Lecture Notes
Math 632, PDE

Sigmund Selberg
Visiting Assistant Professor
Johns Hopkins University

Spring semester 2001

Chapter 1

Week 1: The basics

1.1 The wave equation on $\mathbb{R}^{1+n}$

We consider the equation

$$\Box u = 0,$$

where $u = u(t, x)$ is a function on $\mathbb{R} \times \mathbb{R}^n = \mathbb{R}^{1+n}$ and $\Box$ is the *wave operator*:

$$\Box = \partial_t^2 - \Delta.$$

Here $\Delta = \partial_1^2 + \cdots + \partial_n^2$ is the Laplacian in $x = (x^1, \dots, x^n)$. Thus, we write

$$\partial_t = \frac{\partial}{\partial t}, \quad \partial_i = \frac{\partial}{\partial x^i}.$$

We also write

$$\nabla u = (\partial_1 u, \dots, \partial_n u), \quad \partial u = (\partial_t u, \nabla u).$$

Occasionally it is convenient to write $t = x^0$, in which case $\partial_0 = \partial_t$.

Remark. For those familiar with Lorentizan geometry, $-\Box$ is just the Laplace-Beltrami operator relative to the Minkowski metric

$$\eta_{\mu\nu} = \operatorname{diag}(-1, 1, \dots, 1)$$

on $\mathbb{R}^{1+n}$. If we apply the summation convention, and raise and lower indices relative to this metric, then $-\Box = \eta^{\mu\nu}\partial_\mu\partial_\nu = \partial^\mu\partial_\mu$, where the indices μ, ν run from 0 to n.

1.2 The Cauchy problem

Given functions (initial data) f, g on $\mathbb{R}^n$, we consider the *Cauchy problem* for the initial hypersurface $\{t = 0\} \times \mathbb{R}^n$:

$$\Box u = 0, \quad u\big|_{t=0} = f, \quad \partial_t u\big|_{t=0} = g. \tag{1.1}$$

We want to show that this problem is *well-posed*:

- The solution u exists for all $t > 0$;
- u is unique;
- u depends continuously on f and g.

This is not very precise; more rigorous statements of this type will be proved later in the course.

In fact, one has explicit formulas for the solution u of (1.1) in terms of the data f, g. Before deriving these formulas in dimensions $n = 1, 2, 3$, we will prove uniqueness of the solution, using the energy method.

Theorem 1. *Suppose $u \in C^2([0,T] \times \mathbb{R}^n)$ solves $\Box u = 0$. Fix $x_0 \in \mathbb{R}^3$ and $0 < t_0 \leq T$, and suppose*

$$u = \partial_t u = 0 \quad \textit{for} \quad t = 0, \quad |x - x_0| \leq t_0.$$

Then

$$u = 0 \quad \textit{in} \quad \Omega = \{(t,x) : 0 \leq t \leq t_0,\, |x - x_0| \leq t_0 - t\}.$$

*(We call Ω the **solid backward light cone with vertex at** (t_0, x_0).)*

Proof. Let $B_t = \{x : |x - x_0| \leq t_0 - t\}$ and define the *energy*

$$e(t) = \frac{1}{2} \int_{B_t} |\partial u(t,x)|^2 \, dx.$$

(Recall ∂u is the space-time gradient.) Differentiate (see Exercise 1 below) to get

$$e'(t) = \int_{B_t} (u_t u_{tt} + \nabla u \cdot \nabla u_t) \, dx - \frac{1}{2} \int_{\partial B_t} |\partial u|^2 \, d\sigma(x).$$

Since

$$\operatorname{div}(u_t \nabla u) = \nabla u_t \cdot \nabla u + u_t \Delta u,$$

it follows from the divergence theorem that

$$\begin{aligned} e'(t) &= \int_{B_t} \operatorname{div}(u_t \nabla u) \, dx - \frac{1}{2} \int_{\partial B_t} |\partial u|^2 \, d\sigma(x) \\ &= \int_{\partial B_t} u_t \nabla u \cdot \mathbf{n} \, d\sigma(x) - \frac{1}{2} \int_{\partial B_t} |\partial u|^2 \, d\sigma(x), \end{aligned}$$

where $\mathbf{n}$ is the outward unit normal of ∂B_t. But

$$|u_t \nabla u \cdot \mathbf{n}| \leq |u_t| \, |\nabla u| \leq \frac{1}{2} \left(|u_t|^2 + |\nabla u|^2 \right),$$

and so we conclude that $e'(t) \leq 0$ for $0 \leq t \leq t_0$. This implies $e(t) \leq e(0) = 0$. But certainly $e(t) \geq 0$, so $e(t) = 0$. It follows that $\partial u = 0$ in Ω, and hence $u = 0$ in Ω. □

Exercise 1. Part (b) below was used in the proof of the uniqueness theorem. Here $B_r(x)$ denotes the open ball in $\mathbb{R}^n$ centered at x with radius r, and $S_r(x)$ denotes its boundary, the sphere of radius r at x.

(a) If f is a continuous function on $\mathbb{R}^n$ and $x \in \mathbb{R}^n$, then

$$\frac{d}{dr}\int_{B_r(x)} f(y)\,dy = \int_{S_r(x)} f(y)\,d\sigma(y)$$

where $d\sigma$ is surface measure. (*Hint:* Use polar coordinates to write $\int_{B_r(x)} f(y)\,dy = \int_0^r \int_{S^{n-1}} f(x+\rho\omega)\,d\sigma(\omega)\rho^{n-1}\,d\rho$, where $S^{n-1} = S_1(0)$ is the unit sphere.)

(b) Now suppose $f = f(r,x)$, where $r \in \mathbb{R}$ and $x \in \mathbb{R}^n$. Fix x and set

$$\phi(r) = \int_{B_r(x)} f(r,y)\,dy.$$

Assuming f and $\partial_r f$ are continuous, show that

$$\phi'(r) = \int_{B_r(x)} \partial_r f(r,y)\,dy + \int_{S_r(x)} f(r,y)\,d\sigma(y).$$

(*Hint:* Write

$$\begin{aligned}\frac{\phi(r+h)-\phi(r)}{h} &= \int_{B_{r+h}(x)} \frac{f(r+h,y)-f(r,y)}{h}\,dy \\ &\quad + \frac{1}{h}\left\{\int_{B_{r+h}(x)} f(r,y)\,dy - \int_{B_r(x)} f(r,y)\,dy\right\}.\end{aligned}$$

On the first term use the dominated convergence theorem, and on the second term use part (a).)

1.3 Solution of the Cauchy problem

1.4 Dimension $n = 1$: D'Alembert's formula

The solution of (1.1) in dimension $n = 1$ is given by D'Alembert's formula:

$$u(t,x) = \frac{1}{2}\left[f(x+t) + f(x-t)\right] + \frac{1}{2}\int_{x-t}^{x+t} g.$$

Theorem 2. *If $f \in C^2(\mathbb{R})$ and $g \in C^1(\mathbb{R})$, then u defined as above is C^2 and solves the Cauchy problem* (1.1).

The proof is a simple calculation.

1.4.1 Dimension $n = 3$: Spherical mean

Define the *spherical mean* of a function $\phi : \mathbb{R}^3 \to \mathbb{R}$ by

$$M_\phi(x,r) = \frac{1}{4\pi} \int_{S^2} \phi(x+ry)\, d\sigma(y) \quad \text{for} \quad x \in \mathbb{R}^3, \quad r \in \mathbb{R}.$$

Observe that

$$M_\phi(x,0) = \phi(x) \tag{1.2}$$

and that M_ϕ is an even function of r:

$$M_\phi(x,-r) = M_\phi(x,r). \tag{1.3}$$

We showed that (see Sogge or Folland) M_ϕ satisfies the *Darboux* equation

$$\Delta_x M_\phi = \left(\partial_r^2 + (2/r)\partial_r\right) M_\phi,$$

provided $\phi \in C^2(\mathbb{R}^3)$.

Next, we set

$$M_u(t,x,r) = M_{u(t,\cdot)}(x,r).$$

Then an easy calculation shows

$$\Box u = 0 \iff \partial_t^2 M_u = \left(\partial_r^2 + (2/r)\partial_r\right) M_u.$$

If we then fix x and set

$$v(t,r) = rM_u(t,x,r),$$

it follows that u solves $\Box u = 0$ on $\mathbb{R}^{1+3}$ if and only if, for each fixed x, v solves the wave equation on $\mathbb{R}^{1+1}$:

$$\partial_t^2 v = \partial_r^2 v.$$

Thus, D'Alembert's formula expresses $v(t,r)$ in terms of the data

$$v(0,r) = rM_f(x,r), \quad \partial_t v(0,r) = rM_g(x,r),$$

and we then obtain a formula for $u(t,x)$ by noting that

$$u(t,x) = M_u(t,x,0) = \lim_{r\to 0} \frac{v(t,r)}{r}.$$

Calculating the right hand side, we finally get

$$\begin{aligned} u(t,x) &= \partial_t \left(tM_f(t,x)\right) + tM_g(x,t) \\ &= \frac{1}{4\pi} \int_{S^2} \left[f(x+ty) + \nabla f(x+ty)\cdot ty + tg(x+ty)\right] d\sigma(y) \\ &= \frac{1}{4\pi t^2} \int_{|y-x|=t} \left[f(y) + \nabla f(y)\cdot(y-x) + tg(y)\right] d\sigma(y) \end{aligned}$$

where the last equality is valid for $t > 0$ by a change of variables.

We then have:

Theorem 3. *If $f \in C^3(\mathbb{R}^3)$ and $g \in C^2(\mathbb{R}^3)$, then u as defined above is C^2 and solves* (1.1) *on $\mathbb{R}^{1+3}$.*

1.4.2 Dimension $n = 2$: Method of descent

We use Hadamard's method of descent to get the solution of the case $n = 2$ from the case $n = 3$.

Assume

$$u = u(t, x_1, x_2)$$

solves (1.1) on $\mathbb{R}^{1+2}$. Define

$$v(t,x) = u(t,x') \quad \text{where} \quad x = (x_1, x_2, x_3), \quad x' = (x_1, x_2).$$

Then v solves a Cauchy problem on $\mathbb{R}^{1+3}$

$$\Box v = 0, \quad v(0,x) = f(x'), \quad \partial_t v(0,x) = g(x').$$

By the formula derived for $n = 3$, it then follows that

$$v(t,x) = \partial_t \left(\frac{t}{4\pi} \int_{S^2} f(x' + ty')\, d\sigma(y) \right) + \frac{t}{4\pi} \int_{S^2} g(x' + ty')\, d\sigma(y).$$

But since the integrands are independent of y_3, we get (see Exercise 2 below),

$$u(t,x) = \partial_t \left(\frac{t}{2\pi} \int_{|y|<1} f(x+ty) \frac{dy}{\sqrt{1-|y|^2}} \right) + \frac{t}{2\pi} \int_{|y|<1} g(x+ty) \frac{dy}{\sqrt{1-|y|^2}}$$

where we *dropped the primes* on x and y (so now $x, y \in \mathbb{R}^2$).

Exercise 2. Prove that

$$\int_{S^2} h(y')\, d\sigma(y) = 2 \int_{|y'|<1} h(y') \frac{dy'}{\sqrt{1-|y'|^2}},$$

where $y' = (y_1, y_2)$ and $y = (y', y_3) \in S^2$. (*Hint:* Parametrize the two hemispheres over the y_1y_2-plane by $(y', \pm\phi(y'))$ for $|y'| < 1$, where $\phi(y') = \sqrt{1-|y'|^2}$.)

1.4.3 Higher dimensions

See Folland. The formulas are as follows.

Theorem 4. *Suppose $n \geq 3$ is odd. If*

$$f \in C^{(n+3)/2}(\mathbb{R}^n), \quad g \in C^{(n+1)/2}(\mathbb{R}^n),$$

then

$$u(t,x) = \gamma_n \left[\partial_t (t^{-1}\partial_t)^{(n-3)/2} \left(t^{n-2} \int_{y \in S^{n-1}} f(x+ty)\, d\sigma(y) \right) + (t^{-1}\partial_t)^{(n-3)/2} \left(t^{n-2} \int_{y \in S^{n-1}} g(x+ty)\, d\sigma(y) \right) \right]$$

is C^2 and solves the Cauchy problem (1.1) on $\mathbb{R}^{1+n}$. Here γ_n is a constant.

By the method of descent one then obtains:

Theorem 5. *Suppose $n \geq 2$ is even. If*

$$f \in C^{(n+4)/2}(\mathbb{R}^n), \quad g \in C^{(n+2)/2}(\mathbb{R}^n),$$

then

$$u(t,x) = \gamma_n \Bigg[\partial_t (t^{-1}\partial_t)^{(n-2)/2} \left(t^{n-1} \int_{|y|<1} f(x+ty) \frac{dy}{\sqrt{1-|y|^2}} \right) + (t^{-1}\partial_t)^{(n-2)/2} \left(t^{n-1} \int_{|y|<1} g(x+ty) \frac{dy}{\sqrt{1-|y|^2}} \right) \Bigg]$$

is C^2 and solves the Cauchy problem (1.1) *on $\mathbb{R}^{1+n}$, where γ_n is a constant.*

1.5 Huygens' principle

It is evident from the formula we derived, that in dimension $n = 3$ (and in higher odd dimensions as well), the value of u at a point (t, x) $(t > 0)$ only depends on the values of the data f, g on the set $\{y : |y - x| = t\}$ (or more precisely in an infinitesimal neighborhood of this sphere, since the formula involves ∇f).

As a consequence, an initial disturbance at the origin, say a flash of light, propagates with unit speed and can only be seen on the *forward light cone* with vertex at the origin, namely the set $\{(t, x) : t = |x|\}$. This is known as the (strong) Hyugens principle.

In dimensions $n = 1, 2$ (and in any even dimension $n \geq 2$) a weaker version of Huygens' principle holds. Then u at (t, x) depends on the values of f, g in the *ball* $\{y : |y - x| \leq t\}$. Consequently, a flash of light at the origin will be visible to an observer at a point x_0 in space, at times $t \geq |x_0|$, and not just at $t = |x_0|$ as in dimensions $n = 3, 5, \ldots$, although the intensity of the light will decay (except in dimension $n = 1$; see the next section).

Exercise 3. (Finite speed of propagation.) Suppose f, g are smooth and compactly supported, say

$$f(x) = g(x) = 0 \quad \text{for} \quad |x| > R$$

for some $R > 0$. Prove that $u(t, \cdot)$ is compactly supported for each $t > 0$, and that in fact

$$u(t,x) = 0 \quad \text{for} \quad |x| > t + R.$$

Moreover, if n is odd and $n \geq 3$, then

$$u(t,x) = 0 \quad \text{unless} \quad t - R \leq |x| \leq t + R.$$

Show also that the energy

$$e(t) = \frac{1}{2} \int_{\mathbb{R}^n} |\partial u(t,x)|^2 \, dx$$

is independent of t.

Chapter 2

Week 2: Weak and even weaker solutions

2.1 Decay as $t \to \infty$

Consider the Cauchy problem on $\mathbb{R}^{1+n}$,

$$\Box u = 0, \quad u\big|_{t=0} = f, \quad \partial_t u\big|_{t=0} = g, \tag{2.1}$$

and assume f, g are smooth and compactly supported:

$$f(x) = g(x) = 0 \quad \text{for} \quad |x| \geq R,$$

for some $R > 0$. We then have:

Theorem 6. $\|u(t,\cdot)\|_{L^\infty} = O(t^{-(n-1)/2})$ *as* $t \to \infty$.

We will prove this in dimensions $n = 1, 2, 3$.

Proof for $n = 1$. By D'Alembert's formula,

$$u(t,x) = \frac{1}{2}[f(x+t) + f(x-t)] + \frac{1}{2}\int_{x-t}^{x+t} g,$$

and so

$$\|u(t,\cdot)\|_{L^\infty} \leq \|f\|_{L^\infty} + \frac{1}{2}\int_{-\infty}^{\infty} |g| \, .$$

Proof for $n = 2$. It suffices to consider the case $f = 0$. Then

$$u(t,x) = \frac{t}{2\pi}\int_{|y|<1} g(x+ty)\frac{dy}{\sqrt{1-|y|^2}}.$$

Since

$$\operatorname{supp} g \subset \{|x| \le R\}, \tag{2.2}$$

we have

$$\operatorname{supp} u(t,\cdot) \subset \{|x| \le R + t\}. \tag{2.3}$$

Now convert to polar coordinates $y = \rho e^{i\theta}$, $0 < \rho < 1$, $0 < \theta < 2\pi$. Then

$$u(t,x) = \frac{t}{2\pi} \int_0^1 \left(\int_0^{2\pi} g(x + t\rho e^{i\theta})\, d\theta \right) \frac{\rho\, d\rho}{\sqrt{1-\rho^2}}.$$

But in view of (2.2),

$$\rho_{\min} \le \rho \le \rho_{\max}, \quad \rho_{\min} = \frac{|x| - R}{t}, \quad \rho_{\max} = \frac{|x| + R}{t},$$

and the angle θ is restricted to an interval of length $\le \frac{C}{1+|x|}$.

We conclude that

$$|u(t,x)| \le \|g\|_{L^\infty} \frac{Ct}{1+|x|} \int_a^b \frac{\rho\, d\rho}{\sqrt{1-\rho}} \tag{2.4}$$

where

$$a = \max(0, \rho_{\min}), \quad b = \min(1, \rho_{\max}). \tag{2.5}$$

(Observe that by (2.3) we may assume $|x| \le t + R$, and so $\rho_{\min} \le 1$ and therefore $a \le b$.)

Since

$$b - a \le \rho_{\max} - \rho_{\min} = \frac{2R}{t},$$

we can now appeal to the following estimate, which we leave as an exercise.

Exercise 1. Assume $0 \le a < b \le 1$ and $b - a \le \frac{1}{4}$. Then

$$\int_a^b \frac{\rho\, d\rho}{\sqrt{1-\rho}} \le \begin{cases} (b-a)^2 & \text{if } a \le \frac{1}{2} \\ \sqrt{b-a} & \text{if } a \ge \frac{1}{2}. \end{cases}$$

(*Hint:* If $a \le \frac{1}{2}$, then $1 - \rho \ge \frac{1}{4}$. If $a \ge \frac{1}{2}$, bound the integrand by $\frac{1}{\sqrt{b-\rho}}$.)

To apply this, note that (with a, b defined by (2.5)):

(i) If $\rho_{\min} \le \frac{1}{2}$, then $a \le \frac{1}{2}$, and so $|u(t,x)| = O(1/t)$ in this case, using (2.4) and the exercise.

(ii) If $\rho_{\min} \ge \frac{1}{2}$, then $a \ge \frac{1}{2}$ and $|x| \sim t$, so $|u(t,x)| = O(1/\sqrt{t})$ by (2.4) and the exercise.

Proof for $n = 3$. Again we may assume $f = 0$, in which case

$$u(t,x) = \frac{t}{4\pi}\int_{S^2} g(x+ty)\,d\sigma(y).$$

From (2.2) it is clear that u vanishes unless

$$|x| - R \le t \le |x| + R,$$

so $t \sim |x|$ on the support of u. Moreover, it is clear that the integrand $g(x+ty)$ ($y \in S^2$) vanishes unless y makes an angle $\lesssim 1/|x|$ with $\omega_0 = -x/|x| \in S^2$. The corresponding region on S^2 has area $\lesssim |x|^{-2} \sim t^{-2}$, and we conclude $|u(t,x)| \le C_R \|g\|_{L^\infty} t^{-1}$.

Notation. The symbol $\lesssim$ stands for $\le$ up to a positive, multiplicative constant, which may depend on parameters that are considered fixed. (For example, in the above the constant depends on R.) The notation $r \sim s$ means that $r \lesssim s \lesssim r$.

2.2 The equation $\Box u = F$ and Duhamel's principle

Consider the Cauchy problem for the *inhomogeneous* wave equation:

$$\Box u = F, \quad u\big|_{t=0} = f, \quad \partial_t u\big|_{t=0} = g, \tag{2.6}$$

which represents waves influenced by a driving force $F = F(t,x)$. By linearity, the solution is

$$u = v + w,$$

where v is the solution of the corresponding homogeneous problem ($F = 0$), and w is the solution of (2.6) with zero data ($f = g = 0$). The idea is that w is a continuous superposition (integral) of solutions of the homogeneous wave equation. This is expressed by *Duhamel's principle*:

Theorem 7. *Suppose $F \in C^{(n+2)/2}(\mathbb{R}^{1+n})$ if n is even, or $F \in C^{(n+1)/2}(\mathbb{R}^{1+n})$ if n is odd. For each $s \in \mathbb{R}$, let $v(t,x;s)$ be the solution of the Cauchy problem*

$$\Box v = 0, \quad v(0,x;s) = 0, \quad \partial_t v(0,x;s) = F(s,x).$$

Then $u(t,x) = \int_0^t v(t-s,x;s)\,ds$ is in C^2 and solves the Cauchy problem (2.6) with $f = g = 0$.

Exercise 2. Prove this. (Observe that if $\phi(t) = \int_0^t \psi(t,s)\,dt$, where ψ and $\partial_t\psi$ are continuous, then $\phi'(t) = \psi(t,t) + \int_0^t \partial_t\psi(t,s)\,ds$.)

2.3 Weak solutions

So far we only considered *classical* solutions—that is, solutions which are at least C^2—of the Cauchy problem (2.6). However, the solution formulas make sense for data f, g and F with very little regularity. For example, D'Alembert's formula for the solution of the homogeneous equation in dimension $n = 1$ makes perfectly good sense for any $f, g \in L^1_{\text{loc}}(\mathbb{R})$—the only question is whether the resulting function u can be said to "solve" the Cauchy problem in some reasonable sense. The answer is affirmative, as we now demonstrate.

To motivate our definition of "weak solution", let us start with a classical solution $u \in C^2$ of the Cauchy problem on a time-strip

$$S_T = [0, T] \times \mathbb{R}^n.$$

Thus, we assume

$$\Box u = F \quad \text{on} \quad S_T, \quad u\big|_{t=0} = f, \quad \partial_t u\big|_{t=0} = g. \tag{2.7}$$

Let ϕ be a test function compactly supported in $(-\infty, T) \times \mathbb{R}^n$. Now multiply the equation by ϕ, and integrate by parts, using the fact that ϕ vanishes near $t = T$, to get

$$\begin{aligned}\int_{S_T} F\phi\,dt\,dx &= \int_{S_T} (\Box u)\phi\,dt\,dx \\ &= \int_{\mathbb{R}^n} \left(-\int_0^T \partial_t u \partial_t \phi\,dt - g(x)\phi(0,x) \right) dx - \int_{S_T} u\Delta\phi\,dt\,dx \\ &= \int_{S_T} u\Box\phi\,dt\,dx + \int_{\mathbb{R}^n} f(x)\partial_t\phi(0,x)\,dx - \int_{\mathbb{R}^n} g(x)\phi(0,x)\,dx.\end{aligned}$$

This leads us to make the following

Definition. Let $f, g \in L^1_{\text{loc}}(\mathbb{R}^n)$ and $F \in L^1_{\text{loc}}(S_T)$. We say $u \in L^1_{\text{loc}}(S_T)$ is a *weak solution* of (2.7) if

$$\int_{S_T} u\Box\phi\,dt\,dx = \int_{S_T} F\phi\,dt\,dx - \int_{\mathbb{R}^n} f(x)\partial_t\phi(0,x)\,dx + \int_{\mathbb{R}^n} g(x)\phi(0,x)\,dx \tag{2.8}$$

for all $\phi \in C^\infty_c$ supported in $(-\infty, T) \times \mathbb{R}^n$.

The next result shows that this definition is reasonable.

Theorem 8. *A weak solution belonging to $C^2(S_T)$ is a classical solution.*

Proof. If ϕ is supported in $(0, T) \times \mathbb{R}^n$, then (2.8) says that

$$\int u\Box\phi\,dt\,dx = \int F\phi\,dt\,dx$$

and integration by parts shows that $\int u\Box\phi = \int(\Box u)\phi$. Thus,[1]

$$\int (\Box u - F)\,\phi\,dt\,dx = 0,$$

and since ϕ was arbitrary, we conclude that $\Box u = F$ on S_T. (After redefining F on a set of measure zero.)

It remains to prove that u takes the initial data f, g. Let us prove that

$$\partial_t u(0,x) = g(x), \tag{2.9}$$

leaving the verification of $u(0,x) = f(x)$ as an exercise (see below). First, (2.9) is equivalent to

$$\int_{\mathbb{R}^n} \partial_t u(0,x)\psi(x)\,dx = \int_{\mathbb{R}^n} g(x)\psi(x)\,dx \quad \text{for all} \quad \psi \in C_c^\infty(\mathbb{R}^n). \tag{2.10}$$

Fix such a ψ. Let a be a smooth function such that

$$a(t) = \begin{cases} 1 & \text{for} \quad t \le 0, \\ 0 & \text{for} \quad t \ge 1. \end{cases}$$

Then set $\theta_k(t) = a(kt)$ for $k \in \mathcal{N}$ and $t \in \mathbb{R}$. Observe that

$$\theta_k'(0) = 0, \qquad \theta_k(t) = 0 \quad \text{for} \quad t \ge 1/k.$$

Now take $\phi = \phi_k$ in (2.8), where

$$\phi_k(t,x) = \theta_k(t)\psi(x).$$

Then the right hand side of (2.8) reads (we take k so large that $1/k < T$)

$$\int_{\mathbb{R}^n} \left(\int_0^{1/k} F(t,x)a(kt)\,dt \right) \psi(x)\,dx + \int_{\mathbb{R}^n} g(x)\psi(x)\,dx. \tag{2.11}$$

Since $F \in C(S_T)$ (we showed $\Box u = F$), the first term is $O(1/k)$ as $k \to \infty$.

We claim that the left hand side of (2.8) equals

$$\int_{\mathbb{R}^n} \partial_t u(0,x)\psi(x)\,dx + O(1/k). \tag{2.12}$$

Equating (2.12) with (2.11) and passing to the limit $k \to \infty$ then gives (2.10).

It remains to prove the claim. But the left hand side of (2.8) is

$$\int_{S_T} u(t,x)\theta_k''(t)\psi(x)\,dt\,dx - \int_{\mathbb{R}^n} \left(\int_0^{1/k} u(t,x)a(kt)\,dt \right) \Delta\psi(x)\,dx.$$

---

[1] We use the fact that if $h \in L^1_{\text{loc}}$ and $\int h\psi = 0$ for all $\psi \in C_c^\infty$, then $h = 0$. (We identify functions in L^1_{loc} which are equal almost everywhere.)

The second term is $O(1/k)$, and after an integration by parts, the first term becomes

$$-\int_{S_T} \partial_t u(t,x)\theta_k'(t)\psi(x)\,dt\,dx.$$

(There are no boundary terms, since $\theta_k'(t) = 0$ for $t = 0, T$.) A second integration by parts transforms this into

$$\int_{\mathbb{R}^n} \left(\int_0^{1/k} \partial_t^2 u(t,x) a(kt)\,dt \right) \psi(x)\,dx + \int_{\mathbb{R}^n} \partial_t u(0,x)\psi(x)\,dx.$$

Again, the first term is $O(1/k)$, proving the claim. □

Exercise 3. Complete the proof of the theorem by showing that $u(0,x) = f(x)$. Proceed as in the proof of (2.10), but now choose $\theta_k(t) = \frac{1}{k} b(kt)$, where b is some smooth, compactly supported function such that

$$b(0) = 0, \quad b'(0) = 1.$$

Thus $\theta_k(0) = 0$ and $\theta_k'(0) = 1$ for all k, and the support of θ_k shrinks to the origin as $k \to \infty$. (You should find that the "error terms" are now $O(1/k^2)$.)

Example. Recall D'Alembert's formula for the solution of the Cauchy problem

$$\Box u = 0, \quad u\big|_{t=0} = f, \quad \partial_t u\big|_{t=0} = g \tag{2.13}$$

on $\mathbb{R}^{1+1}$:

$$u(t,x) = \frac{1}{2}\left[f(x+t) + f(x-t)\right] + \frac{1}{2}\int_{x-t}^{x+t} g. \tag{2.14}$$

Clearly, this defines a function $u \in L^1_{\text{loc}}(\mathbb{R}^2)$ if $f, g \in L^1_{\text{loc}}(\mathbb{R})$, and we claim that u is a weak solution of (2.13) on $[0,\infty) \times \mathbb{R}$.

There are two ways to do this: by direct calculation or approximation by smooth function. Let us briefly outline both procedures.

First method: Direct calculation. The key is to show, for all $A \in L^1_{\text{loc}}(\mathbb{R})$ and $\phi \in C_c^\infty(\mathbb{R}^2)$,

$$\int_{\mathbb{R}} \int_0^\infty A(x+t)\Box\phi(t,x)\,dt\,dx = -\int_{\mathbb{R}} A(x)\left[\partial_t\phi(0,x) + \partial_x\phi(0,x)\right]\,dx, \tag{2.15}$$

$$\int_{\mathbb{R}} \int_0^\infty A(x-t)\Box\phi(t,x)\,dt\,dx = \int_{\mathbb{R}} A(x)\left[-\partial_t\phi(0,x) + \partial_x\phi(0,x)\right]\,dx \tag{2.16}$$

Exercise 4. Prove these formulas by changing variables $(t,x) \to (\xi,\eta)$, where

$$\xi = x + t, \quad \eta = x - t.$$

Thus,

$$t = \frac{\xi - \eta}{2}, \quad x = \frac{\xi + \eta}{2}, \quad 2dt\,dx = d\xi\,d\eta,$$

and the region $t > 0$ is transformed into $\xi > \eta$. Moreover, if we set

$$\psi(\xi, \eta) = \phi(t, x),$$

then $\Box\phi = -4\partial_\xi \partial_\eta \psi$.

Applying (2.15) and (2.16), we obtain

$$\int_{\mathbb{R}} \int_0^\infty \frac{1}{2} \left[f(x+t) + f(x-t) \right] \Box\phi(t,x)\, dt\, dx = -\int_{\mathbb{R}} f(x) \partial_t \phi(0,x)\, dx.$$

Similarly, since $\int_{x-t}^{x+t} g = \int_0^{x+t} g - \int_0^{x-t} g$, we find that

$$\int_{\mathbb{R}} \int_0^\infty \frac{1}{2} \left(\int_{x-t}^{x+t} g \right) \Box\phi(t,x)\, dt\, dx = -\int_{\mathbb{R}} \left(\int_0^x g \right) \partial_x \phi(0,x)\, dx,$$

and an integration by parts shows that the right hand side equals

$$\int_{\mathbb{R}} g(x) \phi(0,x)\, dx.$$

This proves that u solves (2.13) in the weak sense.

Second method: Smooth approximation. Here we employ an approximation technique to prove that u is a weak solution. Recall that C_c^∞ is dense in L^1. Therefore, if we fix a compact interval $[-a, a]$, we can find sequences f_j, g_j in C_c^∞ such that

$$\int_{-a}^{a} |f - f_j|\, dx, \int_{-a}^{a} |g - g_j|\, dx \longrightarrow 0 \quad \text{as} \quad j \to \infty. \tag{2.17}$$

Let u_j be given by (2.14) with f, g replaced by f_j, g_j. Then u_j solves the corresponding Cauchy problem in the classical sense, and hence also in the weak sense:

$$\int_{\mathbb{R}} \int_0^\infty u_j \Box\phi\, dt\, dx = -\int_{\mathbb{R}} f_j(x) \partial_t \phi(0,x)\, dx + \int_{\mathbb{R}} g_j(x) \phi(0,x)\, dx \tag{2.18}$$

Fix $\phi \in C_c^\infty(\mathbb{R}^2)$. We want to pass to the limit and conclude that u solves (2.13) in the weak sense:

$$\int_{\mathbb{R}} \int_0^\infty u \Box\phi\, dt\, dx = -\int_{\mathbb{R}} f(x) \partial_t \phi(0,x)\, dx + \int_{\mathbb{R}} g(x) \phi(0,x)\, dx \tag{2.19}$$

Clearly, RHS(2.18) converges to RHS(2.19), if we choose a in (2.17) so large that the support of ϕ is contained in $[-a, a] \times \mathbb{R}$. But since u is given by D'Alembert's formula (2.14), it follows from (2.17) that $u_j \to u$ in the L^1 norm on the support of ϕ, if we take a sufficiently large. In fact, we have to take a so large that for any backward light cone with vertex at a point $(t, x) \in \operatorname{supp} \phi$, $t > 0$, its base in the plane $t = 0$ is contained in $[-a, a]$. (Draw a picture). It then follows that also LHS(2.18) converges to LHS(2.19).

2.4 Even weaker solutions

In the previous section we defined the concept of weak solution of the Cauchy problem for locally integrable initial data. Now we weaken the regularity assumptions on then data f and g even further: we merely assume they are distributions on $\mathbb{R}^n$. We consider the case $n = 1$ in detail, and leave the higher dimensional cases as exercises.

Let us start by recalling some basic properties of distributions.

2.4.1 Distributions

A good reference for this material is Folland's book on real analysis.

Let Ω be an open set in $\mathbb{R}^n$. Let $C_c^\infty(\Omega)$ be the set of smooth functions $\phi : \mathbb{R}^n \to \mathbb{C}$ compactly supported in Ω. Such functions are called test functions. Convergence in the space of test functions is defined as follows:

$$\phi_j \to \phi \quad \text{in} \quad C_c^\infty(\Omega)$$

means that (i) there is a compact $K \subset \Omega$ such that $\operatorname{supp} \phi_j \subset K$ for all j, and (ii) $\partial^\alpha \phi_j \to \partial^\alpha \phi$ uniformly for all multi-indices α.

Exercise 5. Let e_j be the j-th standard basis vector on $\mathbb{R}^n$, let $\phi \in C_c^\infty(\Omega)$, and define the difference quotient

$$\Delta_h^j \phi(x) = \frac{\phi(x + he_j) - \phi(x)}{h}.$$

Prove that $\Delta_h^j \phi \to \partial_j \phi$ as $h \to 0$ in the sense of $C_c^\infty(\Omega)$. (That is, $\partial^\alpha \Delta_h^j \phi \to \partial^\alpha \partial_j \phi$ uniformly as $h \to 0$, for all multi-indices α. Note that since $\partial^\alpha \phi$ is again a test function, you may without loss of generality take $\alpha = 0$. Now use the mean value theorem.)

A *distribution* on Ω is a linear functional u on $C_c^\infty(\Omega)$ which is continuous, in the sense that

$$\phi_j \to \phi \quad \text{in} \quad C_c^\infty(\Omega) \qquad \Longrightarrow \qquad \langle u, \phi_j \rangle \to \langle u, \phi \rangle .$$

Note that one writes $\langle u, \phi \rangle$ rather than $u(\phi)$ for the value of u at a test function ϕ. We denote by $\mathcal{D}'(\Omega)$ the set of distributions u on Ω, and we equip it with the topology of pointwise convergence. Thus, sequential convergence in $\mathcal{D}'(\Omega)$ has the following meaning:

$$u_j \to u \quad \text{in} \quad \mathcal{D}'(\Omega) \qquad \Longleftrightarrow \qquad \langle u_j, \phi \rangle \to \langle u, \phi \rangle \quad \text{for all} \quad \phi \in C_c^\infty(\Omega). \tag{2.20}$$

Exercise 6. Show that any $u \in L^1_{\text{loc}}(\Omega)$ defines a distribution by setting

$$\langle u, \phi \rangle = \int u(x)\phi(x)\, dx.$$

(The linearity is obvious; the point is to check that u is continuous.) Show also that $u, v \in L^1_{\text{loc}}(\Omega)$ define the *same* distribution iff $u = v$ (almost everywhere). Thus we may consider $L^1_{\text{loc}}(\Omega)$ as a subset of $\mathcal{D}'(\Omega)$.

It is customary to write

$$\langle u, \phi \rangle = \int u(x)\phi(x)\, dx \tag{2.21}$$

even when u is *not* a function. This convenient abuse of notation clarifies many operations on distributions (translation for example).

Exercise 7. For any $x_0 \in \Omega$ and multi-index α, show that the map $\phi \to \partial^\alpha \phi(x_0)$ is a distribution on Ω and that it is not given by a locally integrable function.

If we take $\Omega = \mathbb{R}^n$, $\alpha = 0$ and $x_0 = 0$ in the previous exercise, we get the famous *delta function*

$$\langle \delta, \phi \rangle = \phi(0),$$

which of course is not a function (it is a measure).

The following characterization of distributions is sometimes useful.

Theorem 9. *Let $u : C_c^\infty(\Omega) \to \mathbb{C}$ be linear. Then $u \in \mathcal{D}'(\Omega)$ iff for every compact set $K \subset \Omega$ there exist $C_K > 0$ and $N_K \in \mathcal{N}$ such that*

$$|\langle u, \phi \rangle| \leq C_K \sum_{|\alpha| \leq N_K} \|\partial^\alpha \phi\|_{L^\infty} \tag{2.22}$$

for all test functions ϕ supported in K.

Proof. Clearly (2.22) implies that $\langle u, \phi_j \rangle \to \langle u, \phi \rangle$ whenever $\phi_j \to \phi$ in $C_c^\infty(\Omega)$. Conversely, assume the condition in the theorem fails to hold. Then there exists a compact set $K \subset \Omega$ and a sequence ϕ_j in $C_c^\infty(K)$ such that

$$|\langle u, \phi_j \rangle| > j \sum_{|\alpha| \leq j} \|\partial^\alpha \phi_j\|_{L^\infty}.$$

By homogeneity we may assume $\langle u, \phi_j \rangle = 1$ for all j. But then $\|\partial^\alpha \phi_j\|_{L^\infty} \to 0$ as $j \to \infty$ for all α, so $\phi_j \to 0$ in $C_c^\infty(\Omega)$. But $\langle u, \phi_j \rangle$ does not converge to 0, so $u \notin \mathcal{D}'(\Omega)$. □

2.4.2 Operations on distributions

Many of the usual operations on functions carry over to distributions. For example, to motivate the definition of differentiation of a distribution, consider smooth functions u, v on Ω, and let α be any multi-index. Then

$$v = \partial^\alpha u \quad \Longleftrightarrow \quad \int v\phi\, dx = (-1)^{|\alpha|} \int u \partial^\alpha \phi\, dx,$$

as follows from repeated integration by parts. Consequently, for any $u \in \mathcal{D}'(\Omega)$, we *define* its partial derivative $\partial^\alpha u$ as a distribution by

$$\langle \partial^\alpha u, \phi \rangle = \left\langle u, (-1)^{|\alpha|} \partial^\alpha \phi \right\rangle.$$

Exercise 8. Show that $\partial^\alpha u$ as defined above is actually a distribution on Ω, and that

$$u_j \to u \quad \text{in} \quad \mathcal{D}'(\Omega) \qquad \Longrightarrow \qquad \partial^\alpha u_j \to \partial^\alpha u \quad \text{in} \quad \mathcal{D}'(\Omega).$$

(See the definition (2.20).)

Multiplication of $u \in \mathcal{D}'(\Omega)$ with a function $\psi \in C^\infty(\Omega)$ (note that ψ is not required to be compactly supported) is defined by

$$\langle \psi u, \phi \rangle = \langle u, \psi\phi \rangle .$$

To see that this defines a distribution, we have to check that

$$\phi_j \to \phi \quad \text{in} \quad C_c^\infty(\Omega) \qquad \Longrightarrow \qquad \psi\phi_j \to \psi\phi \quad \text{in} \quad C_c^\infty(\Omega).$$

This is immediate from the product rule:

$$\partial^\alpha(\psi\phi) = \sum_{\beta+\gamma=\alpha} \frac{\alpha}{\beta!\gamma!} \partial^\beta\psi\partial^\gamma\phi. \tag{2.23}$$

(Here $\alpha! = \alpha_1!\alpha_2!\cdots\alpha_n!$ for a multi-index $\alpha = (\alpha_1, \dots, \alpha_n)$.)

Exercise 9. Prove that the product rule holds for the product of a distribution u with a smooth function ψ.

Next, we recall the definitions of reflection, translation and convolution with a smooth function. For simplicity, we take $\Omega = \mathbb{R}^n$ from now on. Also, we write C_c^∞ instead of $C_c^\infty(\mathbb{R}^n)$ and $\mathcal{D}'$ instead of $\mathcal{D}'(\mathbb{R}^n)$.

The *reflection* of a test function ϕ is the function $\phi\tilde{}(x) = \phi(-x)$. Note that

$$\int u(-x)\phi(x)\,dx = \int u(x)\phi(-x)\,dx.$$

This holds for any locally integrable function, and we take it as the definition of $u\tilde{}$ for any $u \in \mathcal{D}'$:

$$\langle u\tilde{}, \phi \rangle = \langle u, \phi\tilde{} \rangle .$$

Let $y \in \mathbb{R}^n$, and define the *translation* of $\phi \in C_c^\infty$ by $\phi_y(x) = \phi(x+y)$. Since

$$\int u(x+y)\phi(x)\,dx = \int u(x)\phi(x-y)\,dx$$

when u is locally integrable, we define u_y for $u \in \mathcal{D}'$ by

$$\langle u_y, \phi \rangle = \langle u, \phi_{-y} \rangle .$$

Finally, we consider the convolution of a distribution with a test function. If u is locally integrable and ψ is a test function, then

$$u * \psi(x) = \int u(y)\psi(x-y)\,dx,$$

and since $\psi(x-y) = \psi\tilde{}(y-x) = (\psi\tilde{})_{-x}(y)$, this leads us to define

$$u * \psi(x) = \langle\, u, (\psi\tilde{})_{-x} \,\rangle$$

for $u \in \mathcal{D}'$, $\psi \in C_c^\infty$. A more suggestive notation is

$$u * \psi(x) = \langle\, u, \psi(x - \cdot) \,\rangle\,.$$

Exercise 10. Show that $u * \psi$ is C^∞ on $\mathbb{R}^n$ and that

$$\partial^\alpha(u * \psi) = u * \partial^\alpha \psi = (\partial^\alpha u) * \psi$$

for all multi-indices α. (*Hints:* The second equality is trivial. To prove the first equality, note that by induction, it suffices to prove $\partial_j(u * \psi) = u * \partial_j \psi$ for $j = 1, \ldots, n$; use exercise 5.)

There is another, equivalent, definition of the convolution. It is suggested by the fact that

$$\int u * \psi(x)\phi(x)\,dx = \int u(y)\psi\tilde{} * \phi(y)\,dy$$

for sufficiently regular functions. It is therefore natural to try to define $\psi * u$ as a *distribution* by

$$\langle\, \psi * u, \phi \,\rangle = \langle\, u, \psi\tilde{} * \phi \,\rangle$$

for any $\psi \in C_c^\infty$ and $u \in \mathcal{D}'$. It is not hard to show that this definition agrees with our previous definition. In other words,

$$\langle\, u, \psi\tilde{} * \phi \,\rangle = \int \langle\, u, \psi(x - \cdot) \,\rangle\, \phi(x)\,dx$$

for all $\phi \in C_c^\infty$. That is to say,

$$\left\langle u, \int \psi(x - \cdot)\phi(x)\,dx \right\rangle = \int \langle\, u, \psi(x - \cdot) \,\rangle\, \phi(x)\,dx.$$

We leave it as an exercise to prove this using Riemann sums (see Folland).

2.4.3 Compactly supported distributions

We denote by $\vec{E}'(\Omega)$ the set of *compactly supported* distributions on Ω. Recall that $u \in \mathcal{D}'(\Omega)$ is compactly supported if there is a compact $K \subset \Omega$ such that $u = 0$ on the complement of K, in the sense that

$$\langle\, u, \phi \,\rangle = 0 \quad \text{for all} \quad \phi \in C_c^\infty(\Omega \setminus K).$$

The intersection of all such K is $\operatorname{supp} u$. (The smallest closed set outside of which u vanishes.)

We remark that u can be extended to a linear functional on $C^\infty(\mathbb{R}^n)$: Choose $\zeta \in C_c^\infty(\Omega)$ such that $\zeta = 1$ on a neighborhood of $\operatorname{supp} u$. Then evidently

$$\langle u, \phi \rangle = \langle u, \zeta\phi \rangle$$

for all $\phi \in C^\infty(\Omega)$, and we can take this identity as the definition of u for all $\phi \in C^\infty(\mathbb{R}^n)$. It is easy to check that this extension is independent of the choice of ζ. In particular, u can be regarded as an element of $\vec{E}'(\mathbb{R}^n)$. We often write $\vec{E}'$ instead of $\vec{E}'(\mathbb{R}^n)$.

Let u and ζ be as above. Applying Theorem 9 on the compact set $\operatorname{supp} \zeta \subset \Omega$, we see that there exist $C > 0$ and $N \in \mathcal{N}$ such that

$$|\langle u, \phi \rangle| \leq C \sum_{|\alpha| \leq N} \|\partial^\alpha(\zeta\phi)\|_{L^\infty}$$

for all $\phi \in C^\infty(R^n)$. By the product rule we then get

$$|\langle u, \phi \rangle| \leq C' \sum_{|\alpha| \leq N} \|\partial^\alpha \phi\|_{L^\infty(K)} \tag{2.24}$$

for all $\phi \in C^\infty(R^n)$, where $K = \operatorname{supp} \zeta$. Thus, every compactly supported distribution is of *finite order.*

Given $u \in \vec{E}'$, extended to $C^\infty(\mathbb{R}^n)$ as above, choose a compact set K which contains the support of u in its interior. Then it follows from

Next, observe that if $\psi \in C_c^\infty$ and $v \in \vec{E}'$, then $\psi * v \in \vec{E}'$. But we know $\psi * v \in C^\infty$, whence $\psi * v \in C_c^\infty$. We can therefore define the convolution $u * v$ of any $u \in \mathcal{D}'$ with any $v \in \vec{E}'$ by (cf. our second definition of the convolution)

$$\langle u * v, \phi \rangle = \langle u, v\tilde{} * \phi \rangle .$$

Exercise 11. With notation as above, prove that:

(a) $u * v \in \mathcal{D}'$ (*Hint:* You have to check that if $\phi_j \to \phi$ in C_c^∞ and $w \in \vec{E}'$, then $w * \phi_j \to w * \phi$ in C_c^∞. Recall that w is of finite order, so we have an estimate of the type (2.24).)

(b) $\partial^\alpha(u * v) = u * \partial^\alpha v = (\partial^\alpha u) * v$ for all multi-indices α.

2.4.4 Smooth approximation

It is an important fact that the test functions are dense in $\mathcal{D}'(\Omega)$.

Theorem 10. $C_c^\infty(\Omega)$ *is dense in* $\mathcal{D}'(\Omega)$.

See Folland for a complete proof. The idea is write Ω as a union of an increasing sequence of compact subsets K_j, and to choose $\zeta_j \in C_c^\infty(\Omega$ such that $\zeta_j = 1$ on K_j. Also let χ_ε be a smooth and compactly supported approximation of the identity. Given $u \in \mathcal{D}'(\Omega)$, one then defines

$$u_j = \chi_{\varepsilon_j} * (\zeta_j u), \tag{2.25}$$

where $\varepsilon_j \to 0$. Clearly $u_j \in C_c^\infty(\Omega)$ if one chooses ε_j sufficiently small, and it is not hard to check that $u_j \to u$ in $\mathcal{D}'(\Omega)$.

It is not hard to check that for the particular sequence u_j constructed above, one has estimates (2.22) with constants C_K and N_K independent of j. This turns out to be true for *any* convergent sequence of distributions however; see the next section.

Chapter 3

Week 3: Some facts about distributions

3.1 Continuity properties of $\mathcal{D}'(\Omega)$

Recall that for a linear functional $u : C_c^\infty(\Omega) \to \mathbb{C}$, the following statements are equivalent:

(a) $u \in \mathcal{D}'(\Omega)$.

(b) $\langle u, \phi_j \rangle \to \langle u, \phi \rangle$ whenever $\phi_j \to \phi$ in $C_c^\infty(\Omega)$.

(c) For every compact $K \subset \Omega$, there exist $C_k > 0$ and $N_K \in \mathcal{N}$ such that

$$|\langle u, \phi \rangle| \le C_K \sum_{|\alpha| \le N_K} \|\partial^\alpha \phi\|_{L^\infty} \quad \text{for all} \quad \phi \in C_c^\infty(K). \tag{3.1}$$

It is natural to pose the following

Question. *Suppose*

$$\begin{aligned} u_j &\to u \quad \textit{in} \quad \mathcal{D}'(\Omega), \\ \phi_j &\to \phi \quad \textit{in} \quad C_c^\infty(\Omega). \end{aligned}$$

Then is it true that

$$\langle u_j, \phi_j \rangle \to \langle u, \phi \rangle\,? \tag{3.2}$$

The answer is yes, but this is far from obvious.

When trying to prove (3.2), the natural course of action is to exploit the bilinearity of the pairing $\langle \cdot, \cdot \rangle$ and write

$$\langle u_j, \phi_j \rangle - \langle u, \phi \rangle = \langle u_j, \phi_j - \phi \rangle + \langle u_j - u, \phi \rangle\,.$$

Since $u_j \to u$, the second term on the right hand side converges to 0, but it is not at all clear that the first term does. Notice, however, that if the estimate (3.1) were to hold for all the u_j, with constants C_K and N_K *independent of* j, then also the first term on the right hand side of the above equation would converge to 0, since $\phi_j \to \phi$. This turns out to be true, which is quite remarkable, considering the fact that $\mathcal{D}'(\Omega)$ has the topology of pointwise convergence.

Theorem 11. *If $u_j \to u$ in $\mathcal{D}'(\Omega)$, then for every compact $K \subset \Omega$, there exist C_K and N_K, independent of j, such that (3.1) holds for all the u_j.*

This striking result comes out of the Uniform Boundedness Principle, which in fact implies a stronger statement (we will need this later):

Theorem 12. *Consider an indexed family $\{u_\lambda\}_{\lambda\in I} \subset \mathcal{D}'(\Omega)$. If*

$$\sup_{\lambda\in I} |\langle u_\lambda, \phi \rangle| < \infty \quad \textit{for all} \quad \phi \in C_c^\infty(\Omega),$$

then for every compact $K \subset \Omega$, there exist C_K and N_K such that

$$\sup_{\lambda\in I} |\langle u_\lambda, \phi \rangle| \le C_K \sum_{|\alpha|\le N_K} \|\partial^\alpha \phi\|_{L^\infty} \quad \textit{for all} \quad \phi \in C_c^\infty(K).$$

The proof is relegated to an appendix.

Let us note the following interesting corollary.

Corollary. *Let $\{u_j\}$ be a sequence in $\mathcal{D}'(\Omega)$. Suppose*

$$\lim_{j\to\infty} \langle u_j, \phi \rangle$$

exists for every $\phi \in C_c^\infty(\Omega)$, and denote the limit by $\langle u, \phi \rangle$. Then the map $u : C_c^\infty \to \mathbb{C}$ so defined belongs to $\mathcal{D}'(\Omega)$.

Proof. Obviously, u is a linear functional on $C_c^\infty(\Omega)$, and the continuity follows immediately from Theorem 12. □

This further implies:

Corollary. *$\mathcal{D}'(\Omega)$ is complete.*

In other words, every Cauchy sequence $\{u_j\}$ in $\mathcal{D}'(\Omega)$ has a limit in $\mathcal{D}'(\Omega)$. Indeed, since $u_j - u_k \to 0$ as $j, k \to \infty$, then $\langle u_j, \phi \rangle$ is a Cauchy sequence in $\mathbb{C}$ for every fixed test function ϕ; hence it converges in $\mathbb{C}$, and we can apply the previous corollary.

3.2 Time-dependent distributions

Consider a map

$$\mathbb{R} \to \mathcal{D}'(\mathbb{R}^n), \quad t \to u(t).$$

Since $\mathcal{D}'(\mathbb{R}^n)$ has a topology—namely the topology of pointwise convergence—it makes sense to talk about continuity or differentiability of u as a function of t. Thus, u is continuous at t_0 iff $t \to \langle u(t), \phi \rangle$ is continuous at t_0 for every $\phi \in C_c^\infty(\mathbb{R}^n)$, and u is differentiable at t_0 iff there exists a $v \in \mathcal{D}'(\mathbb{R}^n)$ such that for every $\phi \in C_c^\infty$,

$$\frac{d}{dt} \langle u(t), \phi \rangle |_{t=t_0} = \langle v, \phi \rangle .$$

We then write $u'(t) = v$. If $u, u', \dots, u^{(k)}$ exist and are continuous on $\mathbb{R}$, we say that u is of class C^k and write $u \in C^k(\mathbb{R}, \mathcal{D}')$ or just $u \in C^k$. If $u \in C^k$ for all k, we say $u \in C^\infty$.

Proposition 1. *Every $u \in C(\mathbb{R}, \mathcal{D}'(\mathbb{R}^n))$ defines a distribution on $\mathbb{R}^{1+n}$ by*

$$\langle u, \psi \rangle = \int_{\mathbb{R}} \langle u(t), \psi(t, \cdot) \rangle \, dt \quad \textit{for} \quad \psi \in C_c^\infty(\mathbb{R}^{1+n}).$$

Proof. First observe that the integrand is a continuous and compactly supported function of t, so the integral exists. The functional u thus defined belongs to $\mathcal{D}'(\mathbb{R}^{1+n})$ iff (3.1) holds for every compact $K \subset \mathbb{R}^{1+n}$. It suffices to take K of the form $I \times K'$ where $I \subset \mathbb{R}$ and $K' \subset \mathbb{R}^n$ are compact. But by Theorem 12 there exist $C_{I,K'} > 0$ and $N_{I,K'} \in \mathcal{N}$ such that

$$\sup_{t \in I} |\langle u(t), \phi \rangle| \le C_{I,K'} \sum_{|\alpha| \le N_{I,K'}} \|\partial^\alpha \phi\|_{L^\infty} \quad \text{for all} \quad \phi \in C_c^\infty(K'),$$

Apply this for each $t \in I$ with $\phi = \psi(t, \cdot)$ and then integrate in t to get the desired estimate. □

Now assume $u \in C^1(\mathbb{R}, \mathcal{D}'(\mathbb{R}^n))$. Then u and u' can both be interpreted as elements of $\mathcal{D}'(\mathbb{R}^{1+n})$, and the question naturally arises whether

$$\partial_t u = u' \tag{3.3}$$

in the sense of $\mathcal{D}'(\mathbb{R}^{1+n})$. The answer is yes. Indeed, (3.3) is equivalent to the condition that, for every $\psi \in C_c^\infty(\mathbb{R}^{1+n})$,

$$\int_{\mathbb{R}} \langle u(t), (-1)\partial_t \psi(t, \cdot) \rangle \, dt = \int_{\mathbb{R}} \langle u'(t), \psi(t, \cdot) \rangle \, dt.$$

But since u is C^1, we have

$$\frac{d}{dt} \langle u(t), \psi(t, \cdot) \rangle = \langle u'(t), \psi(t, \cdot) \rangle + \langle u(t), \partial_t \psi(t, \cdot) \rangle , \tag{3.4}$$

and integration in t gives the desired identity.

Exercise 12. Prove (3.4). Note that if $f(t) = \langle u(t), \psi(t, \cdot) \rangle$, then

$$\frac{1}{h}\{f(t+h) - f(t)\}$$
$$= \left\langle \frac{1}{h}\{u(t+h) - u(t)\}, \psi(t, \cdot) \right\rangle + \left\langle u(t+h), \frac{1}{h}\{\psi(t+h, \cdot) - \psi(t, \cdot)\} \right\rangle .$$

Use Theorem 12 on the last term (cf. the discussion in section 1).

3.3 Distributional solutions of $\Box u = 0$.

We consider again the Cauchy problem on $\mathbb{R}^{1+n}$,

$$\Box u = 0, \quad u\big|_{t=0} = f, \quad \partial_t u\big|_{t=0} = g. \tag{3.5}$$

First observe that the equation $\Box u = 0$ makes sense for $u \in \mathcal{D}'(\mathbb{R}^{1+n})$. In fact,

$$\Box u = 0 \iff \langle u, \Box \phi \rangle = 0 \quad \text{for all} \quad \phi \in C_c^\infty(\mathbb{R}^{1+n}).$$

We could also replace $\mathbb{R}^{1+n}$ by any open subset. However, it does not make sense to say that u satisfies the initial conditions, since in general we cannot restrict u to the plane $t = 0$. But if u is a time-dependent distribution of class C^1 as defined in the previous section, then the initial condition is clearly meaningful (cf. also (3.3)).

Theorem 13. *For all $f, g \in \mathcal{D}'(\mathbb{R}^n)$ there exists a time-dependent distribution $u \in C^\infty(\mathbb{R}, \mathcal{D}'(\mathbb{R}^n))$ which solves the Cauchy problem* (3.5).

The solution is also unique; we leave this as an exercise for next week.

To prove the theorem, one simply takes the representation formulas for smooth initial data, and see that they make sense also when f and g are distributions. In fact, in every dimension n, the solution formulas can be written

$$u(t) = W'(t) * f + W(t) * g, \tag{3.6}$$

where

$$W \in C^\infty\big(\mathbb{R}, \vec{E}'(\mathbb{R}^n)\big), \quad W(0) = W''(0) = 0, \quad W'(0) = \delta. \tag{3.7}$$

(Here δ is the point mass at the origin in $\mathbb{R}^n$.) Recall that $\vec{E}'(\mathbb{R}^n)$ is the space of compactly supported distributions. In fact, we shall see that

$$\operatorname{supp} W(t) \subset \{x : |x| \le |t|\},$$

and, moreover,

$$\operatorname{supp} W(t) \subset \{x : |x| = |t|\} \quad \text{if } n \text{ is odd, } n \ge 3,$$

which is merely a statement of Huygens' principle.

Let us for the moment simply assume the existence of $W(t)$ with the stated properties, and show that (3.6) defines a solution of (3.5) for arbitrary $f, g \in \mathcal{D}'(\mathbb{R}^n)$.

First, since $W(t)$ is compactly supported for each t, the convolutions in (3.6) are well-defined for all $f, g \in \mathcal{D}'(\mathbb{R}^n)$. Thus, (3.6) defines a smooth time-dependent distribution (see Exercise 2 below).

Secondly, the initial conditions are satisfied, since

$$u(0) = W'(0) * f + W(0) * g = \delta * f = f$$
$$u'(0) = W''(0) * f + W'(0) * g = \delta * g = g.$$

Finally, to see that $\Box u = 0$ in the sense of distributions, we use a simple approximation argument. Choose sequences $f_j, g_j \in C_c^\infty(\mathbb{R}^n)$ such that $f_j \to f$ and $g_j \to g$ in $\mathcal{D}'(\mathbb{R}^n)$. Let u_j be the solution of (3.5) with initial data f_j, g_j. Then

$$u_j(t) = W'(t) * f_j + W(t) * g_j,$$

whence (see Exercise 2)

$$u_j \to u \quad \text{in} \quad \mathcal{D}'(\mathbb{R}^{1+n}).$$

But then

$$\Box u_j \to \Box u \quad \text{in} \quad \mathcal{D}'(\mathbb{R}^{1+n}),$$

and since $\Box u_j = 0$ for all j, we must have $\Box u = 0$.

Exercise 13. Given $v \in C^\infty\big(\mathbb{R}, \vec{E}'(\mathbb{R}^n)\big)$ and $f \in \mathcal{D}'(\mathbb{R}^n)$, define

$$u(t) = v(t) * f$$

and prove:

(a) $u \in C^\infty\big(\mathbb{R}, \mathcal{D}'(\mathbb{R}^n)\big)$.

(b) If $f_j \to f$ in $\mathcal{D}'(\mathbb{R}^n)$, and we set $u_j(t) = v(t) * f_j$, then $u_j \to u$ in the sense of $\mathcal{D}'(\mathbb{R}^{1+n})$.

(*Hint:* Use the fact [see the next exercise] that the convolution product of two distributions is continuous in both arguments. For part (b), use the Dominated Convergence Theorem, recalling that Theorem 12 furnishes estimates of the type (3.1) which are uniform on compact intervals in t.)

Exercise 14. Recall that for $u \in \mathcal{D}'$ and $v \in \vec{E}'$, we defined $u * v \in \mathcal{D}'$ by

$$\langle u * v, \phi \rangle = \langle u, v\tilde{} * \phi \rangle \quad \text{for} \quad \phi \in C_c^\infty. \tag{3.8}$$

Recall also that if $v \in C_c^\infty$, then $u * v$ agrees with the C^∞ function $u * v(x) = \langle u, v(x - \cdot) \rangle$. In other words, we have the identity

$$\int \langle u, v(x - \cdot) \rangle \, \phi(x) \, dx = \langle u, v\tilde{} * \phi \rangle . \tag{3.9}$$

The purpose of this exercise is to *extend* the convolution product to

$$(\mathcal{D}' \times \vec{E}') \cup (\vec{E}' \times \mathcal{D}')$$

and prove that it is commutative, so that $u * v = v * u$. Moreover, we want to prove that the product is continuous in both arguments.

We break this into a number of steps. At the outset, let us record some facts that will be used:

(i) Any $u \in \vec{E}'$ can be extended to a functional on C^∞ by $\langle u, \phi \rangle = \langle u, \zeta\phi \rangle$ where $\zeta \in C_c^\infty$ and $\zeta = 1$ on $\operatorname{supp} u$.

(ii) Assume $u_j, u \in \vec{E}'$. The statement $u_j \to u$ in $\vec{E}'$ means that $\langle u_j, \phi \rangle \to \langle u, \phi \rangle$ for all $\phi \in C^\infty$. Note that if the u_j are all supported in a fixed compact set K, then this is equivalent to saying that $\langle u_j, \phi \rangle \to \langle u, \phi \rangle$ for all $\phi \in C_c^\infty$. (Choose a cut-off $\zeta \in C_c^\infty$ such that $\zeta = 1$ on K.)

(iii) If $u \in \vec{E}'$ and $\phi \in C^\infty$, then $u * \phi$ is a C^∞ function, given by $u * \phi(x) = \langle u, \phi(x - \cdot) \rangle$.

(iv) Assume $\chi \in C_c^\infty$ and $\int_{\mathbb{R}^n} \chi(x)\, dx = 1$. Set $\chi_\varepsilon = \varepsilon^{-n}\chi(x/\varepsilon)$. If $u \in \vec{E}'$, then $u * \chi_\varepsilon \to u$ in $\vec{E}'$ as $\varepsilon \to 0$.

Now prove the following:

(a) Equation (3.8) defines a distribution $u * v$ whenever $u \in \vec{E}'$ and $v \in D'$. Thus, the convolution product is extended from $\mathcal{D}' \times \vec{E}'$ to $(\mathcal{D}' \times \vec{E}') \cup (\vec{E}' \times \mathcal{D}')$

(*Hint:* In view of facts (i) and (iii) above, the right hand side of (3.8) makes sense. Proving that the linear functional $u * v$ so defined is a distribution, boils down to the following: If $w \in \mathcal{D}'$ and $\phi_j \to \phi$ in C_c^∞, then $w * \phi_j$ converges uniformly on *compact sets* to $w * \phi$.)

(b) Prove that if $u \in \vec{E}'$, $\phi \in C^\infty$ and $\psi \in C_c^\infty$, then

$$(u * \phi) * \psi = u * (\phi * \psi).$$

Note that both sides are C^∞ functions. (*Hint:* This can be reduced to the identity (3.9).)

(c) If $u, v \in \mathcal{D}'$ and at least one of them has compact support, then

$$\langle u, v\tilde{} * \phi \rangle = \langle v, u\tilde{} * \phi \rangle \quad \text{for} \quad \phi \in C_c^\infty. \tag{3.10}$$

(*Hint:* By symmetry, we may assume $u \in \mathcal{D}'$ and $v \in \vec{E}'$. Show first that (3.10) holds when $v \in C_c^\infty$; this can be reduced to the identity (3.9), since $v\tilde{} * \phi = \phi * v\tilde{}$. For general $v \in \vec{E}'$, use fact (iv) and part (b).)

(d) Convolution is commutative. That is, $u * v = v * u$ whenever $u, v \in \mathcal{D}'$ and at least one of them has compact support. (*Hint:* Show that this is equivalent to (3.10).)

(e) The convolution product is continuous in both arguments. (*Hint:* By commutativity, it suffices to prove two things. First, if $u_j \to u$ in $\mathcal{D}'$ and $v \in \vec{E}'$, then $u_j * v \to u * v$ in $\mathcal{D}'$. This follows easily from (3.8). Secondly, one has to show the same thing if $u_j \to u$ in $\vec{E}'$ and $v \in \mathcal{D}'$. Again this follows from (3.8), since the assumption $u_j \to u$ in $\vec{E}'$ means that [see fact (ii)] $\langle u_j - u, \phi \rangle \to 0$ as $j \to \infty$ for all $\phi \in C^\infty$.)

3.4 What is $W(t)$?

It remains to find $W(t)$ satisfying (3.7) and such that the solution of (3.5) is given by (3.6) for all $f, g \in C_c^\infty(\mathbb{R}^n)$. Of course, $W(t)$ depends on the space dimension n. We only consider $n = 1, 2, 3$, leaving the higher dimensional cases as an exercise.

Dimension $n = 1$. In this case we have d'Alembert's formula:

$$u(t,x) = \frac{1}{2}[f(x+t) + f(x-t)] + \frac{1}{2}\int_{x-t}^{x+t} g.$$

Define $W(t) \in \mathcal{D}'(\mathbb{R})$ by

$$\langle\, W(t), \phi\,\rangle = \frac{1}{2}\int_{-t}^{t} \phi \quad \text{for} \quad \phi \in C_c^\infty(\mathbb{R}).$$

It is easy to check that $[W(t) * g](x) = \frac{1}{2}\int_{x-t}^{x+t} g$ for $g \in C_c^\infty$.

Since

$$\frac{d}{dt}\left(\int_{-t}^{t} \phi\right) = \phi(t) + \phi(-t),$$

we have

$$W'(t) = \frac{1}{2}[\delta(\cdot + t) + \delta(\cdot - t)].$$

In particular, this implies

$$[W'(t) * f](x) = \frac{1}{2}[f(x+t) + f(x-t)]$$

for $f \in C_c^\infty$. The higher derivatives are given by

$$\left\langle W^{(k)}(t), \phi \right\rangle = \frac{1}{2}\left[\phi^{(k-1)}(t) + (-1)^{k-1}\phi^{(k-1)}(-t)\right].$$

We conclude that $W(t)$ has all the required properties.

Dimension $n = 3$. In this case the solution of (3.5) for $f, g \in C_c^\infty$ is

$$u(t,x) = \partial_t\left(\frac{t}{4\pi}\int_{y\in S^2} f(x - ty)\, d\sigma(y)\right) + \frac{t}{4\pi}\int_{y\in S^2} g(x - ty)\, d\sigma(y).$$

(Here we have changed variables $y \to -y$, which does not affect the value of the integral.) The solution is therefore of the form (3.6), if we take

$$W(t) = t\Sigma(t),$$

where $\Sigma \in C^\infty(\mathbb{R}, \vec{E}'(\mathbb{R}^3))$ is defined by

$$\langle\, \Sigma(t), \phi\,\rangle = \frac{1}{4\pi}\int_{y\in S^2} \phi(ty)\, d\sigma(y).$$

Then

$$\langle \Sigma'(t), \phi \rangle = \frac{1}{4\pi} \int_{y \in S^2} \nabla\phi(ty) \cdot y \, d\sigma(y),$$

and the higher derivatives are given by

$$\left\langle \Sigma^{(k)}(t), \phi \right\rangle = \frac{1}{4\pi} \int_{y \in S^2} D^k\phi(ty)(y, \ldots, y) \, d\sigma(y),$$

where

$$D^k\phi(ty)(y, \ldots, y) = \sum_{j_1, \ldots, j_k = 1}^{n} \partial_{j_1} \cdots \partial_{j_k} \phi(ty) \, y_{j_1} \cdots y_{j_k}.$$

Observe that

$$\Sigma(0) = \delta, \quad \Sigma'(0) = 0.$$

(The latter holds by the Divergence Theorem.)

Since $\Sigma \in C^\infty(\mathbb{R}, \vec{E}'(\mathbb{R}^3))$, it follows that $W \in C^\infty(\mathbb{R}, \vec{E}'(\mathbb{R}^3))$, and we have

$$\begin{aligned} W'(t) &= \Sigma(t) + t\Sigma'(t) && \Longrightarrow W'(0) = \delta, \\ W'(t) &= 2\Sigma'(t) + t\Sigma''(t) && \Longrightarrow W''(0) = 0. \end{aligned}$$

Thus (3.7) is satisfied.

Dimension $n = 2$. The solution is now given by

$$u(t,x) = \partial_t \left(\frac{t}{2\pi} \int_{|y|<1} f(x - ty) \frac{dy}{\sqrt{1 - |y|^2}} \right) + \frac{t}{2\pi} \int_{|y|<1} g(x - ty) \frac{dy}{\sqrt{1 - |y|^2}},$$

which is of the form (3.6) if we set

$$W(t) = t\Theta(t),$$

where $\Theta \in C^\infty(\mathbb{R}, \vec{E}'(\mathbb{R}^2))$ is given by

$$\langle \Theta(t), \phi \rangle = \frac{1}{2\pi} \int_{|y|<1} \phi(ty) \frac{dy}{\sqrt{1 - |y|^2}}.$$

Remark. If we consider ϕ as a function on $\mathbb{R}^3$ which is independent of y_3, then $\langle \Theta(t), \phi \rangle$ is just $\langle \Sigma(t), \phi \rangle$.

For $k \geq 1$,

$$\left\langle \Theta^{(k)}(t), \phi \right\rangle = \frac{1}{2\pi} \int_{|y|<1} D^k\phi(ty)(y, \ldots, y) \frac{dy}{\sqrt{1 - |y|^2}}.$$

Moreover,

$$\Theta(0) = \delta, \quad \Theta'(0) = 0.$$

(The latter follows by the above remark, since $\Sigma'(0) = 0$.)

Since $\Theta \in C^\infty(\mathbb{R}, \vec{E}'(\mathbb{R}^2))$, it follows that $W \in C^\infty(\mathbb{R}, \vec{E}'(\mathbb{R}^2))$, and we have

$$\begin{aligned} W'(t) &= \Theta(t) + t\Theta'(t) &&\implies W'(0) = \delta, \\ W'(t) &= 2\Theta'(t) + t\Theta''(t) &&\implies W''(0) = 0. \end{aligned}$$

so (3.7) holds.

Exercise 15. Find $W(t)$ for $n = 5, 7, \ldots$ and $n = 4, 6, \ldots$.

3.5 Proof of Theorem 2

We are given an indexed family $\{u_\lambda\}_{\lambda \in I} \subset \mathcal{D}'(\Omega)$ such that

$$\sup_{\lambda \in I} |\langle u_\lambda, \phi \rangle| < \infty \quad \text{for all} \quad \phi \in C_c^\infty(\Omega),$$

and we want to prove that for every compact $K \subset \Omega$, there exist $C_K > 0$ and $N_K \in \mathcal{N}$ such that

$$\sup_{\lambda \in I} |\langle u_\lambda, \phi \rangle| \le C_K \sum_{|\alpha| \le N_K} \|\partial^\alpha \phi\|_{L^\infty} \quad \text{for all} \quad \phi \in C_c^\infty(K).$$

To prove this, fix the set K. The space $C_c^\infty(K)$ is a *Fréchet space*: a complete Hausdorff topological vector space whose topology is induced by a countable family of seminorms. In this case the seminorms are

$$\|\phi\|_{(\alpha)} = \|\partial^\alpha \phi\|_{L^\infty},$$

where α runs over the set of multi-indices. (See Folland's Real Analysis, section 5.4, for a brief discussion of topological vector spaces.) In fact, every Fréchet space is a complete metric space. Convergence in this topology on $C_c^\infty(K)$ means the following:

$$\phi_j \to \phi \quad \text{in} \quad C_c^\infty \quad \iff \quad \|\phi_j - \phi\|_{(\alpha)} \to 0 \quad \text{for all} \quad \alpha.$$

As a consequence, we see that a linear functional $u : C_c^\infty(\Omega) \to \mathbb{C}$ belongs to $\mathcal{D}'(\Omega)$ iff $u|C_c^\infty(K)$ is continuous for every compact $K \subset \Omega$.

Returning to the proof of the theorem, then, we have that $u_\lambda|C_c^\infty(K)$ is continuous for every $\lambda \in I$. But the conclusion of the theorem then follows immediately from

Theorem. ***(Uniform Boundedness Principle.)*** *Assume that*

- *$\mathcal{X}$ is a Fréchet space whose topology is given by a countable family of seminorms $\{p_n\}_1^\infty$;*
- *$\mathcal{Y}$ is a normed vector space;*

- $\{T_\lambda\}_{\lambda\in I}$ *is a family of continuous linear operators from* $\mathcal{X}$ *into* $\mathcal{Y}$ *which is pointwise bounded, in the sense that*

$$\sup_{\lambda\in I} \|T_\lambda x\| < \infty \quad \textit{for all} \quad x \in \mathcal{X}.$$

Then there exist $C > 0$ *and* $N \in \mathcal{N}$ *such that*

$$\sup_{\lambda\in I} \|T_\lambda x\| \le C \sum_1^N p_n(x) \quad \textit{for all} \quad x \in \mathcal{X}.$$

The theorem is usually stated for the special case where $\mathcal{X}$ is a Banach space.

The proof is based on the Baire Category Theorem, which states that if X is a complete metric space and $X = \bigcup_1^\infty E_j$ where each E_j is closed in X, then at least one E_j has a nonempty interior. Recall that every Fréchet space is a complete metric space.

In the present situation, one defines the sets E_j by

$$E_j = \left\{ x \in \mathcal{X} : \sup_{\lambda\in I} \|T_\lambda x\| \le j \right\}.$$

Since the T_λ are continuous, each E_j is closed, and since the family $\{T_\lambda\}$ is pointwise bounded, we see that $\mathcal{X} = \bigcup_1^\infty E_j$. Therefore, some E_j has nonempty interior, by the Baire Category Theorem.

One can show that for every $x_0 \in \mathcal{X}$, the finite intersections of the sets

$$U_{x_0 n\varepsilon} = \{x \in \mathcal{X} : p_n(x - x_0) < \varepsilon\}, \quad n \in \mathcal{N}, \quad \varepsilon > 0,$$

form a neighborhood base at x_0. (See Folland.)

Since E_j has nonempty interior, it then follows that there exist $x_0 \in \mathcal{X}$, $N \in \mathcal{N}$ and $\varepsilon > 0$ such that

$$\bigcap_{n=1}^N U_{x_0 n\varepsilon} \subset E_j.$$

But this implies that E_{2j} contains an open neighborhood of the origin in $\mathcal{X}$, essentially by translation.

In fact, if

$$x \in \bigcap_{n=1}^N U_{0n\varepsilon},$$

that is to say, if $p_n(x) < \varepsilon$ for $n = 1, \ldots, N$, then obviously

$$x_0 + x \in \bigcap_{n=1}^N U_{x_0 n\varepsilon},$$

whence

$$\|T_\lambda x\| \le \|T_\lambda(x_0 + x)\| + \|T_\lambda x_0\| \le 2j$$

for all λ. Thus $x \in E_{2j}$.

Fix $\delta > 0$ and $x \in \mathcal{X}$. Then

$$z = \varepsilon \left(\delta + \sum_1^N p_n(x) \right)^{-1} x \in \bigcap_{n=1}^N U_{0n\varepsilon},$$

because $p_n(z) < \varepsilon$ for $n = 1, \dots, N$. Therefore,

$$\sup_\lambda \|T_\lambda z\| \leq 2j,$$

which by linearity and homogeneity implies

$$\sup_\lambda \|T_\lambda x\| \leq \frac{2j}{\varepsilon} \left(\delta + \sum_1^N p_n(x) \right).$$

Since $\delta > 0$ was arbitrary, we get the desired conclusion.

Chapter 4

Week 4: More on distributional solutions

4.1 Uniqueness of distributional solutions

Fix $T > 0$, and assume that $u \in C^1([0,T], \mathcal{D}'(\mathbb{R}^n))$ solves

$$\begin{cases} \Box u = F \quad \text{on} \quad S_T = (0,T) \times \mathbb{R}^n, \\ u\big|_{t=0} = f, \quad \partial_t u = g, \end{cases} \tag{4.1}$$

where $F \in \mathcal{D}'(S_T)$ and $f, g \in \mathcal{D}'(\mathbb{R}^n)$. (The equation $\Box u = F$ is understood in the sense of distributions on S_T.)

Theorem 14. *The solution of* (4.1) *is unique in the class* $C^1([0,T], \mathcal{D}'(\mathbb{R}^n))$.

Since $\Box$ is a linear operator, it suffices to prove that if $f = g = 0$ and $F = 0$, then u must vanish on S_T. We leave the proof as an exercise:

Exercise 16. Prove the following statements.

(a) If $u \in C^2([0,T], \mathcal{D}'(\mathbb{R}^n))$, then for any $\phi \in C^\infty(\mathbb{R}^{1+n})$ such that

$$\phi(t, \cdot) \in C_c^\infty(\mathbb{R}^n) \quad \text{for all} \quad t$$

we have

$$\int_0^T \langle\, u(t), \Box\phi(t,\cdot) \,\rangle \; dt = \int_0^T \langle\, \Box u(t), \phi(t,\cdot) \,\rangle \; dt + \langle\, u(t), \partial_t \phi(t,\cdot) \,\rangle \,|_0^T - \langle\, u'(t), \phi(t,\cdot) \,\rangle \,|_0^T,$$

where we write $a(t)|_0^T = a(T) - a(t)$.

Hint: Since $\Box u(t) = u''(t) - \Delta u(t)$, it suffices to show

$$\int_0^T \left\langle\, u(t), \partial_t^2 \phi(t,\cdot) \,\right\rangle \, dt = \int_0^T \langle\, u''(t), \phi(t,\cdot) \,\rangle \, dt + \langle\, u(t), \partial_t \phi(t,\cdot) \,\rangle \,|_0^T - \langle\, u'(t), \phi(t,\cdot) \,\rangle \,|_0^T .$$

Integrate by parts twice to prove this (see equation (4) in the notes from last week).

(b) If $u \in C^1([0,T], \mathcal{D}'(\mathbb{R}^n))$ and $\Box u = 0$ on S_T, then $u \in C^2([0,T], \mathcal{D}'(\mathbb{R}^n))$. [In fact, $u \in C^\infty([0,T], \mathcal{D}'(\mathbb{R}^n))$.]

Hint: $\partial_t^2 u = \Delta u$, and the right hand side belongs to $C^1([0,T], \mathcal{D}'(\mathbb{R}^n))$. By equation (3) from last week we have $\partial_t^2 u = \partial_t(u')$. Now apply Proposition 2 below.

(c) Assume that $u \in C^1([0,T], \mathcal{D}'(\mathbb{R}^n))$ solves (4.1) with $f = g = 0$ and $F = 0$. Prove that u vanishes on S_T.

Hint: By part (b), $u \in C^2(\mathbb{R}, \mathcal{D}')$, so we can apply part (a), with ϕ the solution of the following Cauchy problem with initial data at $t = T$:

$$\begin{cases} \Box \phi = 0 \quad \text{on} \quad \mathbb{R}^{1+n} \\ \phi(T,\cdot) = \phi_0, \quad \partial_t \phi(T,\cdot) = \phi_1, \end{cases}$$

for arbitrary $\phi_0, \phi_1 \in C_c^\infty$. Conclude that $u(T) = u'(T) = 0$.

Proposition 2. *Let $u, v \in C([0,T], \mathcal{D}'(\mathbb{R}^n))$ and set $S_T = (0,T) \times \mathbb{R}^n$. If*

$$\partial_t u = v \quad \textit{in the sense of} \quad \mathcal{D}'(S_T),$$

then $u \in C^1([0,T], \mathcal{D}'(\mathbb{R}^n))$ and $u' = v$.

Proof. We are given that

$$\int \langle\, u, (-1)\partial_t \psi(t,\cdot) \,\rangle \, dt = \int \langle\, v, \psi(t,\cdot) \,\rangle \, dt$$

for all $\psi \in C_c^\infty(S_T)$. Now take

$$\psi(t,x) = \theta(t)\phi(x)$$

where $\theta \in C_c^\infty((0,T))$ and $\phi \in C_c^\infty(\mathbb{R}^n)$. Considering ϕ to be fixed, the above equation then reads

$$\int \{ f(t)\theta'(t) + g(t)\theta(t) \} \, dt = 0 \quad \text{for all} \quad \theta \in C_c^\infty\big((0,T)\big),$$

where $f, g \in C([0,T])$ are given by

$$f(t) = \langle\, u(t), \phi \,\rangle \,, \quad g(t) = \langle\, v(t), \phi \,\rangle \,.$$

But this means that

$$f' = g \quad \text{in the sense of} \quad \mathcal{D}'\big((0,T)\big),$$

and since f and g are both continuous, it follows from Proposition 3 below that $f' = g$ in the classical sense, that is,

$$\frac{d}{dt}\langle\, u(t), \phi \,\rangle = \langle\, v(t), \phi \,\rangle\,.$$

Since $\phi \in C_c^\infty(\mathbb{R}^n)$ was arbitrary, this finishes the proof. □

Proposition 3. *Let $\Omega \subset \mathbb{R}^n$ be open. If $u \in C(\Omega)$ and the distributional derivatives $\partial^\alpha u$ are also in $C(\Omega)$ for all $|\alpha| \le k$, then $u \in C^k(\Omega)$.*

Proof. By induction, it suffices to consider $k = 1$. The problem is then to prove that $\partial_1 u, \dots, \partial_n u$ exist in the classical sense at every point of Ω, and agree with the distributional derivatives, which are assumed to be continuous. Since we only consider partial derivatives in the coordinate directions and at a given point, it suffices to take $n = 1$ and assume that u is a function in $C([a,b])$ whose distributional derivative on (a,b) agrees with a function $v \in C([a,b])$. Thus,

$$\int u(x)(-1)\phi'(x)\,dx = \int v(x)\phi(x)\,dx \quad \text{for all} \quad \phi \in C_c^\infty\big((a,b)\big).$$

Now define $U : [a,b] \to \mathbb{C}$ by

$$U(x) = \int_a^x v(t)\,dt + u(a).$$

Note that $U' = v$ in the classical sense. Thus $(U - u)' = 0$ in the sense of distributions, so by Proposition 4 below, $U - u$ is a constant, and since $U(a) = u(a)$ it follows that $u = U \in C^1$. □

Proposition 4. *If $u \in \mathcal{D}'(I)$, where $I \subset \mathbb{R}$ is an open interval. Then*

$$u' = 0 \implies u = \textit{const.}$$

Proof. Fix $\phi \in C_c^\infty(I)$ with $\int \phi = 1$. If u were a constant, that constant would necessarily be $\langle\, u, \phi \,\rangle$. Thus, we want to prove that

$$\langle\, u - \langle\, u, \phi \,\rangle\,, \psi \,\rangle = 0 \quad \text{for all} \quad \psi \in C_c^\infty(I).$$

First note that the left hand side equals $\langle\, u, \theta \,\rangle$, where

$$\theta = \psi - \left(\int \psi\right)\phi.$$

But $\int \theta = 0$, so if we set

$$\eta(x) = \int_{-\infty}^x \theta,$$

then $\eta \in C_c^\infty(I)$ and $\eta' = \theta$. Finally, then,

$$\langle\, u - \langle\, u, \phi \,\rangle\,, \psi \,\rangle = \langle\, u, \eta' \,\rangle = -\langle\, u', \eta \,\rangle = 0,$$

since $u' = 0$. This concludes the proof. □

4.2 A fundamental solution for $\Box$

Recall that a *fundamental solution* for a constant-coefficient linear differential operator

$$L = \sum_{|\alpha|\le N} a_\alpha \partial^\alpha$$

is a distribution w such that $Lw = \delta$, where δ is the point-mass at the origin. Observe that if $f \in C_c^\infty$, then

$$L(w * f) = (Lw) * f = \delta * f = f.$$

Conversely, if w is a distribution such that $L(w * f) = f$ for every $f \in C_c^\infty$, then $(Lw) * f = f$ for all $f \in C_c^\infty$, and this implies $Lw = \delta$.

Now consider $\Box$ on $\mathbb{R}^{1+n}$. Define $W_+ \in C(\mathbb{R}, \mathcal{D}'(\mathbb{R}^n))$ by

$$W_+(t) = \begin{cases} W(t) & t \ge 0, \\ 0 & t < 0, \end{cases}$$

where $W(t)$ is the "wave propagator" defined last week. Thus $W_+ \in \mathcal{D}'(\mathbb{R}^{1+n})$ with action

$$\langle\, W_+, \psi \,\rangle = \int_0^\infty \langle\, W(t), \psi(t, \cdot) \,\rangle \; dt$$

on a test function $\psi \in C_c^\infty(\mathbb{R}^{1+n})$.

Theorem 15. *W_+ is a fundamental solution for $\Box$.*

Equivalently,

$$\Box(W_+ * F) = F \quad \text{for all} \quad F \in C_c^\infty(\mathbb{R}^{1+n}). \tag{4.2}$$

(Convolution product in $\mathbb{R}^{1+n}$.)

Exercise 17. Prove (4.2) and hence the theorem. (*Hint:* Fix $F \in C_c^\infty(\mathbb{R}^{1+n})$ and choose t_0 so that $F(t, x)$ vanishes for $t \le t_0$. Now use Duhamel's principle with initial time $t = t_0$ instead of $t = 0$. That is, show that

$$u(t, x) = \int_{t_0}^t W(t - s) * F(s, x)\, ds$$

solves $\Box u = F$. Then show that $u = W_+ * F$.)

4.3 Loss of classical derivatives

In space dimensions $n > 1$, the solution of

$$\Box u = 0, \quad u\big|_{t=0} = f, \quad \partial_t u\big|_{t=0} = g$$

loses up to $n/2$ degrees of differentiability from time $t = 0$ to any time $t > 0$. More precisely, to ensure

$$u(t) \in C^2(\mathbb{R}^n) \quad \text{for} \quad t > 0,$$

we must assume, in general,

$$\begin{aligned} &f \in C^{2+n/2}, && g \in C^{1+n/2} && (n \text{ even}), \\ &f \in C^{2+(n-1)/2}, && g \in C^{1+(n-1)/2} && (n \text{ odd}). \end{aligned}$$

Intuitively, this is because "weak" singularities at $t = 0$ propagate on forward light cones, thus interacting at times $t > 0$ to create "stronger" singularities. We will not try to make this precise.

In contrast, there is *no loss of L^2-differentiability* from $t = 0$ to $t > 0$. That is, if the initial data have weak derivatives in L^2 up to some order k, then so does the solution $u(t, \cdot)$ for all times $t > 0$. The *energy identity*, which we consider next, illustrates this principle.

4.4 Energy identity

Recall that ∂u denotes the spacetime gradient:

$$\partial u = (u_t, \nabla_x u).$$

Theorem 16. *Suppose $u \in C^2([0,T] \times \mathbb{R}^n)$ solves $\Box u = 0$ and that $u(t, \cdot)$ is compactly supported for every t. Then*

$$\|\partial u(t, \cdot)\|_{L^2} = \|\partial u(0, \cdot)\|_{L^2}$$

for all $0 \le t \le T$.

Proof. Consider the energy

$$e(t) = \frac{1}{2} \int_{\mathbb{R}^n} |\partial u(t,x)|^2 \, dx = \frac{1}{2} \|\partial u(t, \cdot)\|_{L^2}^2 .$$

Differentiate $e(t)$, and integrate by parts, to get

$$e'(t) = \int_{\mathbb{R}^n} u_t \Box u \, dx. \tag{4.3}$$

Since $\Box u = 0$, we get $e'(t) = 0$, hence $e(t)$ is constant. $\Box$

With a little more work we can prove the following *energy inequality* for a solution of the inhomogeneous wave equation. (The assumptions of compact support and smoothness of u in these theorems can be removed by approximation arguments similar to ones we will encounter later on.)

Theorem 17. *Suppose $u \in C^2([0,T] \times \mathbb{R}^n)$ and $u(t,\cdot)$ is compactly supported for every t. Then*

$$\|\partial u(t,\cdot)\|_{L^2} \leq \|\partial u(0,\cdot)\|_{L^2} + \int_0^t \|\Box u(s,\cdot)\|_{L^2} \, ds$$

for all $0 \leq t \leq T$.

Proof. Define the energy as in the proof of the previous theorem. Applying the Cauchy-Schwarz inequality to (4.3), we have

$$e'(t) \leq \|\partial u(t,\cdot)\|_{L^2} \|\Box u(t,\cdot)\|_{L^2} = \sqrt{2e(t)} \, \|\Box u(t,\cdot)\|_{L^2} .$$

Thus, whenever $e'(t) \neq 0$,

$$\frac{d}{dt}\sqrt{e(t)} = \frac{e'(t)}{2\sqrt{e(t)}} \leq \frac{1}{\sqrt{2}} \|\Box u(t,\cdot)\|_{L^2} .$$

It follows that

$$\sqrt{e(t)} \leq \sqrt{e(0)} + \frac{1}{\sqrt{2}} \int_0^t \|\Box u(s,\cdot)\|_{L^2} \, ds,$$

which is exactly what we want. □

4.5 The L^2 theory

4.5.1 The Fourier transform

The Fourier transform of $f \in L^1(\mathbb{R}^n)$ is

$$\mathcal{F}f(\xi) = \widehat{f}(\xi) = \frac{1}{(2\pi)^{n/2}} \int_{\mathbb{R}^n} e^{-ix\cdot\xi} f(x) \, dx. \tag{4.4}$$

We let $\mathcal{S}(\mathbb{R}^n)$ be the class of Schwartz functions:

$$\mathcal{S}(\mathbb{R}^n) = \left\{ f \in C^\infty(\mathbb{R}^n) : \|f\|_{(N,\alpha)} < \infty \text{ for all } N,\, \alpha \right\},$$

where

$$\|f\|_{(N,\alpha)} = \sup_{x\in\mathbb{R}^n} (1+|x|)^N \, |\partial^\alpha f(x)|$$

is a seminorm for every $N \in \mathcal{N}$ and every multi-index α. When equipped with the topology induced by this countable family of seminorms, $\mathcal{S}$ is a Fréchet space, and in particular a complete metric space. The topology is characterized by the notion of sequential convergence:

$$f_j \to f \quad \text{in} \quad \mathcal{S} \quad \Longleftrightarrow \quad \|f_j - f\|_{(N,\alpha)} \to 0 \quad \text{for all} \quad N, \alpha.$$

Recall the following basic facts concerning the Fourier transform on $\mathcal{S}$:

(i) $\int \widehat{f} g = \int f \widehat{g}$.

(ii) $\mathcal{F}(f * g) = (2\pi)^{n/2} \widehat{f}\,\widehat{g}$.

(iii) $\mathcal{F}(\partial^\alpha f)(\xi) = (i\xi)^\alpha \widehat{f}(\xi)$.

(iv) $\partial^\alpha \widehat{f}(\xi) = \mathcal{F}\left[(-ix)^\alpha f\right]$.

(v) $\mathcal{F}$ maps $\mathcal{S}$ isomorphically onto itself, with inverse $\mathcal{F}^{-1} : \mathcal{S} \to \mathcal{S}$ given by

$$\mathcal{F}^{-1} g(x) = \frac{1}{(2\pi)^{n/2}} \int_{\mathbb{R}^n} e^{ix\cdot\xi} g(\xi)\, d\xi.$$

(vi) **(Plancherel's Theorem.)** $\mathcal{F} : \mathcal{S} \to \mathcal{S}$ extends to a unitary isomorphism $\mathcal{F} : L^2 \to L^2$. Thus,

$$\|f\|_{L^2} = \left\|\widehat{f}\right\|_{L^2}.$$

Also, if $f \in L^1 \cap L^2$, the $\widehat{f}$ is given by the integral (4.4).

The dual of $\mathcal{S}$, denoted $\mathcal{S}'$, consists of all continuous linear maps $u : \mathcal{S} \to \mathbb{C}$. Clearly,

$$\phi_j \to \phi \quad \text{in} \quad C_c^\infty \qquad \Longrightarrow \qquad \phi_j \to \phi \quad \text{in} \quad \mathcal{S},$$

and since C_c^∞ is dense in $\mathcal{S}$, we may consider $\mathcal{S}'$ to be a subset of $\mathcal{D}'$. It is then easy to see that $\mathcal{S}'$ in turn contains the set of compactly supported distributions. We summarize:

$$\vec{E}' \subset \mathcal{S}' \subset \mathcal{D}'.$$

The elements of $\mathcal{S}'$ are called *tempered distributions*. The main advantage of tempered distributions is that they have a Fourier transform: In view of (i) and (v) above, the Fourier transform extends to an isomorphism of $\mathcal{S}'$ onto itself, given by

$$\langle\, \widehat{u}, \phi \,\rangle = \left\langle\, u, \widehat{\phi} \,\right\rangle.$$

Recall also that if $u \in \vec{E}'$, then $\widehat{u}$ is a C^∞ function, given by

$$\widehat{u}(\xi) = (2\pi)^{-n/2} \langle\, u, E_\xi \,\rangle\,,$$

where $E_\xi \in C^\infty(\mathbb{R}^n)$ is the function $E_\xi(x) = e^{-ix\cdot\xi}$.

We say that a function $\psi \in C^\infty(\mathbb{R}^n)$ is *slowly increasing* if ψ and all its partial derivatives have at most polynomial growth at infinity:

$$|\partial^\alpha \psi(x)| \leq C_\alpha (1 + |x|)^{N(\alpha)} \quad \text{for all} \quad \alpha.$$

Then $\psi\phi \in \mathcal{S}$ whenever $\phi \in \mathcal{S}$, and this gives a continuous map from $\mathcal{S}$ into itself. Therefore, if $u \in \mathcal{S}'$ and ψ is slowly increasing, we can define their product $\psi u \in \mathcal{S}'$ by

$$\langle\, \psi u, \phi \,\rangle = \langle\, u, \psi\phi \,\rangle \quad \text{for} \quad \phi \in \mathcal{S}.$$

4.5.2 Solution of wave equation by Fourier analysis

Let u be the solution of

$$\Box u = 0 \quad \text{on} \quad \mathbb{R}^{1+n}, \quad u\big|_{t=0} = f, \quad \partial_t u\big|_{t=0} = g,$$

where $f, g \in \mathcal{S}(\mathbb{R}^n)$. Now apply the Fourier transform in the space variable:

$$\widehat{u}(t,\xi) = \frac{1}{(2\pi)^{n/2}} \int_{\mathbb{R}^n} e^{-ix\cdot\xi} u(t,x)\, dx.$$

Then $\Box u = 0$ transforms to, using property (iii) above,

$$\partial_t \widehat{u}(t,\xi) + |\xi|^2 \, \widehat{u}(t,\xi) = 0, \quad \widehat{u}(0,\xi) = \widehat{f}(\xi), \quad \partial_t \widehat{u}(0,\xi) = \widehat{g}(\xi).$$

But for fixed ξ, this is an initial value problem for a second order ODE in time, whose solution is

$$\widehat{u}(t,\xi) = \cos(t\,|\xi|)\widehat{f}(\xi) + \frac{\sin(t\,|\xi|)}{|\xi|}\widehat{g}(\xi). \tag{4.5}$$

On the other hand, we know that

$$u(t,\cdot) = W'(t) * f + W(t) * g.$$

Since $W(t) \in \vec{E}'$, we can apply the Fourier transform to the last equation, which gives

$$\widehat{u}(t,\xi) = (2\pi)^{n/2}\widehat{W'(t)}(\xi)\widehat{f}(\xi) + (2\pi)^{n/2}\widehat{W(t)}(\xi)\widehat{g}(\xi).$$

Comparing this equation with (4.5), we conclude:

$$\widehat{W(t)}(\xi) = (2\pi)^{-n/2}\frac{\sin(t\,|\xi|)}{|\xi|}, \quad \widehat{W'(t)}(\xi) = (2\pi)^{-n/2}\cos(t\,|\xi|). \tag{4.6}$$

Now recall Duhamel's formula for the solution of $\Box u = F$ with vanishing initial data at $t = 0$:

$$u(t,x) = \int_0^t W(t-s) * F(s,x)\, ds.$$

Applying the Fourier transform and using (4.6) we then have

$$\widehat{u}(t,\xi) = \int_0^t \frac{\sin(t-s)\,|\xi|}{|\xi|}\widehat{F}(s,\xi)\, ds.$$

Thus, the solution of the full inhomogeneous Cauchy problem

$$\Box u = F \quad \text{on} \quad \mathbb{R}^{1+n}, \quad u\big|_{t=0} = f, \quad \partial_t u\big|_{t=0} = g,$$

is given in Fourier space by

$$\widehat{u}(t,\xi) = \cos(t\,|\xi|)\widehat{f}(\xi) + \frac{\sin(t\,|\xi|)}{|\xi|}\widehat{g}(\xi) + \int_0^t \frac{\sin(t-s)\,|\xi|}{|\xi|}\widehat{F}(s,\xi)\, ds.$$

4.5.3 The Sobolev spaces H^s

Fix $s \in \mathbb{R}$. Observe that the function $\xi \to (1+|\xi|^2)^{s/2}$ is C^∞ and slowly increasing. We can therefore define $\Lambda^s : \mathcal{S}' \to \mathcal{S}'$ by

$$\Lambda^s f = \mathcal{F}^{-1}(1+|\xi|^2)^{s/2}\mathcal{F}f.$$

Thus Λ^s is a composition of three continouous maps, so it is itself continouous. Moreover, it is an isomorphism, since its inverse is just Λ^{-s}.

By the same reasoning, Λ^s restricted to $\mathcal{S}$ is an isomorphism $\mathcal{S}$ onto itself.

Now set $H^s(\mathbb{R}^n) = \Lambda^{-s}\big(L^2(\mathbb{R}^n)\big)$ with norm

$$\|f\|_{H^s} = \|\Lambda^s f\|_{L^2}\,.$$

In other words, $H^s = \{f \in \mathcal{S}' : \Lambda^s f \in L^2\}$. By Plancherel's theorem,

$$\|f\|_{H^s} = \left\|(1+|\xi|^2)^{s/2}\widehat{f}\right\|_{L^2}.$$

Of course, H^0 is just L^2. Observe that $s < t \implies H^t \subset H^s$.

Thus, $\Lambda^s : H^s \to L^2$ is an isometric isomorphism (in particular, H^s is a Hilbert space) and since $\mathcal{S}$ is dense in L^2 and $\Lambda^s(\mathcal{S}) = \mathcal{S}$, it follows that $\mathcal{S}$ is dense in H^s. Note also that $\Lambda^s : H^t \to H^{t-s}$ is an isometric isomorphism for all $s, t \in \mathbb{R}$.

If $s \in \mathcal{N}$, then

$$H^s = \left\{f \in L^2 : \partial^\alpha f \in L^2 \text{ for } |\alpha| \le s\right\},$$

and the norm $\|f\|_{H^s}$ is equivalent to

$$\left(\sum_{|\alpha|\le s} \|\partial^\alpha f\|_{L^2}^2\right)^{1/2}$$

These assertions follow easily from Plancherel's theorem.

We will study these spaces in more detail later on.

4.5.4 L^2 estimates for solutions of $\Box u = 0$

Recall that by the energy identity, if $\Box u = 0$, then the L^2 norm of the spacetime gradient, $\|\partial u(t,\cdot)\|_{L^2}$, is a conserved quantity. The estimates for the L^2 norm of u itself are less favorable. In fact,

$$\|u(t,\cdot)\|_{L^2} = O(t) \quad \text{as} \quad t \to \infty,$$

and this is essentially sharp, since in general $\|u(t,\cdot)\|_{L^2}$ fails to be $O(t^\theta)$ for any $\theta < 1$ (see Exercise 3 below).

The key estimates are contained in the following

Lemma 1. *For any $s \in \mathbb{R}$,* *(i)* $\|W'(t) * f\|_{H^s} \le \|f\|_{H^s}$,

(ii) $\|W(t) * g\|_{H^s} \le \sqrt{2}(1 + |t|)\, \|g\|_{H^{s-1}}$.

Proof. Without loss of generality we may assume $s = 0$. Using (4.6), we have

$$\|W'(t) * f\|_{L^2} = \left\|\cos(t\,|\xi|)\widehat{f}\right\|_{L^2} \le \left\|\widehat{f}\right\|_{L^2} = \|f\|_{L^2}.$$

To prove (b), write

$$\|W(t) * g\|_{L^2}^2 = \int \left|\frac{\sin(t\,|\xi|)}{|\xi|}\widehat{g}(\xi)\right|^2 d\xi = \int_{|\xi|<1} + \int_{|\xi|\ge 1}.$$

Since

$$\frac{\sin(t\,|\xi|)}{|\xi|} \le \begin{cases} t & \text{(use when } |\xi| < 1) \\ \frac{1}{|\xi|} & \text{(use when } |\xi| \ge 1) \end{cases}$$

and

$$|\xi| < 1 \implies 1 < \frac{2}{1 + |\xi|^2}, \qquad |\xi| \ge 1 \implies \frac{1}{|\xi|^2} \le \frac{2}{1 + |\xi|^2},$$

we obtain

$$\|W(t) * g\|_{L^2}^2 \le 2(1 + t^2) \int \frac{|\widehat{g}(\xi)|^2}{1 + |\xi|^2}\, d\xi = 2(1 + t^2)\, \|g\|_{H^{-1}}^2.$$

□

Exercise 18. Construct a function g such that $u(t, \cdot) = W(t) * g$ satisfies

$$\|u(t, \cdot)\|_{L^2} \ge \frac{Ct}{\log t} \quad \text{as} \quad t \to \infty,$$

for some $C > 0$. Conclude that $\|u(t, \cdot)\|_{L^2}$ fails to be $O(t^\theta)$ for any $\theta < 1$.

Extended hint: Let $\widehat{g}(\xi) = h(|\xi|)$, where

$$h(r) = \begin{cases} r^{-n/2}(-\log r)^{-1} & 0 < r < \frac{1}{2}, \\ 0 & r \ge \frac{1}{2}. \end{cases}$$

Now show

$$\|u(t, \cdot)\|_{L^2}^2 = c_n t^2 \int_0^{\frac{1}{2}} \frac{\sin^2(tr)}{(tr)^2} \cdot \frac{dr}{r \log^2 r} \ge c_n' t^2 \int_0^{\frac{1}{2t}} \frac{dr}{r \log^2 r}$$

and change variables to $y = -\log r$.

Chapter 5

Week 5: The L^2 theory

5.1 Existence and uniqueness in H^s

Consider the Cauchy problem on $\mathbb{R}^{1+n}$ for the linear wave equation:

$$\square u = F, \tag{5.1a}$$

$$u\big|_{t=0} = f, \quad \partial_t u\big|_{t=0} = g. \tag{5.1b}$$

We shall prove:

Theorem 18. *Let $s \in \mathbb{R}$. Let $f \in H^s$, $g \in H^{s-1}$ and $F \in L^1([0,T], H^{s-1})$. Then for every $T > 0$, there is a unique u which belongs to*

$$C([0,T], H^s) \cap C^1([0,T], H^{s-1}) \tag{5.2}$$

and solves (5.1) on $S_T = (0,T) \times \mathbb{R}^n$. Moreover, u satisfies

$$\|u(t)\|_{H^s} + \|\partial_t u(t)\|_{H^{s-1}} \le C_T \left(\|f\|_{H^s} + \|g\|_{H^{s-1}} + \int_0^t \|F(t')\|_{H^{s-1}}\, dt' \right) \tag{5.3}$$

for all $0 \le t \le T$, where $C_T = C(1+T)$ and C only depends on s.

Remarks. (i) $F \in L^1([0,T], H^{s-1})$ means that the function

$$[0,T] \to \mathbb{R}, \quad t \to \|F(t)\|_{H^{s-1}}$$

is in $L^1([0,T])$. Any such F defines a distribution on S_T by

$$\langle F, \psi \rangle = \int_0^T \langle F(t), \psi(t,\cdot) \rangle \, dt \quad \text{for} \quad \psi \in C_c^\infty(S_T). \tag{5.4}$$

The key facts needed to prove that this integral converges, and defines an element of $\mathcal{D}'(S_T)$, are as follows. First, we have the inequality

$$|\langle u, \phi \rangle| \le \|u\|_{H^s} \|\phi\|_{H^{-s}} \quad \text{for all} \quad s \in \mathbb{R}, u \in H^s, \phi \in \mathcal{S}. \tag{5.5}$$

Secondly, the inclusions

$$C_c^\infty \subset \mathcal{S} \subset H^s \subset \mathcal{S}' \subset \mathcal{D}' \tag{5.6}$$

are all sequentially continuous. We leave the details as good exercises.

(ii) By (5.6) one has, for every interval $I \subset \mathbb{R}$, continuous inclusions

$$C(I, H^s) \subset C(I, \mathcal{S}') \subset C(I, \mathcal{D}'),$$

and the same holds for every C^k. Thus, if u is in (5.2), then $u \in C^1([0,T], \mathcal{D}')$.

(iii) The equation (5.1a) is of course understood in the sense of $\mathcal{D}'(S_T)$, which makes sense in view of the previous remarks.

(iv) It is interesting to note that if the theorem holds for some $s_0 \in \mathbb{R}$, then it follows immediately that it holds for every $s \in \mathbb{R}$. This is easy to see, since we have isometric isomorphisms

$$\Lambda^{s_0 - s} : H^{s_0} \to H^s, \quad \Lambda^{s - s_0} : H^s \to H^{s_0}$$

which commute with the wave operator $\square$. That is,

$$\square \Lambda^s u = \Lambda^s \square u$$

for any s. To see this, apply the Fourier transform in the space variable.

(v) We know that the solution of (5.1) is given by

$$u(t) = W'(t) * f + W(t) * g + \int_0^t W(t - t') * F(t')\, dt', \tag{5.7}$$

at least if F is smooth. In fact, this formula makes sense for any F in the space $L^1([0,T], H^{s-1})$ if the last term is interpreted as a Hilbert space-valued integral, the Hilbert space being H^{s-1}, which is separable. Equivalently, the integral can be understood in the weak sense, that is to say, as a distribution whose action on a test function $\phi \in \mathcal{S}$ is

$$\int_0^t \langle W(t - t') * F(t'), \phi \rangle \; dt'.$$

We leave it as an exercise to show that this defines a time-dependent distribution $t \to w(t)$ which belongs to (5.2) and solves $\square w = F$ on S_T with vanishing initial data. (Use the inequality (5.5), Lemma 1 from last week, and the result in Exercise 1(c).)

Let us now prove Theorem 18.

Uniqueness. In view of remark (ii) above, uniqueness follows from Theorem 1 in last week's lecture notes. Alternatively, we can use remark (iv) to conclude that it suffices to prove uniqueness in the space (5.2) for *some* $s \in \mathbb{R}$. If we take s very large, then by Sobolev embedding, any solution of (5.1) with $f = g = 0$ and $F = 0$, and belonging to the space (5.2), must be C^2, so we can appeal to the uniqueness theorem for smooth solutions proved in the first week.

Existence. We use a standard approximation argument. (Alternatively, one could work out the details in remark (v) above.) We first check that (5.3) holds when f, g and F are smooth, specifically $f, g \in \mathcal{S}$ and $F \in C^\infty([0,T], \mathcal{S})$. Then the solution is given by (5.7), and in view of Lemma 1 from last week,

(5.8)

$$\begin{aligned} \|u(t)\|_{H^s} &\le \|f\|_{H^s} + C(1+t)\,\|g\|_{H^{s-1}} + \int_0^t \|W(t-t') * F(t',\cdot)\|_{H^{s-1}}\,dt' \\ &\le \|f\|_{H^s} + C(1+t)\,\|g\|_{H^{s-1}} + C(1+t)\int_0^t \|F(t',\cdot)\|_{H^{s-1}}\,dt'. \end{aligned}$$

Moreover, by Theorem 4 from last week (cf. also remark (iv) above),

$$\|\partial_t u(t)\|_{H^{s-1}} \le C\left(\|f\|_{H^s} + \|g\|_{H^{s-1}} + \int_0^t \|F(t')\|_{H^{s-1}}\,dt'\right). \tag{5.9}$$

Combining (5.8) and (5.9) gives (5.3).

Now assume $f \in H^s$, $g \in H^{s-1}$ and $F \in L^1([0,T], H^{s-1})$, and choose sequences $f_j, g_j \in \mathcal{S}$ and $F_j \in C^\infty([0,T], \mathcal{S})$ (cf. Exercise 4) such that

$$\|f_j - f\|_{H^s} \to 0, \quad \|g_j - g\|_{H^{s-1}} \to 0, \quad \int_0^T \|F_j(t') - F(t')\|_{H^{s-1}}\,dt' \to 0.$$

Let u_j be the corresponding solutions, given by (5.7) with f_j, g_j, F_j replacing f, g, F. Denote by X_T the space (5.2). Then X_T is a Banach space with norm

$$\|u\|_{X_T} = \sup_{0\le t\le T}\big(\|u(t)\|_{H^s} + \|\partial_t u(t)\|_{H^{s-1}}\big).$$

Since (5.3) holds for smooth solutions, we have

$$\|u_j - u_k\|_{X_T} \lesssim \|f_j - f\|_{H^s} + \|g_j - g\|_{H^{s-1}} + \int_0^T \|F_j(t') - F(t')\|_{H^{s-1}}\,dt',$$

so u_j is Cauchy in X_T and hence converges in that space to a limit u. But then u_j also converges to u in the sense of $\mathcal{D}'(S_T)$, whence $\Box u_j \to \Box u$ in $\mathcal{D}'(S_T)$. On the other hand, $\Box u_j = F_j \to F$ in $\mathcal{D}'(S_T)$ [cf. Exercise 3(b)], so u solves (5.1). This completes the proof.

5.1.1 Exercises

Exercise 19. (i) Show that, for any $u \in H^s$ and $\phi \in \mathcal{S}$,

$$\langle\, u, \phi \,\rangle = \int \widehat{u}(\xi)\widehat{\phi}(-\xi)\,d\xi = \int (1+|\xi|^2)^{s/2}\widehat{u}(\xi)(1+|\xi|^2)^{-s/2}\widehat{\phi}(-\xi)\,d\xi.$$

(ii) Prove (5.5).

(iii) Prove that $\|u\|_{H^s} = \sup\big\{|\langle u, \phi \rangle| : \phi \in \mathcal{S},\ \|\phi\|_{H^{-s}} = 1\big\}$. (*Hint:* Recall that L^2 is self-dual.)

Exercise 20. Prove that the inclusions in (5.6) are sequentially continuous.

Exercise 21. Let $s \in \mathbb{R}$, $T > 0$. Set $S_T = (0, T) \times \mathbb{R}^n$.

(i) Prove that (5.4) defines an element of $\mathcal{D}'(S_T)$ for every $F \in L^1([0, T], H^s)$.

(ii) Prove that the inclusion

$$L^1([0, T], H^s) \subset \mathcal{D}'(S_T)$$

is sequentially continuous.

Exercise 22. Prove that

$$C^\infty([0, T], \mathcal{S}) \quad \text{is dense in} \quad L^1([0, T], H^s) \tag{5.10}$$

by completing the following steps.

(i) Show that if $C_c^\infty(S_T)$ is dense in $L^1([0, T], L^2)$, then (5.10) follows. (*Hint:* Use the isomorphism $\Lambda^s : H^s \to L^2$ and its inverse Λ^{-s}.)

(ii) $C_c(S_T)$ is dense in $L^1([0, T], L^2)$. (*Hint:* First approximate by simple functions.)

(iii) If $F \in C_c(S_T)$, $\phi \in C_c^\infty(\mathbb{R}^{1+n})$ and $\int \phi = 1$, then

$$\int_0^T \|F_\varepsilon(t, \cdot) - F(t, \cdot)\|_{L^2}\, dt \to 0 \quad \text{as} \quad \varepsilon \to 0,$$

where $F_\varepsilon(t, x) \in C_c^\infty$ is given by

$$F_\varepsilon(t, x) = \int_{\mathbb{R}^{1+n}} F(t - \varepsilon t', x - \varepsilon x')\phi(t', x')\, dt'\, dx' = F * \phi_\varepsilon(t, x),$$

and $\phi_\varepsilon(t, x) = \varepsilon^{-1-n}\phi(t/\varepsilon, x/\varepsilon)$.

5.2 Nonlinear equations

5.2.1 An example of blow-up in finite time

We look at the nonlinear Cauchy problem

$$\Box u = (\partial_t u)^2, \tag{5.11a}$$

$$u\big|_{t=0} = 0, \quad \partial_t u\big|_{t=0} = g. \tag{5.11b}$$

Theorem 19. *For any $T > 0$, there exists $g \in C_c^\infty(\mathbb{R}^n)$ such that (5.11) does not admit a C^2 solution past time T.*

Step 1. Take g in (5.11b) to be a constant $c > 0$. Then (5.11a) reduces to an ODE in time, $u_{tt} = u_t^2$, or in terms of $y = u_t$,

$$y' = y^2, \quad y(0) = c.$$

The solution of this initial value problem is

$$y(t) = \frac{c}{1 - ct},$$

which blows up as $t \to 1/c$. Thus $u(t,x) = -\log(1 - ct)$ solves (5.11) and $u \to \infty$ as $t \to 1/c$.

Step 2. We prove a uniqueness theorem for the equation (5.11a).

Theorem. *If $u \in C^2(\Omega)$ solves $\Box u = (\partial_t u)^2$ in the solid backward light cone*

$$\Omega = \{(t,x) : 0 \le t < T,\ |x - x_0| < T - t\}, \tag{5.12}$$

with base $B_0 = \{x : |x - x_0| < T - t\}$, then u is uniquely determined by its data

$$u|_{B_0}, \quad \partial_t u|_{B_0}.$$

In other words, if $u, v \in C^2(\Omega)$ are two solutions of (5.11a) in Ω, with the *same data* in B_0, then $u = v$ in Ω. To see this, observe that $u - v$ solves the *linear* equation

$$\Box(u - v) = a(t,x)\partial_t(u - v) \quad \text{in} \quad \Omega,$$

where $a = \partial_t u + \partial_t v \in C^1(\Omega)$, and that $u - v$ and $\partial_t(u - v)$ vanish in B_0. Then by uniqueness of solutions of the linear equation $\Box w = a(t,x)\partial_t w$ (see Theorem 21 below) we conclude that $u = v$ in Ω.

Step 3. By applying Step 2 to the solution obtained in Step 1, we conclude that if $u \in C^2(\Omega)$ solves (5.11a) in Ω, with initial data $u = 0$ and $\partial_t u = 1/T$ in B_0, where Ω is given by (5.12) and B_0 is its base, then

$$u(t,x) = -\log(1 - t/T) \quad \text{for} \quad (t,x) \in \Omega.$$

To finish the proof of Theorem 19, we only have to cut off the constant $1/T$ smoothly outside a sufficiently large ball to produce $g \in C_c^\infty$ with the claimed property.

5.3 A uniqueness theorem for nonlinear equations

We generalize the argument used in Step 2 above to prove uniqueness of smooth solutions of an equation

$$\Box u = F(u, \partial u), \tag{5.13}$$

where $F : \mathbb{R}^{n+2} \to \mathbb{R}$ is a given C^∞ function.

Theorem 20. *If $u \in C^2(\Omega)$ solves (5.13) in the solid backward light cone*

$$\Omega = \{(t,x) : 0 \le t < T,\, |x - x_0| < T - t\},$$

with base $B_0 = \{x : |x - x_0| < T - t\}$, then u is uniquely determined by its data

$$u|_{B_0}, \quad \partial_t u|_{B_0}.$$

Proof. Suppose $u, v \in C^2(\Omega)$ both solve the equation, and with identical data in B_0. Then

$$\Box(u - v) = F(u, \partial u) - F(v, \partial v).$$

But if $w, z \in R^{n+2}$, then

$$\begin{aligned} F(w) - F(z) &= \int_0^t \frac{d}{dt} F\big([1-t]z + tw\big)\, dt \\ &= \int_0^t \nabla F\big([1-t]z + tw\big) \cdot (w - z)\, dt \\ &= G(w, z) \cdot (w - z), \end{aligned}$$

where $G : R^{2(n+2)} \to \mathbb{R}$ is C^∞. Thus

$$\begin{aligned} F(u, \partial u) - F(v, \partial v) &= G(u, \partial u, v, \partial v) \cdot (u - v, \partial u - \partial v) \\ &= a(t,x)(u - v) + b(t,x) \cdot \partial(u - v), \end{aligned}$$

where $a, b \in C^1(\Omega)$, so $w = u - v$ solves the linear equation

$$\Box w = a(t,x)w + b(t,x) \cdot \partial w \quad \text{in} \quad \Omega$$

with vanishing initial data in B_0. The next theorem shows that $u = v$ in Ω. $\Box$

Using the energy method, we now prove a uniqueness theorem for linear equations of the form

$$\Box u = a(t,x)u + b(t,x) \cdot \partial u \tag{5.14}$$

with continuous coefficents a and b (the latter is of course $\mathbb{R}^{1+n}$-valued).

Theorem 21. *Let $a, b \in C(\Omega)$, where Ω as before is the solid cone*

$$\Omega = \{(t,x) : 0 \le t < T,\, |x - x_0| < T - t\},$$

with base $B_0 = \{x : |x - x_0| < T - t\}$. If $u \in C^2(\Omega)$ solves (5.14) in Ω and

$$u|_{B_0} = \partial_t u|_{B_0} = 0,$$

then $u = 0$ in Ω.

Proof. Fix $\varepsilon > 0$. We will show that $u = 0$ in

$$K = \{(t, x) : 0 \le t \le T - \varepsilon,\ |x - x_0| \le T - \varepsilon - t\}.$$

We use a slight modification of the energy method employed to prove uniqueness for $\Box u = 0$ in the first week (see Theorem 1 in the lecture notes for that week). Let B_t be the time-slices of K:

$$B_t = \{x : |x - x_0| \le T - \varepsilon - t\}.$$

Set

$$E(t) = \frac{1}{2} \int_{B_t} \left(|u(t, x)|^2 + |\partial u(t, x)|^2 \right) dx.$$

Then

$$E'(t) = \int_{B_t} (u u_t + u_t u_{tt} + \nabla u \cdot \nabla u_t)\, dx \\ - \frac{1}{2} \int_{\partial B_t} \left(|u(t, x)|^2 + |\partial u(t, x)|^2 \right) d\sigma(x).$$

Since

$$u_{tt} = \Delta u + au + b \cdot \partial u, \quad \operatorname{div}(u_t \nabla u) = \nabla u_t \cdot \nabla u + u_t \Delta u,$$

we get $E'(t) = I(t) + II(t)$, where

$$I(t) = \int_{B_t} u_t \left[(1 + a)u + b \cdot \partial u\right] dx,$$

$$II(t) = \int_{B_t} \operatorname{div}(u_t \nabla u)\, dx - \frac{1}{2} \int_{\partial B_t} \left(|u(t, x)|^2 + |\partial u(t, x)|^2 \right) d\sigma(x).$$

By the divergence theorem (cf. proof of Theorem 1, week 1),

$$II(t) \le -\frac{1}{2} \int_{\partial B_t} |u(t, x)|^2\, d\sigma(x) \le 0,$$

so $E'(t) \le I(t)$. Since a and b are uniformly bounded on K, and since

$$|u u_t| \le \frac{1}{2} \left(|u|^2 + |u_t|^2 \right),$$

we finally get

$$E'(t) \le I(t) \le C E(t).$$

This implies $\frac{d}{dt}\left[E(t) e^{-Ct}\right] \le 0$, so $E(t) \le E(0) e^{Ct}$ for $0 \le t \le T - \varepsilon$. Since $E(0) = 0$, we conclude that $u = 0$ in K. □

5.3.1 Exercises

Exercise 23. Show that the proof of Theorem 21 works for a *system* of equations, of the form

$$\Box u^I = \sum_J a_{IJ}(t,x)u^J + \sum_J b_{IJ}(t,x) \cdot \partial u^J, \quad 1 \le I \le N,$$

where $u = (u^1, \dots, u^N)$ takes values in R^N and the a_{IJ} and b_{IJ} are continuous. In this case, set

$$E(t) = \frac{1}{2} \int_{B_t} \left(|u(t,x)|^2 + |\partial u(t,x)|^2 \right) dx,$$

where

$$|u|^2 = \sum_I (u^I)^2, \quad |\partial u|^2 = \sum_I \left| \partial u^I \right|^2 .$$

Then check that the argument used to prove the nonlinear uniqueness result (Theorem 20) generalizes to a nonlinear system

$$\Box u^I = F^I(u, \partial u), \quad 1 \le I \le N,$$

where the F^I are given smooth functions (real-valued).

5.4 Local existence and uniqueness

We consider a nonlinear Cauchy problem

$$\Box u = F(u, \partial u), \tag{5.15a}$$

$$u\big|_{t=0} = f, \quad \partial_t u\big|_{t=0} = g. \tag{5.15b}$$

Here F is a given smooth function satisfying $F(0) = 0$. For simplicity we think of u and F as being real-valued, but the theorems that follow are true for systems as well (that is, for $\mathbb{R}^N$-valued u and F), with identical proofs.

Theorem 22. ***(Local Existence and Uniqueness.)*** *Let $s > \frac{n}{2} + 1$. Then for all $(f, g) \in H^s \times H^{s-1}$, there exist $T > 0$ and a unique*

$$u \in C([0,T], H^s) \cap C^1([0,T], H^{s-1}) \tag{5.16}$$

solving (5.15) *on $S_T = (0,T) \times \mathbb{R}^n$. Moreover, the time T can be chosen to depend continuously on $\|f\|_{H^s} + \|g\|_{H^{s-1}}$.*

We emphasize that uniqueness holds for solutions in the space (5.16) for any $T > 0$, as is clear from the proof below.

The next theorem says that a solution u on S_T belonging to (5.16) can be extended to a time $T' > T$ provided $\partial u \in L^\infty(S_T)$.

Theorem 23. ***(Continuation.)*** *Let $s > \frac{n}{2} + 1$. Fix f, g as in Theorem 22. Let $T^* = T^*(f, g)$ be the supremum of all $T > 0$ such that* (5.15) *has a solution on S_T satisfying* (5.16). *Then either $T^* = \infty$ or*

$$\partial u \notin L^\infty(S_{T^*}).$$

From the two previous theorems we will deduce the following.

Theorem 24. ***(Smooth Solutions.)*** *If $f, g \in C_c^\infty(\mathbb{R}^n)$, then there exist $T > 0$ and a unique*

$$u \in C^\infty([0, T] \times \mathbb{R}^n)$$

solving (5.15) *on S_T.*

Using the uniqueness of smooth solutions in backward solid light cones, we obtain the following corollary.

Corollary. *If $f, g \in C^\infty(\mathbb{R}^n)$, there exists a set*

$$A = \{(t, x) : 0 \le t \le T(x)\},$$

where T is a continuous and strictly positive function on $\mathbb{R}^n$, and a unique solution $u \in C^\infty(A)$ of (5.15).

Proof of Corollary. Pick a smooth cutoff function $\chi \in C_c^\infty(\mathbb{R}^n)$ such that $\chi = 1$ on the unit ball centered at the origin, and set

$$f_j(x) = \chi\left(\frac{x}{j}\right) f(x), \quad g_j(x) = \chi\left(\frac{x}{j}\right) g(x), \quad j = 1, 2, \dots,$$

so that $f = f_j$ and $g = g_j$ in the unit ball at the origin with radius j. By Theorem 24 there exist $T_j > 0$ and $u_j \in C^\infty([0, T_j] \times \mathbb{R}^n)$ solving (5.15) with data f_j, g_j. We may of course assume $T_{j+1} \le T_j$. Now let K_j be the truncated backward solid light cone

$$K_j = \{(t, x) : 0 \le t \le T_j,\, |x| \le j - t\}.$$

Define

$$u : \bigcup_1^\infty K_j \to \mathbb{R}$$

by

$$u = u_j \quad \text{in} \quad K_j.$$

To see that this is well-defined and solves (5.15), it is enough to check that $u_j = u_k$ on $K_j \cap K_k$. But if $(t, x) \in K_j \cap K_k$, then the base of the backward solid light cone with vertex at (t, x) will be contained in the intersection of the bases of K_j and K_k, where $f_j = f_k = f$ and $g_j = g_k = g$. Therefore, $u_j(t, x) = u_k(t, x)$ by Theorem 20. Finally, it is easy to see that $\bigcup_1^\infty K_j$ contains a set of the form $A = \{(t, x) : 0 \le t \le T(x)\}$ where $T > 0$ is continuous. □

5.5 The model equation $\Box u = (\partial_t u)^2$

For simplicity, we shall first prove Theorems 22, 23 and 24 for the equation

$$\Box u = (\partial_t u)^2. \tag{5.17}$$

Here are the key facts we shall make use of:

(i) The ***energy inequality*** (5.3).

(ii) ***Sobolev's Lemma*** on $\mathbb{R}^n$, and in particular the inequality

$$\|f\|_{L^\infty} \le C_s \|f\|_{H^s} \quad \text{for all} \quad f \in H^s,\ s > \frac{n}{2}. \tag{5.18}$$

(iii) The following ***calculus inequality***,

$$\|fg\|_{H^s} \le C_s\big(\|f\|_{L^\infty} \|g\|_{H^s} + \|f\|_{H^s} \|g\|_{L^\infty}\big), \tag{5.19}$$

valid for all $s \ge 0$ and all $f, g \in H^s \cap L^\infty$.

(iv) A special case of ***Gronwall's Lemma***: If $C_1, C_2 \ge 0$ are given constants, and A is a continuous, non-negative function on $[0, T]$ such that

$$A(t) \le C_1 + C_2 \int_0^t A(\tau)\, d\tau \quad \text{for} \quad 0 \le t \le T,$$

then

$$A(t) \le C_1 e^{C_2 t} \quad \text{for} \quad 0 \le t \le T.$$

Sobolev's Lemma is stated and proved below. The Calculus inequality and Gronwall's Lemma will be proved later in the course. Note, however, that (5.19) is elementary for $s = 1$, since for $f, g \in \mathcal{S}$,

$$\begin{aligned}
\|fg\|_{H^1} &\approx \|fg\|_{L^2} + \sum_1^n \|\partial_j(fg)\|_{L^2} \\
&\le \|f\|_{L^\infty} \|g\|_{L^2} + \sum_1^n \big(\|(\partial_j f)g\|_{L^2} + \|f\partial_j g\|_{L^2}\big) \\
&\le \|f\|_{L^\infty} \|g\|_{L^2} + \sum_1^n \big(\|\partial_j f\|_{L^2} \|g\|_{L^\infty} + \|f\|_{L^\infty} \|\partial_j g\|_{L^2}\big) \\
&\le C\big(\|f\|_{L^\infty} \|g\|_{H^1} + \|f\|_{H^1} \|g\|_{L^\infty}\big).
\end{aligned}$$

To prove Theorems 22, 23 and 24 for a general nonlinearity $F(u, \partial u)$, we shall need to replace (5.19) by the more general ***Moser inequality***, but aside from this technicality the arguments are really identical. Thus, in studying first the model equation $\Box u = (\partial_t u)^2$, we gain transparency without any real loss of generality.

Before we start the proofs of the main theorems, let us prove Sobolev's Lemma.

Theorem 25. ***(Sobolev's Lemma.)*** *If $s > k + \frac{n}{2}$, where k is a non-negative integer, then*

$$H^s(\mathbb{R}^n) \subset C^k(\mathbb{R}^n) \cap L^\infty(\mathbb{R}^n),$$

where the inclusion is continuous. In fact,

$$\sum_{|\alpha|\le k} \|\partial^\alpha f\|_{L^\infty} \le C_s \|f\|_{H^s},$$

where C_s is independent of f.

Proof. The key observation is that if $f \in H^s$, where $s > \frac{n}{2}$, then by the Cauchy-Scwharz inequality,

$$\int |\widehat{f}(\xi)|\, d\xi = \int (1+|\xi|^2)^{-s/2}(1+|\xi|^2)^{s/2}|\widehat{f}(\xi)|\, d\xi \le C_s \|f\|_{H^s},$$

where

$$C_s^2 = \int (1+|\xi|^2)^{-s}\, d\xi < \infty.$$

Thus $f \in C^0 \cap L^\infty$ by Fourier inversion and the Riemann-Lebesgue Lemma, and

$$\|f\|_{L^\infty} \le \int |\widehat{f}(\xi)|\, d\xi \le C_s \|f\|_{H^s}.$$

This establishes the theorem for $k = 0$. Now suppose $f \in H^s$, $s > k + \frac{n}{2}$, $k \in \mathcal{N}$. Apply the special case just proved to

$$\partial^\alpha f \in H^{s-|\alpha|} \quad \text{for} \quad |\alpha| \le k$$

to see that $\partial^\alpha f \in C^0 \cap L^\infty$ and

$$\|\partial^\alpha f\|_{L^\infty} \le C_{s-|\alpha|} \|\partial^\alpha f\|_{H^{s-|\alpha|}} \le C_s \|f\|_{H^s}.$$

It now follows from Proposition 2 in last week's notes that $f \in C^k$. ☐

5.5.1 Proof of Theorem 22 (Will be covered on March 5)

Let $s > \frac{n}{2} + 1$. The problem under consideration is

$$\Box u = (\partial_t u)^2, \tag{5.20a}$$

$$u\big|_{t=0} = f \in H^s, \quad \partial_t u\big|_{t=0} = g \in H^{s-1} \tag{5.20b}$$

Existence

We use the following iteration scheme. First set

$$u_{-1} = 0.$$

Then define $u_0, u_1, \ldots$ inductively by

$$\Box u_j = (\partial_t u_{j-1})^2, \quad u_j\big|_{t=0} = f, \quad \partial_t u_j\big|_{t=0} = g. \tag{5.21}$$

For $T > 0$ we denote by X_T the space (5.16). Then X_T is a Banach space with norm

$$\|u\|_{X_T} = \sup_{0 \le t \le T} \left(\|u(t)\|_{H^s} + \|\partial_t u(t)\|_{H^{s-1}} \right).$$

Now observe that

$$u_{j-1} \in X_T \implies u_j \in X_T.$$

Indeed, by Sobolev's Lemma and the calculus inequality (5.19) we have, since $s - 1 > \frac{n}{2}$,

$$\|vw\|_{H^{s-1}} \le C_s \|v\|_{H^{s-1}} \|w\|_{H^{s-1}} \tag{5.22}$$

for all $v, w \in H^{s-1}(\mathbb{R}^n)$. Thus

$$u_{j-1} \in X_T \implies (\partial_t u_{j-1})^2 \in C([0,T], H^{s-1}),$$

and it follows from Theorem 1 that (5.21) has a unique solution in X_T.

The sequence of iterates is therefore well-defined in X_T for any $T > 0$. Our aim is to prove that this sequence is Cauchy provided $T > 0$ is taken sufficiently small. Then the limit $u \in X_T$ will be a solution of (5.20a) on $S_T = (0,T) \times \mathbb{R}^n$, since

$$\begin{aligned} u_j \to u \quad \text{in} \quad X_T \quad &\implies \quad u_j \to u \quad \text{in} \quad \mathcal{D}'(S_T) \\ &\implies \quad \Box u_j \to \Box u \quad \text{in} \quad \mathcal{D}'(S_T), \end{aligned}$$

and since, by (5.22),

$$\begin{aligned} u_j \to u \quad \text{in} \quad X_T \quad &\implies \quad (\partial_t u_j)^2 \to (\partial_t u)^2 \quad \text{in} \quad C([0,T], H^{s-1}) \\ &\implies \quad (\partial_t u_j)^2 \to (\partial_t u)^2 \quad \text{in} \quad \mathcal{D}'(S_T). \end{aligned}$$

The initial condition (5.20b) is evidently satisfied, so u solves the Cauchy problem.

We now prove that $\{u_j\}$ is Cauchy in X_T for $T > 0$ sufficiently small. This is done in two steps.

Step 1. The sequence is bounded:

$$\|u_j\|_{X_T} \le 2CE_s \quad \text{for} \quad j = 0, 1, \ldots \tag{5.23}$$

provided $0 < T \le 1$ is so small that

$$T \le \frac{1}{8C^2 E_s}, \tag{5.24}$$

where

$$E_s = \|f\|_{H^s} + \|g\|_{H^{s-1}}$$

and $C > 0$ is a constant which only depends on s and n. If (5.23) holds for $j-1$, then applying the energy inequality (5.3) to equation (5.21) and using the estimate (5.22), we get

$$\|u_j\|_{X_T} \le CE_s + CT\|u_{j-1}\|_{X_T}^2 \le CE_s + CT(2CE_s)^2$$

where C only depends on s and n if $T \le 1$ [C comes from the energy inequality and the estimate (5.22)]. If (5.24) holds, then the right hand side is bounded by $2CE_s$, so we obtain (5.23) for all j by induction, since the case $j = -1$ holds trivially.

Step 2. The sequence satisfies

$$\|u_{j+1} - u_j\|_{X_T} \le \frac{1}{2}\|u_j - u_{j-1}\|_{X_T},$$

and is therefore Cauchy. To prove this inequality, note that

$$\Box(u_{j+1} - u_j) = \partial_t(u_j - u_{j-1})\partial_t(u_j + u_{j-1})$$

with vanishing initial data at $t = 0$. Thus, if we apply the energy inequality and use (5.22) as well as the uniform bound (5.23), we get

$$\begin{aligned}\|u_{j+1} - u_j\|_{X_T} &\le CT\left(\|u_j\|_{X_T} + \|u_{j-1}\|_{X_T}\right)\|u_j - u_{j-1}\|_{X_T}\\ &\le CT4CE_s\|u_j - u_{j-1}\|_{X_T}.\end{aligned}$$

Thus we get the desired bound using (5.24), and this completes the proof of existence.

Uniqueness

Suppose $u, v \in X_T$ for some $T > 0$ and solve (5.20a) on $S_T = (0,T) \times \mathbb{R}^n$ with identical initial data at $t = 0$. Then

$$\Box(u - v) = \partial_t(u + v)\partial_t(u - v)$$

and setting

$$A(t) = \|(u - v)(t)\|_{H^{s-1}} + \|\partial_t(u - v)(t)\|_{H^{s-1}},$$

we have by the energy inequality and (5.22),

$$A(t) \le C\int_0^t A(t')\,dt' \quad \text{for} \quad 0 \le t \le T$$

for some constant C independent of t. (Note that C depends on the norms $\|u\|_{X_T}$ and $\|v\|_{X_T}$, but this is not a problem since u and v are considered fixed for this argument.) By Gronwall's Lemma, $E(t) = 0$ for $0 \le t \le T$, whence $u = v$ in S_T.

5.5.2 Proof of Theorem 23

Let $s > \frac{n}{2} + 1$. Let $f \in H^s$, $g \in H^{s-1}$. It suffices to prove that if $0 < T < \infty$ and

$$u \in C([0,T), H^s) \cap C^1([0,T), H^{s-1})$$

solves (5.20) on $S_T = (0,T) \times \mathbb{R}^n$ and satisfies

$$\partial u \in L^\infty(S_T),$$

then

$$\sup_{0 \le t < T} \left(\|u(t)\|_{H^s} + \|\partial_t u(t)\|_{H^{s-1}} \right) < \infty. \tag{5.25}$$

For we can then extend the solution to a time strip $[0, T+\varepsilon] \times \mathbb{R}^n$, for some $\varepsilon > 0$, in view of Theorem 22.

To prove (5.25), set

$$A(t) = \|u(t)\|_{H^s} + \|\partial_t u(t)\|_{H^{s-1}}.$$

By the energy inequality and (5.19), we have

$$\begin{aligned} A(t) &\le C_{s,T} \left(\|f\|_{H^s} + \|g\|_{H^{s-1}} + \int_0^t \|\partial_t u(t')\|_{L^\infty} \|\partial_t u(t')\|_{H^{s-1}} \, dt' \right) \\ &\le C_{s,T} \left(\|f\|_{H^s} + \|g\|_{H^{s-1}} + \|\partial u\|_{L^\infty(S_T)} \int_0^t A(t') \, dt' \right), \end{aligned}$$

and Gronwall's Lemma gives (5.25).

5.5.3 Proof of Theorem 24

Let $f, g \in C_c^\infty(\mathbb{R}^n)$. Then $f, g \in H^s$ for every $s \in \mathbb{R}$. Fix $s_0 > \frac{n}{2} + 1$. By Theorem 22, there exist $T > 0$ and a unique solution

$$u \in C([0,T], H^{s_0}) \cap C^1([0,T], H^{s_0-1}) \tag{5.26}$$

of (5.15) on $S_T = (0,T) \times \mathbb{R}^n$. By the same theorem, for every $s > s_0$ there exists $T_s > 0$ such that

$$u \in C([0,T_s], H^s) \cap C^1([0,T_s], H^{s-1}).$$

We claim that we can take $T_s = T$. Indeed, this follows from Theorem 23, since

$$\partial u \in L^\infty(S_T)$$

by Sobolev's Lemma and (5.26). Thus

$$u \in C([0,T], H^s) \cap C^1([0,T], H^{s-1})$$

for *every* $s \geq s_0$, so by Sobolev's Lemma,

$$\partial_t^j \partial_x^\alpha u \in C([0,T] \times \mathbb{R}^n) \tag{5.27}$$

for $j = 0, 1$ and all α. But u solves (5.15a), so

$$\partial_t^2 u = \Delta u + (\partial_t u)^2 \tag{5.28}$$

From this and (5.27) it follows that

$$\partial_t^2 \partial_x^\alpha u \in C([0,T] \times \mathbb{R}^n),$$

so (5.27) holds for $j = 0, 1, 2$ and all α. Applying ∂_t to both sides of (5.28) yields

$$\partial_t^3 u = \Delta \partial_t u + 2\partial_t u \partial_t^2 u,$$

and from this we see that (5.27) holds for $j = 0, 1, 2, 3$. If we keep taking successive time derivatives of the equation, we obtain (5.27) for all j by induction. It then follows from Proposition 2 in last week's notes that $u \in C^\infty([0,T] \times \mathbb{R}^n)$.

5.5.4 Exercises

Exercise 24. Prove analogues of the main theorems in section 4 for the equation $\Box u = u^2$. Show that the condition on s can now be weakened to $s > \frac{n}{2}$ and that the condition $\partial u \in L^\infty(S_T)$ in the continuation argument (Theorem 23) can be replaced by $u \in L^\infty(S_T)$.

Chapter 6

Week 6: Littlewood-Paley theory

Throughout, all functions and distributions are understood to be defined on $\mathbb{R}^n$ unless otherwise stated.

6.1 Littlewood-Paley decomposition

Fix a radial cut-off function $\chi \in C_c^\infty(\mathbb{R}^n)$ such that

$$0 \le \chi \le 1, \qquad \chi(\xi) = \begin{cases} 1 & \text{if } |\xi| \le \frac{1}{2}, \\ 0 & \text{if } |\xi| \ge 1. \end{cases} \tag{6.1}$$

Define $S_j : \mathcal{S}' \to \mathcal{S}'$ by

$$\widehat{S_j f} = \chi(2^{-j}\xi)\widehat{f}$$

for $j \in \mathcal{N}_0$. In other words,

$$S_j = \mathcal{F}^{-1}\chi(2^{-j}\cdot)\mathcal{F}.$$

Lemma 2. *For every $\phi \in \mathcal{S}$, $\chi(2^{-j}\cdot)\phi \to \phi$ in $\mathcal{S}$ as $j \to \infty$.*

We leave the proof as an easy exercise. Since $\mathcal{F}$ is an isomorphism of $\mathcal{S}$ onto itself, we obtain immediately:

$$S_j f \to f \quad \text{in} \quad \mathcal{S}$$

for every $f \in \mathcal{S}$. This, in turn, implies that $S_j f \to f$ in $\mathcal{S}'$ for every $f \in \mathcal{S}'$.

Next, define

$$\Delta_0 = S_0, \qquad \Delta_j = S_j - S_{j-1} \quad \text{for} \quad j \in \mathcal{N}.$$

Thus $f = \sum_0^\infty \Delta_j f$ in $\mathcal{S}$ (resp. $\mathcal{S}'$) for every $f \in \mathcal{S}$ (resp. $f \in \mathcal{S}'$).

Definition. $\{\Delta_j f\}_0^\infty$ is called the ***Littlewood-Paley decomposition*** of f.

Observe that $\widehat{\Delta_j f} = \beta_j(\xi)\widehat{f}$, where

$$\beta_0(\xi) = \chi(\xi), \qquad \beta_j(\xi) = \chi(2^{-j}\xi) - \chi(2^{1-j}\xi) \quad \text{for} \quad j \in \mathcal{N}.$$

From (6.1) it is evident that $0 \le \beta_j \le 1$ for all j, whence

$$\sum_0^\infty \beta_j^2 \le \sum_0^\infty \beta_j = 1. \tag{6.2}$$

For later use we record some elementary properties of the Littlewood-Paley decomposition:

(i) $\operatorname{supp} \widehat{S_j f} \subset \{|\xi| \le 2^j\}$.

(ii) $\operatorname{supp} \widehat{\Delta_j f} \subset \{2^{j-2} \le |\xi| \le 2^j\}$.

(iii) $S_j f = 2^{jn}\psi(2^j \cdot) * f$ where $\widehat{\psi} = \chi$.

(iv) $\Delta_j f = 2^{jn}\phi(2^j \cdot) * f$ where $\widehat{\phi} = \chi(\xi) - \chi(2\xi)$.

(v) $\|S_j f\|_{L^p} \le C\|f\|_{L^p}$ where $C = \|\psi\|_{L^1}$.

(vi) $\|\Delta_j f\|_{L^p} \le C\|f\|_{L^p}$ where $C = \|\phi\|_{L^1}$.

(vii) $S_j f \in C^\infty$ if $f \in L^p$, $1 \le p \le \infty$.

(viii) $\Delta_j f \in C^\infty$ if $f \in L^p$, $1 \le p \le \infty$.

Property (v) [resp. (vi)] follows from (iii) [resp. (iv)] and ***Young's Inequality***:

$$\|f * g\|_{L^p} \le \|f\|_{L^1}\|g\|_{L^p}, \quad 1 \le p \le \infty. \tag{6.3}$$

6.2 Littlewood-Paley decomposition of H^s

Recall that

$$H^s = \left\{f \in \mathcal{S}' : \Lambda^s f \in L^2\right\},$$

where $\Lambda^s = (I - \Delta)^{s/2}$. Thus H^s is a Hilbert space with inner product

$$\langle f, g \rangle_{H^s} = \langle \Lambda^s f, \Lambda^s g \rangle_{L^2} = \int (1 + |\xi|^2)^s \widehat{f}(\xi)\overline{\widehat{g}(\xi)}\, d\xi \tag{6.4}$$

and norm $\|f\|_{H^s} = \|\Lambda^s f\|_{L^2}$, and $\Lambda^s : H^s \to L^2$ is a unitary isomorphism. Observe that if the supports of $\widehat{f}$ and $\widehat{g}$ are almost disjoint, then $f \perp g$ in H^s.

Being a Hilbert space, H^s is self-dual via the map $g \mapsto \langle \cdot, g \rangle_{H^s}$. A more interesting fact is that the dual of H^s can be identified with H^{-s} via the pairing $\langle \cdot, \cdot \rangle : \mathcal{S}' \times \mathcal{S} \to \mathbb{C}$. Indeed,

$$|\langle f, g \rangle| = \left|\left\langle \Lambda^s f, \Lambda^{-s} g \right\rangle\right| = \left|\int \Lambda^s f \Lambda^{-s} g\right| \le \|f\|_{H^s}\|g\|_{H^{-s}} \tag{6.5}$$

for all $f \in H^s$ and $g \in \mathcal{S}$. Thus, the map $g \mapsto \langle \cdot, g \rangle$ extends to a linear map from H^{-s} into $(H^s)^*$, which is a surjective isometry in view of the self-duality of L^2.

6.2.1 Characterization of H^s

We want to characterize H^s in terms of the Littlewood-Paley decomposition. We begin with some elementary observations.

Observation 1. $S_j f \to f$ in H^s for every $f \in H^s$. Indeed,

$$\|S_j f - f\|_{H^s}^2 = \int \left[1 - \chi(2^{-j}\xi)\right]^2 (1 + |\xi|^2)^s \left|\widehat{f}(\xi)\right|^2 d\xi \to 0$$

by the Dominated Convergence Theorem.

Observation 2. If $2^{j-2} \leq |\xi| \leq 2^j$, then

$$(1 + |\xi|^2)^{s/2} \approx 2^{js}$$

in the sense that there is a constant $C_s > 0$, independent of j, such that

$$C_s^{-1} 2^{js} \leq (1 + |\xi|^2)^{s/2} \leq C_s 2^{js}.$$

It then follows that

$$\|\Delta_j f\|_{H^s} \approx 2^{js} \|\Delta_j f\|_{L^2} . \tag{6.6}$$

Observation 3. In view of (ii) in the previous section,

$$\Delta_j f \perp \Delta_k f \quad \text{if} \quad |j - k| \geq 2, \tag{6.7}$$

relative to the inner product on H^s.

We can now prove the following:

Proposition 5. *Let $s \in \mathbb{R}$. Then for every $f \in H^s$,*

$$\|f\|_{H^s}^2 \approx \sum_0^\infty 2^{2js} \|\Delta_j f\|_{L^2}^2 .$$

Proof. In view of (6.6), it suffices to show that

$$\|f\|_{H^s}^2 \approx \sum_0^\infty \|\Delta_j f\|_{H^s}^2 .$$

Using the orthogonality property (6.7) and the Cauchy-Schwarz inequality,

$$\begin{aligned}
\|f\|_{H^s}^2 &= \langle f, f \rangle_{H^s} = \Big\langle \sum \Delta_j f, \sum \Delta_k f \Big\rangle_{H^s} \\
&= \sum \sum \langle \Delta_j f, \Delta_k f \rangle_{H^s} = \sum_{l=-1}^{1} \sum_{j=0}^{\infty} \langle \Delta_j f, \Delta_{j+l} f \rangle_{H^s} \\
&\le \sum_{l=-1}^{1} \sum_{j=0}^{\infty} \|\Delta_j f\|_{H^s} \|\Delta_{j+l} f\|_{H^s} \\
&\le \sum_{l=-1}^{1} \Big(\sum_{j=0}^{\infty} \|\Delta_j f\|_{H^s}^2 \Big)^{\frac{1}{2}} \Big(\sum_{j=0}^{\infty} \|\Delta_{j+l} f\|_{H^s}^2 \Big)^{\frac{1}{2}} \le 3 \sum_{j=0}^{\infty} \|\Delta_j f\|_{H^s}^2 .
\end{aligned}$$

On the other hand, we have by (6.2),

$$\begin{aligned}
\sum_0^{\infty} \|\Delta_j f\|_{H^s}^2 &= \sum_0^{\infty} \int (1+|\xi|)^s \big|\beta_j(\xi) \widehat{f}(\xi)\big|^2 \, d\xi \\
&= \int \Big[\sum_0^{\infty} \beta_j^2(\xi) \Big] (1+|\xi|)^s \big|\widehat{f}(\xi)\big|^2 \, d\xi \\
&\le \int (1+|\xi|)^s \big|\widehat{f}(\xi)\big|^2 \, d\xi = \|f\|_{H^s}^2 ,
\end{aligned}$$

and the proof is complete. □

Using the ideas in the proof above, it is easy to show that if $\{f_j\}$ is a sequence in L^2 such that for some $R > 0$,

$$\operatorname{supp} \widehat{f_j} \subset \{R^{-1} 2^j \le |\xi| \le R 2^j\}$$

and $\sum 2^{2js} \|f_j\|_{L^2}^2 < \infty$, then $f = \sum f_j$ converges in H^s, and

$$\|f\|_{H^s}^2 \approx \sum 2^{2js} \|f_j\|_{L^2}^2 .$$

It is interesting that when $s > 0$, we can get essentially the same result if we only assume

$$\operatorname{supp} \widehat{f_j} \subset \{|\xi| \le R 2^j\}. \tag{6.8}$$

Proposition 6. *Let $s > 0$. Suppose $f_j \in L^2$ satisfies (6.8), for some $R \ge 1$ and that*

$$\sum_0^{\infty} 2^{2js} \|f_j\|_{L^2}^2 < \infty.$$

Then $f = \sum f_j$ converges in H^s, and

$$\|f\|_{H^s}^2 \le C_{s,R} \sum_0^{\infty} 2^{2js} \|f_j\|_{L^2}^2 . \tag{6.9}$$

Remark. If $R = 2^q$, then the constant $C_{s,R}$ is of the form $C_s 2^{2qs}$. We shall need this remark in the proof of the Moser inequality.

To prove this proposition, we shall need the following elementary estimate:

Lemma 3. *Let $s < 0$ and $R \geq 1$. If $f \in H^s$ and*

$$\operatorname{supp} \widehat{f} \subset \{|\xi| \leq R\},$$

then

$$\|f\|_{H^s} \leq 2^s R^s \, \|f\|_{L^2} .$$

Proof. We only have to observe that

$$|\xi| \leq R \implies (1 + |\xi|^2)^s \leq (2R^2)^s.$$

□

We now turn to the proof of Proposition 6. It suffices to show that

$$\left\langle \sum_M^N f_j, \sum_M^N f_k \right\rangle_{H^s} \leq C_{s,R} \sum_M^\infty 2^{2js} \, \|f_j\|_{L^2}^2 \tag{6.10}$$

whenever $M < N$. For then it follows that the sequence of partial sums is Cauchy in H^s, and hence the series converges to some $f \in H^s$. Then, taking $M = 1$ and letting $N \to \infty$, we obtain (6.9).

The left hand side of (6.10) equals

$$\sum_M^N \sum_M^N \langle f_j, f_k \rangle_{H^s} = \sum_{j<k} + \sum_{j=k} + \sum_{j>k} .$$

It suffices to estimate the first two sums on the right hand side. For the second sum we have, using Lemma 3,

$$\sum_{j=k} = \sum_M^N \langle f_j, f_j \rangle_{H^s} = \sum_M^N \|f_j\|_{H^s}^2 \leq C_s R^s \sum_M^N 2^{2js} \, \|f_j\|_{L^2}^2 .$$

Next we consider

$$\sum_{j<k} = \sum_{j=M}^{N-1} \sum_{k=j+1}^{N} \langle f_j, f_k \rangle_{H^s} .$$

Let $q \in \mathcal{N}$ be the smallest number satisfying $R \leq 2^q$. In view of (6.8) and property (i) in the previous section, we then have $f_j = S_{j+q+1} f_j$. Thus [cf. (6.4)]

$$\langle f_j, f_k \rangle_{H^s} = \langle S_{j+q+1} f_j, f_k \rangle_{H^s} = \langle f_j, S_{j+q+1} f_k \rangle_{H^s} .$$

Using Lemma 3, we then obtain

$$\begin{aligned}
\sum_{j<k} &= \sum_{j=M}^{N-1} \sum_{k=j+1}^{N} \langle f_j, S_{j+q+1} f_k \rangle_{H^s} \leq \sum \sum \|f_j\|_{H^s} \|S_{j+q+1} f_k\|_{H^s} \\
&\leq C_s \sum \sum R^s 2^{js} \|f_j\|_{L^2} 2^{(j+q+1)s} \|f_k\|_{L^2} \\
&= C_{s,R} \sum_{j=M}^{N-1} \sum_{k=j+1}^{N} 2^{js} \|f_j\|_{L^2} 2^{js} \|f_k\|_{L^2} \\
&\leq C_{s,R} \sum_{l=1}^{\infty} 2^{-ls} \sum_{j=M}^{\infty} 2^{js} \|f_j\|_{L^2} 2^{(j+l)s} \|f_{j+l}\|_{L^2} \\
&\leq C_{s,R} \sum_{l=1}^{\infty} 2^{-ls} \left(\sum_{j=M}^{\infty} 2^{2js} \|f_j\|_{L^2}^2 \right)^{\frac{1}{2}} \left(\sum_{j=M}^{\infty} 2^{2(j+l)s} \|f_{j+l}\|_{L^2}^2 \right)^{\frac{1}{2}} \\
&\leq C_{s,R} \sum_{l=1}^{\infty} 2^{-ls} \sum_{j=M}^{\infty} 2^{2js} \|f_j\|_{L^2}^2 .
\end{aligned}$$

This concludes the proof since $\sum_{l=1}^{\infty} 2^{-ls} < \infty$.

6.2.2 The Calculus Inequality

Here we employ the machinery just developed to prove the inequality

$$\|fg\|_{H^s} \leq C_s \left(\|f\|_{H^s} \|g\|_{L^\infty} + \|f\|_{L^\infty} \|g\|_{H^s} \right), \tag{6.11}$$

where $s \geq 0$ and $f, g \in H^s \cap L^\infty$. Observe that if $s = 0$ this follows from Hölder's inequality, so we assume $s > 0$ from now on. The proof is split into three steps.

Step 1. Write

$$fg = \sum \sum \Delta_j f \Delta_k g = \sum_{j \leq k-3} + \sum_{|j-k| \leq 2} + \sum_{j \geq k+3} = \sum\nolimits_1 + \sum\nolimits_2 + \sum\nolimits_3 .$$

By symmetry it suffices to estimate $\sum_1$ and $\sum_2$.

Remark. The above equation holds in the sense of L^1, since $f = \sum \Delta_j f$ and $g = \sum \Delta_k g$ in the sense of L^2. Observe also that $\Delta_j f$ and $\Delta_k g$ are smooth and bounded functions, by properties (vi) and (viii) from the previous section.

Step 2. We estimate $\|\sum_2\|_{H^s}$. By property (ii), we see that if $|j-k| \leq 2$, then

$$\operatorname{supp} \mathcal{F}(\Delta_j f \Delta_k g) \subset \{|\xi| \leq 2^{j+3}\}.$$

In view of Proposition 6, it therefore suffices to check that

$$\sum 2^{2js} \|\Delta_j f \Delta_{j+l} g\|_{L^2}^2 \leq C_s \|f\|_{H^s}^2 \|g\|_{L^\infty}^2 \tag{6.12}$$

for $l = -2, \ldots, 2$. But

$$\|\Delta_j f \Delta_{j+l} g\|_{L^2} \leq \|\Delta_j f\|_{L^2} \|\Delta_{j+l} g\|_{L^\infty} \leq \|\Delta_j f\|_{L^2} \|g\|_{L^\infty},$$

so the left hand side of (6.12) is bounded by

$$\left(\sum 2^{2js} \|\Delta_j f\|_{L^2}^2\right) \|g\|_{L^\infty}^2 \approx \|f\|_{H^s}^2 \|g\|_{L^\infty}^2 .$$

Step 3. We estimate $\|\sum_1\|_{H^s}$. Write

$$\sum\nolimits_1 = \sum_{k=3}^{\infty} \sum_{j=0}^{k-3} \Delta_j f \Delta_k g = \sum_{k=3}^{\infty} S_{k-3} f \Delta_k g.$$

Observe that

$$\operatorname{supp} \mathcal{F}(S_{k-3} f \Delta_k g) \subset \{2^{k-3} \leq |\xi| \leq 2^{k+1}\}.$$

Thus

$$S_{j-3} f \Delta_j g \perp S_{k-3} f \Delta_k g \quad \text{if} \quad |j-k| \geq 4.$$

Consequently, with the convention that $S_j = 0$ for $j = -1, -2, \ldots$, we have

$$\begin{aligned}
\left\|\sum\nolimits_1\right\|_{H^s}^2 &= \left\langle \sum S_{j-3} f \Delta_j g, \sum S_{k-3} f \Delta_k g \right\rangle \\
&= \sum_{|j-k| \leq 3} \langle S_{j-3} f \Delta_j g, S_{k-3} f \Delta_k g \rangle \\
&\leq \sum_{|j-k| \leq 3} \|S_{j-3} f \Delta_j g\|_{H^s} \|S_{k-3} f \Delta_k g\|_{H^s} \\
&= \sum_{l=-3}^{3} \sum_{j=0}^{\infty} \|S_{j-3} f \Delta_j g\|_{H^s} \|S_{j+l-3} f \Delta_{j+l} g\|_{H^s} \\
&\leq \sum_{l=-3}^{3} \left(\sum_{j=0}^{\infty} \|S_{j-3} f \Delta_j g\|_{H^s}^2\right)^{\frac{1}{2}} \left(\sum_{j=0}^{\infty} \|S_{j+l-3} f \Delta_{j+l} g\|_{H^s}^2\right)^{\frac{1}{2}} \\
&\leq 7 \sum \|S_{j-3} f \Delta_j g\|_{H^s}^2 \\
&\leq C_s \sum 2^{2js} \|S_{j-3} f \Delta_j g\|_{L^2}^2 \\
&\leq C_s \sum 2^{2js} \|S_{j-3} f\|_{L^\infty}^2 \|\Delta_j g\|_{L^2}^2 \\
&\leq C_s \|f\|_{L^\infty}^2 \sum 2^{2js} \|\Delta_j g\|_{L^2}^2 ,
\end{aligned}$$

which completes the proof of the inequality (6.11).

6.3 Moser's inequality

The following is known as the ***Moser inequality***.

Theorem 26. *Assume that $F : \mathbb{R}^N \to \mathbb{R}$ is C^∞ and $F(0) = 0$. Then for all $s \geq 0$ there is a continuous function $\gamma = \gamma_s : \mathbb{R} \to \mathbb{R}$ such that*

$$\|F(f)\|_{H^s} \leq \gamma(\|f\|_{L^\infty}) \|f\|_{H^s}$$

for all $\mathbb{R}^N$-valued $f \in H^s \cap L^\infty$.

It is then relatively easy to prove:

Corollary. *If $s > \frac{n}{2}$, the map $f \mapsto F(f)$, $H^s \to H^s$ is C^∞.*

To prove the Moser inequality we shall need the following generalization of Lemma 3. Recall that the *spectrum* of f is the support of its Fourier transform.

Lemma 4. ***(Bernstein's Lemma.)*** *Assume that the spectrum of $f \in L^p$, $1 \leq p \leq \infty$, is contained in the ball $|\xi| \leq 2^j$. Then*

$$\|\partial^\alpha f\|_{L^p} \leq C_\alpha 2^{j|\alpha|} \|f\|_{L^p}$$

for any multi-index α. Moreover, if the spectrum is contained in $2^{j-2} \leq |\xi| \leq 2^j$, then

$$C_k^{-1} 2^{jk} \|f\|_{L^p} \leq \sup_{|\alpha|=k} \|\partial^\alpha f\|_{L^p} \leq C_k 2^{jk} \|f\|_{L^p}$$

for any $k \in \mathcal{N}_0$.

We omit the proof (essentially an application of Young's inequality).

We now have the necessary tools to prove the Moser inequality. For $s = 0$, the proof is trivial, so we assume $s > 0$. We shall in fact prove that

$$\|F(S_p f) - F(S_q f)\|_{H^s} \leq \gamma(\|f\|_{L^\infty}) \|S_p f - S_q f\|_{H^s} \tag{6.13}$$

for all $p > q \geq -1$, where it is understood that $S_{-1} \equiv 0$, and where γ depends on the partial derivatives of F of order $1, \dots, M$, with M is the smallest integer greater than s. (Observe that $S_p f$ is smooth and bounded for $f \in H^s \cap L^\infty$.)

Assuming this, since $S_j f \to f$ in H^s, it follows that $F(S_j f)$ is Cauchy in H^s, and

$$\|F(S_j f)\|_{H^s} \leq \gamma(\|f\|_{L^\infty}) \|f\|_{H^s}$$

for all j. But the limit of $F(S_j f)$ in H^s must necessarily be $F(f)$, by an a.e. argument. Thus we obtain Moser's inequality.

To prove (6.13), we write

$$F(S_p f) - F(S_q f) = \sum_{q+1}^{p} [F(S_j f) - F(S_{j-1} f)] = \sum_{q+1}^{p} m_j \Delta_j f,$$

where

$$m_j = \int_0^1 F'(S_{j-1} f + \lambda \Delta_j f)\, d\lambda.$$

We claim that for all $j, k, M \in \mathcal{N}_0$,

$$\|\Delta_{j+k} m_j\|_{L^\infty} \le c_M(\|f\|_{L^\infty}) 2^{-Mk} \tag{6.14}$$

where c_M is continuous.

Indeed, by the chain rule, $\partial^\alpha m_j$ is a linear combination of terms

$$\int_0^1 F^{(l+1)}(g_\lambda)(\partial^{\alpha_1} g_\lambda, \ldots, \partial^{\alpha_l} g_\lambda)\, d\lambda,$$

where $l \le |\alpha|$, $\alpha_1 + \cdots + \alpha_l = \alpha$ and $g_\lambda = S_{j-1} f + \lambda \Delta_j f$, so by Bernstein's Lemma,

$$\|\partial^\alpha m_j\|_{L^\infty} \le c_\alpha(\|f\|_{L^\infty}) 2^{j|\alpha|}, \tag{6.15}$$

where c_α is a continuous function which depends on α and on the derivatives of F' up to order $|\alpha|$. But by Bernstein's Lemma again,

$$\|\Delta_{j+k} m_j\|_{L^\infty} \le C_M 2^{-M(j+k)} \sup_{|\alpha|=M} \|\partial^\alpha \Delta_{j+k} m_j\|_{L^\infty},$$

which combined with (6.15) proves (6.14).

Writing $m_j = S_j m_j + \sum_1^\infty \Delta_{j+k} m_j$, we have

$$F(S_p f) - F(S_q f) = A + \sum_1^\infty B_k,$$

where $A = \sum_{j=q+1}^p S_j m_j \Delta_j f$ and $B_k = \sum_{j=q+1}^p \Delta_{j+k} m_j \Delta_j f$. Using Proposition 6 and the remark following it, the decay estimates (6.14) and (6.15), and the fact that

$$\|S_p f - S_q f\|_{H^s}^2 = \left\|\sum_{q+1}^p \Delta_j f\right\|_{H^s}^2 \sim \sum_{q+1}^p \|\Delta_j f\|_{H^s}^2,$$

it then follows that

$$\begin{aligned} \|A\|_{H^s} &\le C_s c_0(\|f\|_{L^\infty}) \|S_p f - S_q f\|_{H^s}, \\ \|B_k\|_{H^s} &\le C_s c_M(\|f\|_{L^\infty}) 2^{-Mk} 2^{ks} \|S_p f - S_q f\|_{H^s} \end{aligned}$$

for any $M \in \mathcal{N}$. By taking $M > s$ and summing over k, we obtain the desired inequality.

Remark. Observe that if the partial derivatives of F of order $1, \ldots, M+1$ are all bounded on $\mathbb{R}^N$, where M is the smallest integer strictly greater than s, then there is in fact no dependence on $\|f\|_{L^\infty}$ in (6.15), and ditto in (6.14). Under this assumption, we therefore have the inequality, for all $f \in H^s$, $s \ge 0$,

$$\|F(f)\|_{H^s} \le C_s \|f\|_{H^s}.$$

As an example, one may consider $F(y) = e^{iy} - 1$ for $y \in \mathbb{R}$.

6.4 Further applications of the Littlewood-Paley theory

We state without proof the following theorem.

Theorem 27. *If $2 \le p < \infty$, then*

$$\|f\|_{L^p}^2 \le C \sum_0^\infty \|\Delta_j f\|_{L^p}^2 ,$$

and if $1 < p \le 2$, then

$$\sum_0^\infty \|\Delta_j f\|_{L^p}^2 \le C \|f\|_{L^p}^2 .$$

Using this result one can easily prove the Sobolev embeddings for H^s.

Embeddings and non-embeddings for H^s

Recall that

$$H^s \hookrightarrow L^{\frac{2n}{n-2s}}, \qquad 0 \le s < \frac{n}{2}, \tag{6.16}$$

$$H^s \hookrightarrow L^\infty, \qquad s > \frac{n}{2}. \tag{6.17}$$

Observe that the point mass δ belongs to H^{-s} when $s > n/2$. Thus

$$|f(x)| = |\langle \delta(\cdot - x), f \rangle| \le \|\delta(\cdot - x)\|_{H^{-s}} \|f\|_{H^s} = \|\delta\|_{H^{-s}} \|f\|_{H^s} ,$$

which proves (6.17). To prove (6.16), first note that by Young's inequality,

$$\|\Delta_j f\|_{L^{\frac{2n}{n-2s}}} \le C 2^{js} \|f\|_{L^2} .$$

Since $\Delta_j f = \Delta_j \left(\sum_{k=j-1}^{j+1} \Delta_k f\right)$, it follows that

$$\|\Delta_j f\|_{L^{\frac{2n}{n-2s}}} \le C \sum_{k=j-1}^{j+1} 2^{ks} \|\Delta_k f\|_{L^2} .$$

Now square both sides, sum over j, and use Theorem 27 on the left hand side.

Next, we show that the embedding (6.17) fails unless $s > \frac{n}{2}$.

Proposition 7. *$H^{n/2}$ is not a subset of L^∞.*

Proof. Assume $H^{n/2} \subseteq L^\infty$. Then by the Closed Graph Theorem, we actually have $H^{n/2} \hookrightarrow L^\infty$, so

$$\|f\|_{L^\infty} \le C \|f\|_{H^{n/2}} .$$

But this implies that $\delta \in (H^{n/2})^* = H^{-n/2}$, which is false. □

Corollary. *$H^{n/2}$ is not an algebra.*

Proof. Assume that it is. Then $\left\|f^k\right\|_{H^{n/2}} \leq C^k \left\|f\right\|_{H^{n/2}}^k$ for all $k \in \mathcal{N}$. Hence, if $\|f\|_{H^{n/2}} < 1/C$, then $f^k \to 0$ in $H^{n/2}$ as $k \to \infty$. But this implies that some subsequence converges to zero a.e on $\mathbb{R}^n$, so we must have $|f| < 1$ a.e. But this means that the ball of radius $1/C$ in $H^{n/2}$ is contained in L^∞, which implies that H^s is a subset of L^∞, contradicting Proposition 7. □

Chapter 7

Week 7: Global existence results

7.1 Statement of main theorem

Consider the nonlinear Cauchy problem on $\mathbb{R}^{1+n}$,

$$\square u = F(\partial u), \tag{7.1}$$

$$u\big|_{t=0} = \varepsilon f, \quad \partial_t u\big|_{t=0} = \varepsilon g, \tag{7.2}$$

where $F : \mathbb{R}^{1+n} \to \mathbb{R}$ is a given C^∞ function which vanishes to second order at the origin:

$$F(0) = 0, \quad DF(0) = 0. \tag{7.3}$$

We shall prove:

Theorem 28. *Let $n \geq 4$. Let $f, g \in C_c^\infty(\mathbb{R}^n)$. Then there exists $\varepsilon_0 > 0$ such that* (7.1)*,* (7.2) *has a solution*

$$u \in C^\infty\big([0,\infty) \times \mathbb{R}^n\big)$$

provided $\varepsilon \leq \varepsilon_0$.

Note carefully that ε_0 depends on f and g, which are considered fixed. Recall also that the solution is unique, by the theory from week 5.

In dimensions $n = 1, 2, 3$ we will obtain (from the proof of the above theorem) asymptotic lower bounds on the lifespan

$$T_\varepsilon = T^*(\varepsilon f, \varepsilon g)$$

as $\varepsilon \to 0$. Recall that the lifespan is the supremum of $T > 0$ such that (7.1), (7.2) has a solution $u \in C^\infty([0,T] \times \mathbb{R}^n)$. Specifically, we shall see that there exists $c > 0$ such that

$$\begin{aligned} T_\varepsilon &\geq e^{c/\varepsilon}, && (n = 3) \\ T_\varepsilon &\geq c/\varepsilon^2, && (n = 2) \\ T_\varepsilon &\geq c/\varepsilon, && (n = 1) \end{aligned}$$

for ε sufficiently small. Again, c depends on f, g.

7.2 The invariant vector fields

Recall that the proof of the local existence theorem for nonlinear equations relied in part on the Sobolev inequality

$$|f(x)| \leq C_s \|f\|_{H^s}, \qquad s > \frac{n}{2},$$

which in particular implies

$$|f(x)| \leq C \sum_{|\alpha| \leq \frac{n+2}{2}} \|\partial^\alpha f\|_{L^2}. \tag{7.4}$$

To prove Theorem 28 we shall need a similar estimate for $|u(t, x)|$ in terms of L^2 norms in space of certain spacetime derivatives of u, multiplied by a decay factor in t. The spacetime derivatives involve the ***invariant vector fields:***

$$\partial_t, \partial_1, \dots, \partial_n, \tag{7.5}$$

$$\Omega_{ij} = x_j \partial_i - x_i \partial_j, \tag{7.6}$$

$$\Omega_{0i} = t\partial_i + x_i \partial_t, \tag{7.7}$$

$$L_0 = t\partial_t + \sum\nolimits_1^n x_i \partial_i, \tag{7.8}$$

where $1 \leq i, j \leq n$. To obtain symmetrical notation we sometimes write $t = x_0$ and $\partial_t = \partial_0$. Note that in (7.6) we can restrict to $1 \leq i < j \leq n$ by skew-symmetry. Thus we have a total of

$$n + 1 + \frac{n(n-1)}{2} + n + 1 = \frac{(n+1)(n+2)}{2} + 1$$

different vector fields, which we enumerate

$$\Gamma = (\Gamma_0, \dots, \Gamma_m), \qquad m = \frac{(n+1)(n+2)}{2}.$$

We use multi-index notation:

$$\Gamma^\alpha = \Gamma_0^{\alpha_0} \cdots \Gamma_m^{\alpha_m}, \qquad \alpha = (\alpha_0, \dots, \alpha_m).$$

The above vector fields are the generators of the transformations of the Minkowski space $\mathbb{R}^{1+n}$ which preserve the equation $\Box u = 0$. In fact, (7.5) correspond to translations in the coordinate directions; (7.6) correspond to rotations in the space variable x; (7.6) and (7.7) taken together correspond to a basis for the Lorentz transformations; finally, (7.8) corresponds to dilations.

Recall that the Lorentz transformations are the invertible linear transformations of $\mathbb{R}^{1+n}$ which are isometries with respect to the Lorentz metric

$$\mathrm{diag}(-1, 1, \ldots, 1).$$

One can then show that $\Box$ is the unique 2^{nd} order differential operator on $\mathbb{R}^{1+n}$ which commutes with all translations and Lorentz transformations.[1] Accordingly, for the vector fields (7.5), (7.6) and (7.7) we have, as one can also check directly,

$$[\Box, \partial_i] = 0, \qquad (0 \le i \le n) \tag{7.9}$$

$$[\Box, \Omega_{ij}] = 0, \qquad (0 \le i < j \le n) \tag{7.10}$$

where $[P, Q] = PQ - QP$. Although $\Box$ does not commute with dilations, the equation $\Box u = 0$ is certainly preserved. For the corresponding vector field (7.8) we have the simple commutation relation

$$[\Box, L_0] = 2\Box. \tag{7.11}$$

We shall also need the fact that for all i, j,

$$[\Gamma_i, \partial_j] = \sum_{k=0}^{n} a_{ijk} \partial_k, \tag{7.12}$$

as one can check by calculating the left hand side for each of the vector fields (7.6), (7.7) and (7.8).

7.3 The Klainerman-Sobolev inequality

We need the following replacement for (7.4).

Theorem 29. ***(Klainerman-Sobolev inequality.)*** *There is a constant C such that*

$$(1 + t + |x|)^{\frac{n-1}{2}} |u(t, x)| \le C \sum_{|\alpha| \le \frac{n+2}{2}} \|\Gamma^\alpha u(t, \cdot)\|_{L^2} \quad \text{for} \quad t > 0,\, x \in \mathbb{R}^n,$$

whenever $u \in C^\infty([0, \infty) \times \mathbb{R}^n)$ and $\mathrm{supp}\, u(t, \cdot)$ is compact for every t.

Clearly this implies the same estimate locally, that is, the estimate holds for $0 < t < T$, with the same constant C, if $u \in C^\infty([0, T) \times \mathbb{R}^n)$.

For the proof we shall require some lemmas which we now state.

[1] This parallels the fact that on $\mathbb{R}^n$, the unique 2^{nd} order operator commuting with all translations and rotations is the Laplacian Δ.

The first of these expresses the fact that at any point outside the light cone

$$\Lambda = \{(t,x) : |t| = |x|\},$$

the *homogeneous* vector fields (7.6), (7.7) and (7.8) span the tangent space of $\mathbb{R}^{1+n}$.

Lemma 5. *For any multi-index $\alpha \neq 0$ and $(t,x) \notin \Lambda$,*

$$\partial^\alpha = \sum_{1\le|\beta|\le|\alpha|} c_{\alpha\beta}(t,x)\Gamma^\beta,$$

where $c_{\alpha\beta}$ are C^∞ and homogeneous of degree $-|\alpha|$ outside the light cone Λ. In fact, the sum on the right hand side only involves the homogeneous vector fields.

Proof. We claim that

$$\left(t^2 - |x|^2\right)\partial_j = -\varepsilon_j x_j L_0 + \varepsilon_j \sum_{i=0}^n x_i \Omega_{ij}, \qquad \left(|x|^2 = \sum_1^n x_i^2\right)$$

where $\varepsilon_0 = -1$, $\varepsilon_1 = \cdots = \varepsilon_n = 1$ and by convention $\Omega_{0i} = \Omega_{i0}$. Recall also that $\Omega_{ij} = -\Omega_{ji}$ for $1 \le i,j \le n$. To prove the claim, note that when $j = 0$,

$$\begin{aligned}
\sum_1^n x_i \Omega_{i0} &= \sum_1^n \left(x_i^2 \partial_t + t x_i \partial_i\right) \\
&= |x|^2 \partial_t - t^2 \partial_t + t\left(t\partial_t + \sum_1^n x_i \partial_i\right) \\
&= \left(|x|^2 - t^2\right)\partial_t + tL_0,
\end{aligned}$$

while for $1 \le j \le n$,

$$\begin{aligned}
\sum_0^n x_i \Omega_{ij} &= t^2 \partial_j + t x_j \partial_t + \sum_1^n \left(x_i x_j \partial_i - x_i^2 \partial_j\right) \\
&= t^2 \partial_j + x_j L_0 - |x|^2 \partial_j.
\end{aligned}$$

Thus we have the result for $|\alpha| = 1$, and the general case follows by induction; the key observation is that if $a(t,x)$ is C^∞ and homogeneous of degree $-k$ outside Λ, then so are $L_0 a$ and $\Omega_{ij} a$ for $0 \le i < j \le n$. □

The next result is just a localized Sobolev inequality.

Lemma 6. *Given $\delta > 0$, there is a constant C_δ such that*

$$|f(0)|^2 \le C_\delta \sum_{|\alpha| \le \frac{n+2}{2}} \int_{|y|<\delta} |\partial^\alpha f(y)|^2 \, dy$$

for all $f \in C^\infty(\mathbb{R}^n)$. Moreover,

$$\sup_{\delta \ge \delta_0} C_\delta < \infty$$

for every $\delta_0 > 0$.

Proof. Fix a cutoff $\chi \in C_c^\infty(\mathbb{R}^n)$ which equals 1 in the unit ball at the origin. Applying (7.4) to the function

$$\chi(y/\delta)f(y),$$

yields the desired inequality with $C_\delta \leq C(1+\delta^{-n-2})$, and the final statement of the Lemma is then evident. □

Finally, we need a Sobolev inequality for a smooth function $v(q,\omega)$ where $q \in \mathbb{R}$ and $\omega \in S^{n-1}$. Observe that the vector fields Ω_{ij}, $1 \leq i < j \leq n$ can be regarded as vector fields on the sphere $S^{n-1} = \{x \in \mathbb{R}^n : |x| = 1\}$. Accordingly, we write

$$\partial_\omega^\alpha = \Omega_{12}^{\alpha_1} \cdots \Omega_{n-1,n}^{\alpha_m}; \quad \alpha = (\alpha_1, \ldots, \alpha_m), \quad m = n(n-1)/2,$$

at any point $\omega \in S^{n-1}$.

Lemma 7. *Given $\delta > 0$, there is a constant C_δ such that*

$$|v(q,\omega)|^2 \leq C_\delta \sum_{j+|\alpha| \leq \frac{n+2}{2}} \int_{|p|<\delta} \int_{\eta \in S^{n-1}} \left|\partial_q^j \partial_\omega^\alpha v(q+p,\eta)\right|^2 \, d\sigma(\eta)\, dp$$

for all $v \in C^\infty(\mathbb{R} \times S^{n-1})$. Moreover,

$$\sup_{\delta \geq \delta_0} C_\delta < \infty$$

for every $\delta_0 > 0$.

Proof. This follows from Lemma 6 if we cover the sphere S^{n-1} by finitely many coordinate charts and choose a subordinate partition of unity, the key observation being that the vector fields

$$\Omega_{12}, \ldots, \Omega_{n-1,n}$$

span $T_\omega S^{n-1}$ for all $\omega \in S^{n-1}$. □

7.3.1 Proof of the Klainerman-Sobolev inequality

If $t + |x| \leq 1$, the inequality follows from the standard Sobolev inequality (7.4), so in what follows we assume $t + |x| > 1$. We use different arguments depending on whether (t,x) is close to the light cone or not.

Case 1. Assume

$$|x| \leq \frac{t}{2} \quad \text{or} \quad |x| \geq \frac{3t}{2} \tag{7.13}$$

and of course

$$R = t + |x| > 1, \tag{7.14}$$

which defines R. We then claim that the coefficents in Lemma 5 satisfy

$$|y| \le \frac{R}{8} \implies |c_{\alpha\beta}(t, x+y)| \le CR^{-|\alpha|} \quad \text{for} \quad 1 \le |\beta| \le |\alpha| \le \frac{n+2}{2}. \tag{7.15}$$

Assuming this, it follows by applying Lemma 6 to the function $z \to u(t, x+Rz)$ that

$$\begin{aligned} |u(t,x)|^2 &\le C \sum_{|\alpha| \le \frac{n+2}{2}} \int_{|z| \le \frac{1}{8}} \left| R^{|\alpha|} \partial_x^\alpha u(t, x+Rz) \right|^2 dz \\ &= CR^{-n} \sum_{|\alpha| \le \frac{n+2}{2}} \int_{|y| \le \frac{R}{8}} \left| R^{|\alpha|} \partial_x^\alpha u(t, x+y) \right|^2 dy, \end{aligned}$$

where we changed variables to $y = Rz$. Since, by Lemma 5,

$$\alpha \ne 0 \implies \partial_x^\alpha u(t, x+y) = \sum_{1 \le |\beta| \le |\alpha|} c_{\alpha\beta}(t, x+y)(\Gamma^\beta u)(t, x+y),$$

we conclude, using (7.15), that

$$R^n |u(t,x)|^2 \le C \sum_{|\alpha| \le \frac{n+2}{2}} \|\Gamma^\alpha u(t, \cdot)\|_{L^2}^2,$$

which proves the inequality in the region (7.13), (7.14).

It only remains to prove the claimed property (7.15). In fact, it is easy to check, using (7.13), that

$$|y| \le \frac{R}{8} \implies \big| t - |x+y| \big| \ge cR \quad \text{for some } c > 0.$$

Moreover,

$$\left| \frac{t + |x+y|}{R} - 1 \right| \le \frac{1}{8},$$

so the point $(t/R, (x+y)/R)$ is in a compact set disjoint from the light cone Λ. Thus (7.15) follows by continuity and homogeneity of $c_{\alpha\beta}$.

Case 2. Assume

$$\frac{t}{2} \le |x| \le \frac{3t}{2} \tag{7.16}$$

as well as $t + |x| > 1$. Introduce polar coordinates

$$x = r\omega \quad \text{where} \quad r > 0,\, \omega \in S^{n-1}.$$

Then write

$$u(t,x) = u(t, r\omega) = v(t, q, \omega),$$

where

$$q = r - t.$$

In other words, we define v by

$$v(t, q, \omega) = u\big(t, (t+q)\omega\big).$$

Observe that

$$\partial_q v = \sum_1^n \omega_j \partial_j u = \partial_r u, \tag{7.17}$$

and that

$$\Omega_{ij} v = \Omega_{ij} u, \qquad (1 \le i < j \le n) \tag{7.18}$$

where on the left hand side we consider Ω_{ij} to act in ω, as a vector field on S^{n-1}.

Thus, if we consider a point (t, x) satisfying (7.16), and apply Lemma 7 to $v(t, q, \omega)$, we obtain

$$\begin{aligned} |u(t,x)|^2 &= |v(t,q,\omega)|^2 \\ &\le C_t \sum_{j+|\alpha| \le \frac{n+2}{2}} \int_{|p|<\frac{t}{4}} \int_{\eta \in S^{n-1}} \left|\partial_q^j \partial_\omega^\alpha v(t, q+p, \eta)\right|^2 \, d\sigma(\eta)\, dp \\ &\le C_t \sum_{j+|\alpha| \le \frac{n+2}{2}} \int_{|p|<\frac{t}{4}} \int_{\eta \in S^{n-1}} \left|\partial_r^j \Gamma^\alpha u\big(t, (t+q+p)\eta\big)\right|^2 \, d\sigma(\eta)\, dp, \end{aligned} \tag{7.19}$$

where we used (7.17) and (7.18) in the last step. Let us remark at this point that $C_t \le C$ where C is independent of t. In fact,

$$|x| \le \frac{3t}{2} \implies 1 < t + |x| \le \frac{5t}{2} \implies t > \frac{2}{5},$$

so that t is bounded away from 0. Thus C_t is bounded above in view of the last statement in Lemma 7.

Next observe that by (7.16), $|q| = |r - t| \le t/2$, and $|p| \le t/4$ then implies

$$\frac{t}{4} \le t + p + q \le 2t.$$

Therefore, by changing variables to $r = t + p + q$ in (7.19), and noting that

$$\partial_r u = \sum_1^n \eta_i \partial_i u,$$

we conclude that

$$\begin{aligned}
|u(t,x)|^2 &\le C \sum_{|\beta|\le\frac{n+2}{2}} \int_{\frac{t}{4}\le r\le 2t} \int_{\eta\in S^{n-1}} \left|\Gamma^\beta u(t,r\eta)\right|^2 \, d\sigma(\eta)\, dr \\
&\le Ct^{1-n} \sum_{|\beta|\le\frac{n+2}{2}} \int_0^\infty \int_{\eta\in S^{n-1}} \left|\Gamma^\beta u(t,r\eta)\right|^2 \, d\sigma(\eta) r^{n-1}\, dr \\
&= Ct^{1-n} \sum_{|\beta|\le\frac{n+2}{2}} \left\|\Gamma^\beta u(t,\cdot)\right\|_{L^2}^2 .
\end{aligned}$$

Since $t \approx |x|$ and $t + |x| > 1$, this proves the Klainerman-Sobolev inequality in the region (7.16), thus finishing the proof of Theorem 29.

7.4 Proof of the main theorem

We begin by making some observations to aid us in the proof.

Observation 1. Since F vanishes to second order at the origin [cf. (7.3)],

$$|F(z)| \le G_2(|z|)\,|z|^2 , \tag{7.20}$$

$$|DF(z)| \le G_2(|z|)\,|z| , \tag{7.21}$$

$$|D^m F(z)| \le G_m(|z|), \qquad (m \ge 2) \tag{7.22}$$

for all $z \in \mathbb{R}^{1+n}$, where $G_2, G_3, \dots$ are continuous, increasing functions and $D^m F$ stands for any $\partial^\alpha F$ with $|\alpha| = m$.

Proof. Set

$$G_m(r) = \sup_{|z|\le r} |D^m F(z)| .$$

Then G_m is continuous and increasing, and (7.22) holds. Now write

$$\partial_j F(z) = \partial_j F(z) - \partial_j F(0) = \int_0^1 \frac{d}{dt}\left[\partial_j F(tz)\right] dt = \left[\int_0^1 \nabla\partial_j F(tz)\, dt\right] \cdot z.$$

Taking absolute values, we get (7.21). Applying the same argument to $F(z)$ then gives

$$|F(z)| \le \sup_{0\le t\le 1} |DF(tz)|\,|z| \le G_2(|z|)\,|z|^2 .$$

□

Observation 2. $\Box\Gamma^\alpha = \sum_{|\beta|\le|\alpha|} c_{\alpha\beta}\Gamma^\beta\Box$ where $c_{\alpha\beta}$ are constants.

Proof. This follows from (7.9), (7.10) and (7.11). □

Observation 3. For $\alpha \neq 0$, $\Gamma^\alpha[F(\partial u)]$ is a linear combination of terms

$$[D^m F](\partial u)\Gamma^{\beta_1}\partial u \cdots \Gamma^{\beta_m}\partial u \quad \text{where} \quad 1 \leq m \leq |\alpha|\,, \quad \sum_1^m |\beta_i| = |\alpha|\,. \tag{7.23}$$

Proof. This is a simple induction. □

Observation 4. In (7.23), at most one β_i can have order $|\beta_i| > |\alpha|/2$. Let us order the β_i so that

$$|\beta_m| = \max_{1\leq i\leq m} |\beta_i|\,. \tag{7.24}$$

Then we have $|\beta_i| \leq |\alpha|/2$ for $1 \leq i \leq m-1$.

Observation 5. Let $N = n+4$. If $|\alpha| \leq N$ and $|\beta_j| \leq |\alpha|/2$, then

$$|\beta_j| + 1 + \frac{n+2}{2} \leq N.$$

Proof. $N/2 + 1 + (n+2)/2 \leq N$ iff $n+4 \leq N$. □

We now turn to the proof of Theorem 28.

Step 1. Some initial reductions. Set $N = n+4$. Define

$$A(t) = \sum_{|\alpha|\leq N} \|\Gamma^\alpha \partial u(t,\cdot)\|_{L^2}\,, \quad 0 \leq t < T$$

whenever $u \in C^\infty([0,T)\times\mathbb{R}^n)$ solves (7.1), (7.2) on $[0,T)\times\mathbb{R}^n$ for some $T > 0$. Observe that by (7.2),

$$A(0) \leq \frac{A\varepsilon}{2}, \tag{7.25}$$

where A depends only on f and g (and their derivatives).

Claim. *There exists $\varepsilon_0 > 0$ such that if $T > 0$ and $u \in C^\infty([0,T)\times\mathbb{R}^n)$ solves (7.1), (7.2) on $[0,T)\times\mathbb{R}^n$ with $\varepsilon \leq \varepsilon_0$, then $A(t) \leq A\varepsilon$ for all $0 \leq t < T$.*

Observe that by the Sobolev inequality (7.4),

$$\|\partial u\|_{L^\infty([0,T)\times\mathbb{R}^n)} \leq C \sup_{0\leq t<T} A(t).$$

Therefore, if the claim holds, it follows from Theorem 6, Week 5, that the lifespan $T_\varepsilon = \infty$ when $\varepsilon \leq \varepsilon_0$, and Theorem 28 will then be proved.

Step 2. Further reductions. To prove the claim we set

$$E = \{t \in [0, T) : A(s) \leq A\varepsilon \text{ for all } 0 \leq s \leq t\}.$$

By (7.25), E is nonempty. Since $A(t)$ is continuous in t, E is relatively closed in $[0, T)$. Thus, if we can show that E is also relatively open in $[0, T)$, it will follow that $E = [0, T)$, and the claim is then proved.

To prove that E is open, we fix $t_0 \in E$ with $t_0 < T$. Since $A(t)$ is continuous, there exists $t_1 > t_0$ such that

$$A(t) \leq 2A\varepsilon \quad \text{for} \quad 0 \leq t \leq t_1. \tag{7.26}$$

We shall prove that this implies

$$A(t) \leq A\varepsilon \quad \text{for} \quad 0 \leq t \leq t_1, \tag{7.27}$$

if ε is sufficiently small. It suffices to prove that

$$A(t) \leq A\varepsilon/2 + C_A\varepsilon \int_0^t \frac{A(s)}{(1+s)^{(n-1)/2}}\, ds. \tag{7.28}$$

For then it follows by Gronwall's Lemma that

$$A(t) \leq \frac{A\varepsilon}{2} \exp\left[C_A\varepsilon \int_0^t \frac{ds}{(1+s)^{(n-1)/2}}\right], \tag{7.29}$$

and since $\int_0^\infty \frac{ds}{(1+s)^{(n-1)/2}} < \infty$ when $n \geq 4$, we only have to choose $\varepsilon > 0$ so small that

$$\exp\left[C_A\varepsilon \int_0^\infty \frac{ds}{(1+s)^{(n-1)/2}}\right] \leq 2,$$

and the proof is complete.

Step 3. Proof of (7.28)**.** In view of (7.12), we can apply the energy inequality (10.13), obtaining

$$\begin{aligned} A(t) &\leq A(0) + C_N \int_0^t \sum_{|\alpha| \leq N} \|\Box\Gamma^\alpha u(s, \cdot)\|_{L^2}\, ds \\ &\leq A\varepsilon/2 + C_N \int_0^t \sum_{|\alpha| \leq N} \|\Gamma^\alpha \Box u(s, \cdot)\|_{L^2}\, ds \\ &= A\varepsilon/2 + C_N \int_0^t \sum_{|\alpha| \leq N} \|\Gamma^\alpha [F(\partial u)](s, \cdot)\|_{L^2}\, ds, \end{aligned}$$

where we used (7.25) and Observation 2 to get the next to last inequality. Here C_N denotes a generic constant which can change from line to line. We now estimate

$$\|\Gamma^\alpha [F(\partial u)](t, \cdot)\|_{L^2}, \qquad |\alpha| \leq N. \tag{7.30}$$

If $\alpha = 0$, we use (7.20) to get

$$\|F(\partial u)(t,\cdot)\|_{L^2} \le G_2\left(\|\partial u(t,\cdot)\|_{L^\infty}\right)\|\partial u(t,\cdot)\|_{L^\infty}\|\partial u(t,\cdot)\|_{L^2}\,. \tag{7.31}$$

The first factor on the right hand side is bounded by a continuous function of A, since

$$\|\partial u(t,\cdot)\|_{L^\infty} \le CA(t) \le 2CA\varepsilon \tag{7.32}$$

by the Sobolev inequality and (7.26), and since we of course can assume that $\varepsilon \le 1$. By the Klainerman-Sobolev inequality, the second factor in (7.31) is bounded by

$$C\frac{A(t)}{(1+t)^{(n-1)/2}},$$

and by (7.26), the third factor in (7.31) is bounded by $2A\varepsilon$.

If $\alpha \ne 0$, we use Observation 3 to write $\Gamma^\alpha u(t,\cdot)$ as a sum of terms of the form (7.23), whose L^2 norms in space we bound by

$$\left\|[D^m F]\big(\partial u(t,\cdot)\big)\right\|_{L^\infty}\prod_{i=1}^{m-1}\left\|\Gamma^{\beta_i}\partial u(t,\cdot)\right\|_{L^\infty}\left\|\Gamma^{\beta_m}\partial u(t,\cdot)\right\|_{L^2}. \tag{7.33}$$

Let us first consider the case $m \ge 2$. Then the first factor is bounded by a continuous function of A, in view of (7.22) and (7.32). Using (7.26) and the fact that $|\beta_m| \le |\alpha| \le N$, we see that the last factor is bounded by $2A\varepsilon$. Since we may assume that (7.24) holds, we have $|\beta_i| + 1 + (n+2)/2 \le N$ for $1 \le i \le m-1$, in view of Observations 4 and 5. Hence the Klainerman-Sobolev inequality and (7.26) imply

$$\prod_{i=1}^{m-1}\left\|\Gamma^{\beta_i}\partial u(t,\cdot)\right\|_{L^\infty} \le C\left[\frac{A(t)}{(1+t)^{(n-1)/2}}\right]^{m-1} \le CA^{m-2}\frac{A(t)}{(1+t)^{(n-1)/2}}.$$

We conclude that (7.33) is bounded by

$$C_A\varepsilon A(t)(1+t)^{-(n-1)/2}$$

when $m \ge 2$. When $m = 1$ we get the same bound if we use (7.21) instead of (7.22). This completes the proof.

Chapter 8

Week 8: Low dimensions

We have proved existence of global smooth solutions in space dimensions $n \geq 4$ of the nonlinear Cauchy problem, on $\mathbb{R}^{1+n}$,

$$\Box u = F(\partial u), \tag{8.1}$$

$$u\big|_{t=0} = \varepsilon f, \quad \partial_t u\big|_{t=0} = \varepsilon g, \tag{8.2}$$

for $\varepsilon > 0$ sufficiently small. Here $F : \mathbb{R}^{1+n} \to \mathbb{R}$ is a given C^∞ function which vanishes to second order at the origin:

$$F(0) = 0, \quad DF(0) = 0. \tag{8.3}$$

Next we want to see what happens in dimensions $n = 1, 2, 3$. Then global existence fails in general (we will give an example later on in the course), but the proof used for $n \geq 4$ gives asymptotic lower bounds on the lifespan[1]

$$T_\varepsilon = T^*(\varepsilon f, \varepsilon g)$$

as $\varepsilon \to 0$. Specifically, we shall see that there exists $c > 0$ such that

$$T_\varepsilon \geq \begin{cases} e^{c/\varepsilon}, & \text{if } n = 3, \\ c/\varepsilon^2, & \text{if } n = 2, \\ c/\varepsilon, & \text{if } n = 1, \end{cases} \tag{8.4}$$

for ε sufficiently small. Again, c depends on f, g.

8.1 Proof of the lower bounds for T_ε

Suppose $u \in C^\infty([0,T) \times \mathbb{R}^n)$ solves (8.1), (8.2) on $[0,T) \times \mathbb{R}^n$ for some $T > 0$. Recall from the proof of global existence for $n \geq 4$ that if we set $N = n + 4$ and

$$A(t) = \sum_{|\alpha| \leq N} \|\Gamma^\alpha \partial u(t, \cdot)\|_{L^2}, \quad 0 \leq t < T,$$

[1] Recall that the lifespan is the supremum of $T > 0$ such that (8.1), (8.2) has a solution $u \in C^\infty([0,T] \times \mathbb{R}^n)$. By uniqueness, the totality of such solutions assemble to a smooth solution on $[0, T_\varepsilon) \times \mathbb{R}^n$.

then there is a constant $A = A(f,g)$ such that the *boot-strap* assumption

$$A(t) \leq 2A\varepsilon, \quad 0 \leq t \leq T' < T$$

implies

$$A(t) \leq A\varepsilon/2 + C_A\varepsilon \int_0^t \frac{A(s)}{(1+s)^{(n-1)/2}}\, ds, \quad 0 \leq t \leq T'. \tag{8.5}$$

Then by Gronwall's Lemma,

$$A(t) \leq \frac{A\varepsilon}{2} \exp\left[C_A\varepsilon \int_0^t \frac{ds}{(1+s)^{(n-1)/2}}\right], \quad 0 \leq t \leq T'. \tag{8.6}$$

When $n \geq 4$, $\int_0^\infty \frac{ds}{(1+s)^{(n-1)/2}} < \infty$, and so we obtain the stronger bound $A(t) \leq A\varepsilon$ on $[0,T']$, provided that $0 < \varepsilon < \varepsilon_0$, where ε_0 is determined by the condition

$$\exp\left[C_A\varepsilon_0 \int_0^\infty \frac{ds}{(1+s)^{(n-1)/2}}\right] = 2.$$

[So ε_0 depends on A, hence on (f,g), but not on T.] Combining this boot-strap argument with the continuity method then gives the *a priori* bound

$$A(t) \leq A\varepsilon, \qquad 0 \leq t < T, \tag{8.7}$$

Once we have this *a priori* bound, we can control $\|\partial u\|_{L^\infty([0,T)\times\mathbb{R}^n)}$ using Sobolev's Lemma. Since $T > 0$ was arbitrary, it follows from the local existence theory of week 5 (see Theorem 6), that $T_\varepsilon = \infty$.

When $n = 1, 2$ or 3, the function $(1+s)^{-(n-1)/2}$ is no longer integrable at infinity, but we still get the bound (8.7), provided ε is sufficiently small and T satisfies

$$T \leq \begin{cases} e^{c/\varepsilon}, & \text{if } n = 3, \\ c/\varepsilon^2, & \text{if } n = 2, \\ c/\varepsilon, & \text{if } n = 1. \end{cases} \tag{8.8}$$

From this we get the statement (8.4) about the lifespan, reasoning as above.

We consider the cases $n = 1, 2, 3$ one by one.

Case 1: $n = 3$. $\int_0^t \frac{ds}{(1+s)} = \log(1+t)$, so (8.6) becomes

$$A(t) \leq \frac{A\varepsilon}{2}(1+t)^{C_A\varepsilon},$$

which implies (8.7) if $(1+T)^{C_A\varepsilon} \leq 2$. This is true for $T < 1$ if $C_A\varepsilon \leq 1$. If on the other hand $T \geq 1$, and we assume $C_A\varepsilon \leq \frac{1}{2}$, then

$$(1+T)^{C_A\varepsilon} \leq \sqrt{2}T^{C_A\varepsilon} \leq 2$$

provided $T^{C_A\varepsilon} \leq \sqrt{2}$, that is, $T \leq 2^{1/2C_A\varepsilon}$.

Case 2: $n = 2$. $\int_0^t \frac{ds}{(1+s)^{1/2}} = 2\sqrt{1+t} - 2 \le C\sqrt{t}$, so (8.6) gives

$$A(t) \le \frac{A\varepsilon}{2} e^{C_A \varepsilon \sqrt{t}},$$

and (8.7) follows if

$$e^{C_A \varepsilon \sqrt{T}} \le 2 \iff \sqrt{T} \le \frac{\log 2}{C_A \varepsilon}.$$

Case 3: $n = 1$. $\int_0^t ds = t$, so (8.6) gives

$$A(t) \le \frac{A\varepsilon}{2} e^{C_A \varepsilon t},$$

and (8.7) follows if

$$e^{C_A \varepsilon T} \le 2 \iff T \le \frac{\log 2}{C_A \varepsilon}.$$

Remark. If F vanishes to higher order than 2 at 0, we can get global results also for $n = 2, 3$ by obvious modifications of the proof of the main theorem from week 7. In fact, if F vanishes to third order at 0, then referring to the notes of week 6, the estimates in Observation 1 are improved to

$$\begin{aligned} |F(z)| &\le G_3(|z|)\,|z|^3, \\ |DF(z)| &\le G_3(|z|)\,|z|^2, \\ \left|D^2 F(z)\right| &\le G_3(|z|)\,|z|, \\ |D^m F(z)| &\le G_3(|z|), \qquad m \ge 3, \end{aligned}$$

and so in Step 3 of the proof we see that we always get one extra power of $\|\partial u(t,\cdot)\|_{L^\infty}$. After estimating these L^∞ norms using the Klainerman-Sobolev inequality, we will then have the integrand

$$\left[\frac{1}{(1+t)^{(n-1)/2}}\right]^2 = \frac{1}{(1+t)^{n-1}} \tag{8.9}$$

in equation (30) of week 5, instead of $(1+t)^{-(n-1)/2}$. Since (8.9) is integrable when $n = 3$, we get global existence. Similarly, if F vanishes to fourth order at 0, then we get the integrand

$$\left[\frac{1}{(1+t)^{(n-1)/2}}\right]^3.$$

Thus the integral converges for $n = 2$, so we get a global result also in this case.

8.2 The null condition and global existence for space dimension $n = 3$

In general, existence of global smooth solutions for small data *fails* in dimension $n = 3$, for equations of the type $\Box u = F(\partial u)$.

Example. F. John proved that every smooth solution of

$$\Box u = (\partial_t u)^2, \qquad t \ge 0,\ x \in \mathbb{R}^3 \tag{8.10}$$

with nonzero data in $C_c^\infty(\mathbb{R}^3)$ blows up in finite time.

Example. (Due to Nirenberg.) In sharp contrast to the previous example, for the superficially similar equation

$$\Box u = (\partial_t u)^2 - \sum_1^3 (\partial_j u)^2, \qquad t \ge 0,\ x \in \mathbb{R}^3 \tag{8.11}$$

we have global smooth solutions for small data:

$$u\big|_{t=0} = \varepsilon f, \quad \partial_t u\big|_{t=0} = \varepsilon g, \tag{8.12}$$

where $f, g \in C_c^\infty(\mathbb{R}^3)$ and $\varepsilon > 0$ is sufficiently small. The key observation is that if we set

$$v(t,x) = 1 - e^{-u(t,x)},$$

then v solves the *linear* Cauchy problem

$$\Box v = 0, \quad v\big|_{t=0} = 1 - e^{-\varepsilon f}, \quad \partial_t v\big|_{t=0} = \varepsilon g e^{-\varepsilon f}, \tag{8.13}$$

which of course has a global smooth solution. The inverse of the transformation $u \to v$ is

$$u(t,x) = -\log[1 - v(t,x)].$$

This is well-defined as long as $|v| < 1$, and u then solves (8.11), (8.12). To ensure that v is *globally* small,

$$\|v(t,\cdot)\|_{L^\infty} < 1 \quad \text{for all} \quad t \ge 0,$$

we only have to take $\varepsilon > 0$ sufficiently small, depending on f and g. Indeed, recall from week 2 that when $n = 3$ we have the decay estimate

$$\|v(t,\cdot)\|_{L^\infty} \le \frac{A}{1+t} \quad \text{for all} \quad t \ge 0,$$

where A is a constant which depends *linearly* on the L^∞ norms of $v\big|_{t=0}$, $\nabla_x v\big|_{t=0}$ and $\partial_t v\big|_{t=0}$. In view of (8.13), therefore, $A < 1$ if $\varepsilon > 0$ is sufficiently small. Then the transformation $v \to u$ is globally defined, giving a global smooth solution of (8.11),(8.12).

These examples suggest that in dimension $n = 3$, the question of global existence of smooth solutions of systems of the type $\Box u = B(\partial u, \partial u)$, where each vector component of B is a bilinear form in ∂u, depends strongly on the algebraic structure of B. More generally, for a system of the form $\Box u = F(\partial u)$, where F vanishes to second order at the origin, it is the *quadratic part* of F that determines the global regularity properties of the equation. The higher order terms are not important. [Recall from the remark at the end of the previous section that we always have global existence for nonlinearities $F(\partial u)$ which vanish to third order at 0.]

8.2.1 Statement of null condition

We now consider a system of N equations

$$\Box u^I = F^I(u, \partial u), \qquad (t, x) \in \mathbb{R}^{1+3}, \tag{8.14}$$

where the unknown u and the given C^∞ function F are $\mathbb{R}^N$-valued:

$$u = (u^1, \ldots, u^N), \quad F = (F^1, \ldots, F^N).$$

Definition. A vector $\xi = (\xi_0, \ldots, \xi_4) \in \mathbb{R}^{1+3}$ is ***null*** if $\xi \neq 0$ and $\xi_0^2 = \xi_1^2 + \xi_2^2 + \xi_3^2$. In other words, ξ lies on the light cone (or ***null cone***) in Minkowski space $\mathbb{R}^{1+3}$.

Definition. The ***quadratic part*** of F^I is

$$F^I_{(2)}(z) = \sum_{|\alpha|=2} \frac{1}{\alpha!} \partial^\alpha F^I(0) z^\alpha,$$

where $z \in \mathbb{R}^{N+(n+1)N}$ corresponds to $(u, \partial u)$.

As motivated above, we have to impose some condition on the quadratic part of F in order to ensure the existence of global smooth solutions of (8.14) for small data. The relevant principle is the so-called null condition of Klainerman and Christodoulou.

Definition. F in (8.14) satisfies the ***null condition*** if:

(i) F vanishes to second order at the origin:

$$F(0) = 0, \quad DF(0) = 0.$$

Thus, by Taylor's theorem, $F(z) = F_{(2)}(z) + R(z)$, where R is C^∞ and vanishes to third order at 0.

(ii) The quadratic part of F is of the form

$$F^I_{(2)}(u, \partial u) = \sum_{J,K=1}^{N} \sum_{\mu,\nu=0}^{3} a^{I\mu\nu}_{JK} \partial_\mu u^J \partial_\nu u^K,$$

where the a's are real constants satisfying, for all $I, J, K = 1, \dots, N$,

$$\sum_{\mu,\nu=0}^{3} a_{JK}^{I\mu\nu} \xi_\mu \xi_\nu = 0 \quad \text{for all null vectors } \xi.$$

Observe that $F_{(2)}$ is only allowed to depend on ∂u, not on u.

Example. The equation (8.11) satisfies the null condition, while (8.10) does not.

Lemma 8. *If B is a real bilinear form on $\mathbb{R}^4 \times \mathbb{R}^4$ such that*

$$B(\xi, \xi) = 0 \quad \textit{for all null vectors } \xi,$$

*then B is a linear combination, with real coefficients, of the so-called **null forms***

$$Q_0(\xi, \eta) = \xi_0 \eta_0 - \sum_1^3 \xi_i \eta_i, \tag{8.15a}$$

$$Q_{\mu\nu}(\xi, \eta) = \xi_\mu \eta_\nu - \xi_\nu \eta_\mu, \quad 0 \le \mu < \nu \le 3. \tag{8.15b}$$

Proof. $B(\xi, \xi) = \xi^T A \xi = \sum a^{\mu\nu} \xi_\mu \xi_\nu$, where $A = (a^{\mu\nu})$ is a real 4×4 matrix and we consider ξ as a column vector with transpose $\xi^T = (\xi_0, \dots, \xi_3)$. Now decompose A into its symmetric and skew-symmetric parts:

$$A = \frac{A + A^T}{2} + \frac{A - A^T}{2} = A_s + A_a.$$

Since $\xi^T A \xi = (\xi^T A \xi)^T = \xi^T A^T \xi$, we see that

$$\xi^T A_s \xi = \xi^T A \xi = 0$$

for null ξ. Using this condition with the null vectors

$$\xi^T = (\pm 1, 1, 0, 0), \quad (\pm 1, 0, 1, 0), \quad (\pm 1, 0, 0, 1),$$

and then with

$$\xi^T = (\sqrt{2}, 1, 1, 0), \quad (\sqrt{2}, 1, 0, 1), \quad (\sqrt{2}, 0, 1, 1),$$

it is not hard to see that A_s must be of the form

$$A_s = a^{00} \operatorname{diag}(1, -1, -1, -1),$$

which of course corresponds to Q_0. On the other hand, the skew-symmetric part A_a gives a combination of the $Q_{\mu\nu}$ in an obvious way. Summing up, we have $B = a^{00} Q_0 + \sum_{0 \le \mu < \nu \le 3} \frac{1}{2}(a^{\mu\nu} - a^{\nu\mu}) Q_{\mu\nu}$. □

That the converse of the above lemma holds is obvious.

Corollary. *F in (8.14) satisfies the null condition iff each component $F^I(u,\partial u)$ is of the form*

$$\sum_{J,K} a^I_{JK} Q_0(\partial u^J, \partial u^K) + \sum_{J,K} \sum_{0\le\mu<\nu\le 3} b^{I\mu\nu}_{JK} Q_{\mu\nu}(\partial u^J, \partial u^K) + R^I(u,\partial u),$$

where the a's and b's are real constants and R^I is C^∞ and vanishes to third order at 0.

We can now state the main theorem.

Consider the system (8.14) with initial data

$$u\big|_{t=0} = \varepsilon f, \qquad \partial_t u\big|_{t=0} = \varepsilon g, \tag{8.16}$$

where $f=(f^1,\dots,f^N)$ and $g=(g^1,\dots,g^N)$ belong to $C_c^\infty(\mathbb{R}^3)$ and $\varepsilon>0$.

Theorem 30. *Assume that F in (8.14) satisfies the null condition. Then there exists $\varepsilon_0=\varepsilon_0(f,g)>0$ such that (8.14),(8.16) has a smooth global solution provided $\varepsilon<\varepsilon_0$.*

8.2.2 Improved decay

The next lemma is of key importance. It quantifies the fact that the null forms have better decay properties, due to cancellations, than generic bilinear forms. To state this result, we need some more notation.

For the invariant vector fields $\Gamma_0,\dots,\Gamma_m$, let $\Gamma_j(t,x;\xi)$ denote the *symbol* of Γ_j, obtained by replacing ∂u by the vector $\xi\in\mathbb{R}^4$. Thus, the symbol of ∂_μ is just ξ_μ, while

$$\begin{aligned}
\Omega_{0j}(t,x;\xi) &= t\xi_j + x_j\xi_0,\\
\Omega_{ij}(t,x;\xi) &= x_j\xi_i - x_i\xi_j, \quad (1\le i<j\le 3)\\
L(t,x;\xi) &= t\xi_0 + \sum_1^3 x_i\xi_i.
\end{aligned}$$

We denote by $\Gamma(t,x;\xi)$ the vector $\big(\Gamma_0(t,x;\xi),\dots,\Gamma_m(t,x;\xi)\big)$. Thus,

$$|\Gamma(t,x;\xi)|^2 = \sum |\Gamma_j(t,x;\xi)|^2 .$$

Lemma 9. *Let B be a bilinear form on $\mathbb{R}^4$. Then there exists a constant C such that*

$$|B(\xi,\eta)| \le \frac{C}{1+|t|+|x|}\,|\Gamma(t,x;\xi)|\,|\Gamma(t,x;\eta)| \tag{8.17}$$

for all $(t,x),\xi,\eta\in\mathbb{R}^{1+3}$, if and only if B satisfies

$$B(\xi,\xi)=0 \quad \textit{for all null vectors } \xi. \tag{8.18}$$

The proof actually shows that if (8.18) is not satisfied, then there is no estimate

$$|B(\xi,\eta)| \le c(t)\,|\Gamma(t,x;\xi)|\,|\Gamma(t,x;\eta)|$$

where $c(t)\to 0$ as $t\to\infty$. Thus, the best one can say is that the trivial estimate $|B(\xi,\eta)|\le C\,|\xi|\,|\eta|$ holds.

Corollary. *For each of the null forms Q defined in* (8.15a)*,*

$$|Q(\partial v,\partial w)| \le \frac{C}{1+|t|+|x|}\,|\Gamma v(t,x)|\,|\Gamma w(t,x)|$$

for all (t,x) and all smooth functions v,w. This estimate does not hold for any other bilinear form Q.

Proof of Lemma 9. We first prove that (8.17) implies (8.18). Fix a null vector ξ. Set $(t,x)=\lambda(\xi_0,-\xi_1,-\xi_2,-\xi_3)$, where $\lambda>0$. Then all the homogeneous symbols vanish at $(t,x;\xi)$. In fact,

$$L_0(t,x;\xi)=\lambda(\xi_0^2-\xi_1^2-\xi_2^2-\xi_3^2)=0$$

since ξ is null, and it is also easy to check that $\Omega_{\mu\nu}(t,x;\xi)=0$ for $0\le\mu<\nu\le 3$. Thus, $|\Gamma(t,x;\xi)|=|\xi|$, and since $|(t,x)|=\lambda\,|\xi|$, it follows from (8.17) that $|B(\xi,\xi)|=O(1/\lambda)$ as $\lambda\to\infty$, so (8.18) holds.

Conversely, assume (8.18) holds. Then by Lemma 8, B is a linear combination of the null forms Q_0 and $Q_{\mu\nu}$, so it suffices to verify (8.17) for these. Since (8.17) holds trivially when $|t|+|x|\le 1$, we will assume $|t|+|x|>1$.

First consider Q_{ij} with $1\le i<j\le 3$. If $|t|\ge|x|$, then (8.17) follows from the identity

$$Q_{ij}(\xi,\eta)=\frac{1}{t}\left[\xi_0\Omega_{ij}(t,x;\eta)+\Omega_{0i}(t,x;\xi)\eta_j-\Omega_{0j}(t,x;\xi)\eta_i\right].$$

If on the other hand $|t|\le|x|$, then we use the identity

$$\xi_i=\sum_{k=1}^{3}\frac{x_ix_k}{|x|^2}\xi_k+\sum_{k=1}^{3}\frac{x_k\Omega_{ik}(t,x;\xi)}{|x|^2}. \tag{8.19}$$

Thus $Q_{ij}(\xi,\eta)=\xi_i\eta_j-\xi_j\eta_i$ equals, if we express ξ_i and ξ_j using this identity,

$$\frac{1}{|x|}\sum_{k=1}^{3}\frac{x_k}{|x|}\left[\xi_k\Omega_{ij}(t,x;\eta)+\Omega_{ik}(t,x;\xi)\eta_j-\Omega_{jk}(t,x;\xi)\eta_i\right].$$

For Q_{0j} we have

$$Q_{0j}(\xi,\eta)=\frac{1}{t}\left[\xi_0\Omega_{0j}(t,x;\eta)-\eta_0\Omega_{0j}(t,x;\xi)\right],$$

which takes care of the case $|t|\ge|x|$. For $|t|\le|x|$ we use (8.19) as well as

$$\xi_0=-\sum_{k=1}^{3}\frac{tx_k}{|x|^2}\xi_k+\sum_{k=1}^{3}\frac{x_k\Omega_{0k}(t,x;\xi)}{|x|^2} \tag{8.20}$$

to express $Q_{0j}(\xi,\eta)=\xi_0\eta_j-\xi_j\eta_i$.

Finally, for Q_0 we have

$$Q_0(\xi,\eta)=\frac{1}{t}\left[\xi_0 L_0(t,x;\eta)-\sum_{i=1}^{3}\Omega_{0i}(t,x;\xi)\eta_i\right],$$

which covers $|t|\geq|x|$. In the other case we substitute (8.19) and (8.20) for the components of ξ in $Q_0(\xi,\eta)=\xi_0\eta_0-\sum_1^3\xi_i\eta_i$, which then equals

$$\frac{1}{|x|}\sum_{k=1}^{3}\frac{x_k}{|x|}\left[-t\xi_k\eta_0-\sum_{i=1}^{3}x_i\xi_k\eta_i+\cdots\right]=\frac{1}{|x|}\sum_{k=1}^{3}\frac{x_k}{|x|}\left[-\xi_k L_0(t,x;\eta)+\cdots\right]$$

where $\cdots$ indicate terms involving $\Omega(t,x;\xi)$ and η, which are OK. □

We will need to use energy norms involving Γ, so we have to calculcate $\Gamma^\alpha[F^I(u,\partial u)]$. We therefore need to calculate $\Gamma Q(\partial v,\partial w)$ for any null form Q. By the product rule for derivatives it is easy to see that

$$\Gamma Q(\partial v,\partial w)=Q(\Gamma\partial v,\partial w)+Q(\partial v,\Gamma\partial w), \tag{8.21}$$

but in order to apply Lemma 9 we need to commute ∂ with Γ in the right hand side. This introduces some error terms, but fortunately these are again combinations of null forms, and are therefore estimable by Lemma 9. It is convenient then to introduce the "commutator"

$$[\Gamma,Q](\partial v,\partial w)=\Gamma Q(\partial v,\partial w)-Q(\partial\Gamma v,\partial w)-Q(\partial v,\partial\Gamma w).$$

The commutation relations we need are easily checked, and we list them in the following lemma.

Lemma 10. *We have*

$$\begin{aligned}[\Omega_{ij},Q_0]&=0,\\ [\Omega_{ij},Q_{ab}]&=\delta_{ia}Q_{jb}-\delta_{ja}Q_{ib}-\delta_{ib}Q_{ja}+\delta_{jb}Q_{ia},\\ [L_0,Q]&=-2Q,\end{aligned}$$

where Q stands for any null form and $0\leq i,j,a,b\leq 3$.

Proof. The idea is to use (8.21) and commute Γ with ∂. To this end, one uses the easily checked commutation relations

$$\begin{aligned}[L_0,\partial_k]&=-\partial_k,\\ [\Omega_{0j},\partial_k]&=-\delta_{0k}\partial_j-\delta_{jk}\partial_0,\\ [\Omega_{ij},\partial_k]&=\delta_{ik}\partial_j-\delta_{jk}\partial_i,\end{aligned}$$

where $0\leq k\leq 3$ and $1\leq i,j\leq 3$. We omit the details. □

From Lemmas 9 and 10 one immediately obtains the following key estimate:

Proposition 8. *For any null form Q, and any integer $M \geq 0$, we have*

$$\begin{aligned}(1+|t|+|x|) &\sum_{|\alpha|\leq M} |\Gamma^\alpha Q(\partial v, \partial w)| \\ &\leq C_M \left(\sum_{1\leq|\alpha|\leq M+1} |\Gamma^\alpha v(t,x)| \right) \left(\sum_{1\leq|\alpha|\leq \frac{M}{2}+1} |\Gamma^\alpha w(t,x)| \right) \\ &+ C_M \left(\sum_{1\leq|\alpha|\leq \frac{M}{2}+1} |\Gamma^\alpha v(t,x)| \right) \left(\sum_{1\leq|\alpha|\leq M+1} |\Gamma^\alpha w(t,x)| \right).\end{aligned}$$

8.2.3 An energy inequality and Hörmander's estimate

We need two more ingredients for the proof. The first is a generalization of the energy inequality

$$\|\partial u(t,\cdot)\|_{L^2} \leq \|\partial u(0,\cdot)\|_{L^2} + \int_0^t \|\Box u(s,\cdot)\|_{L^2}\, ds. \tag{8.22}$$

Recall that this is proved by noticing that $u_t \Box u$ is a spacetime divergence:

$$u_t \Box u = \operatorname{div}_{t,x}(e_0, e'), \tag{8.23}$$

where $e_0 = \frac{1}{2}|\partial u|^2$ and $e' = -u_t \nabla_x u$. Integrating this identity for fixed t gives

$$\int u_t \Box u\, dx = \frac{d}{dt}\int e_0\, dx - \int \operatorname{div}_x(u_t \nabla_x u)\, dx,$$

and the last term vanishes by the divergence theorem, if we assume that u decays sufficiently fast as $|x| \to \infty$. For the *energy* $E(t) = \int e_0(t,x)\, dx$ one then obtains, after applying the Cauchy-Schwarz inequality to the left hand side of the above identity,

$$E'(t) \leq \sqrt{E(t)}\, \|\Box u(t,\cdot)\|_{L^2},$$

and (8.22) follows readily.

We now want to generalize this method by replacing (8.23) with

$$X(\partial)u \cdot \Box u = \operatorname{div}_{t,x}(e_0, e') \tag{8.24}$$

where $X(\partial)$ is some first order differential operator and (e_0, e') is some spacetime vector involving u, such that the associated energy is non-negative:

$$E(t) = \int e_0(t,x)\, dx \geq 0.$$

Then by integrating (8.24) one obtains a generalized energy inequality. Set

$$X(\partial) = \vec{X} \cdot \partial + 2t,$$

where as usual ∂ is the spacetime gradient and

$$\vec{X} = (1 + t^2 + |x|^2, 2tx_1, 2tx_2, 2tx_3).$$

Let m be the matrix $\mathrm{diag}(1, -1, -1, -1)$ (the Minkowski metric), and let $\mathbf{1} = (1, 0, 0, 0)$. It then turns out that (8.24) holds with (here we consider ∂u as a column vector for the purposes of matrix multiplication)

$$(e_0, e') = X(\partial)u \cdot m(\partial u) - \frac{1}{2}(\partial u)^T m(\partial u)\vec{X} - v^2\mathbf{1}.$$

Integration of (8.24) then yields the identity

$$\frac{d}{dt}E(t) = \int X(\partial)u \cdot \Box u \, dx$$

for the energy $E(t) = \int e_0 \, dx$, and the right hand side is bounded by

$$\left\|(1 + t + |\cdot|)^{-1} X(\partial)u(t, \cdot)\right\|_{L^2} \|(1 + t + |\cdot|)\Box u(t, \cdot)\|_{L^2}.$$

One then shows that $\left\|(1 + t + |\cdot|)^{-1} X(\partial)u(t, \cdot)\right\|_{L^2} \le C\sqrt{E(t)}$. Putting all this together, one obtains

$$\sqrt{E(t)} \le C\sqrt{E(0)} + C \int_0^t \|(1 + s + |\cdot|)\Box u(s, \cdot)\|_{L^2} \, ds.$$

Moreover, it turns out that

$$E(t) \approx \sum_{|\alpha| \le 1} \|\Gamma^\alpha u(t, \cdot)\|_{L^2}^2,$$

and recalling the commutation relations between $\Box$ and the invariant vector fields, one finally obtains:

Proposition 9. *For any integer $M \ge 0$, there is a constant C such that*

$$\sum_{|\alpha| \le M+1} \|\Gamma^\alpha u(t, \cdot)\|_{L^2} \le C \sum_{|\alpha| \le M+1} \|\Gamma^\alpha u(0, \cdot)\|_{L^2} + C \sum_{|\alpha| \le M} \int_0^t \|(1 + s + |\cdot|)\Gamma^\alpha \Box u(s, \cdot)\|_{L^2} \, ds$$

for all $t > 0$ and all $u \in C^\infty([0, \infty) \times \mathbb{R}^3)$ with compact support in x for each t.

See Sogge's book for the details.

We need one more ingredient for the proof of the main theorem:

Theorem 31. *(Hörmander.) There exists C such that if $F \in C^2([0, \infty) \times \mathbb{R}^3)$ and $\Box u = F$ with vanishing initial data at $t = 0$, then*

$$(1 + t + |x|) \le C \sum_{|\alpha| \le 2} \int_0^t \int_{\mathbb{R}^3} |\Gamma^\alpha F(s, y)| \frac{dy \, ds}{1 + s + |y|}.$$

See Sogge's book or the lecture notes of Hörmander for the proof, which is based on the special form of the fundamental solution for the wave operator in space dimension three.

8.2.4 Proof of the main theorem

Since F is assumed to satsify the null condition, we know that the system (8.14) takes the form

$$\square u^I = \sum_{J,K} a^I_{JK} Q_0(\partial u^J, \partial u^K) + \sum_{J,K,\mu,\nu} b^{I\mu\nu}_{JK} Q_{\mu\nu}(\partial u^J, \partial u^K) + R^I(u, \partial u), \tag{8.25}$$

where R^I vanishes to third order at 0. We specify initial data

$$u\big|_{t=0} = \varepsilon f, \qquad \partial_t u\big|_{t=0} = \varepsilon g. \tag{8.26}$$

For simplicity we will ignore the higher order term R^I.

We shall want to apply Theorem 31 to $\Gamma^\alpha u$, where u solves the above Cauchy problem, but in order to do this we must subtract off the solution $w_\alpha = (w^1_\alpha, \dots, w^N_\alpha)$ of the linear Cauchy problem

$$\square w_\alpha = 0, \quad w_\alpha\big|_{t=0} = (\Gamma^\alpha u)\big|_{t=0}, \quad \partial_t w_\alpha\big|_{t=0} = (\partial_t \Gamma^\alpha u)\big|_{t=0}. \tag{8.27}$$

Thus $\Gamma^\alpha u - w_\alpha$ has vanishing initial data, so we may apply Theorem 31 to it. But then we also need to estimate $|w_\alpha|$.

Observation 1. If u solves (8.25),(8.26) and w_α solves (8.27), then

$$|w_\alpha(t,x)| \le \frac{C_\alpha \varepsilon}{1+t} \quad \text{for all} \quad t \ge 0,\ x \in \mathbb{R}^3 \tag{8.28}$$

where C_α is independent of t and ε.

Proof. Let u_0 solve $\square u_0 = 0$ with initial data (8.26). Then $\square \Gamma^\alpha u = 0$, and so by the decay estimate for solutions of the homogeneous wave equation (week 2),

$$|\Gamma^\alpha u_0(t,x)| \le \frac{C_\alpha \varepsilon}{1+t} \quad \text{for all} \quad t \ge 0,\ x \in \mathbb{R}^3. \tag{8.29}$$

Thus, to get (8.28), it is enough to show that

$$|w_\alpha(t,x) - \Gamma^\alpha u_0(t,x)| \le \frac{C_\alpha \varepsilon^2}{1+t} \quad \text{for all} \quad t \ge 0,\ x \in \mathbb{R}^3.$$

But $\square(w_\alpha - \Gamma^\alpha u_0) = 0$, so this estimate again follows from the decay property used above, if we just observe that the initial data are $O(\varepsilon^2)$. In fact, the data are

$$\Gamma^\alpha u(0,x) - \Gamma^\alpha u_0(0,x), \quad \partial_t \Gamma^\alpha u(0,x) - \partial_t \Gamma^\alpha u_0(0,x).$$

To express $\Gamma^\alpha u(0,x)$ in terms of ε, f, g we just use (8.25) and (8.26) (whenever there are two or more time derivatives we need to use the equation). If we do the same for $\Gamma^\alpha u(0,x)$ and subtract, all terms which are linear in ε cancel out, and we are left with terms arising from the nonlinearity in (8.25), and which therefore are at least quadratic in ε. ∎

We split the proof of the main theorem into a number of steps.

Step 1. Let $0 < T_0 < \infty$ and set $S_{T_0} = [0, T_0) \times \mathbb{R}^3$. Suppose $u \in C^\infty(S_{T_0})$ solves (8.25), (8.26) (setting $R^I = 0$ for simplicity). We shall prove the existence of $\varepsilon_0 > 0$, independent of T_0, such that

$$0 < \varepsilon < \varepsilon_0 \implies u, \partial u \in L^\infty(S_{T_0}). \tag{8.30}$$

Once we know this, it follows from the local existence theory of week 5 that the lifespan $T_\varepsilon = \infty$.

Step 2. (8.30) will follow if we can prove the *a priori* estimate

$$\sum_{|\alpha| \le k} \|\Gamma^\alpha u(t, \cdot)\|_{L^\infty} \le \frac{A\varepsilon}{1+t} \tag{8.31}$$

for $0 \le t < T_0$, provided $\varepsilon < \varepsilon_0$. Here A is a constant independent of T_0 and ε (A will depend on f and g), and k is a sufficiently large integer ($k = 4$ will do).

Step 3. The plan is to prove (8.31) using the continuity method. Thus, we define

$$E = \{T \in [0, T_0) : (8.31) \text{ holds for all } 0 \le t \le T\}.$$

Clearly $0 \in E$ if we take A sufficently large, and E is evidently closed. If we can show that E is open in $[0, T_0)$, it will therefore follow that $E = [0, T_0)$, finishing the proof.

To this end, fix $T \in E$. By continuity of the left hand side of (8.31) ($\Gamma^\alpha u$ is C^∞ and compactly supported in x on $[0, T']$, by Huygens' principle), there certainly exists $T' > T$ such that

$$\sum_{|\alpha| \le k} \|\Gamma^\alpha u(t, \cdot)\|_{L^\infty} \le \frac{2A\varepsilon}{1+t} \quad \text{for} \quad 0 \le t \le T'. \tag{8.32}$$

The idea is then to use a boot-strap argument to show that we have the stronger estimate (8.31) on $[0, T']$. It then follows that $T' \in E$, proving that E is open.

To prove that (8.32) implies (8.31) when ε is sufficiently small, we first show that (8.32) implies

$$\sum_{|\alpha| \le k+3} \|\Gamma^\alpha u(t, \cdot)\|_{L^2} \le C_0 (1+t)^{C_1 \varepsilon} \sum_{|\alpha| \le k+3} \|\Gamma^\alpha u(0, \cdot)\|_{L^2} \tag{8.33}$$

for $0 \le t \le T'$, where C_0 and C_1 are absolute constants. Then we prove that (8.33) implies (8.31), if ε is sufficiently small.

Step 4. We prove (8.32) $\implies$ (8.33). Define $A(t)$ to be the left hand side of (8.33). Then by Proposition 9,

$$A(t) \le CA(0) + C \sum_{|\alpha| \le k+2} \int_0^t \|(1 + s + |\cdot|) \Gamma^\alpha \Box u(s, \cdot)\|_{L^2} \, ds.$$

If we ignore R^I, then $\Box u^I$ is just a linear combination of $Q(\partial u^J, \partial u^K)$ for the null forms Q, and so from Proposition 8 we get, taking the highest order derivatives in L^2, and the lowest order in L^∞,

$$A(t) \leq CA(0) + C\int_0^t A(s)\left(\sum_{|\alpha|\leq \frac{k+2}{2}+1} \|\Gamma^\alpha u(s,\cdot)\|_{L^\infty}\right) ds.$$

Since $\frac{k+2}{2}+1 = \frac{k}{2}+2 \leq k$ if $k \geq 4$, the second factor in the integrand is bounded by $2A\varepsilon/(1+s)$ according to (8.32), and so we get

$$A(t) \leq CA(0) + C'\varepsilon \int_0^t \frac{A(s)}{1+s}\, ds, \quad 0 \leq t \leq T'.$$

Then by Gronwall's Lemma,

$$A(t) \leq CA(0) \exp\left[C'\varepsilon \int_0^t (1+s)^{-1}\, ds\right] = CA(0)(1+t)^{C'\varepsilon},$$

proving (8.33).

Step 5. We prove (8.33) $\Longrightarrow$ (8.31). We can of course choose A so large that (8.28) implies

$$\sum_{|\alpha|\leq k} \|w_\alpha(t,\cdot)\|_{L^\infty} \leq \frac{A\varepsilon/2}{1+t}$$

for all $t \geq 0$. Thus (8.31) follows if we can show that

$$\sum_{|\alpha|\leq k} \|\Gamma^\alpha u(t,\cdot) - w_\alpha(t,\cdot)\|_{L^\infty} \leq \frac{A\varepsilon/2}{1+t} \tag{8.34}$$

for $0 \leq t \leq T'$. To prove this we apply Theorem 31, which gives

$$(1+t)\sum_{|\alpha|\leq k} \|\Gamma^\alpha u(t,\cdot) - w_\alpha(t,\cdot)\|_{L^\infty} \leq C\sum_{|\beta|\leq 2}\int_0^t\int_{\mathbb{R}^3}\sum_{|\alpha|\leq k}\left|\Gamma^\beta\Box\Gamma^\alpha u(s,y)\right| \frac{dy\, ds}{1+s}.$$

Using the commutation relations between $\Box$ and the invariant vector fields, we may bound the right hand side by

$$C\sum_{|\alpha|\leq k+2}\int_0^t\int_{\mathbb{R}^3} |\Gamma^\alpha \Box u(s,y)| \frac{dy\, ds}{1+s},$$

and ignoring again the R^I it suffices to estimate this with $\Box u$ replaced by $Q(\partial u^J, \partial u^K)$ for each of the null forms Q. Then if we apply Proposition 8 we can bound the last expression by

$$C\sum_{|\alpha|\leq k+3}\int_0^t\int_{\mathbb{R}^3} |\Gamma^\alpha u(s,y)|^2 \frac{dy\, ds}{(1+s)^2}.$$

This equals

$$C \sum_{|\alpha| \le k+3} \int_0^t \|\Gamma^\alpha u(s,\cdot)\|_{L^2}^2 \frac{ds}{(1+s)^2},$$

and using (8.33) we bound this by

$$CA(0)^2 \int_0^t (1+s)^{2C_1\varepsilon - 2}\, ds.$$

If $2C_1\varepsilon < 1$, the integral is uniformly bounded in t, and since $A(0) = O(\varepsilon)$, we finally obtain the bound

$$(1+t) \sum_{|\alpha| \le k} \|\Gamma^\alpha u(t,\cdot) - w_\alpha(t,\cdot)\|_{L^\infty} \le C\varepsilon^2$$

for $0 \le t \le T'$, where C is an absolute constant. If $C\varepsilon \le A/2$, then (8.34) follows, and the proof is now complete.

Chapter 9

Week 10: Well-posedness

9.1 Local well-posedness

9.1.1 Introduction and definitions

Consider again a system

$$\Box u^I = F^I(u, \partial u), \qquad (t,x) \in \mathbb{R}^{1+n}, \tag{9.1}$$

where the unknown u and the given C^∞ function F are $\mathbb{R}^N$-valued:

$$u = (u^1, \dots, u^N), \quad F = (F^1, \dots, F^N).$$

Moreover, we assume that $F(0) = 0$. Now specify initial data

$$u\big|_{t=0} = f \in H^s, \qquad \partial_t u\big|_{t=0} = g \in H^{s-1}. \tag{9.2}$$

Here $f = (f^1, \dots, f^N)$ with each $f^I \in H^s$, and similarly for g. The norm on $H^s \times H^{s-1}$ is denoted

$$\|(f,g)\|_{(s)} = \|f\|_{H^s} + \|g\|_{H^{s-1}}.$$

Recall that we have the following local existence and uniqueness result.

Theorem 32. ***(Classical Local Existence Theorem.)*** *The Cauchy problem* (9.1), (9.2) *is locally well-posed for initial data in* $H^s \times H^{s-1}$ *for all* $s > \frac{n}{2} + 1$.

Here, *locally well-posed* (abbreviated LWP henceforth) means:

(i) ***(Local existence.)*** Given $(f,g) \in H^s \times H^{s-1}$, there exist:

- $T = T(f,g) > 0$, depending continuously on $\|(f,g)\|_{(s)}$;
- $u = u(f,g) \in X^s_T = C([0,T], H^s) \cap C^1([0,T], H^{s-1})$ solving (9.1), (9.2) on $S_T = (0,T) \times \mathbb{R}^n$. [Here (9.1) holds in the sense of $\mathcal{D}'(S_T)$.]

(ii) ***(Uniqueness.)*** Solutions are unique in X^s_T, for *any* $T > 0$.

(iii) ***(Continuous dependence on data.)*** The solution $u = u(f, g)$ depends continuously on the data (f, g), in the following sense: If $u(f, g)$ exists up to some time $T > 0$, then there are constants $C, \delta > 0$ such that whenever

$$\|(f - f', g - g')\|_{(s)} \leq \delta,$$

the solution $u(f', g')$ exists up to time T also, and

$$\|u(f', g') - u(f, g)\|_{X^s_T} \leq C \, \|(f - f', g - g')\|_{(s)} \, .$$

(iv) ***(Persistence of higher regularity.)*** If the data have some additional Sobolev regularity $(f, g) \in H^\sigma \times H^{\sigma-1}$, where $\sigma > s$, then the solution in part (a) is in the space $C([0, T], H^\sigma) \cap C^1([0, T], H^{\sigma-1})$. In particular, if the data are in C^∞_c, then the solution is smooth. (One obtains smoothness in time by using the equation to express time derivatives of order two and higher.)

Remark. Since $F(0) = 0$, the unique solution in X^s_T (any $T > 0$) with data $f = g = 0$ is $u = 0$. Then (c) says that for any $T > 0$, $u(f, g)$ exists up to time T if $\|(f, g)\|_{(s)}$ is sufficiently small (possibly depending on T).

Notation. We often write $u(t)$ instead of $u(t, \cdot)$. This is natural since u solves a time evolution problem. In fact, X^s_T as defined above is just the space of continuous curves from $[0, T]$ into $H^s \times H^{s-1}$.

Recall that the proof of Theorem 32 relies on:

(i) The *energy inequality* for the linear wave equation.

(ii) *Sobolev's Lemma* (the special case $H^r \subset L^\infty$ iff $r > \frac{n}{2}$).

(iii) The *Moser inequality*, which for F as in (9.1) says that there is a continuous function $\phi_s : [0, \infty) \to [0, \infty)$ such that

$$\|F(u, \partial u)(t)\|_{H^{s-1}} \leq \phi_s\big(\|(u, \partial u)(t)\|_{L^\infty}\big) \, \|(u, \partial_t u)(t)\|_{(s)} \, ,$$

provided $s \geq 1$.

The reason for the lower bound on s in Theorem 32 is then clear: After applying (i) and (iii) we need to control

$$\|(u, \partial u)(t)\|_{L^\infty} \, ,$$

and by (ii) this norm is dominated by

$$\|u\|_{X^s_T} = \sup_{0 \leq t \leq T} \|(u, \partial_t u)(t)\|_{(s)}$$

precisely when $s > \frac{n}{2} + 1$.

This lower bound on s is sharp in general. In fact we shall see later that the scalar equation $\square u = (\partial_t u)^k$ $(k \geq 2)$ is *not* well-posed (thus we say it is *ill-posed*) for data in $H^s \times H^{s-1}$ with

$$s < \frac{n}{2} + 1 - \frac{1}{k-1},$$

and this approaches $n/2 + 1$ as $k \to \infty$.

On the other hand, one can show that the equation $\square u = (\partial_t u)^2$ is LWP for $s > 2$ in dimension $n = 3$, whereas Theorem 32 requires $s > 5/2$, so there is a gap.

We shall be interested in improving the lower bound in Theorem 32 for certain equations. Thus we ask the following:

Question. *For a given F in* (9.1)*, what is the minimal s for which the conclusion of Theorem 32 holds for data in $H^s \times H^{s-1}$?*

Remark. We may have to replace the "energy space" X_T^s by some subspace (still a Banach space) in the definition of LWP above. This will be clear from examples to follow.

9.1.2 Scaling

For all the equations we are interested in (namely [models for] classical field equations from physics), there is a natural lower bound for s imposed by scaling properties (homogeneity) of F and the data space $\dot{H}^s$, given by the norm

$$\|f\|_{\dot{H}^s} = \left\| |\xi|^s \widehat{f} \right\|_{L^2}.$$

This lower bound we call the *critical well-posedness exponent* and denote s_c; it is the unique $s \in \mathbb{R}$ such that the homogeneous data space $\dot{H}^s \times \dot{H}^{s-1}$ is invariant (dimensionless) under the natural scaling of the equation (9.1). This is best illustrated by some simple examples.

We shall use the easily proved fact that

$$\|f(\lambda x)\|_{\dot{H}^s} = \lambda^{s-n/2} \|f\|_{\dot{H}^s} \quad \text{for} \quad \lambda > 0. \tag{9.3}$$

Examples. (A) Consider $\square u = (\partial_t u)^2$ on $\mathbb{R}^{1+n}$. If u is a solution, then so is u_λ $(\lambda > 0)$ given by

$$u_\lambda(t, x) = u(\lambda t, \lambda x).$$

Since by (9.3) we have

$$\|u_\lambda(0)\|_{\dot{H}^s} = \lambda^{s-n/2} \|u(0)\|_{\dot{H}^s},$$

we conclude that $s_c = n/2$.

(B) For $\square u = u\partial_t u$ the scaling is

$$u_\lambda(t, x) = \lambda u(\lambda t, \lambda x).$$

Thus $\|u_\lambda(0)\|_{\dot{H}^s} = \lambda^{1+s-n/2}\|u(0)\|_{\dot{H}^s}$, and $s_c = n/2 - 1$.

(C) For $\Box u = u^2$ we find similarly $s_c = n/2 - 2$.

We now formulate:

General WP Conjecture. *(i) For all classical field theories the Cauchy problem is LWP for data in $H^s \times H^{s-1}$, $s > s_c$.*

(ii) For smooth data with small $H^{s_c} \times H^{s_c-1}$ norm, there exists a global smooth solution.

(iii) The Cauchy problem is ill-posed for data in $H^s \times H^{s-1}$, $s < s_c$.

Part (i) has been verified for several important equations in the last few years. Very recently, part (ii) was verified for so-called Wave Maps (analogue of the wave equation for functions with values in a sphere) through the work of T. Tao.

We will content ourselves with looking at some simple examples which give at least some motivation for this conjecture. Let us start with item (iii).

Some terminology: the regimes $s > s_c$, $s = s_c$ and $s < s_c$ are called *subcritical*, *critical* and *supercritical*, respectively.

9.1.3 Blowup and nonexistence in the supercritical range

According to item (iii) of the WP Conjecture, we expect ill-posedness in the supercritical range $s < s_c$. The following result gives some motivation for this.

Theorem 33. *If there exist data $f, g \in C_c^\infty(\mathbb{R}^n)$ such that the solution of $\Box u = F(u, \partial u)$ with data (f, g) blows up at finite time in some open ball, then there is nonexistence in the supercritical range.*

Remark. The above applies to Wave Maps in dimensions $n \geq 3$ (Shatah).

For simplicity let us assume $f = 0$. The idea is that when we scale u (and hence g) in the natural way with a parameter $\lambda \to \infty$, then the blowup time goes to zero, and the size of the $\dot{H}^{s-1}$ norm of g also goes to zero by supercriticality. Moreover, the support of g shrinks to a point, so by letting $\lambda \to \infty$ through an appropriate sequence, and adding up the scaled g's, translated so as to make the supports well separated, we get a series converging in H^{s-1} to some $\widetilde{g}$. Then by a domain of dependence argument, there is no local existence for $\Box u = F(u, \partial u)$ with initial data $(0, \widetilde{g})$. Furthermore, $\widetilde{g}$ can be made to have arbitrarily small norm and arbitrarily small support.

Let us see how this works for a concrete example.

Example. Consider again $\Box u = (\partial_t u)^k$ with $k \geq 2$ an integer. To determine the scaling, we set

$$u_\lambda(t, x) = \lambda^\beta u(\lambda t, \lambda x),$$

where β must be determined. It is easily seen that if both u and u_λ are solutions of the equation, then

$$\beta + 2 = (\beta + 1)k \implies \beta = \frac{2-k}{k-1}.$$

We conclude that $s_c = n/2 - \beta = n/2 + 1 - 1/(k-1)$.

Using the fact that the ODE

$$y' = y^k, \qquad y(0) = y_0 > 0$$

blows up in finite time, we conclude that there exists $g \in C_c^\infty$ so that the solution of $\Box u = (\partial_t u)^k$ with data $(0, g)$ blows up in the unit ball $|x| \le 1$ at time $t = 1$, say. In fact, we can just start with g constant and then cut it off smoothly outside a sufficiently large ball. The blowup then follows by domain of dependence (uniqueness in backwards light cones).

Given such g, let us see how to construct $\widetilde{g}$ with the properties described above. Corresponding to the scaling $u \to u_\lambda$ found above, the initial datum scales as follows:

$$g \to g_\lambda, \qquad g_\lambda(x) = \lambda^{1+\beta} g(\lambda x).$$

Now fix $1 \le s < s_c$ and observe the following:

(i) Since u blows up at time $T = 1$ in $|x| \le 1$ and g vanishes outside some ball $|x| \le R$, it follows that u_λ blows up at time $1/\lambda$ in $|x| \le 1/\lambda$ and g_λ vanishes outside $|x| \le R/\lambda$.

(ii) By supercriticality, $\|g_\lambda\|_{H^{s-1}} \to 0$ as $\lambda \to 0$. In fact, it is easy to see that

$$\|g_\lambda\|_{H^{s-1}} \le C\lambda^{s-s_c} \|g\|_{H^{s-1}}.$$

It is easy to choose a sequence $\lambda_j \to \infty$ and a convergent sequence of disjoint points x_j in $\mathbb{R}^n$ such that $\sum_0^\infty \lambda_j^{s-s_c} < \infty$ and the supports of

$$h_j(x) = g_{\lambda_j}(x - x_j) = \lambda_j^{1+\beta} g(\lambda_j[x - x_j])$$

are mutually disjoint. Then set

$$\widetilde{g} = \sum_0^\infty h_j(x).$$

This converges absolutely in H^{s-1} in view of the above. Now let $\widetilde{u}$ be a solution with data $(0, \widetilde{g})$. Then by a domain of dependence argument, $\widetilde{u}$ must blow up in the ball $|x - x_j| \le 1/\lambda_j$ at time $T = 1/\lambda_j$, for every j. Hence there is no local existence near $\lim x_j$. Finally, by replacing the sequences by their N-tails for N large, we can make the norm and support of $\widetilde{g}$ as small as we like. Note that $\widetilde{g}$ is smooth except at $\lim x_j$. This concludes the example.

9.1.4 LWP implies domain of dependence valid

Observe that the proof of nonexistence in the previous section relied on a domain of dependence argument, that is, the data in a ball uniquely determine the solution in the backwards light cone over that ball. So what we really showed is that there is no local solution obeying this principle. A key fact, proved below, is that any solution obtained in a LWP framework satisfies domain of dependence.

Thus, under the hypotheses of Theorem 33 there is ill-posedness in the supercritical range, verifying item (iii) of the General WP Conjecture for equations with blowup for smooth data.

Assume $\Box u = F(u, \partial u)$ is LWP for data in $H^s \times H^{s-1}$. Suppose u and v both solve the equation up to time T with data (f, g) and (f', g') respectively, and assume that

$$f = f' \quad \text{and} \quad g = g' \quad \text{in the ball} \quad |x - x_0| \le r.$$

Let Ω be the cone over this ball:

$$\Omega = \{(t, x) : t > 0,\, t + |x - x_0| < r\}.$$

Then

$$u = v \quad \text{in} \quad \Omega \cap S_T,$$

where $S_T = (0, T) \times \mathbb{R}^n$. It suffices to prove this in a smaller cone Ω' defined as Ω but with r replaced by a slightly smaller r'. Using cutoffs and a smooth approximation of the identity, we can find sequences $f_j, g_j, f_j', g_j' \in C_c^\infty$ such that $(f_j, g_j) \to (f, g)$ and $(f_j', g_j') \to (f', g')$ in $H^s \times H^{s-1}$ and

$$f_j = f_j' \quad \text{and} \quad g_j = g_j' \quad \text{in the ball} \quad |x - x_0| \le r'$$

for all j. Let u_j and u_j' be the solutions corresponding to the data (f_j, g_j) and (f_j', g_j'), respectively. It follows by continuous dependence on the data that $u_j \to u$ and $u_j' \to u'$ in $\mathcal{D}'(S_T)$.

Moreover, both u_j and u_j' are C^∞ by persistence of higher regularity. From the uniqueness theorem for smooth solutions in backwards light cones (week 5, Theorem 3) it then follows that $u_j = u_j'$ in $\Omega' \cap S_T$. Passing to the limit we conclude that $u = u'$ in $\Omega' \cap S_T$, and this completes the proof, since $r' < r$ was arbitrary.

9.1.5 Nonuniqueness in the supercritical case

Here we give an example due to Hans Lindblad of nonuniqueness of a nonlinear wave equation in the supercritical range. In particular, domain of dependence fails.

Example. Consider $\Box u = u^3$ on $\mathbb{R}_+^{1+3} = (0, \infty) \times \mathbb{R}^3$. Then $s_c = 1/2$, and one can show (we will do this later using Strichartz' inequality) that the equation is globally well-posed for data in $\dot H^{1/2} \times \dot H^{-1/2}$ with small norm.

We want to show that uniqueness fails in the space

$$C([0,\infty),\dot H^s)\cap C^1([0,\infty),\dot H^{s-1})$$

when the regularity is supercritical, that is, $s<1/2$. In fact, we give an example of a nonzero solution of $\Box u = u^3$ in the above space, with initial data

$$u\big|_{t=0} = \partial_t u\big|_{t=0} = 0. \tag{9.4}$$

Since $u \equiv 0$ also is a solution (the only reasonable one), we have nonuniqueness.

Define

$$u(t,x) = \frac{\sqrt{2}H(t-|x|)}{t}, \qquad (t,x)\in\mathbb{R}^{1+3}_+.$$

Here H is the Heaviside function. Thus, $H(t-|x|)$ is just the characteristic function of the solid light cone $\{(t,x) : t>0, |x|<t\}$. Recall that $H'=\delta$. Using the Chain Rule (this is justified; see section 6.1 in Hörmander's Linear Partial Differential Operators Vol. I, 2nd ed.) it is easy to calculate the first and second order partial derivatives of u in $\mathcal{D}'$, and one finds that u solves $\Box u = u^3$.

It only remains to check that

$$\lim_{t\to0^+}\|u(t)\|_{\dot H^s} = \lim_{t\to0^+}\|\partial_t u(t)\|_{\dot H^{s-1}} = 0$$

when $s<1/2$. It then follows that $(u,\partial_t u)$ extends continuously to $t=0$ in $\dot H^s\times\dot H^{s-1}$ and (9.4) holds.

It suffices to prove that $\|u(t=1)\|_{\dot H^s}<\infty$. Then we can exploit the supercritical scaling to conclude that the limit as $t\to0$ is 0. Thus, we have to show $\|\chi_B\|_{\dot H^s}<\infty$, where χ_B is the characteristic function of the unit ball $B=\{x : |x|<1\}$ in $\mathbb{R}^3$. We have to calculate the Fourier transform. Using polar coordinates $x=r\omega$ we have

$$\widehat{\chi_B}(\xi) = \int_0^1\int_{S^2} e^{-ir\omega\cdot\xi}\,d\sigma(\omega)\,r^2\,dr = \int_0^1 \widehat{\sigma}(r\xi)r^2\,dr,$$

where σ is surface measure on S^2. So we need to calculate $\widehat{\sigma}(\xi)$.

By rotational symmetry it suffices to take $\xi=(0,0,\rho)$, $\rho=|\xi|$. Then using spherical coordinates on $S^2=\{(x,y,z) : x^2+y^2+z^2=1\}$,

$$\omega = \begin{cases} x=\sin\phi\cos\theta \\ y=\sin\phi\sin\theta \\ z=\cos\phi \end{cases} \qquad 0<\phi<\pi, \quad 0<\theta<2\pi,$$

we have

$$\int_{S^2} f(\omega)\,d\sigma(\omega) = \int_0^\pi\int_0^{2\pi} f\,|\omega_\phi\times\omega_\theta|\,d\theta\,d\phi = \int_0^\pi\int_0^{2\pi} f\sin\phi\,d\theta\,d\phi,$$

and we conclude that

$$\widehat{\sigma}(0,0,\rho) = \int_0^\pi\int_0^{2\pi} e^{-i\rho\cos\phi}\sin\phi\,d\theta\,d\phi = 2\phi\int_{-1}^1 e^{i\rho r}\,dr = 4\pi\frac{\sin\rho}{\rho}.$$

Therefore

$$\widehat{\chi_B}(\xi) = 4\pi \int_0^1 \frac{\sin(r\,|\xi|)}{r\,|\xi|} r^2\, dr = \frac{4\pi}{|\xi|^3} \int_0^{|\xi|} \lambda \sin\lambda \, d\lambda = \frac{4\pi}{|\xi|^3} \left(\sin|\xi| - |\xi| \cos|\xi|\right),$$

whence $|\widehat{\chi_B}(\xi)| \leq \frac{C}{1+|\xi|^2}$. Thus $\int_{\mathbb{R}^3} |\xi|^{2s} \left|\widehat{\chi_B}(\xi)\right|^2 \, d\xi < \infty$ iff $s < 1/2$.

It remains to work out the scaling of $\|u(t)\|_{\dot{H}^s}$ with respect to $t > 0$. In fact it is easy to see, by using the Fourier transform, that

$$\|u(t)\|_{\dot{H}^s}^2 = t^{1-2s} \|u(1)\|_{\dot{H}^s}^2 ,$$

so the limit as $t \to$ is 0 iff $s < 1/2$. This concludes the example.

Chapter 10

Week 11: Strichartz type estimates

10.1 Introduction

Our next topic is Strichartz estimates. These are certain spacetime integrability properties of solutions to the linear Cauchy problem

$$\Box u = F, \qquad (u, \partial_t u)\big|_{t=0} = (f, g), \tag{10.1}$$

and are intimately connected with a well-known problem in harmonic analysis, namely the (L^p, L^2) restriction problem for the Fourier transform: Consider the Fourier transform $f \to \widehat{f}$ on $\mathbb{R}^n$, and let $S \subset \mathbb{R}^n$ be a hypersurface. The question is then for which exponents $1 \le p < 2$ the map

$$f \to \widehat{f}\big|_S \qquad (f \in \mathcal{S})$$

extends to a bounded map from $L^p(\mathbb{R}^n)$ to $L^2(S)$.

For $p = 1$ the map is bounded by the Riemann-Lebesgue lemma, while if $p = 2$, then $\widehat{f}$ can be any L^2 function on $\mathbb{R}^n$, so $\widehat{f}\big|_S$ is meaningless, since S has measure zero in $\mathbb{R}^n$. Hence the restriction to $1 \le p < 2$.

The solution to the restriction problem depends on the curvature of S. If S is a plane, then no $p > 1$ is allowed, while if S is the unit sphere S^{n-1}, then the admissible range of p is

$$1 \le p \le \frac{2(n+1)}{n+3}.$$

This is due to Stein and Tomas. subsequently, Strichartz realized that restriction theorems of this type imply—via a duality argument—estimates for the wave and Schrödinger equations (with the hypersurface being a cone or a paraboloid, respectively). His estimates have been extensively generalized by many people, but still go under the name of Strichartz estimates.

As an initial example, let us consider the following estimate for the solution of (10.1) on $\mathbb{R}^{1+3}$, proved by Strichartz in his original paper:

$$\|u\|_{L^4(\mathbb{R}^{1+3}_+)} + \|u(t)\|_{\dot{H}^{\frac{1}{2}}} + \|\partial_t u(t)\|_{\dot{H}^{-\frac{1}{2}}} \leq C \left(\|f\|_{\dot{H}^{\frac{1}{2}}} + \|g\|_{\dot{H}^{-\frac{1}{2}}} + \|F\|_{L^{\frac{4}{3}}(\mathbb{R}^{1+3}_+)} \right), \tag{10.2}$$

for all $t \geq 0$. We shall see later that this inequality is equivalent to a restriction theorem for the light cone in $\mathbb{R}^{1+3}_+$.

Example. Let us apply the above inequality to prove global existence for $\square u = u^3$ on $\mathbb{R}^{1+3}_+$ with data $(u, \partial_t u)\big|_{t=0} = (f, g) \in \dot{H}^{\frac{1}{2}} \times \dot{H}^{-\frac{1}{2}}$, provided

$$E_0 = \|f\|_{\dot{H}^{\frac{1}{2}}} + \|g\|_{\dot{H}^{-\frac{1}{2}}}$$

is sufficiently small. To see this, denote by $X(u)$ the left hand side of (10.2), with supremum over all $t \geq 0$. We now iterate in this norm. As usual, the iterates are defined inductively by $u_{-1} \equiv 0$ and

$$\square u_j = u_{j-1}^3,$$

with data (f, g), for $j \geq 0$. Then by (10.2), using the fact that

$$\|uvw\|_{L^{4/3}} \leq \|u\|_{L^4} \|v\|_{L^4} \|w\|_{L^4},$$

we have

$$X(u_j) \leq CE_0 + CX(u_{j-1})^3.$$

So if $X(u_{j-1}) \leq 2CE_0$, then so is $X(u_j)$, provided $C(2CE_0)^2 \leq \frac{1}{2}$.

Then, since

$$\square(u_{j+1} - u_j) = u_j^3 - u_{j-1}^3 = (u_j - u_{j-1})u_j^2 + u_{j-1}(u_j + u_{j-1})(u_j - u_{j-1})$$

with vanishing initial data, we have

$$X(u_{j+1} - u_j) \leq C'[X(u_j) + X(u_{j-1})]^2 X(u_j - u_{j-1}) \leq C'(4CE_0)^2 X(u_j - u_{j-1}),$$

so $\{u_j\}$ is Cauchy provided $C' 16 C^2 E_0^2 \leq \frac{1}{2}$.

10.2 Proof of the estimates for $\square u = 0$

Here we prove the Strichartz type estimates for solutions of

$$\square u = 0 \quad \text{on} \quad \mathbb{R}^{1+n}, \qquad u\big|_{t=0} = f, \qquad \partial_t u\big|_{t=0} = g. \tag{10.3}$$

We assume $n \geq 2$ throughout. The estimates are of the form

$$\|u\|_{L^q_t(L^r_x)} \leq C\big(\|f\|_{\dot{H}^s} + \|g\|_{\dot{H}^{s-1}}\big), \tag{10.4}$$

where

$$\|u\|_{L_t^q(L_x^r)} = \left(\int_{\mathbb{R}} \left(\int_{\mathbb{R}^n} |u(t,x)|^r \, dx \right)^{q/r} dt \right)^{1/q},$$

with the obvious modifications if q or $r = \infty$.

By scaling, we must have

$$s = \frac{n}{2} - \frac{n}{r} - \frac{1}{q} \tag{10.5}$$

Definition. We say that a pair (q, r) is *wave admissible* if

$$2 \le q \le \infty, \quad 2 \le r < \infty \quad \text{and} \quad \frac{2}{q} \le \frac{n-1}{2}\left(1 - \frac{2}{r}\right).$$

Theorem 34. *The estimate* (10.4) *holds for all solutions of* (10.3) *if and only if* (q, r) *is wave admissible and* s *is given by* (10.5)*.*

Remark. There are counterexamples which show that the conditions are optimal, but we will not discuss these. Also, we will not prove the so-called endpoint estimate, where

$$1 = \frac{2}{q} \le \frac{n-1}{2}\left(1 - \frac{2}{r}\right).$$

In other words, $q = 2$ and $r = 2(n-1)/(n-3)$. Since we require $r < \infty$, the endpoint is only allowed when $n > 3$.

10.3 The truncated cone operator

Instead of (10.4), it suffices to prove

$$\left\| e^{it\sqrt{-\Delta}} f \right\|_{L_t^q(L_x^r)} \le C \|f\|_{\dot{H}^s}. \tag{10.6}$$

This is because $\widehat{u}(t,\xi)$ is a linear combination of $e^{\pm it|\xi|}\widehat{f}(\xi)$ and $e^{\pm it|\xi|}\widehat{g}(\xi)/|\xi|$. By a density argument, it suffices to prove the estimate for $f \in \mathcal{S}$.

We first prove (10.6) for frequency-localized f, and then obtain the general case using Littlewood-Paley theory. Thus, we fix a radial cutoff function $\beta \in C_c^\infty$ supported away from zero, and consider the *truncated cone operator*

$$Tf(t,x) = \int_{\mathbb{R}^n} e^{ix\cdot\xi} e^{it|\xi|} \beta(\xi) \widehat{f}(\xi) \, d\xi \qquad (f \in \mathcal{S}) \tag{10.7}$$

Now that the frequency has been localized, the $\dot{H}^s$ norm behaves like an L^2 norm, and the problem is then to prove

$$\|Tf\|_{L_t^q(L_x^r)} \le C \|f\|_{L^2} \tag{10.8}$$

for wave admissible (q, r).

10.4 The formal adjoint

The formal adjoint of T is the operator $F(t,x) \to T^*F(x)$ determined by

$$\langle Tf, F\rangle = \langle f, T^*F\rangle \quad \text{for} \quad f \in \mathcal{S}(\mathbb{R}^n),\ F \in \mathcal{S}(\mathbb{R}^{1+n})$$

where $\langle \cdot, \cdot \rangle$ denotes the L^2 inner product. In other words, the condition is that

$$\int Tf \cdot \overline{F}\, dt\, dx = \int f \cdot \overline{T^*F}\, dx.$$

Let us calculate T^*F. Using the definition of Tf we have

$$\begin{aligned}
\int Tf \cdot \overline{F}\, dt\, dx &= \int \widehat{Tf}(t,\xi)\overline{\widehat{F}(t,\xi)}\, d\xi\, dt \\
&= \int e^{it|\xi|}\beta(\xi)\widehat{f}(\xi)\overline{\widehat{F}(t,\xi)}\, d\xi\, dt \\
&= \int f(x) \left(\int e^{-ix\cdot\xi} e^{it|\xi|}\beta(\xi)\overline{\widehat{F}(t,\xi)}\, d\xi\, dt \right) dx.
\end{aligned}$$

We conclude that

$$T^*F(x) = \int e^{i(x\cdot\xi - t|\xi|)}\overline{\beta(\xi)}\widehat{F}(t,\xi)\, d\xi\, dt = \int e^{ix\cdot\xi}\overline{\beta(\xi)}\widetilde{F}(|\xi|,\xi)\, d\xi, \tag{10.9}$$

where $\widetilde{F}$ is the spacetime Fourier transform.

Remark. The above gives the connection with the Fourier restriction problem for the forward light cone $\Lambda = \{(\tau,\xi) : \tau = |\xi| > 0\}$ in $\mathbb{R}^{1+n}$. In fact, from (10.9) we see that

$$\widehat{T^*F}(\xi) \simeq \overline{\beta(\xi)}\widetilde{F}(|\xi|,\xi) = RF(\xi) \tag{10.10}$$

is just the restriction of the spacetime Fourier transform of F to Λ, multiplied by a smooth cutoff. The cone is the graph of $\xi \to (|\xi|,\xi)$, and with respect to this parametrization, surface measure $d\sigma$ on the cone is just $d\xi$ up to a constant. Thus, in view of Plancherel's theorem, $\|T^*F\|_{L^2} \simeq \|RF\|_{L^2(\Lambda,d\sigma)}$. Consequently, the estimate (10.8) is equivalent to the following restriction theorem:

Theorem. *$R : L_t^{q'}(L_x^{r'}) \to L^2(\Lambda, d\sigma)$ is bounded if (q,r) is wave admissible.*

10.5 Duality and the TT^* principle

Recall that for all $1 \le p \le \infty$,

$$\|f\|_{L^p} = \sup\left\{|\langle f, g\rangle| : g \in \mathcal{S},\ \|g\|_{L^{p'}} \le 1\right\},$$

where p' denotes the conjugate exponent. Similarly one has for the mixed norms, for all $1 \le q, r \le \infty$,

$$\|F\|_{L_t^q(L_x^r)} = \sup\left\{|\langle F, G\rangle| : G \in \mathcal{S},\ \|G\|_{L_t^{q'}(L_x^{r'})} \le 1\right\}. \tag{10.11}$$

Using this fact we prove:

Lemma 11. *The following statements are equivalent:*

(i) $T : L^2(\mathbb{R}^n) \to L_t^q(L_x^r)$ *is bounded,*

(ii) $T^* : L_t^{q'}(L_x^{r'}) \to L^2(\mathbb{R}^n)$ *is bounded,*

(iii) $TT^* : L_t^{q'}(L_x^{r'}) \to L_t^q(L_x^r)$ *is bounded.*

Proof. Since

$$|\langle Tf, F\rangle| = |\langle f, T^*F\rangle| \leq \|f\|_{L^2} \|T^*F\|_{L^2},$$

it follows from (10.11) that (ii) implies (i), and the converse follows from

$$|\langle f, T^*F\rangle| = |\langle Tf, F\rangle| \leq \|Tf\|_{L_t^q(L_x^r)} \|F\|_{L_t^{q'}(L_x^{r'})}.$$

Obviously, (i) and (ii) together imply (iii), so it remains to prove that (iii) implies (ii). To see this, observe that

$$\|T^*F\|_{L^2}^2 = \langle T^*F, T^*F\rangle = \langle F, TT^*F\rangle \leq \|F\|_{L_t^{q'}(L_x^{r'})} \|TT^*F\|_{L_t^q(L_x^r)}.$$

□

It turns out that TT^* is a convolution operator. In fact, using (10.7) and (10.10) we see that

$$\widehat{TT^*F}(t,\xi) \simeq e^{it|\xi|}\beta(\xi)\widehat{T^*F}(\xi) \simeq \int e^{i(t-s)|\xi|} |\beta(\xi)|^2 \widehat{F}(s,\xi)\, ds$$

and we conclude that $TT^*F = K * F$, where

$$K(t,x) = \int e^{i(x\cdot\xi + t|\xi|)} |\beta(\xi)|^2 \, d\xi. \tag{10.12}$$

We also define $K_t(x) = K(t,x)$. Observe that $f(x) \to K_t * f(x)$ is essentially the operator T. The only difference is that in the latter we have β and not $|\beta|^2$.

10.6 Estimates for the kernel

We shall prove two estimates at fixed t for the operator $f \to K_t * f$, which we recall is essentially the same as T. We then interpolate between these two fixed-time estimates, and finally we apply either the Hardy-Littlewood inequality or Young's inequality to get a spacetime estimate.

We first prove

$$\|K_t * f\|_{L^2} \leq C \|f\|_{L^2}, \tag{10.13}$$

$$\|K_t * f\|_{L^\infty} \leq \frac{C}{(1+|t|)^{(n-1)/2}} \|f\|_{L^1}. \tag{10.14}$$

Observe that Riesz-Thorin interpolation between these two estimates gives

$$\|K_t * f\|_{L^r} \le \frac{C}{(1+|t|)^{\gamma(r)}} \|f\|_{L^{r'}} \tag{10.15}$$

for all $2 \le r \le \infty$, where

$$\gamma(r) = \frac{n-1}{2}\left(1 - \frac{2}{r}\right). \tag{10.16}$$

The estimate (10.13) is just an energy inequality, and is a trivial consequence of Plancherel's theorem, since $\widehat{K_t}(\xi) \simeq e^{it|\xi|} |\beta(\xi)|^2$. Inequality (10.14) is called the *dispersive inequality.* To prove it, note that by Young's inequality,

$$\|K_t * f\|_{L^\infty} \le \|K_t\|_{L^\infty} \|f\|_{L^1},$$

so it suffices to show that

$$|K(t,x)| \le \frac{C}{(1+|t|)^{(n-1)/2}} \tag{10.17}$$

holds uniformly on $\mathbb{R}^{1+n}$. This is an instance of the following general fact:

Theorem. *Suppose $S \subset \mathbb{R}^{1+n}$ is a hypersurface with at least k nonvanishing principal curvatures at each point. Then the Fourier transform of the surface measure $d\sigma$ on S multiplied by a function $\phi \in C_c^\infty(S)$ satisfies the decay estimate*

$$\widehat{\phi d\sigma}(\xi) = O(|\xi|^{-k/2})$$

as $|\xi| \to \infty$.

Now observe that the forward light cone Λ in $\mathbb{R}^{1+n}$ has exactly $n-1$ non-vanishing principal curvatures at each point, and from (10.12) we have

$$K(t,x) = \int e^{i(t,x)\cdot(\tau,\xi)} |\beta(\xi)|^2 \,\delta(\tau - |\xi|)\, d\tau\, d\xi$$

But $\delta(\tau - |\xi|)\, d\tau\, d\xi$ is surface measure on Λ, up to a constant, so the above theorem applies, and gives (10.17).

However, instead of relying on this general argument, we will prove (10.17) using a very special case of the above theorem, namely for surface measure σ on the sphere S^{n-1} in $\mathbb{R}^n$. We then have

$$|\widehat{\sigma}(\xi)| \le C(1+|\xi|)^{-(n-1)/2}. \tag{10.18}$$

In fact, we proved this for $n = 3$, which is the dimension we shall be concerned with in applications, last week.

Armed with this fact we prove (10.17). Recall that β is assumed to be radial and supported away from zero. Using polar coordinates $\xi = \rho\omega$ we then have

$$K(t,x) = \int_0^\infty \int_{S^{n-1}} e^{i\rho(x\cdot\omega + t)} a(\rho)\, d\sigma(\omega)\, d\rho = \int_0^\infty \widehat{\sigma}(\rho x) e^{it\rho} a(\rho)\, d\rho,$$

where $a(\rho)$ is smooth and compactly supported away from zero.

Case 1: $|t| \geq 2|x|$**.** Integrate by parts N times in $I = \int_0^\infty e^{i\rho(x\cdot\omega+t)} a(\rho)\, d\rho$, using the fact that

$$e^{i\rho(x\cdot\omega+t)} = \frac{(d/d\rho) e^{i\rho(x\cdot\omega+t)}}{i(t + x\cdot\omega)},$$

to get $|I| \leq C_N |t + x\cdot\omega|^{-N} \leq C_N 2^N |t|^{-N}$ uniformly.

Case 2: $|t| < 2|x|$**.** Using (10.18) we have

$$\begin{aligned} |K(t,x)| &\leq \int_0^\infty |\widehat{\sigma}(\rho x)| \, |a(\rho)| \, d\rho \\ &\leq C \int_0^\infty |\rho x|^{-(n-1)/2} |a(\rho)| \, d\rho \leq C |x|^{-(n-1)/2} \leq C |t|^{-(n-1)/2}, \end{aligned}$$

where C as usual can change from line to line. This concludes the proof of (10.17).

10.7 Conclusion of the frequency localized case

We now finish the proof of (10.8). According to Lemma 11 it suffices to prove boundedness of TT^*,

$$\|K * F\|_{L_t^q(L_x^r)} \leq C \|F\|_{L_t^{q'}(L_x^{r'})}, \tag{10.19}$$

for (q, r) wave admissible. Since we ignore the endpoint case, this means that $2 \leq q \leq \infty$, $2 \leq r < \infty$, $2/q \leq \gamma(r)$ and $\big(2/q, \gamma(r)\big) \neq (1, 1)$.

By Minkowski's integral inequality and (10.15),

$$\|K * F(t)\|_{L^r} \leq \int \|K(t-s) * F(s)\|_{L^r} \, ds \leq C \int \frac{\|F(s)\|_{L^{r'}}}{(1 + |t-s|)^{\gamma(r)}} \, ds. \tag{10.20}$$

We claim that this implies (10.19). To see this, we consider separately the cases $2/q < \gamma(r)$ and $2/q = \gamma(r)$.

Case 1: $2/q < \gamma(r)$**.** Then $(1 + |t|)^{-\gamma(r)}$ belongs to $L^{q/2}(\mathbb{R})$, so in this case (10.19) follows from Young's inequality, applied to (10.20). Recall that Young's inequality says that

$$\|f * g\|_{L^q} \leq \|f\|_{L^a} \|g\|_{L^b},$$

provided $1 \leq a, b, q \leq \infty$ and $1 + 1/q = 1/a + 1/b$. In this case we take $a = q/2$ and $b = q'$.

Case 2: $2/q = \gamma(r)$**.** Since we exclude the endpoint case, we must have $2/q = \gamma(r) < 1$. Now $(1 + |t|)^{-\gamma(r)}$ just fails to belong to $L^{q/2}(\mathbb{R})$, so we cannot apply Young's inequality. However, recall the Hardy-Littlewood inequality:

Theorem. ***(Hardy-Littlewood inequality)*** *Let $0 < \alpha < 1$. Assume that $1 < p < q < \infty$ and $1 + 1/q = 1/p + \alpha$. Set*

$$Tf(t) = \int_{\mathbb{R}} \frac{f(s)}{|t-s|^\alpha}\, ds.$$

Then T is bounded from $L^q(\mathbb{R})$ into $L^p(\mathbb{R})$.

Remark. Note that T is convolution with the kernel $|t|^{-\alpha}$, which just fails to belong to $L^{1/\alpha}$; if it *did* belong to $L^{1/\alpha}$, the boundedness of T would follow from Young's inequality.

Apply this theorem to (10.20) with $0 < \alpha = \gamma(r) = 2/q < 1$. (If $\alpha = 0$, we have $q = \infty$ and $r = 2$, so we just have the energy inequality, which is trivial.) Since

$$1 + \frac{1}{q} = 1 - \frac{1}{q} + \gamma(r) = \frac{1}{q'} + \alpha,$$

we then obtain (10.19), and this concludes the proof of (10.8).

10.8 Littlewood-Paley decomposition and conclusion of proof

Having obtained (10.8), we now scale to put the $\dot{H}^s$ norm back in the right hand side, and apply Littlewood-Paley theory to obtain (10.6). Let us write $W(t)f = e^{it\sqrt{-\Delta}}f$, so that

$$\widehat{W(t)f}(\xi) = e^{it|\xi|}\widehat{f}(\xi).$$

Choose a radial $\beta \in C_c^\infty(\mathbb{R}^n)$ supported away from zero such that[1]

$$\sum_{j\in\mathbb{Z}} \beta(\xi/2^j) = 1 \quad \text{for all} \quad \xi \neq 0.$$

Define the frequency projections Δ_j by $\widehat{\Delta_j f}(\xi) = \beta(\xi/2^j)\widehat{f}(\xi)$. Then

$$f = \sum \Delta_j f \quad \text{and} \quad W(t)f = \sum W(t)\Delta_j f.$$

It is readily checked that $W(t)\Delta_j f \simeq T[f(\cdot/2^j)](2^j t, 2^j x)$, so by (10.8),

(10.21)
$$\begin{aligned} \|W(t)\Delta_j f\|_{L_t^q(L_x^r)} &\simeq 2^{j(-n/r-1/q)} \left\|T[f(\cdot/2^j)]\right\|_{L_t^q(L_x^r)} \\ &\lesssim 2^{j(-n/r-1/q)} \left\|f(\cdot/2^j)\right\|_{L^2} = 2^{j(n/2-n/r-1/q)} \|f\|_{L^2}. \end{aligned}$$

[1]Start with a radial bump function χ such that $\chi(0) = 1$ and $\operatorname{supp}\chi \subset \{|\xi| \le 2\}$. Set $\beta(\xi) = \chi(\xi) - \chi(2\xi)$. Then $\sum_{-M}^N \beta(\xi/2^j) = \chi(\xi/2^N) - \chi(2^{M+1}\xi) \to 1$ as $M, N \to \infty$.

Now $\Delta_j \Delta_k = 0$ unless $|j-k| \le 3$, so

$$\Delta_j f = \Delta_j(\sum \Delta_k f) = \sum_{|k-j|\le 3} \Delta_j \Delta_k f.$$

Applying (10.21) then gives

$$\begin{aligned} \|W(t)\Delta_j f\|_{L^q_t(L^r_x)} &\lesssim \sum_{|k-j|\le 3} \|W(t)\Delta_j\Delta_k f\|_{L^q_t(L^r_x)} \\ &\lesssim \sum_{|k-j|\le 3} 2^{js}\|\Delta_k f\|_{L^2} \lesssim \sum_{|k-j|\le 3} \|\Delta_k f\|_{\dot H^a} \end{aligned} \tag{10.22}$$

where $s = n/2 - n/r - 1/q$. Now apply the following result from Littlewood-Paley theory:

Theorem 35. *For $2 \le p < \infty$, we have*

$$\|f\|_{L^p} \lesssim \sqrt{\sum_{j\in\mathbb{Z}} \|\Delta_j f\|^2_{L^p}}.$$

Thus,

$$\|W(t)f\|_{L^q_t(L^r_x)} \lesssim \left\| \sqrt{\sum \|W(t)\Delta_j f\|^2_{L^r_x}} \right\|_{L^q_t} \lesssim \sqrt{\sum \|W(t)\Delta_j f\|^2_{L^q_t(L^r_x)}},$$

where the last inequality follows by Minkowski's integral inequality, since $q \ge 2$. Combining this with (10.22), we conclude that (10.6) holds, since

$$\|f\|_{\dot H^s} \approx \sqrt{\sum \|\Delta_j f\|^2_{\dot H^s}}.$$

To conclude, let us remark that Theorem 35 is an immediate corollary of Minkowski's inequality (this is where we need the condition $p \ge 2$) and the following fundamental fact:

Theorem 36. *If $1 < p < \infty$, then*

$$\|f\|_{L^p} \approx \left\| \sqrt{\sum |\Delta_j f|^2} \right\|_{L^p}.$$

Remark. We conclude with the remark that if $0 < T < \infty$ and $S_T = (0,T) \times \mathbb{R}^n$, then we have the following variant of (10.4) with inhomogeneous data norms:

$$\|u\|_{L^q_t L^r_x(S_T)} \le C_T\big(\|f\|_{H^s} + \|g\|_{H^{s-1}}\big), \tag{10.23}$$

where $C_T \lesssim 1 + T^2$. This is obvious if $g = 0$, since $s \ge 0$. It is also obvious if $\widehat{g}$ is supported in $|\xi| \ge 1$. Thus, we may assume $f = 0$ and $\operatorname{supp}\widehat{g} \subset \{|\xi| \le 1\}$. Then

$$|\widehat{u}(t,\xi)| = \frac{|\sin(t\,|\xi|)|}{|\xi|}\,|\widehat{g}(\xi)| \le |t|\,|\widehat{g}(\xi)|\,,$$

and by Hölder's inequality and Sobolev embedding, we have

$$\begin{aligned}\|u\|_{L_t^q L_x^r(S_T)} \leq T^{1/q} \sup_{0\leq t\leq T} \|u(t)\|_{L^r} &\leq CT^{1/q} \sup_{0\leq t\leq T} \|u(t)\|_{\dot{H}^{n(1/2-1/r)}} \\ &\leq CT^{1+1/q} \|g\|_{L^2} \leq CT^{1+1/q} \|g\|_{H^{s-1}} ,\end{aligned}$$

proving our claim.

Chapter 11

Week 12: Application to Maxwell-Klein-Gordon

Our final objective is to prove global existence of smooth solutions to the Maxwell-Klein-Gordon equations (abbreviated MKG henceforth) on $\mathbb{R}^{1+3}$. This is a nonlinear system of equations resulting from a coupling of Maxwell's equations with a Klein-Gordon equation.

The main new tool will be bilinear generalizations of Strichartz' L^4 spacetime estimate for solutions of $\Box u = 0$ on $\mathbb{R}^{1+3}$. Once again the null condition surfaces.

11.1 Presentation of the equations

The usual conventions apply:

- Coordinates on $\mathbb{R}^{1+3}$ are denoted (t, x) or $(x^\alpha)_{\alpha=0,1,2,3}$. We write $\partial_\alpha = \partial/\partial x^\alpha$.
- Indices are raised and lowered using the Minkowski metric $\mathrm{diag}(-1, 1, 1, 1)$.
- The summation convention is in effect. Roman indices run over $1, 2, 3$ and Greek indices over $0, 1, 2, 3$.

Example. With the above conventions, $\Delta = \partial^i \partial_i$ and $\Box = \partial^\alpha \partial_\alpha$.

The MKG system is a classical field theory, derived from a variational principle. The fields involved are:

- The *electromagnetic field* $F = dA$, an exact two-form on $\mathbb{R}^{1+3}$. Here $A = A_\alpha dx^\alpha$ is a one-form, the *gauge potential*, whose components are real-valued functions $A_\alpha : \mathbb{R}^{1+3} \to \mathbb{R}$. Note that

$$F_{\alpha\beta} = \partial_\alpha A_\beta - \partial_\beta A_\alpha.$$

- A *scalar field* ϕ. This is just a function $\phi : \mathbb{R}^{1+3} \to \mathbb{C}$.

Associated to these fields we have the *Lagrangian density*

$$L = -\frac{1}{4}F^{\alpha\beta}F_{\alpha\beta} - \frac{1}{2}D^\mu\phi\overline{D_\mu\phi},$$

where $D_\mu = \partial_\mu + \sqrt{-1}A_\mu$ is the *covariant derivative*. Integrating over spacetime, we get the *action integral*

$$\mathcal{L}[A,\phi] = \int L\,dt\,dx.$$

Now consider smooth, compactly supported variations $(A^\varepsilon, \phi^\varepsilon)$ of (A,ϕ). By this we mean that $(A^\varepsilon, \phi^\varepsilon)$ depends smoothly on $\varepsilon \in \mathbb{R}$ and equals (A,ϕ) if $\varepsilon = 0$ or if (t,x) is outside some compact set. If (A,ϕ) is a stationary point for $\mathcal{L}$, that is, if

$$\frac{d}{d\varepsilon}\mathcal{L}[A^\varepsilon,\phi^\varepsilon]\big|_{\varepsilon=0} = 0$$

for all variations, then (A,ϕ) must satisfy the PDE

$$\text{(MKG)}\qquad \begin{aligned} \partial^\alpha F_{\alpha\beta} &= -\Im\left(\phi\overline{D_\beta\phi}\right), \\ D^\mu D_\mu\phi &= 0. \end{aligned}$$

We use the notation $\Re z$ and $\Im z$ for the real and imaginary parts of a complex number z.

Let us derive (MKG). We use the notation $\dot{f} = \frac{d}{d\varepsilon}f^\varepsilon\big|_{\varepsilon=0}$. Note that

$$(D^\mu\phi)^\cdot = D^\mu\dot{\phi} + \sqrt{-1}\dot{A}^\mu\phi.$$

We calculate

$$\begin{aligned} \mathrm{L}^\cdot &= -\frac{1}{2}\int\left\{F^{\alpha\beta}\dot{F}_{\alpha\beta} + (D^\mu\phi)^\cdot\,\overline{D_\mu\phi} + D^\mu\phi\overline{(D_\mu\phi)^\cdot}\right\}dt\,dx \\ &= \int\left\{-\frac{1}{2}F^{\alpha\beta}\left(\partial_\alpha\dot{A}_\beta - \partial_\beta\dot{A}_\alpha\right) - \Re\left[(D^\mu\phi)^\cdot\,\overline{D_\mu\phi}\right]\right\}dt\,dx \\ &= \int\left\{-F^{\alpha\beta}\partial_\alpha\dot{A}_\beta - \Re\left[\left(D^\mu\dot{\phi} + \sqrt{-1}\dot{A}^\mu\phi\right)\overline{D_\mu\phi}\right]\right\}dt\,dx \\ &= \int\left\{-F^{\alpha\beta}\partial_\alpha\dot{A}_\beta + \Im\left(\phi\overline{D_\beta\phi}\right)\dot{A}^\mu - \Re\left(D^\mu\dot{\phi}\overline{D_\mu\phi}\right)\right\}dt\,dx \\ &= \int\left\{\left[\partial_\alpha F^{\alpha\beta} + \Im\left(\phi\overline{D^\beta\phi}\right)\right]\dot{A}_\beta + \Re\left(\dot{\phi}\overline{D^\mu D_\mu\phi}\right)\right\}dt\,dx, \end{aligned}$$

where the last equality follows after an integration by parts. Varying A and ϕ separately then gives (MKG). In fact, given arbitrary C_c^∞ functions $\dot{A}_\alpha$ (real-valued) and $\dot{\phi}$ on $\mathbb{R}^{1+3}$, we can construct compactly supported variations simply by setting

$$A^\varepsilon = A + \varepsilon\dot{A}, \qquad \phi^\varepsilon = \phi + \varepsilon\dot{\phi}.$$

The *electric field* $\vec{E}$ and *magnetic field* $\vec{H}$ are three-vectors given by the matrix identity

$$\begin{pmatrix} 0 & F_{01} & F_{02} & F_{03} \\ & 0 & F_{12} & F_{13} \\ & & 0 & F_{23} \\ & & & 0 \end{pmatrix} = \begin{pmatrix} 0 & E_1 & E_2 & E_3 \\ & 0 & H_3 & -H_2 \\ & & 0 & H_1 \\ & & & 0 \end{pmatrix}.$$

Splitting A into its temporal and spatial components,

$$A = (A_0, \vec{A}), \qquad \vec{A} = (A_1, A_2, A_3),$$

we thus have

$$\vec{E} = \partial_t \vec{A} - \nabla A_0, \qquad \vec{H} = \operatorname{curl} \vec{A}.$$

Associated to the Lagrangian is an *energy-momentum* tensor $T_{\alpha\beta}$ satisfying $\partial_\alpha T^{\alpha\beta} = 0$. In particular,

$$\partial_0 T^{00} + \partial_i T^{i0} = 0.$$

Integrating this over $\mathbb{R}^3$ for fixed t and using the divergence theorem, we obtain, assuming sufficient decay as $|x| \to \infty$,

$$\frac{d}{dt} \int_{\mathbb{R}^3} T^{00}(t,x)\, dx = 0.$$

It turns out that

$$T^{00} = \frac{1}{2} \left(\left|\vec{E}\right|^2 + \left|\vec{H}\right|^2 + |D_0\phi|^2 + \sum_1^3 |D_i\phi|^2 \right),$$

and we define the *energy* $\mathcal{E}(t) = \int T^{00}(t,x)\, dx$. Then by the above we have conservation of energy:

$$\mathcal{E}(t) = \mathcal{E}(0)$$

for all $t > 0$, provided the solution is smooth and decays sufficently fast as $|x| \to \infty$.

11.2 Gauge ambiguity

The gauge potential A is not uniquely determined, which is a problem since we have a PDE involving A. Suppose $\chi : \mathbb{R}^{1+n} \to R$ is smooth, and consider the *gauge transformation*

$$A \to \widetilde{A} = A + d\chi,$$
$$\phi \to \widetilde{\phi} = e^{-\sqrt{-1}\chi}\phi.$$

Clearly, F is invariant under this transformation, since $d^2\chi = 0$. It is also not hard to check that if (A, ϕ) solves (MKG), then so does $(\widetilde{A}, \widetilde{\phi})$. (Observe that the covariant derivative changes when A changes!)

Because of this *gauge ambiguity*, we must understand a solution of (MKG) as an equivalence class of gauge equivalent pairs.

To fix the potential A, we stipulate an additional *gauge condition.* The traditional ones are:

- *temporal:* $A_0 = 0$,
- *Lorentz:* $\partial^\alpha A_\alpha = 0$,
- *Coulomb:* $\partial^i A_i = 0$.

11.3 MKG in Lorentz gauge

Under the Lorentz condition,

$$\partial^\alpha F_{\alpha\beta} = \partial^\alpha \left(\partial_\alpha A_\beta - \partial_\beta A_\alpha\right) = \Box A_\beta.$$

Also, $\partial^\mu(A_\mu \phi) = A_\mu \partial^\mu \phi$, whence

$$\begin{aligned} D^\mu D_\mu \phi &= \Box\phi + \sqrt{-1}\partial^\mu(A_\mu\phi) + \sqrt{-1}A^\mu\partial_\mu\phi - A^\mu A_\mu \phi \\ &= \Box\phi + 2\sqrt{-1}\partial^\mu(A_\mu\phi) - A^\mu A_\mu\phi. \end{aligned}$$

Since, moreover,

$$\phi\overline{D_\beta\phi} = \phi\overline{\partial_\beta\phi} - \sqrt{-1}A_\beta\left|\phi\right|^2,$$

we conclude that (MKG) under the Lorentz condition reduces to the following system of nonlinear wave equations:

$$\text{(MKGL)} \qquad \begin{aligned} \Box A_\alpha &= -\Im\left(\phi\overline{\partial_\beta\phi}\right) + A_\alpha\left|\phi\right|^2, \\ \Box\phi &= -2\sqrt{-1}A_\mu\partial^\mu\phi + A^\mu A_\mu\phi. \end{aligned}$$

Schematically, this is of the form, setting $\Phi = (A, \phi)$,

$$\Box\Phi = \Phi\partial\Phi + \Phi^3.$$

This equation has a simple structure, but there is a problem: In order to exploit conservation of energy and get a global existence result, we need to prove local well-posedness in the data norm $H^1 \times L^2$, but this fails to be true for generic equations of the form $\Box u = u\partial u$ on $\mathbb{R}^{1+3}$, as proved by Lindblad.

The cubic term Φ^3 is not a problem. It is easy to prove local well-posedness for $\Box u = u^3$ in $H^1 \times L^2$ by just using the energy inequality and the following Sobolev inequality:

$$\left\|f\right\|_{L^6(\mathbb{R}^3)} \le C\left\|\nabla f\right\|_{L^2(\mathbb{R}^3)}. \qquad (11.1)$$

11.4 MKG in Coulomb gauge

Splitting A into its temporal part A_0 and its spatial part $\vec{A}$ as before, the Coulomb condition says that

$$\operatorname{div} \vec{A} = 0.$$

Assuming this, we have

$$\begin{aligned} \partial^\alpha F_{\alpha 0} &= \partial^i \left(\partial_i A_0 - \partial_t A_i \right) = \Delta A_0, \\ \partial^\alpha F_{\alpha i} &= \partial^\alpha \left(\partial_\alpha A_i - \partial_i A_\alpha \right) = \Box A_i + \partial_i \partial_t A_0. \end{aligned}$$

The (MKG) system then becomes a mixed hyperbolic/elliptic system

$$\text{(MKGC)} \qquad \begin{aligned} \operatorname{div} \vec{A} &= 0, \\ \Delta A_0 &= -\Im\big(\phi\overline{\partial_t \phi}\big) + |\phi|^2 A_0, \\ \Box \vec{A} + \partial_t \nabla A_0 &= -\Im\big(\phi \overline{\nabla \phi}\big) + |\phi|^2 \vec{A}, \\ \Box \phi &= -2\sqrt{-1}\vec{A} \cdot \nabla \phi + 2\sqrt{-1} A_0 \partial_t \phi \\ &\quad + \sqrt{-1}(\partial_t A_0)\phi + \big|\vec{A}\big|^2 \phi - A_0^2 \phi. \end{aligned}$$

We study the Cauchy problem with initial data

$$\text{(11.2a)} \qquad \vec{A}\big|_{t=0} = \vec{a}, \qquad \partial_t \vec{A}\big|_{t=0} = \vec{b},$$

$$\text{(11.2b)} \qquad \phi\big|_{t=0} = \phi_0, \qquad \partial_t \phi\big|_{t=0} = \phi_1,$$

for the *dynamical variables* $(\vec{A}, \phi)$. (Then the *nondynamical* variable A_0 at $t = 0$ is uniquely determined by solving the elliptic equation in (MKGC).) In view of the Coulomb condition $\operatorname{div} \vec{A} = 0$, we must require

$$\text{(11.3)} \qquad \operatorname{div} \vec{a} = \operatorname{div} \vec{b} = 0.$$

Observe that the equations for the dynamical variables are of the form

$$\begin{aligned} \Box \vec{A} &= -\Im\left(\phi\overline{\nabla \phi}\right) + \mathfrak{C} + \mathfrak{E}, \\ \Box \phi &= -2\sqrt{-1}\vec{A} \cdot \nabla \phi + \mathfrak{C} + \mathfrak{E}, \end{aligned}$$

where $\mathfrak{E}$ denotes terms involving A_0 or $\partial_t A_0$, and $\mathfrak{C}$ denotes cubic terms involving $\vec{A}$ and ϕ.

The terms in $\mathfrak{C}$ and $\mathfrak{E}$ will be relatively easy to handle. The most important terms are $\phi\overline{\nabla\phi}$ and $\vec{A}\cdot\nabla\phi$. Recall that this type of expression is what caused the problems in the Lorentz gauge, so it seems we have gained nothing by going from Lorentz to Coulomb (we have only made the system a lot more complicated, it would seem).

The remarkable fact, however, is that the terms $\phi\overline{\nabla\phi}$ and $\vec{A} \cdot \nabla\phi$ can be expressed, due the Coulomb condition $\operatorname{div} \vec{A} = 0$, in terms of the null forms

$$Q_{ij}(\partial u, \partial v) = \partial_i u \partial_j v - \partial_j u \partial_i v, \qquad 1 \le i, j \le 3.$$

This is discussed next.

11.5 The null structure

Let $\mathcal{P}$ be the projection onto the divergence-free vector fields on $\mathbb{R}^3$:

$$\mathcal{P} = (-\Delta)^{-1} \operatorname{curl} \operatorname{curl}.$$

To motivate this, recall the vector identity

$$\operatorname{curl} \operatorname{curl} \vec{X} = \operatorname{grad} \operatorname{div} \vec{X} - \Delta \vec{X}. \tag{11.4}$$

Thus

$$\vec{X} = (-\Delta)^{-1} \operatorname{curl} \operatorname{curl} \vec{X} - (-\Delta)^{-1} \operatorname{grad} \operatorname{div} \vec{X}.$$

Since $\operatorname{div} \operatorname{curl} = 0$ and $\operatorname{curl} \operatorname{grad} = 0$, this expresses $\vec{X}$ as the sum of its divergence-free and curl-free parts. It also shows that

$$\operatorname{div} \vec{X} = 0 \implies \mathcal{P}\vec{X} = \vec{X}.$$

Thus, if we apply $\mathcal{P}$ to the equation for $\Box \vec{A}$ in (MKGC), we get

$$\Box \vec{A} = -\Im \mathcal{P} \left(\phi \overline{\nabla \phi} \right) + \mathcal{P} \left(|\phi|^2 \vec{A} \right).$$

The term $\partial_t \nabla \vec{A}$ disappears because $\operatorname{curl} \operatorname{grad} = 0$.

To state the key lemma, we need some definitions. We write $|D|^\gamma = (-\Delta)^{\gamma/2}$. The *Riesz transforms* are defined by $R_i = |D|^{-1} \partial_i$. Note that the Fourier symbol of R_i is $\xi_i / |\xi|$, modulo a multiplicative constant. Thus, R_i is bounded on every H^s. In fact, R_i is bounded on every L^p, $1 < p < \infty$, but this is a much deeper fact. From the identity (11.4) and the fact that $R^j R_j = -\mathrm{Id}$, we see that

$$(\mathcal{P}\vec{X})^i = X^i + R^i R_j X^j = R_j (R^i X^j - R^j X^i). \tag{11.5}$$

In view of the above remarks about the boundedness of R_i, this implies, in particular, that $\mathcal{P}$ is bounded on every H^s (and on every L^p with $1 < p < \infty$).

We now state the main result of this section.

Lemma 12. *We have the identities*

(i) $\mathcal{P}(u \nabla v)_i = R^j |D|^{-1} Q_{ij}(\partial u, \partial v)$.

(ii) $2 \nabla u \cdot \mathcal{P}\vec{X} = Q_{ij}\big(\partial u, |D|^{-1} [R^i \partial X^j - R^j \partial X^i]\big)$.

Proof. (i) says that

$$\operatorname{curl} \operatorname{curl} (u \nabla v)_i = \partial^j (\partial_i u \partial_j v - \partial_j u \partial_i v).$$

To prove this, recall the vector identities

$$\operatorname{curl}(u \nabla v) = \nabla u \times \nabla v,$$
$$\operatorname{curl}(\vec{X} \times \vec{Y}) = -(\vec{X} \cdot \nabla)\vec{Y} + (\vec{Y} \cdot \nabla)\vec{X} + (\operatorname{div} \vec{Y})\vec{X} - (\operatorname{div} \vec{X})\vec{Y}.$$

Thus

$$\begin{aligned}\operatorname{curl}\operatorname{curl}(u\nabla v)_i &= \operatorname{curl}(\nabla u \times \nabla v)_i \\ &= -\partial^j u \partial_j \partial_i v + \partial^j v \partial_j \partial_i u + \Delta v \partial_i u - \Delta u \partial_i v \\ &= \partial^j(\partial_i u \partial_j v - \partial_j u \partial_i v),\end{aligned}$$

as desired.

To prove (ii), we use (11.5) to write

$$2\nabla u \cdot \mathcal{P}\vec{X} = 2\partial_i u R_j (R^i X^j - R^j X^i). \tag{11.6}$$

But expanding the right hand side of (ii) gives

$$\partial_i u R_j (R^i X^j - R^j X^i) - \partial_j u R_i (R^i X^j - R^j X^i)$$

which is seen to equal (11.6) after relabeling. □

From this lemma we immediately obtain:

Corollary. *(i)* $\mathcal{P}\left(\phi\overline{\nabla\phi}\right)_i = 2R^j \left|D\right|^{-1} Q_{ij}(\partial\Re\phi, \partial\Im\phi)$.

(ii) If $\operatorname{div}\vec{A} = 0$, *and hence* $\mathcal{P}\vec{A} = \vec{A}$, *then*

$$2\vec{A}\cdot\nabla\phi = Q_{ij}\big(\partial\phi, \left|D\right|^{-1}[R^i\partial A^j - R^j \partial A^i]\big).$$

11.6 Reformulation of MKG in Coulomb gauge

Using the corollary to Lemma 12, we obtain the following equivalent formulation of (MKGC):

$$\Delta A_0 = -\Im\big(\phi\overline{\partial_t\phi}\big) + \left|\phi\right|^2 A_0, \tag{11.7a}$$

$$\Delta\partial_t A_0 = -\Im\operatorname{div}\big(\phi\overline{\nabla\phi}\big) + \operatorname{div}\big(\left|\phi\right|^2 \vec{A}\big) \tag{11.7b}$$

$$\Box A_i = 2R^j\left|D\right|^{-1} Q_{ij}(\partial\Re\phi, \partial\Im\phi) + \mathcal{P}\big(\left|\phi\right|^2 A_i\big) \tag{11.7c}$$

$$\begin{aligned}\Box\phi &= -\sqrt{-1}Q_{ij}\left(\phi, \left|D\right|^{-1}\left[R^i\partial A^j - R^j\partial A^i\right]\right) \\ &\quad + 2\sqrt{-1}A_0\partial_t\phi + \sqrt{-1}(\partial_t A_0)\phi + \big|\vec{A}\big|^2\phi - A_0^2\phi.\end{aligned} \tag{11.7d}$$

In fact, as noted already, applying $\mathcal{P}$ to the equation for $\Box\vec{A}$ in (MKGC) gives (11.7c). To get (11.7b), apply div to the same equation, using the Coulomb condition $\operatorname{div}\vec{A} = 0$. This proves one half of the following:

Lemma 13. *The systems* (MKGC) *and* (11.7) *are equivalent for initial data* (11.2) *satisfying* (11.3).

Conversely, to prove that (11.7) implies (MKGC), first observe that by formula (a) in the corollary to Lemma 12, (11.7c) just says that

$$\Box \vec{A} = \mathcal{P}\left(\phi \overline{\nabla \phi} + |\phi|^2 \vec{A}\right). \tag{11.7c$'$}$$

Thus $\Box \operatorname{div} \vec{A} = 0$, and since the initial data of $\vec{A}$ are divergence free, it follows by uniqueness for solutions of the homogeneous wave equation that the Coulomb condition $\operatorname{div} \vec{A} = 0$ holds for all time. Then, in view of formula (b) in the corollary to Lemma 12, we see that (11.7d) is equivalent to the equation for $\Box\phi$ in (MKGC). Finally, from (11.7c′) we have

$$\mathcal{P}\left(\Box \vec{A} + \partial_t \nabla A_0 + \Im\big(\phi \overline{\nabla \phi}\big) - |\phi|^2 \vec{A}\right) = 0.$$

But in view of (11.7b),

$$\operatorname{div}\left(\Box \vec{A} + \partial_t \nabla A_0 + \Im\big(\phi \overline{\nabla \phi}\big) - |\phi|^2 \vec{A}\right) = 0,$$

and we conclude that

$$\Box \vec{A} + \partial_t \nabla A_0 + \Im\big(\phi \overline{\nabla \phi}\big) - |\phi|^2 \vec{A} = 0,$$

which is exactly the equation for $\Box \vec{A}$ in (MKGC).

11.7 The main result

Our aim is to prove the following:

Theorem 37. *If the data* (11.2) *belong to* $C_c^\infty(\mathbb{R}^3)$ *and satisfy* (11.3)*, then* (MKGC)*, or equivalently* (11.7)*, has a unique smooth solution*

$$(A_0, \vec{A}, \phi) \in C^\infty(\mathbb{R}^{1+3}_+).$$

The strategy for proving this is as follows:

- Prove local well-posedness for data in $H^1 \times L^2$.
- Use conservation of energy to deduce that the lifespan $= +\infty$.

We would like to obtain a local solution of (11.7) by iterating in the energy space

$$E_T = C([0,T], H^1) \cap C^1([0,T], L^2).$$

To do this, we first eliminate the nondynamical variables by solving the elliptic equations (11.7a) and (11.7b).

Definition. We denote by $A_0(\phi)$ and $B_0(\vec{A}, \phi)$ the solutions of (11.7a) and (11.7b), respectively.

Remark. One can show (in fact this essentially follows from estimates we prove later) that (11.7a) [resp. (11.7b)] has a unique solution in $\dot{H}^1$ [resp. L^2] for every fixed time t, provided $\vec{A}, \phi \in E_T$. The above definition therefore makes sense.

Replacing A_0 and $\partial_t A_0$ by the nonlinear operators $A_0(\phi)$ and $B_0(\vec{A}, \phi)$ in equation (11.7d), we obtain a system of nonlinear wave equations for the dynamical variables:

$$\begin{aligned} \square \vec{A} &= \mathcal{M}(\vec{A}, \phi), \\ \square \phi &= \mathcal{N}(\vec{A}, \phi), \end{aligned} \tag{11.8}$$

where

$$\mathcal{M}(\vec{A}, \phi)_i = 2R^j \left|D\right|^{-1} Q_{ij}(\partial \Re\phi, \partial \Im\phi) + \mathcal{P}\big(\left|\phi\right|^2 A_i\big), \qquad 1 \le i \le 3, \tag{11.9}$$

$$\begin{aligned} \mathcal{N}(\vec{A}, \phi) = &-\sqrt{-1} Q_{ij} \left(\phi, \left|D\right|^{-1} \left[R^i \partial A^j - R^j \partial A^i\right]\right) \\ &+ 2\sqrt{-1} A_0(\phi) \partial_t \phi + \sqrt{-1} B_0(\vec{A}, \phi)\phi + \left|\vec{A}\right|^2 \phi - A_0(\phi)^2 \phi. \end{aligned} \tag{11.10}$$

We shall then prove:

Theorem 38. *The system* (11.8) *is locally well-posed for data in* $H^1 \times L^2$.

In fact, we will only prove local existence, but with a little more work one can show uniqueness, continuous dependence on data and persistence of higher regularity; in particular, smooth data gives a smooth solution.

Corollary. *The system* (11.7)*, and hence also* (MKGC)*, is LWP for data* (11.2) *in* $H^1 \times L^2$ *satisfying* (11.3).

Let us merely sketch the proof of this corollary. Assuming $(\vec{A}, \phi)$ solves (11.8), we define $A_0 = A_0(\phi)$. Then one shows that

$$\partial_t A_0 = B_0(\vec{A}, \phi)$$

in the sense of distributions, and the corollary follows. To prove the last identity, one shows by a straightforward calculation that

$$\Delta(\partial_t A_0 - B_0) = \left|\phi\right|^2 (\partial_t A_0 - B_0),$$

and use the fact that the the unique solution in $\dot{H}^1(\mathbb{R}^3)$ of the nonlinear elliptic equation $\Delta u = \left|\phi\right|^2 u$ is $u = 0$. This argument works if the data are sufficiently smooth, say C_c^∞, and for general data one chooses smooth approximating sequences and exploit persistence of higher regularity and continuous dependence on the data to pass to the limit.

We postpone the proof of the local existence statement of Theorem 38 and consider the next step, namely how to exploit energy conservation to see that the lifespan is infinite.

11.8 Data norm controlled by energy

Here we combine Theorem 38, or rather its corollary, with energy conservation, to obtain global existence.

Suppose $0 < T < \infty$ and $(A_0, \vec{A}, \phi) \in C^\infty([0,T) \times \mathbb{R}^3)$ solves (MKGC) with C_c^∞ data (11.2). We claim that

$$\sup_{0 \le t < T} \left(\left\| \vec{A}(t) \right\|_{H^1} + \left\| \partial_t \vec{A}(t) \right\|_{L^2} + \|\phi(t)\|_{H^1} + \|\partial_t \phi(t)\|_{L^2} \right) < \infty. \tag{11.11}$$

It then follows by the corollary to Theorem 38 that the solution extends beyond time T, and we conclude that the lifespan $= +\infty$.

Let us prove the claim. Recall that the energy

$$\mathcal{E}(t) = \frac{1}{2} \int \left(\left|\vec{E}\right|^2 + \left|\vec{H}\right|^2 + |D_0\phi|^2 + \sum_1^3 |D_i\phi|^2 \right) dx$$

is conserved:

$$\mathcal{E}(t) = \mathcal{E}(0) \quad \text{for} \quad 0 \le t < T. \tag{11.12}$$

We have to control the $L^2(\mathbb{R}^3)$ norms of $\vec{A}$, $\nabla\vec{A}$, $\partial_t\vec{A}$, ϕ, $\nabla\phi$ and $\partial_t\phi$ uniformly in $0 \le t < T$.

Estimate for $\left\|\nabla\vec{A}\right\|_{L^2}$. Since $\vec{H} = \operatorname{curl}\vec{A}$ and $\operatorname{div}\vec{A} = 0$, (11.4) implies

$$\operatorname{curl}\vec{H} = \operatorname{curl}\operatorname{curl}\vec{A} = -\Delta\vec{A}.$$

Hence, using Plancherel's theorem,

$$\left\|\nabla\vec{A}(t)\right\|_{L^2} \le C\left\|\vec{H}(t)\right\|_{L^2} \le C\sqrt{\mathcal{E}(t)} = C\sqrt{\mathcal{E}(0)} \tag{11.13}$$

Estimate for $\left\|\partial_t\vec{A}\right\|_{L^2}$. Since $\vec{E} = \partial_t\vec{A} - \nabla A_0$, we have

$$\operatorname{curl}\operatorname{curl}\vec{E} = \partial_t \operatorname{curl}\operatorname{curl}\vec{A} = -\partial_t\Delta\vec{A} = -\Delta\partial_t\vec{A},$$

whence $\mathcal{P}\vec{E} = \partial_t\vec{A}$. Consequently,

$$\left\|\partial_t\vec{A}(t)\right\|_{L^2} \le C\left\|\vec{E}(t)\right\|_{L^2} \le C\sqrt{\mathcal{E}(0)}. \tag{11.14}$$

Estimate for $\left\|\vec{A}\right\|_{L^2}$. Observe that

$$\frac{d}{dt}\int \frac{1}{2}\left|\vec{A}(t,x)\right|^2 dx = \int \vec{A}\cdot\partial_t\vec{A}\, dx \le \left\|\vec{A}(t)\right\|_{L^2}\left\|\partial_t\vec{A}(t)\right\|_{L^2}.$$

Thus, $(d/dt)\left\|\vec{A}(t)\right\|_{L^2} \le \left\|\partial_t\vec{A}(t)\right\|_{L^2}$ whenever $\vec{A}(t,\cdot) \ne 0$, whence

$$\left\|\vec{A}(t)\right\|_{L^2} \le \left\|\vec{A}(0)\right\|_{L^2} + \int_0^t \left\|\partial_t\vec{A}(t')\right\|_{L^2} dt' \le \|\vec{a}\|_{L^2} + Ct\sqrt{\mathcal{E}(0)}. \tag{11.15}$$

Estimate for $\|\phi\|_{L^2}$. Since

$$\frac{d}{dt}\int \frac{1}{2}|\phi|^2\,dx = \Re\int \phi\overline{\partial_t\phi}\,dx = \Re\int \phi\overline{D_0\phi}\,dx \le \|\phi\|_{L^2}\sqrt{\mathcal{E}(0)},$$

we conclude that

$$\|\phi(t)\|_{L^2} \le \|\phi_0\|_{L^2} + Ct\sqrt{\mathcal{E}(0)}. \tag{11.16}$$

Estimate for $\|\partial_t\phi\|_{L^2}$. Since $\partial_t\phi = D_0\phi - \sqrt{-1}A_0\phi$, we have

$$\|\partial_t\phi\|_{L^2} \le \|D_0\phi\|_{L^2} + \|A_0\phi\|_{L^2} \le \sqrt{\mathcal{E}(0)} + \|A_0\|_{L^6}\|\phi\|_{L^3}. \tag{11.17}$$

Taking the divergence of $\vec{E} = \partial_t\vec{A} - \nabla A_0$ gives $\operatorname{div}\vec{E} = -\Delta A_0$, and using (11.1) we conclude that

$$\|A_0\|_{L^6} \le C\|\nabla A_0\|_{L^2} \le C\|\vec{E}\|_{L^2} \le C\sqrt{\mathcal{E}(0)}. \tag{11.18}$$

It remains to estimate $\|\phi\|_{L^3}$. By Hölder's inequality,

$$\|\phi\|_{L^3} = \left\|\phi^2\right\|_{L^{3/2}}^{\frac{1}{2}} \le \|\phi\|_{L^2}^{\frac{1}{2}}\|\phi\|_{L^6}^{\frac{1}{2}} \le C\left(\|\phi_0\|_{L^2} + t\sqrt{\mathcal{E}(0)}\right)^{\frac{1}{2}}\|\nabla\phi\|_{L^2}^{\frac{1}{2}}, \tag{11.19}$$

where we used (11.16) and (11.1) to get the last inequality. Now

$$\|\partial_i\phi\|_{L^2} \le \|D_i\phi\|_{L^2} + C\|\nabla A_i\|_{L^2}\|\phi\|_{L^3} \le C\sqrt{\mathcal{E}(0)}\,(1 + \|\phi\|_{L^3}),$$

where we used (11.1) and (11.13). This together with (11.19) gives

$$\|\phi\|_{L^3} \le C\sqrt{(1+t)(1+\|\phi\|_{L^3})},$$

where C depends on the initial data. Since the last inequality clearly continues to hold if we add 1 to the left hand side, we conclude that

$$\sqrt{1+\|\phi\|_{L^3}} \le C\sqrt{1+t} \implies \|\phi\|_{L^3} \le C(1+t).$$

Combining this with (11.17) and (11.18) we get

$$\|\partial_t\phi\|_{L^2} \le C(1+t),$$

where again C depends on the data. Since the above also shows that the L^2 norm of $\nabla\phi$ is under control, the proof of (11.11) is complete.

11.9 Local existence

Here we prove that (11.8) has a local solution for initial data in $H^1 \times L^2$. We employ the usual iteration scheme: Set $\vec{A}_{-1}, \phi_{-1} \equiv 0$ and define $\vec{A}_j$ and ϕ_j inductively for $j \ge 0$ by

$$\begin{aligned}\Box\vec{A}_j &= \mathcal{M}(\vec{A}_{j-1}, \phi_{j-1}),\\ \Box\phi_{j+1} &= \mathcal{N}(\vec{A}_{j-1}, \phi_{j-1}),\end{aligned}$$

with initial data (11.2).

We start by obtaining estimates for the iterates in the space

$$E_T = C([0,T],H^1) \cap C^1([0,T],L^2)$$

with norm

$$E_T(u) = \sup_{0 \le t \le T} \left(\|u(t)\|_{H^1} + \|\partial_t u(t)\|_{L^2} \right).$$

By the energy inequality, we have[1]

$$E_T(u) \le CE_0(u) + C\int_0^T \|\Box u(t)\|_{L^2}\, dt,$$

so we need to control $\left\|\mathcal{M}(\vec{A}_j,\phi_j)\right\|_{L^1_t L^2_x(S_T)}$ and $\left\|\mathcal{N}(\vec{A}_j,\phi_j)\right\|_{L^1_t L^2_x(S_T)}$. We divide the terms we need to estimate into three categories:

(i) ***Bilinear in $\vec{A}$ and ϕ.*** There are two terms of this type:

$$|D|^{-1} Q(\Re\partial\phi, \Im\partial\phi), \tag{11.20}$$

$$Q(\partial\phi, |D|^{-1}\partial\vec{A}). \tag{11.21}$$

Here we ignore the Riesz operators, which is justified since these are bounded on L^2.

(i) ***Elliptic terms.*** That is, terms containing A_0 or B_0. There are three terms of this type:

$$A_0(\phi)\partial_t\phi, \tag{11.22}$$

$$[A_0(\phi)]^2\phi, \tag{11.23}$$

$$B_0(\vec{A},\phi)\phi. \tag{11.24}$$

(i) ***Cubic terms in $\vec{A}$ and ϕ.*** There are two terms of this type:

$$|\phi|^2 \vec{A}, \tag{11.25}$$

$$\left|\vec{A}\right|^2 \phi. \tag{11.26}$$

Here we ignore the projection $\mathcal{P}$ acting on (11.25). Again this is justified because $\mathcal{P}$ is bounded on L^2.

11.10 Estimates for cubic terms

By Hölder's inequality,

$$\left\||\phi|^2 \vec{A}\right\|_{L^1_t L^2_x(S_T)} \le T\left\||\phi|^2 \vec{A}\right\|_{L^\infty_t L^2_x(S_T)} \le T\|\phi\|^2_{L^\infty_t L^6_x(S_T)} \left\|\vec{A}\right\|_{L^\infty_t L^6_x(S_T)}.$$

[1] Here the constant C grows linearly as $T \to \infty$, but since we are interested in a local existence result, we can assume $T \le 1$, say.

Now apply the Sobolev inequality (11.1) to get

$$\|\phi\|_{L_t^\infty L_x^6(S_T)} \le C \|\nabla\phi\|_{L_t^\infty L_x^2(S_T)} \le CE_T(\phi),$$

and similarly for $\vec{A}$. Thus

$$\big\| |\phi|^2 \vec{A} \big\|_{L_t^1 L_x^2(S_T)} \le CTE_T(\vec{A},\phi)^3. \tag{11.27}$$

The term (11.26) is treated in the same way, yielding

$$\Big\| \left|\vec{A}\right|^2 \phi \Big\|_{L_t^1 L_x^2(S_T)} \le CTE_T(\vec{A},\phi)^3. \tag{11.28}$$

11.11 Estimates for elliptic terms

We first prove some basic estimates for A_0.

Definition. Let $\dot{H}^1(\mathbb{R}^3)$ be the Hilbert space such that the Fourier transform $\mathcal{F}$ maps $\dot{H}^1$ unitarily onto $L^2(|\xi|^2\, d\xi)$.

In other words, we are using the fact that $L^2(|\xi|^2\, d\xi) \subset L^1_{\mathrm{loc}}(\mathbb{R}^3) \subset \mathcal{S}'(\mathbb{R}^3)$, and defining

$$\dot{H}^1 = \mathcal{F}^{-1}\left[L^2(|\xi|^2\, d\xi)\right].$$

The norm on this space is $\|f\|_{\dot{H}^1} = \big\| |\xi|\, \widehat{f}(\xi) \big\|_{L^2} \simeq \|\nabla f\|_{L^2}$. Observe that $\mathcal{S}$ is dense in $L^2(|\xi|^2\, d\xi)$, hence also in $\dot{H}^1$.

Now consider the elliptic equation

$$\Delta u - |\phi|^2 u = -\Im(\phi\psi). \tag{11.29}$$

Lemma 14. *If $\phi \in H^1$ and $\psi \in L^2$, then (11.29) has a unique solution $u \in \dot{H}^1$. Moreover, u is real valued and satisfies the estimates*

(i) $\|\nabla u\|_{L^2} + \|u\phi\|_{L^2} \le C \|\psi\|_{L^2}$,

(ii) $\|u\|_{L^\infty} \le C \|\psi\|_{L^2} (1 + \|\phi\|_{L^8})$ *if* $\phi \in L^8$.

Proof. Assume $u \in \dot{H}^1$ satisfies (11.29) in the sense of distributions. By definition, this means that

$$\int \nabla u \cdot \overline{\nabla v} + |\phi|^2 u\overline{v}\, dx = \Im \int \overline{v}\phi\overline{\psi}\, dx \tag{11.30}$$

for all $v \in \mathcal{S}$. By density, this identity then holds for all $v \in \dot{H}^1$. Taking $v = u$, we have

$$\int |\nabla u|^2 + |\phi u|^2\, dx = \Im \int \overline{u}\phi\overline{\psi}\, dx$$

Applying Hölder's inequality on the right hand side, we get

$$\|u\|_{L^2}^2 + \|u\phi\|_{L^2}^2 \le \|u\phi\|_{L^2} \|\psi\|_{L^2}.$$

Setting $N = \|u\|_{L^2} + \|u\phi\|_{L^2}$, and using the fact that $2ab \le a^2 + b^2$, and hence $(a+b)^2 \le 2(a^2+b^2)$, we conclude that

$$N^2 \le 2N \|\psi\|_{L^2} \implies N \le 2\|\psi\|_{L^2},$$

which proves (a). The above argument also shows existence and uniqueness in $\dot{H}^1$. In fact, the left hand side of (11.30) defines an inner product on $\dot{H}^1$ whose associated norm is equivalent to the one associated to the standard inner product $\int \nabla u \cdot \overline{\nabla v}\, dx$, since

$$\int |\phi|^2 u\overline{v}\, dx \le \|\phi\|_{L^2} \|\phi\|_{L^6} \|u\|_{L^6} \|v\|_{L^6} \le C\|\phi\|_{H^1}^2 \|u\|_{\dot{H}^1} \|v\|_{\dot{H}^1}$$

where we used Hölder's inequality and (11.1). Moreover, the right hand side of (11.30) defines a bounded linear functional on $\dot{H}^1$, since

$$\left| \Im \int \overline{v}\phi\overline{\psi}\, dx \right| \le \|\phi\|_{L^3} \|\psi\|_{L^2} \|v\|_{L^6} \le C\|v\|_{\dot{H}^1},$$

where

$$C = \|\phi\|_{L^2}^{\frac{1}{2}} \|\phi\|_{L^6}^{\frac{1}{2}} \|\psi\|_{L^2} \le C\|\phi\|_{H^1}^2 \|\psi\|_{L^2}.$$

By Riesz' Representation Theorem it follows that there is a unique $u \in \dot{H}^1$ satisfying (11.30) for all $v \in \dot{H}^1$. It follows from (11.29) that $\Im u$ solves the same equation with $\psi = 0$, so by the estimate in (a), we must have $\Im u = 0$.

We now prove (b). We apply the following estimate (see remark following the proof) valid on $\mathbb{R}^3$ for any $\varepsilon > 0$,

$$\|u\|_{L^\infty} \le C_\varepsilon \left(\|\Delta u\|_{L^{3/2+\varepsilon}} + \|u\|_{L^6}\right). \tag{11.31}$$

Taking $\varepsilon = 1/10$ and noting that $3/2 + 1/10 = 8/5$ and $5/8 = 1/2 + 1/8$,

$$\|\Delta u\|_{L^{3/2+\varepsilon}} \le \left\||\phi|^2 u\right\|_{L^{8/5}} + \|\phi\psi\|_{L^{8/5}} \le \|\phi u\|_{L^2} \|\phi\|_{L^8} + \|\phi\|_{L^8} \|\psi\|_{L^2}.$$

Combining this with (11.31), the Sobolev inequality (11.1) and estimate (a), we conclude that

$$\|u\|_{L^\infty} \le C\|\psi\|_{L^2}(1 + \|\phi\|_{L^8}).$$

□

Remark. To prove (11.31), we can apply Sobolev's Lemma, concluding that

$$\|u\|_{L^\infty} \le C_\delta \left\|(I-\Delta)^{1/4+\delta/2} u\right\|_{L^6} \qquad (\delta > 0).$$

Now use the following fact (see Stein, *Singular integrals and diffentiability properties of functions*, Lemma 2(ii) in Chapter V):

Lemma 15. *For $s > 0$ and $1 \le p \le \infty$,*

$$\left\|(I-\Delta)^{s/2} u\right\|_{L^p} \le C_s \left(\|u\|_{L^p} + \left\|(-\Delta)^{s/2} u\right\|_{L^p}\right).$$

Thus, it only remains to see that

$$\left\|(-\Delta)^{1/4+\delta/2} u\right\|_{L^6} \le C_\varepsilon \|\Delta u\|_{L^{3/2+\varepsilon}},$$

but this follows by Sobolev embedding, if we choose $\delta = 2 - 6/(3 + 2\varepsilon)$.

11.11.1 Estimate for (11.22).

Applying (b) of Lemma 14, we see that

$$\|A_0\partial_t\phi\|_{L^1_tL^2_x(S_T)} \le \|A_0\|_{L^1_tL^\infty_x(S_T)} \|\partial_t\phi\|_{L^\infty_tL^2_x(S_T)} \le CE_T(\phi)^2\left(T+\|\phi\|_{L^1_tL^8_x(S_T)}\right) \le CE_T(\phi)^2\left(T+T^{7/8}\|\phi\|_{L^8(S_T)}\right). \tag{11.32}$$

To estimate $\|\phi\|_{L^8(S_T)}$, we apply the following corollary to the Strichartz estimates:

Theorem 39. *If (q,r) is wave admissible for $n=3$, that is, if*

$$2\le q\le\infty,\quad 2\le r<\infty,\quad \frac{2}{q}\le 1-\frac{2}{r},$$

then setting $s=3/2-3/r-1/q$, we have the estimate

$$\|u\|_{L^q_tL^r_x(S_T)} \le C\left(\|u(0)\|_{H^s}+\|\partial_t u(0)\|_{H^{s-1}}+\int_0^T \|\Box u(t)\|_{H^{s-1}}\,dt\right).$$

(Here C increases with T, but since we assume $T\le 1$, this is not an issue.)

Proof. Write $u=u_0+v$ where $\Box u_0=0$ with the same data as u at $t=0$, and $\Box v=\Box u$ with vanishing data. Then by Duhamel's principle,

$$v(t)=-\int_0^t W(t-t')\Box u(t')\,dt',$$

and applying the Strichartz estimate

$$\|W(t-t_0)g\|_{L^q_tL^r_x(S_T)}\le C\|g\|_{H^{s-1}},$$

where C is independent of t_0, we obtain, using Minkowski's integral inequality,

$$\begin{aligned}\|v\|_{L^q_tL^r_x(S_T)} &\le \int \left\|\chi_{\{0<t'<t<T\}}W(t-t')\Box u(t')\right\|_{L^q_tL^r_x(S_T)}\,dt'\\ &\le \int_0^T \|W(t-t')\Box u(t')\|_{L^q_tL^r_x(S_T)}\,dt'\\ &\le C\int_0^T \|\Box u(t')\|_{H^{s-1}}\,dt',\end{aligned}$$

which together with the Strichartz estimate for u_0 proves the theorem. ☐

Applying this theorem with $q=r=8$, we have $s=1$, hence

$$\|\phi\|_{L^8(S_T)}\le CE_0(\phi)+C\int_0^T\|\Box\phi(t)\|_{L^2}\,dt.$$

Combining this with (11.32) gives

$$\|A_0\partial_t\phi\|_{L^1_tL^2_x(S_T)}\le C\sqrt{T}E_T(\phi)^2\left(E_0(\phi)+\int_0^T\|\Box\phi\|_{L^2}\,dt\right). \tag{11.33}$$

11.11.2 Estimate for (11.23).

We have

$$\begin{aligned}\left\|A_0^2\phi\right\|_{L_t^1 L_x^2(S_T)} &\le T\left\|A_0^2\phi\right\|_{L_t^\infty L_x^2(S_T)} \\ &\le T\left\|A_0\right\|_{L_t^\infty L_x^6(S_T)}^2 \left\|\phi\right\|_{L_t^\infty L_x^6(S_T)} \le CT\left\|\nabla A_0\right\|_{L_t^\infty L_x^2(S_T)}^2 \left\|\nabla\phi\right\|_{L_t^\infty L_x^2(S_T)},\end{aligned}$$

and since $\|\nabla A_0\|_{L_t^\infty L_x^2(S_T)} \le CE_T(\phi)$ by Lemma 14(a), we conclude that

$$\left\|A_0^2\phi\right\|_{L_t^1 L_x^2(S_T)} \le CTE_T(\phi)^2. \tag{11.34}$$

11.11.3 Estimate for (11.24).

Fix t. Recall that B_0 is the solution of

$$\Delta B_0 = -\Im \operatorname{div}\big(\phi\overline{\nabla\phi}\big) + \operatorname{div}\big(|\phi|^2 \vec{A}\big).$$

Applying the Sobolev embedding $L^{6/5} \subset \dot{H}^{-1}$, which is the dual of (11.1), we get

$$\|B_0\|_{L^2} \le C\left(\|\phi\nabla\phi\|_{L^{6/5}} + \left\|\phi^2\vec{A}\right\|_{L^{6/5}}\right). \tag{11.35}$$

Now $\|\phi\nabla\phi\|_{L^{6/5}} \le \|\phi\|_{L^3}\|\nabla\phi\|_{L^2}$, and

$$\|\phi\|_{L^3} \le \|\phi\|_{L^2}^{\frac{1}{2}}\|\phi\|_{L^6}^{\frac{1}{2}} \le C\|\phi\|_{L^2}^{\frac{1}{2}}\|\nabla\phi\|_{L^2}^{\frac{1}{2}},$$

whence

$$\|\phi\nabla\phi\|_{L^{6/5}} \le CE_T(\phi)^2. \tag{11.36}$$

On the other hand,

$$\left\|\phi^2\vec{A}\right\|_{L^{6/5}} \le \|\phi\|_{L^6}^2\left\|\vec{A}\right\|_{L^2} \le C\|\nabla\phi\|_{L^2}^2\left\|\vec{A}\right\|_{L^2}.$$

Combining this with (11.35) and (11.36), and integrating in time, we get

$$\|B_0\|_{L_t^1 L_x^2(S_T)} \le CTE_T(\phi)^2\left(1 + E_T(\vec{A})\right). \tag{11.37}$$

11.12 Bilinear estimates

We shall prove the following estimates for the bilinear terms (11.20) and (11.21):

$$\begin{aligned}\left\||D|^{-1}Q(\partial u,\partial v)\right\|_{L_t^1 L_x^2(S_T)} &\le C\sqrt{T}\left(E_0(u) + \|\Box u\|_{L_t^1 L_x^2(S_T)}\right) \\ &\quad\times\left(E_0(v) + \|\Box v\|_{L_t^1 L_x^2(S_T)}\right),\end{aligned} \tag{11.38}$$

$$\begin{aligned}\left\|Q(|D|^{-1}\partial u,\partial v)\right\|_{L_t^1 L_x^2(S_T)} &\le C\sqrt{T}\left(E_0(u) + \|\Box u\|_{L_t^1 L_x^2(S_T)}\right) \\ &\quad\times\left(E_0(v) + \|\Box v\|_{L_t^1 L_x^2(S_T)}\right),\end{aligned} \tag{11.39}$$

where Q can be any of the null forms Q_{ij}, $1 \le i, j \le 3$.

Let us postpone the proofs of these until the end, and finish the iteration argument.

11.13 Modification of the iteration space

We were hoping to be able to iterate in the energy space E_T, but in view of (11.33) and the bilinear estimates stated above, we have to modify the space. Thus, we define a new norm

$$X_T(u) = E_T(u) + \|\Box u\|_{L_t^1 L_x^2(S_T)}, \tag{11.40}$$

and we let X_T be the corresponding subspace of E_T.

Applying the energy inequality and then using the estimates (11.27), (11.28), (11.33), (11.34), (11.37), (11.38) and (11.39), we obtain

(11.41)
$$\begin{aligned} X_T(\vec{A}_{j+1}, \phi_{j+1}) &\le CE_0 + C \int_0^T \left(\left\|\Box \vec{A}_{j+1}\right\|_{L^2} + \|\Box \phi_{j+1}\|_{L^2} \right) dt \\ &\le CE_0 + C \int_0^T \left(\left\|\mathcal{M}(\vec{A}_j, \phi_j)\right\|_{L^2} + \left\|\mathcal{N}(\vec{A}_j, \phi_j)\right\|_{L^2} \right) dt \\ &\le CE_0 + C\sqrt{T} \left(1 + X_T(\vec{A}_j, \phi_j)\right) X_T(\vec{A}_j, \phi_j)^2, \end{aligned}$$

where E_0 is the norm of the initial data (11.2). We assume $T \le 1$ to avoid having C dependent on T.

The rest of the argument is as usual:

- Assume $X_T(\vec{A}_j, \phi_j) \le 2CE_0$ (induction hypothesis).
- Choose $T = T(E_0) \le 1$ so small that $C\sqrt{T}(1 + 2CE_0)(2CE_0) \le 1/2$.

Then it follows from (11.41) and the induction hypothesis that also

$$X_T(\vec{A}_{j+1}, \phi_{j+1}) \le 2CE_0.$$

Having obtained this bound, the next step is to estimate

$$X_T(\vec{A}_{j+1} - \vec{A}_j, \phi_{j+1} - \phi_j)$$

in order to show that the sequence of iterates is Cauchy in X_T. However, by multilinearity etc., these estimates can be reduced to the estimates we have already proved (the only exception is for the equation (11.29), but for this it is easy to prove estimates for differences), so we ignore this issue.

11.14 Proof of the bilinear estimates

To wrap things up, we prove the estimates (11.38) and (11.39). By Duhamel's principle, it suffices to prove the estimates when $\Box u = \Box v = 0$. By Hölder's inequality,

$$\|F\|_{L_t^1 L_x^2(S_T)} \leq \sqrt{T}\,\|F\|_{L^2(S_T)} \leq \sqrt{T}\,\|F\|_{L^2(\mathbb{R}^{1+3})}\,,$$

so we can further reduce to proving the following:

Theorem 40. *Suppose* $\Box u = \Box v = 0$ *on* $\mathbb{R}^{1+3}$ *with initial data*

$$(u, \partial_t u)\big|_{t=0} = (u_0, u_1) \quad \textit{and} \quad (v, \partial_t v)\big|_{t=0} = (v_0, v_1).$$

Then

$$\left\| |D|^{-1} Q(\partial u, \partial v) \right\|_{L^2(\mathbb{R}^{1+3})} \leq C(\|u_0\|_{\dot{H}^1} + \|u_1\|_{L^2})(\|v_0\|_{\dot{H}^1} + \|v_1\|_{L^2}) \tag{11.42}$$

$$\left\| Q(|D|^{-1} \partial u, \partial v) \right\|_{L^2(\mathbb{R}^{1+3})} \leq C(\|u_0\|_{\dot{H}^1} + \|u_1\|_{L^2})(\|v_0\|_{\dot{H}^1} + \|v_1\|_{L^2}) \tag{11.43}$$

where Q *can be any of the null forms* Q_{ij}*,* $1 \leq i, j \leq 3$*.*

By the usual argument, it is enough to prove this with $u_1 = v_1 = 0$ and

$$u = e^{it|D|}u_0, \qquad \widehat{u}(t,\xi) = e^{it|\xi|}\widehat{u_0}(\xi), \tag{11.44a}$$

$$v = e^{\pm it|D|}v_0, \qquad \widehat{v}(t,\xi) = e^{\pm it|\xi|}\widehat{v_0}(\xi). \tag{11.44b}$$

The spacetime Fourier transforms are then

$$\widehat{u}(\tau,\xi) = \delta(\tau - |\xi|)\widehat{u}_0(\xi), \qquad \widehat{v}(\lambda,\eta) = \delta(\lambda \mp |\eta|)\widehat{v}_0(\eta). \tag{11.45}$$

These Fourier transforms are distributions, in fact measures, supported on the light cone: $\sqrt{2}\delta(\tau - |\xi|)\,d\tau\,d\xi$ [respectively $\sqrt{2}\delta(\tau + |\xi|)\,d\tau\,d\xi$] is surface measure on the forward [respectively backward] light cone. These statements are special cases of the following useful fact:

Proposition 10. *Let* $\phi : \mathbb{R}^m \to \mathbb{R}$ *be smooth. Set* $S = \{\eta : \phi(x) = 0\}$*. If* $\nabla\phi(x) \neq 0$ *for all* $x \in S$*, then*

$$\delta\big(\phi(x)\big) = \frac{d\sigma(x)}{|\nabla\phi(x)|},$$

where $d\sigma$ *is surface measure on the hypersurface* S*.*

Proof. The proof is a simple calculation. Let $f \in C_c^\infty(\mathbb{R}^m)$. Since

$$\delta = \lim_{\varepsilon \to 0^+} (2\varepsilon)^{-1}\chi_{(-\varepsilon,\varepsilon)},$$

we have

$$\int f(\eta)\delta(\phi(x))\,dx = \lim_{\varepsilon \to 0^+} (2\varepsilon)^{-1} \int_{|\phi(x)|<\varepsilon} f(x)\,dx,$$

so it suffices to show

$$\lim_{\varepsilon\to 0^+}(2\varepsilon)^{-1}\int_{|\phi(x)|<\varepsilon} f(x)\,dx = \int_S f(x)\frac{d\sigma(x)}{|\nabla\phi(x)|}.$$

Fix $p \in S$. Relabeling the axes, we may assume $\partial_m\phi(p) \neq 0$. Split the coordinates $x = (x', x_m)$, and change variables $x \to y = F(x)$, where

$$F(x) = (x', \phi(x)).$$

Then

$$|\det DF(x)| = |\partial_m\phi(x)|\,, \tag{11.46}$$

so by the Inverse Function Theorem, F maps some neighbourhood U of p diffeomorphically onto an open set V. Denote by $G : V \to U$ the inverse map. Then, using (11.46),

$$\int f(x)g(\phi(x))\,dx = \int f(G(y))g(y_m)\frac{dy}{|\partial_m\phi(G(y))|}.$$

Apply this with $g = (2\varepsilon)^{-1}\chi_{(-\varepsilon,\varepsilon)}$ and let $\varepsilon \to 0$, to obtain, for $f \in C_c^\infty(U)$,

(11.47)
$$\begin{aligned}\int f(x)\delta(\phi(x))\,dx &= \lim_{\varepsilon\to 0}\int\left[\frac{1}{2\varepsilon}\int_{-\varepsilon}^{\varepsilon} f(G(y',y_m))\frac{dy_m}{|\partial_m\phi(G(y',y_m))|}\right]dy' \\ &= \int f(G(y',0))\frac{dy'}{|\partial_m\phi(G(y',0))|}.\end{aligned}$$

To simplify the notation, we write $z = y'$. Then $G(z,0) = (z, h(z))$, where

$$\phi(z, h(z)) = 0.$$

Differentiating this, we get

$$\partial_i\phi(z,h(z)) + \partial_m\phi(z,h(z))\partial_i h(z) = 0$$

for $1 \le i \le m-1$. It then follows easily that

$$\sqrt{1+|\nabla h(z)|^2} = \frac{|\nabla\phi(z,h(z))|}{|\partial_m\phi(z,h(z))|}. \tag{11.48}$$

But $S \cap U$ is the graph $z \to (z, h(z))$, so surface measure is $\sqrt{1+|\nabla h(z)|^2}\,dz$. Thus, from (11.47) and (11.48) we get

$$\int f(x)\delta(\phi(x))\,dx = \int f(z,h(z))\frac{\sqrt{1+|\nabla h(z)|^2}}{|\nabla\phi(z,h(z))|}\,dz = \int_S f(x)\frac{d\sigma(x)}{|\nabla\phi(x)|}.$$

This holds for all $f \in C_c^\infty(U)$, and since $p \in S$ was arbitrary, it follows by a partition of unity argument that it holds for all $f \in C_c^\infty(\mathbb{R}^m)$. □

11.15 Proof of (11.42)

We shall prove (11.42) by reducing it to Strichartz' estimate

$$\|u\|_{L^4(\mathbb{R}^{1+3})} \leq C \|u_0\|_{\dot{H}^{\frac{1}{2}}}. \tag{11.49}$$

First note that $(\partial_j f \partial_j g)\hat{} = (\xi_j \widehat{f}\,) * (\eta_j \widehat{g})$ for functions f and g on $\mathbb{R}^3$. Using this fact, we see that the spacetime Fourier transform of $|D|^{-1} Q(\partial u, \partial v)$, $Q = Q_{ij}$, is

$$\frac{1}{|\xi|} \int Q(\xi - \eta, \eta) \widehat{u}(\tau - \lambda, \xi - \eta) \widehat{v}(\lambda, \eta) \, d\lambda \, d\eta.$$

The absolute value is bounded by

$$\frac{1}{|\xi|} \int |Q(\xi - \eta, \eta)| \, \widehat{U}(\tau - \lambda, \xi - \eta) \widehat{V}(\lambda, \eta) \, d\lambda \, d\eta,$$

where U and V are defined by

$$\widehat{U}(t, \xi) = e^{it|\xi|} \left|\widehat{u_0}(\xi)\right|, \qquad \widehat{V}(t, \xi) = e^{\pm it|\xi|} \left|\widehat{v_0}(\xi)\right|.$$

Next observe that $|Q(\xi, \eta)| \leq |\xi \times \eta|$. Since $\xi \times \eta = (\xi + \eta) \times \eta = \xi \times (\xi + \eta)$, we have

$$|\xi \times \eta| \leq \begin{cases} |\xi + \eta| \, |\eta| \, , \\ |\xi| \, |\xi + \eta| \, , \end{cases}$$

and hence $|\xi + \eta| \leq |\xi + \eta| \, |\xi|^{\frac{1}{2}} \, |\eta|^{\frac{1}{2}}$. We conclude that

$$\left| [|D|^{-1} Q(\partial u, \partial v)]\hat{}(\tau, \xi) \right| \leq [|D|^{\frac{1}{2}} U \, |D|^{\frac{1}{2}} V]\hat{}(\tau, \xi).$$

Therefore,

$$\left\| |D|^{-1} Q(\partial u, \partial v) \right\|_{L^2} \leq \left\| |D|^{\frac{1}{2}} U \right\|_{L^4} \left\| |D|^{\frac{1}{2}} V \right\|_{L^4} \leq C \|u_0\|_{\dot{H}^1} \|v_0\|_{\dot{H}^1},$$

where we used (11.49).

11.16 Proof of (11.43)

First some motivational remarks.

Remark. Observe that $Q(|D|^{-1} \partial u, \partial v)$ is schematically of the form $u \nabla v$. However, the estimate

$$\|u \nabla v\|_{L^2} \leq C \|u_0\|_{\dot{H}^1} \|v_0\|_{\dot{H}^1} \tag{11.50}$$

is false (barely). This is related to the false endpoint case of the Strichartz estimates for $n = 3$:

$$\|u\|_{L^2_t(L^\infty_x)} \leq C \|u_0\|_{\dot{H}^1}. \tag{11.51}$$

If this were true, we could write

$$\|u\nabla v\|_{L^2} \le \|u\|_{L_t^2(L_x^\infty)} \|\nabla v\|_{L_t^\infty(L_x^2)} \le \|u_0\|_{\dot H^1} \|v_0\|_{\dot H^1},$$

proving (11.50). (Incidentally, (11.51) is true for radially symmetric data, hence so is (11.50).)

The bad case in (11.50) is when u is at low frequency relative to v. If not, the estimate is true. Indeed, if $|\xi| \ge \frac{1}{2}|\eta|$ for $\xi \in \operatorname{supp}\widehat{u}_0$ and $\eta \in \operatorname{supp}\widehat{v}_0$, then $|\eta| \le \sqrt{2}\,|\xi|^{\frac12}|\eta|^{\frac12}$, hence

$$\|u\nabla v\|_{L^2} \le \sqrt{2}\big\||D|^{\frac12} U\,|D|^{\frac12} V\big\|_{L^2} \le C\big\||D|^{\frac12} U\big\|_{L^4}\big\||D|^{\frac12} V\big\|_{L^4},$$

and (11.50) follows after applying (11.49).

Let us now prove (11.43) with $u_1 = v_1 = 0$:

$$\Big\|Q(|D|^{-1}\partial u, \partial v)\Big\|_{L^2(\mathbb{R}^{1+3})} \le C\,\|u_0\|_{\dot H^1}\|v_0\|_{\dot H^1}.$$

Denote by $I(\tau,\xi)$ the spacetime Fourier transform of $Q(|D|^{-1}\partial u, \partial v)$. In view of (11.45) we have

$$I(\tau,\xi) = \int \frac{Q(\eta,\xi-\eta)}{|\eta|^2\,|\xi-\eta|} f(\eta) g(\xi-\eta)\delta(\tau - |\eta| \mp |\xi-\eta|)\,d\eta,$$

where $f(\eta) = |\eta|\,\widehat{u_0}(\eta)$ and $g(\eta) = |\eta|\,\widehat{v}_0(\eta)$. Apply Cauchy-Schwarz with respect to the measure $\delta(\dots)\,d\eta$:

$$|I(\tau,\xi)|^2 \le J_\pm(\tau,\xi)\int |f(\eta)|^2\,|g(\xi-\eta)|^2\,\delta(\tau-|\eta|\mp|\xi-\eta|)\,d\eta, \tag{11.52}$$

where

$$J_\pm(\tau,\xi) = \int \frac{|Q(\eta,\xi-\eta)|^2}{|\eta|^4\,|\xi-\eta|^2}\delta(\tau-|\eta|\mp|\xi-\eta|)\,d\eta.$$

Thus, it suffices to prove that

$$\sup_{\tau,\xi} J_\pm(\tau,\xi) < \infty, \tag{11.53}$$

for then the estimate follows after integrating (11.52) with respect to $d\tau\,d\xi$.

Observe that it suffices to prove (11.53) for (τ,ξ) in the complement of the light cone, which has measure zero in $\mathbb{R}^{1+3}$. Thus we assume $|\tau| \ne |\xi|$.

Switching to polar coordinates $\eta = \rho\omega$, where $\rho > 0$ and $\omega \in S^2$, and using the fact that

$$|Q(\eta,\xi-\eta)| \le |\eta\times(\xi-\eta)| = |\eta|\,|\xi-\eta|\sin\theta,$$

where $\theta = \theta(\eta, \xi - \eta)$ is the angle between η and $\xi - \eta$, we see that

$$
\begin{aligned}
J_\pm(\tau,\xi) &\le \int \frac{\sin^2\theta}{|\eta|^2}\delta(\tau - |\eta| \mp |\xi-\eta|)\, d\eta \\
&= \int_{S^2}\int_0^\infty \sin^2\theta\delta(\tau - \rho \mp |\xi - \rho\omega|)\, d\rho\, d\sigma(\omega),
\end{aligned} \tag{11.54}
$$

so it suffices to show that the last integral is uniformly bounded for $|\tau| \neq |\xi|$. We consider two cases depending on the choice of sign in $J_\pm$.

Case I. Estimate for J_+. $\delta(\tau - |\eta| - |\xi - \eta|)\, d\eta$ is a measure on the ellipsoid

$$E(\tau,\xi) = \{\eta : |\eta| + |\xi - \eta| = \tau\}.$$

This is empty if $|\xi| > \tau$, so we may assume

$$|\xi| < \tau. \tag{11.55}$$

Fix such τ, ξ, as well as $\omega \in S^2$. Then [2]

$$
\begin{aligned}
\eta = \rho\omega \in E(\tau,\xi) &\iff \tau - \rho = |\xi - \rho\omega| \\
&\iff (\tau-\rho)^2 = |\xi - \rho\omega|^2 \iff \rho = \frac{\tau^2 - |\xi|^2}{2(\tau - \xi\cdot\omega)}.
\end{aligned} \tag{11.56}
$$

Thus, applying Proposition 10 to the function

$$\rho \to \tau - \rho - |\xi - \rho\omega|$$

(smooth near its zero set) we get after a simple calculation,

$$\delta(\tau - \rho - |\xi - \rho\omega|)\, d\rho = \frac{2\rho(\tau-\rho)}{\tau^2 - |\xi|^2}\delta\left(\rho - \frac{\tau^2 - |\xi|^2}{2(\tau - \xi\cdot\omega)}\right) d\rho. \tag{11.57}$$

On the other hand,

$$
\begin{aligned}
\sin^2\theta &= 1 - \cos^2\theta \le 2(1 - \cos\theta) \\
&= \frac{2}{|\eta|\,|\xi-\eta|}[|\eta|\,|\xi - \eta| - \eta\cdot(\xi - \eta)] = \frac{\tau^2 - |\xi|^2}{|\eta|\,|\xi - \eta|},
\end{aligned} \tag{11.58}
$$

where we used the identity $\tau = |\eta| + |\xi - \eta|$ to get the last equality.

Combining (11.54), (11.57) and (11.58) gives

$$J_+(\tau,\xi) \le \int_{S^2}\int_0^\infty 2\delta\left(\rho - \frac{\tau^2 - |\xi|^2}{2(\tau - \xi\cdot\omega)}\right) d\rho\, d\sigma(\omega) = 2\,\mathrm{Area}(S^2).$$

[2] Note that $(\tau-\rho)^2 = |\xi - \rho\omega|^2$ implies $\tau = \rho + |\xi - \rho\omega|$ or $\tau = \rho - |\xi - \rho\omega|$; however, the second alternative can be discounted, since it implies $|\tau| \le |\xi|$, contradicting (11.55).

Case II. Estimate for J_-. $\delta(\tau - |\eta| + |\xi - \eta|)\, d\eta$ is a measure on

$$H(\tau, \xi) = \{\eta : |\eta| - |\xi - \eta| = \tau\}$$

(one sheet of a hyperboloid). $H(\tau, \xi)$ is empty if $|\xi| < |\tau|$, so we may assume

$$|\tau| < |\xi|\,.$$

Fix such τ, ξ, as well as $\omega \in S^2$. Then reasoning as in the previous case,

$$\eta = \rho\omega \in H(\tau, \xi) \iff \rho - \tau = |\xi - \rho\omega| \iff \rho = \frac{\tau^2 - |\xi|^2}{2(\tau - \xi \cdot \omega)}. \tag{11.59}$$

Applying Proposition 10 to the function

$$\rho \to \tau - \rho + |\xi - \rho\omega|$$

then gives

$$\delta(\tau - \rho + |\xi - \rho\omega|)\, d\rho = \frac{2\rho(\rho - \tau)}{|\xi|^2 - |\xi|^2} \delta \left(\rho - \frac{\tau^2 - |\xi|^2}{2(\tau - \xi \cdot \omega)} \right) d\rho. \tag{11.60}$$

On the other hand,

$$\begin{aligned} \sin^2\theta &= 1 - \cos^2\theta \le 2(1 + \cos\theta) \\ &= \frac{2}{|\eta|\,|\xi - \eta|} [|\eta|\,|\xi - \eta| + \eta \cdot (\xi - \eta)] = \frac{|\xi|^2 - \tau^2}{|\eta|\,|\xi - \eta|}, \end{aligned} \tag{11.61}$$

where we used the identity $\tau = |\eta| - |\xi - \eta|$ to get the last equality.

Combining (11.54), (11.60) and (11.61) gives

$$J_-(\tau, \xi) \le \int_{S^2} \int_0^\infty 2\delta \left(\rho - \frac{\tau^2 - |\xi|^2}{2(\tau - \xi \cdot \omega)} \right) d\rho\, d\sigma(\omega) \le 2\,\mathrm{Area}(S^2).$$

This concludes the proof.