^{\frac{1}{p}} f^*(t) \right\|_{L^q(t^{-1}\mathrm{d}t)}$$

as desired. □

4.56 Remark
Observe that $t^{-1}\mathrm{d}t$ is essentially the Haar measure on the Lie group $(\mathbb{R}_+,\cdot)$, the mapping

$$f^*(t) \mapsto t^{-1}\int_0^t f^*(\tau)\,\mathrm{d}\tau$$

can actually be expressed as a convolution. One way to see this would be to take the exponential change of variables $\sigma = \ln t$ to rewrite f^* as a function of $\sigma \in \mathbb{R}$. Then the fact that $\|g\|_{L^q} \lesssim \left\|t^{1/p} f^*(t)\right\|_{L^q}$ can be seen alternatively as a consequence of Young's inequality Corollary 4.27 for convolutions. ■

4.57 Corollary (Marcinkiewicz)

1. For $p \in (1,\infty)$, the space $L^p_p = L^p$.

2. As consequence: suppose $p_0 \neq p_1$ and both are in $(1,\infty]$ and $q_0, q_1 \in (1,\infty]$. Take T to be a sublinear mapping (see Remark 4.43) that is bounded from $L^{p_i} \to L^{q_i}_\infty$ $(i = 0,1)$. Then for any $\theta \in (0,1)$ such that

$$\frac{1}{p} = \frac{1-\theta}{p_0} + \frac{\theta}{p_1}$$
$$\frac{1}{q} = \frac{1-\theta}{q_0} + \frac{\theta}{q_1}$$
$$p \leq q$$

   we have T is bounded from L^p to L^q. ■

4.58 Remark
In the statement of the previous corollary we took $L^\infty_\infty = L^\infty$ by convention.

The requirement that $p \leq q$ is due to the fact that our theorem on real interpolation would actually imply that, for any $r \in (1,\infty)$ that the mapping T is bounded from the Lorentz space L^p_r to L^q_r, without the restriction on $p \leq q$. However, we further have that when $r \leq q$ the embedding $L^q_r \subset L^q_q$ is continuous. So when $p \leq q$ we can take $r = p$ and get that $T : L^p = L^p_p \to L^q_p \subset L^q_q = L^q$.

Lastly, the lower bound $p > 1$ is not necessary; one can define the Lorentz space L^p_q by interpolation between L^{p_0} and L^∞, with $q > p_0$, for any $p_0 > 0$ (in particular $p_0 < 1$), using essentially the same definition as was given above. With minor changes all of the above computations go through (just need to keep track of more indices). This allows extending the definition of Lorentz spaces to L^p_q with $p \leq 1$. The problem however is that the L^{p_0} space in the definition is not a normed space: the L^{p_0} "norm" fails the triangle inequality and is only a quasinorm. As a consequence, the Lorentz spaces with $p \leq 1$ are also only quasi-normed spaces (as opposed to the case $p > 1$ where the functionals given in Corollary 4.54 are norms when $q < \infty$). We will in general not be working with these quasi-normed spaces, except that the Marcinkiewicz Theorem with the edge case of mappings $T : L^1 \to L^1_\infty$ is frequently useful. ■

4.59 Definition (Weak-L^p spaces)
The (quasi-normed) spaces L^p_∞ for $p \in (0,\infty)$ or usually referred to as the *weak-L^p spaces*. They are characterized by

$$\sup_{t\in[0,\infty)} t^{\frac{1}{p}} f^*(t) < \infty.$$

■

Let us move on to the interpolation of another frequently used family of spaces.

4.60 (Sequence spaces) Fix $\lambda > 1$. Let $p_0, p_1 \in [1,\infty)$ and consider as the space X_0 the space of sequences $a = (a_n)_{n\in\mathbb{Z}}$ such that

$$\|a\|_{X_0} \stackrel{\text{def}}{=} \left(\sum_{n\in\mathbb{Z}} \lambda^{np_0} |a_n|^{p_0} \right)^{\frac{1}{p_0}} < \infty,$$

and X_1 the space of sequences such that

$$\|a\|_{X_1} \stackrel{\text{def}}{=} \left(\sum_{n\in\mathbb{Z}} |a_n|^{p_1} \right)^{\frac{1}{p_1}} < \infty.$$

What is $X_{\theta,q}$?

Let us begin with the case $p_0 = p_1 = 1$. We claim that

$$K(t,a;X_0,X_1) = \sum_{n=-\infty}^{\lfloor \log_\lambda t \rfloor} \lambda^n |a_n| + \sum_{n=\lceil \log_\lambda t \rceil}^{\infty} t|a_n| = \sum_{n\in\mathbb{Z}} \min(\lambda^n, t)|a_n|.$$

For the $\leq$, observe that we can write $a = b + c$, where $b_n = a_n$ if $n \leq \log_\lambda t$ and zero otherwise. Then putting $b \in X_0$ and $c \in X_1$ we clearly have the desired inequality. For the $\geq$, suppose $a = b + c$ with $b \in X_0$ and $c \in X_1$, then by the triangle inequality we have

$$\sum_{n=-\infty}^{\lfloor \log_\lambda t \rfloor} \lambda^n |a_n| + \sum_{n=\lceil \log_\lambda t \rceil}^{\infty} t|a_n| \leq$$
$$\sum_{n=-\infty}^{\lfloor \log_\lambda t \rfloor} \lambda^n |b_n| + \sum_{n=\lceil \log_\lambda t \rceil}^{\infty} t|b_n| + \sum_{n=-\infty}^{\lfloor \log_\lambda t \rfloor} \lambda^n |c_n| + \sum_{n=\lceil \log_\lambda t \rceil}^{\infty} t|c_n|$$
$$\leq \|b\|_{X_0} + t\|c\|_{X_1}$$

and the claim follows after taking infimums.

We can discretize K by taking comparisons: we have that for $t \in [\lambda^k, \lambda^{k+1}]$ that

$$\sum \lambda^{\min(n,k)} |a_n| \leq K(t, a; X_0, X_1) \leq \lambda \sum \lambda^{\min(n,k)} |a_n|.$$

So we can approximate

$$\int_0^\infty \left(t^{-\theta} K(t,a)\right)^q \frac{\mathrm{d}t}{t} \approx \sum_{k\in\mathbb{Z}} \lambda^{-kq\theta} \left(\sum_{n\in\mathbb{Z}} \lambda^{\min(n,k)} |a_n| \right)^q.$$

The inner sum we can rewrite to be

$$\sum_{n\in\mathbb{Z}} \lambda^{\min(n,k)} |a_n| = \lambda^k \sum_{n\in\mathbb{Z}} \min(\lambda^n, 1) |a_{n+k}|.$$

This gives us

$$\int_0^\infty \left(t^{-\theta} K(t,a)\right)^q \frac{\mathrm{d}t}{t} \approx \sum_{k\in\mathbb{Z}} \lambda^{kq(1-\theta)} \left(\sum_{n\in\mathbb{Z}} \lambda^{\min(n,0)} |a_{n+k}| \right)^q.$$

We claim that in fact

$$\int_0^\infty \left(t^{-\theta} K(t,a)\right)^q \frac{\mathrm{d}t}{t} \approx \sum_{k\in\mathbb{Z}} \lambda^{kq(1-\theta)} |a_k|^q. \tag{4.61}$$

The computations here are directly analogous to the computations in the proof of Lemma 4.55. The "convolution-like" structure is clearer, since the discretization effectively implements the required change of variables from a multiplicative group to an additive group.

For $\geq$ it suffices to note that $|a_k| \leq \sum_{n\in\mathbb{Z}} \lambda^{\min(n,0)} |a_{n+k}|$. For $\lesssim$, we use Minkowski's inequality to get

$$\left[\sum_{k\in\mathbb{Z}} \lambda^{kq(1-\theta)} \left(\sum_{n\in\mathbb{Z}} \lambda^{\min(n,0)} |a_{n+k}| \right)^q \right]^{\frac{1}{q}} \leq \sum_{n\in\mathbb{Z}} \lambda^{\min(n,0)} \left[\sum_{k\in\mathbb{Z}} \lambda^{kq(1-\theta)} |a_{n+k}|^q \right]^{\frac{1}{q}}.$$

Another change of variable in the inner sum gives us

$$= \sum_{n\in\mathbb{Z}} \lambda^{\min(n,0)} \lambda^{n(\theta-1)} \left[\sum_{k\in\mathbb{Z}} \lambda^{kq(1-\theta)} |a_k|^q\right]^{\frac{1}{q}}.$$

Noting that for $n < 0$ we have $\min(n,0) + n(\theta-1) = n\theta < 0$ and for $n > 0$ we have $\min(n,0) + n(\theta-1) = n(\theta-1) < 0$ we get that

$$\sum_{n\in\mathbb{Z}} \lambda^{\min(n,0)+n(\theta-1)}$$

converges for every $\theta \in (0,1)$. And our claim is proved.

Now, define the homogeneous sequence spaces $\mathring{\ell}^p_s$ for $p \in [1,\infty)$ and $s \in \mathbb{R}$ to be the space of sequences $a = (a_n)_{n\in\mathbb{Z}}$ such that

$$\|a\|_{\mathring{\ell}^p_s} \stackrel{\text{def}}{=} \left(\sum_{n\in\mathbb{Z}} 2^{spn} |a_n|^p\right)^{\frac{1}{p}} < \infty. \tag{4.62}$$

Our discussion above shows that

$$(\mathring{\ell}^1_0, \mathring{\ell}^1_s)_{\theta,q} = \mathring{\ell}^q_{s\theta}.$$

So by the Reiteration Lemma we conclude that

$$(\mathring{\ell}^{p_0}_{s_0}, \mathring{\ell}^{p_1}_{s_1})_{\theta,q} = \mathring{\ell}^q_s \tag{4.63}$$

where $s = (1-\theta)s_0 + \theta s_1$. Minor modifications of the above argument shows also that for the norm

$$\|a\|_{\mathring{\ell}^\infty_s} \stackrel{\text{def}}{=} \sup_{n\in\mathbb{Z}} 2^{sn} |a_n| \tag{4.64}$$

the equality of (4.63) can be extended for p_0, p_1, and possibly q taking ∞ as value. ¶

The above discussion on sequence spaces implies immediately that the homogeneous Besov spaces $\mathring{B}^{s,p}_q$ have the interpolation relationship

$$(\mathring{B}^{s_0,p}_{q_0}, \mathring{B}^{s_1,p}_{q_1})_{\theta,q} = \mathring{B}^{s,p}_q$$

whenever $s = (1-\theta)s_0 + \theta s_1$. It should be clear that instead of using sequences indexed by $\mathbb{Z}$, if we index by $\mathbb{N}$, the analogous sequence space estimates will also hold, thereby implying the same result for the inhomogeneous Besove spaces.

More refined computations (combining the results concerning Lebesgue and Lorentz spaces with that of sequence spaces, together with some other powerful technical results from general interpolation theory) gives us the following interpolation relationships between Sobolev and Besov spaces, which we state without proof for the inhomogeneous Sobolev and Besov spaces. We include also several embedding theorems (including Sobolev embedding). The analogous interpolation results are also true for the homogeneous versions.

See Chapter 6 of Bergh and Löfström, Interpolation spaces. An introduction *for proofs of the Sobolev and Besov space interpolation results.*

4.65 Theorem (Relation between Besov and Sobolev spaces) Throughout, $s, s_0, s_1 \in \mathbb{R}$ and $p, p_0, p_1, q, q_0, q_1 \in [1, \infty]$ unless otherwise specified. θ is always in $(0,1)$. When used, the notations s^*, p^*, q^* denote the values

$$s^* = (1-\theta)s_0 + \theta s_1, \quad \frac{1}{p^*} = \frac{1-\theta}{p_0} + \frac{\theta}{p_1}, \quad \text{and } \frac{1}{q^*} = \frac{1-\theta}{q_0} + \frac{\theta}{q_1}.$$

All function spaces are defined over $\mathbb{R}^d$.

- $B^{s,p}_{q_0} \subset B^{s,p}_{q_1}$ if $q_0 < q_1$.
- $B^{s,p}_{p} \subset W^{s,p} \subset B^{s,p}_{2}$ if $1 < p \leq 2$.
- $B^{s,p}_{2} \subset W^{s,p} \subset B^{s,p}_{p}$ if $2 \leq p < \infty$.
- $B^{s_0,p_0}_{q_0} \subset B^{s_1,p_1}_{q_1}$ provided that $p_0 \leq p_1$, $q_0 \leq q_1$, and $s_0 - \frac{d}{p_0} = s_1 - \frac{d}{p_1}$.
- $(W^{s_0,p}, W^{s_1,p})_{\theta,q} = B^{s^*,p}_{q}$ if $s_0 \neq s_1$.
- $(W^{s,p_0}, W^{s,p_1})_{\theta,p^*} = W^{s,p^*}$.
- $(W^{s_0,p_0}, W^{s_1,p_1})_{[\theta]} = W^{s^*,p^*}$ if $s_0 \neq s_1$ and $p_0, p_1 \in (1,\infty)$. (Note that this one uses *complex* interpolation.)
- $(B^{s_0,p}_{q_0}, B^{s_1,p}_{q_1})_{\theta,q} = B^{s^*,p}_{q}$ if $s_0 \neq s_1$.
- $(B^{s,p}_{q_0}, B^{s,p}_{q_1})_{\theta,q^*} = B^{s,p}_{q^*}$.
- $(B^{s_0,p_0}_{q_0}, B^{s_1,p_1}_{q_1})_{\theta,p^*} = B^{s^*,p^*}_{q^*}$ provided $s_0 \neq s_1$ and $p^* = q^*$.

- $(B^{s_0,p_0}_{q_0}, B^{s_1,p_1}_{q_1})_{[\theta]} = B^{s^*,p^*}_{q^*}$ provided $s_0 \neq s_1$. (Note that this one uses *complex* interpolation.) ■

4.66 Exercise (L^p decay of wave)
Consider the solution to the linear wave equation

$$-\partial^2_{tt}\phi + \triangle\phi = 0$$

with initial data $\phi(0,x) = \phi_0(x)$ and$\partial_t\phi(0,x) = 0$. Recall from Theorem 3.46 that

Ref. 3.46: "Summary of decay of solutions to the wave equation"

$$\left|\phi(t,x)\right| \leq C|t|^{-\frac{d-1}{2}}\left\|\phi_0\right\|_{\mathring{B}_1^{\frac{d+1}{2},1}}.$$

1. Using the fundamental solution $G_t^{(\text{wave})}$ (refer also to Exercise 2.49) and Plancherel's identity (Proposition 2.23), prove that

$$\left\|\phi(t,\bullet)\right\|_{L^2} \leq \left\|\phi_0\right\|_{L^2}.$$

2. Using Exercise 3.45 and Theorem 4.65, together with the previous part, prove that for any $p \in (1,2)$, there exists a constant C_p such that

$$\left\|\phi(t,\bullet)\right\|_{L^{p'}} \leq C_p|t|^{-(d-1)\left(\frac{1}{p}-\frac{1}{2}\right)}\left\|\phi_0\right\|_{\mathring{B}_p^{(d+1)\left(\frac{1}{p}-\frac{1}{2}\right),p}}$$

where p' is given by $(p')^{-1} + p^{-1} = 1$. ■

Multilinear interpolation and Hardy-Littlewood-Sobolev

The two multi-linear theorems can can be found as Theorem 4.4.1 and Exercise 3.13.5b in Bergh and Löfström, Interpolation spaces. An introduction *respectively. For the latter theorem, we will only be using cases where the Banach spaces involved are weighted L^p spaces and Lorentz spaces, and in this case a complete proof is given in O'Neil, "Convolution operators and $L(p,q)$ spaces".*

So far the theory covered have been only applicable to linear and sublinear mappings. The same techniques can be extended to *multilinear operators*, as well as interpolation with more than 2 endpoints. We quote without proof two results in this direction; the first using complex interpolation, and the second using real interpolation.

4.67 Theorem (Multilinear, two-endpoint interpolation)
Let $A_0, A_1, B_0, B_1, C_0, C_1$ be Banach spaces, and suppose that T is a *bilinear mapping* satisfying the bounds

$$\|T(a,b)\|_{C_0} \leq M_0\|a\|_{A_0}\|b\|_{B_0} \qquad (T : A_0 \times B_0 \to C_0)$$
$$\|T(a,b)\|_{C_1} \leq M_1\|a\|_{A_1}\|b\|_{B_1} \qquad (T : A_1 \times B_1 \to C_1)$$

hold, then for any $\theta \in [0,1]$,

$$\|T(a,b)\|_{C_{[\theta]}} \leq M_0^{(1-\theta)} M_1^{\theta} \|a\|_{A_{[\theta]}} \|b\|_{B_{[\theta]}} \qquad (T : A_{[\theta]} \times B_{[\theta]} \to C_{[\theta]}). \qquad \blacksquare$$

4.68 Theorem (Multilinear, three-endpoint interpolation)
Let $A_0, A_1, B_0, B_1, C_0, C_1$ be Banach spaces, and suppose that T is a *bilinear mapping* such that the bounds

$$\begin{aligned} \|T(a,b)\|_{C_0} &\lesssim \|a\|_{A_0} \|b\|_{B_0} & \qquad (T : A_0 \times B_0 \to C_0) \\ \|T(a,b)\|_{C_1} &\lesssim \|a\|_{A_0} \|b\|_{B_1} & \qquad (T : A_0 \times B_1 \to C_1) \\ \|T(a,b)\|_{C_1} &\lesssim \|a\|_{A_1} \|b\|_{B_0} & \qquad (T : A_1 \times B_0 \to C_1) \end{aligned}$$

hold, then for $\theta, \theta_A, \theta_B \in (0,1)$, and $p, q, r \in [1,\infty]$ satisfying

$$\theta = \theta_A + \theta_B, \quad 1 \leq \frac{1}{p} + \frac{1}{q},$$

we have the estimate

$$\|T(a,b)\|_{C_{\theta,r}} \lesssim \|a\|_{A_{\theta_A,pr}} \|b\|_{B_{\theta_B,qr}} \qquad (T : A_{\theta_A,pr} \times B_{\theta_B,qr} \to C_{\theta,r}). \qquad \blacksquare$$

Let's look at some of the consequence of the second theorem. A first consequence of this theorem is the Hölder inequality for Lorentz spaces. Letting $A_0 = B_0 = C_0 = L^\infty$ and $A_1 = B_1 = C_1 = L^1$ in the above theorem, we see that the mapping $(f,g) \mapsto fg$ is indeed from $L^1 \times L^\infty \to L^1$ and $L^\infty \times L^\infty \to L^\infty$. The above theorem implies, together with our Definition 4.45, that the multiplication operation also satisfies

$$\|fg\|_{L^p_q} \lesssim \|f\|_{L^{p_0}_{q_0}} \|g\|_{L^{p_1}_{q_1}} \tag{4.69}$$

whenever

$$\frac{1}{p} = \frac{1}{p_0} + \frac{1}{p_1}, \quad \frac{1}{q} \leq \frac{1}{q_0} + \frac{1}{q_1}.$$

A second consequence is a remarkably simple proof of the Hardy-Littlewood-Sobolev fractional integration estimate. This same estimate is frequently proven in textbooks using maximal functions estimates, which uses the geometric structure of Euclidean spaces through the Vitali covering lemma. Here we can obtain the estimate as a real interpolation version of Young's inequality Corollary 4.27.

The usual proof of the Hardy-Littlewood-Sobolev estimate can be found in Chapter 1, Section 3 of Stein, Harmonic analysis: real-variable methods, orthogonality, and oscillatory integrals.

4.70 Theorem (Hardy-Littlewood-Sobolev)
Given $f \in L^{p_0}_{q_0}$ and $g \in L^{p_1}_{q_1}$, such that $(q_0)^{-1} + (q_1)^{-1} = q^{-1} < 1$, and $p^{-1} = (p_0)^{-1} + (p_1)^{-1} - 1$ satisfies $p, p_0, p_1 \in (1, \infty)$, we have

$$\|f * g\|_{L^p_q} \lesssim \|f\|_{L^{p_0}_{q_0}} \|g\|_{L^{p_1}_{q_1}}.$$ ■

Proof Observe that the convolution mapping $(f, g) \mapsto f * g$ is symmetric in f and g and satisfies trivially

$$\|f * g\|_{L^1} \leq \|f\|_{L^1} \|g\|_{L^1}, \quad \|f * g\|_{L^\infty} \leq \|f\|_{L^1} \|g\|_{L^\infty}.$$

So we can apply the multilinear interpolation theorem with $A_0 = B_0 = C_0 = L^1$ and $A_1 = B_1 = C_1 = L^\infty$ with $\theta = 1 - 1/p$ to get the desired conclusion. □

Now, consider the function $k_\gamma(x) = |x|^{-\gamma}$ for $\gamma \in (0, d)$. The function just barely fails to be in $L^{d/\gamma}$, but it is in fact in $L^{d/\gamma}_\infty$ (in other words in weak-$L^{d/\gamma}$). The Hardy-Littlewood-Sobolev inequality immediately implies

$$\left\| k_\gamma * f \right\|_{L^p_q} \lesssim \|f\|_{L^r_q} \tag{4.71}$$

where $q \in (1, \infty]$ and

$$\frac{1}{p} = \frac{1}{r} + \frac{\gamma}{d} - 1 > 0.$$

Noticing that since $\frac{\gamma}{d} - 1 < 0$, necessarily $p > r$ and so we have

$$\left\| k_\gamma * f \right\|_{L^p} = \left\| k_\gamma * f \right\|_{L^p_p} \lesssim \left\| k_\gamma * f \right\|_{L^p_r} \lesssim \|f\|_{L^r_r} = \|f\|_{L^r} \tag{4.72}$$

which is the classical statement of the fractional integration inequality of Hardy-Littlewood-Sobolev.

4.73 Exercise (Hardy-Littlewood-Sobolev: lower bound is sharp)
In the statement (4.72), which follows from Theorem 4.70, the restriction on p_0 in the hypothesis of the theorem translates to $\gamma \in (0, d)$. The restriction on p_1 requies $r \in (1, \infty)$. Show that the inequality cannot be extended to $r = 1$. That is to say, find a function $f \in L^1$ such that $k_\gamma * f \notin L^{d/\gamma}$. ■

4.74 Remark (Fractional integration)
The operation of convolving a function against k_γ is called a *fractional integration*. Observe first that in the case $d > 2$ and $\gamma = d - 2$, the function k_γ is proportional to the Newton potential, so that up to a constant $\Delta(k_{d-2} * f) =$

cf. In other words, one can think of convolving by k_{d-2} as the same as acting by the inverse of the Laplacian.

More generally, for $\gamma \in (0, d)$, we have that k_γ is a locally integrable function, and so for any function $f \in \mathscr{S}$ the integral $\int_{\mathbb{R}^d} k_\gamma(x) f(x)\, dx$ is well-defined and k_γ in fact represents an element of $\mathscr{S}'$. In this case, what is its Fourier transform? Going back to Proposition 2.7 we see that $k_\gamma(x)$ is homogeneous, and so $S_\lambda k_\gamma(x) = \lambda^{d/2}|\lambda x|^{-\gamma} = \lambda^{d/2-\gamma} k_\gamma(x)$. And so we expect

Ref. 2.7: "Fourier transform properties: scaling, translation, modulation"

$$\lambda^{-d/2}\widehat{k_\gamma}(\lambda^{-1}x) = \lambda^{d/2-\gamma}\widehat{k_\gamma}(x).$$

For this to hold we must have $\widehat{k_\gamma}(x) \propto |x|^{\gamma-d}$. For $\gamma \in (0, d)$ this can in fact be justified rigorously, and therefore we have that

$$\mathscr{F}[k_\gamma * f](\xi) \propto \widehat{k_\gamma}(\xi)\widehat{f}(\xi) \propto |x|^{\gamma-d}\widehat{f}(\xi),$$

so k_γ corresponds to a Fourier multiplier that behaves like the inverse of taking $d - \gamma$ derivatives, and hence they are called fractional integration operators. ∎

Strichartz estimates

We now turn to one of the most useful results in the study of dispersive equations. Returning to the statements in Exercise 4.29 and Exercise 4.66, we see that measuring the solution at a fixed time in L^q for $q \in [2, \infty)$, we get different decay rates depending on q. The decay rates are however homogeneous: that they are of the power type $|t|^{-\gamma}$. Now, functions of the form $|t|^{-\gamma}$ are *not* in $L^p(\mathbb{R})$ for any p; the closest we get is that for $p = \gamma^{-1}$ the function is almost in L^p, in fact, in weak-L^p. Recalling however that the L^q decay estimates are obtained by interpolation from the conservation of mass and the L^1–L^∞ decay estimates, one can ask whether we can do the interpolation better, in view of weak-type interpolation results such as Corollary 4.57, to get results where the decay in t is measured in an integral (L^p) norm, rather than a pointwise norm. This turns out to be possible, through combining the pointwise-in-time decay results with the Hardy-Littlewood-Sobolev inequality, together with an abstract functional analytic technique called the TT^* method.

For further developments of the ideas presented in this section, as well as for a different exposition, please refer to Chapter 8 of Bahouri et al., Fourier analysis and nonlinear partial differential equations.

4.75 (The TT^* method) Let H denote a Hilbert space, and Y some reflexive Banach space. Suppose we are given a bounded linear operator

$$T : H \to Y.$$

Its adjoint (transpose) operator is then the bounded linear operator

$$T^* : Y^* \to H^* = H$$

defined by $\langle T^* y, h\rangle = \langle y, Th\rangle$; we used that Hilbert spaces are their own duals, and abuse the notation $\langle , \rangle$ to denote both the duality pairing and also the inner product for the Hilbert space. This means that

$$TT^* : Y^* \to Y$$

is again a bounded linear operator. We can compute the norms of these three operators.

The definition of operator norm is given by

$$\|T\|_{H\to Y} = \sup_{\|h\|_H = 1} \|Th\|_Y.$$

Putting in the dual characterisation of the Y norm we get

$$\|T\|_{H\to Y} = \sup_{\|h\|_H=1} \sup_{\|y\|_{Y^*}=1} \langle y, Th\rangle = \sup_{\|h\|_H=1} \sup_{\|y\|_{Y^*}=1} \langle T^* y, h\rangle = \|T^*\|_{Y^*\to H},$$

showing the well-known result that the operator norm of an operator and that of its adjoint are the same. By taking the composition we have that $\|TT^*\|_{Y^*\to Y} \leq \|T\|^2_{H\to Y}$. On the other hand,

$$\begin{aligned} \|T\|_{Y^*\to Y} = \sup_{\|y\|_{Y^*}=1=\|z\|_{Y^*}} \langle z, TT^* y\rangle &= \sup_{\|y\|_{Y^*}=1=\|z\|_{Y^*}} \langle T^* z, T^* y\rangle \\ &\geq \sup_{\|y\|_{Y^*}=1} \langle T^* y, T^* y\rangle = \|T^*\|^2_{Y^*\to H}. \end{aligned}$$

So we have shown that in fact, $\|TT^*\|_{Y^*\to Y} = \|T\|^2_{H\to Y}$.

The basic idea of the TT^* method is this: in many situations, we are given an operator T that acts on some Hilbert space H. In this situation, in order to show that it maps H boundedly into a reflexive Banach space Y, it suffices to study the operator TT^* and show that it acts boundedly from $Y^* \to Y$. Moreover, this can be reduced to proving the *bilinear estimate*

$$\langle z, TT^* y\rangle \lesssim \|z\|_{Y^*} \|y\|_{Y^*}.$$ ¶

In what follows we will assume that our Hilbert space $H = L^2(\mathbb{R}^d)$, and all duality pairings will be with respect to the L^2 pairing. Given a Banach space X of functions on $\mathbb{R}^d$ in which $\mathscr{S}(\mathbb{R}^d)$ is dense, we denote by the space L^pX (for $1 \leq p < \infty$) the space of functions on $\mathbb{R}^{d+1}$ given by the closure of $\mathscr{S}(\mathbb{R}^{d+1})$ with respect to the norm

$$\|\Phi\|_{L^pX} \stackrel{\text{def}}{=} \left[\int_{\mathbb{R}} \|\Phi(t,\bullet)\|_X^p \, \mathrm{d}t\right]^{\frac{1}{p}}. \tag{4.76}$$

We state an abstract Strichartz-type estimate.

4.77 Theorem ("Abstract Strichartz")
Let $U(t)$ for $t \in \mathbb{R}$ denote an uniformly bounded one parameter family of linear operators mapping $L^2 \to L^2$. Let X be a reflexive Banach space of functions on $\mathbb{R}^d$, such that $\mathscr{S}(\mathbb{R}^d)$ is dense in both X and X^*. Suppose there exists $\gamma \in (0,1)$ such that for every $f \in \mathscr{S}(\mathbb{R}^d)$, and every $t, t' \in \mathbb{R}$, the inequality

$$\left\|U(t)U^*(t')f\right\|_{X^*} \lesssim \frac{C}{|t-t'|^\gamma}\|f\|_X$$

holds. Then

$$\|U(t)f\|_{L^{\frac{2}{\gamma}}X^*} \lesssim \|f\|_{L^2}$$

$$\left\|\int_{\mathbb{R}} U^*(t)\Phi(t,\bullet)\,\mathrm{d}t\right\|_{L^2} \lesssim \|\Phi\|_{L^{\frac{2}{2-\gamma}}X} \qquad \blacksquare$$

Proof As usual it suffices to prove things using Schwartz functions. We begin by noting that

$$\Big|\iint_{\mathbb{R}^2} \langle U(t)U^*(t')\Psi(t'),\Phi(t)\rangle \,\mathrm{d}t\,\mathrm{d}t'\Big| \lesssim \iint_{\mathbb{R}^2} \frac{1}{|t-t'|^\gamma}\left\|\Psi(t')\right\|_X \|\Phi(t)\|_X \,\mathrm{d}t\,\mathrm{d}t'.$$

We wish to apply (4.72) on fractional integration. To do so we need to use the case where p and r in the inequality are conjugate exponents. That is to say, we need both $p^{-1}+r^{-1}=1$ with $p>r$ and also $p^{-1}=r^{-1}+\gamma-1$ from Hardy-Littlewood-Sobolev. This means that $p=2/\gamma$ and $r=2/(2-\gamma)$. With these values we have the bilinear estimate

$$\Big|\iint_{\mathbb{R}^2} \langle U(t)U^*(t')\Psi(t'),\Phi(t)\rangle \,\mathrm{d}t\,\mathrm{d}t'\Big| \lesssim \|\Psi\|_{L^{\frac{2}{2-\gamma}}X}\|\Phi\|_{L^{\frac{2}{2-\gamma}}X}. \tag{4.78}$$

Now, take the operator T to be the mapping

$$(Tf)(t,x) = [U(t)f](x)$$

and T^* to be

$$(T^*\Phi)(x) = \int_{\mathbb{R}} U^*(t)\Phi(t,x)\,\mathrm{d}t.$$

The estimate (4.78) reads that for every $\Phi, \Psi \in L^{2/(2-\gamma)}X \cap \mathscr{S}(\mathbb{R}^{d+1})$, the inequality $\langle TT^*\Psi, \Phi\rangle \lesssim \|\Psi\|_{L^{2/(2-\gamma)}X}\|\Phi\|_{L^{2/(2-\gamma)}X}$ holds. By density it holds for all such functions, and setting $\Psi = \Phi$ we recover, through the TT^* method that $T^* : L^{2/(2-\gamma)}X \to L^2(\mathbb{R}^d)$ is bounded, and so T is bounded from L^2 to $L^{2/\gamma}X^*$. □

4.79 Remark

Note that the proof of (4.78) is based on essentially a pointwise comparison. And hence we also have the following version: let $\varpi : \mathbb{R}^2 \to \mathbb{C}$ be any function that satisfies $|\varpi| \leq 1$, then

$$\Big|\iint_{\mathbb{R}^2} \varpi(t',t)\langle U(t)U^*(t')\Psi(t'), \Phi(t)\rangle\,\mathrm{d}t\,\mathrm{d}t'\Big| \lesssim \|\Psi\|_{L^{\frac{2}{2-\gamma}}X}\|\Phi\|_{L^{\frac{2}{2-\gamma}}X}. \tag{4.80}$$

Taking, for example, $\varpi(t',t) = 1$ when $t' \in [0,t]$ and 0 otherwise, we conclude from this

$$\Big\|\int_0^t U(t)U^*(t')\Psi(t')\,\mathrm{d}t'\Big\|_{L^{\frac{2}{\gamma}}X^*} \lesssim \|\Psi\|_{L^{\frac{2}{2-\gamma}}X}, \tag{4.81}$$

an inequality that can be useful for applications to inhomogeneous and nonlinear problems. ■

4.82 Corollary (Strichartz for Schrödinger)

Let $\phi(t,x)$ solve Schrödinger's equation with initial data $\phi_0(x)$. Then

$$\big\|\phi\big\|_{L_t^{\frac{4p}{d(p-2)}}L_x^p} \lesssim \big\|\phi_0\big\|_{L^2}$$

for any $p \in (2, \frac{2d}{d-2})$. ■

Proof By Exercise 4.29, we have that

$$\big\|\phi(t,\bullet)\big\|_{L^p} \leq C|t|^{-\frac{d}{2}(1-\frac{2}{p})}\big\|\phi_0\big\|_{L^{p'}}.$$

So we can take in Theorem 4.77 that $X = L^{p'}$ and $\gamma = \frac{d}{2}(1 - \frac{2}{p})$. This implies the result immediately. □

The application to Schrödinger equation is rather straightforward, as the natural dispersive estimate already interpolates to statements about mappings from L^p to $L^{p'}$. For the wave equation the situation is however more complicated: we need to rephrase our dispersive estimate in a more suitable form than that of Theorem 3.46. Let $\phi_0, \phi_1 \in \mathscr{S}$, and let $\phi(t,x)$ solve the linear wave equation with initial data $\phi(0,x) = \phi_0(x)$ and $\partial_t \phi(0,x) = \phi_1(x)$. Observe that in the statement of Theorem 4.77, we need to get some control on $U(t)$ as mapping from some space X to its dual. The initial value problem for the wave equation however has two initial data points, the initial position ϕ_0 and the initial velocity ϕ_1. So to make it amenable to our abstract theorem, instead of considering simply the solution operator $(\phi_0, \phi_1) \mapsto \phi$, it is better to consider the mapping of pairs

Ref. 3.46: "Summary of decay of solutions to the wave equation"

$$U(t) : (\phi_0, \phi_1) \mapsto (\phi(t), \partial_t \phi(t)). \tag{4.83}$$

We get our estimate for the operator $U(t)$ from Remark 3.43.

Ref. 3.43: "Decay estimates for Littlewood-Paley projected wave kernel"

Noting that the Littlewood-Paley projectors commute with the fundamental solution $G_t^{(\text{wave})}$, we have that the function $\Delta_k \phi(t,x)$ solves the linear wave equation with initial data $\Delta_k \phi_0$ and $\Delta_k \phi_1$. Therefore we have that

$$\left\|\Delta_k \phi(t,\bullet)\right\|_{L^\infty} \lesssim 2^{k(d+1)/2} |t|^{-(d-1)/2} \left(\left\|\Delta_k \phi_0\right\|_{L^1} + 2^{-k} \left\|\Delta_k \phi_1\right\|_{L^1} \right).$$

Conservation of energy gives however

$$\left\|\Delta_k \phi(t,\bullet)\right\|_{L^2} \approx \left\|\Delta_k \phi_0\right\|_{L^2} + 2^{-k} \left\|\Delta_k \phi_1\right\|_{L^2}.$$

Similarly, for $\partial_t \phi$, we find

$$\begin{aligned} \left\|\Delta_k \partial_t \phi(t,\bullet)\right\|_{L^\infty} &\lesssim 2^{k(d+1)/2} |t|^{-(d-1)/2} \left(2^k \left\|\Delta_k \phi_0\right\|_{L^1} + \left\|\Delta_k \phi_1\right\|_{L^1} \right), \\ \left\|\Delta_k \partial_t \phi(t,\bullet)\right\|_{L^2} &\approx 2^k \left\|\Delta_k \phi_0\right\|_{L^2} + \left\|\Delta_k \phi_1\right\|_{L^2}. \end{aligned}$$

Complex interpolation between the two sets gives that with $\theta = 2/p$ and $1/p' + 1/p = 1$ we have

Maybe at this point I should introduce Bernstein's inequality and use this to get the L^p instead of the Besov versions?

$$\begin{aligned} &\left\|\Delta_k \phi(t,\bullet)\right\|_{L^p} + 2^{-k} \left\|\Delta_k \partial_t \phi(t,\bullet)\right\|_{L^p} \lesssim \\ &\qquad 2^{\frac{k}{2}(d+1)(1-\theta)} |t|^{-(d-1)(1-\theta)/2} \left(\left\|\Delta_k \phi_0\right\|_{L^{p'}} + 2^{-k} \left\|\Delta_k \phi_1\right\|_{L^{p'}} \right). \end{aligned}$$

Therefore we have that for

$$\gamma = (d-1)(\frac{1}{2}-\frac{1}{p}), \quad \frac{1}{p}+\frac{1}{p'}=1 \quad s = \frac{d+1}{2}\left(\frac{1}{2}-\frac{1}{p}\right), \tag{4.84}$$

we have

$$2^{-ks}\left\|\Delta_k\left(\phi(t,\bullet),|\nabla|^{-1}\partial_t\phi(t,\bullet)\right)\right\|_{L^p} \lesssim 2^{ks}|t|^{-\gamma}\left\|\Delta_k(\phi_0,|\nabla|^{-1}\phi_1)\right\|_{L^{p'}}.$$

To simplify notation, we can write $\Phi_0 = (\phi_0, |\nabla|^{-1}\phi_1)$ and $\Phi = (\phi, |\nabla|^{-1}\partial_t\phi)$. Our decay estimate can be written then as

In particular, Bernstein says that for the $2^{-ks}\|\Delta_k\Phi\|_p \lesssim \|\Delta_k\Phi\|_q$ for $s = d(1/q-1/p)$. This actually gives the correct proof for the Exercise below; and this gives the correct range for the Strichartz estimates for wave; the corollary here is a bit off.

$$\|\Phi(t,\bullet)\|_{\mathring{B}^{-s,p}_q} \lesssim |t|^{-\gamma}\|\Phi_0\|_{\mathring{B}^{s,p'}_q} \tag{4.85}$$

for every $q \in [1,\infty]$. Now we are ready to state the Strichartz estimate for wave equations.

4.86 Corollary (Strichartz for wave)
Let s and γ be defined as in (4.84), such that $\gamma \in (0,1)$ (in other words, $p \in (2, \frac{2(d-1)}{d-3})$ for $d > 3$ and $p \in (2,\infty)$ for $d = 2,3$), then

$$\left\|\phi\right\|_{L^{\frac{2}{\gamma}}\mathring{B}^{-s,p}_2} + \left\|\partial_t\phi\right\|_{L^{\frac{2}{\gamma}}\mathring{B}^{-s-1,p}_2} \lesssim \left\|\phi_0\right\|_{\mathring{B}^{0,2}_2} + \left\|\phi_1\right\|_{\mathring{B}^{-1,2}_2}.$$

■

Proof We apply the abstract Strichartz estimate Theorem 4.77 with $X = \mathring{B}^{s,p'}_2$. We use the fact that for $p,q \in (1,\infty)$ we have that the dual space fo $\mathring{B}^{s,p'}_q$ is $\mathring{B}^{-s,p}_{q'}$ where $1/p + 1/p' = 1/q + 1/q' = 1$. The theorem then follows from (4.85) after noting that

$$\|\Phi_0\|_{L^2} = \|\Phi_0\|_{\mathring{B}^{0,2}_2} \approx \left\|\phi_0\right\|_{\mathring{B}^{0,2}_2} + \left\|\phi_1\right\|_{\mathring{B}^{-1,2}_2}$$

by definition of Φ_0. □

4.87 Exercise

Check Bergh and Lofstrom to see if the homogeneous versions of the Sobolev theorems actually incorporates the Bernstein versions.

1. Combining Corollary 4.86 and Theorem 4.65, we can pick up as a consequence
$$\left\|\phi\right\|_{L^rL^q} \lesssim \left\|\phi_0\right\|_{\mathring{B}^{0,2}_2} + \left\|\phi_1\right\|_{\mathring{B}^{-1,2}_2};$$
determine the range of allowable (r,q) in the above inequality based on (4.84).

2. Do the same as the previous part, except instead of bounding ϕ in $L^r L^q$, bound $\partial_t \phi$ in $L^r L^q$. ■

In the remainder of this chapter we will discuss the so-called *end-point* Strichartz estimates. Going back to the discussion of the abstract Strichartz Theorem 4.77, we see that we only allow $\gamma \in (0,1)$. We ask: can there be analogous estimates for $\gamma = 0$ or $\gamma = 1$?

The case $\gamma = 0$ is simple: looking at, for example, the case of Schrödinger equation, we see that plugging in formally $\gamma = 0$ will require $p = 2$, and then the estimate is nothing more than the conservation of mass $\|\phi(t,\bullet)\|_{L^2} \leq \|\phi_0\|_{L^2}$ we already know about the solution. More generally, we expect the case of a uniform bound in t to correspond to some sort of L^2 based conservation laws.

What about the case $\gamma = 1$? In fact frequently one can prove exactly such an estimate. The proof relies on a clever use of both of the multilinear interpolation theorems given in the previous section. We demonstrate the case for Schrödinger equations, but similar results are also available for other dispersive equations.

4.88 Theorem (End-point Strichartz for Schrödinger)
Let $d > 2$. Let $\phi(t,x)$ solve Schrödinger's equation with initial data $\phi_0(x)$. Then

$$\|\phi\|_{L^2_t L_x^{\frac{2d}{d-2}}} \lesssim \|\phi_0\|_{L^2}.$$

■

Proof Denote by $U(t)$ the solution operator to Schrödinger equation. Let $\mu = (\mu_j)_{j\in\mathbb{Z}}$ be the sequence defined by

$$\mu_j = \iint\limits_{|t-t'|\in[2^j,2^{j+1})} \langle U(t)U^*(t')\Psi(t'), \Phi(t)\rangle \, \mathrm{d}t \, \mathrm{d}t'.$$

Let's estimate μ_j in two different ways.

First, by the uniform decay estimate in Corollary 3.29, we have

$$\langle U(t)U^*(t')\Psi(t'), \Phi(t)\rangle \lesssim \left|t-t'\right|^{-d/2} \left\|\Psi(t')\right\|_{L^1} \|\Phi(t)\|_{L^1}.$$

And so by Cauchy-Schwarz we get

$$\left|\mu_j\right| \lesssim 2^{j(1-\frac{d}{2})} \|\Phi\|_{L^2_t L^1} \|\Psi\|_{L^2_t L^1}. \tag{4.89}$$

Similarly, from the conservation of energy we have that

$$\langle U(t)U^*(t')\Psi(t'), \Phi(t)\rangle \leq \left\|\Psi(t')\right\|_{L^2} \|\Phi(t)\|_{L^2}$$

and so

$$\left|\mu_j\right| \lesssim 2^j \|\Phi\|_{L^2_t L^2} \|\Psi\|_{L^2_t L^2}. \tag{4.90}$$

Next, rewrite the integration for μ_j, bringing in the integral inside the bilinear pairing to get

$$\begin{aligned}\mu_j &= \iint_{\mathbb{R}^2} \chi_j(t-t')\langle U(t)U^*(t')\Psi(t'), \Phi(t)\rangle \, \mathrm{d}t \, \mathrm{d}t' \\ &= \int_{\mathbb{R}} \left\langle U^*(t')\Psi(t'), \int_{\mathbb{R}} U^*(t)[\chi_j(t-t')\Phi(t)] \, \mathrm{d}t \right\rangle \mathrm{d}t'\end{aligned}$$

where $\chi_j(s)$ is equal to 1 when $|s| \in [2^j, 2^{j+1})$ and 0 otherwise. Now evaluate the pairing as the L^2-L^2 pairing, and we have

$$\left|\mu_j\right| \leq \int_{\mathbb{R}} \left\|U^*(t')\Psi(t')\right\|_{L^2} \left\| \int_{\mathbb{R}} U^*(t)[\chi_j(t-t')\Phi(t)] \, \mathrm{d}t \right\|_{L^2} \mathrm{d}t'.$$

Now, noting that $U^*(t')$ acts boundedly on L^2, we have that $\|U^*(t')\Psi(t')\|_{L^2} \lesssim \|\Psi(t')\|_{L^2}$ (the constant is in fact 1 in the case of Schrödinger). So we can perform a Cauchy-Schwarz to get

$$\left|\mu_j\right| \leq \|\Psi\|_{L^2_t L^2} \cdot \left\| \int_{\mathbb{R}} U^*(t)[\chi_j(t-\bullet)\Phi(t)] \, \mathrm{d}t \right\|_{L^2_{t'} L^2}.$$

Let us examine the second term on the right. By Theorem 4.77, we have for any $p \in (2, \frac{2d}{d-2})$ the estimate

$$\left\| \int_{\mathbb{R}} U^*(t) g(t) \, \mathrm{d}t \right\|_{L^2} \lesssim \|g\|_{L^{2/(2-\gamma)} L^{p'}}$$

where $\gamma = d(\frac{1}{2} - \frac{1}{p}) \in (0,1)$. In particular, $2/(2-\gamma) \in (1,2)$. Applying this to our expression we note that we should set $g(t) = \chi(t-t')\Phi(t)$ which has compact support in t. And hence we can apply Hölder's inequality to get

$$\left\| \int_{\mathbb{R}} U^*(t)\chi(t-t')\Phi(t)\,\mathrm{d}t \right\|_{L^2} \lesssim \left\|\chi(t-t')\right\|_{L_t^{2/(1-\gamma)}} \left\|\chi(t'-\bullet)\Phi\right\|_{L_t^2 L^{p'}}$$
$$\lesssim 2^{j(1-\gamma)/2} \left(\int_{\mathbb{R}} \chi(s) \left\|\Phi(s+t')\right\|^2_{L^{p'}}\,\mathrm{d}s \right)^{\frac{1}{2}}. \qquad (4.91)$$

Integrating this quantity squred in t' gives

$$\left\| \int_{\mathbb{R}} U^*(t)[\chi_j(t-\bullet)\Phi(t)]\,\mathrm{d}t \right\|^2_{L^2_{t'}L^2} \lesssim 2^{j(1-\gamma)} \iint_{\mathbb{R}^2} \chi(s)\left\|\Phi(s+t')\right\|^2_{L^{p'}}\,\mathrm{d}s\,\mathrm{d}t'.$$

Performing the integral in t' first (by Fubini) we get finally

$$\left\| \int_{\mathbb{R}} U^*(t)[\chi_j(t-\bullet)\Phi(t)]\,\mathrm{d}t \right\|_{L^2_{t'}L^2} \lesssim 2^{j(2-\gamma)/2}\|\Phi\|_{L^2_t L^{p'}}.$$

And we finally conclude that

$$\left|\mu_j\right| \lesssim 2^{j(2-\gamma)/2}\|\Psi\|_{L^2_t L^2}\|\Phi\|_{L^2_t L^{p'}}. \qquad (4.92)$$

Quite clearly the same estimate holds with the roles of Ψ and Φ swapped.

Summarizing (4.89), (4.90), and (4.92), we have that the mapping

$$(\Psi,\Phi) \mapsto \mu_j$$

maps to $\mathbb{R}$ from

- $L^2_t L^1 \times L^2_t L^1$ with norm $2^{j(1-\frac{d}{2})}$;
- $L^2_t L^2 \times L^2_t L^q$, where $q \in (\frac{2d}{d+2}, 2]$, with norm $2^{j(1+\frac{d}{4}-\frac{d}{2q})}$;
- similarly from $L^2_t L^q \times L^2_t L^2$.

Interpolating using Theorem 4.67 we get that (when $d > 2$) for any $q \in (\frac{2d}{d+2}, 2]$, we also have that the mapping maps

$$L^2_t L^{p_0} \times L^2_t L^{p_1} \to \mathbb{R} \text{ with norm } 2^{j\beta}$$

where

$$p_0 = \frac{1}{1-\frac{\theta}{2}}, \quad p_1 = \frac{1}{1-\theta+\frac{\theta}{q}}$$

and

$$\beta = 1 - \frac{d}{2} + \theta\frac{d}{4} + \theta\frac{d}{2} - \theta\frac{d}{2q} = 1 + \frac{d}{2}\left(1 - \frac{1}{p_0} - \frac{1}{p_1}\right).$$

In particular, we see that there exists some $\epsilon > 0$ such that for every

$$(p_0)^{-1}, (p_1)^{-1} \in \left[\frac{d+2}{2d} - \epsilon, \frac{d+2}{2d} + \epsilon\right] \tag{4.93}$$

we have the estimate

$$\left|\mu_j\right| \lesssim 2^{j\beta(p_0,p_1)} \|\Psi\|_{L_t^2 L^{p_0}} \|\Phi\|_{L_t^2 L^{p_1}}, \quad \beta(p_0,p_1) = 1 + \frac{d}{2}\left(1 - \frac{1}{p_0} - \frac{1}{p_1}\right). \tag{4.94}$$

Note in particular that $\beta(\frac{2d}{d+2}, \frac{2d}{d+2}) = 0$.

Now, consider the bilinear mapping $(\Psi,\Phi) \mapsto \mu$, where μ is treated as an element of a sequence space. Our estimate (4.94) implies that for p_0, p_1 within the box defined by (4.93) that the mapping is bounded from $L_t^2 L^{p_0} \times L_t^2 L^{p_1} \to \mathring{\ell}^\infty_{-\beta(p_0,p_1)}$. And here we can take advantage of Theorem 4.68 and use a three-end-point interpolation.

We do so by first choosing two values

$$(p_0)^{-1} = \frac{d+2}{2d} + \frac{\epsilon}{2}, \qquad (p_1)^{-1} = \frac{d+2}{2d} - \epsilon$$

within the allowed box. We have that the mapping $(\Psi,\Phi) \mapsto \mu$ is bounded when consider as

$$\begin{aligned} L_t^2 L^{p_0} \times L_t^2 L^{p_0} &\to \mathring{\ell}^\infty_{d\epsilon/2} \\ L_t^2 L^{p_0} \times L_t^2 L^{p_1} &\to \mathring{\ell}^\infty_{-d\epsilon/4} \\ L_t^2 L^{p_1} \times L_t^2 L^{p_0} &\to \mathring{\ell}^\infty_{-d\epsilon/4} \end{aligned} \tag{4.95}$$

So if we apply Theorem 4.68 with $r = 1$, $p = q = 2$, and $\theta = \frac{2}{3}$ with $\theta_A = \theta_B = \frac{1}{3}$, we get that the mapping is a bounded mapping

$$(L_t^2 L^{p_0}, L_t^2 L^{p_1})_{\frac{1}{3},2} \times (L_t^2 L^{p_0}, L_t^2 L^{p_1})_{\frac{1}{3},2} \to \mathring{\ell}^1_0 \tag{4.96}$$

using what we know about real interpolation of sequence spaces (see Thought 4.60).

Observe that by construction

$$\frac{1-\theta_A}{p_0} + \frac{\theta_A}{p_1} = \frac{d+2}{2d}$$

for our choice of $\theta_A = \frac{1}{3}$. And so by our definition of Lorentz spaces we have

$$(L^2_t L^{p_0}, L^2_t L^{p_1})_{\frac{1}{3},2} = L^2_t L^{2d/(d+2)}_2 \supset L^2_t L^{2d/(d+2)}_{2d/(d+2)} = L^t_2 L^{2d/(d+2)}$$

where we used that $2d/(d+2) < 2$ and so we have the inclusion inequality for Lorentz spaces. Finally, since

$$\sum \mu_j = \iint_{\mathbb{R}^2} \langle U(t)U^*(t')\Psi(t'), \Phi(t)\rangle \,\mathrm{d}t \,\mathrm{d}t',$$

the boundedness (4.96) implies that

$$\left\| \int_{\mathbb{R}} U^*(t)\Phi(t) \,\mathrm{d}t \right\|_{L^2} \lesssim \|\Phi\|_{L^2_t L^{2d/(d+2)}}$$

and by duality (seeing that $(d+2)/2d + (d-2)/2d = 1$) implies the desired estimate. □

Chapter 5

Vector Field Method: Another Approach to Decay

As we have already seen in the discussion between Theorem 1.10 and Theorem 1.21, there are multiple ways to approach proving dispersive estimates. The Fourier integral method explained in the previous chapters roughly corresponds to the "fundamental solution" method used to prove Theorem 1.10. In this chapter we will apply the ideas of the proof of Theorem 1.21 to study the decay properties for the Schrödinger and wave equations.

Our general strategy, following our discussion on the Vlasov equation, consists of:

1. First, examine the symmetries of the equation in terms of its commuting vector fields. More precisely, we seek vector fields W such that if ϕ solves the equation of interest, so does $W\phi$.
2. Second, examine the symmetries of the equation in terms of its conservation laws.
3. Third, prove space-time weighted Sobolev inequalities to convert higher derivative estimates in integral norms to lower derivative estimates in pointwise norms.

This method was originally developed by Klainerman to systematically treat the decay properties of the wave equation. In the presentation below we instead start with a version adapted to the Schrödinger equations. There are various technical simplifications in this situation which, it is hoped, can make the overall strategy more transparent. The technically more complicated case of the wave and Klein-Gordon equations is treated in the second half of this chapter. Our treatment of the wave equation is non-standard compared to what is available in the literature; in particular our presentation relegates the wave equation estimates as a degenerate case of those available for the Klein-Gordon equations. The advantage lies in the somewhat shorter and unified presentation. The reader however is encouraged to consult the textbooks of F. John and C. Sogge for the standard presentations of the estimates available for the wave equations, for that is the formulation used in most extant proofs of small-data global well-posedness theorems in the context of quasilinear wave equations. The method presented, as applied to the Klein-Gordon equations, is however standard.

John, Nonlinear wave equations, formation of singularities; *Sogge*, Lectures on non-linear wave equations

Schrödinger

The first step in the vector field method is finding differential operators that commute with the equation. For the linear Schrödinger equation, which we re-write as

$$\partial_t \phi + i\triangle\phi = 0, \tag{5.1}$$

where $\phi : \mathbb{R}^{1+d} \to \mathbb{C}$, we will use the operators

$$W_j = t\partial_{x^j} + \frac{i}{2}x^j. \tag{5.2}$$

These operators are related to the Galilean boosts G_j use in the proof of Theorem 1.21 and introduced in Thought 1.19. If we employ the Wigner transformation $v \mapsto i\nabla^{(x)}$ as in the discussion surrounding (1.29), we would have obtained $G_j = t\partial_{x^j} + \partial_{v^i} \mapsto t\partial_{x^j} + ix^j$, which is off from our vector field by a factor of two in the second term.

We can check by direct computation that

$$W_j(\partial_t + i\triangle)\phi = (\partial_t + i\triangle)W_j\phi.$$

This commutation relation can be understood in terms of the quantum phase space described in the discussion about (1.29). For the linear Schrödinger equation, we can think of it as the superposition of a bunch of waves satisfying the dispersion relation

$$\omega = |k|^2.$$

That is to say, a solution to the linear Schrödinger equation can be viewed as a measure defined on the hypersurface $\{\omega = |k|^2 \mid (\omega,k) \in \mathbb{R}^{1+d}\}$. Therefore any transformation of the (ω,k) space that preserves the hypersurface $\{\omega = |k|^2\}$ gives rise to a transformation sending solutions to the linear Schrödinger equation to other solutions.

Particular examples of such transformations are the infinitesimal transformations given by vector fields tangent to the hypersurface, especially those vector fields $\widehat{W}$ satisfying $\widehat{W}(\omega - |k|^2) = 0$. Examples of these include the vector fields

$$\widehat{W_j} = k_j \partial_\omega + \frac{1}{2}\partial_{k^j}.$$

As the superposition principle (1.27) is nothing more than the Fourier transform, we see that by Proposition 2.9 the frequency space operation $\widehat{W_j}$ corresponds precisely to the physical space operation W_j.

Ref. 2.9: "Fourier transform properties: differentiation"

5.3 Exercise
Another example of transformations that preserve solutions are multiplication by scalar functions in frequency space. Interpret these operations in terms of Thought 1.17. Furthermore, consider the coordinate scalars ω and k_j; what are these operations in physical space? ■

Using the W_j operators, we can prove an analogue of (1.24) for Schrödinger equations. Observe that by the fundamental theorem of calculus we have

$$\left|\phi(t,x)\right|^2 = \int_{R_x} \partial_{x^1}\partial_{x^2}\cdots\partial_{x^d}(\phi\overline{\phi})\,\mathrm{d}y$$

where $R_x = \{y \in \mathbb{R}^d \mid y^i \leq x^i\}$ as before. We can rewrite in the form

$$\left|\phi(t,x)\right|^2 = \frac{1}{t^d}\int_{R_x} (t\partial_{x^1})(t\partial_{x^2})\cdots(t\partial_{x^d})(\phi\overline{\phi})\,\mathrm{d}y.$$

Next we use that

$$\begin{aligned} t\partial_{x^j}(\phi\overline{\psi}) &= \phi\cdot t\partial_{x^j}\overline{\psi} + \overline{\psi}\cdot t\partial_{x^j}\phi \\ &= \phi\cdot t\partial_{x^j}\overline{\psi} - \phi\cdot\frac{i}{2}x^j\overline{\psi} + \overline{\psi}\cdot\frac{i}{2}x^j\phi + \overline{\psi}\cdot t\partial_{x^j}\phi \\ &= \phi\overline{W_j\psi} + \overline{\psi}W_j\phi. \end{aligned}$$

And hence we have

$$\left|\phi(t,x)\right|^2 = \frac{1}{t^d} \sum_{\alpha+\beta=(1,1,\dots 1)} \int_{R_x} W^\alpha \phi \overline{W^\beta \phi} \, \mathrm{d}y;$$

we note that α, β are multi-indices and that $[W_i, W_j] = 0$ so that the multi-index notation is well-defined. Finally, by Cauchy-Schwarz we obtain the global Sobolev inequality

$$\left|\phi(t,x)\right|^2 \leq \frac{1}{t^d} \sum_{\alpha+\beta=(1,1,\dots 1)} \left\|W^\alpha \phi\right\|_{L^2(\mathbb{R}^d)} \left\|W^\beta \phi\right\|_{L^2(\mathbb{R}^d)}. \tag{5.4}$$

5.5 Remark
The estimate (5.4) holds for any function with suitable regularity and spatial decay, and not ust for solutions to the linear Schrödinger equation. Its usefulness in studying the Schrödinger equation lies in the fact that W^α commutes with the Schrödinger flow, and so if ϕ solves (5.1) so does $W^\alpha \phi$. Therefore by the conservation of mass (L^2) for the Schrödinger equation (see Exercise 4.29), the norm quantities on the right hand side of (5.4) are constant in time, and are determined entirely by the initial data. This in particular shows that, similar to the conclusion of Corollary 3.29, that $\left|\phi(t,x)\right|$ decays like $|t|^{d/2}$. ■

For the remainder of this section we focus on sharpening the right hand side of (5.4).

At first glance, one may think that (5.4) is much worse compared to Corollary 3.29, as in the latter the bound is in terms of merely the L^1 norm of the initial data and in the former W_j are first order differential operators. However, notice that at $t = 0$ we have $W_j = \frac{i}{2} x^j$, and so writing ϕ_0 for the initial data of the solution ϕ, we have that

$$\left|\phi(t,x)\right|^2 \leq \frac{1}{(2t)^d} \sum_{\alpha+\beta=(1,1,\dots 1)} \left\|x^\alpha \phi_0\right\|_{L^2(\mathbb{R}^d)} \left\|x^\beta \phi_0\right\|_{L^2(\mathbb{R}^d)}.$$

Having dealt with the regularity, we next consider the weight: for the right hand side in the above expression to be finite, in the case when $\alpha = (1, 1, \dots, 1)$ we expect to require $|x|^d \phi_0(x)$ be L^2 integrable, which would indicate an asymptotic decay rate of $\phi_0 \approx o(|x|^{-3d/2})$ as $|x| \nearrow \infty$. This decay is stronger than a typical L^1 decay rate of $\approx o(|x|^{-d})$, and suggests that our estimate is weaker, insofar as spatial decay is concerned, than Corollary 3.29.

However, we can improve the estimate by noting that by our assumption, when $|\alpha|$ is large, the corresponding $\left|\beta\right|$ must be small. We can balance the two by taking advantage of the fact that we are studying a linear equations, and using a dyadic decomposition in physical space. Let χ_k denote the characteristic function of the set $\{|x| \in [2^k, 2^{k+1})$. We can write

$$\phi_0 = \sum_{k\in\mathbb{Z}} \chi_k \phi_0.$$

Denote by $\phi^{(k)}$ the solution to (5.1) with initial data $\chi_k\phi_0$; we have that

$$\phi = \sum_{k\in\mathbb{Z}} \phi^{(k)}.$$

For each individual piece, however, we can apply our estimate to obtain

$$\left|\phi^{(k)}(t,x)\right|^2 \lesssim \frac{1}{t^d} \sum_{a+b=d} \left\||x|^a \chi_k \phi_0\right\|_{L^2} \left\||x|^b \chi_k \phi_0\right\|_{L^2}.$$

On the support of χ_k, however, $|x| \approx 2^k$; so we arrive at

$$\left|\phi^{(k)}(t,x)\right| \lesssim \frac{2^{kd/2}}{t^{d/2}} \left\|\chi_k \phi_0\right\|_{L^2}$$

with the implicit constant independent of k. Summing over k we get finally

$$\left|\phi(t,x)\right| \lesssim \frac{1}{t^{d/2}} \sum_{k\in\mathbb{Z}} 2^{kd/2} \left\|\chi_k \phi_0\right\|_{L^2}. \tag{5.6}$$

Equation (5.6) has the correct scaling: as $|x| \nearrow \infty$, we expect ϕ_0 to decay like $|x|^{-d}$ or better. For convenience we can define the following notation:

$$\begin{aligned} \|f\|_{Y^{s,q}(\mathbb{R}^d)} &\stackrel{\text{def}}{=} \left(\sum_{k\in\mathbb{Z}} 2^{skq} \|\chi_k f\|_{L^2}^q\right)^{\frac{1}{q}} \qquad s \in \mathbb{R}, q \in [1,\infty); \\ \|f\|_{Y^{s,\infty}(\mathbb{R}^d)} &\stackrel{\text{def}}{=} \sup_{k\in\mathbb{Z}} 2^{sk} \|\chi_k f\|_{L^2} \qquad s \in \mathbb{R}. \end{aligned} \tag{5.7}$$

The space $Y^{s,q}$ is a sequence space (see Thought 4.60); and has some similarities to the homogeneous Besov spaces with the index $p = 2$ (see Definition 3.44).

5.8 Exercise
Let's make the concept of *scaling* precise. Given a function f on $\mathbb{R}^d$, for $\lambda \in \mathbb{R}_+$ denote by $f_\lambda(x) = \lambda^d f(\lambda x)$.

1. Show that $\|f_\lambda\|_{L^1} = \|f\|_{L^1}$.
2. Show that $\|f_\lambda\|_{Y^{d/2,1}} = \|f\|_{Y^{d/2,1}}$ when $\lambda = 2^\ell$ for some $\ell \in \mathbb{Z}$. Conclude that there exists a universal constant C such that

$$C^{-1}\|f\|_{Y^{d/2,1}} \leq \|f_\lambda\|_{Y^{d/2,1}} \leq C\|f\|_{Y^{d/2,1}}. \qquad \blacksquare$$

5.9 Exercise
Now, let $a \in \mathbb{R}$ be a constant to be determined, and denote by $f_\lambda(x) = \lambda^a f(\lambda x)$.

1. Given $p \in [1,\infty]$, find a such that $\|f_\lambda\|_{L^p} = \|f\|_{L^p}$.
2. Given a, find all pairs $(s,q) \in \mathbb{R}\times[1,\infty]$ such that $\|f_\lambda\|_{Y^{s,q}}$ is uniformly (independently of f) comparable with $\|f\|_{Y^{s,q}}$. ■

5.10 Exercise (Other decay scales)
Prove the following assertions by interpolating between (5.6) and L^2 conservation. Suppose ϕ solves the linear Schrödinger equation with initial data ϕ_0, then

1. $\left\|\phi(t)\right\|_{L^p(\mathbb{R}^d)} \lesssim \frac{1}{|t|^\sigma}\left\|\phi_0\right\|_{Y^{\sigma,2}(\mathbb{R}^d)}$, where $\sigma = \frac{d}{2} - \frac{d}{p}$.
2. $\left\|\phi(t)\right\|_{Y^{-\sigma,2}(\mathbb{R}^d)} \lesssim \frac{1}{|t|^\sigma}\left\|\phi_0\right\|_{Y^{\sigma,2}(\mathbb{R}^d)}$, where $\sigma \in (0,d/2)$.

(*Hint:* the expression (4.63) is useful; for the second part you will need to prove, as a first step, the embedding relation $L^\infty \hookrightarrow Y^{-d/2,\infty}$.) ■

5.11 Remark
The second result in the previous exercise can be interpreted as a statement concerning "decay of local mass". Observe that the norm $\|f\|_{Y^{s,2}}$ is comparable with the norm $\left\||x|^s f\right\|_{L^2}$. When s is negative $Y^{s,2}$ is measuring how concentrated the mass density is around the origin, while when s is positive $Y^{s,2}$ is measuring how diffused the initial mass distribution is over large scales. So the conclusion of the previous exercise states that for initial mass distribution that is concentrated (not distributed over large scales), the solution cannot remain concentrated forever. This is exactly in concert with our physical expectations: by the Heisenberg uncertainty principle,

an initial mass distribution that is concentrated near the origin must have a frequency distribution that is spread out; so the entire distribution will quickly disperse leaving only small amounts near the origin. ■

5.12 Exercise (Integrated local energy decay)
Suppose ϕ solves the linear Schrödinger equation with initial data ϕ_0, prove that for every $\sigma \in (0,1)$,

$$\left\|\phi\right\|_{L_t^{2/\sigma} Y^{-\sigma,2}} \lesssim \left\|\phi_0\right\|_{L^2}.$$

■

In the last part of this section, we show how to recover the L^1–L^∞ decay of Corollary 3.29 starting from the estimate $\left|\phi(t,x)\right| \lesssim |t|^{-d/2}\left\|\phi_0\right\|_{Y^{d/2,1}}$. First we observe that this implication is not trivial, as $L^1(\mathbb{R}^d)$ does not embed into $Y^{d/2,1}(\mathbb{R}^d)$.

5.13 Exercise

1. Show that $Y^{d/2,1}(\mathbb{R}^d)$ embeds into $L^1(\mathbb{R}^d)$; that is to say, show that there exists a universal constant C such that

$$\|f\|_{L^1(\mathbb{R}^d)} \leq C\|f\|_{Y^{d/2,1}(\mathbb{R}^d)}.$$

2. Show that there exists a function $f \in L^1(\mathbb{R}^d)$ such that $f \notin Y^{d/2,1}(\mathbb{R}^d)$.
■

The key observation to make is that L^∞ is translation invariant, but not $Y^{d/2,1}$. So if we optimize via translations, and exploit the linearity of the solution operator, we can hope to further sharpen the inequality.

This is a variation of the "polarization argument".

Let τ_y be the operator $\tau_y f(x) = f(x+y)$. Exploiting the translation invariance of the L^∞ norm we have that (5.6) implies

$$\left\|\phi(t)\right\|_{L^\infty} \lesssim \frac{1}{|t|^{d/2}} \inf_{y\in\mathbb{R}^d}\left\|\tau_y\phi_0\right\|_{Y^{d/2,1}}.$$

Thus if ϕ, ψ are two solutions with data ϕ_0, ψ_0, we have

$$\left\|\phi(t)+\psi(t)\right\|_{L^\infty} \lesssim \frac{1}{|t|^{d/2}}\left(\inf_{y\in\mathbb{R}^d}\left\|\tau_y\phi_0\right\|_{Y^{d/2,1}} + \inf_{y'\in\mathbb{R}^d}\left\|\tau_{y'}\psi_0\right\|_{Y^{d/2,1}}\right).$$

Notice that the terms inside the brackets is bounded above by $\left\|\phi_0+\psi_0\right\|_{Y^{d/2,1}}$, and is generally somewhat smaller, especially if ϕ_0 and ψ_0 have disjoint support.

Now approximate a given $\phi_0 \in L^1$ by *simple functions*, which here we mean a finite linear combination of characteristic functions of cubes; given a simple function which we can write as

$$\sum_{k=1}^{N} a_k \mathbf{1}_{Q_k}(x)$$

we see that it suffices to estimate

$$\inf_{y\in\mathbb{R}^d} \left\| \tau_y \mathbf{1}_{Q_k} \right\|_{Y^{d/2,1}}.$$

But this can be directly computed to be bounded by $\left\|\mathbf{1}_{Q_k}\right\|_{L^1}$. And hence for ϕ_0 being a simple function we have

$$\left\|\phi(t)\right\|_{L^\infty} \lesssim \frac{1}{|t|^{d/2}} \left\|\phi_0\right\|_{L^1}.$$

Taking limits we arrive at our conclusion.

5.14 Exercise
Let Q be a cube. Prove that there exists a universal constant C such that

$$\inf_{y\in\mathbb{R}^d} \left\| \tau_y \mathbf{1}_{Q} \right\|_{Y^{d/2,1}} \leq C \left\|\mathbf{1}_{Q}\right\|_{L^1}.$$

(*Hint*: Can you guess a good y?) ■

Klein-Gordon and Wave

Next we consider the Klein-Gordon family of equations

$$\partial^2_{tt}\phi - \Delta\phi + M^2\phi = 0; \qquad (5.15)$$

where $M \in \mathbb{R}$ is the particle mass. In the case $M = 0$ the equation reduces to the linear wave equation. We introduce the D'Alembertian symbol

$$\square \stackrel{\text{def}}{=} -\partial^2_{tt} + \Delta\phi \qquad (5.16)$$

for the principal part of our operator. Geometrically the D'Alembertian is the *Laplace-Beltrami operator* associated to the Minkowski metric on $\mathbb{R}^{1+d}$. Our computations and constructions below are rooted in an understanding of Lorentzian geometry; however the calculations themselves can be verified without this background knowledge.

5.17 Lemma
The following vector fields commute with the linear evolution (5.15):

- Space-time translations $\partial_t, \partial_{x^1}, \ldots, \partial_{x^d}$.
- Spatial rotations $\Omega_{ij} = x^i \partial_{x^j} - x^j \partial_{x^i}$.
- Lorentz boosts $L^i = t\partial_{x^i} + x^i \partial_t$. ■

5.18 Exercise

1. Verify the commutation properties of Lemma 5.17; that is to say, letting V denote any of the vector fields listed in the Lemma, show that $V(\Box\phi - m^2\phi) = (\Box - M^2)V\phi$.
2. Write down the frequency-space dispersion relation for (5.15).
3. Verify that the operators in Lemma 5.17 correspond to *frequency space* transformations that preserve measures whose support are given by the dispersion relation found in the previous part. ■

5.19 Remark
In addition to interpreting the vector fields in Lemma 5.17 as symmetry properties of the physical laws, and as operators tangent to the Fourier support, we can also interpret the vector fields as symmetries of the underlying Minkowski metric. That is to say, on Minkowski space they are *Killing vector fields*. ■

Now recall the method as we discussed for the Vlasov and Schrödinger equation: to prove dispersive estimates we rely on two ingredients: first is a conservation law and second a weighted Sobolev inequality. For the Vlasov and Schrödinger equations, using that the G_i and W_i operators are essentially tangent to the level sets of the time function t, we use adapted Sobolev inequalities on such slices. Note, however, the natural candidate to provide temporal decay in the wave and Klein-Gordon equation cases are the t-weighted vector field corresponding to Lorentz boosts; they are not tangent to the constant t hypersurfaces. Instead, we are led to consider Sobolev inequalities on hypersurfaces to which L^i are tangent.

5.20 (Geometric setup) Using that L^i is geometrically a symmetry of Minkowski space, we expect it to be tangent to level sets of the Minkowski distance. Indeed, consider the function $t^2 - |x|^2$, then one easily checks that $L^i(t^2 - |x|^2) = 0$. We will focus our attention on the level sets with $t^2 - |x|^2 > 0$, that is, the future of the expanding light cone emanating from the origin of Minkowski space. On this set we define the function $\tau \stackrel{\text{def}}{=} \sqrt{t^2 - |x|^2}$, and denote by Σ_τ its level sets; note that they are hyperboloids that asymptote to the cone $t^2 = |x|^2$.

We note that the vector fields $\{L^i\}$ are linearly independent, and since there are d of them they must span the tangent space $T\Sigma_\tau$.

For computation, we choose a coordinate system for the set $\{t^2 > |x|^2\}$. Let

$$(\tau, \rho, \theta) \in \mathbb{R}_+ \times \mathbb{R}_+ \times \mathbb{S}^{d-1} \tag{5.21}$$

(where we identify $\mathbb{S}^{d-1}$ canonically as a submanifold of $\mathbb{R}^d$) be the coordinate system defined by

$$t = \tau \cosh(\rho), \tag{5.22}$$

$$x = \tau \sinh(\rho) \cdot \theta. \tag{5.23}$$

Relative to this coordinate system, the Minkowski metric takes the warped-product form

$$m = -\,\mathrm{d}t^2 + \sum_{i=1}^{d} \mathrm{d}(x^i)^2 \mapsto -\,\mathrm{d}\tau^2 + \tau^2\,\mathrm{d}\rho^2 + \tau^2 \sinh(\rho)^2\,\mathrm{d}\theta^2 \tag{5.24}$$

where by $\mathrm{d}\theta^2$ we refer to the standard metric on $\mathbb{S}^{d-1}$.

We will write h_τ and $(h_\tau)^{-1}$ the induced Riemannian metric and its inverse on Σ_τ; they have the coordinate expressions

$$h_\tau = \tau^2(\mathrm{d}\rho^2 + \sinh(\rho)^2\,\mathrm{d}\theta^2), \tag{5.25}$$

$$(h_\tau)^{-1} = \frac{1}{\tau^2}(\partial_\rho \otimes \partial_\rho + \frac{1}{\sinh(\rho)^2}\partial_\theta \otimes \partial_\theta), \tag{5.26}$$

where $\partial_\theta \otimes \partial_\theta$ is the inverse standard metric on $\mathbb{S}^{d-1}$. ¶

5.27 (Representation of the metric) Observe that

$$\begin{aligned} L^i \otimes L^i &= (x^i)^2 \partial_t \otimes \partial_t + x^i t(\partial_t \otimes \partial_{x^i} + \partial_{x^i} \otimes \partial_t) + t^2 \partial_{x^i} \otimes \partial_{x^i} \\ \implies \sum_{i=1}^{d} L^i \otimes L^i &= r^2 \partial_t \otimes \partial_t + rt(\partial_t \otimes \partial_r + \partial_r \otimes \partial_t) + t^2 \partial_r \otimes \partial_r \\ &\quad + t^2 \left(\sum_{i=1}^{d} \partial_{x^i} \otimes \partial_{x^i} - \partial_r \otimes \partial_r \right) \\ &= \partial_\rho \otimes \partial_\rho + \frac{t^2}{r^2} \partial_\theta \otimes \partial_\theta \end{aligned}$$

and hence

$$\sum_{i=1}^{d} L^i \otimes L^i = \partial_\rho \otimes \partial_\rho + \frac{\cosh^2(\rho)}{\sinh^2(\rho)} \partial_\theta \otimes \partial_\theta. \tag{5.28}$$

On the other hand we have

$$\begin{aligned}\sum_{i<j} \Omega_{ij} \otimes \Omega_{ij} &= \sum_{i<j} (x^i)^2 \partial^2_{x^j} + (x^j)^2 \partial^2_{x^j} - x^i x^j (\partial_{x^i} \otimes \partial_{x^j} + \partial_{x^j} \otimes \partial_{x^i}) \\ &= r^2 \sum_i \partial_{x^i} \otimes \partial_{x^i} - r^2 \partial_r \otimes \partial_r\end{aligned}$$

which implies

$$(\tau^{-2} h_\tau)^{-1} + \sum_{i<j} \Omega_{ij} \otimes \Omega_{ij} = \sum_{i=1}^{d} L^i \otimes L^i. \tag{5.29}$$

We remark that $\tau^{-2} h_\tau$ is the metric on *standard hyperbolic space* $\mathbb{H}^d$. Furthermore, the coercivity properties above implies that, with ∇ the Levi-Civita connection on Σ_τ relative to the metric h_τ, we have

$$\langle \nabla f, \nabla f \rangle_{\tau^{-2} h_\tau} \leq \sum_{i=1}^{d} \left| L^i f \right|^2. \tag{5.30}$$

We will make use of this inequality in our formulation of the generalized Sobolev inequalities. ¶

5.31 **(Commutators)** We will let $\mathcal{Z}$ denote the set $\{L^i, \Omega_{ij}\}_{i,j \in \{1,\dots,d\}}$, and Z a generic element. We note that under commutation $\mathcal{Z}$ forms an algebra:

$$\begin{aligned}[L^i, L^j] &= \Omega_{ij}, \\ [\Omega_{ij}, \Omega_{jk}] &= \Omega_{ik}, \\ [L^i, \Omega_{ij}] &= L^j.\end{aligned}$$

¶

An application of the Sobolev inequality on the unit disk gives the following version on a bounded region of hyperbolic space.

5.32 Proposition (Sobolev inequality)
Let f be a function defined on hyperbolic space $\mathbb{H}^d$, which we represent in polar coordinates $(\rho, \theta) \in \mathbb{R}_+ \times \mathbb{S}^{d-1}$. Then we have

$$\sup_{\rho < \frac{5}{3}} \left| f(\rho, \theta) \right|^2 \lesssim \sum_{k \leq \lfloor \frac{d}{2} \rfloor + 1} \int_0^2 \int_{\mathbb{S}^{d-1}} \left| \nabla^k f \right|_h^2 \sinh(\rho)^{d-1} \, \mathrm{d}\theta \, \mathrm{d}\rho;$$

h is the standard hyperbolic metric, and ∇ is the Levi-Civita connection. ■

Combining this estimate with (5.30) as well as the fact that $\cosh(\rho)$ is bounded above and below on $\rho \in (0,2)$ we have the following corollary.

5.33 Corollary
Let f be a function defined on $\Sigma_\tau \subset \mathbb{R}^{1+d}$. Let $\ell \in \mathbb{R}$. Then

$$\sup_{\rho<\frac{5}{3}} \left| f(\tau,\rho,\theta)\right|^2 \cosh(\rho)^\ell \lesssim \tau^{-d} \sum_{|\alpha|\leq \lfloor \frac{d}{2}\rfloor+1} \int_{\Sigma_\tau} \cosh(\rho)^\ell |L^\alpha f|^2 \mathrm{dvol}_{h_\tau}.$$

■

5.34 Remark
We note that the L^i do not pairwise commute, as $[L^i, L^j] = \Omega_{ij}$; so here α is not really a multi-index. We've abused notations and the sum is in fact over *all possible permutations* of strings of $L^i L^j \cdots L^k$ of no more than $\lfloor \frac{d}{2}\rfloor + 1$ symbols. ■

Proof We note first that the intrinsic metric on Σ_τ is conformal to the standard hyperbolic metric, and so they have the same Levi-Civita connection. This implies that by Proposition 5.32 we get

$$\sup_{\rho<\frac{5}{3}} \left| f(\tau,\rho,\theta)\right|^2 \lesssim \sum_{k\leq \lfloor \frac{d}{2}\rfloor+1} \int_{\Sigma_\tau\cap\{\rho<2\}} \left|\nabla^k f\right|^2_{\tau^{-2}h_\tau} \mathrm{dvol}_{\tau^{-2}h_\tau}.$$

By (5.30) we can replace the metric norm

$$\left|\nabla^k f\right|^2_{\tau^{-2}h_\tau} \leq \sum_{|\alpha|\leq k} |L^\alpha f|^2;$$

here we implicitly used that when $\rho < 2$ the quantities in the lower-order terms $\left|\nabla_{L^i} L^j\right|_{\tau^{-2}h_\tau}$ have universal bounds, and the L^i are close to orthogonal.

The upper-and-lower boundedness of $\cosh(\rho)$ implies we can insert those factors into our estimates with impunity, and finally we can expand the domain of integration to the whole of Σ_τ and rescale to the induced metric to deduce the desired inequality. □

On $\rho > 1$ we can also use a different estimate. The Sobolev inequality on the half-infinite cylinder states that

$$\sup_{\rho>43} \left| f(\rho,\theta)\right|^2 \lesssim \sum_{|\alpha|\leq\lfloor d/2\rfloor+1} \int_1^\infty \int_{\mathbb{S}^{d-1}} |\partial^\alpha f|^2 \,\mathrm{d}\theta\,\mathrm{d}\rho; \tag{5.35}$$

here we use multi-index notation α. This implies

5.36 Proposition
Let f be a function defined on $\Sigma_\tau \subset \mathbb{R}^{1+d}$. Let $\ell \in \mathbb{R}$. Then

$$\sup_{\rho > \frac{4}{3}} \left| f(\tau,\rho,\theta)^2 \right| \cosh(\rho)^{\ell+d-1} \lesssim \tau^{-d} \sum_{|\alpha| \leq \lfloor \frac{d}{2} \rfloor + 1} \int_{\Sigma_\tau} \cosh(\rho)^\ell |L^\alpha f|^2 \mathrm{dvol}_{h_\tau}.$$

■

Proof We apply (5.35) to the function $f \cosh(\rho)^{\ell/2} \sinh(\rho)^{(n-1)/2}$. We observe that when $\rho > 1$ both $\cosh(\rho)$ and $\sinh(\rho)$ are uniformly comparable to e^ρ, and hence we have the weighted version

$$\sup_{\rho > 43} \left| f(\rho,\theta) \right|^2 \cosh(\rho)^\ell \sinh(\rho)^{d-1} \lesssim$$
$$\sum_{|\alpha| \leq \lfloor d/2 \rfloor + 1} \int_1^\infty \int_{\mathbb{S}^{d-1}} \cosh(\rho)^\ell |\partial^\alpha f|^2 \underbrace{\sinh(\rho)^{d-1} \, \mathrm{d}\theta \, \mathrm{d}\rho}_{\mathrm{dvol}_{\tau^{-2} h_\tau}}.$$

Next, observing that both ∂_ρ and ∂_θ can be expressed in terms of $\frac{x}{r} L$, and that the coefficients x/r have bounded derivatives (along Σ_τ) to all orders away from $\rho = 0$, we conclude that

$$\left| \partial^\beta f \right|^2 \lesssim \sum_{|\alpha| \leq |\beta|} |L^\alpha f|^2.$$

So doing another rescaling after expanding the domain of integration to cover the entirety of Σ_τ, we recover the claimed Sobolev estimate. □

Putting everything together we have the following theorem.

5.37 Theorem (Global Sobolev inequality)
Let f be a function define on the region $\{t^2 > |x|^2\} \subset \mathbb{R}^{1+d}$ and $\ell \in \mathbb{R}$, then

$$\left| f(\tau,\rho,\theta) \right|^2 \lesssim \tau^{-d} \cosh(\rho)^{1-d-\ell} \sum_{|\alpha| \leq \lfloor \frac{d}{2} \rfloor + 1} \int_{\Sigma_\tau} \cosh(\rho)^\ell |L^\alpha f|^2 \mathrm{dvol}_{h_\tau}$$

provided the integrals converge. ■

Having obtained the requisite Sobolev inequality, we turn our attention to the derivation of the conservation laws that will enable our use of Theorem 5.37. For the Klein-Gordon equation, we can define the associated *stress-energy tensor*

$$Q^{[M]}_{\alpha\beta} \stackrel{\text{def}}{=} \partial_\alpha \phi \partial_\beta \phi - \frac{1}{2} m_{\alpha\beta} \left| \mathrm{d}\phi \right|^2_m - \frac{1}{2} m_{\alpha\beta} M^2 \phi^2. \qquad (5.38)$$

Here m is the Minkowski metric defined in (5.24). If we compute its divergence we get

$$(m^{-1})^{\alpha\gamma}\partial_\gamma Q^{(M)}_{\alpha\beta} = \Box\phi\partial_\beta\phi + (m^{-1})^{\alpha\gamma}\partial_\alpha\phi\partial^2_{\beta\gamma}\phi - \frac{1}{2}\underbrace{(m^{-1})^{\alpha\gamma}m_{\alpha\beta}\partial_\gamma}_{=\partial_\beta}\left(\left|\mathrm{d}\phi\right|^2_m + M^2\phi^2\right)$$
$$= (\Box\phi + M^2\phi)\partial_\beta\phi.$$

Hence if ϕ solves (5.15), then $Q^{(M)}_{\alpha\beta}$ is divergence free. As ∂_t is a constant (parallel) vector field on Minkowski space, we have then that, defining

$$^{(\partial_t)}P^\gamma \stackrel{\text{def}}{=} (m^{-1})^{\alpha\gamma}Q^{[M]}_{\alpha\beta}(\partial_t)^\beta = Q(\partial_t,\bullet)^\sharp \tag{5.39}$$

the vector field $^{(\partial_t)}P$ is divergence free.

5.40 Lemma
Let X be a future causal vector, that is to say, $X = X_0\partial_t + \sum X_i\partial_i$ with $X_0 > 0$ and $(X_0)^2 \geq \sum(X_i)^2$, then $\left\langle X, {}^{(\partial_t)}P\right\rangle_m \geq 0$. ■

5.41 Exercise
Prove the preceding lemma. (It is an exercise in completing the square.) ■

Consider the region

$$\mathcal{D}^\tau \stackrel{\text{def}}{=} \{(t,x)\in\mathbb{R}^{1+d} \mid t > 0,\ \ t^2 \leq \tau^2 + |x|^2\}$$

sandwiched between $\{t = 0\}$ and Σ_τ. Applying the divergence theorem to $\int_{\mathcal{D}^\tau}\operatorname{div}{}^{(\partial_t)}P\,\mathrm{dvol}$, we obtain the following estimate.

5.42 Proposition
If ϕ solves (5.15) with initial data in $\mathscr{S}$, then

$$\int_{\Sigma_\tau} Q(\partial_t,\partial_\tau)\mathrm{dvol}_{h_\tau} \leq \int_{\{0\}\times\mathbb{R}^d} Q(\partial_t,\partial_t)\,\mathrm{d}x.$$
■

Proof Consider the subregions

$$\mathcal{D}^\tau_\mu \stackrel{\text{def}}{=} \mathcal{D}^\tau \cap \{(t,x)\in\mathbb{R}^{1+d} \mid |x| < \mu - t\}.$$

Their boundaries we denote by

$$\Sigma_{\tau,\mu} \stackrel{\text{def}}{=} \{t^2 - |x|^2 = \tau^2, \ \ |x| < \mu - t\},$$
$$R_\mu \stackrel{\text{def}}{=} \{(0,x) \mid |x| < \mu\},$$
$$C_{\tau,\mu} \stackrel{\text{def}}{=} \{t^2 \leq \tau^2 - x^2, \ \ 0 \leq t = \mu - |x|\}.$$

The divergence theorem states that

$$\int\limits_{\Sigma_{\tau,\mu}+R_\mu+C_{\tau,\mu}} \iota_{(\partial_t)P}\mathrm{dvol}_m = 0.$$

(Observe that the interior product $\iota_{(\partial_t)P}\mathrm{dvol}_m$ is an n-form and can be integrated over an n dimensional submanifold carrying the induced orientation.)

Along Σ_τ, the space-time volume form dvol_m can be expressed as

$$\mathrm{d}\tau \wedge \mathrm{dvol}_{h_\tau}$$

and so the integral along $\Sigma_{\tau,\mu}$ can be written as

$$\int\limits_{\Sigma_{\tau,\mu}} \iota_{(\partial_t)P}\mathrm{dvol}_m = \int\limits_{\Sigma_{\tau,\mu}} {}^{(\partial_t)}P(\tau)\mathrm{dvol}_{h_\tau} = \int\limits_{\Sigma_{\tau,\mu}} Q(\partial_\tau, \partial_t)\mathrm{dvol}_{h_\tau}.$$

Similarly, factoring $\mathrm{dvol}_m = \mathrm{d}t \wedge \mathrm{d}x$, and observing that change of orientation which contributes a minus sign, we have that

$$\int\limits_{R_\mu} \iota_{(\partial_t)P}\mathrm{dvol}_m = -\int\limits_{|x|<\mu} Q(\partial_t, \partial_t)\,\mathrm{d}x.$$

Finally, we observe that as a consequence of Lemma 5.40, the integral

$$\int\limits_{C_{\tau,\mu}} \iota_{(\partial_t)P}\mathrm{dvol}_m \geq 0.$$

Therefore we have the uniform estimate

$$\int\limits_{\Sigma_{\tau,\mu}} Q(\partial_\tau, \partial_t)\mathrm{dvol}_{h_\tau} \leq \int\limits_{t=0,|x|<\mu} Q(\partial_t, \partial_t)\,\mathrm{d}x \leq \int\limits_{\{0\}\times\mathbb{R}^d} Q(\partial_t, \partial_t)\,\mathrm{d}x.$$

Taking the limit $\mu \nearrow +\infty$ concludes the proof. □

For convenience, given ϕ a solution to (5.15), we denote by

$$\mathcal{E}[\phi] \stackrel{\text{def}}{=} \left[\int_{\{0\}\times\mathbb{R}^d} Q(\partial_t, \partial_t) \,\mathrm{d}x \right]^{\frac{1}{2}} \tag{5.43}$$

the total initial energy. Note that if $\phi(0,x) = \phi_0(x)$ and $\partial_t\phi(0,x) = \phi_1(x)$, we have

$$\mathcal{E}[\phi] \lesssim \left\|\nabla\phi_0\right\|_{L^2} + \left\|\phi_1\right\|_{L^2} + M\left\|\phi_0\right\|_{L^2}.$$

It remains to see what $Q(\partial_t, \partial_\tau)$ evaluates to; its coercivity is crucial in applying the global Sobolev inequality. For this, it is convenient to express

$$\partial_t = \cosh(\rho)\partial_\tau - \tau^{-1}\sinh(\rho)\partial_\rho \tag{5.44}$$

using the hyperboloidal coordinate system introduced earlier. Then we have

$$Q(\partial_t, \partial_\tau) = \cosh(\rho) Q(\partial_\tau, \partial_\tau) - \tau^{-1}\sinh(\rho) Q(\partial_\rho, \partial_\tau).$$

The second term is easy: since ∂_ρ and ∂_τ are orthogonal we have that

$$Q(\partial_\rho, \partial_\tau) = \partial_\rho\phi\partial_\tau\phi.$$

For the first term, as ∂_τ is a unit time-like vector, we find

$$Q(\partial_\tau, \partial_\tau) = \frac{1}{2}(\partial_\tau\phi)^2 + \frac{1}{2\tau^2}(\partial_\rho\phi)^2 + \frac{1}{2\tau^2\sinh(\rho)^2}\left|\partial_\theta\phi\right|^2 + \frac{1}{2}M^2\phi^2.$$

Combining the two and completing the square we get

$$\begin{aligned} Q(\partial_\tau, \partial_t) = {} & \frac{\cosh(\rho)}{2\tau^2\sinh(\rho)^2}\left|\partial_\theta\phi\right|^2 + \frac{1}{2\tau^2\cosh(\rho)}(\partial_\rho\phi)^2 \\ & + \frac{1}{2}\cosh(\rho)M^2\phi^2 + \frac{1}{2}\cosh(\rho)\left(\partial_\tau\phi - \frac{\sinh(\rho)}{\tau\cosh(\rho)}\partial_\rho\phi\right)^2. \end{aligned}$$

Now, going back to (5.28), and the decomposition (5.44), we have that

$$\begin{aligned} Q(\partial_\tau, \partial_t) = {} & \frac{1}{2\tau^2\cosh(\rho)}\sum_{i=1}^{d}(L^i\phi)^2 \\ & + \frac{1}{2\cosh(\rho)}(\partial_t\phi)^2 + \frac{1}{2}\cosh(\rho)M^2\phi^2. \end{aligned} \tag{5.45}$$

5.46 Exercise (Klein-Gordon decay)
Suppose ϕ solves (5.15) with $M \neq 0$. Using (5.45) as well as Theorem 5.37, prove that within the region $\{t^2 > |x|^2\}$ we have the uniform estimate

$$\left|\phi(t,x)\right| \lesssim \frac{1}{t^{d/2}} \sum_{|\alpha| \leq \lfloor \frac{d}{2} \rfloor + 1} \mathcal{E}[L^\alpha \phi]$$

provided that the right hand side is finite. ∎

5.47 Remark
We note that in the previous exercise, on the level of initial data we need to control up to $d/2 + 2$ derivatives. ∎

In the case of wave equation, the energy does not (obviously) control the L^2 integral (though as we will see later some control is possible) due to the lack of an M term. Instead we find the decay through the derivative terms in (5.45).

5.48 Theorem (Wave equation, derivative decay)
Let ϕ solve (5.15) with $M = 0$. Then within the region $\{t^2 > |x|^2\}$ we have the uniform estimates

$$\left|L^i \phi(t,x)\right| \lesssim \frac{1}{t^{d/2-1}} \sum_{|\alpha| \leq \lfloor \frac{d}{2} \rfloor + 1} \mathcal{E}[L^\alpha \phi],$$

$$\left|\partial_t \phi(t,x)\right| \lesssim \frac{1}{\sqrt{t+|x|}\sqrt{t-|x|}\, t^{d/2-1}} \sum_{|\alpha| \leq \lfloor \frac{d}{2} \rfloor + 1} \mathcal{E}[L^\alpha \phi],$$

provided the energies on the right hand side are finite. ∎

5.49 Remark
We can compare the rates to what is proved in Theorem 3.46. Aside from the fact that our estimate only prove decay at the level of the first derivative, and not at the level of the solution itself, we see that we are, morally speaking, using the same number of derivatives on the initial data. In terms of the decay rate, first examine that for $\partial_t \phi$: the decay proved here is slightly sharper than Theorem 3.46. In addition to the expected $t^{(d-1)/2}$ decay (contributed by the factors $t^{n/2-1}\sqrt{t+|x|}$), we see that there is an *additional* decay away from the light cone, when $t - |x|$ is large.

Ref. 3.46: "Summary of decay of solutions to the wave equation"

For the L^i derivatives, we note that since $L = t\partial_x + x\partial_t$, it has "length" at least t when regarded in Euclidean coordinates. So naively we would expect, based on the Theorem 3.46, that $L\phi$ should decay like $t^{(d-3)/2}$. What we proved is the better decay rate of $t^{(d-2)/2}$; this phenomenon is well-known for the wave equation, that while "generic" derivatives of the solution decay at the rate $t^{(d-1)/2}$, the derivatives in certain "tangential" directions decay faster like $t^{d/2}$.

Ref. 3.10: "Van der Corput: stationary case, $k = 2$"

Ref. 3.8: "Van der Corput: non-stationary case"

This can be explained also from the stationary phase point of view. Going back to Lemma 3.10, we see that the obstruction to not being able to run Lemma 3.8 lies in not being able to generically divide the amplitude ϕ by the gradient of the phase function η' when integrating by parts, when η' has no lower bound. If, however, ϕ vanishes *precisely* on the zero set of η', then the division can be carried out and we can improve the $1/\sqrt{\lambda}$ decay rate in Lemma 3.10 to $1/\lambda$ since we can now integrate by parts once with ϕ/η' still bounded.

Recalling that differentiation in physical space is the same as multiplication by coordinate functions in Fourier space, we see that if the derivative is such that the corresponding multiplier vanishes precisely at the critical points of the phase function when evaluating the stationary phase argument, we expect a gain of at least $t^{1/2}$ decay rate for the solutions of the wave equation in the same manner as described above. ■

5.50 Remark

To be slightly more precise: consider the "angular derivatives" $r^{-1}\Omega_{ij}$; these are "length one" derivatives. Using that

$$\Omega_{ij} = \frac{x^i}{r}L^j - \frac{x^j}{r}L^i$$

we see that

$$\left|\frac{1}{r}\Omega_{ij} f(t,x)\right| \lesssim \frac{1}{t^{d/2-1}r}.$$

Furthermore, using that at $\rho = 0$ we have $L^i = t\partial_{x_i}$, we have that

$$|\partial f(t,0)| \lesssim \frac{1}{t^{d/2}}.$$

(Note that this is a bit weaker than the rate found in Exercise 3.27.) ■

Proof (Theorem 5.48) First let us consider the decay estimates for $L^i\phi$. Observe that the conserved energy controls

$$\int_{\Sigma_\tau} \frac{1}{2\tau^2\cosh(\rho)} \sum (L^i\phi)^2 \mathrm{dvol}_{h_\tau} \leq \int_{\Sigma_\tau} Q(\partial_t, \partial_\tau)\mathrm{dvol}_{h_\tau}. \tag{5.51}$$

Considering the energy estimates for $L^\alpha \phi$ we have control over

$$\int_{\Sigma_\tau} \frac{1}{\tau^2 \cosh(\rho)} \sum_{i=1}^{d} (L^i L^\alpha \phi)^2 \mathrm{dvol}_{h_\tau}.$$

If we sum over all α with length at most $\lfloor d/2 \rfloor + 1$, the above will also control

$$\int_{\Sigma_\tau} \frac{1}{\tau^2 \cosh(\rho)} \sum_{i=1}^{d} \sum_{|\beta| \leq \lfloor \frac{d}{2} \rfloor + 1} (L^\beta L^i \phi)^2 \mathrm{dvol}_{h_\tau},$$

without needing to commute operators (and thus simplifies the argument). Therefore we have that, by Theorem 5.37,

$$\frac{1}{\tau^2} \left| L^i \phi(\tau, \rho, \theta) \right|^2 \lesssim \frac{1}{\tau^d \cosh(\rho)^{d-2}} \sum_{|\beta| \leq \lfloor \frac{d}{2} \rfloor + 1} \mathcal{E}[L^\beta \phi]^2.$$

Our claim follows after noting $\tau \cosh(\rho) = t$.

The estimate for $\partial_t \phi$ is slightly more involved. The conserved energy controls

$$\sum_{|\alpha| \leq \lfloor \frac{d}{2} \rfloor + 1} \int_{\Sigma_\tau} \frac{1}{2\cosh(\rho)} (\partial_t L^\alpha \phi)^2 \mathrm{dvol}_{h_\tau} \leq \sum_{|\alpha| \leq \lfloor \frac{d}{2} \rfloor + 1} \mathcal{E}[L^\alpha \phi]^2.$$

In order to apply the global Sobolev inequality, we need instead to control

$$\sum_{|\alpha| \leq \lfloor \frac{d}{2} \rfloor + 1} \int_{\Sigma_\tau} \frac{1}{\cosh(\rho)} (L^\alpha \partial_t \phi)^2 \mathrm{dvol}_{h_\tau},$$

with the position of L^α and ∂_t swapped. Therefore we need to consider the commutators. We can first compute

$$[L^i, \partial_t] = -\partial_{x^i},$$

this requires us to then further compute

$$[L^i, \partial_{x^j}] = -\delta_{ij} \partial_t.$$

By induction we see then

$$\left| L^\alpha \partial_t \phi \right| \lesssim \sum_{|\beta| \leq |\alpha|} \left| \partial_t L^\beta \phi \right| + \sum_{|\beta| \leq |\alpha| - 1} \sum_{i=1}^{d} \left| \partial_{x^i} L^\beta \phi \right|.$$

Using that

$$\partial_{x^i} = \frac{1}{t}(L^i - x^i \partial_t)$$

and $x^i/t < 1$ in our region of consideration, we get

$$\left|L^\alpha \partial_t \phi\right| \lesssim \sum_{\left|\beta\right| \leq |\alpha|} \left|\partial_t L^\beta \phi\right| + \frac{1}{\tau \cosh(\rho)} \left|L^\beta \phi\right|.$$

The latter term is controlled by $\frac{1}{\tau}\left|L^\beta \phi\right|$, using that $\cosh(\rho) \geq 1$. And hence we conclude that

$$\sum_{|\alpha| \leq \lfloor \frac{d}{2} \rfloor + 1} \int_{\Sigma_\tau} \frac{1}{\cosh(\rho)} (L^\alpha \partial_t \phi)^2 \mathrm{dvol}_{h_\tau} \lesssim$$

$$\sum_{|\alpha| \leq \lfloor \frac{d}{2} \rfloor + 1} \int_{\Sigma_\tau} \frac{1}{\cosh(\rho)} (\partial_t L^\alpha \phi)^2 + \frac{1}{\tau^2 \cosh(\rho)} \sum_{i=1}^{d} (L^i L^\alpha \phi)^2 \mathrm{dvol}_{h_\tau}$$

$$\lesssim \sum_{|\alpha| \leq \lfloor \frac{d}{2} \rfloor + 1} \mathcal{E}[L^\alpha \phi]^2.$$

So by the global Sobolev inequality we get that

$$\left|\partial_t \phi\right|^2 \lesssim \frac{1}{\tau^d \cosh(\rho)^{d-2}} \sum_{|\alpha| \leq \lfloor \frac{d}{2} \rfloor + 1} \mathcal{E}[L^\alpha \phi]^2$$

as claimed. □

It is also possible to get some decay for ϕ itself, without any derivatives, when working in dimension $d \geq 3$. We will require the following Hardy-type inequality.

5.52 Lemma (Hardy's inequality)
Let $f : [0, \infty) \to \mathbb{R}$ be a function with compact support (f is allowed to be non-vanishing at 0). Then

$$\int_0^\infty f(\rho)^2 \cosh(\rho) \sinh(\rho)^{\alpha - 1} \, \mathrm{d}\rho \leq \frac{4}{\alpha^2} \int_0^\infty f'(\rho)^2 \frac{\sinh(\rho)^{\alpha+1}}{\cosh(\rho)} \, \mathrm{d}\rho.$$

■

Proof We note that

$$\int_0^\infty \partial_\rho (f(\rho)^2 \sinh(\rho)^\alpha)\, \mathrm{d}\rho = 0$$

by the fundamental theorem of calculus. So we have

$$\alpha \int_0^\infty f(\rho)^2 \cosh(\rho) \sinh(\rho)^{\alpha-1}\, \mathrm{d}\rho \leq \left| 2 \int_0^\infty f(\rho) f'(\rho) \sinh(\rho)^\alpha\, \mathrm{d}\rho \right|.$$

Cauchy-Schwarz on the right implies

$$\alpha \int_0^\infty f(\rho)^2 \cosh(\rho) \sinh(\rho)^{\alpha-1}\, \mathrm{d}\rho \leq$$

$$2 \left(\int_0^\infty f(\rho)^2 \cosh(\rho) \sinh(\rho)^{\alpha-1}\, \mathrm{d}\rho \right)^{\frac{1}{2}} \left(\int_0^\infty f'(\rho)^2 \frac{\sinh(\rho)^{\alpha+1}}{\cosh(\rho)}\, \mathrm{d}\rho \right)^{\frac{1}{2}},$$

which simplifies to the claimed inequality. □

An immediate consequence of this inequality is the following spectral gap property:

On $\mathbb{R}^d$, the spectrum of the Laplacian goes all the way to zero, and hence there are no uniform estimates of the L^2 norm of a function by its $\mathring{H}^1$ norm. On $\mathbb{H}^d$, on the other hand, the spectrum of the Laplacian has a "gap", and so $\mathring{H}^1$ is strongly coercive on L^2.

5.53 Corollary
Let $d \geq 3$, then

$$\int_{\Sigma_\tau} \frac{1}{\cosh(\rho)} \phi^2 \mathrm{dvol}_{h_\tau} \leq \frac{4}{(d-2)^2} \int_{\Sigma_\tau} \frac{1}{\cosh(\rho)} \sum_{i=1}^d \left| L^i \phi \right|^2 \mathrm{dvol}_{h_\tau}.$$

■

This in turn implies

5.54 Theorem (Wave equation, solution decay)
Let ϕ solve (5.15) with $M = 0$, then within the region $\{t^2 > |x|^2\}$ we have the uniform estimate

$$\left|\phi\right| \lesssim \frac{1}{t^{d/2-1}} \sum_{|\alpha| \leq \lfloor \frac{d}{2} \rfloor + 1} \mathcal{E}[L^\alpha \phi].$$

■

5.55 Remark
This rate is not quite optimal; we expect a decay rate of $t^{(d-1)/2}$ which is $\frac{1}{2}$ better. We will deal with this in the next section. ■

5.56 Exercise
Prove Theorem 5.54. ■

Wave equation: improved decay

For the wave equation case where $M = 0$, and where the initial data has compact support, we can in fact do slightly better by replacing the vector field ∂_t in the definition of the conserved energy by the *inverted time translation* vector field

$$K \stackrel{\text{def}}{=} (t^2 + |x|^2)\partial_t + \sum_{i=1}^{d} 2tx^i \partial_{x^i}. \tag{5.57}$$

The K vector field represent a *conformal symmetry* of the Minkowski space: consider the mapping

$$(t,x) \mapsto \frac{1}{-t^2 + |x|^2}(t,x)$$

which is the Minkowski-space analogue of the conformal inversion. (On Euclidean space the conformal inversion is the mapping $x \mapsto x/|x|^2$ interchanging the interior and exterior of the unit sphere.) Note that this mapping squares to the identity. The K vector field is the pushforward of the ∂_t vector field under this mapping, and by conformality is also globally future-causal.

This is not a coincidence. On a general Lorentzian space-time, for every conformal Killing vector field X there is associated a conserved quantity called a "modified current". See, e.g. Klainerman, "A commuting vectorfields approach to Strichartz-type inequalities and applications to quasi-linear wave equations".

It turns out there is a conservation law associated to K. Define the vector field

$$^{(K)}P^\gamma \stackrel{\text{def}}{=} (m^{-1})^{\alpha\gamma}\left[Q_{\alpha\beta}K^\beta + \frac{d-1}{2}t\partial_\alpha(\phi^2) - \frac{d-1}{2}\phi^2\partial_\alpha t\right]. \tag{5.58}$$

5.59 Exercise (Modified Morawetz current)
Check that if $\Box\phi = 0$, then the vector field $^{(K)}P$ is divergence free. ■

Therefore we can integrate $\operatorname{div}{}^{(K)}P$ over a space-time region in $\mathbb{R}^{1+d}$ and use the divergence theorem to conclude that we have conserved fluxes.

Because of the presence of the lower order terms in the modified current, the positivity of the associated energy flux is not simply an application of a variant of Lemma 5.40. For convenience we will assume that we are prescribing initial data *at* $t = 2$, and the initial data is such that the support of ϕ and $\partial_t\phi$ at $t = 2$ are both contained in the ball of radius 1. By finite-speed of propagation, the space-time support of ϕ, in the half-space $\{t \geq 2\}$, is contained within the set $\{(t-1)^2 > |x|^2\}$. This implies that on each Σ_τ with $\tau \geq 2$ the solution ϕ and its derivatives have compact support, and hence we conclude the following exact conservation law without needing to argue with an approximation procedure as we did in Proposition 5.42.

5.60 Proposition

Suppose that $\Box\phi = 0$, and such that when $t = 2$, the support of $\phi(2, x)$ and $\partial_t(2, x)$ are both contained within the unit ball. Then for every $\tau \geq 2$

$$\int_{\Sigma_\tau} \left\langle \partial_\tau, {}^{(K)}P \right\rangle_m \mathrm{dvol}_{h_\tau} = \int_{t=2,|x|<1} \left\langle \partial_t, {}^{(K)}P \right\rangle_m \,\mathrm{d}x.$$

■

The term on the right is an "initial data quantity" which we will refer to as

$$\tilde{\mathcal{E}}[\phi] \stackrel{\text{def}}{=} \left(\int_{t=2,|x|<1} \left\langle \partial_t, {}^{(K)}P \right\rangle_m \,\mathrm{d}x \right)^{\frac{1}{2}}. \tag{5.61}$$

We note that $\tilde{\mathcal{E}}[\phi]$ is controlled by $\int_{t=2} (\partial_t\phi)^2 + \left|\nabla\phi\right|^2 + \phi^2$.

It remains to see what quantities are controlled by $\left\langle \partial_\tau, {}^{(K)}P \right\rangle_m$. Expanding from the definition we have

$$\left\langle \partial_\tau, {}^{(K)}P \right\rangle_m = Q(K, \partial_\tau) + \frac{d-1}{2}\tau\cosh(\rho)\partial_\tau(\phi^2) - \frac{d-1}{2}\phi^2\cosh(\rho).$$

We observe that

$$K = \tau^2\cosh(\rho)\partial_\tau + \tau\sinh(\rho)\partial_\rho,$$

with which we rewrite

$$\begin{aligned} \left\langle \partial_\tau, {}^{(K)}P \right\rangle_m &= Q(K, \partial_\tau) + \frac{d-1}{2\tau}K(\phi^2) - \frac{d-1}{2}\left[\sinh(\rho)\partial_\rho(\phi^2) + \phi^2\cosh(\rho)\right] \\ &= Q(K, \partial_\tau) + \frac{d-1}{2\tau}K(\phi^2) - \frac{d-1}{2}\partial_\rho\left[\sinh(\rho)\phi^2\right]. \end{aligned}$$

Integrating over Σ_τ with the induced measure we get

$$\int_{\Sigma_\tau} \left\langle \partial_\tau, {}^{(K)}P \right\rangle_m \mathrm{dvol}_{h_\tau} = \int_0^\infty \int_{\mathbb{S}^{d-1}} \Big(Q(K,\partial_\tau) + \frac{d-1}{2\tau} K(\phi^2) - \frac{d-1}{2} \partial_\rho \big[\sinh(\rho)\phi^2\big] \Big) \tau^d \sinh(\rho)^{d-1} \,\mathrm{d}\theta\,\mathrm{d}\rho.$$

Integrating the final term in the bracket by parts, we have that ∂_ρ hits the volume form to give

$$= \int_0^\infty \int_{\mathbb{S}^{d-1}} \Big(Q(K,\partial_\tau) + \frac{d-1}{2\tau} K(\phi^2) + \frac{(d-1)^2}{2} \cosh(\rho)\phi^2 \Big) \tau^d \sinh(\rho)^{d-1} \,\mathrm{d}\theta\,\mathrm{d}\rho.$$

The integrand we can now factor, using

$$\begin{aligned} Q(K,\partial_\tau) &= \tau^2 \cosh(\rho) Q(\partial_\tau,\partial_\tau) + \tau \sinh(\rho) Q(\partial_\tau,\partial_\rho) \\ &= \frac{1}{2}\cosh(\rho)\Big[\tau^2(\partial_\tau\phi)^2 + (\partial_\rho\phi)^2 + \frac{1}{\sinh(\rho)^2}\big|\partial_\theta\phi\big|^2\Big] \\ &\qquad + \tau\sinh(\rho)\partial_\tau\phi\partial_\rho\phi \\ &= \frac{1}{2\cosh(\rho)}\Big[(\tau\cosh(\rho)\partial_\tau\phi + \sinh(\rho)\partial_\rho\phi)^2 + (\partial_\rho\phi)^2 \\ &\qquad + \frac{\cosh(\rho)^2}{\sinh(\rho)^2}\big|\partial_\theta\phi\big|^2\Big] \\ &= \frac{1}{2\cosh(\rho)}\Big[\frac{1}{\tau^2}(K\phi)^2 + \sum_{i=1}^d (L^i\phi)^2\Big]. \end{aligned}$$

We obtain that

$$\begin{aligned} &Q(K,\partial_\tau) + \frac{d-1}{2\tau}K(\phi^2) + \frac{(d-1)^2}{2}\cosh(\rho)\phi^2 \\ &\quad = \frac{1}{2\cosh(\rho)}\sum_{i=1}^d (L^i\phi)^2 + \frac{1}{2\tau^2\cosh(\rho)}\Big(K\phi + (d-1)\tau\cosh\rho\phi\Big)^2. \end{aligned} \tag{5.62}$$

This shows

$$\int_{\Sigma_\tau} \frac{1}{2\cosh(\rho)} \sum_{i=1}^{d} (L^i \phi)^2 \,\mathrm{dvol}_{h_\tau} \leq \tilde{\mathcal{E}}[\phi]. \tag{5.63}$$

Comparing (5.63) against (5.51), we see that the estimates derived from the K vector field has a gain of additional factor of τ^2.

As a consequence, we can improve Theorem 5.48 and Theorem 5.54 to read

5.64 Theorem (Wave equation, improved decay)
Let ϕ solve $\Box\phi = 0$ and take $d \geq 2$. Suppose when $t = 2$, the support of $\phi(2,x)$ and $\partial_t\phi(2,x)$ are both contained within the unit ball. Then for every $t > 2$ we have the uniform decay estimates

$$(d-2)\left|\phi\right| + \left|L^i \phi(t,x)\right| \lesssim \frac{1}{t^{(d-1)/2}\sqrt{t-|x|}} \sum_{|\alpha| \leq \lfloor \frac{d}{2} \rfloor + 1} \tilde{\mathcal{E}}[L^\alpha \phi].$$

■

5.65 Remark
Note that not only did we recover the uniform $t^{(d-1)/2}$ decay rate for ϕ, we also have a gain of extra $t^{1/2}$ when moving toward the interior. Furthermore, recalling that L^i is essentially a "length t" derivative, we see that the above theorem guarantees that tangential derivatives decay one full t^{-1} better than the solution itself. ■

Chapter 6

Nonlinear Applications: a Tour of Wellposedness

In this chapter, we will discuss the meaning of well-posedness of an initial value problem, and apply the theory developed thus far to study the well-posedness of some nonlinear examples. For simplicity of the computations we will limit our discussion to Schrödinger equations, but one can easily imagine analogous results for other dispersive equations.

We will limit our discussion here to *semilinear* problems. The landscape of nonlinear partial differential equations can be largely classified in accordance to some measure of how nonlinear the problem is. Consider the general form of a nonlinear partial differential equation of degree k on $\mathbb{R}^d$:

$$F(x, u, Du, \ldots, D^k u) = 0 \tag{6.1}$$

where F is some smooth function of its arguments; $Du, \ldots, D^k u$ represents the arrays of higher order derivatives of u. Typically we say that the partial differential equation (6.1) is *fully nonlinear* when the dependence of F on its final group of arguments (those that corresponding to $D^k u$) is nonlinear, even with the rest of the arguments held fixed. The partial differential equation is said to be *quasilinear* when F can be schematically factored as

$$F(x, u, Du, \ldots, D^k u) = F_1(x, u, \ldots, D^{k-1} u) + F_2(x, u, \ldots, D^{k-1} u) \cdot D^k u;$$

in other words, it has linear dependence on its final argument, but the coefficient may depend on the lower order terms. Finally, the partial differential equation is said to be *semilinear* when F can be written as

$$F(x,u,Du,\cdots,D^k u) = F_1(x,u,\ldots,D^{k-1}u) + F_2(x)\cdot D^k u;$$

that is, it has linear dependence on its final argument, whose coefficients are independent of the lower order derivatives of the unknown u.

One of the main simplifications afforded by semilinear equations is that *Duhamel's principle* may be used to reformulate its corresponding initial value problem as an integral equation. And this will be our starting point of discussion.

Duhamel and the contraction mapping argument

Let us start with the formulation for ordinary differential equations. Let $X : \mathbb{R}^k \to \mathbb{R}^k$ be some linear operator. A typical problem is to study the linear, *inhomogeneous* ordinary differential equation

$$\phi'(t) = X\phi(t) + F(t) \tag{6.2}$$

and examine solutions $\phi : \mathbb{R} \to \mathbb{R}^k$; here $F : \mathbb{R} \to \mathbb{R}^k$ is some given function that we will call the "source term" or the "inhomogeneity". In the case of the homogeneous equation, where F vanishes, the solution to the corresponding equation can be written in terms of the matrix exponential

$$\phi(t) = e^{tX}\phi(0)$$

where

$$e^{tX} = \sum_{j=0}^{\infty} \frac{t^j}{j!} X^j$$

converges absolutely for all t given any fixed X.

In the inhomogeneous case, the matrix exponential can be used in method of "variation of constants", which is also called *Duhamel's principle*. Observing that

$$\left[e^{-tX}\phi(t)\right]' = e^{-tX}\phi'(t) - e^{-tX}X\phi(t),$$

we can rewrite our original equation as

$$\left[e^{-tX}\phi(t)\right]' = e^{-tX}F(t).$$

Integrating in t on both sides we get

$$e^{-tX}\phi(t) = \phi(0) + \int_0^t e^{-sX} F(s)\, \mathrm{d}s,$$

which we can rearrange to read

$$\phi(t) = e^{tX}\phi(0) + \int_0^t e^{(t-s)X} F(s)\, \mathrm{d}s. \tag{6.3}$$

In the case of ordinary differential equations, the two formulations (6.2) and (6.3) are equivalent. This allows us to reformulate the theory of semi-linear ordinary differential equations in terms of the corresponding integral equation. In particular, suppose $F : \mathbb{R}^k \to \mathbb{R}^k$ is an arbitrary (continuous) function, then solving

$$\phi'(t) = X\phi(t) + F(\phi(t))$$

is the same as solving the integral formulation

$$\phi(t) = e^{tX}\phi(0) + \int_0^t e^{(t-s)X} F(\phi(s))\, \mathrm{d}s. \tag{6.4}$$

Using this formulation, we can prove a simple version of the Picard existence theorem.

6.5 Theorem (Picard)
Given the operator X; suppose $F : \mathbb{R}^k \to \mathbb{R}^k$ is Lipschitz continuous with constant L. Then there exists a $T > 0$ such that for any ϕ_0, there exists a unique continuous function $\phi : [0, T] \to \mathbb{R}^k$ satisfying (6.4). ■

Proof Given ψ a continuous function with values in $\mathbb{R}^k$, consider the mapping

$$Q\psi(t) = e^{tX}\phi_0 + \int_0^t e^{(t-s)X} F(\psi(s))\, \mathrm{d}s.$$

Then the function ψ solves (6.4) if and only if $Q\psi = \psi$.

Considering ψ_1, ψ_2 two functions. We have that

$$\left|Q\psi_1(t) - Q\psi_2(t)\right| \leq \left|\int_0^t e^{(t-s)X}[F(\psi_1(s)) - F(\psi_2(s))]\,\mathrm{d}s\right|.$$

Writing $|X| \in \mathbb{R}$ the operator norm of X, we have

$$\left|Q\psi_1(t) - Q\psi_2(t)\right| \leq \int_0^t e^{(t-s)|X|}\left|F(\psi_1(s)) - F(\psi_2(s))\right|\,\mathrm{d}s.$$

By Lipschitz continuity we get

$$\left|Q\psi_1(t) - Q\psi_2(t)\right| \leq L\int_0^t e^{(t-s)|X|}\left|\psi_1(s) - \psi_2(s)\right|\,\mathrm{d}s.$$

Now, choose T sufficiently small such that

$$e^{T|X|} < 1 + \frac{|X|}{2L}.$$

Then, taking supremum of $t \in [0, T]$ we get

$$\sup_{t\in[0,T]}\left|Q\psi_1(t) - Q\psi_2(t)\right| \leq \frac{1}{2}\sup_{t\in[0,T]}\left|\psi_1(t) - \psi_2(t)\right|.$$

Thus Q acts as a contraction mapping on the space $C^0([0,T];\mathbb{R}^k)$, and by Banach's fixed point theorem has a unique fixed point. □

6.6 (Lipschitz dependence) Now suppose ϕ_0 and φ_0 are two different initial data. Let ϕ, φ be the corresponding solutions. We can estimate, similarly to the proof above,

$$\begin{aligned}\left|\phi(t) - \varphi(t)\right| &\leq \left|e^{tX}(\phi_0 - \varphi_0)\right| + \int_0^t e^{(t-s)|X|}\left|F(\phi(s)) - F(\varphi(s))\right|\,\mathrm{d}s \\ &\leq e^{t|X|}\left|\phi_0 - \varphi_0\right| + L\int_0^t e^{(t-s)|X|}\left|\phi(s) - \varphi(s)\right|\,\mathrm{d}s.\end{aligned}$$

by Grönwall's inequality this implies

$$\left|\phi(t)-\varphi(t)\right| \leq \left|\phi_0-\varphi_0\right| \cdot \exp\left[t|X| + \frac{L}{|X|}(e^{t|X|}-1)\right].$$

In particular, we conclude that the *solution mapping*

$$\mathbb{R}^k \ni \phi_0 \mapsto \phi \in C^0([0,T];\mathbb{R}^k)$$

is Lipschitz continuous. ¶

6.7 Remark

Grönwall's inequality in integral form states that if u, α, β are continuous functions, with $\beta \geq 0$ and α non-decreasing, such that

$$u(t) \leq \alpha(t) + \int_0^t \beta(s) u(s) \, \mathrm{d}s,$$

then

$$u(t) \leq \alpha(t) \cdot \exp\left(\int_0^t \beta(s) \, \mathrm{d}s\right).$$

■

The discussion above can be generalized to *partial differential equations*, especially in the semilinear case, by thinking of the partial differential equations as an ordinary differential equation on a Banach space. We outline the method in the case of Schrödinger equations. Instead of the linear Schrödinger equation, we would be interested in the equation

$$i\partial_t \phi = \triangle\phi + F(\phi, \overline{\phi}) \tag{6.8}$$

By focussing on a concrete equation for which we have a theory developed for its corresponding linear solutions (for data in $\mathscr{S}$), we can sidestep some of the issues of functional analysis that can come up for abstract evolution equations in Banach spaces. See Cazenave and Haraux, An introduction to semilinear evolution equations *for a more detailed treatment of these problems.*

where F is some pointwise function of the unknown ϕ and its complex conjugate $\overline{\phi}$. Formally, we can take the inhomogeneity to be the function $-iF$, and the linear operator $X = -i\triangle$. The solution operator e^{tX} is given by the solution operator $U(t)$ to the linear Schrödinger equation:

$$U(t)\phi_0 = G_t^{(\mathrm{Sch})} * \phi_0.$$

Observe that the solution operator is formally a one-parameter group, with $U(t)U(s) = U(t+s)$. So we interpret "solving (6.8) with initial data ϕ_0" to be the same as solving the corresponding integral equation

$$\phi(t) = U(t)\phi_0 - i\int_0^t U(t-s) F(\phi(s), \overline{\phi}(s)) \, \mathrm{d}s. \tag{6.9}$$

6.10 Remark (A word on the physics)
The nonlinear Schrödinger equation does not *directly* model any quantum behavior; in fact from the physical point of view, quantum evolution has to obey the principle of superposition, and cannot be nonlinear. The equation, however, arises from many physical systems. One of the main applications in which properties of nonlinear Schrödinger equations are studied is signal propagation in optical fibers. Another is the modeling of multi-particle quantum ensembles, such as the Bose-Einstein condensate, via the mean-field approximation. ■

6.11 Exercise

1. Let $X : \mathbb{R}^k \to \mathbb{R}^k$ be a linear mapping.

   (a) Derive an explicit formula for the solution to the equation

   $$\phi''(t) = X\phi(t)$$

   for $\phi : \mathbb{R} \to \mathbb{R}^k$, with initial data $\phi(0) = \phi_0$ and $\phi'(0) = \phi_1$.

   (b) Apply Duhamel's principle to obtain an integral expression for the solution of the inhomogeneous equation

   $$\phi''(t) = X\phi(t) + F(t).$$

2. Applying the result from the previous part, combined with the discussion surrounding Exercise 2.49, write down the integral form of the equation

   $$\partial_t^2 \phi(t,x) = \triangle\phi(t,x) + F(t,x),$$

   in analogy to (6.9) for the Schrödinger equation. ■

6.12 (Wellposedness) In this setting, the main questions to be addressed are those treated above for the case of the ordinary differential equation, namely the *existence* and the *uniqueness* of solutions, and whether the solutions depend *continuously* (or *in a Lipschitz fashion* when available; typically for quasilinear problems only continuity is available, while semilinear problems often allow Lipschitz continuity) on the initial data.

Unlike the case of the ordinary differential equations, where at every time t the solution lives in $\mathbb{R}^k$ and thus have essentially a canonical topology with respect to which to measure continuity, in the case of the partial differential equation the choice of function spaces is a big issue. And a theorem on wellposedness of an initial value problem requires, as part of the statement, a choice of a function space which contains the allowable initial data, as well as a choice of function space in which the solution exists. This issue will be discussed a bit later in our examples.

There is also a distinction between *local-in-time* and *global-in-time* results. The former are stated in a similar fashion as Theorem 6.5, that for any initial data there is a corresponding minimum time of existence T of the solution. The global-in-time results are precisely those in which, with the possibility of added hypotheses, the time of existence is infinite. ¶

6.13 (Local-in-time results) The key idea to proving local-in-time results is that of *approximate solutions*. Typically one builds a solution by constructing a sequence of approximate solutions such that the limiting object would be, if it exists, a bona fide solution. Many choices of approximation schemes are possible. Here we describe a few.

1. In the semilinear situation, Picard's iteration scheme, which is used in the proof of Theorem 6.5 and is based on the contraction mapping principle, is commonly used. In this scheme one solves a corresponding *linear homogeneous* problem, and builds the sequence of approximated solutions by using, as the inhomogeneity in the nth step, the approximation from the $n-1$st step. The smallness that is required for the contraction mapping is extracted from the smallness of the time of existence.

Picard's scheme is the only one we will use in these notes.

2. In the quasilinear setting, Picard's scheme is often too rough; while in the semilinear case one can get away with using the same linear solution operator, in the quasilinear case one usually needs to linearize the equation about the current approximation to best compute the next approximation (think: Newton's scheme). Convergence issues is tricky: depending on the situation sometimes one can use compactness arguments, and other times careful regularization would be required.

For a demonstration of the compactness based argument, see the proof of local wellposedness for quasilinear wave equations in Sogge, Lectures on non-linear wave equations.

3. Another method is that of the Nash-Moser implicit function theorem. Roughly speaking, with a suitable rescaling, frequently a local existence result for a nonlinear problem can be re-written as a result stating that "for all sufficiently small initial data, there exists a solution that exists up to time $T = 1$". Rewriting the small initial data as ϵ times a size 1 data, one can rewrite solving

$$F(x, u, Du) = 0$$

with initial data ϵu_0 as the family of problems

$$G(\epsilon, x, u, Du) = 0$$

For more on the Nash-Moser theorem and applications, see Alinhac and Gérard, Pseudo-differential operators and the Nash-Moser theorem.

with initial data u_0. The particularity of this scheme is that the proper choice of scale will render $G(0, x, u, Du) = 0$ a simple-to-solve equation (frequently linear). Then provided G is sufficiently "regular", one can expect to apply some version of the implicit function theorem to obtain solutions for all ϵ small.

An example of a finite-dimensional approximation is in Glimm's proof of his namesake existence theorem. A presentation is given in Hörmander, Lectures on nonlinear hyperbolic differential equations.

4. Another useful method is that of finite-dimensional approximation. Instead of proving existence directly for what appears to be an ordinary differential equation on some Banach space, one first exhausts the Banach space by an increasing sequence of finite dimensional subspaces. On each of these subspaces, one finds an approximate equation, which is a bona fide ordinary differential equation on a finite dimensional space. Then existence and uniqueness on the subspaces follows from Picard's existence theorem; the difficulty is in obtaining a uniform time of existence, and demonstrating that one can suitably interpret these solutions of the approximate equations as converging to a solution of the original partial differential equation. When these methods work, they are very amenable to numerical modeling.
5. An alternative to finite-dimensional approximation that is well suited for "hyperbolic" partial differential equations is frequency truncation. Recall by Exercise 2.11 that functions with compact frequency support are real analytic. For hyperbolic partial differential equations, when the initial data is real analytic one can formally solve the equation by Taylor expansion; the convergence of this expansion is known as the Cauchy-Kovalevskaya Theorem. Similar to the case of the finite dimensional approximation, what is required to show is that for this sequence of solutions associated to frequency-truncated initial data, we have uniform time of existence and convergence in a weaker function space norm. ¶

6.14 (Global-in-time results) Similar to the local-in-time case, there are many different ways to approach the global-in-time problem for nonlinear equations. Here we describe two classical approaches with broad applications, using the ordinary differential equation (6.4) as our base model.

1. The first approach is that of *conservation laws*. Suppose that our equation (6.4) is *Hamiltonian*; in particular, suppose for simplicity that the linear mapping X and the function F are such that

$$X\phi \perp \phi, \quad F(\phi) \perp \phi.$$

Then, returning to the differential form we see that the right hand side of the equation satisfies

$$\langle X\phi + F(\phi), \phi\rangle = 0. \tag{6.15}$$

By extension, then, $\langle \phi', \phi\rangle = 0$ and hence $|\phi|$ is constant in time. Now, returning to Theorem 6.5 for local existence, we can argue thus: Take M to be a constant $M > |\phi_0|$. Then our local existence theorem guarantees a continuous solution ϕ on the interval $[0, T]$. By our conservation law, however, $\phi(T)$ also satisfies $M > |\phi(T)|$, so we can use that as the initial data and solve up to time $2T$. By induction our solution in fact can be extended as a bounded continuous function on $[0, \infty)$, proving global existence.

This method can be further generalized to the case where one doesn't necessarily have strict conservation laws, but only some *a priori* estimates. For example, if we can show that there is a universal constant C such that for any time T and any solution $\phi : [0, T] \to \mathbb{R}^k$ of the equation that the estimate

$$\frac{\sup_{t\in[0,T]}|\phi(t)|}{\inf_{t\in[0,T]}|\phi(t)|} \leq C$$

holds, then a similar argument as above will give global existence of the solutions.

2. The second approach is that of *stability*. The stability concept we will describe here is different from the usual stability of ordinary differential equations. In the set-up of the usual stability discussion, the equation will satisfy a dissipative condition

$$\langle X\phi, \phi\rangle \leq -\epsilon|\phi|^2$$

and a condition that

$$\forall\phi \text{ s.t. } |\phi| < \delta, \quad |F(\phi)| \leq \frac{1}{2}\epsilon|\phi|.$$

Under these conditions we can show that $\langle\phi', \phi\rangle < 0$ whenever $|\phi| < \delta$, and this gives a basin of attraction around the origin, which is then a stable fixed point of the system.

The dissipative condition however is atypical for dispersive equations: by virtue of energy conservation and time reversibility, we cannot expect such decay. (The dissipation is more typical of *parabolic* partial differential equations.) In fact, one typically expects that the linear evolution satisfies $\langle X\phi, \phi\rangle = 0$. In our setup, what we can rely on is dispersive decay. In the abstract, this is a property of the interaction between the linear flow and the nonlinearity. Typically this manifests in a boundedness statement of the form

$$\int_0^t \left|e^{(t-s)X} F(\phi(s))\right| \leq \Lambda \sup_{s\in[0,t]} \left|\phi(s)\right|^2 \tag{6.16}$$

where Λ is *independent of* t. Then returning to (6.4), we see that we can estimate

$$\left|\phi(t)\right| \leq \left|\phi(0)\right| + \Lambda \sup_{\tau\in[0,t]} \left|\phi(\tau)\right|^2.$$

An inequality of this form allows us to prove the *bootstrapping inequality*:

$$\left\{ \begin{array}{c} \left|\phi(0)\right| \leq \frac{1}{3\Lambda} \\ \sup_{\tau\in[0,t]}\left|\phi(\tau)\right| \leq 3\left|\phi(0)\right| \end{array} \right\} \implies \sup_{\tau\in[0,t]} \left|\phi(\tau)\right| \leq 2\left|\phi(0)\right|. \tag{6.17}$$

This inequality states that for all initial data sufficiently small, the solution can never take values in the annulus with the inner radius $2\left|\phi(0)\right|$ and outer radius $3\left|\phi(0)\right|$. Therefore by the continuity of our solutions, for any initial data $\left|\phi_0\right| \leq (3\Lambda)^{-1}$, the corresponding solution can never escape the ball of radius $2\left|\phi_0\right|$ and hence we can apply our local existence theorem inductively to get global existence.

A major difficulty in this argument, however, is that estimates of the form (6.16) are typically *not* true for arbitrary $\phi(s)$, since it depends on analyzing how the linear flow and the nonlinearity interact. (In particular, for example, if $\phi(s)$ is chosen such that $F(\phi(s)) = e^{-sX}\psi_0$ for some ψ_0, then the estimate (6.16) cannot hold with uniform Λ.) And hiding behind the general statements above is an implicit assumption that $\phi(s)$ behaves almost like a solution to the homogeneous linear equation. The argument as a whole is necessarily perturbative. ¶

6.18 Example (Mass conservation in Schrödinger)
To illustrate the conservation law in the context of nonlinear Schrödinger equations (6.8), suppose the nonlinearity $F : \mathbb{C}^2 \to \mathbb{C}$ is such that for every

$z \in \mathbb{C}$, $F(z,\overline{z})\overline{z} \in \mathbb{R}$. Then *formally* if ϕ is a smooth (in both t and x) solution of the equation that decays suitably fast as $|x| \to \infty$, we have that

$$\begin{aligned} \partial_t \left|\phi\right|^2 &= \overline{\phi}\partial_t\phi + \phi\partial_t\overline{\phi} \\ &= -i\overline{\phi}\,[\triangle\phi + F(\phi,\overline{\phi})] + i\phi\left[\triangle\overline{\phi} + \overline{F(\phi,\overline{\phi})}\right] \\ &= 2i\,\mathrm{Im}\Big[\phi\triangle\overline{\phi} - \underbrace{F(\phi,\overline{\phi})\overline{\phi}}_{\in\mathbb{R}}\Big]. \end{aligned}$$

In the computation we used that by taking the complex conjugate of (6.8) that, $\overline{\phi}$ must satisfy the equation

$$-i\partial_t\overline{\phi} = \triangle\overline{\phi} + \overline{F(\phi,\overline{\phi})}.$$

Integrating over $\mathbb{R}^d$ we have that, after integrating by parts,

$$\int_{\mathbb{R}^d} \phi\triangle\overline{\phi}\,\mathrm{d}x = -\int_{\mathbb{R}^d} \nabla\phi\cdot\nabla\overline{\phi}\,\mathrm{d}x \in \mathbb{R}_-.$$

From this we can conclude that $\left\|\phi(t,\bullet)\right\|_{L^2}$ is constant in time. ■

6.19 Exercise (Mass conservation in Schrödinger)
The computation in the above example is only formal, since it assumes that $\partial_t\phi$ is well-defined and that the equation (6.8) is satisfied in the classical sense. In this exercise, suppose that $\phi \in C([0,T];L^2(\mathbb{R}^d))$ solves the integral equation (6.9), and F is such that $F(\phi(t),\overline{\phi}(t)) \in L^2(\mathbb{R}^d)$.

1. Prove the identity

$$\begin{aligned} \langle\phi(t),\phi(t)\rangle = \langle\phi_0,\phi_0\rangle &- \int_0^t \langle\phi(s), iF(\phi(s),\overline{\phi}(s))\rangle\,\mathrm{d}s \\ &- \int_0^t \langle iF(\phi(r),\overline{\phi}(r)),\phi(r)\rangle\,\mathrm{d}r. \qquad (6.20) \end{aligned}$$

   (*Hint*: You need to use that the solution operator for Schrödinger's equation, $U(t)$, acts by isometry on L^2.)

2. Using the identity to argue that if $F(z,\bar{z})\bar{z} \in \mathbb{R}$, then $\|\phi\|_{L^2}$ is constant in time. ■

6.21 Remark

In general the assumption that $F(\phi,\overline{\phi}) \in L^2$ will not hold for $\phi \in L^2$, the typical cases of a nonlinearity will have $|F(z,\bar{z})| \approx |z|^\alpha$ for some $\alpha > 1$; however, an integrated-in-time version of this statement will often still hold as a consequence of Strichartz estimates, and so a similar identity can in fact be justified. ■

Strichartz estimates with inhomogeneity

One of the things we saw in the discussion of the case of the ordinary differential equations is that we need to also be able to control the contributions from the inhomogeneities in (6.8). To this end we revisit the Strichartz estimates from Corollary 4.82. Observe that the Abstract Strichartz Theorem 4.77 actually gives more information: writing $U(t)$ the solution operator for the linear Schrödinger equation, the conclusion of the Corollary 4.82, together with our discussion of the endpoints, actually implies that for every $p \in (2, \frac{2d}{d-2})$ and $r = \frac{4p}{d(p-2)}$, the estimates

$$\left\|U(t)\phi_0\right\|_{L^r_t L^p_x} \lesssim \left\|\phi_0\right\|_{L^2} \tag{6.22}$$

$$\left\|\int_{\mathbb{R}} U^*(t)\Phi(t,\bullet)\,\mathrm{d}t\right\|_{L^2} \lesssim \|\Phi\|_{L^{r'}_t L^{p'}_x} \tag{6.23}$$

where p', r' are, as already defined many times, the Hölder conjugates of p and r. Now, we can chain together two such estimates with different p and r: let $p_0, p_1 \in (2, \frac{2d}{d-2})$ and r_0, r_1 defined correspondingly, we in fact have that

$$\left\|\int_{\mathbb{R}} U(t)U^*(t')\Phi(t',\bullet)\,\mathrm{d}t'\right\|_{L^{r_0}_t L^{p_0}_x} \lesssim \|\Phi\|_{L^{r_1'}_t L^{p_1'}_x}. \tag{6.24}$$

Furthermore, in view of (4.81), we also have

$$\left\|\int_0^t U(t)U^*(t')\Phi(t',\bullet)\,\mathrm{d}t'\right\|_{L^{r_0}_t L^{p_0}_x} \lesssim \|\Phi\|_{L^{r_1'}_t L^{p_1'}_x}. \tag{6.25}$$

We summarize these computations in the following proposition for solutions of inhomogeneous linear Schrödinger equations.

6.26 Proposition
Let $\phi : [0,T] \times \mathbb{R}^d \to \mathbb{C}$ satisfy

$$i\partial_t \phi = \triangle \phi + F$$

where F is some function on $[0,T] \times \mathbb{R}^d$, then for any $p_0, p_1 \in (2, \frac{2d}{d-2})$ and r_0, r_1 with $r_i = \frac{4p_i}{d(p_i - 2)}$ we have

$$\left\|\phi\right\|_{L^{r_0}([0,T];L^{p_0}(\mathbb{R}^d))} \lesssim \left\|\phi(0,\bullet)\right\|_{L^2(\mathbb{R}^d)} + \|F\|_{L^{r_1'}([0,T];L^{p_1'}(\mathbb{R}^d))}. \qquad \blacksquare$$

For convenience, we record the following technical lemma here:

6.27 Lemma (Admissible powers)
Consider the system for r_0, r_1, p_0, p_1

$$\begin{aligned} \frac{K}{r_0} &\leq \left(1 - \frac{1}{r_1}\right) = \frac{1}{r_1'} \\ \frac{K}{p_0} &= \left(1 - \frac{1}{p_1}\right) = \frac{1}{p_1'} \\ r_0 &= \frac{4p_0}{d(p_0 - 2)} \\ r_1 &= \frac{4p_1}{d(p_1 - 2)} \end{aligned} \tag{6.28}$$

where K, d are parameters.

1. The system has *no solutions* when $K > 1 + \frac{4}{d}$.
2. When $K = 1 + \frac{4}{d}$, for every p_0, the system can be solved, with the first inequality necessarily an equality.
3. When $K < 1 + \frac{4}{d}$, for every p_0, the system can be solved, with the first inequality necessarily strict.

Furthermore, in the latter two cases, if additionally $K > 1$, the solution p_0 and p_1 can be chosen to lie within the admissible interval for Proposition 6.26. $\blacksquare$

Proof The relation between r_i and p_i can be rewritten as

$$\frac{1}{r_i} = \frac{d}{4} - \frac{d}{2p_i}.$$

Using that $\frac{1}{p_1} = 1 - \frac{k}{p_0}$ we get

$$\frac{1}{r_1} = \frac{d}{4} - \frac{d}{2} + \frac{d}{2}\frac{k}{p_0}$$

so that

$$1 - \frac{1}{r_1} = 1 + \frac{d}{4} - \frac{d}{2}\frac{k}{p_0}.$$

Similarly we have

$$\frac{k}{r_0} = \frac{kd}{4} - \frac{d}{2}\frac{k}{p_0}.$$

From this we conclude that the existence of a solution to the system forces, by the first inequality, that $k \leq 1 + \frac{4}{d}$.

To show the existence of compatible solutions, it suffices to observe that if $p_0, p_1 \in (2, \frac{2d}{d-2})$, we have $(p_0)^{-1}, (p_1)^{-1} \in (\frac{1}{2} - \frac{1}{d}, \frac{1}{2})$. And hence we can arrange for $[1-(p_1)^{-1}]/(p_0)^{-1}$ to take any value between $(1, (\frac{1}{2}+\frac{1}{d})/(\frac{1}{2}-\frac{1}{d}))$. The upper bound of this interval is strictly larger than $1 + \frac{4}{d}$. And hence for any $K \in (1, 1+\frac{4}{d}]$ we can find p_0, p_1 satisfying the second expression in the system, for which our computations above indicate that the corresponding r_0 and r_1 will satisfy the first inequality. □

Examples of wellposedness results

We will concentrate on the case of the semilinear Schrödinger equation with power nonlinearity. These are the equations

$$i\partial_t \phi = \triangle\phi + P_K(\phi, \overline{\phi}) \tag{6.29}$$

where P_K is a homogeneous polynomial of degree K. (For example, $P_3(\phi, \overline{\phi})$ can be written in the form

$$a_0\phi^3 + a_1\phi^2\overline{\phi} + a_2\phi\overline{\phi}^2 + a_3\overline{\phi}^3$$

where $a_0, \ldots, a_3$ are complex constants.) As we will see, the wellposedness of these equations depends on

1. the number of spatial dimensions d;
2. the degree of the polynomial K;
3. in which function space are we considering the initial data and the solution.

Let's start with local wellposedness.

6.30 Theorem (L^2-LWP for Schrödinger; sub-critical case)
Suppose $K \in (1, 1 + \frac{4}{d})$. Then for every $M \geq 0$ there exists a $T > 0$ such that for every initial data $\phi_0 \in L^2(\mathbb{R}^d)$ with $\|\phi_0\|_{L^2} < M$, there exists a solution $\phi \in C([0,T]; L^2(\mathbb{R}^d))$ to (6.29) (in the sense of (6.9)) satisfying $\phi(0) = \phi_0$. ∎

Proof We proceed by Picard iteration using Strichartz estimates.

By Lemma 6.27, there exists p_0, p_1, r_0, r_1 with $p_0, p_1 \in (2, \frac{2d}{d-2})$ such that the system (6.28) can be solved with the first inequality strict. Let $\mathfrak{X} = C([0,T]; L^2(\mathbb{R}^d)) \cap L^{r_0}([0,T]; L^{p_0}(\mathbb{R}^d))$, where the T remains to be determined. Consider the operator

$$Q\psi(t) = U(t)\phi_0 + \int_0^t U(t)U^*(s) P_K(\psi(s), \overline{\psi(s)}) \,\mathrm{d}s.$$

It suffices to show that Q is a contraction mapping. First we show that $Q : \mathfrak{X} \to \mathfrak{X}$. By our Strichartz inequality Proposition 6.26 we have that

$$\|Q\psi\|_{\mathfrak{X}} = \max(\|Q\psi\|_{L^\infty_t L^2}, \|Q\psi\|_{L^{r_0}_t L^{p_0}}) \lesssim \|\phi_0\|_{L^2} + \|P_K(\psi, \overline{\psi})\|_{L^{r_1'}_t L^{p_1'}}.$$

Now, by definition of the inhomogeneity P_K, we have

$$\|P_K(\psi, \overline{\psi})\|_{L^{r_1'}_t L^{p_1'}} \lesssim \|\psi\|^K_{L^{Kr_1'}_t L^{Kp_1'}}.$$

Our original choice of p_0, p_1 satisfies (6.28) with the first inequality being strict, and therefore $Kp_1' = p_0$ and $Kr_1' < r_0$. Using that $T < \infty$, by Hölder's inequality we then have that, for some $\epsilon > 0$ depending on p_0 and p_1,

$$\|P_K(\psi, \overline{\psi})\|_{L^{r_1'}_t L^{p_1'}} \lesssim T^{\epsilon K} \cdot \|\psi\|^K_{L^{r_0}_t L^{p_0}}.$$

This shows that

$$\|Q\psi\|_{\mathfrak{X}} \lesssim \|\phi_0\|_{L^2} + T^{\epsilon K} \|\psi\|^K_{\mathfrak{X}}$$

and indeed Q maps $\mathfrak{X}$ to itself. Furthermore, letting C denote the implicit constant (which is independent of ϕ_0 or ψ, or the choice of T) in the inequality above, and taking $\tilde{M} = 2CM$, we see that for T such that $T^{\epsilon K} C\tilde{M}^{K-1} < \frac{1}{2}$, that Q in fact maps the ball of radius $\tilde{M}$ in $\mathfrak{X}$ to itself.

It remains to prove the contraction mapping property. Let $\psi_1, \psi_2 \in \mathfrak{X}$, both with norm bounded by $\tilde{M}$. We have that

$$Q\psi_1 - Q\psi_2 = \int_0^t U(t)U^*(s)\left[P_K(\psi_1(s), \overline{\psi_1(s)}) - P_K(\psi_2(s), \overline{\psi_2(s)})\right] \,\mathrm{d}x.$$

So we also have, as above,

$$\left\|Q\psi_1 - Q\psi_2\right\|_{\mathfrak{X}} \lesssim \left\|P_K(\psi_1, \overline{\psi_1}) - P_K(\psi_2, \overline{\psi_2})\right\|_{L_t^{r_1'} L^{p_1'}}. \tag{6.31}$$

Since P_K is a homogeneous polynomial of degree K, we have that

$$\left|P_K(\psi_1, \overline{\psi_1}) - P_K(\psi_2, \overline{\psi_2})\right| \lesssim \left(\left|\psi_1\right|^{K-1} + \left|\psi_2\right|^{K-1}\right) \cdot \left|\psi_1 - \psi_2\right|.$$

And Hölder's inequality implies

$$\left\|P_K(\psi_1, \overline{\psi_1}) - P_K(\psi_2, \overline{\psi_2})\right\|_{L_t^{r_1'} L^{p_1'}} \lesssim \left(\left\|\psi_1\right\|^{K-1}_{L_t^{Kr_1'} L^{Kp_1'}} + \left\|\psi_2\right\|^{K-1}_{L_t^{Kr_1'} L^{Kp_1'}}\right) \cdot \left\|\psi_1 - \psi_2\right\|_{L_t^{Kr_1'} L^{Kp_1'}}.$$

As before, the choice of p_0, p_1 allows us to write this in the form

$$\left\|Q\psi_1 - Q\psi_2\right\|_{\mathfrak{X}} \lesssim T^{\epsilon K} \tilde{M}^{K-1} \cdot \left\|\psi_1 - \psi_2\right\|_{\mathfrak{X}}$$

where the implicit constant is independent of T, ψ_1, or ψ_2. And hence taking T sufficiently small (compared to structural constants and the value $\tilde{M}$) we can guarantee that Q acts as a contraction mapping on the (closed) ball of radius $\tilde{M}$ in $\mathfrak{X}$. This concludes our proof. □

Observe that as a consequence of our proof, we have that the solution constructed is such that $P_K(\phi, \overline{\phi}) \in L_t^{r_1'} L^{p_1'}$, while Proposition 6.26 implies that $\phi \in L_t^{r_1} L^{p_1}$ (in addition to being in $L_t^{r_0} L^{p_0}$. This means that we can interpret the identity (6.20) with the inhomogeneity coming in through the space-time pairing of $L_t^{r_1} L^{p_1}$ with its dual. And in particular, if $P_K(z, \overline{z})\overline{z} \in \mathbb{R}$ for all $z \in \mathbb{C}$, we can conclude that the L^2 norm of the solution is constant in time.

6.32 Corollary (Subcritical L^2-GWP by mass conservation)
Let $K \in (1, 1 + \frac{4}{d})$, and suppose P_K is such that $P_K(z, \bar{z})\bar{z} \in \mathbb{R}$ for all $z \in \mathbb{C}$. Then for every $\phi_0 \in L^2(\mathbb{R}^d)$, there exists a solution $\phi \in C(\mathbb{R}; L^2(\mathbb{R}^d))$ to (6.29) satisfying $\phi(0) = \phi_0$. ■

Proof Let $M > \|\phi_0\|_{L^2}$. By the previous theorem there exists $T > 0$ depending on M such that a solution exists in $C([0,T]; L^2(\mathbb{R}^d))$. Note that for this solution $\|\phi(t)\|_{L^2} < M$ still, and hence the solution can be extended to one that exists on the time interval $[0, 2T]$. By time reversibility of the Schrödinger equation and induction, we can cover the whole of $\mathbb{R}$. □

Furthermore, note that for both of the above results, the same argument given for the Picard theorem Theorem 6.5 and Thought 6.6 implies that the solution is unique, and depends Lipschitz-continuously on the initial data.

Next let us treat the case $K = 1 + \frac{4}{d}$. The main difference in this case is that, whereas in the case $K < 1 + \frac{4}{d}$ we have a time of existence that depends only on the L^2 norm of the initial data, here the dependence is more subtle. (When reading the theorem, pay attention to the order of quantifiers.)

6.33 Theorem (L^2-existence for Schrödinger; critical case)
Suppose $K = 1 + \frac{4}{d}$. Then for every $\phi_0 \in L^2(\mathbb{R}^d)$ there exists a $T > 0$ and a solution $\phi \in C([0,T]; L^2(\mathbb{R}^d))$ to (6.29) satisfying $\phi(0) = \phi_0$. ■

Proof We have to set-up and run our iteration slightly differently.

By Lemma 6.27, there exists p_0, p_1, r_0, r_1 with $p_0, p_1 \in (2, \frac{2d}{d-2})$ such that the system (6.28) can be solved with the first inequality being an equality. Define as before $\mathfrak{X} = C([0,T]; L^2(\mathbb{R}^d)) \cap L^{r_0}([0,T]; L^{p_0}(\mathbb{R}^d))$ with T to be determined. This time we consider the operator

$$Q\psi(t) = \int_0^t U(t)U^*(s)P_K(U(s)\phi_0 + \psi(s), \overline{U(s)\phi_0 + \psi(s)}) \,\mathrm{d}s$$

then if ψ is a fixed point of Q we have $U(t)\phi_0 + \psi(t)$ is the solution that we seek.

By our Strichartz inequality we have that

$$\|Q\psi\|_{\mathfrak{X}} \lesssim \|P_K(U(\bullet)\phi_0 + \psi, \overline{U(\bullet)\phi_0 + \psi})\|_{L_t^{r_1'} L^{p_1'}}.$$

Since P_K is polynomial, we can bound, analogously to before,

$$\|P_K(U(\bullet)\phi_0 + \psi, \overline{U(\bullet)\phi_0 + \psi})\|_{L_t^{r_1'} L^{p_1'}} \lesssim \|U(\bullet)\phi_0\|_{L_t^{r_0} L^{p_0}}^K + \|\psi\|_{L_t^{r_0} L^{p_0}}^K,$$

where we used the fact that by construction $Kr_1' = r_0$ and $Kp_1' = p_0$. Now, by Strichartz we know that $\|U(\bullet)\phi_0\|_{L^{r_0}(\mathbb{R};L^{p_0}(\mathbb{R}^d))} \lesssim \|\phi_0\|_{L^2}$, so we know that by taking T small we can make $\|U(\bullet)\phi_0\|_{L^{r_0}([0,T];L^{p_0}(\mathbb{R}^d))}$ as small as we want. Therefore examining the bound

$$\|Q\psi\|_{\mathfrak{X}} \leq C\left(\|U(\bullet)\phi_0\|^K_{L^{r_0}_t L^{p_0}} + \|\psi\|^K_{L^{r_0}_t L^{p_0}}\right)$$

where the constant C depends on the universal constant in Strichartz estimates (which depend on the dimension d, the powers r_0, r_1, p_0, p_1) as well as the structural constants in the definition of P_K, we can first choose ϵ_0 sufficiently small such that $(\epsilon_0)^{K-1}C < \frac{1}{2}$ and then T sufficiently small such that $C\|U(\bullet)\phi_0\|^K_{L^{r_0}([0,T];L^{p_0}(\mathbb{R}^d))} < \frac{1}{2}\epsilon_0$. Then the above inequality implies that Q maps the ball of radius ϵ_0 in $\mathfrak{X}$ to itself.

A similar argument, starting from the fact that

$$\|Q\psi_1 - Q\psi_2\|_{\mathfrak{X}} \lesssim \left(\|U(\bullet)\phi_0\|^{K-1}_{L^{r_0}_t L^{p_0}} + \|\psi_1\|^{K-1}_{L^{r_0}_t L^{p_0}} + \|\psi_2\|^{K-1}_{L^{r_0}_t L^{p_0}}\right)\|\psi_1 - \psi_2\|_{L^{r_0}_t L^{p_0}},$$

shows we can choose ϵ_0 small enough, and T small enough, such that Q acts as a contraction mapping on the ball of radius ϵ_0 in $\mathfrak{X}$, thereby proving the theorem. □

6.34 Remark

The fact that we used the Banach fixed-point theorem in the proof guarantees that the solution is unique. However, unlike the previous subcritical case, we do not have Lipschitz-dependence on the initial data. In fact, from the proof given here it is not even guaranteed that the time of existence T depends continuously on ϕ_0 in the $L^2(\mathbb{R}^d)$ topology! ■

Observe the only limit on T is in making $\|U(\bullet)\phi_0\|_{L^{r_0}([0,T];L^{p_0}(\mathbb{R}^d))}$ small enough to run the contraction mapping argument. One can alternatively make the norm $\|U(\bullet)\phi_0\|_{L^{r_0}(\mathbb{R};L^{p_0}(\mathbb{R}^d))}$ sufficiently small by imposing a smallness condition on ϕ_0. This implies the following global theorem:

6.35 Theorem (Critical small data L^2-GWP)

Suppose $K = 1 + \frac{4}{d}$. There exists $\epsilon > 0$ such that for every $\phi_0 \in L^2(\mathbb{R}^d)$ satisfying $\|\phi_0\|_{L^2} < \epsilon$, there exists a solution $\phi \in C(\mathbb{R}; L^2(\mathbb{R}^d))$ to (6.29) satisfying $\phi(0) = \phi_0$. ■

6.36 Remark
Note that in practice, for $K \in \mathbb{N}$, the case $K = 1 + \frac{4}{d}$ only occurs for $K = 2$ and $d = 4$, or $K = 3$ and $d = 2$, or $K = 5$ and $d = 1$. ■

6.37 Remark
The global theorem in the critical case, in contrast to the local theorem, is indeed a wellposedness result. In addition to existence and uniqueness of the solution, since the space in which the iteration argument is run is fixed (by the small data limit ϵ) and does not depend on the choice of the data (which is the case in the local theorem), the same argument as in the case of the ordinary differential equation case Theorem 6.5 guarantees Lipschitz dependence on the initial data.

Interestingly, the proof for Theorem 6.35 cannot easily go through when $K < 1 + \frac{4}{d}$. The fact that $Kr_1' < r_0$, which is useful in allowing us to use a Hölder inequality to exploit the small time interval in the local wellposedness theorem, turns out to be detrimental in proving any sort of global existence. ■

6.38 Exercise
Prove Theorem 6.35. ■

As mentioned before, the wellposedness results depend on a choice of Banach spaces. In the previous few theorems we have fixed $L^2(\mathbb{R}^d)$ as the basic space, but other choices are also possible. One particularly useful one is that of $H^1(\mathbb{R}^d)$. Start again with (6.29), we want to treat it as an evolution equation in the space $H^1(\mathbb{R}^d)$. For this purpose we will consider the *system* of equations

$$\begin{aligned} i\partial_t \phi &= \triangle\phi + P_K(\phi, \overline{\phi}) \\ i\partial_t(\nabla\phi) &= \triangle(\nabla\phi) + Q_{K-1}(\phi, \overline{\phi}) \cdot (\nabla\phi, \nabla\overline{\phi}) \end{aligned} \tag{6.39}$$

where the second equation is obtained from taking the spatial gradient of the first one (and hence Q_{K-1} is a vector-valued homogeneous polynomial of degree $K-1$ that is obtained by taking the derivative of P_K). Treating $\Phi = (\phi, \nabla\phi)$ the vector, we can think of the evolution equation for ϕ in $C([0,T]; H^1(\mathbb{R}^d))$ as the same as the evolution equation for Φ in $C([0,T]; L^2(\mathbb{R}^d))$.

What we gain from the additional derivative in H^1 is *Sobolev's inequality*: if we can measure both ϕ and $\nabla\phi$ in the Strichartz norm $L^r([0,T]; L^p(\mathbb{R}^d))$, then Sobolev's inequality means that we can measure ϕ in the mixed norm $L^r([0,T]; L^q(\mathbb{R}^d))$ for any $q \in [p, \frac{dp}{d-p}]$ when $p < d$ (and similarly for the appropriate exponents when $p \geq d$).

Now let us examine the second of the equations in (6.39). Define the mapping

$$Q\psi(t) = U(t)\phi_0 + \int_0^t U(t)U^*(s)P_K(\psi,\overline{\psi})\,\mathrm{d}s.$$

We have as a consequence

$$\nabla Q\psi(t) = U(t)(\nabla\phi_0) + \int_0^t U(t)U^*(s)Q_{K-1}(\psi,\overline{\psi})\cdot(\nabla\psi,\nabla\overline{\psi})\,\mathrm{d}s.$$

Applying Proposition 6.26 to it, we get the estimate

$$\left\|\nabla Q\psi\right\|_{L_t^{r_0}L^{p_0}} \lesssim \left\|\nabla\phi\right\|_{L^2} + \left\|Q_{K-1}(\psi,\overline{\psi})\cdot(\nabla\psi,\nabla\overline{\psi})\right\|_{L_t^{r_1'}L^{p_1'}}. \tag{6.40}$$

As in the proof of the theorems above, we seek to estimate the final term in (6.40) in terms of the Strichartz norm on the left. Observe that

$$\left\|Q_{K-1}(\psi,\overline{\psi})\cdot(\nabla\psi,\nabla\overline{\psi})\right\|_{L_t^{r_1'}L^{p_1'}} \lesssim \left\|\left|\psi\right|^{K-1}\cdot\left|\nabla\psi\right|\right\|_{L_t^{r_1'}L^{p_1'}}.$$

By the Gagliardo-Nirenberg-Sobolev inequality we have that $\left\|\psi\right\|_{L^{\frac{dp_0}{d-p_0}}(\mathbb{R}^d)} \lesssim \left\|\nabla\psi\right\|_{L^{p_0}}$. So we have that, as long as

$$(K-1)\cdot\left(\frac{1}{p_0}-\frac{1}{d}\right)+\frac{1}{p_0}=\frac{1}{p_1'}, \tag{6.41}$$

by Hölder's inequality we have

$$\left\|\left|\psi\right|^{K-1}\cdot\left|\nabla\psi\right|\right\|_{L^{p_1'}} \leq \left\|\psi\right\|^{K-1}_{L^{\frac{K-1}{1/p_1'-1/p_0}}}\left\|\nabla\psi\right\|_{L^{p_0}} \lesssim \left\|\nabla\psi\right\|^K_{L^{p_0}}.$$

The relation (6.41) is the replacement of the second line of (6.28) for the case where we work with H^1 instead of L^2.

Similarly, under (6.41) we can also check that

$$\left\|\phi^K\right\|_{L^{p_1'}} \lesssim \left\|\phi\right\|^K_{W^{1,p_0}}.$$

Together we conclude that

$$\left\|Q\psi\right\|_{L_t^{r_0}W^{1,p_0}} \lesssim \left\|\phi_0\right\|_{H^1} + \left\|\psi\right\|^K_{L_t^{Kr_1'}W^{1,p_0}} \tag{6.42}$$

and therefore as long as

$$K r_1' \leq r_0$$

(as in (6.28) also) we can run the argument exactly as in the L^2 cases. The solvability of the system of exponents can be summarised in the following lemma, the analogue of Lemma 6.27.

6.43 Lemma

The system, for $d \geq 3$,

$$\begin{aligned} \frac{K}{r_0} &\leq \left(1 - \frac{1}{r_1}\right) = \frac{1}{r_1'} \\ \frac{K}{p_0} - \frac{K-1}{d} &= \left(1 - \frac{1}{p_1}\right) = \frac{1}{p_1'} \\ r_0 &= \frac{4p_0}{d(p_0 - 2)} \\ r_1 &= \frac{4p_1}{d(p_1 - 2)} \end{aligned}$$

can be solved if and only if $K \leq 1 + \frac{4}{d-2} = \frac{d+2}{d-2}$, with equality if and only if the first line in the system above evaluates to equality. When $K > 1$ additionally, the solution can be chosen with $p_0, p_1 \in (2, \frac{2d}{d-2})$. ∎

Putting together the above discussion and emulating the proofs in the L^2 case, we obtain the following theorems concerning the solvability of the initial value problem for (6.29) with initial data in $H^1(\mathbb{R}^d)$.

6.44 Theorem (H^1-LWP for Schrödinger; subcritical case)

Suppose $K \in (1, \frac{d+2}{d-2})$, and $d \geq 3$. The for every $M \geq 0$ there exists $T > 0$ such that for every initial data $\phi_0 \in H^1(\mathbb{R}^d)$ with $\|\phi_0\|_{H^1} < M$, there exists a solution $\phi \in C([0,T]; H^1(\mathbb{R}^d))$ to (6.29) satisfying $\phi(0) = \phi_0$. Furthermore, the solution is unique, and depends Lipschitz-continuously on the initial data. ∎

6.45 Exercise

Prove Theorem 6.44. ∎

6.46 Exercise

Above we stated Theorem 6.44 only for $d \geq 3$. Formulate and prove the correct, analogous statements for $d = 1, 2$. (Note: when $d = 1, 2$, and $p > 2$, the Sobolev embedding theorem tells us that $W^{1,p} \hookrightarrow L^q$ for any $q \in [p, \infty)$.) ∎

6.47 Theorem (H^1-existence for Schrödinger; critical case)
Suppose $d \geq 3$ and $K = \frac{d+2}{d-2}$. For every $\phi_0 \in H^1(\mathbb{R}^d)$, there exists a $T > 0$ and a function $\phi \in C([0,T]; H^1(\mathbb{R}^d))$ solving (6.29) and satisfying $\phi(0) = \phi_0$. ■

6.48 Theorem (Critical small data H^1-GWP)
Suppose $d \geq 3$ and $K = \frac{d+2}{d-2}$. There exists $\epsilon > 0$ such that for every $\phi_0 \in H^1(\mathbb{R}^d)$ with $\|\phi_0\|_{H^1} < \epsilon$, there exists a solution $\phi \in C(\mathbb{R}; L^2(\mathbb{R}^d))$ to (6.29) satisfying $\phi(0) = \phi_0$. The solution is unique and depends Lipschitz-continuously on the initial data. ■

6.49 Exercise
Prove Theorem 6.47 and Theorem 6.48. ■

We conclude this tour of the wellposedness results with a discussion of the global wellposedness of (6.29) in the subcritical case of $K \in (1, \frac{d+2}{d-2})$. Here the fact that the time of existence T in the local wellposedness Theorem 6.44 depends only on the H^1 norm of the initial data allows us to hope for some sort of conservation law argument, similar to Corollary 6.32 in the L^2 case. In addition to the conservation of L^2 norm as discussed in Example 6.18, we need the following additional structure on the nonlinearities.

6.50 (Energy conservation; formal computations) Starting with (6.29), we can multiply the equation with $\partial_t \overline{\phi}$ to obtain

$$i\left|\partial_t \phi\right|^2 = \triangle\phi \partial_t \overline{\phi} + P_K(\phi, \overline{\phi})\partial_t \overline{\phi}.$$

Adding to its complex conjugate we have

$$0 = \triangle\phi\partial_t\overline{\phi} + \triangle\overline{\phi}\partial_t\phi + P_K(\phi,\overline{\phi})\partial_t\overline{\phi} + \overline{P_K(\phi,\overline{\phi})}\partial_t\phi.$$

Integrating by parts we get

$$0 = -\int_{\mathbb{R}^d} \partial_t\left|\nabla\phi\right|^2 + P_K(\phi,\overline{\phi})\partial_t\overline{\phi} + \overline{P_K(\phi,\overline{\phi})}\partial_t\phi \,\mathrm{d}x.$$

Now, if $P_K(z,\overline{z})\,\mathrm{d}\overline{z} + \overline{P_K(z,\overline{z})}\,\mathrm{d}z$ is an exact form, we can find some $V : \mathbb{C}^2 \to \mathbb{C}$ such that

$$0 = -\int_{\mathbb{R}^d} \partial_t\left[\left|\nabla\phi\right|^2 + V(\phi,\overline{\phi})\right] \mathrm{d}x$$

and hence we can conclude that

$$\int_{\mathbb{R}^d} \left|\nabla\phi\right|^2 + V(\phi,\overline{\phi})\,\mathrm{d}x$$

is a conserved quantity.

If additionally we have that V takes values only in $\mathbb{R}_+$, then this conservation law would give us a global *a priori* bound on $\|\phi(t,\bullet)\|_{H^1}$ by the initial data, which would allow us to conclude, as a corollary of our local wellposedness statement Theorem 6.44 that the corresponding solution can in fact extend globally to a function in $C(\mathbb{R};H^1(\mathbb{R}^d))$.

It is worth noting that the condition of $P_K(z,\bar{z})\,\mathrm{d}\bar{z} + \overline{P_K(z,\bar{z})}\,\mathrm{d}z$ being an exact form is independent of the condition that $P_K(z,\bar{z})\bar{z} \in \mathbb{R}$, which was used in L^2 conservation. To wit: the function $(z,\bar{z}) \mapsto z\bar{z} + z^2$ satisfy the latter condition, but

$$\mathrm{d}\Big[(z\bar{z}+z^2)\,\mathrm{d}\bar{z} + (z\bar{z}+\bar{z}^2)\,\mathrm{d}z\Big] = (z-\bar{z})\,\mathrm{d}z\wedge\mathrm{d}\bar{z} \not\equiv 0.$$

Similarly, the function $(z,\bar{z}) \mapsto 2z\bar{z} + z^2$ fails the condition for L^2 conservation, but the corresponding one form

$$(2z\bar{z}+z^2)\,\mathrm{d}\bar{z} + (2z\bar{z}+\bar{z}^2)\,\mathrm{d}z = \mathrm{d}(z^2\bar{z}+\bar{z}^2 z)$$

is exact.

A particular family of nonlinearities that exhibit both conservation laws is given by

$$P_K(z,\bar{z}) = \lambda|z|^{K-1}z, \qquad \lambda\in\mathbb{R}. \tag{6.51}$$

The corresponding potentials are

$$V = -\frac{2\lambda}{K+1}|z|^{K+1}. \tag{6.52}$$

Note that when $\lambda < 0$ we have that the conserved energy is *coercive*. ¶

6.53 Corollary (H^1-GWP; subcritical defocussing)
Let $d \geq 3$ and $K \in (1, \frac{d+2}{d-2})$. Then for every $\phi_0 \in H^1(\mathbb{R}^d)$, there exists a unique solution $\phi \in C(\mathbb{R};H^1(\mathbb{R}^d))$ to the equation

$$i\partial_t\phi = \triangle\phi - \left|\phi\right|^{K-1}\phi$$

with $\phi(0) = \phi_0$. ■

Proof Noting that $1 + \frac{d+2}{d-2} = \frac{2d}{d-2}$, we have that for initial data $\phi_0 \in H^1(\mathbb{R}^d)$, both $\left|\nabla\phi_0\right|^2$ and $\left|\phi_0\right|^{K+1}$ are integrable, and so the conserved energy is finite and well-defined, using also the computations in Thought 6.50. Therefore, by iterating Theorem 6.44 and using time reversibility of the equation, we obtain a globally existing solution. □

In the above corollary we used that the conserved energy is coercive on the $\mathring{H}^1$ norm of the solution ϕ. The next corollary shows that even in the case where the conserved energy is not globally coercive, sometimes one can still exploit some sort of "small data coercivity" to get the desired result. In the statement a lower bound on the admissible K is given: this is not too surprising as in the small data regime the higher the power of the nonlinearity, the smaller the resulting perturbation. Notice however that the lower endpoint of $K = 1 + \frac{4}{d}$ is precisely where we have small data global existence using merely the L^2 theory, and below that the L^2 theory give global existence for even large initial data by mass conservation.

6.54 Corollary (Small data H^1-GWP; subcritical focussing) Let $d \geq 3$ and $K \in (1+\frac{4}{d}, \frac{d+2}{d-2})$. Then there exists $\epsilon > 0$ such that for every $\phi_0 \in H^1(\mathbb{R}^d)$ with $\|\phi_0\|_{H^1} < \epsilon$, there exists a unique solution $\phi \in C(\mathbb{R}; H^1(\mathbb{R}^d))$ to the equation

$$i\partial_t \phi = \triangle\phi + |\phi|^{K-1}\phi$$

with $\phi(0) = \phi_0$. ■

Proof Denote by

$$E = \int_{\mathbb{R}^d} |\nabla\phi|^2 - \frac{2}{K+1}|\phi|^{K+1} \,\mathrm{d}x$$

the conserved energy for our equation. Note that this quantity is *not* generally coercive on the $\mathring{H}^1$ norm of ϕ. What we will prove, however, is the small-data coercivity of the conserved energy. Observe that by the Gagliardo-Nirenberg-Sobolev inequality we have

$$\|\phi\|_{L^{K+1}} \lesssim \|\phi\|_{L^2}^{1-\theta} \|\nabla\phi\|_{L^2}^{\theta}$$

where

$$\frac{1}{K+1} = \frac{1-\theta}{2} + \theta \cdot \left(\frac{1}{2} - \frac{1}{d}\right) = \frac{1}{2} - \frac{\theta}{d}.$$

When $K \in (1 + \frac{4}{d}, 1 + \frac{4}{d-2})$, we can check that $\theta \cdot (K+1) > 2$ as a result. This implies that for some $\eta \in (2, K+1)$ we have

$$\int_{\mathbb{R}^d} |\phi|^{K+1} \,\mathrm{d}x \lesssim \|\phi\|_{L^2}^{K+1-\eta} \|\nabla\phi\|_{L^2}^{\eta}.$$

Now, by assumption the L^2 norm of our solution is conserved. So for $\|\phi_0\|_{L^2}$ sufficiently small, the above computation implies that

$$E \geq \|\nabla\phi\|_{L^2}^2 - \frac{1}{2}\|\nabla\phi\|_{L^2}^\eta.$$

Since E is conserved, for initial $\|\nabla\phi\|_{L^2}$ small, E will remain small for all time. When $E \approx 0$, either $\|\nabla\phi\|_{L^2} \approx 0$ or $\|\nabla\phi\|_{L^2}^{\eta-2} > 2 - \delta$. Since $\eta > 2$ the two regimes are disjoint. So by continuity of $\|\phi(t)\|_{H^1}$ if the initial data is such that $\|\nabla\phi\|_{L^2}$ is small, the same holds for the solution. This established the *a priori* boundedness of the H^1 norm of the solution, and hence by our previous results we obtain global existence of the solution. □

6.55 Exercise
Formulate and prove the analogues of the two above corollaries for the cases when $d = 1, 2$. ■

It turns out that the small-data assumption in the final corollary above is necessary. In fact, for large initial data we can prove that solutions *cannot* exist for all time. This final result is based on an argument originally due to Glassey.

6.56 Proposition (Glassey's Virial Identity)
Let $K \geq 1 + \frac{4}{d}$, and let ϕ solve

$$i\partial_t\phi = \triangle\phi + |\phi|^{K-1}\phi.$$

Denote by

$$\mathcal{V}(t) = \int_{\mathbb{R}^d} |x|^2|\phi|^2 \, \mathrm{d}x;$$

then

$$\partial_{tt}^2\mathcal{V} \leq 8E$$

where E is the conserved energy. ■

Proof Directly taking the derivatives we get

$$\partial_t \mathcal{V} = \int_{\mathbb{R}^d} |x|^2 [\overline{\phi}\partial_t\phi + \phi\partial_t\overline{\phi}] \,\mathrm{d}x$$
$$= -i\int_{\mathbb{R}^d} |x|^2 [\overline{\phi}\triangle\phi - \phi\triangle\overline{\phi}] \,\mathrm{d}x$$
$$= 2i\int_{\mathbb{R}^d} x\cdot[\overline{\phi}\nabla\phi - \phi\nabla\overline{\phi}] \,\mathrm{d}x.$$

Now take another derivative and repeatedly integrate by parts

$$\partial^2_{tt}\mathcal{V} = 2\int_{\mathbb{R}^d} x\cdot[i\overline{\phi}_t\nabla\phi + \overline{\phi}\nabla i\phi_t - i\phi_t\nabla\overline{\phi} - \phi\nabla i\overline{\phi}_t] \,\mathrm{d}x$$
$$= 2\int_{\mathbb{R}^d} x\cdot[2i\overline{\phi_t}\nabla\phi - 2i\phi_t\nabla\overline{\phi}] - d\overline{\phi}i\phi_t + d\phi i\overline{\phi}_t \,\mathrm{d}x$$
$$= -2\int_{\mathbb{R}^d} 2x\cdot\Big[\triangle\overline{\phi}\nabla\phi + \triangle\phi\nabla\overline{\phi} + \big|\phi\big|^{K-1}\nabla\big|\phi\big|^2\Big]$$
$$+ d\Big[\overline{\phi}\triangle\phi + \phi\triangle\overline{\phi} + 2\big|\phi\big|^{K+1}\Big] \,\mathrm{d}x$$
$$= 4\int_{\mathbb{R}^d} 2\big|\nabla\phi\big|^2 - d\frac{K-1}{K+1}\big|\phi\big|^{K+1} \,\mathrm{d}x$$
$$= 8E - \frac{4d[K-(1+\frac{4}{d})]}{K+1}\int_{\mathbb{R}^d}\big|\phi\big|^{K+1} \,\mathrm{d}x.$$

The final term is signed when $K \geq 1 + \frac{4}{d}$. □

6.57 Theorem
Let $K \geq 1 + \frac{4}{d}$, and suppose $\phi_0 \in \mathscr{S}(\mathbb{R}^d)$ is such that the corresponding energy

$$E[\phi_0] = \int_{\mathbb{R}^d} \big|\nabla\phi_0\big|^2 - \frac{2}{K+1}\big|\phi\big|^{K+1} \,\mathrm{d}x < 0.$$

Then there does not exist any global-in-time solution to

$$i\partial_t\phi = \triangle\phi + \big|\phi\big|^{K-1}\phi$$

with initial data $\phi(0) = \phi_0$. ■

Proof Consider the quantity $\mathcal{V}(t)$. By conservation of energy and the Virial identity we have that $\mathcal{V}'' \leq 8E < 0$. Assume for contradiction that a global-in-time solution exists. Then for some positive time T_0 we have that $\mathcal{V}(T_0) < 0$ by concavity. But by definition $\mathcal{V}$ is manifestly non-negative. □

6.58 Remark
The assumption of the blow-up theorem is non-empty. Observe that for any non-trivial initial data ϕ_0, we have

$$E[\lambda\phi_0] = \lambda^2 \|\nabla\phi_0\|_{L^2}^2 - \lambda^{K+1} \frac{2}{K+1} \|\phi_0\|_{L^{K+1}}^{K+1}.$$

So taking $\lambda \nearrow +\infty$ we can certainly find large initial data such that the corresponding energy is negative. ■

Further reading

Adams et al.: Sobolev spaces

Robert A. Adams and John J. F. Fournier. *Sobolev spaces*. Second. Vol. 140. Pure and Applied Mathematics (Amsterdam). Elsevier/Academic Press, Amsterdam, 2003, pp. xiv+305. ISBN: 0-12-044143-8.

Alinhac et al.: Pseudo-differential operators and the Nash-Moser theorem

Serge Alinhac and Patrick Gérard. *Pseudo-differential operators and the Nash-Moser theorem*. Vol. 82. GSM. American Mathematical Society, 2007.

Bahouri et al.: Fourier analysis and nonlinear partial differential equations

Hajer Bahouri, Jean-Yves Chemin, and Raphaël Danchin. *Fourier analysis and nonlinear partial differential equations*. Vol. 343. Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences]. Springer, Heidelberg, 2011, pp. xvi+523. ISBN: 978-3-642-16829-1. DOI: 10.1007/978-3-642-16830-7. URL: http://dx.doi.org/10.1007/978-3-642-16830-7.

Bergh et al.: Interpolation spaces. An introduction

Jöran Bergh and Jörgen Löfström. *Interpolation spaces. An introduction*. Grundlehren der Mathematischen Wissenschaften, No. 223. Springer-Verlag, Berlin-New York, 1976, pp. x+207.

Calderón: Intermediate spaces and interpolation, the complex method

A.-P. Calderón. "Intermediate spaces and interpolation, the complex method". In: *Studia Math.* 24 (1964), pp. 113–190. ISSN: 0039-3223.

Cazenave et al.: An introduction to semilinear evolution equations

Thierry Cazenave and Alain Haraux. *An introduction to semilinear evolution equations*. Vol. 13. Oxford Lecture Series in Mathematics and its Applications. Translated from the 1990 French original by Yvan Martel and revised by the authors. The Clarendon Press, Oxford University Press, New York, 1998, pp. xiv+186. ISBN: 0-19-850277-X.

Cohen-Tannoudji et al.: Quantum Mechanics

C. Cohen-Tannoudji, B. Diu, and F. Laloë. *Quantum Mechanics*. A Wiley interscience publication v. 1. Wiley, 1978. ISBN: 9780471164340.

Hörmander: Lectures on nonlinear hyperbolic differential equations

Lars Hörmander. *Lectures on nonlinear hyperbolic differential equations*. Vol. 26. Mathématiques & Applications (Berlin) [Mathematics & Applications]. Springer-Verlag, Berlin, 1997, pp. viii+289. ISBN: 3-540-62921-1.

Hörmander: The analysis of linear partial differential operators. I

Lars Hörmander. *The analysis of linear partial differential operators. I*. Classics in Mathematics. Distribution theory and Fourier analysis, Reprint of the second (1990) edition [Springer, Berlin; MR1065993 (91m:35001a)]. Springer-Verlag, Berlin, 2003, pp. x+440. ISBN: 3-540-00662-1. DOI: 10.1007/978-3-642-61497-2. URL: http://dx.doi.org/10.1007/978-3-642-61497-2.

John: Nonlinear wave equations, formation of singularities

Fritz John. *Nonlinear wave equations, formation of singularities*. Vol. 2. University Lecture Series. Seventh Annual Pitcher Lectures delivered at Lehigh University, Bethlehem, Pennsylvania, April 1989. Providence, RI: American Mathematical Society, 1990, pp. viii+64. ISBN: 0-8218-7001-7.

Klainerman: A commuting vectorfields approach to Strichartz-type inequalities and applications to quasi-linear wave equations

S. Klainerman. “A commuting vectorfields approach to Strichartz-type inequalities and applications to quasi-linear wave equations”. In: *Int. Math. Res. Not. IMRN* 2001.5 (2001), pp. 221–274. ISSN: 1073-7928. DOI: 10.1155/S1073792801000137. URL: http://dx.doi.org/10.1155/S1073792801000137.

Koch et al.: Dispersive Equations and Nonlinear Waves

Herbert Koch, Daniel Tataru, and Monica Vişan. *Dispersive Equations and Nonlinear Waves*. Vol. 45. Oberwolfach Seminars. Birkhäuser, 2014.

O’Neil: Convolution operators and $L(p, q)$ spaces

Richard O’Neil. “Convolution operators and $L(p, q)$ spaces”. In: *Duke Math. J.* 30 (1963), pp. 129–142. ISSN: 0012-7094. URL: http://projecteuclid.org/euclid.dmj/1077374532.

Smulevici: Small data solutions of the Vlasov-Poisson system and the vector field method

Jacques Smulevici. “Small data solutions of the Vlasov-Poisson system and the vector field method”. In: *Annals of PDE* (2016). URL: arXiv:1504.02195.

Sogge: Lectures on non-linear wave equations

Christopher D. Sogge. *Lectures on non-linear wave equations*. Second. Boston, MA: International Press, 2008, pp. x+205. ISBN: 978-1-57146-173-5.

Stein: Harmonic analysis: real-variable methods, orthogonality, and oscillatory integrals

Elias M. Stein. *Harmonic analysis: real-variable methods, orthogonality, and oscillatory integrals*. Vol. 43. Princeton Mathematical Series. With the assistance of Timothy S. Murphy, Monographs in Harmonic Analysis, III. Princeton, NJ: Princeton University Press, 1993, pp. xiv+695. ISBN: 0-691-03216-5.

Stein et al.: Fourier analysis

Elias M. Stein and Rami Shakarchi. *Fourier analysis*. Vol. 1. Princeton Lectures in Analysis. An introduction. Princeton University Press, Princeton, NJ, 2003, pp. xvi+311. ISBN: 0-691-11384-X.

Stein et al.: Real analysis

Elias M. Stein and Rami Shakarchi. *Real analysis*. Princeton Lectures in Analysis, III. Measure theory, integration, and Hilbert spaces. Princeton University Press, Princeton, NJ, 2005, pp. xx+402. ISBN: 0-691-11386-6.

Stein et al.: Introduction to Fourier analysis on Euclidean spaces

Elias M. Stein and Guido Weiss. *Introduction to Fourier analysis on Euclidean spaces*. Princeton Mathematical Series, No. 32. Princeton, N.J.: Princeton University Press, 1971, pp. x+297.

Tao: Nonlinear dispersive equations: local and global analysis

Terence Tao. *Nonlinear dispersive equations: local and global analysis*. CBMS Regional Conference Series in Mathematics. American Mathematical Society and Conference Board of the Mathematical Sciences, 2006.

Teschl: Mathematical methods in quantum mechanics

Gerald Teschl. *Mathematical methods in quantum mechanics*. Second. Vol. 157. Graduate Studies in Mathematics. With applications to Schrödinger operators. American Mathematical Society, Providence, RI, 2014, pp. xiv+358. ISBN: 978-1-4704-1704-8.

Trèves: Topological Vector Spaces, Distributions and Kernels

François Trèves. *Topological Vector Spaces, Distributions and Kernels*. Republication of Academic Press (1967) edition. Dover, 2006.

Whitney: Singularities of Mappings of Euclidean Spaces

Hassler Whitney. "Singularities of Mappings of Euclidean Spaces". In: *Collected papers. Vol. I*. Ed. by James Eells and Domingo Toledo. Contemporary Mathematicians. Boston, MA: Birkhäuser Boston, 1992, pp. 436–452.