Math 222B Lecture Notes
Partial Differential Equations

Professor: Sung-Jin Oh
Scribe: Daniel Raban

Spring 2022

Contents

1 Introduction to Sobolev Spaces

The main reference for our material on Sobolev spaces will be Ch 5 of Evans' PDE book.

1.1 Sobolev spaces

Definition 1.1. Let $u \in \mathcal{D}'(U)$, where $U \subseteq \mathbb{R}^d$ is open. The **k-th order L^p-based Sobolev norm** of u is

$$\|u\|_{W^{k,p}(U)} := \sum_{\alpha:|\alpha \leq k} \|D^\alpha u\|_{L^p},$$

where we are using the distributional derivative and assume that D^α is an L^p function.

Remark 1.1. Expressions of the form $\|D^\alpha u\|_{L^p}$ arise in the energy method for PDEs with $p = 2$.

Definition 1.2. The **L^p-based Sobolev space of order k on U** is

$$W^{k,p}(U) = \{u \in \mathcal{D}'(U) : \|u\|_{W^{k,p}(U)} < \infty\}.$$

Note that $C_c^\infty(U) \subseteq W^{k,p}(U)$. This allows us to make the following definition:

Definition 1.3. The set of $u \in W^{k,p}(U)$ that vanish (to appropriate orders) on ∂U is

$$W_0^{k,p}(U) = \overline{C_0^\infty(U)}^{\|\cdot\|_{W^{k,p}}}.$$

When $p = 2$, we introduce the notation

$$H^k(U) = W^{k,2}(U), \qquad H_0^k = W_0^{k,2}(U).$$

We can define an inner product on H^k by

$$\langle u, v\rangle_{H^k} = \sum_{\alpha:|\alpha|\leq k} \langle D^\alpha u, D^\alpha v\rangle_{L^2}.$$

Proposition 1.1.

(i) For all $k \in \mathbb{Z}_{\geq 0}$ and $1 \leq p \leq \infty$, $(W^{k,p}(U), \|\cdot\|_{W^{k,p}})$ and $(W_0^{k,p}(U), \|\cdot\|_{W^{k,p}})$ are Banach spaces.

(ii) For all $k \in \mathbb{Z}_{\geq 0}$, $(H^k(U), \langle\cdot,\cdot\rangle_{H^k})$ and $(H_0^k(U), \langle\cdot,\cdot\rangle_{H^k})$ are Hilbert spaces.

(iii) (Fourier-analytic characterization of H^k) Given $u \in H^k(U)$,

$$\begin{aligned}\|u\|_{H^k} &\simeq \|\widehat{u}\|_{L^2} + \||\xi|^k \widehat{u}\|_{L^2} \\ &\simeq \|(1+|\xi|^2)^{k/2}\widehat{u}\|_{L^2},\end{aligned}$$

where $A \simeq B$ means $A \lesssim B$ and $B \lesssim A$.

1.2 Duality and Sobolev spaces of negative order

First, we will give a proposition, and then we will explain what is going on.

Proposition 1.2. *For $k \in \mathbb{Z}_{\geq 0}$ and $1 < p < +\infty$,*

$$(W_0^{k,p}(U))^* \simeq W^{-k,p'}(U),$$

where $\frac{1}{p'} + \frac{1}{p} = 1$.

Definition 1.4. For $k \in \mathbb{Z}_+$ and $1 < p < +\infty$, the **negative order Sobolev space norm** is

$$\|u\|_{W^{-k,p}(U)} = \inf \left\{ \sum_{\alpha:|\alpha|\leq k} \|g_\alpha\|_{L^p} : u = \sum_{\alpha:|\alpha|\leq k} D^\alpha g_\alpha \right\}.$$

The **negative order Sobolev space** is

$$W^{-k,p}(U) = \{u \in \mathcal{D}'(U) : u = \sum_{\alpha:|\alpha|\leq k} A^\alpha g_\alpha, g_\alpha \in L^p(U)\}.$$

Remark 1.2. If $g \in L^p$, then $D_{x^1} g \in W^{-1,p}(U)$. Compare this with the property of Sobolev spaces that if $u \in W^{k,p}(U)$, then $D_{x^j} u \in W^{k-1,p}(U)$.

Here is the proof of the proposition:

Proof. $(W_0^{k,p}(U))^* \supseteq W^{-k,p'}(U)$: Take $v \in W^{-k,p'}(U)$, so $v = \sum_{\alpha:|\alpha|\leq k} D^\alpha g_\alpha$; we can also take this decomposition so that $\| \sum_{\alpha:|\alpha|\leq k} D^\alpha g_\alpha \|_{L^p} \leq 2\|v\|_{W^{k,p'}(U)}$. Then for $u \in W_0^{k,p}(U)$, we can treat v as a linear functional by

$$\langle v, u \rangle = \int vu \, dx$$

To show that this is bounded,

$$= \sum_{\alpha:|\alpha|\leq k} \int D^\alpha g_\alpha u \, dx$$

First assuming $u \in C_c^\infty$ and then applying a density argument,

$$\begin{aligned} &= \sum_{\alpha:|\alpha|\leq k} \int (-1)^{|\alpha|} g_\alpha D^\alpha u \, dx \\ &\leq \sum_{\alpha:|\alpha|\leq k} |g_\alpha\|_{L^{p'}} \|D^\alpha u\|_{L^p} \\ &\leq C\|v\|_{W^{-k,p'}} \|u\|_{W^{k,p}}. \end{aligned}$$

$(W_0^{k,p}(U))^* \subseteq W^{-k,p'}(U)$: The idea is to use the Hahn-Banach theorem. If X is a nomed vector space and $Y \subseteq X$ with a linear functional $\ell : Y \to \mathbb{R}$ such that $|\ell(u)| \leq C\|u\|$, then there exists an extension $\widetilde{\ell} : X \to \mathbb{R}$ such that $|\widetilde{\ell}(u)| \leq c\|u\|$ and $\widetilde{\ell}|_Y = \ell$.

Let $\ell : W_0^{k,p}(U) \to \mathbb{R}$ be bounded. Define a linear map $C_0^\infty(U) \to L^p(U)^{\oplus K(k)}$ sending $u \mapsto (u, D_{x^1}u, \dots, D_{x^\alpha}u, \dots, D^\alpha u)$, ranging over all multiindices α with $|\alpha| \leq k$. Then $\|Y(u)\| \leq C\|u\|_{W^{k,p}}$, T is injective, and T is an isomorphism of $(C_c^\infty(U), \|\cdot\|_{W^{k,p}})$ with its image $(T(C_c^\infty(U)), \|\cdot\|)$. This gives a bounded map $\widetilde{\ell} : T(C_c^\infty(U)) \to \mathbb{R}$ by $\widetilde{\ell}(Tu) = \ell(u)$. By the Hahn-Banach theorem, $\widetilde{\ell}$ extends to a bounded map $\widetilde{\widetilde{\ell}} : L^p(U)^{\oplus K} \to \mathbb{R}$. That is, $\widetilde{\widetilde{\ell}} \in (L^p(U)^{\oplus K})^* = \{\widetilde{v} = \sum_\alpha \widetilde{g}_\alpha : \widetilde{g}_\alpha \in L^{p'}(U)$. In this picture, for $\widetilde{u} \in L^p(U)^{\oplus K}$, $\langle \widetilde{v}, \widetilde{u} \rangle = \sum_\alpha \langle \widetilde{g}_\alpha, \widetilde{u}_\alpha \rangle$. This means that $\widetilde{\widetilde{\ell}}(\widetilde{v}) = \sum_\alpha \langle \widetilde{g}_\alpha, \widetilde{u}_\alpha \rangle$ for some $\widetilde{g}_\alpha \in L^{p'}(U)$.

This gives

$$\ell(u) = \widetilde{\ell}(Tu) = \widetilde{\widetilde{\ell}}(Tu) = \sum_\alpha \langle \widetilde{g}_\alpha, (Tu)_\alpha \rangle = \sum_\alpha \langle \widetilde{g}_\alpha, D^\alpha u \rangle.$$

Now set $g_\alpha = (-1)^{|\alpha|}\widetilde{g}_\alpha$. $\square$

1.3 Duality in relation to existence and uniqueness

Here is some motivation for our functional analysis. Let X, Y be Banach spaces, and let $P : X \to Y$ be bounded and linear.

- For a given $f \in Y$, does there exists a $u \in X$ such that $Pu = f$? This is the question of existence of a solution to $Pu = f$.
- We can also ask about uniqueness: If $u, u' \in X$ and $Pu = Pu'$, is $u = u'$? That is, if $Pu = 0$, is $u = 0$?

In our course, we usually take P to be a linear differential operator, such as $P = -\Delta$ or $P = \Box$. If we want to solve

$$\begin{cases} -\Delta u = f & \text{in } U \\ u = 0 & \text{on } \partial U \end{cases}$$

Then we have

$$\int_U |Du|^2 = \int_U -\Delta u \cdot u = \int_U f.$$

This gives

$$\|Du\|_{L^2} \lesssim \left| \int f u \, dx \right|.$$

We are assuming that we have a solution and inferring information about u. This is called an **a priori estimate**.

In energy methods for PDEs, you usually prove a priori estimates, which at first sight, only pertain to uniqueness. However, in fact, a priori estimates are also useful for proving existence because existence vs uniqueness are related to each other using duality. This is the phenomenon in linear algebra where if $A \in \mathbb{R}^{n \times m}$, then A is injective iff A^* is surjective.

2 A Priori Estimates and Approximation Theorems

2.1 Relationship between a priori estimates, existence, and uniqueness

Last time, we were investigating the question "Why study Sobolev spaces as Banach spaces?" We made a digression into functional analysis:

If X and Y are Banach spaces and $P : X \to Y$ is bounded and linear, we had 2 concerns:

- (Existence) Given $f \in T$, does there exists a $u \in X$ such that $Pu = f$?
- (Uniqueness) Given $u \in X$ such that $Pu = 0$, does $u = 0$?

These two problems are related to each other by duality.

Remark 2.1. Here is a concrete thing to keep in mind: Often, we prove a priori estimates for a PDE, i.e. if $u \in X$ with $Pu = f$, then $\|u\|_X \leq C\|f\|_Y$.

Proposition 2.1. *Let X, Y be Banach spaces, and let $P : X \to Y$ be a bounded, linear operator. Denote by $P^* : Y^* \to X^*$ the adjoint of P, i.e. $\langle v, Pu \rangle = \langle P^*v, u \rangle$ for all $u \in X, v \in Y^*$. Suppose there exists a constant $C > 0$ such that $\|u\|_X \leq C\|Pu\|_Y$ for all $u \in X$. Then*

1. *(Uniqueness for $Pu = f$) If $u \in X$ and $Pu = 0$, then $u = 0$.*
2. *(Existence for $P^*v = g$) For all $g \in X^*$, there exists a $v \in Y^*$ such that $P^*v = g$ and $\|v\|_{Y^*} \leq C\|g\|_{X^*}$.*

Proof. Here is the proof of 2, via the Hahn-Banach theorem. We want to find $v \in Y^*$ such that $P^*v = g$, which is equivalent to $\langle P^*v, u \rangle = \langle g, u \rangle$ for all $u \in X$. The left side is $\langle v, Pu \rangle$, so we will start with a subspace of elements of the form Pu.

Define $\ell : P(X) \to \mathbb{R}$ by the relation

$$\ell(Pu) = \langle g, u \rangle.$$

Note that since P is injective, this ℓ is well-defined. This is bounded because or $\|Pu\|_Y \leq 1$,

$$\begin{aligned} \ell(Pu)| = |\langle g, u \rangle| &\leq \|g\|_{X^*}\|u\|_X \\ &\leq C\|g\|_{X^*}\|Pu\|_Y \\ &\leq C\|g\|_{X^*} \end{aligned}$$

So Hahn-Banach says that there is a $v \in Y^*$ such that

$$\langle v, Pu \rangle = \ell(pu) = \langle g, u \rangle \qquad \forall u \in X$$

and $\|v\|_{Y^*} \leq C\|g\|_{X^*}$. □

What about existence for the original problem $Pu = f$? Let us take an easy way out and assume that X is **reflexive** ($X \to (X^*)^*$ sending $u \mapsto (u \mapsto \langle v, u\rangle)$ is an isomorphism).

Proposition 2.2. *Let X, Y be Banach spaces, and let $P : X \to Y$ be a bounded, linear operator. Suppose $\|v_{Y^*}\| \leq C\|P^*v\|_{X^*}$. Then*

1. *(Uniqueness for $P^*v = g$) If $v \in Y^*$ and $P^*v = 0$, then $v = 0$.*
2. *(Existence for $Pu = f$) For all $f \in Y$, there exists a $u \in X$ such that $Pu = f$ and $\|u\|_X \leq C\|f\|_Y$.*

Proof. Same as before. Construct $u \in X$ by constructing a bounded linear functional on X^* (because $X = (X^*)^*$ by reflexivity. □

Remark 2.2. All Sobolev spaces $W_0^{k,p}(U)$ with $1 < p < \infty$ are reflexive.

Remark 2.3.

$$(\operatorname{ran} P)^\perp = \ker P^*, \qquad \ker P =^\perp (\operatorname{ran} P^*).$$

Here, we mean *annihilators.*

Definition 2.1. Given $U \subseteq Y$, the **annihilator** of U is $U^\perp = \{v \in Y^* : \langle v, f\rangle = 0 \forall f \in U\}$. Given $V \subseteq X^*$, the **annihilator** of U is $U^\perp = \{u \in X : \langle g, u\rangle = 0\, \forall g \in V\}$.

As a consequence, if $\ker P^* = \{0\}$, then by Hahn-Banach,

$$(\operatorname{ran} P)^\perp = \{0\} \iff \overline{\operatorname{ran} P} = Y.$$

In the finite dimensional case, $\overline{\operatorname{ran} P} = \operatorname{ran} P$. Therefore, we get the well-known fact from linear algebra concerning the solvability of the problem $Ax = b$ with A a possibly non-square matrix:

$$(\text{for all } b, \text{ there exists an } x \text{ such that } Ax = b) \iff (A^*y = 0 \implies y = 0),$$

$$(\text{for all } c, \text{ there exists an } y \text{ such that } A^*y = c) \iff (Ax = 0 \implies x = 0).$$

However, in the infinite dimensional case, $\overline{\operatorname{ran} P} = \operatorname{ran} P$, so we can think of the annihilator as measuring how close these are.

Remark 2.4 (Qualitative vs quantitative)**.** There is no loss of generality in deriving existence for $Pu = f$ from the quantitative bound $\|v\|_{Y^*} \leq C\|P^*v\|_{X^*}$.

Proposition 2.3. *Let X, Y be Banach spaces and $P : X \to Y$ be a bounded linear operator. If $P(X) = T$, then there exists some $C > 0$ such that $\|v\|_{Y^*} \leq C\|P^*v\|_{X^*}$.*

Proof. By the open mapping theorem, $P(B_X)$, the image of the unit ball in X, is open and contains the origin. So there exists a $C > 0$ such that $P(B_X) \supseteq cB_Y$. Then

$$\begin{aligned}
\|P^*v\|_{X^*} &= \sup_{u:\|u\|_X \leq 1} |\langle P^*v, u\rangle| \\
&= \sup_{u \in \overline{B_X}} |\langle v, Pu\rangle| \\
&= \sup_{f \in P(\overline{B_X})} |\langle v, f\rangle| \\
&\geq \sup_{f \in cB_Y} |\langle v, f\rangle| \\
&\geq C\|v\|_{Y^*}.
\end{aligned}$$

□

Example 2.1. Let's try to solve the 1-dimensional Laplace equation

$$\begin{cases} -u'' = f & \text{in } (0,1) \\ u = 0 & \text{at } x = 0, 1. \end{cases}$$

We will investigate solvability in $H_0^1((0,1)) = \overline{C_c^\infty(0,1)}^{\|\cdot\|_{H^1}}$, where $\|u\|_{H^1}^2 = \|u_{L^2}^2 + \|u'\|_{L^2}^2$. Recall that $(H_0^1((0,1)))^* = H^{-1}(0,1)$. Then we have $Pu = -u''$ with domain $X = H_0^1((0,1))$ and codomain $Y = H^{-1}(0,1)$.

We claim that if $Pu = f$ for some $u \in X$ then $\|u\|_X \leq C\|f\|_Y$. This means that if $u \in H_0^1(0,1)$ satisfies the equation $-u'' = f$, then $\|u\|_{H^1} \leq C\|f\|_{H^{-1}}$.

Proof. To prove this bound, it suffices by density to consider $u \in C_c^\infty((0,1))$. Multiply both sides by u and integrate:

$$\int fu\,dx = \int -u''u\,dx$$

Since $u \in C_c^\infty((0,1))$ there are no boundary terms. So we may integrate by parts.

$$= \int (u')^2\,dx.$$

But how about $\|u\|_{L^2}$? Use the fact that u vanishes on the boundary:

$$u(x) = \int_0^x u'(x)\,dx.$$

Then for any $x \in (0,1)$, we can say

$$|u(x)| \leq \int_0^1 |u'(x')|\,dx' \overset{\text{Cauchy-Schwarz}}{\leq} \|u'\|_{L^2}^2.$$

We now have that

$$\begin{aligned}\|u\|_{H^1}^2 &\le C|\langle f, u\rangle| \\ &\le C\|f\|_{H^{-1}}\|u\|_{H^1}.\end{aligned}$$

Cancelling one factor of $\|u\|_{H^1}$ on each side gives $\|u\|_{H^1} \le C\|f\|_{H^{-1}}$. $\square$

Combined with proposition 1 gives us that if $-u'' = 0$ and $u \in H_0^1((0,1))$, then $u = 0$. To use proposition 2, we need to compute P^*:

$$\langle P^* v, u\rangle = \langle v, Pu\rangle \qquad \forall v \in (H^{-1})^*, u \in H_0^1.$$

Note that by reflexivity of H_0^1, $(H^{-1})^* = H_0^1$. Let's write this out:

$$\langle v, Pu\rangle = \int_0^1 v(-u'')\,dx$$

To use integration by parts, do another density argument.

$$\begin{aligned} &= \int_0^1 v'u'\,dx \qquad (v \in H_0^1) \\ &= \int_0^1 -v''u\,dx \qquad (u \in H_0^1) \\ &= \langle P^* v, u\rangle. \end{aligned}$$

This tells us that $P^* v = -v''$ with domain $Y^* = H_0^1((0,1))$ and codomain $X^* = H_0^{-1}((0,1))$, so the problem is self-dual. So we get existence: for all $f \in H^{-1}$, there is a $u \in H_0^1$ such that $Pu = f$.

This is a pretty high-powered approach that works for a variety of problems. To prove quantitative estimates, we will in general use **Poincaré inequlities**.

2.2 Approximation by smooth functions and smooth partition of unity

There are two main tools we will use: convolution and mollifiers.

Lemma 2.1. *Let φ be smooth, compactly supported, and have $\int \varphi\,dx = 1$. Let $u \in L^p(\mathbb{R}^d)$ with $1 \le p < \infty$. Denote **mollifiers** $\varphi_\varepsilon(x) = \frac{1}{\varepsilon^d}\varphi(x/\varepsilon)$ (so $\int \varphi_\varepsilon$). Then*

$$\|\varphi_\varepsilon u - u\|_{L^p} \xrightarrow{\varepsilon \to 0} 0,$$

where $\varphi_\varepsilon * u = \int \varphi_\varepsilon(x-y)u(y)\,dy$.

Proof. The key ingredient is the continuity of the translation operator on L^p. Define for $z \in \mathbb{R}^d$ and $u \in L^p$ the translation operator $\tau_z u(x) = u(x - z)$. Then

$$\lim_{|z| \to 0} \|\tau_z u - u\|_{L^p} = 0,$$

which you can check. Now

$$\varphi_\varepsilon * u(x) - u(x) = \int u(x-y)\varphi_\varepsilon(y)\,dy - u(x)$$

Since $\int \varphi_\varepsilon = 1$,

$$= \int (u(x-y) - u(x))\varphi_\varepsilon(y)\,dy.$$

Taking the L^p norm, we have

$$\begin{aligned} \|\varphi_\varepsilon * u(x) - u(x)\|_{L^p} &= \left\| \int (u(x-y) - u(x))\varphi_\varepsilon(y)\,dy \right\|_{L^p} \\ &\le \int \|u(\cdot - y) - u(\cdot)\|_{L^p} |\varphi_\varepsilon(y)|\,dy \end{aligned}$$

Since φ has compact support, $\operatorname{supp}\varphi_\varepsilon \to \{0\}$ as $\varepsilon \to \{0\}$. Thus, the integrand goes to 0 as $\varepsilon \to 0$. So we may apply the dominated convergence theorem to get

$$\xrightarrow{\varepsilon \to 0} 0.$$

□

This approximation is useful because $\varphi_\varepsilon * u$ is smooth.

Another useful tool is a smooth partition of unity:

Lemma 2.2. *Suppose $\{U_\alpha\}_{\alpha \in A}$ be an open covering of U in $\mathbb{R}^d$. There exists a **smooth partition of unity** $\{\chi_\alpha\}_{\alpha \in A}$ on U subordinate to $\{U_\alpha\}_{\alpha \in A}$, i.e.*

1. *$\sum_\alpha \chi_\alpha(x) = 1$ on U and for all $x \in U$ there exist only finitely many nonzero $\chi_\alpha(x)$*
2. *$\operatorname{supp}\chi_\alpha \subseteq U_\alpha$*
3. *χ_α is smooth.*

Proof. Start from a continuous partition of unity and apply the previous lemma to approximate by smooth functions. □

3 Approximation in Bounded Domains and the Extension Theorem

Today, our goals are

- Prove approximation (or density) theorems for Sobolev spaces.
- Prove extension theorems and the trace theorem (tools for dealing with $W^{k,p}(U)$ when U is a bounded domain).

3.1 Approximation theorems in bounded domains

Given $u \in W^{k,p}(U)$, we want to approximate it by something that is "better" (e.g. u is smooth or has a nice support property). Last time, we discussed two tools:

1. Convolution and mollification: If $f, g : \mathbb{R}^d \to \mathbb{R}$, then

   $$f * g(x) = \int f(x-y)g(y)\, dy.$$

   This has the property that

   $$\partial_{x_j}(f * g)(x) = \partial_{x_j} f * g(x) = f * \partial_{\partial_{x_j}} g(x).$$

   This means that you only need one of the functions to be smooth to get a smooth result.

   For $\varphi \in C_c^\infty(\mathbb{R}^d)$, if we denote $\varphi_\varepsilon = \frac{1}{\varepsilon^d}\varphi(\cdot/\varepsilon)$, then

   $$\varphi_\varepsilon f \xrightarrow{\varepsilon \to 0} f,$$

   where the left hand side is smooth. If $f \in \mathcal{D}'(\mathbb{R}^d)$, this convergence is convergence of distributions, and if $f \in L^p(\mathbb{R}^d)$, this convergence is in L^p.

2. Smooth partition of unity: If $\{U_\alpha\}_\alpha \in A$ is a collection of open sets (usually $U \subseteq U_{\alpha \in A} U_\alpha$) then there exist functions $\chi_\alpha(x)$ ($\alpha \in A$) such that

   (i) χ_α is smooth.

   (ii) $\sum_{\alpha \in A} \chi_\alpha = 1$ on U, where for all $x \in U$, $\chi_\alpha(x) = 0$ except for finitely many α.

   (iii) $\operatorname{supp} \chi_\alpha \subseteq U_\alpha$.

Theorem 3.1. *Let $k \geq 0$ be an integer and $1 \leq p < \infty$.*

(i) $C^\infty(\mathbb{R}^d)$ is dense in $W^{k,p}(\mathbb{R}^d)$.

(ii) $C_c^\infty(\mathbb{R}^d)$ *is dense in* $W^{k,p}(\mathbb{R}^d)$.

Proof.

(a) This is an application of mollification

(b) Approximate by $f\chi(1/R)$, letting $R \to \infty$, where $\chi \in C_c^\infty(\mathbb{R}^d)$ is such that $\chi(0) = 1$. □

Theorem 3.2. *Let* $k \geq 0$ *be an integer,* $1 \leq p < \infty$*, and* U *an open subset of* $\mathbb{R}^d$*. Then* $C^\infty(U)$ *is dense in* $W^{k,p}(U)$.

Proof. Let $u \in W^{k,p}(U)$, and fix $\varepsilon > 0$. We want to find $v \in C^\infty(U)$ such that $\|u - v\|_{W^{k,p}} \leq \varepsilon$.

Define $U_j = \{x \in U : \operatorname{dist}(x, \partial U) > 1/j\}$, and let $V_j = U_j \setminus \overline{U_{j+1}}$

Then $U \subseteq \bigcup_{j=1}^\infty V_j$, so there is a smooth partition of unity χ_j subordinate to V_j. Now split

$$u = \sum_{j=1}^\infty \underbrace{u\chi_j}_{:=u_j}.$$

Then, as $\operatorname{supp}\chi_j \subseteq V_j$, we have that $\operatorname{supp} u_j = \operatorname{supp}(u\chi_j) \subseteq V_j$. Moreover, $u_j \in C_c^\infty(\mathbb{R}^d)$.

If we let $\varphi \in C_c^\infty(\mathbb{R}^d)$ with $\int \varphi = 1$ and $\operatorname{supp}\varphi \subseteq B_1(0)$ is a mollifier, let $v_j = \varphi_{\varepsilon_j} * u_j$, where ε_j is chosen to achieve

$$\|u_j - v_j\|_{W^{k,p}} \leq 2^{-j\varepsilon}, \qquad \operatorname{supp} v_j \subseteq \widetilde{V}_j = U_{j-1} \setminus \overline{U_{j+2}}.$$

Here, we make use of the fact that $\operatorname{supp} f * g \subseteq \operatorname{supp} f + \operatorname{supp} g = \{x + y \in \mathbb{R}^d : x \in \operatorname{supp} f, y \in \operatorname{supp} g\}$. Now take $v = \sum_{j=1}^\infty v_j$. This is well-defined, as $\widetilde{V}_j$ is locally finite. This is also smooth, so $v \in C^\infty(U)$. On the other hand,

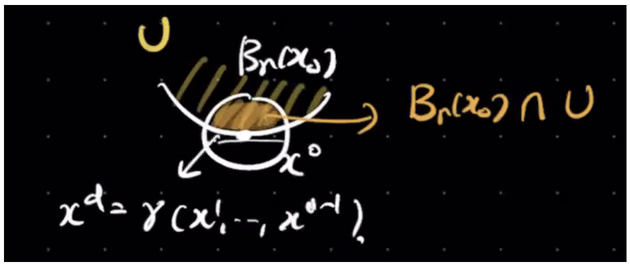
$$\|v - u\|_{W^{k,p}} \sum_{j=1}^\infty \|v_j - u_j\|_{W^{k,p}} \leq \sum_{k=1}^\infty 2^{-j}\varepsilon = \varepsilon.$$

□

Theorem 3.3. *Let $k \geq 0$ be an integer, $1 \leq p < \infty$, and U a bounded open set with ∂U of class C^1. Then $C^\infty(\overline{U})$ is dense in $W^{k,p}$.*

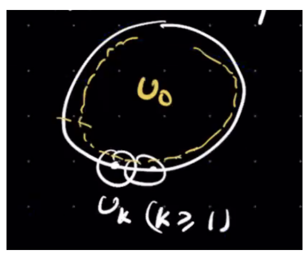
Here, $C^\infty(\overline{U})$ is the set of functions $u : U \to \mathbb{R}$ such that u is the restriction to U of a smooth function $\widetilde{u} \in C^\infty(\widetilde{U})$, where $\widetilde{U} \supseteq \overline{U}$ is open.

Definition 3.1. We say that ∂U is **of class** C^k if for all $x_0 \in \partial U$, there exists a radius $r = r(x_0) > 0$ such that, up to relabeling the variables, $B_r(x_0) \cap U = \{x \in B_r(x_0) : x^d > \gamma(x^1, \ldots, x^{d-1})\}$ for some C^k function $\gamma = \gamma(x^1, \ldots, x^{d-1})$ on $B_r(x_0) \cap (\mathbb{R}^{d-1} \times \{x_0^d\})$.



For the proof, we want to apply mollification, but the difficulty is what happens near the boundary. The idea is to first look at a small piece of the boundary at a time.

Proof. Step 1: Let $u \in W^{k,p}(U)$. By the definition of C^1-regularity of ∂U, ∂U can be covered by balls $\{B_{r_k}(x_k)\}_{k=1}^K$, in each of which U can be represented as the region above some C^1 graph. The number of such balls, K, is finite by the compactness of ∂U. We may add to $U_k = B_{r_k}(x_k)$ an open set U_0 which contains $U \setminus \bigcup_{k=1}^K U_k$, so that $\{U_0, U_1, \ldots, U_k\}$ is an open covering of U.



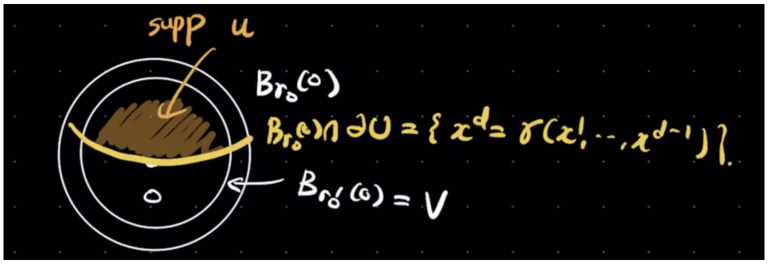
Let $\{\chi_k\}_{k=0}^K$ be a smooth partition of unity subordinate to $\{K_k\}_{k=0}^K$, and split

$$u = \sum_{k=0}^{\infty} u\chi_k =: u_0 + \sum_{k=1}^{K} u_k.$$

Here, u_0 is compactly supported, and $u \in W^{k,p}(\mathbb{R}^d)$, so we can use mollification, as before.

To deal with the u_k with $k \geq 1$, it suffices to consider the case where $U = B_{r_0}(x_0)$ and $\operatorname{supp} u_k \subseteq V \subseteq U$, where V is a smaller ball $B_{r_0'}(x_0)$, in which $B_{r_0}(x_0) \cap \partial U$ is more concrete.

Step 2: Without loss of generality, assume $x_0 = 0$.



We use a two-step approximation. Let $\varepsilon > 0$.

1. Let $w_\eta(x) = u(x + \eta e_d)$, where $e_d = (0, 0, \dots, 0, 1)$, and η will be chosen. Then $\operatorname{supp} w_\eta$ is the support of u shifted by 1. For η small enough, we have

$$\|u - w_\eta\|_{W^{k,p}(U \cap B_{r_0}(0))} < \frac{1}{2}\varepsilon.$$

   Moreover, ε is defined on $B_{r_0}(0) \cap U - \eta e_d$

2. Let $v = \varphi_\delta * w_\eta$, and if $\delta \ll \eta$ (and $\operatorname{supp} \varphi \subseteq B_1(0)$), then v is well-defined on $V \cap \{x^d > \gamma(x^1, \dots, x^{d-1})\}$. And if δ is sufficiently small, then

$$\|v - w_\eta\|_{W^{k,p}(U \cap B_{r_0}(x_0))} < \frac{1}{2}\varepsilon.$$

   This gives us

$$\|u - v\|_{W^{k,p}(U)} \leq \frac{1}{2}\varepsilon + \frac{1}{2}\varepsilon = \varepsilon.$$

   Moreover, $v \in C^\infty(\overline{V \cap \{x^d > \gamma(x^1, \dots, x^{d-1})\}})$, which is acceptable. $\square$

3.2 The extension theorem

The extension theorem is a tool to deal with $u \in W^{k,p}(U)$, where U is a bounded domain, by producing an extension of $u \in W^{k,p}(\mathbb{R}^d)$ with quantitative bounds on the extension.

Theorem 3.4 (Extension theorem)**.** *Let $k \geq 0$ be a nonnegative integer, $1 \leq p < \infty$, U a bounded domain with with C^k boundary. Let V be an open set such that $V \supseteq \overline{U}$. Then there exists an operator $\mathcal{E} : W^{k,p}(U) \to W^{k,p}(\mathbb{R}^d)$ such that*

(i) (Extension) $\mathcal{E}u|_U = u$.

(ii) (Linear and bounded) $\mathcal{E}$ is linear, and $\|\mathcal{E}u\|_{W^{k,p}(\mathbb{R}^d)} \leq C\|u\|_{W^{k,p}(U)}$.

(iii) (Support prescription) $\operatorname{supp} \mathcal{E}u \subseteq V$.

Proof. Observe that, by the previous approximation theorem, it suffices to consider $u \in C^\infty(\overline{U})$ (by density and the boundedness property (ii)).

Step 1: (Reduction to the half-ball case) As in Step 1 in the proof of the previous theorem, construct the open sets $U_0, U_1, \dots, U_K$ and the partition of unity $\chi_0, \chi_1, \dots, \chi_k$. Define $u_k = \chi_k u$, and observe that

- u_0 is already in $W^{k,p}(\mathbb{R}^d)$ and $\operatorname{supp} u_0 \subseteq U_0 \subseteq V$,
- $u_k \in C^\infty(\overline{U})$, and $\operatorname{supp} u_k \subseteq B_{r_0} \subseteq U_k \cap U$.

Observe that if we change variables

$$\begin{cases} y^j = x^j - x_0^j & \text{for } j = 1, \dots, d-1, \\ y^d = x^d - \gamma(x^1, \dots, x^{d-1}), \end{cases}$$

then $U_k \cap U$ gets mapped into $\{y \in B_{\tilde{r}}(0) : y^d > 0\}$.

Note that the change of variables $x \mapsto y$ is C^k, and u_k is smooth, so $u_k(y) = u_k(x(y))$ satisfies, by the chain rule,

$$\|u_k(y)\|_{W_y^{k,p}(\widetilde{U})} \leq C \|u_k(x)\|_{W_x^{k,p}}.$$

Step 2: (Extension in the half-ball case) Now we have $U = B_r^+(0)$, $W = B_{r/2}^+(0)$, and $\operatorname{supp} u \subseteq W$, and we want to extend u. The idea is the *higher order reflection method.* Define

$$\widetilde{u} = \widetilde{u} = \begin{cases} u & x^d > 0 \\ \sum_{j=0}^K \alpha_j u(x^1, \dots, x^{d-1}, -\beta_j x^d) & x^d < 0, \end{cases}$$

where the scaling factor $0 < \beta_j < 1$ is chosen so that $(x^1, \dots, x^{d-1}, -\beta_j x^d) \in B_r^+(0)$. We need to match the normal derivatives on $\{x^d = 0\}$ up to order k. Observe that

$\partial^j_{x^d}(u(x^1,\ldots,x^{d-1}-\beta_j x^d)) = (-1)^j\beta_j^j(\partial^j_{x^d}u)(x^1,\ldots,x^{d-1},-\beta_j x^d)$. We get

$$\begin{cases} u(x^1,\ldots,x^{d-1},0+) = \sum_{j=0}^K \alpha_j u(x^1,\ldots,x^{d-1},0-), \\ \partial_{x^d}u(x^1,\ldots,x^{d-1},0+) = \sum_{j=0}^k \alpha_j(-\beta_j)(\partial_{x^d}u)(x^1,\ldots,x^{d-1},0+) \\ \vdots \\ \partial^k_{x^d}u(x^1,\ldots,x^{d-1},0+) = \sum_{j=0}^k \alpha_j(-\beta_j)^k(\partial^k_{x^d}u)(x^1,\ldots,x^{d-1},0+). \end{cases}$$

This is equivalent to

$$\begin{cases} 1 = \sum_{j=0}^K \alpha_j \\ 1 = \sum_{j=0}^K \alpha_j(-\beta_j) \\ \vdots \\ 1 = \sum_{j=0}^K \alpha_j(-\beta_j)^K. \end{cases}$$

Written in matrix form, this is a linear system involving a **Vandermonde matrix**

$$\begin{bmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{bmatrix} = \begin{bmatrix} 1 & 1 & \cdots & 1 \\ -\beta_0 & -\beta_2 & \cdots & -\beta_K \\ \vdots & \vdots & \vdots & \vdots \\ (-\beta_0)^K & (-\beta_2)^K & \cdots & (-\beta_K)^K \end{bmatrix} \begin{bmatrix} \alpha_0 \\ \alpha_1 \\ \vdots \\ \alpha_K \end{bmatrix}.$$

Now use that fact that if all the β_j are distinct, then this matrix is invertible. This means that there is a choice of $(\alpha_0,\ldots,\alpha_K)$ so that these equations hold. This defines $\widetilde{u}$ on $B_r(x)$ which extends u and matches all derivatives up to order K on the boundary $\{x^d = 0\}$. Finally, put an appropriate smooth cutoff $\chi_V = 1$ on U with $\operatorname{supp}\chi_V \subseteq V$ to define $\mathcal{E}u$, i.e. $\mathcal{E}u = \chi_V\widetilde{u}$. □

4 Trace and Extension Theorems and Introduction to Sobolev Inequalities

Today, we will discuss

(i) trace and extension (from the boundary) theorems

(ii) Sobolev inequalities.

4.1 The trace theorem

Let U be an open subset of $\mathbb{R}^d$ with ∂U being C^1 and $1 < p < \infty$. Recall that for any integer $k \geq 0$, $C^\infty(\overline{U})$ is dense in $W^{k,p}$. In particular, $C^\infty(\overline{U})$ is dense in $W^{1,p}(U)$. Our aim is to discuss the restriction of $u \in W^{1,p}(U)$ to ∂U. Since the boundary is a measure 0 set, this is hard to specify directly (as L^p functions are only well-defined modulo null sets), so we will achieve this by appealing to the dense subset $C^\infty(\overline{U})$.

Definition 4.1. For $u \in C^1(\overline{U})$, we define the **trace** to be $\operatorname{tr}_{\partial U} u = u|_{\partial U}$.

We wish to extend this operation to all of $W^{1,p}(U)$. Note that $\operatorname{tr}_{\partial U}$ is linear, so we can extend it if we know it is bounded.

Theorem 4.1 (Trace theorem, non-sharp)**.** *Let U be a bounded, open subsets of $\mathbb{R}^d$ with C^1 boundary ∂U, and let $1 < p < \infty$. Then for $u \in C^1(\overline{U})$, we have*

$$\| \operatorname{tr}_{\partial U} u\|_{L^p(\partial U)} \leq C \|u\|_{W^{1,p}(U)}.$$

(i) As a consequence, $\operatorname{tr}_{\partial U}$ is extended (uniquely) by continuity (and density of $C^1(\overline{U}) \subseteq W^{1,p}(U)$) to $\operatorname{tr}_{\partial U} : W^{1,p}(U) \to L^p(\partial U)$.

(ii) Moreover, $u \in W_0^{1,p}(U) \iff \operatorname{tr}_{\partial U} u = 0$.

Remark 4.1. The the map $\operatorname{tr}_{\partial U} : W^{1,p}(U) \to L^p(\partial U)$ is not surjective.

Instead of proving this theorem (and you can check the proof in section 5.5 of Evans' book), we will understand the sharp trace theorem in a restricted setting.

The setting we have in mind is $p = 2$. The advantage here is that we may use the theory of the Fourier transform and Plancherel's theorem. We will also focus on the domain $U = \mathbb{R}^d_+ = \{x \in \mathbb{R}^d : x^d > 0\}$ with boundary $\{(x', 0) \in \mathbb{R}^d\} \cong \mathbb{R}^{d-1}$, where $x' := (x^1, \dots, x^{d-1})$.

Recall the Fourier transform characterization of the H^k norm:

$$\|u\|_{H^k}^2 \simeq \|(1 + |\xi|^2)^{k/2} \widehat{u}\|_{L^2_\xi}^2, \qquad k \geq 0 \text{ an integer.}$$

If we replace k with any $s \in \mathbb{R}$, we can talk about **fractional (L^2-based) Sobolev spaces**.

Theorem 4.2 (Sharp trace theorem)**.** *For* $u \in C^1(\overline{\mathbb{R}^d_+}) \cap H^1(\mathbb{R}^d_+)$,

$$\| \operatorname{tr}_{\partial U} u\|_{H^{1/2}(\mathbb{R}^{d-1})} \leq C\|u\|_{H^1(\mathbb{R}^d_+)}.$$

Proof. Take $u \in C^1(\overline{\mathbb{R}^d_+}) \cap H^1(\mathbb{R}^d_+)$. Using the extension procedure from last time, we can find a $\widetilde{u} \in C^1(\mathbb{R}^d)$ such that

$$\|\widetilde{u}\|_{H^1(\mathbb{R}^d)} \leq C\|u\|_{H^1(\mathbb{R}^d_+)}.$$

Then

$$\begin{aligned}
\operatorname{tr}_{\partial U} u(x') &= u(x',0) \\
&= \widetilde{u}(x',0) \\
&= \int \mathcal{F}_{x^d}\widetilde{u}(x',\xi_d)\frac{1}{2\pi}d\xi_d.
\end{aligned}$$

On the other hand,

$$\mathcal{F}_{x'} \operatorname{tr}_{\partial U} u(\xi') = \int \mathcal{F}\widetilde{u}(\xi',\xi_d)\frac{1}{2\pi}\,d\xi_d.$$

For now, let us not assume $s = 1/2$ so we can see where this choice comes from.

$$\begin{aligned}
\| \operatorname{tr}_{\partial U} u\|_{H^s(\mathbb{R}^{d-1})} &\sim \|(1+|\xi'|^2)^{s/2}\mathcal{F}_{x'} \operatorname{tr} u(\xi')\|_{L^2_{\xi'}} \\
&= \left\|(1+|\xi'|^2)^{s/2}\int \mathcal{F}\widetilde{u}(\xi',\xi_d)\frac{1}{2\pi}\,d\xi_d\right\|_{L^2_{\xi'}}
\end{aligned}$$

Writing $|\xi|^2 = |\xi'|^2 + \xi_d^2$,

$$= \left\|\int \frac{(1+|\xi'|^2)^{s/2}}{(1+|\xi'|^2+\xi_d^2)^{1/2}}((1+|\xi'|^2+\xi_d^2)^{1/2}\mathcal{F}\widetilde{u})\frac{1}{2\pi}\,d\xi_d\right\|_{L^2_{\xi'}}$$

Applying Cauchy-Schwarz,

$$\begin{aligned}
&\leq \left\|\left(\int \frac{(1+|\xi'|^2)^s}{1+|\xi'|^2+\xi_d^2}\,d\xi_d\right)^{1/2} \|(1+|\xi|'|^2+\xi_d^2)^{1/2}\mathcal{F}\widetilde{u}\|_{L^2_{\xi_d}}\right\|_{L^2_{\xi'}} \\
&\leq \sup_{\xi'\in\mathbb{R}^{d-1}} \left(\int \frac{(1+|\xi'|^2)^s}{1+|\xi'|^2+\xi_d^2}\,d\xi_d\right)^{1/2} \underbrace{\|\|(1+|\xi|'|^2+\xi_d^2)^{1/2}\mathcal{F}\widetilde{u}\|_{L^2_{\xi_d}}\|_{L^2_{\xi}}}_{\|u\|_{H^1}}.
\end{aligned}$$

For what s is this supremum $< +\infty$? This is $s \leq 1/2$. $\square$

4.2 Extension from the boundary

It turns out that the image of $\mathrm{tr}_{\partial U}$ is *exactly* $H^{1/2}$.

Theorem 4.3 (Extension from ∂U)**.** *There exists a bounded linear map*

$$\mathrm{ext}_{\partial U} : H^{1/2}(\mathbb{R}^{d-1}) \to H^1(\mathbb{R}^d_+)$$

such that $\mathrm{tr}_{\partial U} \circ \mathrm{ext}_{\partial U} = \mathrm{id}$.

Proof. We will use the Poisson semigroup. Suppose we are given $g \in \mathcal{S}(\mathbb{R}^{d-1})$, and let $\eta \in C_c^\infty(\mathbb{R})$ be such that $\eta = 1$ for $|s| < 1$ and $\eta = 0$ for $|s| > 2$. Define $u = \mathrm{ext}_{\partial U}\, g$ by

$$\mathcal{F}_{x'} u(\xi', x^d) = \eta(x^d) e^{-x^d|\xi'|} \widehat{g}(\xi).$$

This right term is the solution to the Laplace equation on the half-space with boundary data g.

We need to show that

$$u \in H^1(\mathbb{R}^d_+) \iff \begin{array}{l} \text{(i) } u, \partial_1 u, \dots, \partial_{d-1} u \in L^2 \\ \text{(ii) } \partial_d u \in L^2. \end{array}$$

(i) implies:

$$\begin{aligned} \|u\|_{L^2}^2 + \|\partial_1 u\|_{L^2}^2 + \cdots + \|\partial_{d-1} u\|_{L^2}^2 &= \|(1 + |\xi'|^2)^{1/2} \mathcal{F}_{x'} u(\xi', x^d)\|^2_{L^2_{\xi'} L_{x^d}} \\ &= \|(1 + |\xi'|^2)^{1/2} \eta(x^d) e^{-x^d|\xi'|} \widehat{g}(\xi')\|^2_{L^2_{\xi'} L^2_{x^d}} \end{aligned}$$

We can integrate in any order, so integrate the x^d integral first.

$$= \| \underbrace{(1 + |\xi'|^2)^{1/4} \|\eta(x^d) e^{-x^d|\xi'|}\|_{L^2_{x^d}}}_{\text{NTS this is unif. bdd. } \xi' \in \mathbb{R}^{d-1}} (1 + |\xi'|^2)^{1/4} \widehat{g}(\xi') \|^2_{L^2_{\xi'}}$$

We can use the bound

$$\|\eta(x^d) e^{-x^d|\xi'|}\|^2_{L^2_{x^d}} \lesssim 1,$$

and the substitution bound

$$\int \eta^2(x^d) e^{-2x^d|\xi'|}\, dx^d \lesssim \frac{1}{|\xi'|}.$$

This gives

$$\|\eta(x^d) e^{-x^d|\xi'|}\|_{L^2_{x^d}} \lesssim \min\left\{1, \tfrac{1}{|\xi'|^{1/2}}\right\} \lesssim (1 + |\xi'|)^{-1/2}.$$

(ii) implies:

$$\partial_{x^d} u = \partial_{x^d}(\eta(x^d) v), \qquad \mathcal{F}_{x'} v = e^{-x^d|\xi'|} \widehat{g}(\xi)$$

$$= \eta'(x^d)v + \eta\partial_{x^d}v.$$

The norm of the first term is bounded by $\|v\|_{L^2(x^d \in \operatorname{supp}\eta)}$, and the norm of the second term is

$$\begin{aligned}
\|\eta\partial_{x^d}v\|_{L^2_{x'}L^2_{\xi_d}} &= \|\eta\partial_{x^d}(e^{-x^d|\xi'|}\widehat{g}(\xi'))\|_{L^2_{\xi'}L^2_{x^d}} \\
&= \||\xi'|\underbrace{e^{-x^d|\xi'|}\widehat{g}(\xi')\eta(x^d)}_{\mathcal{F}_{\xi'}u}\|_{L^2_{\xi'}L^2_{x^d}}
\end{aligned}$$

Using (i),

$$\leq C\|g\|_{H^{1/2}}. \qquad \square$$

Remark 4.2. In fact, by the usual smooth partition of unity argument with boundary straightening, one can define $H^{1/2}(\partial U)$ for ∂U of class C^1 and prove the sharp trace theorem. The independence of this space from the smooth partition of unity and boundary straightening follows from interpolation theory (which you can find in the 1970 textbook of Stein).

Remark 4.3. For $p \neq 2$, $\operatorname{im}(\operatorname{tr}_{\partial U} W^{1,p}(U)) = B_p^{1-1/p,p}(\partial U)$. This is called the **$L^p$-Besov space of order $1-1/p$ and summability index p**. This is also covered in Stein's book.

4.3 The Gaglierdo-Nirenberg-Sobolev inequality and the Loomis-Whitney inequality

In a nutshell, Sobolev inequalities are a quantitative generalization of the fundamental theorem of calculus; we know the size of the derivative of a function, and we want to control the size of the function.

Theorem 4.4 (Gaglierdo-Nirenberg-Sobolev inequality)**.** *Let $d \geq 2$. For $u \in C_c^\infty(\mathbb{R}^d)$, we have*

$$\|u\|_{L^{\frac{d}{d-1}}}(\mathbb{R}^d) \leq C_d\|Du\|_{L^1(\mathbb{R}^d)},$$

where C_d is a constant depending only on d.

Remark 4.4. The exponent on the left hand side need not be remembered because it can be derived from scaling considerations (dimensional analysis). In particular, first observe that both sides are homogeneous: if $u \mapsto u_\lambda(x) = u(x/\lambda)$ for $\lambda > 0$, then

$$\begin{aligned}
\|u_\lambda\|_{L^p} &= \left(\lambda^d \underbrace{\int \left|u\left(\frac{x}{\lambda}\right)\right|^p \frac{1}{\lambda^d}\,dx}_{=\int |u|^p\,dx'}\right)^{1/p} \\
&= \lambda^{d/p}\|u\|_{L^p}.
\end{aligned}$$

On the other hand, $D(u_\lambda) = \frac{1}{\lambda}(Du)_\lambda$, so

$$\|D(u_\lambda)\|_{L^p} = \frac{1}{\lambda}\lambda^{d/p}\|Du\|_{L^p}.$$

Now compare these:

$$\begin{aligned}
\|u_\lambda\|_{L^p} \le c\|Du_\lambda\|_{L^1} \qquad \forall \lambda > 0 &\iff \lambda^{d/p}\|u\|_{L^p} \le c\lambda^{-1+d}\|Du\|_{L^1} \qquad \forall \lambda > 0 \\
&\iff \frac{d}{p} = d - 1 \\
&\iff p = \frac{d}{d-1}.
\end{aligned}$$

All we are doing here is changing the unit of length and requiring that the inequality is invariant under our unit of length.

We will prove this next time. The key ingredient is another inequality. Denoting $(x^1, \dots, \widehat{x}^j, \cdots, x^d) = (x^1, \dots, x^{j-1}, x^{j+1}, \dots, x^d)$, we have the following.

Lemma 4.1 (Loomis-Whitney inequality)**.** *Let $d \ge 2$. For $j = 1, \dots, d$, suppose $f_j = f_j(x^1, \dots, \widehat{x}^j, \dots, x^d)$. Then*

$$\left\| \prod_{j=1}^d f_j \right\|_{L^1(\mathbb{R}^d)} \le \prod_{j=1}^d \|f_j\|_{L^{d-1}(\mathbb{R}^{d-1})}.$$

Proof. Integrate in each variable and apply Hölder:

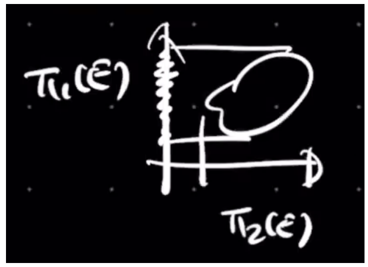
$$\begin{aligned}
\int \left| \prod_{j=1}^d f_j \right| dx^1 &= |f_1| \int |f_2| \cdots |f_d| \, dx^1 \\
&\le |f_1| \|f_2\|_{L^{d-1}_{x^1}} \cdots \|f_d\|_{L^{d-1}_{x^1}}
\end{aligned}$$

This is a function of $(x^2, \dots, x^d)$. Now integrate with respect to the next variable:

$$\begin{aligned}
\iint \left| \prod_{j=1}^d f_j \right| dx^1 \, dx^2 &\le \int |f_1| \|f_2\|_{L^{d-1}_{x^1}} \cdots \|f_d\|_{L^{d-1}_{x^1}} \, dx^2 \\
&= \|f_2\|_{L^{d-1}_{x^1}} \|f_1\|_{L^{d-1}_{x^2}} \|f_3\|_{L^{d-1}_{x^1,x^2}} \cdots \|f_d\|_{L^{d-1}_{x^1,x^2}}.
\end{aligned}$$

Iterating this gives the inequality. $\square$

Remark 4.5. The Loomis-Whitney inequality answers the following geometric question. Suppose $E \subseteq \mathbb{R}^d$, and know the projections $\pi_j(E)$. Can we bound the measure of E by $|\pi_j(E)|$?



Yes!

$$
\begin{aligned}
|E| &= \int \mathbb{1}_E \, dx \\
&\leq \int \prod_{j=1}^{d} \mathbb{1}_{\pi_j(E)}(x^1, \ldots, \widehat{x^j}, \ldots, x^d) \, dx \\
&\overset{\text{L-W}}{\leq} \prod_{j=1}^{d} |\pi_j(E)|^{\frac{1}{d-1}}.
\end{aligned}
$$

5 Sobolev Inequalities

5.1 The Gagliardo-Nirenberg-Sobolev inequality

We have been discussing Sobolev inequalities. Last time, we stated the following theorem.

Theorem 5.1 (Gagliardo-Nirenberg-Sobolev inequality)**.** *Let $d \geq 2$. For $u \in C_c^\infty(\mathbb{R}^d)$, we have*

$$\|u\|_{L^{\frac{d}{d-1}}}(\mathbb{R}^d) \leq \|Du\|_{L^1(\mathbb{R}^d)}.$$

To approach this, we proved a lemma:

Lemma 5.1 (Loomis-Whitney inequality)**.** *Let $d \geq 2$. For $j = 1, \dots, d$, suppose $f_j = f_j(x^1, \dots, \widehat{x^j}, \dots, x^d)$. Then*

$$\left\| \prod_{j=1}^d f_j \right\|_{L^1(\mathbb{R}^d)} \leq \prod_{j=1}^d \|f_j\|_{L^{d-1}(\mathbb{R}^{d-1})}.$$

This answers the geometric question of controlling the measure of a set in $\mathbb{R}^d$ using the measure of its projections, by applying the lemma to $f_j = \mathbb{1}_{\pi_{x^j}(E)}$. Now let's prove the GNS inequality.

Proof. Observe that if we take a point $x \in \mathbb{R}^d$, then we can write

$$u(x) = \int_{-\infty}^{x^j} \partial_{x^j} u(x^1, \dots, x^{j-1}, y, x^{j+1}, \dots, x^d)\, dy,$$

using the fundamental theorem of calculus. Here, we use the compact support assumption to be sure this converges. This means that

$$|u(x)| \leq \int_{-\infty}^{x^j} |\partial_{x^j} u(x^1, \dots, x^{j-1}, y, x^{j+1}, \dots, x^d)| dy.$$

We can upper bound this by replacing x^j by ∞ and ∂_{x^j} by D:

$$|u(x)| \leq \underbrace{\int_{-\infty}^{\infty} |Du(x^1, \dots, x^{j-1}, y, x^{j+1}, \dots, x^d)| dy}_{\widetilde{f}_j(x^1, \dots, \widehat{x}^j, \dots, x^d)}.$$

This means that we have

$$|u(x)| \leq \left(\prod_{j=1}^d \widetilde{f}_j \right),$$

which we can write as

$$|u(x)|^{\frac{d}{d-1}} \le \left(\prod_{j=1}^{d} \widetilde{f}_j^{\frac{1}{d-1}} \right),$$

Using the Loomis-Whitney inequality,

$$\begin{aligned}
\|u\|_{L^{\frac{d}{d-1}}}^{\frac{d}{d-1}} &= \int |u|^{\frac{d}{d-1}}\, dx \\
&\le \int \prod_{j=1}^{d} f_j \, dx \\
&\le \prod_{j=1}^{d} \|f_j\|_{L^{d-1}} \\
&= \prod_{j=1}^{d} \left(\int |f_j|^{d-1}\, dx^1 \cdots \widehat{dx^j} \cdots dx^d \right)^{\frac{1}{d-1}}
\end{aligned}$$

Observe that $|f_j|^{d-1} = \int_{-\infty}^{\infty} |Du(x^1, \ldots, x^j, \ldots, x^d)|\, dx^j = \int |Du|\, dx$, so

$$\le \|Du\|_{L^1}^{\frac{d}{d-1}}. \qquad \square$$

Remark 5.1. GNS is the functional counterpart of the isoperimetric inequality. Given a function, we can make a layer cake decomposition in the y axis and apply the isoperimetric inequality to each part. This is useful for functions on manifolds where we have some geometric information.

5.2 Sobolev inequalities for L^p-based spaces with $p < d$

Now we will upgrade this to the case where we have other L^p spaces on the right hand side.

Theorem 5.2 (Sobolev inequalities for L^p-based spaces)**.** *Let $d \ge 2$, and assume that $1 < p < d$. For $u \in C_c^\infty(\mathbb{R}^d)$, we have*

$$\|u\|_{L^q}(\mathbb{R}^d) \le C\|Du\|_{L^p(\mathbb{R}^d)},$$

where $q = \frac{dp}{d-p}$.

What is q? We do dimensional analysis to figure out the exponent. On the left hand side, we have $[x]^{d/q}$, and on the right hand side, we have $[x]^{-1+d/p}$. If we solve for q, we get $q = \frac{dp}{d-p}$. This also gives us the restriction that $p < d$.

Proof. Take $v = |u|^{\widetilde{q}}$, where $\widetilde{q} = \frac{q}{d/(d-1)}$. Its derivative is $|Dv| = q|u|^{q-1}|Dv|$. This can be justified using approximation: approximate $|x|$ by $(\varepsilon^2 + x^2)^{1/2}$v and let $\varepsilon \to 0$. Then

$$\int |u|^{\widetilde{q}}\, dx = \int |v|^{\frac{d}{d-1}}\, dx$$

Using the GNS inequality,

$$\leq \left(\int |Dv|\, dx \right)^{\frac{d-1}{d}}.$$

It is at this point that we need the above approximation. But it works, using the dominated convergence theorem.

$$= \left(\int |u|^{\widetilde{q}-1} |Du|\, dx \right)^{\frac{d-1}{d}}$$

Using Hölder's inequality, we can put $|Du|$ into L^p, which puts $|u|^{\widetilde{}-1}$ in $L^{p'}$. By dimensional analysis, it must happen that

$$\leq \|u\|_{L^q}^{\frac{d-1}{d}(q-1)} \|Du\|_{L^p}^{\frac{d-1}{d}}.$$

This completes the proof. □

Now we will upgrade this to every element in the abstract Sobolev space and to situations where we have a function which is bounded on an abstract domain.

Theorem 5.3. *Let $d \geq 2$, and assume that $1 \leq p < d$.*

(i) For $u \in W^{1,p}(\mathbb{R}^d)$,

$$\|u\|_{L^q}(\mathbb{R}^d) \leq C\|Du\|_{L^p(\mathbb{R}^d)},$$

where $q = \frac{dp}{d-p}$.

(ii) Let U be a bounded domain. For $u \in W_0^{1,p}(U)$,

$$\|u\|_{L^q}(U) \leq C\|Du\|_{L^p(U)},$$

where $q = \frac{dp}{d-p}$.

(iii) Let U be a bounded domain with C^1 boundary ∂U. Then for $u \in W^{1,p}(U)$,

$$\|u\|_{L^q}(U) \leq C\|Du\|_{W^{1,p}(U)},$$

where $q = \frac{dp}{d-p}$.

Proof.

(i) This is by density of $C_c^\infty(\mathbb{R}^d)$.

(ii) This is by density, as well.

(iii) This follows from extension and approximation. $\square$

Remark 5.2. In (iii), we need both $\|u\|_{L^p}$ and $\|Du\|_{L^p}$ in the extension procedure. Compare this to the case (ii), where no information of u was needed, since "$u|_{\partial U} = 0$." By this reason, (ii) is called a **Poincaré inequality** or **Friedrich inequality**.

5.3 Sobolev inequalities for L^p-based spaces with $p > d$

Next, we investigate: What does $\|u\|_{W^{1,p}}$ tell us when $p \geq d$? This will be based on another way to relate u with its derivative, Du. Start with $u \in C^\infty(\mathbb{R}^d)$, and write down what we get by applying the fundamental theorem of calculus:

$$u(x) - u(y) = \int_0^1 = \frac{d}{ds} u(x + s(y-x))\, dx.$$

The key idea is to average to take advantage of the fact that we are in multiple dimensions. Take absolute values and average this in y: Fix $r > 0$, so

$$\frac{1}{|B_r(x)|} \int_{B_r} |u(x) - u(y)|\, dy \leq \frac{1}{|B_r(x)|} \int_{B_r(x)} \int_0^1 \left| \frac{d}{ds} u(x + s(y-x)) \right| dx\, dy$$

By the chain rule, this derivative is $(y - x) \cdot Du(x + s(y-x))$.

$$\leq C \frac{1}{r^d} \int_{B_r(x)} \int_0^1 |x - y||Du(x + s(y-x))|\, dx\, dy$$

Let $\rho\omega = y - x$, so that $\rho = |y - x|$.

$$= C \frac{1}{r^d} \int_0^r \int_{\mathbb{S}^{d-1}} \int_0^1 \rho |Du(x + s\rho\omega)|\, ds \rho^{d-1}\, d\omega\, d\rho$$

Make another change of variables, so we can make $x + s\rho\omega$ into an actual radius and then evaluate on of the integrals. We do $t = s\rho$

$$= C \frac{1}{r^d} \int_0^r \int_{\mathbb{S}^{d-1}} \int_0^1 \frac{t^d}{s^d} \frac{1}{s} |Du(x + t\omega)|\, ds\, d\omega\, dt$$

Simplify the s integral and upper bound $t \leq r$:

$$\leq C \int_0^r \int_{\mathbb{S}^{d-1}} |Du(x + t\omega)|\, d\omega\, dt$$

$$= C \int_{B_r(x)} \frac{|Du|}{|x - y|^{d-1}}\, dy.$$

We can summarize this as a lemma:

Lemma 5.2. *Let $p > d$, let $d \geq 2$, and let $u \in C^\infty(\mathbb{R}^d)$. Then*

$$\frac{1}{|B_r(x)|} \int_{B_r} |u(x) - u(y)| \, dy \leq C \int_{B_r(x)} \frac{|Du|}{|x-y|^{d-1}} \, dy.$$

Theorem 5.4. *Let $p > d$ with $d \geq 2$, and take $u \in C^\infty(\mathbb{R}^d)$. Then*

$$|u(x) - u(y)| \leq C|x-y|^\alpha \|Du\|_{L^p(\mathbb{R}^d)},$$

where $\alpha = 1 - \frac{d}{p}$.

Again, we can find the value of α by dimensional analysis: $1 = \alpha + (-1) + \frac{d}{p}$ gives $\alpha = 1 - \frac{d}{p}$.

Proof. We will use the lemma. The idea is to introduce an auxiliary variable z and take the average over z on some domain U:

$$\frac{1}{|U|} \int_U |u(x) - u(y)| \, dz \leq \frac{1}{|U|} \int_U |u(x) - u(z)| \, dz + \frac{1}{|U|} \int_U |u(y) - u(y)| \, dz$$

Since $\frac{|B_r(x)|}{|U|} \simeq 1$,

$$\begin{aligned}
&\lesssim \frac{|B_r(x)|}{|U|} \int_{B_r(x)} |u(x) - u(z)| \, dz + \frac{|B_r(y)|}{|U|} \int_{B_r(y)} |u(y) - u(z)| \, dz \\
&\lesssim \int_{B_r(x)} \frac{|Du|}{|x-z|^{d-1}} \, dz + \int_{B_r(y)} \frac{|Du|}{|y-z|^{d-1}} \, dz \\
&\lesssim \|Du\|_{L^p} \left\| \frac{1}{|x-z|^{d-1}} \right\|_{L^{p'}(B_r(x))} + \|Du\|_{L^p} \left\| \frac{1}{|y-z|^{d-1}} \right\|_{L^{p'}(B_r(y))}
\end{aligned}$$

Now we just need to evaluate

$$\int_{B_r(0)} \frac{1}{|z|^{(d-1)p'}} \, dz \simeq r^\alpha. \qquad \square$$

5.4 Sobolev inequalities for L^p-based spaces with $p = d$

What about when $p = d$ (and $d \geq 2$)? In this case, the inequality $\|u\|_{L^\infty(U)} \leq \|u\|_{W^{1,d}(U)}$ fails.

Example 5.1. Here is a counterexample to the above inequality when $p = d = 2$. Take $U = B_1(0) \subseteq \mathbb{R}^2$ and

$$u(x) = \log\log\left(10 + \frac{1}{|x|}\right).$$

A popular remedy for $p = d$ is to think about bounded mean oscillation:

Definition 5.1. $u \in C^\infty$ has **bounded mean oscillation (BMO)** if

$$\|u\|_{\mathrm{BMO}} = \sup_{\substack{x\in\mathbb{R}^d \\ r>0}} \frac{1}{|B_r(x)|}\int_{B_r(x)} \left| u(y) - \frac{1}{|B_r(x)|}\int_{B_r(x)} u \right| \, dy < \infty.$$

We can check that $\|u\|_{\mathrm{BMO}} \leq C\|Du\|_{L^d}$. We will discuss this next time and also introduce the concept of **Hölder space** to recontextualize the theorem we have just proven.

6 Hölder Spaces, Bounded Mean Oscillation, and Compact Operators

6.1 Hölder spaces

Let's continue our discussion of Sobolev inequalities. We want to know: What does $\|u\|_{W^{1,p}}$ say about u when $p \geq d$? We proved a lemma:

Lemma 6.1. *Suppose $u \in C^\infty(\mathbb{R}^d)$ with $d \geq 2$. Then*

$$\frac{1}{|B_r(x)|} \int_{B_r(x)} |u(x) - u(z)| \, dz \leq C \int_{B_r(x)} \frac{|Du(z)|}{|z - x|^{d-1}} \, dz$$

From this lemma, we saw the following theorem:

Theorem 6.1. *Let $u \in C^\infty(\mathbb{R}^d)$ with $d \geq 2$, and let $x, y \in B_R$. Then*

$$|u(x) - u(y)| \leq C|x - y|^\alpha \|Du\|_{L^p(B_R)},$$

where $\alpha = 1 - \frac{d}{p}$.

We want to rephrase this as an inequality for $u \in W^{1,p}(U)$. To do this, we need a space that has a regularity property relating to the theorem above.

Definition 6.1. Let $u \in C(I)$. The **Hölder seminorm** of order α is

$$[u]_{C^\alpha(U)} = \sup_{x \neq y} \frac{|u(x) - u(y)|}{|x - y|^\alpha}.$$

By a **seminorm**, we mean that $[\cdot]_{C^\alpha(U)}$ satisfies all the properties of a norm except the property that $[u]_{C^\alpha(U)} = 0 \implies u = 0$. Instead, this implies that u is constant. Here is how we make it into a norm

Definition 6.2. The **Hölder norm** of order α is

$$\|u\|_{C^\alpha(U)} = [u]_{C^\alpha(U)} + \|u\|_{L^\infty}.$$

The **Hölder space** of order α is

$$C^\alpha(U) = \{u \in C(U) : \|u\|_{C^\alpha} < \infty\}.$$

Theorem 6.2 (Morrey's inequality[1])**.** *Let $d \geq 2$, let $p > d$, and let U be a bounded domain in $\mathbb{R}^d$ with C^1 boundary ∂U. If $u \in W^{1,p}(U)$, then $u \in C^\alpha(U)$ with $\alpha = 1 - \frac{d}{p}$. Moreover,*

$$\|u\|_{C^\alpha(U)} \leq C\|u\|_{W^{1,p}(U)}.$$

[1] This is sometimes called Morey's embedding.

Proof. By extension and density theorems, it suffices to consider $u \in C^\infty(\mathbb{R}^d)$ with $\operatorname{supp} u \subseteq V$, where V is a bounded, open set with $V \supseteq \overline{U}$ (chosen independently of u). By the previous theorem,

$$[u]_{C^\alpha(V)} \leq C\|u\|_{W^{1,p}}.$$

So all that remains is to bound $\|u\|_{L^\infty}$ in terms of $\|u\|_{W^{1,p}}$. For this purpose, we will again use the lemma to approximate u by its average. Let $x \in V$. Then

$$\begin{aligned}
\left| u(x) - \frac{1}{|B_r(x)|}\int_{B_r(x)} u\,dz \right| &\leq \left| \frac{1}{|B_r(x)|}\int_{B_r(x)} u(x) - u(z)\,dz \right| \\
&\leq \frac{1}{|B_r(x)|}\int_{B_r(x)} |u(x) - u(z)|\,dz \\
&\leq C\int_{B_r(x)} \frac{|Du(z)|}{|z-x|^{d-1}}\,dz \\
&\leq Cr^\alpha \|Du\|_{L^p(B_r(x))}.
\end{aligned}$$

Take $r = 1$. Then

$$\begin{aligned}
|u(x)| &\leq C \underbrace{\left| \int_{B_r(x)} u\,dz \right|}_{\leq \int_{B_1(x)} |u|\,dz \leq C\|u\|_{L^p(B_1(0))}} + C\|Du\|_{L^p} \\
&\leq C(\|u\|_{L^p} + \|Du\|_{L^p}).
\end{aligned}$$

□

6.2 Bounded mean oscillation

When $p = d$, $W^{1,d}$ does not embed into L^∞.

Example 6.1. For $d = 2$, let $U = B_1(0)$, and consider

$$u = \log\log\left(10 + \frac{1}{|x|}\right).$$

A useful substitute for the above failure involves the space of bounded mean oscillation (BMO).

Definition 6.3. Let $u \in L^1_{\mathrm{loc}}(U)$. The **BMO seminorm** is

$$[u]_{\mathrm{BMO}} = \sup_{B_r(x_0)\subseteq U} \frac{1}{|B_r(x_0)|}\int_{B_r(x_0)} \left| u(z) - \frac{1}{|B_r(x_0)|}\int_{B_r(x_0)} u \right| dz.$$

Theorem 6.3. *Let $d \geq 2$, $U \subseteq \mathbb{R}^d$, and $u \in W^{1,d}(\mathbb{R}^d)$. Then $[u]_{\mathrm{BMO}} < \infty$, and*

$$[u]_{\mathrm{BMO}} \leq C\|Du\|_{L^d}.$$

Remark 6.1. As an exercise, you can show that $L^\infty \subsetneq \mathrm{BMO}$. The function $u = \mathbb{1}_{B_1(0)} \log|x|$ shows that these spaces are nor equal.

Proof. Assume $u \in C^\infty(\mathbb{R}^d)$. We want to show that

$$[u]_{\mathrm{BMO}} \leq C\|Du\|_{L^d}.$$

Fix $B_r(x)$. We want to show that

$$\frac{1}{|B_r(x)|}\int_{B_r(x)}\left|u(z) - \frac{1}{|B_r(x)|}\int_{B_r(x)} u(y)\,dy\right|dz \leq C\|Du\|_{L^d}.$$

with some fixed constant C. We can rewrite the left hand side to get

$$\begin{aligned}\frac{1}{|B_r(x)|}\int_{B_r(x)}\left|\frac{1}{|B_r(x)|}u(z)\,dy - \frac{1}{|B_r(x)|}\int_{B_r(x)} u(y)\,dy\right|dz& \\ \leq \frac{1}{|B_r(x)|^2}\int_{B_r(x)}\int_{B_r(x)}|u(z)-u(y)|\,dy\,dz&\end{aligned}$$

Since $B_r(x) \subseteq B_{2r}(y)$,

$$\leq \frac{1}{|B_r(x)|^2}\int_{B_r(x)}\int_{B_{2r}(y)}|u(z)-u(y)|\,dy\,dz$$

Using the lemma,

$$\leq \frac{1}{|B_r(x)|^2}\int_{B_r(x)}\underbrace{\int_{B_{2r}(y)}\frac{|Du(z)|}{|z-y|^{d-1}}\,dz}_{F(y)}\,dy$$

This is a convolution, so you might be tempted to use Young's inequality: $\|f * g\|_{L^r} \leq \|f\|_{L^p}\|g\|_{L^q}$, where $1 \leq p \leq q \leq r \leq \infty$ and $1 + \frac{1}{r} = \frac{1}{p} + \frac{1}{q}$. However, this barely fails, since $\frac{1}{|z-x|^{d-1}} \notin L^q$. Instead, we use the following theorem:

Theorem 6.4 (Hardy-Littlewood)**.** *Let $u \in L^1_{\mathrm{loc}}$, and define*

$$\mathcal{M}u(x) = \sup_{r>0}\frac{1}{|B_r(x)|}\int_{B_r(x)}|u|.$$

(Note that $|\mathcal{M}u| \leq \|u\|_{L^\infty}$). For $1 < p \leq \infty$,

$$\|\mathcal{M}u\|_{L^p} \leq C\|u\|_{L^p}.$$

Whenever you are faced with something that is hard to understand, it is a good idea to decompose the region into pieces where the function is mostly constant. The power function $|y|^\alpha$ has the property that if $2^{k-1} \leq |y|, |y'| \leq 2^k$, then $|y|^\alpha \simeq |y'|^\alpha$. For our problem, write $A_k = \{2^{k-1} \leq |z-y| \leq 2^k\}$, so

$$\begin{aligned}
\int_{B_{2r}(y)} \frac{|Du(z)|}{|z-y|^{d-1}} \, dz &\leq C \sum_{2^k \leq 2r+c} \int_{A_k} \frac{1}{(2^k)^{d-1}} |Du(z)| \, dz \\
&\leq C \sum_{2^k \leq 2cr} \frac{1}{(2^k)^{d-1}} \int_{B_{2^k}(y)} |Du(z)| \, dz \\
&\leq C \sum_{2^k \leq 2cr} 2^k \mathcal{M}(|Du|)(y).
\end{aligned}$$

It now suffices to bound

$$\left\| \sum_{2^k \leq 2cr} 2^k \mathcal{M}(|Du|)(y) \right\|_{L^1} \leq Cr \|\mathcal{M}|Du|\|_{L^d} \|1\|_{L^{\frac{d}{d-r}}(B_r(x))}$$

Using the theorem,

$$\leq Cr^d \|Du\|_{L^d}. \qquad \square$$

6.3 Compact operators and embeddings

We will discuss two more topics involving Sobolev spaces:

1. Compactness of Sobolev embedding
2. Poincaré-type inequalities (how to get information about u from $\|Du\|_{L^p}$ given some extra condition for normalizing the function).

Let's set up the discussion for the first topic.

Definition 6.4. Let X, Y be normed spaces, and let $T : X \to Y$ be linear. We say that T is a **compact operator** if $T(B_X)$, the image of the unit ball in X, is compact in Y. Equivalently, we may require that for all bounded $\{x_n\} \subseteq X$, $\{Tx_n\}$ has a convergent subsequence.

Definition 6.5. Suppose that we have an embedding (i.e. a bounded, linear, injective map) $\iota : X \to Y$. We say the embedding $X \subseteq Y$ is **compact** if ι is compact.

We are interested in writing something like this: $W^{1,p}(U) \subseteq L^q(U)$. If we think of $W^{1,p}(U)$ as a subspace of functions, then this embedding will be compact.

What is the basic compactness theorem in the setting of function spaces? We will use the Arzelà-Ascoli theorem:

Theorem 6.5 (Arzelà-Ascoli)**.** *Let K be a compact set and $\mathcal{A} \subseteq C(K)$. Suppose that*

1. *$\mathcal{A}$ is* ***locally bounded****, i.e. for any $x \in K$, there is an $M(x)$ such that for all $f \in \mathcal{A}$, $|f(x)| \leq M(x)$.*

2. *$\mathcal{A}$ is* ***equicontinuous****, i.e. for all $\varepsilon > 0$, there is a $\delta > 0$ such that for all $f \in \mathcal{A}$,*

$$|x - y| < \delta \implies |f(x) - f(y)| < \varepsilon, \qquad \forall x, y \in K.$$

Then $\mathcal{A}$ is compact.

7 Compactness of Sobolev Embeddings and Poincaré-Type Inequalities

7.1 Compactness of embeddings of Hölder spaces into Hölder spaces

Last time we defined the notion of compact operators.

Definition 7.1. Let X, Y be normed spaces, and let $T : X \to Y$ be linear. We say that T is a **compact operator** if $T(B_X)$, the image of the unit ball in X, is compact in Y. Equivalently, we may require that for all bounded $\{x_n\} \subseteq X$, $\{Tx_n\}$ has a convergent subsequence.

The proof will resemble the proof of the Arzelà-Ascoli theorem.

Theorem 7.1 (Arzelà-Ascoli)**.** *Let K be a compact set and $\mathcal{A} \subseteq C(K)$. Suppose that*

1. *$\mathcal{A}$ is **locally bounded**, i.e. for any $x \in K$, there is an $M(x)$ such that for all $f \in \mathcal{A}$, $|f(x)| \leq M(x)$.*
2. *$\mathcal{A}$ is **equicontinuous**, i.e. for all $\varepsilon > 0$, there is a $\delta > 0$ such that for all $f \in \mathcal{A}$,*

$$|x - y| < \delta \implies |f(x) - f(y)| < \varepsilon, \qquad \forall x, y \in K.$$

Then $\mathcal{A}$ is compact.

There is a weaker notion of convergence in $C(K)$, pointwise convergence. The link between pointwise and uniform convergence is given by the equicontinuity assumption. In short, we use extra regularity to help us prove compactness.

Theorem 7.2 (Compactness of $C^{0,\alpha}(U) \subseteq C^{0,\alpha'}(U)$)**.** *Let U be a bounded open subset of $\mathbb{R}^d$, and assume $0 < \alpha' < \alpha < 1$ (so that $C^{0,\alpha}(U) \subseteq C^{0,\alpha'}(U)$). The embedding $C^{0,\alpha}(U) \to C^{0,\alpha'}(U)$ is compact.*

Here is a sketch of the proof.

Proof.

(i) The first observation is to note that the embedding $C^{0,\alpha}(U) \to C(U)$ is compact (this is by Arzelà-Ascoli).

(ii) By (i), if $\{u_n\} \subseteq C^{0,\alpha}(U)$ is bounded: $\|u_n\}_{C^{0,\alpha}} \leq M$, then there is a subsequence u_{n_j} such that $\{u_{n_j}\}$ is convergent in $C(U)$ (to u_∞). We claim that in fact,

$$\|u_{n_j} \to u_\infty\|_{C^{0,\alpha'}(U)} \to 0.$$

The key idea here is **interpolation**. Because

$$\|v\|_{C^{0,\alpha'}} = \|v\|_{L^\infty} + [v]_{C^{0,\alpha'}},$$

we need to show that

$$[v]_{C^{0,\alpha'}} \leq \|v\|_{L^\infty} [v]_{C^{0,\alpha}}^{\alpha'/\alpha},$$

where the α'/α exponent comes from dimensional analysis concerns. If we have this, then

$$[u_{n_j} - u_\infty]_{C^{0,\alpha'}} \leq \underbrace{\|u_{n_j} - u_\infty\|^{1-\alpha'/\alpha}}_{\to 0 \text{ by (i)}} \underbrace{[u_n - u_\infty]_{C^{0,\alpha}}^{\alpha'/\alpha}}_{\text{bdd}}.$$

To prove this inequality, write

$$\frac{|v(x) - v(y)|}{|x-y|^{\alpha'}} \leq (|v(x)| + |v(y)|)^{1-\alpha'/\alpha} \left(\frac{|v(x) - v(y)|}{|x-y|^{\alpha'}} \right)^{\alpha'/\alpha}.$$

Then take the sup over $x, y \in U$ with $x \neq y$ on both sides. □

7.2 Rellich-Kondrachov compactness of embedding Sobolev spaces into L^p spaces

Theorem 7.3 (Rellich-Kondrachov)**.** *Let $d \geq 2$, and let U be a bounded domain in $\mathbb{R}^d$ with C^1 boundary ∂U. (Recall that if $1 \leq p < d$, we have the embedding $W^{1,p}(U) \to L^{p^*}(U)$, where $\frac{d}{p^*} = \frac{d}{p} - 1$.) Let $1 \leq p < d$, and let $1 \leq q < p^*$. Then the embedding $W^{1,p}(U) \to L^q(U)$ is compact.*

As in the discussion of Arzelà-Ascoli, we will approximate a bounded sequence by a part which is compact and leverage some sort of uniform control. Here is a property of mollifiers that will be useful for us: Recall that if $v \in L^p(\mathbb{R}^d)$ and $\varphi \in C_c^\infty(\mathbb{R}^d)$ with $\int \varphi = 1$, $\varphi_\varepsilon * v \to v$ in $L^p(\mathbb{R}^d)$. This is a qualitative statement that doesn't tell us how fast this converges with respect to ε. However, if we have more information, we can rectify this.

Lemma 7.1 (Accelerated convergence of modifiers)**.** *Let $1 \leq p < \infty$, and suppose $v \in W^{k,p}$. Choose $\varphi \in C_c^\infty(\mathbb{R}^d)$ such that $\int \varphi \, dx = 1$ and $\int x^\alpha \varphi \, dx = 0$ for all $1 \leq |\alpha| < k$.*[2] *Then*

$$\|\varphi_\varepsilon * v - v\|_{L^p} \leq C\varepsilon^k \|\partial^{(k)} v\|_{L^p}.$$

Here is the proof of this lemma when $k = 2$. The argument is the same for other values of k.

[2]The conditions $\int x^\alpha \varphi \, dx = 0$ are called **moment conditions**.

Proof. First, write

$$\int \varphi_\varepsilon(y)v(x-y)\,dy - \underbrace{v(x)}_{=\int \varphi_\varepsilon(y)v(x)\,dy} = \int \varphi_\varepsilon(y)(v(x-y)-v(x))\,dy.$$

Here, we should think of $|y| \lesssim \varepsilon$. To quantify the convergence of the v part, we Taylor expand in y. We will be using the integral form of the Taylor expansion with remainder.[3]Write

$$\begin{aligned}\int_0^1 \frac{d}{ds}v(x-sy)\,ds &= -\int \frac{d}{ds}(1-s)\frac{d}{ds}v(x-sy)\,dx \\ &= \frac{d}{ds}v(x-sy)\Big|_{s=0} + \int_0^1 (1-s)\frac{d^2}{ds^2}v(x-sy)\;ds.\end{aligned}$$

The first term gves $y \cdot \nabla v(x)$, and the second term gives $y^i y^j \int_0^1 (1-s)\partial_i \partial_j v(x-sy)\,ds$. The contribution of the first term is 0 by the moment condition, and we are left with the remainder, which we can control. In all, we get

$$\left| \int \varphi_\varepsilon(y)v(x-y)\,dy - v(x) \right| \leq \int |\varphi_\varepsilon(y)||y|^2 \int_0^1 |\partial^2 \varphi(x-sy)|\,ds\,dy.$$

This tells us that

$$\begin{aligned}\|\cdot\|_{L^p} &\leq \|\partial^2 v\|_{L^p} \int |\varphi_\varepsilon(y)| \underbrace{|y|^2}_{\lesssim \varepsilon^2}\,dy \\ &\lesssim \varepsilon^2 \|\partial^2 v\|_{L^p}.\end{aligned}$$

□

Now let's prove the theorem.

Proof.

Step 1: Reduce to the compactness of $W^{1,p}(U) \to L^p(U)$. This is sufficient because of the following two cases:

Case 1: $W^{1,p} \to L^q(U)$ with $1 \leq q \leq p$. In this case, if U is bounded, then Hölder gives $\|v\|_{L^q(U)} \leq |U|^{1/q-1/p}\|v\|_{L^p}$, and we already have control in L^p.

Case 2: $W^{1,p} \to L^q(U)$ with $p < q < p^*$. Again by Hölder, we have

$$\|v\|_{L^q} \leq \|v\|_{L^p}^{\theta} \|v\|_{L^{p^*}}^{1-\theta},$$

where $\frac{d}{q} = \frac{d}{p}\theta + \frac{d}{p^*}(1-\theta)$. The condition that $p < q < p^*$ tells us that $0 < \theta < 1$. The L^p term goes ti 0 by compactness of $W^{1,p} \to C^p$, and the L^{p^*} term goes to 0 by the Sobolev inequality.

[3]Sung-Jin Oh says that this is the only version of Taylor's theorem you should ever use; this is a lesson he learned later than he would have liked.

Step 2: Prove compactness of $W^{1,p}(U) \to L^p(U)$: Given $\{u_n\} \subseteq W^{1,p}(U)$ with $\|u_n\|_{W^{1,p}(U)} \leq M < \infty$, by extension, we can find a sequence of extensions $\widetilde{u}_n$ of u_n defined on $\mathbb{R}^d$ such that

$$\|\widetilde{u}_n\|_{W^{1,p}(\mathbb{R}^d)} \leq C\|u_n\|_{W^1,p(U)} \leq CM$$

and $\operatorname{supp} \widetilde{u}_n \subseteq V$, where V is a bounded open set containing $\overline{U}$. It suffices to find a subsequence of $\widetilde{u}_n$ that converges in L^p. Introduce $\varphi \in C_c^\infty(\mathbb{R}^d)$ with $\int \varphi \, dx = 1$, and write

$$\widetilde{u}_n = \underbrace{\varphi * \widetilde{u}_n}_{v_{n\varepsilon}} + \underbrace{(\widetilde{u}_n - \varphi * \widetilde{u}_n)}_{e_{n,\varepsilon}}.$$

By the lemma,

$$\|e_{n\varepsilon}\|_{L^p} \leq C\varepsilon M,$$

independent of n. Also, note that using Hölder's inequality (specifically using that $\int |\widetilde{u}_n(x-y)\varphi_\varepsilon(x-y)| \, dy \leq \|\widetilde{u}_n\|_{L^p} \|\varphi_\varepsilon\|_{L^{p'}}$),

$$\|v_{n,\varepsilon}\|_{L^\infty} + \|\nabla v_{n,\varepsilon}\|_{L^\infty} \leq C_\varepsilon.$$

For each ℓ, there exists a subsequence $\widetilde{u}_{n_\ell}$ such that

$$\|e_{n_\ell,\varepsilon}\| < 2^{-\ell}$$

and such that

$$\|v_{n_{\ell'},\varepsilon} - v_{n_{\ell'},\varepsilon}\|_{L^p} < 2^{-\ell} \qquad \forall \ell', \ell'' > \ell.$$

(The second statement is by Arzelà-Ascoli. Now use a diagonal argument to extract a convergent subsubsequence; i.e. apply this recursively to subsequences and then extract a diagonal subsequence that converges. $\square$

7.3 Poicaré-type inequalities

A **Poincaré-type inequality** refers to any inequality that controls u in terms of information on Du, along with some additional condition to fix the ambiguity.

Theorem 7.4 (Poincaré inequality)**.** *Let $1 \leq p < \infty$, and let U be a bounded domain in $\mathbb{R}^d$ with C^1 boundary ∂U. For $u \in W^{1,p}(U)$ with $\int_U u \, dx = 0$,*

$$\|u\|_{L^p} \leq C_U \|Du\|_{L^p}.$$

Remark 7.1. For $p = 1$, the proof requires a bit more effort than what we will say.

Here is a proof from Evans' book. This is a typical application of Rellich-Kondrachov compactness.

Proof. We argue by contradiction. For contradiction, assume that for each $n \geq 1$, there exists $u_n \in W^{1,p}(U)$ such that $\int u_n = 0$ and

$$\|u_n\|_{L^p} \geq n\|\nabla u_n\|_{L^p}.$$

By normalization, we may assume that $\|u_n\|_{L^p} = 1$. Then it follows that

$$\|\nabla u_n\|_{L^p} \leq \frac{1}{n}.$$

In particular, this means that $\|u_n\|_{W^{1,p}(U)} \leq 2$, and by Rellich-Kondrachov compactness, there is a subsequence such that $u_n \to u_\infty$ in L^p. Moreover, $1 = \|u_n\|_{L^p} \to \|u_\infty\|_{L^p}$. Since $Du_n \to Du$ weakly in L^p, we must have $Du = 0$. That is, u is constant on U. But $0 = \int u_n \to \int u$, which tells us that $u = 0$ on U. However, this contradicts $\|u\|_{L^p} = 1$. $\square$

In most applications of this compactness arguments, u will satisfy linear relations that imply that it equals 0. Then you can show that it's not 0.

Remark 7.2. Another popular form of the Poincaré inequality is

$$\left\| u - \frac{1}{|U|} u \right\|_{L^p} \leq C_U \|Du\|_{L^p}.$$

Here are some other examples of Poincaré-type inequalities:

Theorem 7.5 (Friedrich inequality)**.** *Let $1 \leq p < \infty$, and let U be a bounded domain in $\mathbb{R}^d$ with C^1 boundary ∂U. For $u \in W^{1,p}(U)$ with $u|_{\partial U} = 0$,*

$$\|u\|_{L^p} \leq C_U \|Du\|_{L^p}.$$

We can prove this in the same way using compactness. On the other hand, we can also prove this just from the Sobolev inequality for $W_0^{1,p}(U)$.

Theorem 7.6 (Hardy's inequality)**.**

(i) If $u \in W^{1,p}(U)$ and $u|_{\partial U} = 0$, then

$$\left\| \frac{1}{\operatorname{dist}(\cdot, \partial U)} u \right\|_{L^p(U)} \leq C \|Du\|_{L^p(U)}.$$

(ii) If $u \in W^{1,p}(\mathbb{R}^d)$ with $p < d$, then

$$\left\| \frac{1}{|x|} u \right\|_{L^p} \leq C \|Du\|_{L^p}.$$

We can view Hardy's inequality as a refinement of Friedrich's inequality. We will discuss the proof next time.

8 Hardy's Inequality and Introduction to Elliptic PDEs

8.1 Hardy's inequality

Last time, we introduced Hardy's inequality.

Theorem 8.1 (Hardy's inequality)**.** *Let $u \in C_c^\infty(\mathbb{R}^d)$ with $d > 2$. Then*

$$\left\| \frac{1}{|x|} u \right\|_{L^2} \leq C \|Du\|_{L^2}.$$

This is a sharp inequality. We will see what the extremizer looks like.

Proof. Switch to polar coordinates (r, ω). It suffices to show that this inequality holds with the radial derivative: For each fixed ω,

$$\int \frac{1}{r^2} u^2 r^{d-1} \, dt \leq C \int |\partial_r u|^2 r^{d-1} \, dr,$$

and then we integrate over ω on both sides. The idea is to complete the square. We will subtract one side from the other and show it is ≥ 0. Without motivation, let's examine

$$(\partial_r u + \tfrac{\alpha}{r} u)^2 = (\partial_r u)^2 + \frac{2\alpha}{r} u \, \partial_r u + \frac{\alpha^2}{r^2} u^2.$$

The left hand side is ≥ 0. Now integrate both sides:

$$\begin{aligned}
0 &\leq \int (\partial_r u + \tfrac{\alpha}{r} u)^2 r^{d-1} \, dr \\
&= \int \left((\partial_r u)^2 + \frac{2\alpha}{r} \underbrace{u \partial_r u}_{\frac{1}{2}\partial_r u^2} + \frac{\alpha^2}{r^2} u^2 \right) r^{d-1} \, dr \\
&= \int (\partial_r u)^2 r^{d-1} \, dr + \alpha^2 \int \frac{1}{r^2} u^2 r^{d-1} \, dt + \alpha \int_0^\infty \partial_r u^2 r^{d-2} \, dr
\end{aligned}$$

We want to integrate by parts. Since $d > 0$, the boundary term will be 0. In particular, $\int_0^\infty \partial_r u^2 r^{d-2} \, dr = \cancel{u^2 r^{d-2}|_0^\infty} - (d-2) \int_0^\infty u^2 r^{d-3} \, dr$.

$$= \int (\partial_r u)^2 r^{d-1} \, dr - ((d-2)\alpha - \alpha^2) \int_0^\infty \frac{1}{r^2} u^2 r^{d-1} \, dr$$

Really, what we need here is $(d-2) > 0$ because we want the coefficient of α in the above quadratic term to be positive. We can upper bound this by plugging in $\alpha = \frac{d-2}{2}$. We can also upper bound $\int_0^\infty \frac{1}{r^2} u^2 r^{d-1} \, dr \leq (\frac{2}{d-2})^2 \int (\partial_r u)^2 r^{d-1} \, dr$. $\square$

Remark 8.1. Not only do we get the inequality, but we also get that

$$\left(\frac{d-2}{2}\right)^2 \int_0^\infty \frac{1}{r^2} u^2 \, dr = \int_0^\infty (\partial_r u)^2 r^{d-1} \, dr - \int_0^\infty \left(\partial_r u + \frac{d-2}{2r} u\right)^2 r^{d-1} \, dr.$$

This tells us that the extremizer is $r^{-(d-2)/2}$. However, this is not an element of H^1, so we can get near extremizers by including appropriate cutoffs.

8.2 Linear elliptic equations

Elliptic PDEs are a generalization of the Laplace equation $-\Delta u = f$.

Definition 8.1. The **symbol** of a partial differential operator is what we get when we replace ∂_j with $i\xi_j$.

It turns out that an important property is that $-\Delta \sum_j \partial_j \partial_j$ has (principal) symbol $-\sum_j (i\xi_j)^2 = |\xi|^2$. What's important is that $|\xi|^2$ is nonzero and thus invertible for $\xi \neq 0$:

$$|\xi|^2 \widehat{u} = \widehat{f} \implies \widehat{u} = \frac{1}{|\xi|^2} \widehat{f}.$$

This leads to the general definition of ellipticity of a partial differential operator.

Suppose that P is a linear partial differential operator such that if $u = (u^I)_{I=1}^N : U \to \mathbb{R}^N$, then (Pu) takes values in $\mathbb{R}^N$ and

$$(Pu)^I = \underbrace{\sum_{\substack{J,\alpha \\ |\alpha|=K}} A^I_{J,\alpha_1,\dots,\alpha_d} \partial^\alpha u^J}_{\text{principal part}} + (\text{lower order terms}).$$

Here, K is called the **order** of P.

Definition 8.2. The **principal symbol** of an operator is

$$\sigma_{\text{prin}}(P) = i^K \sum_{\substack{\alpha \\ \alpha=K}} A^I_{J,\alpha_1,\dots,\alpha_d}(x) \xi_1^{\alpha_1} \cdots \xi_d^{\alpha_d}.$$

Here, we allow the coefficients to be functions of x. We say that P is **elliptic** if $\sigma_{\text{prin}}(P)$ is invertible for all $x \in U$ and $\xi \neq 0$.

The case $N = 1$ is called the **scalar case**, where this looks like

$$Pu = \sum_{|\alpha|=K} a_\alpha(x) \partial^\alpha u.$$

Then the principal symbol is

$$\sigma_{\text{prin}}(P) = i^K \sum_\alpha a_\alpha(x)\xi^\alpha$$

The first nontrivial example is when $K = 2$, so

$$Pu = a^{i,j}\partial_i\partial_j + b^i\partial_j + c.$$

In this case, ellipticity is equivalent to $a^{i,j}\xi_i\xi_j \neq 0$ for all $x \in U$ and $\xi \neq 0$. This is equivalent to $a = [a^{i,j}]$ being a positive definite matrix for all $x \in U$.

We will assume that a is a symmetric matrix and require the following property.

Definition 8.3. **Uniform ellipticity**, is the property that there exists a uniform constant $\lambda > 0$ such that $a^{i,j}\xi_i\xi_j \geq \lambda$ for all $x \in U$ and $|\xi| = 1$.

This is equivalent to saying that all eigenvalues of the matrix $a(x)$ are bounded below by λ.

Why do we care about elliptic PDEs?

1. These arise naturally in optimization problems in math, physics, etc. In the latter part of the course, we will discuss these in the context of calculus of variations.
2. They also often arise as a part of evolutionary problems.

Example 8.1 (Incompressible Euler equations)**.** Let $u : \mathbb{R}_t \times \mathbb{R}^3 \to \mathbb{R}^3$ represent the velocity of a fluid element at each point in time and space. This follows the equation

$$\begin{cases} \partial_t + u \cdot \nabla u + \nabla p = 0 \\ \nabla \cdot u = 0 \end{cases}$$

This is one of the most infamous PDEs because of how difficult it is to understand.

How do we figure out p? Take the divergence of the first equation to get that

$$-\Delta p = \nabla(u \cdot \nabla u).$$

This is the **pressure equation**.

We will cover:

- Boundary value problems for elliptic PDEs, existence, and uniqueness.
- Regularity properties of solutions to elliptic PDEs. If $Pu = f$, where P is elliptic of order K, then we will have **elliptic regularity**[4]: If f has regularity of order k (so $f \in H^k$), then u has regulariity $k + K$.

[4] Elliptic regularity holds even for systems.

- Maximum principles (mostly for the scalar case, $N = 1$).

If we have time, we will discuss topics such as

- Unique continuation.
- Spectral theory.

We will mostly follow Evans' textbook, but we will deviate sometimes on a few topics.

8.3 Boundary value problems and a priori estimates for elliptic PDEs

Assume $d \geq 2$ and $N = 1$ (scalar case). Also assume uniform ellipticity of P and some "nice" regularity for the coefficients a, b, c. We will focus mostly on the case where U is a bounded domain in $\mathbb{R}^d$ with "nice" boundary.

When it comes to boundary value problems, you cannot prescribe both function values and values of the normal derivative at the boundary; this stems from the various uniqueness properties that arise for these PDEs. We will mostly focus on **Dirichlet boundary problems**,

$$\begin{cases} Pu = f & \text{in } U \\ u = g & \text{on } \partial U. \end{cases}$$

We will focus less on boundary problems such as **Neumann boundary problems**,

$$\begin{cases} Pu = f & \text{in } U \\ \frac{\partial}{\partial \nu} u = g & \text{on } \partial U. \end{cases}$$

We will study solvability for $u \in H^1(U)$. We will first study the Dirichlet boundary value problem ($u|_{\partial u} = g$ is okay due to the trace theorem). We will later discuss the Neumann boundary value problem, which needs to be studied in H^2 because we need to use the trace theorem on the derivative.

The standard reduction is that it suffices to understand $g = 0$. This is because if we take any extension (with correct regularity) $\widetilde{g} : \overline{U} \to \mathbb{R}$ of g, then we can work with $v = u0\widetilde{g}$ and solve the problem

$$\begin{cases} Pv = f + P\widetilde{g} = \widetilde{f} & \text{in } U \\ v = 0 & \text{on } \partial U. \end{cases}$$

Definition 8.4. P is in **divergence form** if

$$Pu = \partial_i(a^{i,j}\partial_j u) + \partial_i(b^i u) + c.$$

Note that if a is smooth, then

$$a = a^{i,j}\partial_i\partial_j + (\partial_j a^{i,j} + b^i)\partial_i u + (\partial_i b^i + c)u.$$

Our discussion of existence and uniqueness of the Dirichlet boundary value problem would be based on a-priori estimates.

Theorem 8.2 (a-priori estimate)**.** *Suppose that $u \in H^1$ solves the Dirichlet boundary problem, and assume that $b, c \in L^\infty$ with $\|b\|_{L^\infty} + \|c\|_{L^\infty} \le A$. Then there exist constants $C > 0$ and $\gamma \ge 0$ such that*

$$\|u\|_{H^1(U)} \le C\|f\|_{H^{-1}} + \gamma\|u\|_{L^2(U)}.$$

Proof. The proof is essentially integration by parts. We can use approximation to justify the integration by parts. Write

$$\begin{aligned}\int_U Pu\,dx &= \int_U (\partial_j(a^{i,j}\partial_i u + b^j u) + cu)u\,dx \\ &= \int -a^{i,j}\partial_i u\partial_j u - b^j u\partial_j u + cuu\,dx\end{aligned}$$

Uniform ellipticity tells us that $\lambda|Du|^2 \le a^{i,j}\partial_i u\partial_j u$; integrate this to take care of the first term. The second term can be dealt with using Cauchy-Schwarz, and the third term is $\gamma\|u\|_{L^2}^2$.

Putting this all together gives

$$\begin{aligned}\lambda\|Du\|_{L^2(U)}^2 &\le C\|f\|_{H^{-1}}\|u\|_{H^1} + \int_U |b||\partial u||u|\,dx + \int_U |c||u|^2\,dx \\ &\le C\|f\|_{H^{-1}}\|u\|_{H^1} + A\underbrace{\int_U |\partial u||u|\,dx}_{\le\|\partial u\|\|u\|_{L^2}} + A\underbrace{\int_U |u|^2\,dx}_{\le\gamma\|u\|_{L^2}^2}.\end{aligned}$$

If we make γ large enough so that we have put an $\|u\|_{L^2}^2$ on the right hand side and abosrb the second term, we get

$$\|u\|_{H^1(U)}^2 \le C\|f\|_{H^{-1}}\|u\|_{H^1} + \gamma\|u\|_{L^2}\|u\|_{H^1}.$$ □

Remark 8.2. We can alter this argument to only require $b \in L^{d+}$ and $c \in L^{d/2+}$.

9 Solvability for Elliptic Operators

9.1 The Dirichlet problem and energy estimates for elliptic operators

We are looking at a second order, scalar, elliptic operator P (we will sometimes use L, which Evans' textbook uses):

$$Pu = -\partial_j a^{j,k}\partial_k u) + b^j\partial_j u + cu.$$

For ellipticity, we will assume that $a = [a^{j,k}]$ is a positive definite matrix, and we will further assume that $a \succeq \lambda I$ for some $\lambda > 0$ (i.e. all eigenvalues of a are $\geq \lambda$). For the purposes of this lecture, we will assume that $a, b, c \in L^\infty(U)$, where U is a bounded domain with C^1 boundary.

Last time, we looked at the Dirichlet boundary value problem

$$\begin{cases} Pu = f & \text{in } U \\ u = g & \text{on } \partial U. \end{cases}$$

Recall that we may assume $g = 0$ be working with u minus some extension of g.

By the regularity assumptions on the coefficients a, b, c, $P : H^1(U) \to H^{-1}(U)$. Recall that $H^{-1}(U) = \{f = f_0 + \sum_{i=1}^d \partial_{x^i} f_i : f_0 f_i \in L^2\}$ and that $W_0^{k,p}(U)^* = W^{-k,p'}(U)$. The norm for this space is

$$\|f\|_{H^{-1}} = \inf_{f = f_0 + \sum_{i=1}^d \partial_{x^i} f_i} \left\{ \left(\|f_0\|_{L^2}^2 + \sum_{i=1}^d \|f_i\|_{L^2}^2 \right)^{1/2} \right\}.$$

To build in the Dirichlet boundary condition $u|_{\partial U} = 0$, restrict P to $P : H_0^1(U) \to H^1(U)$ (here, H_0^1 is the set of H^1 functions with 0 trace). To understand the solvability of P (i.e. existence and uniqueness), we want to understand if P is 1 to 1 and onto. We will use a priori estimates.

Last time, we proved the following a priori estimate.

Lemma 9.1 (Energy estimate)**.** *There exist $C > 0, \gamma > 0$ such that for $u \in H_0^1(U)$,*

$$\|u\|_{H^1(U)} \leq C\|Pu\|_{H^{-1}(U)} + \gamma\|u\|_{L^2(U)}.$$

The proof was by integration by parts.

Recall that in order to prove existence statements with a priori estimates, we also needed to think about the dual problem for the adjoint P^*. (In finite dimensional linear algebra, $Ax = y$ has a solution x if and only if $r \in \operatorname{ran} A = {}^\perp(\ker A^*)$. For P as above, let's compute P^* with respect to $\langle u, v\rangle = \int uv\, dx$:

$$\int \partial_j uv\, dx = -\int u\partial_j v\, dx,$$

so

$$P^* = -\partial_j(a^{j,k}\partial_k u) - \partial_j(b^j u) + cu,$$

where we are assuming everything is real-valued. Note that the energy estimate also applies to P^*.

9.2 Case 1: Both P and P^* obey good a priori estimates

In our discussion of Sobolev spaces, we introduced the following lemma from functional analysis.

Lemma 9.2. *Let X, Y be Banach spaces, and let $P : X \to Y$ be a bounded, linear operator. If $\|u\|_X \leq C\|Pu\|_Y$, then*

(i) $\ker P = \{0\}$

(ii) For every $g \in X^*$, there exists a $v \in Y^*$ such that $P^*v = g$ ($\operatorname{ran} P^* = X^*$) and $\|v\|_{X^*} \leq C\|g\|_{X^*}$.

If $\|v\|_{Y^*} \leq C'\|P^*v\|_{X^*}$, then

(i) $\ker P^* = \{0\}$

(ii) For every $f \in Y$, there is a $u \in X$ such that $Pu = f$ ($\operatorname{ran} P = Y$) and $\|u\|_X \leq C'\|f\|_Y$.

Remark 9.1. In our previous proof, we assumed that X is reflexive to reduce (ii) to (i), but this assumption can be dropped. To see this argument, look for the "closed range theorem." The key idea is that $\overline{\operatorname{ran} P} = {}^\perp(\ker P^*)$.

We want to apply this lemma to our P, $X = H_0^1$, and $Y = H^{-1}(U)$. In this setting, $X^* = H^{-1}(U) = Y$, and $Y^* = H_0^1(U) = X$.

In the energy estimate, we have an extra term $\gamma\|u\|_{L^2(U)}$ in the bound. For now, we will get rid of it by cheating. We will deal with it in full later. Here is when we have the energy estimate with $\gamma = 0$:

Lemma 9.3. *If $b = 0$ and $c = 0$, i.e. $Pu = -\partial_j(a^{j,k}\partial_j u)$, then the energy estimate holds with $\gamma = 0$.*

Proof. By density, $u \in C_0^\infty$.

$$\begin{aligned}\int_U Puu\,dx &= \int_U -\partial_j(a^{j,k}\partial_k u)u\,dx \\ &= \int_U a^{j,k}\partial_j u\partial_k u\,dx\end{aligned}$$

$$\geq \lambda \int)U|Du|^2\,dx$$

Using Friedrich's inequality,

$$\geq C\int_U |u|^2\,dx.$$

As in the proof of the energy estimate, we cancel a factor of $\|u\|_{H^1}$ on both sides of the inequality to get the result. □

Remark 9.2. Since P^* has the same form with the same constants, this condition gives the energy estimate with $\gamma = 0$ for P^*, as well.

Theorem 9.1. *For every $f \in H^{-1}(U)$, there exixsts a unique $u \in H_0^1(U)$ such that $-\partial_j(a^{j,k}\partial_j u) = f$ in U.*

Remark 9.3. For the proof of this, Evans' textbook uses the Lax-Milgram lemma, but our lemma is actually stronger.

9.3 Case 2: General P

To obtain stronger results for our general problem, we will develop tools which are specifically useful for this problem. In particular, we will discuss Fredholm theory.

Recall the notion of a compact operator $K : X \to Y$ from functional analysis: $K(\overline{B}_X)$ is compact, where $B_X = \{x \in X : \|x\| < 1\}$.

Lemma 9.4.

(o) *For $K : X \to Y$, K is compact if and only if K^* is compact.*

(i) *(Solvability of $(I + K)x = y$): Let $K : X \to X$ be compact, and let $T = I + K$.*

- *(a)* $\ker(I + K)$ *is finite dimensional.*
- *(b)* *There exists an $n_0 \geq 1$ such that $\ker(I + K)^n = \ker(I + K)^{n_0}$ for $n \geq n_0$.*
- *(c)* $\operatorname{ran}(I + K)$ *is closed, so* $\operatorname{ran}(I + K) = {}^{\perp}(\ker(! + K^*))$.
- *(d)* $\dim\ker(I + K) = \dim\ker(I + K^*)$.

Remark 9.4. Part (d) is the general equivalent of the fact that in finite dimensional linear algebra, the row rank of a matrix is equal to the column rank of a matrix. This statement is that $\operatorname{index}(I + K) = 0$, where the index of an operator is the difference of these two quantities. The index tends to be very stable under perturbation.

Proof. For the proof when X is a Hilbert space, see the appendix of Evans' textbook. What is the idea? Here is how to think about compact operators: Notice that if A has $\dim \operatorname{ran} A < \infty$, then A is compact. Also notice that if $K_n \to K$ in the operator norm topology on $\mathcal{L}(X, Y)$, then K is compact. Combining these two facts tells us that the closure of the set of finite rank operators is a subset of the compact operators; in separable Hilbert spaces, this is what all compact operators look like. $\square$

Why is this lemma relevant for us? Take any general

$$Pu = -\partial_j(a^{j,k}\partial_k u) + b^j partial_j u + cu.$$

In general, the energy estimate gives

$$\|u\|_{H_0^1(U)} \leq C\|Pu\|_{H^{-1}(U)} + \gamma\|u\|_{L^2(U)}.$$

But if we consider instead $(P + \mu I)u = -\partial_j(a^{j,k}\partial_k u) + b^j$
$partial_j u + (c + \mu)u$ with $\mu \gg 1$, then we can remove γ on the right hand side.

Indeed,

$$\int (P + \mu)u \, dx = \underbrace{\int -\partial_j a^{i,k}\partial_k u \, dx}_{\geq \lambda \int |Du|^2 \, dx} + b, c \text{ terms} + \int \mu u^2 \, dx,$$

where the $\int \mu u^2 \, dx$ term is favorable if $\mu > 0$. By case 1, for μ sufficiently positive, for all $f \in H^{-1}$, ther exists a unique $u \in H_0^1$ such that

$$(P + \mu I)u = f.$$

We then have a well-defined map $(P + \mu I)^{-1} : H^{-1}(U) \to H_0^1(U)$. Now go back to

$$(P + \mu)u - \mu u = Pu = f.$$

Apply $(P + \mu)^{-1}$ to get

$$u - \mu(P + \mu)^{-1}u = (P + \mu)^{-1}f.$$

By Rellich-Kondrachov (recalling that U is bounded), the embedding $\iota : H_0^1(U) \to L^2$ is compact. From this, it follows that

$$(P + \mu)^{-1} : L^2(U) \to H^{-1}(U) \xrightarrow{(P+\mu)^{-1}} H^1(U) \to L^2(U)$$

is compact (since $A \circ K$ or $K \circ A$ is compact whenever A is bounded and linear and K is compact). Thus, $-\mu(P + \mu)^{-1} : L^2(U) \to L^2(U)$ is compact. Thus, our repackaging of the problem,

$$u - \mu(P + \mu)^{-1}u = (P + \mu)^{-1}f,$$

is of the form $(I + K)x = y$.

Theorem 9.2 (Fredholm alternative)**.** *Let P be as before, and let U be a bounded domain with C^1 boundary.*

(i) Exactly one of the following holds:

(a) (Solvability) For all $f \in H^{-1}(U)$, there exists a unique $u \in H_0^1(U)$ such that $Pu = f$, and there exists a $C > 0$ independent of u, f such that $\|u\|_{H^1(U)} \leq C\|f\|_{H^{-1}(U)}$.

(b) (Existence of nonzero homogeneous solution) There exists a nonzero $u \in H_0^1(U)$ (or equivalently in $L^2(U)$) such that $Pu = 0$.

(ii) If (b) holds, then $\dim \ker P < \infty$ and $\dim \ker P^* < \infty$. Given $f \in H^{-1}(U)$, there exists a $u \in H_0^1(U)$ such that $Pu = f$ if and only if $\langle f, v\rangle = 0$ for all $v \in \ker P^*$.

Remark 9.5. While our initial approach didn't really care about boundedness, this approach essentially relies on this condition.

Remark 9.6. Part (ii) is a statement about norms. This will be an exercise and follows from compactness.

Remark 9.7. Here is a very nice consequence of this theorem. Take

$$\widetilde{P}u = -\partial_j(a^{j,k}\partial_k u) + b^j\partial_j u.$$

There is a weak maximum principle which says that

$$\sup_{\overline{U}} |u| = \sup_{\partial U} |u|.$$

This gives uniqueness in this Dirichlet problem. Then the Fredholm alternative gives us solvability from the uniqueness. We will properly discuss this later, when we go over maximum principles.

10 L^2-Based Elliptic Regularity

10.1 Regularity theory for the Poisson equation

Last time, we discussed solvability for elliptic PDEs. Now we will talk about the regularity of solutions to elliptic PDEs. Here is a prototypical example.

Example 10.1. Consider the Poisson equation $-\Delta u = f$ in U, where $f \in H^k(U)$ or $C^{k,\alpha} = \{u \in C^k(U) : \partial^\alpha u \in C^{0,\alpha}(U)\ \forall |\alpha| = k\}$. The idea is that u should be more regular than f by order 2. Interior regularity says that for all $V \subseteq\subseteq U$ (notation meaning V is bounded and $\overline{V} \subseteq U$),

$$\|u\|_{H^{k+2}(V)} \leq C\|f\|_{H^k(V)} + C\|u\|_{L^2(U)}.$$

Similarly,

$$\|u\|_{C^{k+2,\alpha}(V)} \leq C\|f\|_{C^{k,\alpha}(V)} + C\|u\|_{L^\infty(U)}.$$

In general, the constant C can depend on the domain V.

The first of these statements is referred to as L^2**-based regularity theory**, and the second is referred to as **Schauder theory**. We will think about L^2-based regularity theory for now and discuss Schauder theory later.

For L^2-based regularity theory, the key idea is integration by parts (the energy method).[5] We will make a simplifying that $u \in H^{k+2}(V)$; this is not assuming everything because from this qualitative fact, we will derive a quantitative bound. This assumption allows us to commute the equation with derivatives. We have not said any assumptions about the boundary, which may seem like an issue with integration by parts, but this is why we are discussing *interior* regularity. We will solve this with a cutoff function.

Let ζ be a nonnegative, smooth cutoff function which equals 1 in V and equals 0 near ∂U. Then (squaring ζ in anticipation of a nice L^2 trick),

$$\begin{aligned}\int_U fu\zeta^2\,dx &= \int_U -\Delta u u\zeta^2 \\ &= \sum_{j=1}^d \int_U \underbrace{\partial_j u \partial_j(u\zeta^2)}_{\partial_j u\zeta^2 + 2u\zeta\partial_j\zeta}\,dx\end{aligned}$$

Note that we have no boundary term in the integration by parts thanks to ζ.

$$= \sum_{j=1}^d \int (\partial_j u)^2\zeta^2 + 2\partial_j u u\zeta\partial_j\zeta\,dx.$$

[5] Fraydoun Rezakhanlou says that he is an analyst, a PDE-ist, and a probabilist. He is an analyst because he uses the Cauchy-Schwarz inequality, a probabilist because he uses Chebyshev's inequality, and a PDE-ist because he uses integration by parts.

Rearrange this to get

$$\int_U |Du|^2\zeta^2\,dx \le \left|\int_U fu\zeta^2\,dx\right| + \underbrace{2\left|\int_U u\zeta Du\cdot D\zeta\,dx\right|}_{\le 2(\int_U |Du|^2\zeta^2)^{1/2}(\int_U u^2|D\zeta|^2\,dx)^{1/2}}$$

To control this right term, we use the AM-GM inequality $ab \le \frac{a}{2} + \frac{b}{2}$. But we can weight this by $\sqrt{\varepsilon}$ on a and $\frac{1}{\sqrt{\varepsilon}}$ on b to get the inequality $ab \le \varepsilon\frac{a^2}{2} + \frac{1}{\varepsilon}\frac{b^2}{2}$. This bounds

$$2\left(\int_U |Du|^2\zeta^2\right)^{1/2}\left(\int_U u^2|D\zeta|^2\,dx\right)^{1/2} \le \varepsilon\int_U |Du|^2\zeta^2\,dx + \frac{1}{\varepsilon}\int_U u^2|D\zeta|^2\,dx.$$

Now set $\varepsilon = 1/2$ to absorb the first term to the right hand side.

This gives

$$\begin{aligned}\frac{1}{2}\int_U |Du|^2\zeta^2\,dx &\le \left|\int_U fu\zeta^2\right| + 2\int_U u^2|D\zeta|^2\,dx\\ &\le \|f\|_{L^2(U)} + \|u\|_{L^2(U)},\end{aligned}$$

and we lower bound the left hand side by $\frac{1}{2}\int_V |Du|^2\,dx$. For the actual result, we could have upgraded the $\|f\|_{L^2(U)}$ to $\|f\|_{H^1(U)}$ by using an additional cutoff argument.

What about higher regularity? Suppose $k+2=2$. Then if $-\Delta u = f$, we get

$$-\Delta\partial_j u = \partial_j f,$$

where $\partial_j u \in H^1$, so we can do integration by parts. Now apply the case $k=1$ to get

$$\int_V |D\partial_j u|^2\,dx \le \left|\int_U \partial_j f\partial_j u\zeta^2\,dx\right| + \|\partial_j u\|_{L^2(U)}$$

Bound the first term by (using the same AM-GM trick)

$$\left|\int_U f\partial_j^2 u\zeta^2\,dx\right| \le \frac{1}{4\varepsilon}\int_U f^2\zeta^2\,dx + \varepsilon\int_U |\partial_j u|^2\zeta^2\,dx.$$

Absorb the second term to the right hand side to get

$$\int_U |D\partial_j u|^2\zeta^2\,dx \le C\int_U f^2\,dx + C\|Du\|^2_{L^2(U)}.$$

We want to change the last term into $\|u\|_{L^2(U)}$. Our tool to do this is the H^1 bound we just proved. But this needs us to have a domain in the interior of U. However, note that if we define $V \subseteq\subseteq W \subseteq\subseteq U$, we can replace this term on the the right hand side by $C\|Du\|_{L^2(W)}$. Then we use the H^1 bound $\|Du\|_{L^2(W)} \le \|f\|_{L^2(U)} + \|u\|_{L^2(U)}$. In conclusion, we get

$$\|D\partial_j u\|_{L^2(V)} \le C\|f\|_{L^2(U)} + C\|u\|_{L^2(U)}$$

for all j. Combined with the H^1 bound, this gives the H^2 bound

$$\|u\|_{H^2(V)} \le C\|f\|_{L^2(U)} + C\|u\|_{L^2(U)}.$$

10.2 L^2-regularity for elliptic operators

For the full L^2-regularity theorem, we have an elliptic operator

$$Pu = -\partial_j(a^{j,k}\partial_k u) + b^j\partial_j u + cu,$$

where $u : U \to \mathbb{R}$ and U is an open subset of $\mathbb{R}^d$. We also assume $a(x) \succ \lambda I$ for some $\lambda > 0$ for all $x \in U$. Also assume $a, b, c \in L^\infty(U)$ (although the natural assumption for $d \geq 3$ is actually $a \in L^\infty, b \in L^d, c \in L^{d/2}$). For the H^2 bound, we also make the assumption that $\partial a \in L^\infty(U)$; this comes from the fact that if we want to commute the derivative as in the argument above, we must be able to deal with the derivative of the coefficients $a^{i,j}$.

Theorem 10.1 (H^2 elliptic regularity)**.** *Let $u \in H^1(U)$ be a weak solution to $Pu = f$ on U, and let $f \in L^2(U)$. Then for all $V \subseteq\subseteq U$, $u \in H^2(V)$, and*

$$\|u\|_{H^2(V)} \leq C(\|f\|_{L^2(U)} + \|u\|_{L^2(U)}).$$

The proof of this theorem is the same as the previous argument but with some minor adjustments. The main step is integration by parts. Formally,

$$\begin{aligned}
\int_U -\partial_j(a^{j,k}\partial_k v)v\zeta^2\,dx &= \int_U a^{j,k}\partial_k v\partial_j v\zeta^2\,dx + \int_U a^{j,k}\partial_k v v\zeta\partial_j\zeta\,dx \\
&\geq \lambda\int_U |Du|^2\zeta^2\,dx - \|a\|_{L^\infty}\cdot \underbrace{\int |Du|\zeta|v||D\zeta|\,dx}_{\leq \frac{\lambda}{2}\frac{1}{\|a\|_{L^\infty}}|Dv|^2\zeta^2 + \frac{1}{\lambda}\|a\|_{L^\infty}|v|^2|D\zeta|^2} \\
&\geq \frac{\lambda}{2}\int_U |Dv|^2\zeta^2\,dx - \frac{\|a\|_{L^\infty}^2}{\lambda}\int_U |v|^2|D\zeta|^2\,dx.
\end{aligned}$$

Since we do not know a priori that $u \in H^2(V)$, need to modify the proof idea to commute the equation with difference quotients instead of derivatives.

Definition 10.1. If $k \in \{1, \ldots, d\}$ and $h \in \mathbb{R} \setminus \{0\}$, the **difference quotient** is

$$D_k^h v(x) = \frac{v(x + hek) - v(x)}{h}.$$

This converges to $\partial_k v(x)$ as $h \to 0$.

Proof. Step 0: Note that for $u \in H^1(U)$,

$$Pu = f \text{ in } U \iff \langle Pu, \varphi\rangle = \langle f, \varphi\rangle \qquad \forall\varphi \in C_c^\infty(U)$$

Here, $Pu \in H^{-1}(U), f \in L^2 \subseteq H^{-1}$.

$$\iff \langle Pu\varphi\rangle = \langle f, \varphi\rangle \qquad \forall\varphi \in H_0^1(U) \quad (= (H^{-1}(U))^*)$$

When we did our a priori estimate last time, we used approximation of u by smooth functions. However, here, we want to show that we have extra regularity, so the equivalent of approximation is this step above.

$$\iff \int_U a^{j,k}\partial_j u \partial_k \varphi + b^j \partial_j u\varphi + cu\varphi \, dx = \int_U f\varphi \, dc \qquad \forall \varphi \in H_0^1(U).$$

Step 1: Now commute the equation with D_j^h. Note that the Leibniz rule holds:

$$D_h^h(uv)(x) = D_j^h u(x)v(x) + u(x+h)D_j^h v(x).$$

This comes from

$$uv(x+h) - uv(x) = (u(x+h) - u(x))v(x) + \underbrace{u(x+h)}_{=:u^h(x)}(v(x+h) - v(x)).$$

Now

$$\begin{aligned} D_j^h f &= D_j^h(-\partial_j a^{j,k}\partial_k u + b^j\partial_j u + cu) \\ &= -\partial_\ell (a^h)^{j,k}\partial_k D_j^h u + (b^h)^j \partial_\ell D_j^h u + c^h D_j^h u - \partial_\ell (D^h a)^{\ell,k}\partial_k u + (D_j^h b)^\ell \partial_\ell u + D_j^h cu. \end{aligned}$$

Rearrange this as

$$-\partial_\ell((a^h)^{\ell,k}\partial_k D_j^h u) = \widetilde{f}_1^h,$$

where f_1^h is everything else. Now

$$\langle -\partial_\ell (a^h)^{\ell,k}\partial_k D_j^h u, \varphi\rangle = \langle \widetilde{f}_1^h, \varphi\rangle \qquad \forall \varphi \in H_0^1(U),$$

where the left hand side equals

$$\int (a^h)^{\ell,k}\partial_k (D_j^h) u \partial_\ell \varphi \, dx.$$

Step 2: "$\varphi = \partial_j u\zeta^2$": Choose $\varphi = D_j^h \zeta^2 \in H_0^1(U)$. By the integration by parts idea, we get

$$\frac{\lambda}{2}\int_U |DD_j^h u|^2 \zeta^2 \; dx \leq \cdots \widetilde{f}_1 D_j^h u.$$

One treats the right hand side like before, treating $D_j^h u$ like $\partial_j u$. To make this precise, we need the following lemma:

Lemma 10.1 (from Ch 5 in Evans)**.** *Let $V \subseteq\subseteq U$.*

1. *If $u \in W^{1,p}$, $\|D_j^h u\|_{L^p(V)} \leq C\|\partial_j u\|_{L^p(U)}$ for $|h| \ll 1$.*
2. *Assume $u \in L^p$. For $h \ll 1$, if $\|D_j^h u\_L^p(V) \leq A$, then $\partial_j \in L^p$, and $\|\partial_j u\|_{L^p(V)} \leq A$.*

This finishes off the proof. $\square$

11 L^2-Based Interior and Boundary Regularity

11.1 H^k elliptic interior regularity

Last time, we were studying L^2-based regularity theory. We were considered with the second order, scalar partial differential operator

$$Pu = -\partial_j(a^{j,k}\partial_k u) + b^j\partial_j u + cu,$$

where $a^{j,k}(x) \succ \lambda I$ for all $x \in U$ for some $\lambda > 0$.

Theorem 11.1 (H^2 interior regularity)**.** *Let U be an open subset of $\mathbb{R}^d$, and suppose $|Da| + |a| + |b| + |c| \leq \Lambda$ for all $x \in U$. Let $u \in H^1(U)$ be a weak solution to $Pu = f$ in U, where $f \in L^2(U)$. Then for all $V \subseteq\subseteq U$ (V bounded with $\overline{V} \subseteq U$), $u \in H^2(V)$, and*

$$\|u\|_{H^2(V)} \leq C(\|f\|_{L^2(U)} + \|u\|_{L^2(U)}).$$

Remark 11.1. The constant C is independent of u and f but dependent on λ, Λ, V, U.

The basic ideas in the proof were:

1. Integration by parts and ellipticity give us control over the highest order term.

2. Commute the equation with ∂_j.

In the proof, we looked at the equation for $\partial_j u$, then applied ellipticity to control $\|\zeta D\partial_j u\|_{L^2}$, where ζ was a smooth curoff which equals 1 on B but is 0 near ∂U. In reality, however, to deduce that $u \in H^2(V)$, we have to work with the difference quotient $D_j(u) = \frac{u(x+he_j)-u(x)}{h}$.

Here is the higher regularity version of this theorem.

Theorem 11.2 (H^k elliptic interior regularity)**.** *Assume the same hypotheses as before, except*

- $|D^\alpha a| \leq A$ *for all* $|\alpha| \leq k-1$, $|D^\alpha b| + |D^\alpha c| \leq A$ *for all* $|\alpha| \leq k-2$,
- $f \in H^{k-2}(U)$.

Then for all $V \subseteq\subseteq U$, $u \in H^k(V)$, and

$$\|u\|_{H^k(V)} \leq C(\|f\|_{H^{k-2}(U)} + \|u\|_{L^2(U)}).$$

Proof. Here is a sketch. The proof follows the same idea, except we commute D^β for $|\beta| \leq k-1$. Then look at the equation for $D^\beta u$:

$$D^\beta f = D^\beta(Pu)$$

$$
\begin{aligned}
&= D^\beta(-\partial_j(a^{j,k}\partial_k u) + b^j\partial_j u + cu) \\
&= -\partial_j(a^{j,k}\partial_k D^\beta u) + D^\beta(b^j\partial_j u) + D^\beta(cu).
\end{aligned}
$$

Multiply both sides by $\zeta^2 D^\beta u$. The first term on the right is

$$
-\sum_{\gamma\le\beta,\gamma\ne\beta} \partial_j(D^{\beta-\gamma}a^{j,k}\partial_k D^\gamma u)c_\gamma.
$$

This gives us control of $\|DD^\beta u\zeta\|_{L^2(U)}$. For the rest of the terms, you do not see more than $k-1$ derivatives of of u and $k-2$ derivatives of b and c after integration by parts.

In reality, the details need to be carried out with difference quotients, using induction to take care of lower derivative terms. The full proof is in Evans' book. □

11.2 L^2-based boundary regularity

Previously, we have been looking at regularity away from the boundary. You may also notice that we have not been putting conditions on boundary behavior of u (we only required, for example, $u \in H^1$ rather than $u \in H^1_0$).

Theorem 11.3. *Assume the same hypotheses as in the H^2 interior regularity theorem, except:*

- *$u \in H^1_0(U)$ (i.e. $u|_{\partial U} = 0$ in the sense of traces).*
- *∂U is C^2.*

Then $u \in H^2(U)$, and

$$
\|u\|_{H^2(U)} \le C(\|f\|_{L^2(U)} + \|u\|_{L^2(U)}).
$$

Proof. Assume for simplicity that $u \in H^2(U)$; we can take care of this by doing the argument with difference quotients instead of derivatives. We will omit the contribution of b and c because they do not contribute much, as we have seen. Start with the equation

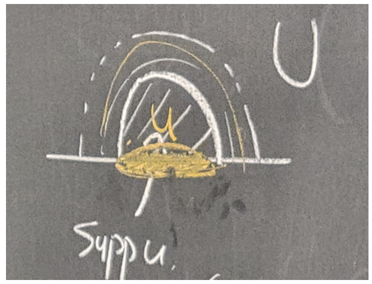
$$
f = \partial_j(a^{j,k}\partial_k u) + \cdots .
$$

We want to take a derivative to say

$$
\partial_\ell f = -\partial_\ell(\partial_j(a^{j,k}\partial_k u)),
$$

but we cannot necessarily take the derivative at the boundary. However, notice that if the boundary is flat (wlog $\{x^d = 0\}$), then all ∂_ℓ exist for $\ell = 1, \dots, d-1$. The only problem is the normal derivative $\partial_{x^d} = -\nu$. In other words only $(d-1)$-many directions (tangential to ∂U) are admissible.

For the sake of simplicity, take the special case when $U = B_1(0) \cap \mathbb{R}^d_+$ and $\operatorname{supp} u \subseteq B_{1/2}(0) \cap \mathbb{R}^d_+$.



In this case, $\ell f = -\partial_j(a^{j,k}\partial_k\partial_\ell u) - \partial_j(\partial_\ell a^{j,k}\partial_k u)$ for $\ell = 1, \ldots, d-1$. For these $d-1$ terms, we can use the cutoff ζ which equals 1 on $B_{1/2}(0)$ and is 0 near $\partial B_1(0)$ to get

$$\|\zeta D\partial_\ell u\|_{L^2} \leq C(\|\zeta f\|_{L^2} + \|u\|_{L^2}).$$

In the integration by parts, there is an additional boundary term from $B_{1/2}(0) \cap \{x^d = 0\}$. However, this contribution is zero because $u|_{\partial U} = 0$, which also implies $\partial_\ell u|_{\partial U} = 0$ for $\ell = 1, \ldots, d-1$.

In this special case, it now remains to control $\|\zeta\partial_{x^d}\partial_{x^d}u\|_{L^2}$. The key observation is that the equation allows us to express $D_\nu D_\nu u$ in terms of everything else. Recall that the original equation is

$$f = -\partial_j(a^{j,k}\partial_k u) + \cdots .$$

The condition that $a \succ \lambda I$ is equivalent to $a^{j,k}\xi_j\xi_k \geq \lambda|\xi|^2$ for all $\xi \in \mathbb{R}^d$. If we take $\xi = e_d$, this tells us that $a^{d,d} \geq \lambda$. Now write the equation as

$$f = \underbrace{-\partial_d(a^{d,d}\partial_d u)}_{=a^{d,d}\partial_d^2 u - (\partial_d a^{d,d})\partial_d u} - \sum_{\substack{j,k \\ j,k \neq d}} \partial_j(a^{j,k}\partial_k u).$$

We can divide the equation by $a^{d,d}$ to get

$$\partial_d^2 u = \frac{1}{-a_{d,d}}(a \cdot D_{\tan} u + (\partial a), b)Du + cu + f.$$

This lets us control $\partial_d^2 u$ by the other derivatives, completing the proof in this special case.

In general, we reduce to this special case by first using a smooth partition of unity and boundary straightening. In particular, for every $x \in \partial U$, there exists a ball $B_r(x)$ such that, after relabeling of the coordinate axes, $U \cap B_r(x) = \{x^d > \gamma(x^1, \ldots, x^{d-1})\}$ for some C^2 function γ. We then take a boundary straightening map y, defined by

$$\begin{cases} y^\ell = x^\ell & \ell = 1, \ldots, d-1 \\ y^d = x^d - \gamma(x^1, \ldots, x^d). \end{cases}$$

By compactness, $U \subseteq (\bigcup_{k=1}^K U_k) \cup U_0$, where U_k are balls covering the boundary and U_0 contains the rest of the interior. Then there exists a smooth partition of unity $\{\chi_k\}_{k=0}^K$ subordinate to this cover, which gives

$$u = \chi_0 u + \sum_{k=1}^K \chi_k u.$$

The first term is supported on the interior of U, so we can apply our interior regularity theorem to it. For each other $\chi_k u$, when we change $x \mapsto y = y(x)$, we are reduced to the half-ball case already covered (both in terms of geometry and support of u). Check that the ellipticity constant of the resulting equation is still $\simeq \lambda$ and that $\partial \widetilde{a}(y)$, $\widetilde{b}(y)$ obey same bounds as before; this comes from writing the equation in terms of derivatives in y and checking that the change of variables formula $a^{j,k} = \frac{\partial x^j}{\partial y^{j'}} \widetilde{a}^{j',k'} \frac{\partial x^k}{\partial y^{k'}}$ preserves the $a \succ \lambda I$ condition. From the H^2 bound for $u\chi_k(y)$, come back to $u\chi_k(x)$ (which needs the C^2 condition on ∂U). $\square$

11.3 High level comparison of L^2-based regularity theory and Schauder theory

L^2-based regularity theory, which deals with weak solutions in H^1, is useful for deriving the existence of the solution. In order to derive the H^1 bound, we only need $a \in L^\infty$, rather than requiring additional regularity. Think of

$$-\partial_j (a^{j,k}(u)\partial_k u) = f,$$

where the coefficients $a^{j,k}$ may be very rough. However, it is wasteful in terms of the regularity required of a for higher regularity of the solution u.

To rectify this, we want another regularity theory that works well in this respect for nonlinear equations. This is achieved by Schauder theory, elliptic regularity theory in $C^{k,\alpha}$. Hölder spaces are naturally algebras; they play well with products, which are generally the problem with nonlinear PDEs. The gap between L^2-based regularity theory and Schauder theory is given by the famous de Giorgi-Nash-Moser estimates, which we will hopefully discuss later in the course.

12 Overview of Schauder Theory

12.1 Main theorems of Schauder theory

Schauder theory can be summarized as "Hölder-based elliptic regularity theory." Here are some of the main theorems.

Theorem 12.1 (Shauder, interior regularity, divergence form)**.** *Let U be an open subset of $\mathbb{R}^d$, and suppose that $Pu = f$, where $Pu = -\partial_j(a^{j,k}\partial_k u)$, $a \succ \lambda I$, and $a \in C^{k-1,\alpha}(\overline{U})$, Assume that $u \in C^{k,\alpha}(\overline{U})$ (with $k \geq 1$ and $0 < \alpha < 1$) and $f \in C^{k-2,\alpha}(\overline{U})$ (if $k = 1$, we assume that $f = f^0 + \sum_{j=1}^d \partial_j f^j$ with $f^0, f^j \in C^{0,\alpha}(\overline{U})$). Then for all $V \subseteq\subseteq U$, there exists a constant $C = C_V$ such that*

$$\|u\|_{C^{k,\alpha}(V)} \leq C(\|u\|_{C^0(U)} + \|f\|_{C^{k-2,\alpha}(U)}).$$

(If $k = 1$, we define $\|f\|_{C^{-1,\alpha}} := \|f^0\|_{C^{0,\alpha}} + \sum_{j=1}^d \|f^j\|_{C^{0,\alpha}}$.)

Remark 12.1. We omit the $b^j + \partial_j u + cu$ parts because they can be easily added, and they are generally dealt with on a case-by-case basis to determine what regularity you need for b and c.

Theorem 12.2 (Schauder, interior regularity, non-divergence form)**.** *Let U be an open subset of $\mathbb{R}^d$, and suppose that $Qu = f$, where $Qu = -a^{j,k}\partial_j\partial_k u$, $a \succ \lambda I$, and $a \in C^{k-2,\alpha}(\overline{U})$, Assume that $u \in C^{k,\alpha}(\overline{U})$ (with $k \geq 2$ and $0 < \alpha < 1$) and $f \in C^{k-2,\alpha}(\overline{U})$. Then for all $V \subseteq\subseteq U$, there exists a constant $C = C_V$ such that*

$$\|u\|_{C^{k,\alpha}(V)} \leq C(\|u\|_{C^0(U)} + \|f\|_{C^{k-2,\alpha}(U)}).$$

Definition 12.1. We say that U has $C^{k,\alpha}$ **boundary** if for all $x \in \partial U$, there exists an $r > 0$ such that (after possibly rearranging the axes)

$$U \cap B_r(x) = \{y \in B_r(x) : y^n > \gamma(y^1, \dots, y^{d-1})\gamma \in C^{k,\alpha}\}.$$

Theorem 12.3 (Schauder, boundary regularity, divergence form)**.** *Assume the same hypotheses in the interior divergence form theorem, and assume that ∂U is $C^{k,\alpha}$ and U is bounded. Take $Pu = f$ with the boundary condition $u|_{\partial U} = 0$. Then there exists a constant C such that*

$$\|u\|_{C^{k,\alpha}(U)} \leq C(\|u\|_{C^0(U)} + \|f\|_{C^{k-2,\alpha}(U)}).$$

Theorem 12.4 (Schauder, boundary regularity, non-divergence form)**.** *Assume the same hypotheses in the interior non-divergence form theorem, and assume that ∂U is $C^{k,\alpha}$ and U is bounded. Take $Qu = f$ with the boundary condition $u|_{\partial U} = 0$. Then there exists a constant C such that*

$$\|u\|_{C^{k,\alpha}(U)} \leq C(\|u\|_{C^0(U)} + \|f\|_{C^{k-2,\alpha}(U)}).$$

12.2 Overall strategies of the proofs

Here are strategies to prove these theorems.

Interior:

1. Prove the result in the constant coefficient case ($a^{j,k}$ constant).
2. Prove the general case using the constant coefficient case by the **method of freezing the coefficients**: Elliptic regularity is local, so we can split the space into small balls and prove the statement on each ball. The regularity of $a^{j,k}$ allows us to approximate the general problem by constant coefficient problems.

Boundary:

0. Locally straighten the boundary to reduce to the case of half balls.

1+2. Use the same method as for interior regularity. Step 0 makes the relevant constant coefficient problems be the half-space case.

We will provide two proofs for the constant coefficient case:

A. Littlewood-Paley theory proof

B. Compactness + contradiction proof.

12.3 Littlewood-Paley proof of Schauder estimates

Theorem 12.5 (Constant coefficient Schauder estimate)**.** *Let $Pu = -\partial_j(a_0^{j,k}\partial_k u) = -a_0^{j,k}\partial_j\partial_k u$, where $a_0^{j,k}$ is constant on $\mathbb{R}^d$, and $a_0 \succ \lambda I$. Assume that $|a_0^{j,k}| \leq \Lambda$, where $\Lambda \geq \lambda > 0$. For $u \in C_c^{k,\alpha}(\mathbb{R}^d)$ and $f \in C^{k-2,\alpha}(\mathbb{R}^d)$ such that $Pu = f$,*

$$\|u\|_{C^{k,\alpha}(\mathbb{R}^d)} \leq C\|f\|_{C^{k-2,\alpha}(\mathbb{R}^d)}.$$

Let us emphasize that we assume that u has *compact support*. We will focus on the case $k = 2$.

Definition 12.2. Define

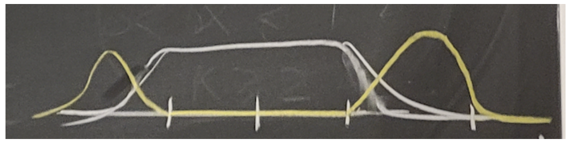
$$\chi_{\leq 0}(\xi) = \begin{cases} 1 & |\xi| \leq 1 \\ 0 & |\xi| > 1 \\ \geq 0 & \forall \xi, \end{cases}$$

$$\chi_{\leq k}(\xi) = \chi_{\leq 0}(\xi/2^k),$$

$$\chi_k(\xi) = \chi_{\leq k+1}(\xi) - \chi_{\leq k}(\xi) \qquad (\text{so } \operatorname{supp}\chi_k \subseteq \{\xi : 2^k \leq |\xi| \leq 2^{k+2}\}).$$

The **Littlewood-Paley projections** are

$$P_k v = \mathcal{F}^{-1}(\chi_k(\xi)\widehat{v}), \qquad P_{\leq k} = \mathcal{F}^{-1}(\chi_{\leq k}(\xi)\widehat{v}).$$



Observe that for all $v \in \mathcal{S}'(\mathbb{R}^d)$,

$$v = P_{\le k_0} v + \sum_{k > k_0} P_k v.$$

If v satisfies certain regularirt conditions in the same norm, $P_{\le k_0} v \to 0$ as $k_0 \to -\infty$. Note that $|\xi| \simeq 2^k$ on $\operatorname{supp} \chi_k$.

Lemma 12.1 (Littlewood-Paley characterization of $C^{0,\alpha}(\mathbb{R}^d)$)**.** *Let $v \in C^{0,\alpha}(\mathbb{R}^d)$. Then*

$$[v]_{C^{0,\alpha}} = \sup_{\substack{x,y \\ x \neq y}} \frac{|v(x) - v(y)|}{|x - y|^\alpha} \simeq \sup_{k \in \mathbb{Z}} 2^{k\alpha} \|P_k v\|_{L^\infty}.$$

Here is the proof of this lemma:

Proof. ($\gtrsim$): Both seminorms are invariant to scaling, so it suffices to consider $k = 0$. So we just have to show that

$$|P_0 v| \lesssim [v]_{C^{0,\alpha}}.$$

Since $\int \overset{\vee}{\chi}_0(y)\, dy = 0$ iff $\chi_0(0) = 0$,

$$\begin{aligned} P_0 v = \int \overset{\vee}{\chi}_0(x - y) v(y)\, dy &= \int \overset{\vee}{\chi}_0(x - y)(v(y) - v(x))\, dy \\ &\le \underbrace{\int \overset{\vee}{\chi}_0(x - y) |x - y|^\alpha\, dy}_{\text{fixed } \mathcal{S}(\mathbb{R}^d) \text{ function}} [v]_{C^{0,\alpha}}. \end{aligned}$$

($\lesssim$): Whenever we work with Littlewood-Paley theory, we should think about what scale we are working on. Let $L = |x - y|$, and choose k_0 so that $L^{-1} \simeq 2^{k_0}$. Decompose

$$v(x) - v(y) = P_{\le k_0} v(k) - P_{\le k_0} v(y) + \sum_{k > k_0} P_k v(x) - P_k v(y)$$

We can bound the latter two terms as

$$\left\| \sum_{k \ge k_0} P_{\le k_0} v \right\|_{L^\infty} \le \sum_{k < k_0} \|P_k v\|_{L^\infty}$$

$$
\begin{aligned}
&\leq \sum_{k\geq k_0} 2^{-k\alpha}[v]_{C^{0,\alpha}} \\
&\simeq L^{\alpha}[v]_{C^{0,\alpha}}.
\end{aligned}
$$

We can bound the first terms using the fundamental theorem of calculus:

$$
\begin{aligned}
|P_{\leq k_0}v(x) - P_{\leq k_0}v(y)| &\leq |\nabla P_{\leq k_0}v\|_{L^\infty}L \\
&\leq \sum_{k\leq k_0} \|\nabla P_k v\|_{L^\infty}L \\
&\lesssim L\sum_{k\leq k_0} 2^k 2^{-k\alpha}[v]_{wtC^{0,\alpha}} \\
&\simeq LL^{-(1-\alpha)}[v]_{\widetilde{C}^{0,\alpha}}.
\end{aligned}
$$

□

Now we can prove the theorem.

Proof. We have $P(P_k u) - P_k f$, so

$$
a^{j,\ell}\xi_j\xi_\ell\widehat{P_k u} = \widehat{P_k f}.
$$

Since $\lambda|\xi|^2 \leq a_0^{j,\ell}\xi_j\xi_k$,

$$
\widehat{P_k u} = \frac{2^{2k}}{a^{j,\ell}\xi_j\xi_\ell}\widehat{P_k f}\widetilde{\chi}_k\frac{1}{2^{2k}} = \frac{1}{2^{2k}}\underbrace{\frac{2^{2k}}{a^{j,\ell}\xi_j\xi_\ell}\widehat{\chi}_k}_{\eta_k(\xi)}\widehat{P_k f},
$$

where $\widetilde{\chi}_k = 1$ on $\operatorname{supp}\chi_k$ and $\operatorname{supp}\widetilde{\chi}_k \subseteq \{|\xi| \simeq 2^k\}$. Then

$$
P_k u = 2^{-2k}\overset{\vee}{\eta}_k * P_k f,
$$

so

$$
\|P_k u\|_{L^\infty} \leq C2^{-2k}\|P_k f\|_{L^\infty} \leq C2^{-2k-\chi k}[f],
$$

which completes the proof. □

12.4 Compactness and contradiction proof of Schauder estimates

Proof. Here are the steps:

1. Assume that the desired inequality fails. Then there exist $a_n^{j,k}, u_n, f_n$ such that (after normalization)

$$
P_n u_n = f_n, \qquad [u_n]_{C^{2\alpha}} = 1, \qquad [f_n]_{C^{0,\alpha}} \leq \frac{1}{n}.
$$

After translation, we may also ensure that for some $\eta_n \in \mathbb{R}^d$,

$$|D^2 u_n(\eta_n) - D^2 u_n(0)| \geq c|\eta_n|^\alpha.$$

Using scaling, we can assume that $|\eta_n| = 1$.

2. Another massaging: Define $v_n(x) = u_n(x) - u_n(0) - xDu_n(0) - \frac{1}{2}x^2 D^2 u_n(0)$ to make $D^2 v_n(0) = 0$. Then

$$P_n v_n = \wr f_n, \qquad \widetilde{f}_n \to 0, [D^2 v_n]_{C^{0,\alpha}} = 1, \qquad |D^2 v_n(\eta_j)| \geq c.$$

3. Take the limit: Let $a_n^{j,k} \to a_\infty^{j,k}$, $\widetilde{f}_n \to 0$, $v_n \to v$, and $\eta_n \to \eta_\infty$. Then $P_\infty v = 0$ on $\mathbb{R}^d$, while

$$[D^2 v]_{C^{0,\alpha}} \leq 1, \qquad D^2 v(\eta_\infty) \neq 0.$$

But now use Liouville's theorem for P_∞ (using Liouville's theorem for the Laplace equation) to get that $D^2 v(\eta_\infty) = 0$, a contradiction. $\square$

13 Maximum Principles for Solutions to Elliptic PDEs

13.1 The weak maximum principle

Today, we will cover maximum principles. This material corresponds to section 6.5 in Evans' textbook. This is a theory for solutions to elliptic PDEs in terms of their pointwise values (inherently *scalar*). Here, it is very important that $u : U \to \mathbb{R}$ is real-valued.

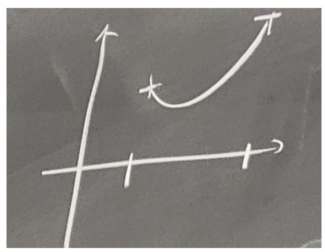
For today's lecture, it is more convenient to consider operators in non-divergence form:

$$Pu = -a^{j,k}\partial_j\partial_k u_k + b^j\partial_j u + cu.$$

We assume the ellipticity condition, that $a \succ \lambda I$ for some $\lambda > 0$, and we assume that $a, b, c \in L^\infty$. (Often, we will start with $c = 0$.)

The theory of maximum principles should be thought of as a generalization of the theory of convex functions on $\mathbb{R}$. In the case of convex functions on $\mathbb{R}$, we have the following theorem.

Theorem 13.1. *Suppose $u : I \to \mathbb{R}$ is convex. Then $\max_I u = \max_{\partial I}$, i.e. the maximum is attained on the boundary.*



One way to generalize 1 dimensional convex functions is to look at convex functions in d dimensions. This is very useful, but it may be too restrictive. Instead, we should think of subsolutions to elliptic PDEs.

Definition 13.1. We say that $u \in C^2(U)$ is a **(classical) subsolution** if $Pu \le 0$.

Remark 13.1. When $d = 1$ and $P = -a\partial_x^2$ with $a > 0$, $Pu \le 0$ if and only if u is convex.

Theorem 13.2 (Weak maximum principle)**.** *Let U be a connected, bounded, open subset of $\mathbb{R}^d$. Let $u \in C^2(U) \cap C(\overline{U})$ with $Pu \le 0$. Assume for now that $c = 0$. Then*

$$\max_{\overline{U}} u = \max_{\partial U} u.$$

Proof. Step 1: Consider strict subsolutions $Pu < 0$. We will show that no interior maximum is possible. Suppose, for contradiction, that $x_0 \in U$ is a (local) maximum. Then $Du(x_0) = 0$, and the second derivative test tells us that $D^2u(x_0) \le 0$. We have

$$0 > Pu(x_0)$$

$$= -a^{j,k}\partial_j\partial_k u|_{x=x_0} + b^j \underbrace{\partial_j u|_{x=x_0}}_{=Du=0} + \underbrace{c}_{=0} u$$

We will interpret the first term as a trace. Call $h = D^2 u$. Since a is positive definite, we can find an orthogonal matrix O such that $OaO^{-1} = D$, where D is diagonal with positive entries e_j. This makes $a^{j,k}\partial_j\partial_k = O_{j,j'}e_{j'}\delta_{j',k'}O_{k,k'}$. Then $a^{j,k}h_{j,k} = P_{j,j'}e_{j'}\delta_{j',k'}O_{k,k}h_{j,k}$.

$$\begin{aligned} &= -\operatorname{tr}(aD^2u) \\ &\geq 0 \end{aligned}$$

This is a contradiction.

Step 2: Upgrade to all subsolutions u. Introduce the approximation

$$u_\varepsilon = u + \varepsilon v,$$

where v is a strict subsolution: $Pv < 0$ with $v \in C^2(U) \cap C(\overline{U})$. Then $u_\varepsilon \to u$ uniformly on $\overline{U}$, and

$$Pu_\varepsilon = Pu + \varepsilon Pv \leq \varepsilon Pv < 0.$$

How do we construct a strict subsolution v? We want something that is convex. A good candidate is $v = e^{x^1}$ because

$$-a^{j,k}\partial_j\partial_k(e^{x^1}) = -a^{1,1}e^{x^1} < 0.$$

We want to introduce a function which has a second order derivative much smaller than a first order derivative. So instead consider $e^{\mu x^1}$, where μ is large. Then

$$-a^{j,k}\partial_j\partial_k(e^{\mu x^1}) = -a^{1,1}e^{\mu x^1} \leq -\lambda\mu^2 e^{\mu x^1},$$

$$|b^j\partial_j e^{\mu x^1}| = |-b^j\mu e^{\mu x^1}| \leq \sup|b| \cdot \mu e^{\mu x^1}.$$

So if μ is large, $Pv < 0$. $\square$

Definition 13.2. We say that $u \in C^2(U)$ is a **(classical) supersolution** if $Pu \geq 0$.

13.2 The weak minimum principle, extension of the weak maximum principle, and the comparison principle

Theorem 13.3 (Weak minimum principle)**.** *Have the same hypotheses except assume that $Pu \geq 0$ and $c = 0$. Then*

$$\min_{\overline{U}} u = \min_{\partial U} u.$$

Remark 13.2. u is a solution if and only if it is a subsolution and a super solution. So under the same hypotheses with $Pu = 0$, we get

$$\max_{\overline{U}} |u| = \max_{\partial U} |u|.$$

Corollary 13.1 (Weak maximum principle, $c \geq 0$)**.** *Suppose U is a bounded, open connected subset of $\mathbb{R}^d$ and $u \in C^2(U) \cap C(\overline{U})$. For $Pu \leq 0$.*

$$Pu \leq 0 \implies \max_{\overline{U}} \leq \max_{\partial U} u^+,$$

$$Pu \geq 0 \implies \min_{\overline{U}} \leq \min_{\partial U} u^-,$$

where

$$u^+ = \begin{cases} u & \text{if } u > 0 \\ 0 & \text{if } u \leq 0, \end{cases} \qquad u^+ = \begin{cases} 0 & \text{if } u \geq 0 \\ -u & \text{if } u < 0. \end{cases}$$

Proof. Here is the max part: Let $V = \{x \in U : u(x) > 0\}$, and let $Qu = Pu - cu$. Q satisfies the hypotheses and has no zero order term: $u \leq -cu \leq 0$ in V. The weak maximum principle for Q on V gives $\max_{\overline{V}} u \leq \max_{\partial V} u$. Note that the maximum of u on ∂V is the maximum of u on ∂U. So we get the claim. □

Theorem 13.4 (Comparison principle)**.** *Let U be an open, bounded, connected subset of $\mathbb{R}^d$. Let P be elliptic with $c \geq 0$. Suppose $u, v \in C^2(U) \cap C(\overline{U})$ with $Pu \leq 0$ in U and $Pv \geq 0$ in U. If $U \leq v$ on ∂U, then $u \leq v$ on U.*

Proof. This is an application of the previous corollary to $u - v$, which is a subsolution. □

13.3 The strong maximum principle

Theorem 13.5 (Strong maximum principle)**.** *Let U be an open, bounded, connected subset of $\mathbb{R}^d$, and let $c = 0$. Let $u \in C^2(U) \cap C(\overline{U})$ be such that $Pu \leq 0$. If u has a maximum at $x_0 \in U$ ($u(x) - \max_{\overline{U}} u$), then u is constant on U.*

Think of the picture of convex functions. The only way to have a maximum in the interior is if the whole function is constant (the graph is a horizontal straight line).

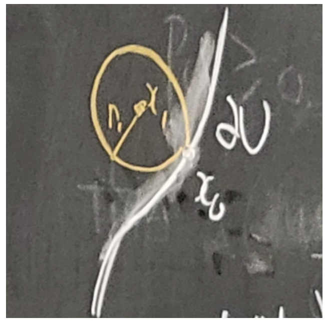
Theorem 13.6 (Hopf's lemma)**.** *Let U be an open, bounded, connected subset of $\mathbb{R}^d$. Suppose that $x_0 \in \partial U$ is such that*

(i) there exists some $x_1 \in U$ and $r_1 > 0$ such that $B_{r_1}(x_1) \subseteq U$ and $\overline{B_{r_1}(x_1)} \cap \partial U = \{x_0\}$,

(ii) $u(x_0) \geq u(x)$ in $\overline{B_{r_1}(x_1)}$,

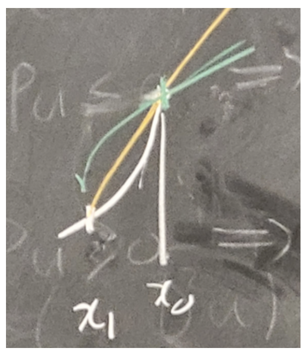
(iii) $u(x_0) > u(x)$ in $B_{r_1}(x_1)$.

Then the normal derivative $\frac{\partial}{\partial \nu}|_{x=x_0} > 0$.



Remark 13.3. We should already be able to tell that $\frac{\partial}{\partial \nu}|_{x=x_0} \geq 0$. The real content of the theorem is the strict positivity.

In the picture of convex functions, take an interior point x_1 and look at the chord connecting x_1 and the boundary point.



The idea is that this chord must have positive slope, so the actual slope of the original function at that point should be greater than the slope of the chord.

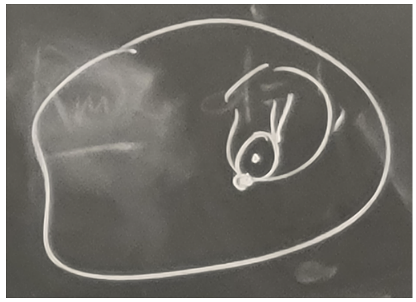
Proof. Without loss of generality, take $x_1 = 0$. Consider $v = e^{-\mu r_1^2} - e^{\mu |x|^2}$ so that $v(x) = 0$ on $\{|x| = r_1\}$. Then $Pv \geq 0$ on $B_{r_1} \setminus B_{r_1/2}$ for large μ (this is the same type of computation as before). Try to compare u to $w = v + u(x_0)$, where

$$Pw = Pv + Pu(x_0) = Pv \geq 0.$$

Let $V = B_{r_1} \setminus B_{r_1/2}$, so $\partial V = \partial B_{r_1} \cup \partial B_{r_1/2}$. On the outer boundary ∂B_{r_1}, $w = u(x_0) \geq u$. On the inner boundary $\partial B_{r_1/2}$, $w = \varepsilon v + u(x_0)$. So for small enough ε, on the inner boundary, $u(x_0) > u(x) + \varepsilon(-v)$. By the comparison principle, $w \geq u$ on $V = B_{r_1} \setminus B_{r_1/2}$. Thus,

$$\left.\frac{\partial u}{\partial \nu}\right|_{x=x_0} \geq \left.\frac{\partial v}{\partial \nu}\right|_{x=x_0} > 0. \qquad \square$$

Proof. Let $V = \{x \in U : u(x) \leq M\}$, where $M = \sup_{\overline{U}} u$. Then for $x_0 \in U$, if $u(x_0) = M$, then $V \subsetneq U$. Assume for contradiction that $V \neq \varnothing$. Find a point x_1 closer to ∂V than ∂U and consider the biggest r_1 such that $B_{r_1}(x_1) \subseteq V$. Let $x_0 \in B_{r_1}(x_1) \cap \partial V$. Let $x_0' \in B_{r_1}(x_1) \cap \partial V$.



We may arrange, by taking x_1 close enough to ∂V, so that Hopf's lemma is applicable. This tells us that $\frac{\partial}{\partial \nu} u|_{x=x_0'} \neq 0$. But this contradicts the facr that $u(x_0') = M$ implies $Du|_{x=x_0'} = 0$ $\square$

14 General Boundary Value Problems for Elliptic PDEs

14.1 How do we make sense of "regular" boundary value problems for eliiptic PDEs?

In this lecture, we will assume that P is an elliptic operator in divergence form:

$$Pu = -\partial_j(a^{j,k}\partial_k u) + b^j\partial_j u + cu.$$

Let U be an open, bounded, connected subset of $\mathbb{R}^d$ with C^1 boundary ∂U. A general boundary value problem might be of the form

$$\begin{cases} Pu = 0 & \text{in } U \\ Bu|_{\partial_U} = g & (\text{on } \partial U) \end{cases}$$

for some operator B.

So far, we have focused on the Dirichlet boundary condition

$$\begin{cases} Pu = 0 & \text{in } U \\ u|_{\partial_U} = g & (\text{on } \partial U) \end{cases}$$

By introducing an extension $\widetilde{g}$ of g to U, we could set, without loss of generality, $g = 0$. With this reduction, the problem we have considered is

$$\begin{cases} Pu = 0 & \text{in } U \\ u|_{\partial_U} = 0 & (\text{on } \partial U). \end{cases}$$

Our goal now is to generalize our elliptic theory to other boundary conditions. This will force us to consider what is a "regular" boundary value problem for PDEs. In order to solve a k-th order ODE, you need k pieces of data on the boundary. For the wave equation, which is a second order PDE, you impose boundary values and normal derivative values. Unlike ODEs, the wave equation, or Cauchy-Kovalevskaya, when we work with an elliptic PDE like $-\Delta u = f$, we do not prescribe the full $u, \frac{\partial}{\partial v}u$ on ∂U. How do we rigorously justify this high level discussion? We will see two approaches.

14.2 Weak formulations of boundary problems

Prove a uniqueness theorem via the energy method.

Example 14.1. If $P = -\Delta$ and we are solving

$$\begin{cases} -\Delta u = 0 & \text{in } U \\ u|_{\partial_U} = g & (\text{on } \partial U) \end{cases}$$

then

$$0 = \int_U -\Delta u u \, dx = \int |Du|^2 \, dx.$$

Note the parallel between this basic consideration and our weak formulation of the Dirichlet problem: $u \in H^1$ solves the Dirichlet problem

$$\begin{cases} Pu = f & \text{in } U \\ u|_{\partial_U} = g & (\text{on } \partial U) \end{cases}$$

if and only if $u \in H_0^1(U)$ and $-\Delta u = f$ in the sense of $\mathcal{D}'(U)$. This is equivalent to

$$\int_U a^{j,k} \partial u \partial_k \varphi + b^j \partial_j u \varphi + cu\varphi \, dx = \int_U f\varphi \, dx \qquad \forall \varphi \in H_0^1(U).$$

We will try to generalize this weak formulation to other boundary conditions.

Example 14.2. Consider the Neumann boundary condition

$$\begin{cases} Pu = f & \text{in } U \\ \nu^j \partial_j u|_{\partial_U} = g & (\text{on } \partial U) \end{cases}$$

We can rewrite this as

$$\begin{cases} Pu = f & \text{in } U \\ a^{j,k} \nu_k \partial_j u|_{\partial_U} = g & (\text{on } \partial U) \end{cases}$$

In the case of the Laplace equation, this is the same. From the point of view of differential geometry, this is a more natural quantity to look at because ν_k is dh, where h is the boundary defining form. The natural Riemannian metric in this problem is a. By an extension procedure, we can write the problem as

$$\begin{cases} Pu = f & \text{in } U \\ a^{j,k} \nu_k \partial_j u|_{\partial_U} = 0 & (\text{on } \partial U) \end{cases}$$

For simplicity, assume $b = c = 0$. Then we have the formal computation

$$\int_U f\varphi \, dx = \int_U -\partial_j (a^{j,k} \partial_j u) \varphi \, dx = \int_U a^{j,k} \partial_j u \partial_k \varphi \, dx - \int_{\partial U} \underbrace{\nu_j a^{j,k} \partial_k u}_{=0} \varphi \, dA.$$

This motivates the following definition:

Definition 14.1. We say that u **satisfies the Neumann boundary problem** if for all $\varphi \in H^1(U)$,

$$\int_U a^{j,k} \partial_j u \partial_k \varphi \, dx = \int_U f\varphi \, dx.$$

Remark 14.1. If $u \in C^1$ then this formulation should be equivalent to the classical one.

Once we formulate the problem like this, the L^2 theory is easy to generalize.

Theorem 14.1. *Suppose ∂U is C^1, $a \succ \lambda I$ in U, and $a \in L^\infty$ (also $b, c \in L^\infty$). Then*

1. *For any $\mu \in \mathbb{R}$, the map $u \mapsto Pu - \mu u$ associated to the Neumann boundary value problem*
$$(\mathrm{NP}_\mu) \begin{cases} Pu - \mu u = f & \text{in } U \\ a^{j,k}\nu_j \partial_j u|_{\partial U} = g & \text{(on } \partial U\text{)} \end{cases}$$
is Fredholm with index 0 from $H^1(U) \to (H^1(U))^* \subseteq H^{-1}(U)$. That is, one of the following holds:
   (i) For all $f \in L^2(U)$, there exists a unique $u \in H^1$ which solves the Neumann boundary problem (NP_μ).
   (ii) There exists a solution $v \neq 0$ to (NP_μ) with $f = 0$. Furthermore, for $\mu \gg 1$, alternative (i) applies.

2. *If ∂U is C^k and $a, b, c \in C^k$, then*
$$\|u\|_{H^{1+k}(U)} \lesssim \|f\|_{H^{k-1}(U)} + \|u\|_{H^k(U)}.$$

Example 14.3. Take $P = -\Delta$ and solve
$$\begin{cases} -\Delta u = 0 \\ u|_{\partial U} = 0. \end{cases}$$
This has a nontrivial solution $v = \text{const} \neq 0$.

This leads to solvability for f orthogonal to the kernel of the adjoint. In this case, this is equivalent to $\int_U f\, dx = 0$.

For other boundary conditions, this weak formulation also makes sense.

Definition 14.2. We say that u **satisfies the Robin boundary problem** if for all $\varphi \in H^1(U)$,
$$\int_U a^{j,k}\partial_j u \partial_k \varphi\, dx + \int_{\partial U} \alpha u \varphi\, dS = \int_U f\varphi\, dx.$$

Example 14.4 (Oblique Dirichlet boundary condition)**.** Suppose $b = c = 0$, and consider the problem
$$(\mathrm{OP}) \begin{cases} Pu = f \\ X^j \partial_j u = 0, \end{cases}$$

where X is transversal to ∂U, outward. Then $X = X^\perp + X^\top$, where $X^\perp$ is parallel to $a^{j,k}\nu_k \vec{e_k}$. Normalize to make $X^\perp = a^{j,k}\nu_j \vec{e_k}$. This tells us that

$$\int_U a^{j,k}\partial_j u \partial_k \varphi + \int_{\partial U} X^\top u\varphi \, dA = \int_U f\varphi \, dx.$$

The second term is trickier to make sense of, since we need to make sense of the trace. As an exercise, check that $\int_{\mathbb{R}^{d-1}} \partial uv \, dx$ is well defined for $u, v \in H^{1/2}(\mathbb{R}^{d-1})$. This is just barely well-defined, however, in the sense of the trace theorem needing $H^{1/2}$.

14.3 The "microlocal" formulation

The reference for this section is volume 1 of Taylor's PDE book, section 5.11. Look at the Laplace equation $-\Delta u = 0$ in the half space $\mathbb{R}^d_+$. Write z for the last variable and x for the remaining $d-1$ variables, so this is $-\partial_z^2 - \Delta_x u = 0$. Suppose we have boundary conditions $u|_{\partial U} =?$ and $\partial_z u|_{\partial U} =?$. We can view this as an evolution equation in the z variable and take the Fourier transform in x to get

$$(-\partial_z^2 + |\xi|^2)\widehat{u} = 0$$

with boundary conditions $\widehat{u}|_{z=0} = g$ and $\partial_z \widehat{u}|_{z=0} = h$. This gives

$$\widehat{(z, \xi)} = a_+(\xi)e^{|\xi|z} + a_-(\xi)e^{-|\xi|z}.$$

However, the first term $e^{|\xi|z}$ is a problem because growth in Fourier space corresponds to a lack of regularity in physical space. So in order to have boundary regularity, we want $a_+(\xi) = 0$. This means that we are only left with half of the full freedom to choose $\widehat{g}$ and $\widehat{h}$.

The claim is that the constant coefficient picture generalizes to the variable coefficient picture. The idea is that using the technique of "freezing the coefficients," we can formulate the notion of a "regular" elliptic boundary value problem, for which we have elliptic regularity and the Fredholm property, based on the constant coefficient computation.

Here, we assume that $a, b, c \in C^\infty(\overline{U})$ and that ∂U is C^∞.

Definition 14.3. For $k \geq 1$, define

$$H^{k-1/2}(\partial U) = \{g = v|_{\partial U} : v \in H^k(U)\},$$

with the norm

$$\|g\|_{H^{k-1/2}(\partial U)} = \inf_{u : u|_{\partial U} = g} \|u\|_{H^k(U)}.$$

Remark 14.2. If we define fractional Sobolev spaces on manifolds, this will actually be the $k - 1/2$ Sobolev space on ∂U.

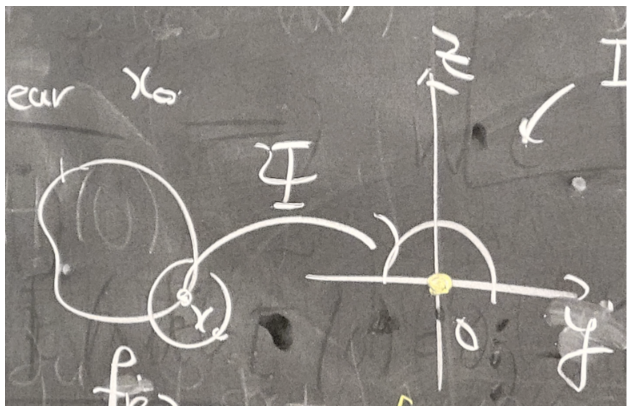
Now consider the boundary problem

$$\begin{cases} Pu = f & \text{in } U \\ Bu|_{\partial U} = g. \end{cases}$$

Here, we assume that $P : C^\infty(U) \to C^\infty(U)$ and $B(\cdot)|_{\partial U} : C^\infty(U) \to C^\infty(\partial U)$. Given $x_0 \in \partial U$, there exists a boundary straightening map near x_0. In these variables, write

$$P = -\partial_z^2 + P_1(y, z, D_y^2)\,\partial_z + P_0(y, z, D_y, D_y^2),$$

$$B = b\partial_z + B_0(y, z, \partial_y).$$



Say x_0 is mapped to 0, and let P_{x_0} be the frozen constant coefficient operator

$$P_{x_0} = -\partial_z^2 + P_1(0, 0, D_y)\partial_z + P_0(0, 0, ; D_y, D_y^2),$$

$$B_{x_0} = b(0, 0)\partial_z + B_0(0, 0, \partial_y).$$

Definition 14.4. A boundary value problem is a **regular elliptic boundary value problem** if for all $x_0 \in \partial U$, for all $\xi \in \mathbb{R}^{d-1}$, and for all ζ, there exists a unique bounded solution to the ODE

$$P_{x_0}\widehat{u}(z, \xi) = 0, \qquad B_{x_0}\widehat{u}(z, \xi) = \zeta.$$

This is called the **Loputinski-Shapiro condition**. This is like if we pretend we take the Fourier transform and replace ∂_y by $c\xi$. This condition gives an ODE in z.

Theorem 14.2. *For a regular elliptic boundary value problem, the map* $H^{k+2}(U) \ni u \mapsto (Pu, Bu) \in H^k(U) \times H^{k-(\text{order } B)-1/2}(\partial U)$ *is Fredholm, and we have elliptic (boundary) regularity*

$$\|u\|_{H^{k+2}(U)} \lesssim \|f\|_{H^k(U)} + \|Bu\|_{H^{k-(\text{order } B)-1/2}(\partial U)} + \|u\|_{H^{k+1}(U)}.$$

15 Unique Continuation for Elliptic PDEs and Introduction to Hyperbolic PDEs

15.1 Unique continuation for elliptic PDEs

The original plan was for this lecture to cover one final topic for elliptic PDEs: unique continuation. Here is the main theorem.

Theorem 15.1 (Aronszajn). *Let $U \subseteq \mathbb{R}^d$ be open and connected, and consider the elliptic partial differential operator P with*

$$Pu = -\partial_j(a^{j,k}\partial_k u) + b^j \partial_j u + cu,$$

where $a^{j,k} b^j, c \in C^\infty(U)$ with $a \succ \lambda I$ in U. Let $u \in H^1(U)$. If $Pu = 0$ in U and $u = 0$ in a nonempty open subset $Q \subseteq U$, then $u = 0$ in U.

For holomorphic functions, the way we prove this is to say that holomorphic functions are analytic and look at the domain of convergence of a Taylor series. The way we prove this for solutions to elliptic PDEs is via an a priori estimate.

Lemma 15.1 (Carleman estimate). *Let $v \in C_c^\infty(\mathbb{R}^d)$. and suppose that $\nabla \psi \neq 0$. Then*

$$t^2 \|e^{t\psi} v\|_{L^2} + t\|e^{t\psi} \nabla v\|_{L^2} \leq C \|e^{t\psi} P v\|_{L^2}.$$

A good reference for this is the book *Carlesman Estimates* by Lerner. This is related to inverse problems and other non-well-posed problems in PDEs.

15.2 Linear hyperbolic PDEs

Instead of formally defining what a hyperbolic PDE is, which is difficult and not entirely productive. Instead, we will give a ‘working definition” of how people think of hyperbolic PDEs.

Definition 15.1. A **hyperbolic PDE** is an evolutionary PDE with two characteristics:

- “#/order of time derivatives” = “#/order of space derivatives”.
- (local) well-posedness of the initial value problem

$$\begin{cases} P\phi = 0 \\ (\phi, \partial_t \phi, \dots, \partial_t^{N-1}\phi)|_{t=0} = (g_0, \dots, g_{N-1}), \end{cases}$$

  where N is the order of the time derivatives.

This second condition is really what people think of when they talk about hyperbolic PDEs.

Example 15.1. The wave equation $(-\partial_t^2 + \Delta)\phi = 0$ is hyperbolic.

Example 15.2. The equation $(-\partial_t + x^j\partial_j)\phi = 0$ is hyperbolic.

Example 15.3 (Non-examples)**.** The heat equation $(\partial_t - \Delta)\phi = 0$ and the Schrödinger equation $(\partial_t - i\Delta)\phi = 0$ are *dispersive* but not hyperbolic.

Example 15.4. The Laplace equation $(\partial_t^2 + \Delta)\phi = 0$ is not hyperbolic because it does not have local well-posedness of the initial value problem.

Local well-posedness of the initial value problem is related to the energy estimate.

Example 15.5 (Linear constant coefficient system)**.** Let

$$\Phi = \begin{bmatrix} \Phi^{(1)} \\ \vdots \\ \Phi^{(n)} \end{bmatrix},$$

and suppose we have a system of linear, constant coefficient PDEs

$$B\partial_t\Phi = A^j\partial_{x^j}\Phi,$$

where A is an $n \times n$ matrix. Without loss of generality, assume we have

$$\partial_t\Phi = A^j\partial_{x^j}\Phi,$$

What guarantees uniqueness of a solution to the initial value problem? That is, what condition do we need on A to guarantee the validity of the energy estimate?

$$\int_{\mathbb{R}^d} \underbrace{\Phi^{(k)}\partial_t\Phi^{(k)}}_{=\frac{1}{2}\partial_t\int\Phi^{(k)}\Phi^{(k)}} + \underbrace{\Phi^{(k)}(A^j)^{(k)}_{(\ell)}\partial_j\Phi^{(\ell)}}_{\frac{1}{2}\int(A^j)^{(k)}_{(\ell)}\Phi^{(k)}\partial_j\Phi^{(\ell)} - \frac{1}{2}(A^j)^{(k)}_{(\ell)}\partial_j\Phi^{(k)}\Phi^{(\ell)}} \, dx = \int \Phi^{(k)}F^{(k)}\,dx.$$

We get the following identity:

$$\frac{1}{2}\partial_t\int|\Phi|^2\,dx + \frac{1}{2}\int((A^j)^{(k)}_{(\ell)} - (A^j)^{(\ell)}_{(k)})\Phi^{(k)}\partial_j\Phi^{(\ell)}\,dx = \int F\cdot\Phi\,dx,$$

where the second term is 0 if A^j is symmetric.

This tells us that if A^j is symmetric, then the energy estimate holds:

$$\int|\Phi|^2(t)\,dx = \int|\Phi|^2(0)\,dx + \int_0^t\int F\cdot\Phi\,dx\,dt.$$

This gives uniqueness

Theorem 15.2. *The linear, constant coefficient system*

$$\partial_t \Phi = A^j \partial_{x^j} \Phi$$

is hyperbolic if and only if the A^j *are symmetric. That is the initial value problem is well-posed in* L^2*, meaning for every* $\Phi_0 \in L^2(\mathbb{R}^d)$*,and* $F \in L^1_t((-\infty,\infty);L^2_x)$*, there exists a unique* $\Phi \in C_t((-\infty,\infty);L^2_x)$ *solving the system.*

We use the notation $\phi \in C_t(I;X)$ to mean that the function $\phi : I \to X$ sending $t \mapsto \phi(t)$ is continuous, where $C_t(I;X)$ has the norm

$$\|\phi\|_{C_t(I;X)} := \sup_{t\in I} \|\phi(t,\cdot)\|_X = \|\phi\|_{L^\infty_t(X)} < \infty.$$

Example 15.6 (1st order formulation of $\Box\phi = f$)**.** Let the **d'Alembertian** be $\Box = -\partial_t^2 + \Delta$. Then

$$\Box\phi = f \iff \partial_t\phi = \psi, \partial_t\psi = \Delta\phi - f.$$

We can write this system as

$$\partial_t \begin{bmatrix} \phi \\ \psi \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ \Delta & 0 \end{bmatrix} \begin{bmatrix} \phi \\ \psi \end{bmatrix} - \begin{bmatrix} 0 \\ f \end{bmatrix}.$$

If we take the Fourier transform of the matrix, we get

$$\begin{bmatrix} 0 & 1 \\ -|\xi|^2 & 0 \end{bmatrix}.$$

and if we diagonalize this, we get

$$\begin{bmatrix} +i|\xi| & 0 \\ 0 & -i|\xi| \end{bmatrix},$$

which is anti-Hermitian. This means that the energy estimate will hold in the diagonalized variables

15.3 Goals for studying hyperbolic PDEs

Here are our goals for studying hyperbolic PDEs:

1. (Local) well-posedness of the initial value problem for variable-coefficient wave equations,

   $$P\phi = \partial_\mu(g^{\mu,\nu}\partial_\nu\phi) + b^\mu\partial_\mu\phi + c\phi,$$

   where g is a **Lorentzian metric**, a non-degenerate symmetric $(d+1)\times(d+1)$ matrix with signature $(-,+,+\cdots,+)$ (meaning that the eigenvalues of g have signs $-,+,+\ldots,+$). This condition can also be stated as: for every (t,x), there exists an invertible matrix M such that $M^{-1}g(t,x)M = \mathrm{diag}(-1,+1,+1,\ldots,+1)$.

Example 15.7. When $g = \text{diag}(-1, +1, +1, \ldots, +1)$ and $b = c = 0$, $P = \Box$.

2. Long-time behavior of the solutions: If we look at this in general, it immediately becomes a research topic.[6] Instead, we will focus on long-time behavior of solutions to equations where P is a small variant of $\Box$.

15.4 Grönwall's inequality

Our treatment for the well-posedness of the initial value problem for variable coefficient wave equations will be closer to Ringström's book *the Cauchy Problem in General Relativity* than it will be to Evans' book.

Our setting is

$$P\phi = \partial_\mu(g^{\mu,\nu}\partial_\nu\phi) + b^\mu\partial_\mu\phi + c\phi.$$

We want to derive energy estimates for

$$\begin{cases} P\phi = f & \text{in } \mathbb{R}_+ \times \mathbb{R}^d, \\ (\phi, \partial_t\phi)|_{t=0} = (g, h) & \text{on } \{t = 0\} \times \mathbb{R}^d. \end{cases}$$

We need the following preliminary tool, which was discussed in Math 222A.

Lemma 15.2 (Grönwall's inequality)**.** *Suppose that $E(t) \in C_t([0,T])$ and $r(t) \in L^1_t([0,T])$ with $E, r \geq 0$ satisfy the inequality*

$$E(t) \leq E_0 + \int_0^t r(t')E(t')\,dt' \qquad \forall 0 \leq t \leq T.$$

Then

$$E(t) \leq E_0 \exp\left(\int_0^t r(t')\,dt'\right) \qquad \forall 0 \leq t \leq T.$$

We give a proof that uses a **bootstrap argument**, i.e. continuous induction on time. First, here is a motivating computation: Take the inequality we are given, and plug in the answer into the right hand side. We get

$$E(t) \leq E_0 + E_0 \int_0^t r(t') \exp\left(\underbrace{\int_0^{t'} r(t'')\,dt''}_{R(t')}\right) dt,$$

where R is just an antiderivative of r.

$$= E_0 + E_0\left(\exp\left(\int_0^t r(t')\,dt'\right) - 1\right)$$

[6] Scattering theory is devoted to studying these problems.

$$= E_0 \exp\left(\int_0^t r(t')\,dt'\right).$$

This tells us that the solution is what we get if we try to find a fixed point when iterating the use of this bound.

Proof. Note that it suffices to prove the inequality

$$E(t) \leq E_0(1+\delta)\exp\left(\int_0^t r(t')\,dt'\right)$$

on $[0,T]$ for all δ. Also note that for some T_0, this inequality holds on $[0,T_0]$ by continuity.

Now assume that

$$E(t) \leq E_0(1+\delta)\exp\left(\int_0^t r(t')\,dt'\right)$$

on $[0,T]$. If we plug this bound into the iteration, we get

$$\begin{aligned}
E(t) &\leq E_0 + E_0(1+\delta)\int_0^t r(t')\exp\left(\int_0^{t'} r(t'')\,dt''\right)\,dt' \\
&= E_0 + E_0(1+\delta)\left(\exp\left(\int_0^t r(t')\,dt'\right) - 1\right) \\
&= E_0(1+\delta)\exp\left(\int_0^t r(t')\,dt'\right) - \underbrace{\delta E_0}_{>0}.
\end{aligned}$$

This means that this bound we assumed holds on $[0, T'+\varepsilon]$ for some ε. The result now holds by trying to do this with the supremum of all T' such that this inequality holds on $[0,T']$. We get that this supremum must be T. $\square$

16 Regularity Estimates for Variable-Coefficient Wave Equations

16.1 Well-posedness of the initial value problem for variable-coefficient wave equations

Today, we are interested in a concrete goal. We will be studying **variable-coefficient wave equations**, PDEs of the form

$$P\phi = \partial_\mu(g^{\mu,\nu}\partial_\nu\phi) + b^\mu\partial_\mu\phi + c\phi,$$

where the key assumption is that g is a symmetric matrix with signature $(-,+,+\dots,+)$. The example we should keep in mind is $g = \mathrm{diag}(-1,1,1,\dots,1)$, $b=0$, $c=0$; this makes $P=\Box$. We are solving the initial value problem

$$\begin{cases} P\phi = f & \text{in } (0,\infty)_t \times \mathbb{R}^d \\ (\phi, \partial_t\phi)|_{t=0} = (g,h) & \text{on } \{t=0\} \times \mathbb{R}^d. \end{cases}$$

We further assume that $g^{\mu,\nu}, b^\mu, c$ are bounded with bounded derivatives of all orders. We also assume a restricted form of g (which we will later show is not much of a restriction): $g^{tt} = -1$ and $g^{t,x^j} = 0$. This means that if we write g as a matrix,

$$g = \begin{bmatrix} -1 & 0_{1\times d} \\ 0_{d\times 1} & \overline{g}, \end{bmatrix}$$

where $\overline{g}$ is **uniformly elliptic** ($\overline{g} \succ \lambda I$).

Our concrete goal is to prove the following theorem:

Theorem 16.1. *The initial value problem is well-posed in $H^k \times H^{k-1}$ for all $k \in \mathbb{Z}$. That is,*

(i) (Existence) Given $(g,h) \in H^k \times H^{k-1}$ and $f \in L^1_t(H^{k-1})$, there exists a solution ϕ to the initial value problem in the class $C_t(\mathcal{H}^k)$.

(ii) (Uniqueness) The solution ϕ in $C_t(\mathcal{H}^k)$ to the initial value problem with (f,g,h) as in (i) is unique.

(iii) (Continuous dependence)

$$\sup_t \|\phi(\phi, \partial_t\phi)\| \le C_k(\|(g,h)\|_{\mathcal{H}^k} + \|f\|_{L^1_t(H^{k-1})}).$$

Here, $\mathcal{H}^k = H^k \times H^{k-1}$, and by $\phi \in C_t(I;\mathcal{H}^k)$, we mean that $\phi \in C_t(I;H^k)$ and $\partial_t\phi \in C_t(I;H^{k-1})$.

We will use the convention that $\mathbb{R}^{1+d} = \{(t = x^0, x^1, \dots, x^d)\}$. The Greek indices μ, ν will range from $0, 1, \dots, d$, while the indices j, k, ℓ will range from $1, \dots, d$. We will also denote $g^{t,t} = g^{0,0}, g^{t,x^j} = g^{0,j}$.

Remark 16.1. The problem is **time reversible**. If we send $t \mapsto -t$, the equation is essentially unchanged.

The reference for this topic is chapters 6-7 of Ringström's book.

16.2 Energy inequality for P

The basic ingredient in this proof is an energy inequality for P. Suppose $P\phi = f$. The idea is to multiply the equation by $\partial_t \phi$ and "integrate by parts." Why should we multiply by $\partial_t \phi$ instead of ϕ? This is a generalization of what we do in the classical wave equation, and we will be able to give a more insightful answer to this once we discuss calculus of variations for problems of this type. The key observation is this integration by parts idea, but in divergence form:

$$\begin{aligned}\partial_\mu (g^{\mu,\nu} \partial_\nu \phi) \partial_t \phi &= -\partial_t^2 \phi \partial_t \phi + \partial_j (\overline{g}^{j,k} \partial_k \phi) \partial_t \phi \\ &= \partial_t \left(-\frac{1}{2} (\partial_t \phi)^2 \right) + \partial_j (\overline{g}^{j,k} \partial_k \phi \partial_t \phi) - \overline{g}^{j,k} \partial_k \phi \partial_j \partial_t \phi\end{aligned}$$

Since g is symmetric, this last term can be written as $-\overline{g}^{j,k} \partial_t (\partial_k \phi \partial_j \phi)$ by symmetrizing. Moving the ∂_t to the outside, we get

$$= \partial_t \left(-\frac{1}{2} (\partial_t \phi)^2 \right) - \frac{1}{2} \overline{g}^{j,k} \partial_j \phi \partial_k \phi + \partial_j (\overline{g}^{j,k} \partial_k \phi \partial_t \phi) + \frac{1}{2} \partial_t \overline{g}^{j,k} \partial_j \phi \partial_k \phi.$$

This form is nice because the terms that have the maximum number of derivatives are all in divergence form, while the terms that don't have the maximum number of derivatives are not in divergence form.

Integrate this on $(t_0, t_1) \times \mathbb{R}^d =: R_{t_0}^{t_1}$ (assuming the boundary term vanishes):

$$\begin{aligned}&\iint_{R_{t_0}^{t_1}} \partial_\mu (g^{\mu,\nu} \partial \nu \phi) \partial_t \phi - \frac{1}{2} \iint_{R_{t_0}^{t_1}} \partial_t \overline{g}^{j,k} \partial_j \phi \partial_k \phi \\ &= - \int_{\Sigma_{t_1}} \frac{1}{2} ((\partial_t \phi)^2 + \overline{g}^{j,k} \partial_j \phi \partial_k \phi) + \int_{\Sigma_{t_0}} \frac{1}{2} ((\partial_t \phi)^2 + \overline{g}^{j,k} \partial_j \phi \partial_k \phi \\ &\qquad + \underbrace{\lim_{R \to \infty} \int_{t_0}^{t_1} \int_{\partial B_R} \nu_j (\overline{g}^{j,k} \partial_k \phi \partial_t \phi) \, dA \, dt,}_{=0}\end{aligned}$$

where $\Sigma_t = \{t\} \times \mathbb{R}^d$.

Denote $\vec{\phi} = (\phi, \partial_t \phi)$, so $(\phi, \partial_t \phi) \in \mathcal{H}^k$ if and only if $\vec{\phi} \in C_t(\mathcal{H}^k)$.

Lemma 16.1. *For $\phi \in C_t(\mathcal{H}^1)$,*

$$\sup_{t\in[0,T]} \|\vec{\phi}\|_{\mathcal{H}^k} \leq C_T \left(\|\vec{\phi}(0)\|_{\mathcal{H}^1} + \int_0^T \|P\phi\|_{L^2}\, dt \right).$$

Proof. We may assume without loss of generality that $\phi \in C^\infty(\overline{R_0^T})$ and $\phi(t, \cdot)$ has compact support for each $t \in [0, T]$. By the computation above, if

$$E[\phi](t) = \frac{1}{2} \int_{\Sigma_t} (\partial_t \phi)^2 + \overline{g}^{j,k} \partial_j \phi \partial_k \phi \, dx,$$

then

$$\mathbb{E}[\phi](t_1) = \mathbb{E}[\phi](0) - \iint_{R_0^{t_1}} \partial_\mu (g^{\mu,\nu} \partial_\nu \phi) + \frac{1}{2} \iint_{R_0^{t_1}} \partial_t \overline{g}^{j,k} \partial_j \partial_k \phi.$$

(Note that $\lim_{R\to\infty} \int_{\partial B_R} = 0$ thanks to the support assumption. Now

$$\partial_\mu (g^{\mu,\nu} \partial_\nu \phi) = P\phi - b^\mu \partial_\mu \phi - c\phi,$$

which tells us that

$$\mathbb{E}[\phi](t_1) = E[\phi](0) + \iint_{R_0^t} P\phi \partial_t \phi \, dx\, dt + \iint_{R_0^t} (b^\mu \partial_\mu \phi \partial_t \phi + c\phi \partial_t \phi + \partial_t \overline{g}^{j,k} \partial_j \phi \partial_k \phi) \, dx\, dt.$$

Call the error

$$\mathcal{E}_0^t = \iint_{R_0^t} |b^\mu \partial_\mu \phi \partial_t \phi + c\phi \partial_t \phi + \partial_t \overline{g}^{j,k} \partial_j \phi \partial_k \phi| \, dx\, dt.$$

We get an inequality:

$$\sup_{t_1\in[0,T]} E[\phi](t_1) \leq E[\phi](0) + \sup_{t\in[0,T} \left| \iint_{R_0^t} P\phi \partial_t \phi \, dx\, dt \right| + \mathcal{E}_0^T.$$

Note that $E[\phi] \geq \frac{1}{2} \int (\partial_t \phi)^2(t)\, dx \geq \frac{\lambda}{2} \int |D_t \phi|^2(t)\, dx$. Using the fundamental theorem of calculus,

$$\int |\phi|^2(t)\, dx = \int_0^t \int \partial\phi\phi \, dx\, dt' + \int |\phi|^2(0)\, dx$$

Using Cauchy-Schwarz,

$$\leq 2 \int E(t')^{1/2} \left(\int |\phi|^2(t')\, dx \right)^{1/2} dt' + \int |\phi|^2(0)\, dx.$$

Skipping a few steps, we get

$$\sup_{t\in[0,T]} \int |\phi|^2(t)\, dt \leq \int |\phi|^2(0)\, dx + CT \sup_{t\in[0,T]} E(t).$$

The point here is that

$$\sup_{t\in[0,T]} \|\vec{\phi}\|_{\mathcal{H}^1} \leq C_T \left(\|\vec{\phi}(0)\|^2_{\mathcal{H}^1} + \sup_{t\in[0,T]} \left| \iint_{R_0^T} P\phi \partial_t \phi \, dx \, dt \right| + \mathcal{E}_0^T \right).$$

If we use Cauchy-Schwarz, we get

$$\begin{aligned}\sup_{t\in[0,T]} \left| \iint_{R_0^t} P\phi \partial_t \phi \, dx \, dt \right| &\leq \int_0^T \|P\phi(t)\|_{L^2} \|\partial_t \phi\|_{L^2} \, dt \\ &\leq C \int_0^T \|P\phi(t)\|_{L^2} E[\phi]^{1/2} \, dt \\ &\leq \int_0^T \|P\phi(t)\|_{L^2} dt \sup_{[0,T]} E[\phi]^{1/2}\end{aligned}$$

We can use Cauchy-Schwarz to absorb the energy term to the left hand side, since $E[\phi] \leq C \int (\partial_t \phi)^2 + (D_x \phi)^2$. We get

$$\sup_{t\in[0,t_1]} \|\vec{\phi}\|^2_{\mathcal{H}^1} \leq C_T \left(\|\vec{\phi}(0)\|^2_{\mathcal{H}^1} + \int_0^T \|P\phi\|_{L^2} \, dt + \int_0^{t_1} \|\phi(t)\|^2_{\mathcal{H}^1} \, dt \right).$$

This means that if we let $\mathcal{D}(t_1)$ be the left hand side and $\mathcal{D}_0$ be the first two terms on the right hand side, we get

$$\mathcal{D}(t_1) \leq \mathcal{D}_0 + \int_0^{t_1} \mathcal{D}(t) \, dt.$$

Using Grönwall's inequality, we get

$$\mathcal{D}(t) \leq \mathcal{D}_0 \exp\left(\int_0^t dt' \right) \leq \mathcal{D}_0 e^T.$$

This finishes the proof. □

16.3 Further regularity estimates for existence and uniqueness

We want to study something like $P : C_t(\mathcal{H}^k) \to L^1_t(H^{k-1})$. This means that we should look at the adjoint $P^* : C_t(H^{-(k-1)}) \to L^1_t(H^{-k})$. The dual problem here includes negative Sobolev spaces.

Lemma 16.2. *For any $k \in \mathbb{Z}$ and $\phi \in C_t(\mathcal{H}^{1+k}) \cap C^\infty_{t,x}$,*

$$\sup_{t\in[0,T]} \|\vec{\phi}(t)\|_{\mathcal{H}^{1+k}} \leq C_{T,k} \left(\|\vec{\phi}(0)\|_{\mathcal{H}^{1+k}} + \int_0^T \|P\phi\|_{H^k} \, dt \right).$$

The positive regularities will give us uniqueness for the initial value problem. The negative regularities will give us existence.

Proof. For $k > 0$, we commute the equation with D^α for $|\alpha| \le k$. Then apply the previous lemma and Grönwall's inequality. (This technique is very similar to our previous proof of higher elliptic regularity bounds. However, we don't need to use a difference quotient.)

For $k < 0$, we work with $\Phi = (1 - \Delta)^{-|k|}\phi$. (This means that we want to look at the solution to the elliptic problem $(1 - \Delta)^{|k|}\Phi = \phi$ in $\mathbb{R}^d$. Another way to write this is $\widehat{\Phi} = (1 - |\xi|^2)^{-|k|}\widehat{\phi}$.) We do this so that we don't have to deal with negative Sobolev spaces; we can study an operator that commutes well with P and use positive Sobolev spaces, instead. The key thing to notice is that $(1 - \Delta)^{-\ell} : H^s \to H^{s+2\ell}$. We also use the following:

Lemma 16.3. *For any $s \in \mathbb{R}$, the H^s norm has the Fourier characterization*

$$\begin{aligned}
\|v\|_{H^s} &= \|(1 + |\xi|^2)^{s/2}\widehat{v}\|_{L^2_\xi}^2 \\
&= \|(1 - \Delta)^{s/2} v\|_{L^2}^2.
\end{aligned}$$

When $s \in 2\mathbb{Z}$, this agrees with our sense of derivatives.

We want to compute

$$\begin{aligned}
\|P\Phi\|_{H^{|k|}}^2 &= \|(1 + \|xi|^2)^{|k|/2}\widehat{P\Phi}\|_{L^2}^2 \\
&= \langle (1 + |\xi|^2)^{|k|/2}\widehat{P\Phi}, (1 + |\xi|^2)^{|k|/2}\widehat{P\Phi}\rangle \\
&= \langle (1 + |\xi|^2)^{|k|/2}\widehat{P\Phi}, \widehat{P\Phi}\rangle \\
&= \langle (1 - \Delta)^{|k|} P\Phi, P\Phi\rangle.
\end{aligned}$$

Now observe that

$$\begin{aligned}
(1 - \Delta)^{|k|} P\Phi &= P((1 - \Delta)^{|k|}\Phi) + [(1 - \Delta)^{|k|}, P]\Phi \\
&= P\phi + \underbrace{[(1 - \Delta)^{|k|}, P]}_{\text{order } 2|k| + 2 - 1}\Phi.
\end{aligned}$$

This tells us that

$$\begin{aligned}
\|\vec{\Phi}(t)\|_{\mathcal{H}^{1+|k|}} &= \|\vec{\phi}(t)\|_{\mathcal{H}^{1+|k|-2|k|}} \\
&= \|\widehat{\phi}(t)\|_{\mathcal{H}^{1+k}}
\end{aligned}$$

for $k < 0$. ∎

17 Local Well-Posedness of the Initial Value Problem for Variable-Coefficient Wave Equations

17.1 Recap: setting and statement of the estimate

We have been looking at linear hyperbolic PDEs $P\phi = f$, where

$$P\phi = \partial_\mu(g^{\mu,\nu}\partial_\nu\phi) + b^\mu\partial_\mu\phi + c\phi.$$

We want to solve the initial value problem

$$\begin{cases} P\phi = f \\ (\phi, \partial_t\phi)|_{t=0} = (g, h). \end{cases}$$

To discuss existence and uniqueness, we made further assumptions on the coefficients:

- $g^{\mu,\nu}$ is a symmetric $(1+d) \times (1+d)$ matrix with signature $(-,+,+,\dots,+)$.
- $g^{0,j}(t,x) = 0$ and $g^{0,0}(t,x) = -1$.
- For $\xi \in \mathbb{R}^d$, $g^{j,k}\xi_j\xi_k \geq \lambda|\xi|^2$ (bottom right $d \times d$ minor is positive definite).
- $g^{\mu,\nu}, b, c$ are uniformly bounded, with uniformly bounded derivatives.

Example 17.1. Set $b = c = 0$, and let $g = \operatorname{diag}(-1, 1, 1, \dots, 1)$. Then $P = \Box$.

We take the convention that $x^0 = t$. We also use Greek indices $\mu, \nu \in \{0, 1, \dots, d\}$ and indices $j, k \in \{1, \dots, d\}$. Last time, we were proving the following theorem.

Theorem 17.1 (Local well-posedness of the initial value problem)**.** *Let $s \in \mathbb{Z}_+$. Given $(g, h) \in H^{s+1} \times H^s(\mathbb{R}^d)$ and $f \in L^1_t([0,t]; H^s(\mathbb{R}^d))$, there exists a unique solution ϕ to the initial value problem with $\phi \in C_t([0,T], H^{s+1})$ and $\partial\phi \in C_t((0,T); H^s)$. Moreover, the unique solution ϕ satisfies the estimate*

$$\|\phi\|_{C_t([0,T];H^{s+1})} + \|\partial_t\phi\|_{C_t([0,T];H^s)} \lesssim_{g^{\mu,\nu},b^\mu,c,T,s} \|(g,h)\|_{H^{s+1}\times H^s} + \|f\|_{L^1_t([0,T];H^s)}.$$

Remark 17.1. Local well-posedness entails continuous dependence of ϕ on (f, g, h). Because of linearity, this a priori estimate implies continuous dependence (and in fact Lipschitz dependence).

17.2 Proof of the a priori estimate

Let's finish the proof. Recall that the idea of the proof is to use the a priori estimate, along with a functional analytic lemma.

Proposition 17.1. *Let $s \in \mathbb{Z}$. Let $\phi \in C_t([0,T];H^{s+1})$ and $\partial_t\phi \in C_t([0,T];H^s)$. Then*

$$\|\phi\|_{C_t([0,T];H^{s+1})} + \|\partial_t\phi\|_{C_t((0,t):H^s)} \lesssim \|(\phi,\partial_t\phi)|_{t=0}\|_{H^{s+1}\times H^s} + \|P\phi\|_{L^1_t([0,T];H^s)}.$$

Proof. ($s \geq 0$): We want to use the energy method. The natural strategy would be to commute $P\phi$ with D^α for $|\alpha| \leq s$ and apply the energy estimate (multiply by $\partial_t\phi$ and integrate by parts). Instead, we vary the multiplier:

$$\langle P\phi, (1-\Delta)^s\partial_t\phi\rangle := \int P\phi(1-\Delta)^s\partial_t\phi\,dx$$

- On one hand, we know by duality that

$$\int_0^T \langle P\phi, (1-\Delta)^s\partial_t\phi\rangle\,dt \lesssim \|P\phi\|_{L^1_t([0,T];H^s)}\|\partial_t\phi\|_{C_t([0,T];H^s)}.$$

  This is basically integrating by parts s times and using Cauchy-Schwarz. We can also think of this as the general bound

$$|\langle f,g\rangle| \lesssim \|f\|_{H^s}\|g\|_{H^{-s}}$$

  In general, if Q is an order r differential operator with that have uniformly bounded derivatives to all order, then (with some Fourier analysis), we can say that

$$\|Qg\|_{H^s} \lesssim \|g\|_{H^{r+s}} \qquad (s\in\mathbb{R}).$$

  For negative s, we get the inequality by duality:

$$\begin{aligned}\|Qf\|_{H^s} &= \sup_{\|g\|_{H^s}=1} ||\langle Qf,g\rangle| \\ &= \sup_{\|g\|_{H^s}=1} ||\langle f,Q^*g\rangle| \\ &\lesssim \|f\|_{H^{s+r}}\|Q^*g\|_{H^{s-r}}.\end{aligned}$$

  We also have the fact that

$$\|(1-\Delta^s)g\|_{L^2} \simeq \|g\|_{H^{2s}}, \langle (1-\Delta)^s g,g\rangle \simeq \|g\|^2_{H^s},$$

  which we get by using the Fourier transform:

$$\langle (1-\Delta)^s g,g\rangle = \langle (1+|\xi|^2)^s,\widehat{g},\widehat{g}\rangle = \|(1+\xi|^2)^{s/2}\widehat{g}\|^2_{L^2}$$

- On the other hand, we have

$$P\phi = \underbrace{\partial_\mu(g^{\mu,\nu}\partial_\nu\phi)}_{-\partial_t^2\phi+\partial_j(g^{j,k}\partial_k\phi)} + b^\mu\partial_\mu\phi + c\phi.$$

Now we can observe that

$$\langle -\partial_t^2\phi, (1-\Delta)^s\partial_t\phi\rangle = -\partial_t\langle\partial_t\phi, (1-\Delta)^s\partial_t\phi\rangle + \langle\partial_t\phi, (1-\Delta)^s\partial_t^2\phi\rangle$$

Since $\langle\partial_t\phi, (1-\Delta)^s\partial_t^2\phi\rangle = \langle(1-\Delta)^s\partial_t\phi, \partial_t^2\phi\rangle$, we get

$$= -\frac{1}{2}\partial_t\langle\partial_t\phi, (1-\Delta)^s\partial_t\phi\rangle$$

For the other term, we have

$$\begin{aligned}\langle\partial_j(g^{j,k}\partial_k\phi), (1-\Delta)^s\partial_t\phi\rangle &= -\langle g^{j,k}\partial_k\phi, (1-\Delta)^s\partial_t\partial_j\phi\rangle \\ &= -\partial_t\langle g^{j,k}\partial_k\phi, (1-\Delta)^s\partial_j, \phi\rangle \\ &\qquad + \langle\partial_t g^{j,k}\partial_k\phi, (1-\Delta)^s\partial_j, \phi\rangle \\ &\qquad + \langle g^{j,k}\partial_k\partial_t\phi, (1-\Delta)^s\partial_j\phi\rangle.\end{aligned}$$

Write the last term as

$$-\langle_t\phi, \partial_k(g^{j,k}(1-\Delta)^s\partial_j\phi)\rangle = -\langle\partial_t\phi\partial_k([g^{j,k}, (1-\Delta)^s]\partial_j\phi)\rangle \underbrace{-\langle\partial_t\phi, \partial_k(1-\Delta)^s(g^{j,k}\partial_j\phi)\rangle}_{=-\langle(1-\Delta)^s\partial_t\phi, \partial_k(g^{j,k}\partial_j\phi\rangle}.$$

Overall, this equals

$$-\frac{1}{2}\partial_t\langle g^{j,k}\partial_k\phi, (1-\Delta)^s\partial_j\phi\rangle + \frac{1}{2}\langle\partial_t g^{j,k}\partial_k\phi, (1-\Delta)^s\partial_j\phi\rangle - \frac{1}{2}\langle\partial_t\phi, \partial_k([g^{j,k}, (1-\Delta)^s]\partial_j\phi)\rangle.$$

The point is of this messy calculation is as follows: for the terms with the highest number of derivatives, we want to put things in to this total derivative form. The other terms will have at least 1 derivative that is not falling on ϕ. This is the purpose of using the commutator. What we get is that

$$\begin{aligned}&\langle P\phi, (1-\Delta)^s\partial_t\phi\rangle \\ &\quad = \underbrace{-\frac{1}{2}\partial_t(\langle\partial_t\phi, (1-\Delta)^s\partial_t\phi\rangle + \langle g^{j,k}\partial_k\phi, (1-\Delta)^s\partial_j\phi\rangle)}_{E_s[\phi](t)} \\ &\quad + \underbrace{O(\langle q_1\partial\phi, \partial^{2s}\partial\phi\rangle) + O(\langle q_2\partial\phi, \partial^{2s-1}\partial\phi\rangle) + \cdots + O(\langle q_{2s+1}\partial\phi, \partial\phi\rangle)}_{R_s},\end{aligned}$$

where $q_1 = \partial g, b$, $q_2 = \partial^2 g\partial bc$, etc.

So our energy argument says

$$\int_0^t \langle P\phi, (1-\Delta)^s\partial_t\phi\rangle\, dt' \geq E_s[\phi](0) - E_s[\phi](t) - C\int_0^t \|\phi\|_{H^{s+1}}^2 + \|\partial_t\phi\|_{H^s}^2\, dt',$$

where we are just using the estimate for the remainder:

$$|R_s(t') \lesssim (\|\phi\|_{H^{s+1}} + \|\partial_t\phi\|_{H^s})^2.$$

Now we have

$$E_s[\phi](t) \leq E_s[\phi](0) + \|P\phi\|_{L^1_t([0,T];H^s}\|\partial_t\phi\|_{C_t((0,T);H^s)} + \int_0^t \|\phi\|^2_{H^{s+1}} + \|\partial_t\phi\|^2_{H^s}\,dt'.$$

Note that $E_s[\phi](t) \simeq \|\phi\|^2_{H^{s+1}} + \|\partial_t\phi\|^2_{H^s}$, so our proprties of H^s and the elliptic estimate for $\partial_j g^{j,k}\partial_k$ gives:

$$E_s[\phi(t) \leq E_s[\phi](0) + \|P\phi\|_{L^1_t([0,T];H^s)}\|\partial_t\phi\|_{C_t((0,T);H^s)} + \int_0^t E_s[\phi(t')\,dt'$$

So Grönwall's inequality tells us that

$$E_s[\Phi](t) \lesssim E_s[\phi](0) + \|P\phi\|_{L^1_t([0,T];H^s)} \sup_{t\in[0,T]} E_s[\phi](t).$$

Now we can take the sup over $t \in [0,T]$ on the left hand side and use the AM-GM inequality with an epsilon to absorb the $\sup_{t\in[0,T]} E_s[\phi](t)$ on the right into the left hand side.

$(s < 0)$: Let $\Phi = (1-\Delta)^{-|s|}\phi$. We have the equivalence

$$\|\Phi\|_{H^{|s|+1}} \simeq \|\phi\|_{H^{-|s|+1}} = \|\phi\|_{H^{s+1}}.$$

Similarly,

$$\|\partial_t\Phi\|_{H^{|s|}} \simeq |\partial_t\|_{H^s}.$$

Now, we do the same argument with s replaced by $|s|$ and ϕ replaced by Φ. The only thing that is different is part 1 above. So we need to estimate

$$\begin{aligned}|\langle P\Phi, (1-\Delta)^{|s|}\partial_t\Phi\rangle| &= |\langle (1-\Delta)^{|s|}P\Phi, \partial_t\Phi\rangle| \\ &= |\langle P\underbrace{(1-\Delta)^{|s|}\Phi}_{\phi}, \partial_t\Phi\rangle| + |\langle[(1-\Delta)^{|s|}, P]\Phi, \partial_t\Phi\rangle|\end{aligned}$$

The right term has order $2|s|+2-1$. Using duality,

$$\lesssim \|P\phi\|_{H^s}\|\partial_t\Phi\|_{H^{|s|}}\|\Phi\|_{H^{|s|+1}}\|\partial_t\Phi\|_{H^{|s|}}.$$

This completes the proof. $\square$

17.3 Proof of well-posedness from the a priori estimate

Now we can quickly conclude the proof existence and uniqueness theorem.

Proof. Note that uniqueness and the a priori estimate follow from the proposition. It remains to prove existence.

Step 1: First, view this as trying to find the inverse of the operator $P : L_t^\infty([0,T], \mathcal{H}^{s+1}) \to L_t^1([0,T]; H^s)$. We want to reduce to the case when the initial data $g, h = 0$; we may achieve this using extension and modifying f.

Step 2: By duality, $\phi \in L_t^\infty([0,T]; H^{s+1}) = (L_t^1([0,T]; H^{-s-1}))^*$. We want

$$\begin{aligned} \int_0^T \langle f, \psi \rangle \, dt &= \int_0^T \langle P\phi, \psi \rangle \, dt \\ &= \int_0^T \langle \phi, P^*\psi \rangle \, dt. \end{aligned}$$

Define $\ell : P^*(L_t^1([0,T]; H^{-s})) \to \mathbb{R}$ by $\ell(P^*\psi) = \int_0^T \langle f, \psi \rangle \, dt$. This is well-defined by our a-priori estimate:

$$\|\ell\| \le \|f\|_{L^1(H^s)} \|\psi\|_{L^\infty(H^{-s})} \le \|f\|_{L^1(H^s)} \|P^*\psi\|_{L^1(H^{-s-1})}.$$

By Hahn-Banach, there exists an extension $\ell^* \in (L_t^1(H^{-s-1}))^*$ which is an extension with the bound $\|\ell^*\| \lesssim \|f\|_{L^1(H^s)}$. Here, $\phi = \ell^* \in L_t^\infty(H^{s+1})$.

Step 3: Upgrade $\phi \in L_t^\infty(H^{s+1})$ to $\phi \in C_t(H^{s+1})$ with $\partial_t \phi \in C_t(H^s)$. The way to do this is to approximate by smooth objects and try to take the limit. The a priori estimate will stay intact through the limit. $\square$

18 Definition of Hyperbolicity

18.1 Working definition of hyperbolicity

Let's return to our general discussion of hyperbolicity. We initially gave a working definition of hyperbolicity:

- Order of t derivatives = order of x-derivatives
- (Local) well-posedness of the initial value problem.

The purpose of the first condition is to ensure that we have a **finite speed of propagation**. This is as opposed to some other equations, where you may have compactly supported initial data, but immediately after $t = 0$, the solution is no longer compactly supported. The finite speed of propagation is very related to Lorentzian geometry.

We would also like to have an algebraic definition of hyperbolicity. Here, we will give the standard definition you may see in a paper or textbook. We gave the working definition first because there are some hyperbolic PDEs which are badly behaved (e.g. you can't prove local well-posedness of the initial value problem without extra assumptions). In what follows, compare with the PDE $P\phi = f$, where

$$P\phi = -\partial_t^2\phi + \partial_j(a^{k,j}\partial_k\phi) + b^j\partial_j\phi + c\phi.$$

This is a special case of hyperbolicity for second order linear PDEs.

18.2 Hyperbolicity for first-order systems

The solution we want will take the form of $\Phi : \mathbb{R}_t \times \mathbb{R}^d \to \mathbb{R}^n$, with the equation

$$\partial_t\Phi^J + (B^j)^J_K\partial_j\Phi^K.$$

We can express this in matrix notation as

$$\partial_t\Phi + B^j\partial_j\Phi = F.$$

Here, $B^j = (B^j)^J_K$ is an $n \times n$ matrix valued function on $\mathbb{R}_t \times \mathbb{R}^d$, and $F : \mathbb{R}_t \times \mathbb{R}^d \to \mathbb{R}^n$. The initial condition is $\Phi|_{t=0} = \Phi_0$.

Definition 18.1. Let the **symbol** of $B^j\partial_j$ be $\sigma(t, x; \xi) = \xi_j B^j(t, x)$, where $\xi \in \mathbb{R}^d$. We say that a first-order equation is **hyperbolic** if $\sigma(t, x; \xi)$ has n real eigenvalues for each t, x, ξ.

Theorem 18.1 (Constant coefficient case)**.** *Assume B^j is independent of (t, x). Then hyperbolicity implies local well-posedness of the initial value problem.*

Proof. The proof involves Fourier analysis, and you can find it in section 7.3 of Evans' textbook. □

If B^j is constant and σ has a non-real eigenvalue, then there exists a **plane wave solution**

$$\Phi = Ae^{i(x\cdot\xi+t\omega)},$$

where $\operatorname{Im}\omega \neq 0$. If $\operatorname{Im}\omega < 0$, this solution experiences exponential growth in time. This can be formalized into an ill-posedness statement. This should motivate our definition of hyperbolicity.

However, in the variable coefficient case, we need stronger conditions to ensure local well-posedness.

Definition 18.2. A first-order equation is **symmetric hyperbolic** if each $B^j(t,x)$ is symmetric for all t,x. If there exists a similarity transformation $P(t,x)$ sending $\Phi \mapsto \widetilde{\Phi} = P\Phi$ such that the transformed equation is symmetric hyperbolic, then the equation is called **symmetrizable hyperbolic**.

Theorem 18.2. *A symmetric hyperbolic first-order equation (with regularity assumptions on B) has local well-posedness of the initial value problem.*

Proof. This proof is by the energy method. You can find a proof in 7.3 in Evans' textbook, but the method we have presented in class is closer to the presentation in Chapter 7 of Ringström's book. □

Definition 18.3. A hyperbolic first-order system is said to be **strictly hyperbolic** if all n real eigenvalues are distinct (for all t,x,ξ)

$$\lambda_1(t,x;\xi) < \cdots < \lambda_n(t,x,;\xi).$$

This is a useful definition when the spatial dimension is $d=1$. In this case, these eigenvalue separation conditions help us use the method of characteristics to solve this system (normally you can only solve scalar equations in this way). This is not discussed in Evans' book, but it is discussed in *Hyperbolic Conservation Laws* by Dafemos.

18.3 Hyperbolicity for second-order, linear, scalar PDEs

Here, we give a notion of hyperbolicity that generalizes our wave equation $\Box\phi = 0$. We have

$$P\phi = \partial_\mu(g^{\mu,\nu}\partial_\nu\phi) + b^\mu\partial_\mu\phi + c\phi.$$

Let's focus on $g^{\mu,\nu}$, the important part. This brings us to the idea of Lorentzian (inverse) metrics.

Let $(g^{-1})^{\mu,\nu}(t,x)$ be a symmetric $(1+d)\times(1+d)$ matrix with signature $(-,+,+,\dots,+)$. (Compare this to if the case where $M = \operatorname{diag}(-1,1,1,\dots,1)$, so the wave equation can be written as $\partial_\mu(m^{\mu,\nu}\partial_\nu\phi) = 0$.) Let $g = g_{\mu,\nu}$.

Definition 18.4. A **Lorentzian manifold** is a pair $(\mathcal{M}, g)$, where $\mathcal{M}$ is a $(1+d)$-dimensional smooth manifold, and g is a symmetric, covariant 2-tensor with signature $(-,+,+,\dots,+)$.

The key difference from Riemannian geometry is due to the following.

Lemma 18.1. *Let $Q(\xi)$ be a quadratic form $q^{\alpha,\beta}\xi_\alpha\xi_\beta$. If q has no zero eigenvalue and has at least one negative and positive eigenvalues, then $\{\xi : Q(\xi) = 0\}$ determines q up to multiplication by a constant.*

The condition $g^{\alpha,\beta}\xi_\alpha\xi_\beta = 0$ determines $g^{\alpha,\beta}$ (up to a constant). This explains why Lorentzian geometry is a natural setting for Einstein's equations. If we have $g_{\alpha,\beta}v^\alpha v^\beta = 0$, the the zero set of P looks like a cone. This lemma tells you that these distinguished directions determine the behavior. In relativity, there are distinguished speeds, such as the speed of light. The tangent space at each point is made of velocity vectors. One way to think of this is that at each point, you get a cone, but you need some way to stitch these together; the Lorentzian metric is a natural way to do this.

Here is an algebraic lemma which connects the restricted class of PDEs that we have considered to this Lorentzian setting.

Lemma 18.2. *Let g be a Lorentzian metric, and let $p \in \mathcal{M}$. There exists a neighborhood $U \ni p$ and local coordinates $(x^o, \dots, x^d)$ in U such that $(g^{-1})^{0,j} = 0$ for all $j = 1, \dots, d$ and $(g^{-1})^{0,0} < -c$ for some $c > 0$. We may also ensure that $(g^{-1})^{j,k}\xi_j\xi_k \geq c_0|\xi|^2$ for some $c_0 > 0$.*

Corollary 18.1. *Locally,*

$$P\phi = \partial_\mu(g^{\mu,\nu}\partial_\nu\phi) + b^\mu\partial_\mu\phi + c\phi$$

can be put in the restricted form discussed earlier ($g^{0,j} = 0$, $g^{0,0} = -1$, and $g^{j,k}\xi_j\xi_k \geq \lambda|\xi|^2$). The condition $g^{0,0} = -1$ can be ensured by normalization at the level of the PDE.

Proof. Take x^0 such that $(g^{-1})^{\mu,\nu}(dx^0)_\mu(dx^0)_\nu < 0$; we say that such a dx^0 is **time-like**. We are looking for hypersurfaces that are transversal to the zero cones at each point. Take any local coordinates x^j near p in $\{x^0 = 0\}$. We want to transport x^j to other level surfaces of x^0 so that $(g^{-1})^{0,j} = 0$. Here is the procedure. Take a 1-form $(dx^0)_\mu$ and form a vector field $(g^{-1})^{\mu,\nu}(dx^0)_\nu = \nabla x^0$. If we write this in coordinates, this vector field is $(g^{-1})^{\mu,\nu}$.

We want to make sure that $(g^{-1})^{0,\mu}\partial_\mu x^j = 0$. So we construct this vector field ∇x^0 and then flow along the vector field. □

18.4 Geometric formulation of local well-posedness of the initial value problem

Here is the geometric idea concerning the initial time we start our initial conditions at. Just as in Riemannian geometry, we can create a Levi-Civita connection, which leads to parallel transport.

Definition 18.5. A C^1 curve γ is a **geodesic** if $\nabla_{\dot\gamma}\dot\gamma = 0$.

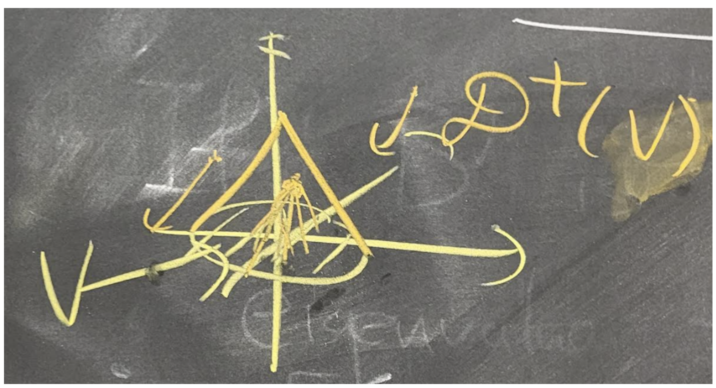
For a geodesic γ, $\frac{d}{dt} g(\dot\gamma, \dot\gamma) = 0$. The curve γ is called

$$\begin{cases} \text{timelike} & \text{if } g(\dot\gamma,\dot\gamma) < 0 \\ \text{null} & \text{if } g(\dot\gamma,\dot\gamma) < 0 \\ \text{spacelike} & \text{if } g(\dot\gamma,\dot\gamma) > 0. \end{cases}$$

Corollary 18.2. *If γ is a geodesic, there is a well-defined **causal** (i.e. timelike or null) character.*

Definition 18.6. A Lorentzian manifold $(\mathcal{M}, g)$ is **time-orientable** if there exists a non-vanishing vector field that is timelike everywhere.

Definition 18.7. Let $\mathcal{M}$ be a Lorentzian manifold, and let $V \subseteq \mathcal{M}$. The **causal future** of V is $\mathcal{J}^+(V) = \{\text{all } q \in \mathcal{M} \text{ with a future causal curve from } p \in V \text{ to } q\}$. We also let $\mathcal{D}^+(V) = \{q : J^-(q) : \text{ all causal past-pointing curves from } q \text{ meet } V\}$.



Here, future means the top half of the cone.

Definition 18.8. A **Cauchy hypersurface** is a spacelike hypersurface; i.e. a hyper surface with all tangent vectors spacelike

In this picture,
$$\mathcal{M} = \mathcal{D}^+(\Sigma) \cup \Sigma \cup \mathcal{D}^-(\Sigma).$$

Definition 18.9. Global hyperbolicity is when $(\mathcal{M}, g)$ is time orientable and there exists a Cauchy hypersurface.

Theorem 18.3. *Let $(\mathcal{M}, g)$ be globally hyperbolic with a Cauchy hypersurface Σ. Then*

$$\begin{cases} \Box_g \phi + B\phi + c\phi = f \\ (\phi, n_\Sigma \phi)|_\Sigma = (g, h) \end{cases}$$

is well-posed (existence and uniqueness). here. B is a vector field, c is a function, n_Σ is the unit normal to Σ, and

$$\Box_g \phi = \mathrm{div}_g(d\phi) = \frac{1}{\sqrt{|d+g|}} \partial_\mu ((g^{-1})^{\mu,\nu} \sqrt{|d+g|} \partial_\nu \phi).$$

The converse is also true.

What is interesting is there is a purely geometric formulation of this. A good reference for this story is Chapters 10 to 12 of Ringström's book.

19 Decay by Dispersion for the Wave Equation

19.1 Motivation: the picture of decay by dispersion for the wave equation

Consider the wave equation in $\mathbb{R}^{1+d}$,

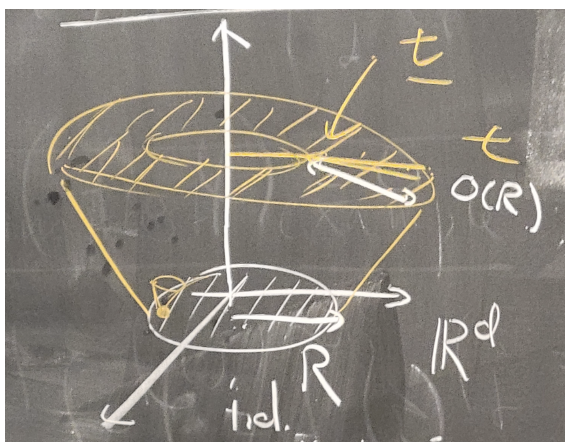
$$\Box\phi = 0, \qquad \Box = -\partial_t^2 + \Delta.$$

We know the conservation of energy:

$$\int ((\partial_t\phi)^2 + |D\phi|^2)|_{t=t_1}\, dx = \int ((\partial_t\phi)^2 + |D\phi|^2)|_{t=0}\, dx \qquad \forall t_1 \in \mathbb{R}.$$

In some sense, the size of ϕ stays constant. Since we are in an infinite dimensional space of functions, we may have a notion of size where the function stays the same in time and another notion of size where the function goes to 0 in time.

Dispersion is a *decay* mechanism for $\Box\phi = 0$ in $\mathbb{R}^{1+d}$, where the amplitude "$|\phi|(t) \to 0$ as $t \to \pm\infty$. " This pointwise estimate will not always hold, but this is the idea. Assume that the initial data is well-localized (compactly supported or at least has strong decay). The solution should try to propagate in every direction (forming a cone in $\mathbb{R}^{1+d}$). At time t, most of the solution should have spread away around $O(R)$.



The quantity $\int |\partial\phi|^2\, dx$ is conserved. We normalize this so that $\int |\partial\phi|^2|_{t=0}\, dx = 1$ and $R = 1$. At time T, $\int |\partial\phi|^2|_{t=T}\, dx = 1$, and ϕ is supported (evenly) in $\{x : t < |x| < t+1\}$. This means that

$$1 = \int |\partial\phi|^2|_{t=T}\, dx \approx a^2 t^{d-1},$$

where a is the amplitude of ϕ at time t. This means that

$$a \approx \frac{1}{t^{(d-1)/2}}.$$

This is also the decay rate for $\|\partial\phi\|_{L^\infty}(t)$ and also for $\|\phi\|_{L^\infty}(t)$. The intuition for the latter statement is that if we know that ϕ is small near 0 at time t, then we can just integrate radially in constant time; this does not give so much up because $R = 1$.

Our goal is to make this decay precise. We will aim to give two proofs of this fact:

1. Using oscillatory integrals: This generalizes to constant foefficient dispersive PDEs such as $(\Box - m^2)\phi = f$ or $(\partial_t + \Delta)\phi = f$.

2. Vector field method: This generalizes better to variable coefficent PDEs and nonlinear PDEs.

19.2 Oscillatory integrals in the solution to the wave equation

The starting point is the solution formula for $\Box\phi = f$ using the Fourier transform. Take the spatial Fourier transform of the equation, $(-\partial_t^2 + \Delta)\phi = f$, which means $\widehat{\phi}(t,\xi) = \mathcal{F}_x[\phi(t,\cdot)]$. This gives

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

In view of Duhamel's formula, it suffices to consider $f = 0$. So we now have $\partial_t^2\widehat{\phi} = -|\xi|^2\widehat{\phi}$. This has the solution $e^{\pm it|\xi|}$, so

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

where a_+, a_- are determined from the initial conditions at $t = 0$. We can then write

$$\phi(t,x) = \frac{1}{(2\pi)^d}\int a_+(\xi)e^{it|\xi||}e^{ix\cdot\xi}\,d\xi + \frac{1}{(2\pi)^d}\int a_-(\xi)e^{-it|\xi||}e^{ix\cdot\xi}\,d\xi.$$

These integrals are essentially the same, so we will concentrate on the + case. Our goal is to analyze the asymptotics of this integral in (t,x).

19.3 General theory for oscillatory integrals

19.3.1 Principle of nonstationary phase

We now take an intermission to study some model oscillatory integrals.

Definition 19.1. An **oscillatory integral** is an integral of the form

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

Here $a : \mathbb{R} \to \mathbb{C}$ is a the **amplitude function**, and $\Phi : \mathbb{R} \to \mathbb{R}$ is called the **phase function**. We assume a and ξ to be small, i.e. $|D^\alpha|$, $|D^\alpha\Phi| \lesssim_\alpha 1$. We also assume that $\operatorname{supp} a$ is compact.

Proposition 19.1 (Principle of nonstationary phase)**.** *If* $|\partial_\xi \Phi| \geq \eta$ *on* $\operatorname{supp} a$*, then*

$$I(\lambda)| \lesssim_k \frac{1}{\lambda^k}.$$

The idea is that the oscillations will make a lot of cancelation, so the size of the integral will be much smaller than if we just integrated a.

Proof. Use integration by parts; the key identity that drives this is $\partial_\xi(e^{i\lambda\Phi(\xi)}) = i\lambda\partial_\xi\Phi(\xi)e^{i\lambda\Phi}$, which gives $e^{i\lambda\Phi} = \frac{1}{i\lambda\partial_\xi\Phi}\partial_\xi(e^{i\lambda\Phi})$. This gives the identity

$$\begin{aligned} I(\lambda) &= \int a(\xi)\frac{1}{i\lambda\partial_\xi\Phi(\xi)}\partial_\xi e^{i\lambda\Phi}\, d\xi \\ &= -\int \partial_\xi\left(a(\xi)\frac{1}{i\lambda\partial_\xi\Phi(\xi)}\right)e^{i\lambda\Phi}\, d\xi. \end{aligned}$$

This is good, as long as $|\partial_\xi\Phi| \geq \eta_0$. The derivative part is $\partial_\xi a\frac{1}{i\lambda\partial_\xi\Phi} - a\frac{1}{i\lambda\partial_\xi\Phi}\frac{\partial_\xi^2\Phi}{\partial_\xi\Phi}$.

$$\lesssim \frac{1}{\lambda}.$$

Integrating by parts k times gives $|I(\lambda)| \leq \cdots \lesssim \frac{1}{\lambda^k}$. □

19.3.2 Principle of stationary phase

In the presence of a critical point ξ_0 of Φ (i.e. $\partial_\xi\Phi(\xi_0) = 0$), we have the **principle of stationary phase** Consider

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

with $\partial_\xi\Phi(0) = 0$ and no other zeros in the support of a. (The general case can be reduced to this by a smooth partition of unity.) In view of Taylor expansion, we would expect that

$$\Phi(\xi) = \Phi_0 + c\xi^n + \cdots, \qquad \text{where } n \geq 2.$$

We can absorb $e^{i\lambda\Phi_0}$ into a, so we may assume that $\Phi_0 = 0$. So our model case is when $\Phi(\xi) = c_n\xi^n$. Here,

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

The principle is that the **stationary phase region** $\{\xi \in \mathbb{R}^n : |\lambda\xi^n| \leq 1\}$ gives you the main contribution:

$$I(\lambda) \sim \int_{\{|\lambda\xi^n|<1\}} d\xi \approx a(0)C\frac{1}{\lambda^{1/n}},$$

where $\frac{1}{\lambda^{1/n}}$ is the volume of the region $\{|\xi| \leq 1/\lambda^{1/n}\}$.

How do we make this precise? Here are two approaches:

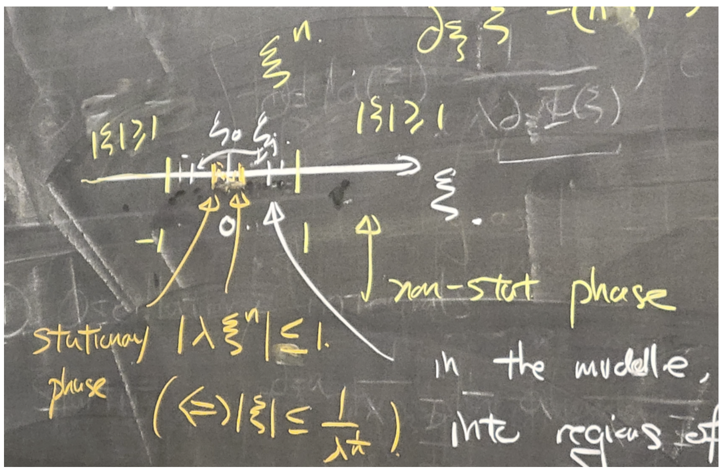
1. Precise algebraic manipulation.

    (a) Change of variables

    (b) Use the Fourier transform of the Gaussian ($n = 2$).

2. (Less precise but more robust) Dyadic decomposition.

We will present the latter approach. In the two regions with $|\xi| \geq 1$, we can use the principle of non-stationary phase. On the other hand, we have stationary phase in the region very close to 0, where $|\xi| \leq 1/\lambda^{1/n}$.



The idea is that in the middle, we can decompose into regions of the form $A_j\{2^{j-1} \leq |\lambda\xi^n| \leq 2^j\}$ for $j \geq 1$. In each of these regions $\lambda\xi^n$ is roughly constant. In particular, we introduce a smooth partition of unity $\{\zeta_j\}_j$ subordinate to $\{A_j\}_j$, and use two estimates for the integral

$$I_j \int \zeta_j(\xi) a(\xi) e^{i\lambda\Phi}\, d\xi,$$

Here, we can use the estimate of stationary phase (ignore $e^{i\cdots}$) and the estimate of nonstationary phase.

Here are the details. Let ζ_0 be adapted to $\{|\lambda\xi^n| \leq 1$. We have

$$I(\lambda) = \underbrace{\int \xi_0 a e^{i\lambda\xi^n}\, d\xi}_{I_0} + I_j.$$

Then

$$|I_0| \lesssim \frac{1}{\lambda^{1/n}}.$$

For I_j, if we do not integrate by parts, we get

$$|I_j| \lesssim |A_j| \lesssim (2^{1/n})^j \lambda^{-1/n},$$

where we have used $A_j \subseteq \{|\xi| \leq 2^{j/n}/\lambda^{1/n}\}$. This is nice when j is small but bad when j is big. If we use integration by parts, we get

$$|I_j| = \int \left| \partial_\xi \left(\zeta_j(\xi) a(\xi) \frac{1}{i\lambda \partial_\xi \Phi} \right) e^{i\lambda\Phi} \right| \, d\xi.$$

Note that $|\lambda \partial_\xi \Phi| \simeq e^j \frac{1}{|\xi|} \simeq 2^{(1-1/n)j} \lambda^{1/n}$. We also have $|a| + |\partial_\xi a| \lesssim 1$,

$$\left| \frac{\partial_\xi^2 \Phi}{\partial_\xi \Phi} \right| \lesssim \frac{1}{|\xi|} \approx 2^{-j/n} \lambda^{1/n}, \qquad |\partial_\xi \zeta_j| \lesssim \frac{1}{|\xi|} \lesssim 2^{-1/n} \lambda^{1/n}.$$

So

$$\left| \zeta_j(\xi) a(\xi) \frac{1}{i\lambda \partial_\xi \Phi(\xi)} \right| \sim 2^{(-1+1/n)j} \lambda^{-1/n},$$

$$\begin{aligned} |\partial_\xi(\cdots)| &\lesssim \frac{1}{|\xi|} 2^{(-1+1/n)j} \lambda^{-1/n} \\ &\lesssim 2^{-j/n} \lambda^{1/n} 2^{(-1+1/n)j} \lambda^{-1/n} \\ &= 2^{-j}. \end{aligned}$$

So

$$\int || \, d\xi| \lesssim 2^{-j} 2^{j/n} \lambda^{-1/n} = 2^{-(1-1/n)j} \lambda^{-1/n}.$$

Putting the these bounds for each $|I_j|$ together, we get

$$\left| \sum_{g \geq 1} I_j \right| \lesssim \sum_{j \geq 1} 2^{-(1-1/n)j} \lambda^{-1/n} \lesssim \lambda^{-1/n}.$$

Remark 19.1. We did not use both our bounds at the end. This is because we picked our dyadic decomposition in a smart way. If we had picked $\widetilde{A}_j = \{|\xi| \simeq 2^j\}$, then we would still be able to proceed with the proof, but we would need both the integration by parts and non-IBP bound.

20 Proof of Decay by Dispersion for the Wave Equation

20.1 Oscillatory integrals and the dispersive inequality for the wave equation

Last time, we were studying decay by dispersion for the wave equation, $\Box\phi = 0$. We saw that, using the Fourier transform, we could write the solution to the equation as

$$\phi(t,x) = \int a_+(\xi)e^{i(t|\xi|+x\cdot\xi)}\,d\xi + \int a_-(\xi)e^{i(-t|\xi|+x\cdot\xi)}\,d\xi.$$

The hope is that studying these integrals will allow us to prove our heuristically derived rate of dispersion for ϕ:

$$|\phi(t,x)| \lesssim \frac{1}{t^{(d-1)/2}}.$$

We studied the model oscillatory integral

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

and proved two general principles:

Theorem 20.1 (Principle of non-stationary phase)**.** *If* $\operatorname{supp} a \subseteq \{|\partial_\xi\Phi| \geq \eta\}$*, then*

$$\left|\int ae^{i\lambda\Phi}\,d\xi\right| \lesssim_{k,\eta} \frac{1}{\lambda^k}$$

for all $k \geq 0$*.*

Theorem 20.2 (Principle of stationary phase)**.** *Suppose there exists one critical point of* Φ *(i.e. one zero of* $\partial_\xi\Phi$*) in* $\operatorname{supp} a$*. Then*

$$\left|\int ae^{i\lambda\Phi}\,d\xi\right| \lesssim_{\eta'} \operatorname{vol}(\{|\lambda\Phi| \leq \eta'\}).$$

In particular, if $\Phi = \xi^n$ *for* $n \geq 2$*, then*

$$|I(\lambda)| \lesssim \operatorname{vol}(\{\lambda|\xi|^n| \leq 1\}) \simeq \lambda^{1/n}.$$

Our justification for the principle of stationary phase was to use the dyadic decomposition. We chose this method because it is robust.

Now, let us return to the wave equation. Let's study

$$I(t,x) = \int a_+(\xi) e^{i(t|\xi|+x\cdot\xi)}\, d\xi,$$

where a_+ is the amplitude and $t|\xi| + x\cdot\xi$ is the phase.

Definition 20.1. Let the **Besov norm** be defined as

$$\|f\|_{B^{s,p}_r} := \left(\sum_k (2^{sk}\|P_k f\|_{L^p})^r\right)^{1/r},$$

where P_k is the Littlewood-Paley projection

$$\widehat{P}_k f = \chi_0(\xi/2^k)\widehat{f}(\xi)$$

with $\operatorname{supp}\chi_0(\cdot/2^k) \subseteq \{|\xi| \simeq 2^k\}$ and $\sum_{k=-\infty}^{\infty}\chi_0(\xi/2^k) = 1$ for $\xi \neq 0$.

Theorem 20.3 (Dispersive inequality for the wave equation)**.** *Consider a solution ϕ to the wave equation*

$$\begin{cases} \Box\phi = 0 \\ (\phi, \partial_t\phi|_{t=0}(g,h). \end{cases}$$

Then

$$\|\phi(t,x)\|_{L^\infty_x} \lesssim t^{-(d-1)/2}(\|g\|_{B_1^{\frac{d+1}{2},1}} + \|h\|_{B_1^{\frac{d-1}{2},1}}).$$

The $\frac{d+1}{2}, \frac{d-1}{2}$ can be determined by dimensional analysis.

20.2 Reduction to an oscillatory integral with projected amplitude

In general, if we want $L^1 \to L^\infty$-type bounds, it usually suffices to just consider fundamental solutions; the idea is that any L^1 data can be split into delta distributions by convolution. A fundamental solution E_+ to the wave equation (with initial data $g = 0$ and $\phi = E_+ * h$) is

$$\begin{cases} \Box E_+ = 0 & t > 0 \\ (E_t, \partial_t E_t)|_{t=0} = (0, \delta_0). \end{cases}$$

In Fourier space, the initial data looks like

$$(\widehat{E}_+, \partial_t\widehat{E}_+)|_{t=0} = (0,1).$$

the constant 1 function has non-compact support, so we want to use a cutoff.

Instead, think of $P_k E_+$, which is the solution to the equation with initial data

$$(\widehat{P_k E_+}, \partial_t\widehat{P_k E_+})|_{t=0} = (0, \chi_0(\xi/2^k)),$$

which will give us a nicer oscillatory integral. This will be enough because we can decompose

$$\begin{aligned}\phi &= E_+ * h \\ &= \sum_k ((P_k E_+) * h\delta_{t=0}) \\ &= \sum_k (\widetilde{P}_k P_k E_+) * h\delta_{t=0},\end{aligned}$$

where $\widetilde{P}_k$ has the same properties as P_k but with $\widetilde{P}_k P_k = 1$. (We saw this in our study of Schauder theory.)

$$= \sum_k (P_k E_+ * \widetilde{P}_k h\delta_{t=0}).$$

We claim that it suffices to prove that

$$\|P_k E_+\|_{L^\infty_x} \lesssim t^{-\frac{d-1}{2}} \|\overset{\vee}{\chi}_0(\cdot/2^k)\|_{L^1} 2^{k\frac{d-1}{2}}.$$

Proof. If this bound holds, then

$$\begin{aligned}\|\phi(t,x)\|_{L^\infty} &= \left\| \sum_k \int P_k E_+(t, x-y)\widetilde{P}_k h(y)\, dy \right\|_{L^\infty} \\ &\lesssim \sum_k \int \|P_k E_+(t,x-y)\|_{L^\infty} |\widetilde{P}_k h(y)|\, dy \\ &\lesssim t^{-\frac{d-1}{2}} \sum_k 2^{k\frac{d-1}{2}} \|\widetilde{P}_k h\|_{L^1}.\end{aligned}$$ □

We now claim that it suffices to take $k = 0$. This is because out bound is invariant under the scaling $(t,x) \mapsto (\lambda t, \lambda x)$. This means that we only need to prove

$$\|P_0 E_+\|_{L^\infty} \lesssim t^{-\frac{d-1}{2}}.$$

$P_0 E_+$ is an oscillatory integral of the form

$$P_0 E_+ = \int a_+(\xi) e^{i(t|\xi| + x\cdot\xi)}\, dx + \int a_-(\xi) e^{i(-t|\xi| + x\cdot\xi)},$$

where $a_\pm$ have support in $\{|\xi| \simeq 1\}$ and obey $|D^\alpha a_\pm| \lesssim_\alpha 1$.

Hence, it suffices to consider

$$I(t,x) = \int a_+(\xi) e^{i(t|\xi| + x\cdot\xi)}\, dx \underbrace{\lesssim t^{-\frac{d-1}{2}}}_{\text{want}}$$

with $\operatorname{supp} a_+ \subseteq \{|\xi| \simeq 1\}$ and $|D^\alpha a_+| \lesssim_\alpha 1$.

20.3 Estimating the size of the oscillatory integrals

To estimate the size of this oscillatory integral, we look at the critical points of the phase

$$\Phi = t|\xi| + x \cdot \xi.$$

When is $\nabla\Phi = 0$? Observe that we have the identity $\partial_{\xi_j} e^{i\Phi} i\partial_{\xi_j}\Phi e^{i\Phi}$, so

$$e^{i\Phi} = \frac{1}{i\partial_{\xi_j}\Phi} e^{i\Phi}.$$

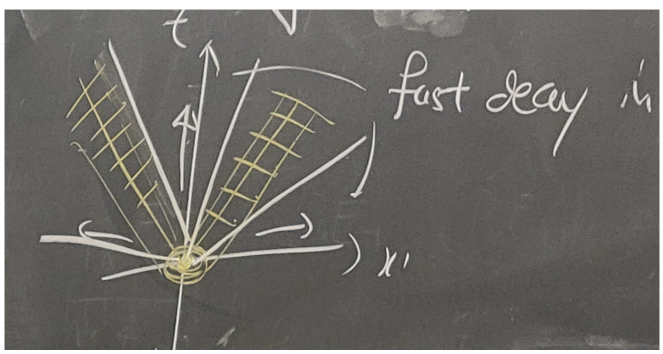
We may assume, by rotation in x-space that x is parallel to the vector e_1. Then

$$\Phi = t|\xi| + x^1\xi_1, \qquad \partial_{\xi_1} = \frac{t\xi_1}{|\xi|} + x^1, \qquad \partial_{\xi_j}\Phi = t\frac{\xi_j}{|\xi|}.$$

for $j \neq 1$. Then $\xi_j = 0$ for $j \neq 1$ occurs when $\frac{\xi_1}{|\xi|} = -\frac{x^1}{t}$. But $\xi_j = 0$ for $j \neq 1$ implies that $|\xi_1| = |\xi|$, so

$$\{\nabla\Phi = 0\} = \begin{cases} \varnothing & \text{if } |\frac{x^1}{t}| \neq 1 \\ \{s(-\frac{x^1}{t}, 0, \dots, 0) : s > 0\} & \text{if } |\frac{x^1}{t}| = 1. \end{cases}$$

If $|\frac{x^1}{t}| > c$ or $|\frac{x^1}{t}| < \frac{1}{c}$, then we the principle of non-stationary phase should apply, and we should be able to get $\frac{1}{\max(|t|,|x|)^k}$. The fundamental solution is a cone, and we smoothed it out with the projection. This says that ifwe look at a cone inside or outside this original cone, we get fast decay in t and $|x|$.



Assume that $|\frac{x^1}{t}| \simeq 1$. We need to look at the domain of Φ near the critical points

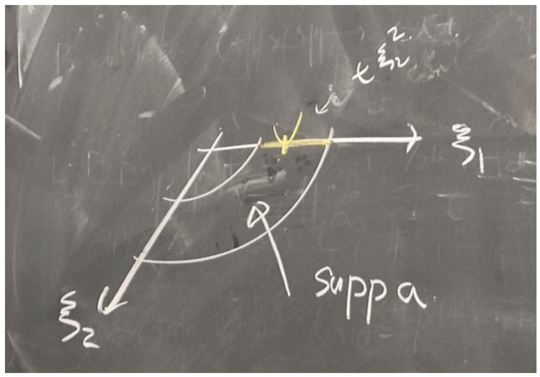
$$\partial_k\Phi = t\frac{\xi_k}{|\xi|} + x^1\delta 1, k,$$

$$\partial_{\xi_j}\partial_{\xi_k}\Phi = -t\frac{\xi_j\xi_k}{|\xi|^3} + t\frac{\delta_{j,k}}{|\xi|} = t\frac{|\xi|^2\delta_{j,k} - \xi_j\xi_k}{|x|^3}.$$

At a critical point, $\xi = (-s\frac{x^1}{t}, 0, \ldots, 0)$,

$$\nabla^2 \Phi = \begin{bmatrix} 0 & 0 \\ 0 & \frac{t}{|\xi|} I \end{bmatrix}.$$

And remember that on the support of a, $|\xi| \simeq 1$. Here is the picture:



So the region of stationary phase is $\{|t(\xi_2^2 + \cdots \xi_d^2)| \lesssim 1\}$, and

$$\mathrm{vol}_{\xi_2,\ldots,\xi_d}(\{|t(\xi_2^2 + \cdots \xi_d^2)| \lesssim 1\}) \lesssim t^{-\frac{d-1}{2}}.$$

By the principle of stationary phase, $t^{-\frac{d-1}{2}}$ dictates the size of $I(t,x)$.

The actual result can be proven via dyadic decomposition into regions of the form $\{t|\xi'|^2 \simeq \alpha\}_{\alpha=2^0,2^1,\ldots}$, where $\xi' = (\xi_2, \ldots, \xi_d)$.

21 The Vector Field Method for Dispersive Decay for the Wave Equation

21.1 Motivation for the vector field method

Today, we will continue discussing dispersive decay for the wave equation

$$\begin{cases} \Box\phi = 0 & \text{in } \mathbb{R}^{1+d} \\ (\phi, \partial_t\phi)|_{t=0} = (g, h). \end{cases}$$

Last time, we applied oscillatory integral techniques to the Fourier-analytic representation of the (frequency localized) fundamental solution. This led to the following dispersive inequality.

Theorem 21.1 (Dispersive inequality)**.** *For a solution ϕ to the wave equation,*

$$\|\phi(t)\|_{L^\infty} \lesssim t^{-\frac{d-1}{2}}(\|g\|_{B_1^{\frac{d+1}{2},1}} + \|h\|_{B_1^{\frac{d-1}{2},1}}).$$

This is the starting point for many estimates that are useful for **semilinear wave equations**, equations of the form $\Box\phi = N(\phi, \nabla\phi)$ with principal term $= \Box\phi$. But many equations of interest may have quasilinear nonlinearity ($g^{\mu,\nu}(\phi, \nabla\phi)\partial_\mu\partial_\nu\phi$) or just $g^{\mu,\nu}(t,x) \neq m^{\mu,\nu}$, for which the previous approach is harder to generalize.

Today, we will cover the vector field method, introduced by Klaineman in the 80s. This is a purely physical space method (as opposed to the Fourier analytic method above).[7] The motivating question is: How do we derive pointwise bounds for $\nabla_{t,x}\phi$ from the energy method?

Step 1: The energy estimate tells us that

$$E[\phi](t) = \int \frac{1}{2}(\partial_t\phi)^2 + \frac{1}{2}|D\phi|^2\,dx$$

is conserved. We can express this as a bound

$$\|\nabla_{t,x}\phi(t)\|_{L^2} \lesssim \|\nabla_{t,x}\phi|_{t=0}\|_{L^2}.$$

Step 2: Notice that if $\Box\phi = 0$, then any derivative satisfies

$$\Box(\partial_\mu\phi) = \partial_\mu\Box\phi = 0.$$

The energy estimate for the derivative then tells us that

$$\|\nabla_{t,x}D^\alpha\phi(t)\|_{L^2} \lesssim_\alpha \|\nabla_{t,x}D^\alpha\phi|_{t=0}\|_{L^2}.$$

[7] In general, Fourier analytic methods work best for constant coefficient, linear equations because when multiplication is involved, it becomes convolution, which can get messy.

Step 3: For $s > \frac{d}{2}$, we can use the Sobolev inequality to get

$$\begin{aligned}\|\nabla_{t,x}\phi(t)\|_{L^\infty_x} &\lesssim \|\nabla_{t,x}\phi(t)\|_{H^s_x}\\ &\lesssim \|\nabla_{t,x}\phi|_{t=0}\|_{H^s_x}.\end{aligned}$$

The basic idea of Klaineman's method was to follow this format to prove dispersive decay. The goal is to derive a pointwise estimate of the form

$$\|\nabla_{t,x}\phi(t)\|_{L^\infty_x} \lesssim t^{-\frac{d-1}{2}}(\text{Initial data}).$$

What parts of the approach should we modify? The first key idea is to modify step 2 of the argument above. The key property is that if $[\partial_\mu, \Box] = 0$, then $\Box\phi = 0$ implies $\Box\partial_\mu\phi = 0$. Thinking about this more geometrically, consider the translation operator $\phi \mapsto \mathcal{T}_{x^\mu,h}\phi = \phi((t,x) + he_\mu)$, where

$$\partial_\mu\phi = \frac{d}{dh}\mathcal{T}_{c^\mu,h}\phi|_{h=0},$$

so $\mathcal{T}_{c^\mu,h}$ is the infinitesimal generator for ∂_μ. The important thing to notice is that $\mathcal{T}_{x^\mu,h}$ is a symmetry for $\Box$:

$$\Box\mathcal{T}_{x^\mu,h}\phi = \mathcal{T}_{x^\mu,h}\Box\phi.$$

This process can be applied to any symmetries of $\Box$!

21.2 Symmetries of the d'Alembertian

Recall that the symmetries of $\Box$ are the linear symmetries $\mathbb{R}^{1+d} \to \mathbb{R}^{1+d}$ that preserve $m(v,w) = m^{\mu,\nu}v_\mu w_\nu$, where $m = \operatorname{diag}(-1,1,1,\dots,1)$. This means we want to look for matrices L_t such that

$$m^{\mu,\nu}(L_t)_\mu^{\mu'}(L_t)_\nu^{\nu'} = m^{\mu,\nu}.$$

If we assume that $L_0 = I$, then differentiating in t gives (denoting $\ell = \frac{d}{dt}L_t|_{t=0}$)

$$m^{\mu,\nu'}\ell_\mu^{\mu'} + m^{\mu',\nu}\ell_\nu^{\nu'} = 0.$$

If we define $\widetilde{\ell}^{\mu,\nu} = m^{\mu,\nu'}\ell_\mu^{\mu'}$, then we get $\widetilde{\ell}^{\nu',\mu'} + \widetilde{\ell}^{\mu',\nu'} = 0$.

The symmetries of $\Box$ turn out to be compositions of the following:

- Translations $\mathcal{T}_{x^\mu,h}$
- Rotations

$$\mathcal{R}_{x^1,x^2,h} = \begin{bmatrix} 1 & 0 & 0 & 0\\ 0 & \cos h & -\sin h & 0\\ 0 & \sin h & \cos h & 0\\ 0 & 0 & 0 & I\end{bmatrix}$$

- Lorentz boosts

$$\mathcal{L}_{x^1,h} = \begin{bmatrix} \frac{1}{\sqrt{1-h^2}} & -\frac{h}{\sqrt{1-h^2}} & 0 \\ -\frac{h}{\sqrt{1-h^2}} & \frac{1}{\sqrt{1-h^2}} & 0 \\ 0 & 0 & I \end{bmatrix}$$

The infinitesimal generators (meaning operators $\frac{d}{dh}\mathcal{S}_h(\cdot)|_{h=0}$ are

$$\Omega_{1,2} = x^1\partial_{x^2} - x^2\partial_{x^1}.$$

$$L_1 = x^1\partial_t + t\partial x^1$$

and the corresponding generators for the other indices. Observe that

$$[\partial_\mu, \Box] = [\Omega_{j,k}, \Box] = [L_j, \Box] = 0.$$

Also consider the scaling operator

$$\mathcal{S}_h\phi = \phi(t/h, t/x).$$

with intinitesimal generator

$$S\phi = -\frac{d}{dh}\mathcal{S}_h\phi|_{h=1} = (t\partial_t + x\partial_x)\phi.$$

This is not a symmetry of $\Box$, because

$$\begin{aligned} \Box\mathcal{S}_h\phi &= \Box\phi(t/h, x/h) \\ &= \frac{1}{h^2}(\Box\phi)(t/h, x/h) \\ &= \frac{1}{h^2}\mathcal{S}_h(\Box\phi). \end{aligned}$$

However, if $\Box\phi = 0$, then $\Box\mathcal{S}_h\phi = 0$. This is a reflection of the fact that

$$S\Box = S\Box - 2\Box,$$

where the -2 represents the homogeneity of $\Box$, a second order operator.

For $\Gamma \in \{\partial_0, \ldots, \partial_d, \Omega_{1,2}, \ldots, \Omega_{(d-1),d}, L_1, \ldots, L_d, S\}$, labeled in order as $\Gamma_1, \Gamma_2, \ldots, \Gamma_K$, we let

$$\Gamma^\alpha\phi = \Gamma_1^{\alpha_1}\cdots\Gamma_K^{\alpha_K}\phi, \qquad \alpha \in \mathbb{R}^K.$$

21.3 Bounds on commuting symmetries with derivatives

Our discussion has told the the following:

Lemma 21.1. *If* $\Box\phi = 0$, *then* $\Box\Gamma^\alpha\phi = 0$ *for all* α.

The energy estimate gives the following.

Corollary 21.1.

$$\|\nabla_{t,x}\Gamma^\alpha\phi(t)\|_{L^2} \lesssim_\alpha \|\nabla_{t,x}\Gamma^\alpha\phi|_{t=0}\|_{L^2}.$$

Lemma 21.2. *Given any smooth function* ψ,

$$|\Gamma^\alpha\nabla_{t,x}\psi| \lesssim \sum_{\beta:|\beta|\le|\alpha|} |\nabla_{t,x}\Gamma^\beta\psi|.$$

Here is the proof of the lemma:

Proof. When $\Gamma \in \partial_0, \dots, \partial_\mu$, there is nothing to do. When $\Gamma \in \{\Omega, L, S\}$, then $[\Gamma, \partial_{x^\mu}] = c^\nu_{\mu,\Gamma}\partial_{x^\nu}$; we can argue this by checking the generators or by claiming that these vector fields form a Lie algebra, so we get information about the Lie bracket. We complete the argument by induction. $\square$

Corollary 21.2. *Fix* s.

$$\sum_{\alpha:|\alpha|\le s} \|\Gamma^\alpha\nabla_{t,x}\psi(t)\|_{L^2} \lesssim \sum_{\alpha:|\alpha|\le s} \|\nabla_{t,x}\Gamma^\alpha\phi|_{t=0}\|_{L^2}.$$

21.4 The Klaineman-Sobolev inequality and proof of the dispersive estimate

The second key idea is to modify step 3, where we used the Sobolev inequality. We first need to understand what control Γ gives us.

Define $\Omega_{\mu,\nu} = x_\mu\partial_\nu - x_\nu\partial_\mu$, where

$$x_\mu = x^\nu m_{\mu,\nu} = \begin{cases} -t & \mu = 0 \\ x^j & m = h \in \{1, \dots, d\}. \end{cases}$$

If we have $\Omega_{j,k}$ as before, then $L_j = \Omega_{j,0}$.

Lemma 21.3.

$$(t^2 - |x|^2)\partial_\mu = x_\mu S - x^\nu\Omega_{\mu,\nu}\frac{x^\nu}{|x|}L_\nu.$$

Proof. Observe that

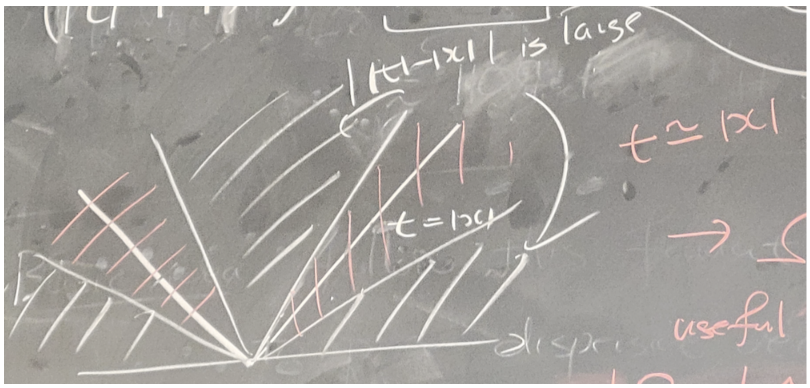
$$\begin{aligned} x^\nu \Omega_{\mu,\nu} &= x^\nu (x_\mu \partial_\nu - x_\nu \partial_\mu) \\ &= x_\mu \underbrace{x^\nu \partial_\nu}_{S} - \underbrace{x^\nu x_\nu}_{(-t^2+|x|^2)} \partial_\mu. \end{aligned}$$

□

This means that

$$(|t| - |x|)\partial_\mu = \underbrace{\frac{x_\mu}{|t|+|x|}}_{\leq 1} S - \underbrace{\frac{x^\nu}{|t|+|x|}}_{\leq 1} \Omega_{\mu,\nu}.$$

Away from the cone $t = |x|$, we get control of the derivatives.



In the region where $t \simeq |x|$, the rotation vector fields $\Omega_{j,k}$ are useful. The size of these rotation vector fields is $|\Omega_{j,k}| \simeq |x|$. We control all angular derivatives ($d-1$ many directions) with weight $|x| \simeq t$; this is why we get $\frac{d-1}{2}$ instead of $\frac{d}{2}$ in the dispersive estimate.

The analytic key to this method is the following inequality.

Theorem 21.2 (Klaineman-Sobolev inequality)**.** *Let ψ be a nice function, and let $s > \frac{d}{2}$. Then for $t > 0$,*

$$|\psi(t,x)| \lesssim \frac{1}{(1+v)^{\frac{d-1}{2}}(1+|u|)^{1/2}} \sum_{|\alpha|\leq s} \|\Gamma^\alpha \psi\|_{L^2},$$

where $v = t - |x|$ and $u = t - |x|$.

If we apply this theorem to $\psi = \nabla_{t,x}\phi$, we get

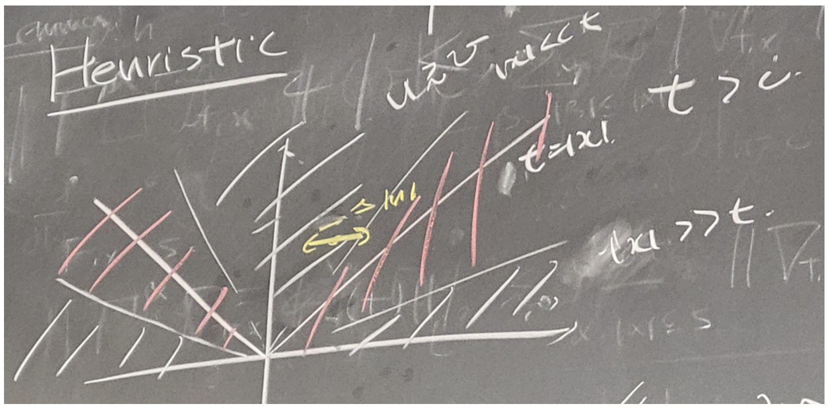
Corollary 21.3.

$$\begin{aligned} |\nabla_{t,x}\phi| &\lesssim \frac{1}{(1+v)^{\frac{d-1}{2}}(1+|u|)^{1/2}} \sum_{|\alpha|\leq s} \|\Gamma^\alpha \nabla_{t,x}\phi(t)\|_{L^2} \\ &\leq \frac{1}{(1+v)^{\frac{d-1}{2}}(1+|u|)^{1/2}} \sum_{|\alpha|\leq s} \|\nabla_{t,x}\Gamma^\alpha \phi|_{t=0}\|_{L^2} \end{aligned}$$

Here, the factor in front is $\lesssim \frac{1}{(1+t)^{\frac{d-1}{2}}}$, so we have something a little better than our original bound.

Here is the idea behind proving the Klaineman-Sobolev inequality.

Proof. The key heuristic is that Γ gives control of $|u|\partial_{\mu,x}$. Now decompose the space into regions where $|x| \ll t$ and $x \simeq t$, and $|x| \gg t$.



Then let $w \simeq \frac{1}{1+|u|)^{d/2}}$. When $|u| \lesssim 1$, the the usual Soboolev inequality works. Otherwise, if $|u| \gtrsim 1$, then $w \simeq \frac{1}{|u|^{d/2}}$.

Lemma 21.4 (Rescaled Sobolev)**.**

$$|\psi(x)| \lesssim \frac{1}{u^{d/2}} \sum_{|\alpha| \le s} \||u|^{|\alpha|} \partial^\alpha \psi\|_{L^2(B_{|u|}(x))}$$

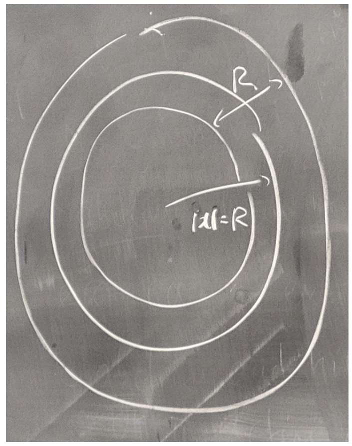
Proof. This follows from rescaling the Sobolev inequality on the unit ball $B_1(0)$. $\square$

When t and $|x|$ are comparable, the weight $w \simeq \frac{1}{(1+v)^{\frac{d-1}{2}}}$. If $|v| \lesssim 1$, the usual Sobolev inequality works. If $|v| \gtrsim 1$, then $w \simeq \frac{1}{v^{\frac{d-1}{2}}} \simeq \frac{1}{|x|^{\frac{d-1}{2}}}$. The final lemma we use is this:

Lemma 21.5 (Rescaled Sobolev on an annulus)**.**

$$|\psi(x)| \lesssim \frac{1}{R^{\frac{d-1}{2}}} \sum_{\alpha,\beta : |\alpha|+|\beta| \le s} \left(\int_{A_R} |\partial_r^\alpha \Omega_x^\beta \psi|^2 \, dx \right)^{1/2},$$

where $A_R = \{||x| - R| \leq cR\}$.



Here, $R^{\frac{d-1}{2}}$ responds to the angular directions that Ω_x^β has control over. $\square$

22 Introduction to Calculus of Variations

22.1 Motivation and general setup

Now, we will begin the final part of this course, where we will study nonlinear PDEs. Calculus of variations gives us a lot of extra structure which is helpful in studying nonlinear PDEs. The reference is sections 8.1, 8.6 in Evans' book, but we will give some more focus on the formalism than Evans.

In the calculus of variations, we are looking for the critical points of a functional $F : X \to \mathbb{R}$; these are necessary to find extrema and is motivated by optimization problems. We will give some more motivations later. For us X will be a set of functions, which differentiates this from an ordinary calculus problem.

Example 22.1 (Energy minimizing curves)**.** Given a curve $\gamma : [0,1] \to \mathbb{R}^d$, we can associate the **energy**

$$E[\gamma] = \int_0^1 |\dot{\gamma}(t)|^2 \, dt.$$

What are the minimizers of $\mathbb{E}[\gamma]$?

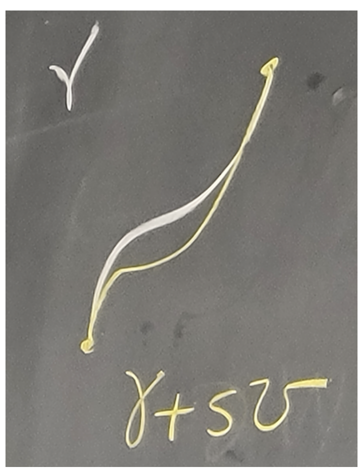
To solve problems like this, we need to generalize what we do in usual calculus: We want to find a way to say something like "$\nabla E[\gamma] = 0$." The idea is to think of directional derivatives instead. We can equivalently find a γ such that

$$\frac{d}{ds} E[\gamma + sv]|_{s=0} = 0$$

for all $v : [0,1] \to \mathbb{R}^d$ in a reasonable class.

22.2 Examples of the Euler-Lagrange equation

For simplicity, we assume $\gamma \in C^\infty([0,1];\mathbb{R}^d)$ and $v \in C_c^\infty((0,1);\mathbb{R}^d) =: \mathscr{A}$.



We can write out

$$\begin{aligned}
\frac{d}{ds}E[\gamma + sv]|_{s=0} &= \frac{d}{ds}\int_0^1 \left|\frac{d}{dt}(\gamma(t) + sv(t))\right|^2 dt\bigg|_{s=0} \\
&= 2\int_0^1 \frac{d}{dt}(\gamma + sv)\frac{d}{dt}v\bigg|_{s=0} dt \\
&= 2\int_0^1 \dot{\gamma}\frac{d}{dt}v\, dt \\
&= -2\int_0^1 \ddot{\gamma}v\, dt.
\end{aligned}$$

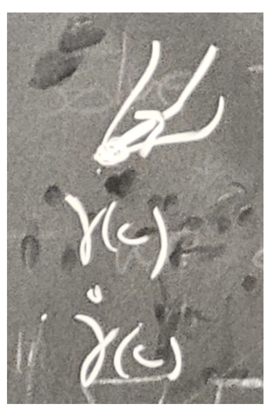
So we see that

$$\begin{aligned}
0 = \frac{d}{ds}E[\gamma + sv]|_{s=0} \qquad \forall v \in \mathscr{A} &\iff 0 = \int_0^1 \ddot{\gamma}v\, dt \qquad \forall v \in \mathscr{A} \\
&\iff \ddot{\gamma} = 0 \qquad \text{on } (0,1).
\end{aligned}$$

Our critical point condition gave us a differential equation. This is called the **Euler-Lagrange equation**. It tells us that energy minimizing curves are straight lines (geodesics). If we take $F[u]$, where $u : U \to \mathbb{R}^N$ and $U \subseteq \mathbb{R}^n$ with $n \geq 2$, we will in general get a PDE for our critical points. There are two ways to generalize this example:

1. If we look for the minimum of F we need some extra condition, such as the idea of **convexity** of F. This leads to elliptic PDEs. Chapter 8 of Evans' book focuses mostly on this approach.

2. We can interpret this as a Lagrangian mechanics problem. This is when there is a natural time variable in the problem. Here, we do not worry about minimizing F; we just look for critical points. In this setting, critical points give an equation (like $\ddot{\gamma} = 0$) which tells us locally how the curve will evolve given initial conditions.



This is known as the principle of **stationary action**, in which case, we call F the **action**.

Here are some examples of Euler-Lagrange equations corresponding to various calculus of variation problems.

Example 22.2 (Dirichlet's principle)**.** Let $u : U \to \mathbb{R}$, where U is an open, bounded C^∞ domain contained in $\mathbb{R}^d$. We take $u \in C^\infty(\overline{U})$ and take our functional to be

$$F[u] = \frac{1}{2} \int_U |Du|^2 \, dx.$$

(Compare this with our equation for geodesics). The critical points satisfy the PDE $-\Delta u = 0$.

Example 22.3 (Action principle for the wave equation)**.** Let $u : O \to \mathbb{R}$, where O is an open subset of $\mathbb{R}^{1+d}_{t,x}$. We take $u \in C^\infty(\overline{U})$ and

$$\mathcal{S}[u] = \int_O (\partial_t u)^2 - |Du|^2 \, dt \, dx.$$

The critical points satisfy the wave equation in O.

22.3 First variation (the Euler-Lagrange equation)

From now on, we restrict our attention to functionals of the form

$$F[u] = \int_U L(Du(x), u(x), x) \, dx,$$

where $L : (p, z, x) : \mathbb{R}^d \times \mathbb{R} \times U \to \mathbb{R}$ is called the **Lagrangian density**. For our notation, we will use brackets when we are talking about u as a function as a whole and parentheses when we are talking about values of u.

To think of first variations, we think of directional derivatives. Take $u \in \mathscr{A}$ and variations $v \in \mathscr{A}_0$. Then we will try to form

$$\frac{d}{ds} F[u + sv]|_{s=0}.$$

Remark 22.1. When $\mathscr{A}_0 = \mathscr{A}$, people in functional analysis call this the **Gâteaux derivative**.

In our case, for simplicity, eww assume $\mathscr{A} = C^\infty(\overline{U})$ and $\mathscr{A}_0 = C_c^\infty(U)$. The assumption on $\mathscr{A}_0$ is okay, and the assumption on $\mathscr{A}$ is restrictive but easily removable. We get

$$D_v F[u] = \frac{d}{ds} F[u + sv] \bigg|_{s=0}$$

$$
\begin{aligned}
&= \frac{d}{ds}\int_U L(D(u+sv), u+sv, x)\,dx\bigg|_{s=0} \\
&= \int_U \frac{d}{ds} L(D(u+sv), u+sv, x)|_{s=0}\,dx \\
&= \int_U \partial_j v \left(\frac{\partial}{\partial_{p_j}} L\right)(Du, u, x) + v(\partial_z L)(Du, u, x)\,dx \\
&= \int_U v\left(-\partial_j\left(\left(\frac{\partial}{\partial p_j} L\right)(Du, u, x)\right) + (\partial_z L)(Du, u, x)\right)\,dx.
\end{aligned}
$$

In particular, if $D_v F[u] = 0$ for all $v \in \mathscr{A}_0$,

$$\left(-\partial_j\left(\frac{\partial}{\partial p_j} L\right) + \partial_z L\right)(Du, u, x) = 0$$

in U. This is the **Euler-Lagrange equation**.

Example 22.4 (Dirichlet's principle)**.** In this example, $L = \frac{1}{2}|p|^2$, so the Euler-Lagrange equation is

$$0 = \partial \underbrace{\left(\frac{\partial}{\partial p_j}\frac{1}{2}|p|^2\right)}_{p_j}|_{p=D_u},$$

which gives us $-\Delta u = 0$.

Example 22.5 (Action principle for the wave equation)**.** In this example, $L = \frac{1}{2}p_0^2 - \frac{1}{2}|p_x|^2$. The Euler-Lagrange equation is

$$0 = -\partial_t \underbrace{\left(\frac{\partial}{\partial p_0} L\right)}_{p_0}|_{p_{t,x}=D_{t,x}u} - \sum_{j=1}^{d} \partial_j \left(\frac{\partial}{\partial p_j} L\right)\bigg|_{p_{t,x}=D_{t,x}u},$$

so we get $-\partial_t^2 u + \Delta u = 0$.

Remark 22.2. In calculus,

$$D_v F[u] = \langle v, \nabla F[u]\rangle.$$

With a choice of inner product, we can define the **gradient** of F. In our case, we have computed that

$$D_v F[u] = \int_U v(\cdots)\,dx.$$

With respect to the L^2 inner product $\langle\cdot,\cdot\rangle = \int_U uv\,dx$, we have

$$D_v F[u] = \langle v, \text{LHS of E-L equation}\rangle.$$

Because of this, the left hand side of the Euler-Lagrange equation is sometimes called the **L^2-gradient of** F, ∇F. Note that ∇F is now an operator $u \mapsto \nabla F[u]$.

22.4 Second order variation

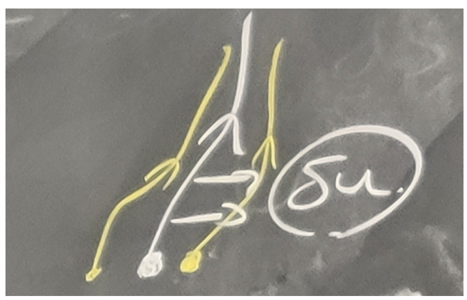
Again, start from directional derivatives. In calculus, the proper way to think about second order directional derivatives is the following:

$$D_{v,w}F[u] = \frac{d}{ds}\frac{d}{dt}F[u + sv + tw]|_{s=0,t=0}.$$

In our case, we define second order directional derivatives of F by this formula. There are two interpretations of the second order variation:

1. In the context of minimization, we can think of this as the Hessian of F at u contracted with two direction vectors v, w. We can then try to come up with a second derivative test to see if a critical point is a maximizer or minimizer.

2. We can think of this as a linearized operator around a critical point. Often, we are not just interested in a single solution but also nearby solutions; this allows us to think about variation through critical points.



In geometry, this is the notion of **Jacobi shifts**. We want $u(x;\lambda)$ such thst $u(x;0) = u(x)$ is a given critical point and $u(x;\lambda)$ are all critical points. We can write this as

$$\nabla F[u(x,\lambda)] = 0,$$

or

$$D_v F[u(x;\lambda)] = 0.$$

We can then differentiate this in λ and get that

$$\frac{d}{d\lambda}\nabla F[u(x,\lambda)]\bigg|_{\lambda=0} = 0,$$

or

$$\frac{d}{d\lambda}D_v F[u(x;\lambda)]\bigg|_{\lambda} = 0, \qquad (v \in \mathscr{A}_v),$$

where

$$u(x;\lambda) = u(x) + \lambda \frac{\partial}{\partial \lambda} u \bigg|_{\lambda=0} = u(x) + \delta u.$$

We can write

$$D_{\delta u} D_v F[u] = 0,$$

which is called the **linearization** of the Euler-Lagrange equation around U for δu.

22.5 Nöther's principle

This principle can be summarized with a slogan: "(continuous) symmetries of the action correspond to conservation laws for solutions." In nonlinear PDEs, conservation laws are very useful but hard to come by. Oftentimes, you have no idea what the solution to an equation is but you know that it's invariant under, say, time translations. This gives you a conserved quantity we can study to understand the solutions to an equation.

Introduce a parameter τ and think about a 1-parameter family of variations.

Definition 22.1. $x \mapsto X(x, \tau)$ si called the **domain variation**, and $u \mapsto u(x, \tau)$ is called the **function variation**.

Example 22.6. We can, for example, take $X(x,\tau) = x - \tau e_1$ and $u(x, \tau) = u(x - \tau e_1)$.

Definition 22.2. F is **invariant** under $X(\cdot, \tau)$ and $a(\cdot, \tau)$ if

$$U(\tau) = X(U, \tau), \qquad u(x, 0) = u(x), X(x, 0) = x,$$

$$\int_U L(Du(x, \tau), u(x, \tau), x)\, dx = \int_{U(\tau)} L(Du, u, x)\, dx.$$

Theorem 22.1 (Nöther's principle)**.** *In this case,*

$$\partial_j (m \partial_{p_j} L - L v^j) = m \cdot \left(\partial_j \frac{\partial}{\partial p_j} L - \partial_z L \right) \bigg|_{p=Du, z=u},$$

where $m(x) = \frac{\partial}{\partial \tau} u(x, \tau)|_{\tau=0}$ *and* $v^j(x) = \frac{\partial}{\partial \tau} X^j(x, \tau)$.

The key idea is that $\partial_j \frac{\partial}{\partial p_j} L - \partial_z L|_{p=Du,z=u}$ is ∇F. We will discuss this in more detail next time.

23 Nöether's Principle and the Energy-Momemtum Tensor

23.1 Nöether's principle

Let's continue our discussion of Nöether's principle with an updated version of the slogan we gave last time. The slogan for the principle is '(continuous) symmetries give rise to conservation laws." The implication in the other direction is not always the case; for more on the reverse, you can see, for example, Carter's constant, which is a "hidden symmetry" for geodesics on Kerr spacetime.

Theorem 23.1. *Consider the Lagrangian action $F[u] = \int_U L(Du, u, x)\,dx$. Suppose there exists a continuous symmetry $(u_\tau(x), X_\tau(x))$ of the action (with $u_\tau : U \to \mathbb{R}$ and $X_\tau : \mathbb{R}^d \to \mathbb{R}^d$ a diffeomorphism for each τ), in the sense that*

$$\int_U L(Du_\tau(x), u_\tau(x), x)\,dx = \int_{U(\tau)} L(Du, u, x)\,dx,$$

where $U(\tau) := X_\tau(U)$. Then

$$\partial_{x^j}(m\partial_{p_j} L(Du, u, x) - L(Du, u, x)V^j) = m\left(\frac{\partial}{\partial x^j}(\partial_{p_j} L(Du, u, x)) - \partial_z L(Du, u, x)\right),$$

where $m = \frac{\partial}{\partial \tau} u|_{\tau=0}$, $u = u_\tau|_{\tau=0}$, $V^j = \frac{\partial}{\partial \tau} X^j_\tau|_{\tau=0}$, and $X_0(x) = x$.

Lemma 23.1. *Let $f_\tau = f_\tau(x)$, and let U_τ be a "smooth" family of C^∞ domains, i.e. there exist a family of diffeomorphisms $X_\tau : \mathbb{R}^d$ to$\mathbb{R}^d$ such that $U_\tau = X_\tau(U)$. Let $V(x) = \frac{\partial}{\partial \tau} X_\tau(x)\Big|_{\tau=0}$ for $x \in \partial U_0$. Then*

$$\frac{d}{d\tau}\int_{U_\tau} f_\tau(x)\,dx\Big|_{\tau=0} = \int_{U_0} \frac{\partial}{\partial \tau} f_\tau(x)\Big|_{\tau=0} dx + \int_{\partial U_0} f_0 V \cdot \nu.$$

Here is the proof of the theorem, assuming the lemma:

Proof.

$$\begin{aligned}
\frac{\partial}{\partial \tau}(\text{LHS})\Big|_{\tau=0} &= \frac{\partial}{\partial \tau}\int_U L(Du_\tau(x), u_\tau(x), x)\,dx\Big|_{\tau=0} \\
&= \int_U \frac{\partial}{\partial \tau} L(Du_\tau(x), u_\tau(x), x)\Big|_{\tau=0} dx
\end{aligned}$$

Using the Euler-Lagrange euqation,

$$= \int_U \frac{\partial}{\partial_j} L \cdot \partial_{x^j} m + \frac{\partial}{\partial z} L \cdot m\,dx$$

Integrating by parts,

$$= \int_U \left(-\partial_{x^j} \left(\frac{\partial}{\partial p_j} L \right) + \frac{\partial}{\partial z} L \right) m \, dx + \int_{\partial U} \frac{\partial}{\partial p_j} L \cdot m \nu_j \, dA.$$

The lemma gives

$$\frac{\partial}{\partial \tau} (\text{RHS}) \Big|_{\tau=0} = \int_{\partial U} L V^j \nu_j \, dA.$$

Putting these together, we get

$$\int_{\partial U} \left(\frac{\partial}{\partial p_j} L \cdot m - L V^j \right) \nu_j \, dA = \int_U \left(-\frac{\partial}{\partial x^j} (\partial_{p_j} L) + \partial_z L \right) m \, dx.$$

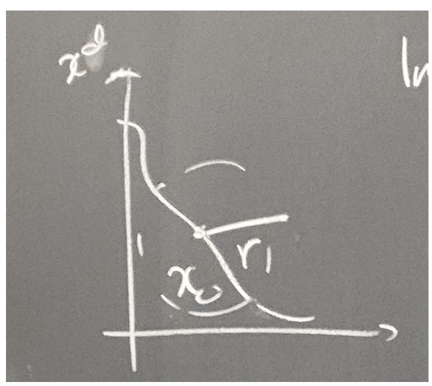
By the divergence theorem, the left hand side is

$$\int_U \partial_{x^j} \left(\frac{\partial}{\partial p_j} L \cdot m - L V^j \right) dx.$$

since U is arbitrary. $\square$

Here is a proof of this lemma, using the fact that the derivative of the Heaviside function is the delta distribution. (A more standard way to prove this is to use a change of variables to turn one of the integrals into a volume integral.)

Proof. Here is a sketch of the idea. Without loss of generality, let $f_\tau = f$, where $f \in C_c^\infty(\mathbb{R}^d)$ and $\operatorname{supp} f \subseteq B_r(x_0)$. Choose x_0 so that $U \cap B_r(x_0) = \{x^d > \gamma(x^1, \ldots, x^{d-1})\}$. So $X^d_{-\tau} - \gamma_\tau(X'_{-\tau})$ is the defining function for ∂U_τ.



Then

$$\int_{U_\tau} f \, dx = \int \mathbb{1}_{U_\tau} f \, dx = \int H(x^n - \gamma(x')) f(x) \, dx.$$

Now we can differentiate

$$\frac{\partial}{\partial\tau}\int H(\underbrace{X_\tau^d-\gamma_\tau(X_\tau')}_{u(X_{-\tau})})f(x)\,dx\bigg|_{\tau=0} = \int H'(\underbrace{X_\tau^n-\gamma_0(x')}_{u(X_{-\tau})})\underbrace{\frac{\partial}{\partial\tau}(x^n-\gamma_\tau(x'))}_{\frac{\partial}{\partial\tau}u(X_{-\tau})|_{\tau=0}}\cdot f(x)\,dx$$

$$= \int \delta_0(U(X_0))\underbrace{\partial_j u\cdot\frac{\partial}{\partial\tau}x^j}_{\nabla u\cdot V^j}\bigg|_{\tau=0}\cdot f(x)\,dx$$

The δ_0 part gives us the surface measure on ∂U times $\frac{1}{|\nabla u(x)|}$

$$= \int_{\partial U} f\underbrace{\frac{(-\nabla u)}{|\nabla u|}}_{\nu}\cdot V\,dA. \qquad \square$$

Remark 23.1. In the view of distribution theory, the divergence theorem is precisely telling us about the derivative of this kind of indicator function.

Example 23.1. Consider the action

$$F[\phi]=\int -\frac{1}{2}|\partial_t\phi|^2+\frac{1}{2}|D_x\phi|^2\,dx,$$

so $L=-\frac{1}{2}p_0^2+\frac{1}{2}|p_x|^2$. Let $\phi:\mathbb{R}^{1+d}\to\mathbb{C}$, and let $\phi_\tau(x)e^{i\tau}u(x)$ and $X_\tau(x)=x$. Then Nöether's principle tells us that there is an associated conservation law for the wave equation: $\partial_\mu J^\mu=0$, where

$$J^0=\mathrm{Im}(\phi\overline{\partial_t\phi}),\qquad J^j=\mathrm{Im}(\phi\overline{\partial_j\phi}).$$

This is called the conservation of the **charge-current vector**. J^0 is the natural change density, and J^j is the natural wave density if we want to couple the wave equation with Maxwell's equations.

In the case of the Schrödinger equation, this type of computation was carried out by Weyl. This gives rise to **gauge theory**. More examples can be found in Evans' book.

23.2 The energy-(stress)-momentum tensor

Here is useful alternate formulation of Nöether's principle. Our setting now is that $U\subseteq\mathcal{M}$, where $\mathcal{M}$ is a manifold with metric g (g may be Riemannian or Lorentzian or pseudo-Riemannian). Assume that

$$L(Du,u,x)=\mathcal{L}(du,u,g)\sqrt{|\det g|},$$

so the action looks like

$$F[u]=\int_U\mathcal{L}(du,u,g)\sqrt{|\det g|}\,dx.$$

This is invariant under change of coordinates, and the claim is that Nöether's principle will give us a conserved quantity that we call the energy-momentum tensor.

Theorem 23.2. *Assume that F is of the above form, and define*

$$T^{\mu,\nu} = \frac{\partial}{\partial g_{\mu,\nu}}\mathcal{L} + \frac{1}{2}(g^{-1})^{\mu,\nu}\mathcal{L}.$$

Then the covariant derivative associated with g satisfies

$$\nabla_\mu T^{\mu,\nu} = 0$$

if u satisfies the Euler-Lagrange equation.

Proof. Consider a compactly supported 1-parameter family of diffeomorphisms $X_\tau : U \to U$ such that $X_0(x) = x$ and such that for all τ, $X_\tau(x) = x$ outside some $K \subseteq\subseteq U$. The invariance looks like

$$\int_U \mathcal{L}(du, u, g)\sqrt{|\det g|}\, dx = \int_U \mathcal{L}(d(u \circ X_\tau), u \circ X_\tau, X_\tau^* g)\sqrt{|\det X_\tau^* g|}\, dx$$

Now

$$\frac{d}{d\tau}(\text{LHS})\bigg|_{\tau=0} = 0,$$

whereas

$$\frac{d}{d\tau}(\text{RHS})\bigg|_{\tau=0} = \underbrace{\left(\frac{\partial}{\partial\tau}\text{falls on } u \circ X_\tau\right)}_{I} + \underbrace{\left(\frac{\partial}{\partial\tau}\text{falls on } X_\tau^* g\right)}_{II}$$

Term I vanishes via the Euler-Lagrange equation. In fact, $I = \int(-\partial_\mu(\frac{\partial}{\partial p_\mu}\mathcal{L}) + \partial_z\mathcal{L})Vu\, dx$, where $V(x) = \frac{\partial}{\partial\tau}X_\tau(x)|_{\tau=0}$. For term II, we have

$$\int \frac{\partial}{\partial g_{\mu,\nu}}\mathcal{L}(du, u, g)\frac{\partial}{\partial\tau}X_\tau^* g_{\mu,\nu}\bigg|_{\tau=0}\sqrt{|\det g|}$$
$$+ \frac{1}{2}\mathcal{L}(du, u, g) \cdot \underbrace{\frac{\det g}{|\det g|}\frac{\frac{\partial}{\partial\tau}\det X_\tau^* g|_{\tau=0}}{|\det g|}}_{\frac{\partial}{\partial\tau}\log\det X_\tau^* g|_{\tau=0}}\sqrt{|\det g|}\, dx,$$

where we have used

$$\begin{aligned}\frac{\partial}{\partial\tau}\sqrt{|\det g_\tau|}\bigg|_{\tau=0} &= \frac{1}{2}\frac{1}{\sqrt{|\det g|}}\frac{\partial}{\partial\tau}|\det g_\tau| \\ &= \frac{1}{2}\frac{\det g}{|\det g|}\partial\tau \det g_\tau|_{\tau=0} \qquad = \frac{1}{2}\frac{1}{\det g}\partial_\tau(\det g_\tau)|_{\tau=0}\sqrt{|\det g|}.\end{aligned}$$

From elementary differential geometry, we have a name for this: this is the Lie derivative

$$\frac{\partial}{\partial\tau}(X_\tau^* g)_{\mu,\nu}\Big|_{\tau=0} = \mathcal{L}v g_{\mu,\nu} = \nabla_\mu V_\nu + \nabla_\nu V_\mu.$$

How do we differentiate the determinant function? First, note that we can differentiate near the identity:

$$\frac{\partial}{\partial\tau}\det(I+\tau A)\Big|_{\tau=0} = \mathrm{tr}(A).$$

Now if we let $B_0 = I$ and $\frac{\partial}{\partial\tau}B_\tau|_{\tau=0} = A$, then

$$\frac{\partial}{\partial\tau}\det(B_\tau)\Big|_{\tau=0} = \frac{\partial}{\partial\tau}\det(I+\tau A + O(\tau^2))\Big|_{\tau=0} = \mathrm{tr}\,A.$$

Now if $C_0 = M$ (which is invertible) and $\frac{\partial}{\partial\tau}C_\tau\Big|_{\tau=0} = A'$, then

$$\begin{aligned}\frac{\partial}{\partial\tau}\det(C_\tau)\Big|_{\tau=0} &= \frac{\partial}{\partial\tau}\det(M^{-1}(\tau))\Big|_{\tau=0}\det M\\ &= \det M\,\mathrm{tr}(M^{-1}A'),\end{aligned}$$

so that

$$\frac{\partial}{\partial\tau}\log\det C_\tau\Big|_{\tau=0} = \mathrm{tr}(M^{-1}A').$$

Now we can deal with the term $\frac{\partial}{\partial\tau}\log\det X_\tau^* g|_{\tau=0}$ as

$$\frac{\partial}{\partial\tau}\log\det X_\tau^* g|_{\tau=0} = \mathrm{tr}(g^{-1}\mathcal{L}_V g) = (g^{-1})^{\mu,\nu}(\mathcal{L}_V g)_{\mu,\nu}$$

All in all, we see that

$$\begin{aligned}II &= \int \underbrace{\left(\frac{\partial}{\partial g_{\mu,\nu}}\mathcal{L}\frac{1}{2}(g^{-1})^{\mu,\nu}\mathcal{L}\right)}_{T^{\mu,\nu}=T^{\nu,\mu}} \underbrace{\mathcal{L}_V g_\mu}_{\nabla_\mu V_\nu + \nabla_\nu V_\mu} \sqrt{|\det g|}\,dx\\ &= 2\int_U T^{\mu,\nu}\nabla_\mu V_\nu\sqrt{|\det g|}\,dx\\ &= -2\int_U(\nabla_\mu T^{\mu,\nu})V_\nu\sqrt{|\det g|}\,dx\\ &= 0\end{aligned}$$

for all X_τ. Thus, $\nabla_\mu T^{\mu,\nu} = 0$. $\square$

Example 23.2 (E-M for Laplace/wave equation)**.** Here, $\mathcal{L} = (g^{-1})^{\mu,\nu}\partial_\mu u \partial \nu u$, so

$$T^{\mu,\nu} = \partial_\mu u \partial_\nu u - \frac{1}{2} g_{\mu,\nu} \partial^\nu u \partial_\nu u.$$

We can see that

$$\frac{\partial}{\partial g_{\mu,\nu}} \mathcal{L} = \frac{\partial}{\partial g^{\mu,\nu}} (g^{-1})^{\mu',\nu'} \frac{\partial}{\partial (g^{-1})^{\mu',\nu'}} \mathcal{L} = -(g^{-1})^{\mu,\mu'} (g^{-1})^{\nu,\nu'} \frac{\partial}{\partial (g^{-1})^{\mu',\nu'}} \mathcal{L}.$$

24 Existence of Minimizers for Lagrangian Actions

24.1 Hilbert's 19th problem

We will now set about giving an answer to Hilbert's 19th problem, which concerns minimizers for certain functionals in the calculus of variations:

$$\mathcal{F}[u] = \int_U L(Du, u, x)\, dx.$$

Under certain conditions (having to do with ellipticity of the Euler-Lagrange equation), there exists a minimizer. Hilbert's 19th problem asks about the regularity of such a minimizer. The minimizers that we find will a priori be in a class of rough functions, but in many situations, they will be solutions to some PDE and will have some smoothness.

This problem was solved by de Giorgi, then later by Nash, and later simplified by Moser. This is called de Giorgi-Nash-Moser theory. Today we will discuss existence, and next time, we will discuss regularity. Since we lost a lecture, we will not have time to discuss our last topic, which is hyperbolic PDEs which arise from calculus of variations. A good reference for this missing topic is *Lectures on nonlinear wave equations* by J. Luk.

24.2 Coercivity

We will basically follow the exposition in Section 8.2 of Evans. Consider a Lagrangian action functional

$$\mathcal{F}[u] = \int_U L(Du, u, x)\, dx.$$

We define the admissible class of functions u we want to minimizer over will be $\mathscr{A} = \{u \in W^{1,q}(U) : u|_{\partial U} = g\}$. The problem is to find

$$\arg\min_{u \in \mathscr{A}} \mathcal{F}[u].$$

We will look for "natural" conditions on L that would guarantee the existence of a minimizer. One pathology that may arise is that $\mathcal{F}$ could decay to 0 if we go to infinity in some direction, so we assume the following condition.

Definition 24.1. The action $\mathcal{F}$ is **coercive** if

$$L(p, z, x) \geq c|p|^q - \beta$$

for some constants $c, \beta > 0$ and $1 < q < \infty$.

Coercivity implies that

$$\mathcal{F}[u] = \int_U L(Du, u, x)\, dx$$

$$\geq c \int |Du|^q \, dx + \beta |U|.$$

Using a Poincaré inequality, we can show that $\int_U |Du|^q \, dx$ controls the $W^{1,q}$ norm. In general, we should first determine the correct q from the action, which then specifies $\mathscr{A}$ accordingly.

24.3 Obstacles to convergence of a minimizing sequence

Let $\ell = \inf_{u \in \mathscr{A}} \mathcal{F}[u]$. There exists a sequence u_k such that $\mathcal{F}[u_k] \searrow \ell$. We want to say that

1. $u_k \to u \in \mathscr{A}$ for some u.
2. $\mathcal{F}[u_k] \to \mathcal{F}[u] = \ell$.

Then u will be a minimizer. In a finite dimensional setting, if we have compactness, we should actually assume that condition 1 is satisfied by a subsequence. But in fact, for $u_k \in \mathscr{A}$, both these conditions fail.

1. Failure of 1: From coercivity and a Poincaré inequality,

$$\begin{aligned} \|u_k\|_{W^{1,q}} &\lesssim \|Du_k\|_{L^q(U)} \\ &\lesssim \mathcal{F}[u_k] + \beta \\ &< \ell + \beta + 1. \end{aligned}$$

   But there does not in general exist a convergent subsequence in $W^{1,q}$. Here are two ideas that may help us to proceed.

   - Rellich-Kondrachov compactness tells us that there exists a subsequence $u_k \to u$ in $L^q(U)$.
   - (weak compactness) Since $1 < q < \infty$, there exists a subsequence with $u_k \to u$ weakly in $W^{1,q}(U)$ (that is, $Du_k \to Du$ weakly in $L^q(U)$).

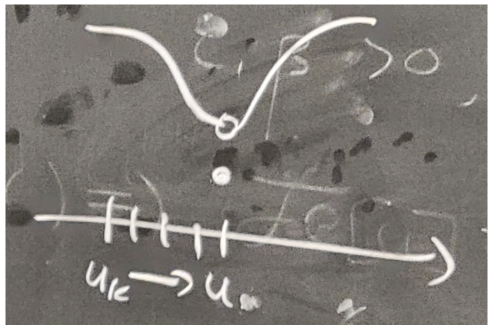
   Without loss of generality, we may assume these are the same subsequence.

2. Failure of 2: Because $Du_k \to Du$ weakly, to ensure that $\mathcal{F}[u_k] \to \mathcal{F}[u]$, we need some sort of continuity of $\mathcal{F}$ under (sequential) weak convergence. It turns out that this is way too restrictive; weak convergence plays very well with linear operators but is in general badly behaved for nonlinear operators. As an example, $e^{ikx} \to 0$ weakly in $\mathcal{D}'(\mathbb{R}^d)$ as $k \to \infty$. On the other hand, $z\overline{z}|_{z=e^{ikx}} = 1 \not\to 0$, so even the simplest nonlinearity can cause issues.

   The fix here is to realize that we only need "half" of the continuity property because $\mathcal{F}[u_k] \searrow \mathcal{F}[u]$.

Definition 24.2. A function f is **(sequentially) weak lower semicontinuous (LSC)** if for $u_k \to u$ weakly in $W^{1,q}(U)$ (i.e. $Du_k \to u$ weakly in $L^q(U)$ and $u_k \to u$ in $L^q(U)$), then

$$\liminf_{k\to\infty} \mathcal{F}[u_k] \geq \mathcal{F}[u].$$



Now, the question is: what is a natural condition on L that guarantees weak LSC of $\mathcal{F}$ on $W^{1,q}(U)$. The answer turns out to be convexity of L in p (Evans motivates this by looking at the Hessian of L):

$$\frac{\partial^2}{\partial p_j \partial p_k} L(p,z,x) \succeq 0 \qquad \forall p, z, x$$

or equivalently,

$$L(p,z,x) \geq L(p_0,z,x) + D_pL(p_0,z,x) \cdot (p - p_0).$$

This is also equivalent to

$$L(\theta p_1 + (1-\theta)p_2, z, x) \leq \theta L(p_1,z,x) + (1-\theta)L(p_2,z,x).$$

Example 24.1. $L = |p|^q$ is convex for $q > 1$.

We will show that this convexity implies weak LSC for $\mathcal{F}$.[8]

[8]It can be shown that these are actually eqiuvalent conditions.

24.4 Lower semicontinuity of the action

Here is the key theorem.

Theorem 24.1. *Assume L is convex in p, and assume coercivity: $L(p,z,x) \geq c|p|^q + \beta$. Then*

$$\mathcal{F}[u] = \int_U L(Du, u, x)\, dx$$

on $\mathcal{W}^{1,q}$ is weak LSC.

Proof. Assume without loss of generality that $\beta = 0$ (by replacing L by $L + \beta$). Take $\{u_k\} \in W^{1,q}(U)$ such that $Du_k \to Du$ weakly in L^q and $u_k \to u$ in $L^q(U)$. This is, up to subsequences, equivalent to $u_k \to u$ weakly in $W^{1,q}(U)$. Also passing to a subsequence, we can assume that $\mathcal{F}[u_k] \to \ell$. The goal is to show that $\ell \geq \mathcal{F}[u]$.

To handle nonlinear expressions in u_k, we use Egorov's theorem. Fix $\varepsilon > 0$. By Egorov's theorem, there exists a set G_ε such that

1. $|U \setminus G_\varepsilon| < \varepsilon$,
2. $u_k \to u$ uniformly on G_ε (up to a subsequence).

Also, define $H_\varepsilon = \{x \in U : |u| < 1/\varepsilon, |Du| \leq 1/\varepsilon\}$. By the monotone convergence theorem, we can arrange that $|U \setminus H_\varepsilon| \lesssim \varepsilon$. On $A_\varepsilon := G_\varepsilon \cap H_\varepsilon$, we have property 2 and $|U \setminus A_\varepsilon| \lesssim \varepsilon$. Now

$$\mathcal{F}[u_k] = \int_U L(Du_k u_k, x)\, dx$$

Since $L \geq c|p|^q$, it is ≥ 0. So we can shrink the domain of integration.

$$\begin{aligned} &\geq \int_{A_\varepsilon} L(Du_k, u_k, x)\, dx \\ &\geq \int_{A_\varepsilon} \underbrace{L(Du, u_k, x)}_{I} + \underbrace{D_p L(Du, u_k, x)(Du_k - Du)}_{II}\, dx. \end{aligned}$$

Take $k \to \infty$, so the left hand side converges to ℓ. By uniform convergence (and continuity of L in p, which we assume),

$$\int_{A_\varepsilon} I\, dx \to \int_{A_\varepsilon} L(du, u, x)\, dx.$$

For the other term,

$$\int_{A_\varepsilon} II\, dx = \int_{A_\varepsilon} \underbrace{(D_p L(Du, u_k, x) - D_p L(Du, u, x))}_{\to 0 \text{ unif.}} \cdot \underbrace{(Du_k - Du)}_{\|\cdot\|_{L^q} \lesssim 1}\, dx$$

$$+\int_{A_\varepsilon} D_p(Du,u,x,)\cdot(Du_k - Du)\,dx,$$

and the latter term goes to 0 thanks to the weak convergence of $Du_k \to Du$. Thus, we have

$$\ell \geq \int_{A_\varepsilon} L(Du,u,x)\,dx.$$

Let $\varepsilon \to 0$ so that "$|U \setminus A_\varepsilon| \to 0$." This gives

$$\ell \geq \int_U L(Du,u,x)\,dx,$$

as desired. □

Remark 24.1. We have been omitting some regularity assumptions on L.

24.5 Proof of existence of minimizers

Theorem 24.2. *In addition to regularity assumptions on L, assume that L is convex in p and $L \geq c|p|^q + \beta$. Consider $\mathscr{A} = \{u \in W^{1,q}(U) : u|_{\partial U} = g\}$ and the action*

$$\mathcal{F}[u] = \int_U L(Du,u,x)\,dx.$$

There exists a minimizer u for $\mathcal{F}[u]$ in $\mathscr{A}$.

Remark 24.2. Uniqueness and regularity conditions require more assumptions on L, which upgrade this convexity property.

Proof. Take a minimizing sequence u_k such that $\mathcal{F}[u_k] \searrow \ell$, where $\ell = \inf_{u\in\mathscr{A}} \mathcal{F}[u] < \infty$ (if this is $\ell = +\infty$, there is nothing to prove). By this and coercivity, $\|Du_k\|_{L^q(U)} \lesssim 1$. There exists some extension $\widetilde{g} \in W^{1,q}$ such that $\widetilde{g}|_{\partial U} = g$, so we can consider $u_k - \widetilde{g} \in W_0^{1.q}(U)$. A Poincaré inequality gives

$$\begin{aligned}\|u_k - \widetilde{g}\|_{W^{1,q}(U)} &\lesssim \|Du_k - D\widetilde{g}\|_{L^q(U)} \\ &\lesssim 1.\end{aligned}$$

By weak compactness of the norm-unit ball in $L^q(U)$, up to a subsequence, we may assume $Du_k \to Du$ weakly in $L^q(U)$. By Rellich-Kondrachov compactness, up to a subsequence, $u_k \to u$ in $L^q(U)$. Now apply the weak LSC theorem to get that

$$\ell = \inf_{v\in\mathscr{A}} \mathcal{F}[v] \leq \mathcal{F}[u] \leq \ell.$$

This gives $\mathcal{F}[u] = \ell$. □

Theorem 24.3. *Let L satisfy*

$$|L| \leq c(|p|^q + |z|^q + 1), \quad |D_p L| \leq c(|p|^{q-1} + |z|^{q-1} + 1), \quad |D_z L| \leq C(|p|^{q-1} + |z|^{q-1} + 1).$$

Then any minimizer u for $\mathcal{F}[u]$ in $\mathscr{A}$ is a weak solution to the Euler-Lagrange equation. That is,

$$\int_U (\partial_{p_j} L(Du, u, x)\partial_{x^j} v + \partial_z L(Du, u, x)v)\, dx \qquad \forall v \in W_0^{1,p}, \qquad \frac{1}{p} + \frac{1}{q} = 1.$$

See Evans for the proof.

25 The de-Giorgi-Nash-Moser Theorem

25.1 How the theorem answers Hilbert's 19th problem

Today, we will be concluding our discussion of the solution to Hilbert's 19th problem, which was posed in 1900. Here is the problem:

Problem 25.1. *Assume $L = L(p)$ is convex and analytic. Prove that minimizers of $\mathcal{F}[u] = \int_U L(Du)\,dx$ are analytic.*

The original problem was stated for $d = 2$ and was solved by Morrey (at Berkeley). Later, Nash solved the problem for $d \geq 3$, but it turns out that de Giorgi solved the problem (with a slightly different theorem) a few years earlier; so both get the credit. Later, Moser simplified the theory and proved a number of other theorems along the way. So this is generally referred to as de Giorgi-Nash-Moser theory.

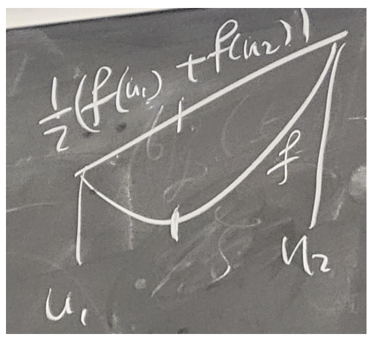
Today, we will be proving the following theorem

Theorem 25.1 (de Giorgi-Nash-Moser)**.** *Assume $L \in C^\infty(\mathbb{R}^d)$ and L is uniformly convex, i.e.*

$$\lambda|\xi|^2 \leq \partial_{p_j}\partial_{p_k} L \xi_j \xi_k \leq \Lambda|\xi|^2.$$

Then for all $V \subseteq\subseteq U$, the minimizer $u \in C^\infty(V)$.

Remark 25.1. With uniform convexity, the uniqueness of the minimizer follows.



Convexity of a function always tells you that

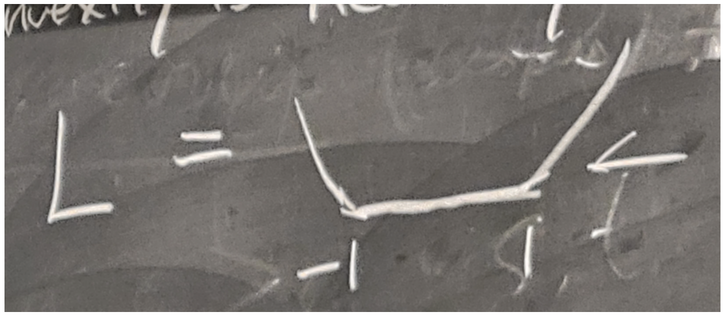
$$f\left(\frac{u_1 + u_2}{2}\right) \leq \frac{f(u_1) + f(u_2)}{2},$$

and strict convexity means that equality holds iff $u_1 = u_2$. So let u_1, u_2 be minimizers for L. Then

$$\int_U L\left(D\left(\frac{u_1 + u_2}{2}\right)\right) \leq \int_U \frac{1}{2}L(Du_1) + \frac{1}{2}L(Du_2).$$

This means that $\frac{u_1+u_2}{2}$ is also a minimizer, so strict convexity gives $Du_1 = \frac{Du_1 + Du_2}{u_2}$. So $Du_1 = Du_2$ in ∂U, and since u_1, u_2 agree on the boudnary, we get $u_1 = u_2$ in U.

Remark 25.2. Strict convexity is necessary for the theorem. Consider the following example in $d = 1$:



Then $|x|$ would be a minimizer, so

$$L(D|x|) = \int L_{\min}\, dx = \ell,$$

but $|x|$ is only Lipschitz.

25.2 Reduction to $u \in C^{1,\alpha}(V)$

Now we will prove a key reduction to $u \in C^{1,\alpha}(V)$. The keyword here is "standard elliptic theory," and in particular L^2 and Schauder theory. The minimizer will satisfy the Euler-Lagrange equation

$$\partial_{x^j}(\partial_{p_j} L(Du)) = 0.$$

The minimizer $u \in H^1(U)$ solves this equation in the weak sense. Let us differentiate this once more. Letting $w_i = \partial_i u$, we will have that each w_i solves the linearized Euler-Lagrange equation

$$\partial_j \left(\frac{\partial^2}{\partial p_j \partial p_k} L \bigg|_{p=Du} \partial_k w_i \right) = 0.$$

The term $\frac{\partial^2}{\partial p_j \partial p_k} L|_{p=Du}$ is uniformly elliptic (i.e. $\lambda|\xi|^2 \leq a^{j,k}\xi_j\xi_k$ with $|a| \leq \Lambda$) and in L^∞. This tells us that $w_i \in H^1(U)$, which follows from standard L^2-elliptic regularity theory (see Evans section 8.3 for details).

But still, all we know is that $a^{j,k} \in L^\infty$. What do we need? All we need is to show that $a^{j,k} \in C^{0,\alpha}$ for some $\alpha > 0$. Remember our equation is

$$\partial_j(a^{j,k}\partial_k w) = 0.$$

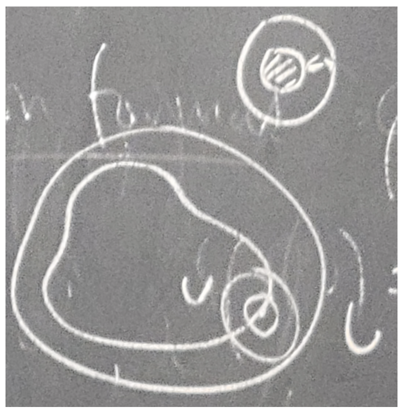
If $a^{j,k} \in C^{0,\alpha}$, then by Schauder theory, $w \in C^{1.\alpha}$. Then we have that $u \in C^{2,\alpha}$, so $a^{j,k} \in C^{1,\alpha}$. Then we get $w \in C^{2,\alpha}$, and we repeat. This is called an (elliptic) bootstrap argument.

The heart of the de Giorgi-Nash-Moser theory is to show that $a^{j,k} \in C^{0,\alpha}$ for some $\alpha > 0$. Now it suffices to show the following theorem.

Theorem 25.2. *Let $w \in H^1(B_1)$ be a solution to $Pw = -\partial_j(a^{j,k}\partial_k w) = 0$. Assume that $a \in L^\infty$ and $\lambda|\xi|^2 \leq a^{j,k}(x)\xi_j\xi_k \leq \Lambda|\xi|^2$. Then*

$$\|w\|_{C^{0,\alpha}(B_{1/2})} \lesssim_{d,\lambda,\Lambda} \|w\|_{L^2(B_1)}.$$

Here we only need to consider a ball because we can cover U will balls. The radius $1/2$ is not important; we could choose any larger number which is < 1.



25.3 Proof of the de Giorgi-Nash-Moser theorem

25.3.1 L^2 to L^∞ bound via Moser iteration

Step 1 of the proof is an L^2 to L^∞ bound.

Proposition 25.1. *Suppose that $Pw \leq 0$ and $w > 0$.*

$$\|w\|_{L^\infty} \lesssim_{d,\lambda,\Lambda} \|w\|_{L^2(B_1)}.$$

These conditions tells us that we cannot have a large peak to contribute to the L^∞ norm without contributing much to the L^2 norm.

Proof. (Moser iteration) Here are the ingredients:

1.

   Lemma 25.1 (Energy estimate for $Pw \leq 0$, $w > 0$)**.** *For all $\theta \in (0,1)$,*

$$\|Dw\|_{L^2(B_{\theta R})} \lesssim \frac{\Lambda}{\lambda}\frac{1}{\theta R}\|w\|_{L^2(B_R)}.$$

Proof. Multiply by a cutoff χ which is 1 in $B_{\theta R}$ and 0 outside B_R and with $|D^\alpha| \lesssim \frac{1}{(\theta R)^{|\alpha|}}$. Then we use the energy method:

$$\begin{aligned} 0 &\geq \int Pw\chi^2 w\,dx \\ &= \int -\partial_k(a^{j,k}\partial_k w)\chi^2 w\,dx \\ &= \int a^{j,k}\partial_k w\chi^2\partial_j w\,dx + 2\int a^{j,k}\partial_j w\chi\partial_k\chi w. \end{aligned}$$

This means that

$$\begin{aligned} \int \chi^2 Dw^2\,dx &\leq \frac{1}{\lambda}\int a^{j,k}\partial_j w\partial_k w\chi^2\,dx \\ &\leq -2\frac{1}{\lambda}\int a^{j,k}\partial_j w\chi w\partial_k\chi \\ &\leq 2\frac{\Lambda}{\lambda}\int \chi|Dw||Dx||w|\,dx \\ &\lesssim \frac{1}{\theta R}\frac{\Lambda}{\lambda}\left(\int x^2|Dw|^2\,dx\right)^{1/2}\left(\int_{B_R}|w|^2\,dx\right)^{1/2}. \end{aligned}$$

Now cancel on both sides to get the result. □

2. Sobolev embedding: For $d \geq 3$, let $p* = \frac{2d}{d-2}$. If $\theta R < 1$,

$$\|w\|_{H^1(B_{\theta R})} \lesssim \frac{\Lambda}{\lambda}\frac{1}{\theta R}\|w\|_{L^2(B_R)}.$$

By the Sobolev ineuqality, we get a better L^p bound:

$$\|w\|_{L^{p*}(B_{\theta R})} \lesssim \frac{\Lambda}{\lambda}\frac{1}{\theta R}\|w\|_{L^2(B_R)}.$$

How do we iterate Step 2? The observation of Moser was that if $\beta > 1$, $Pw \leq 0$, and $w > 0$, then w^β satisfies $Pw \leq 0$ and $w > 0$; this is because the map $s \mapsto s^\beta$. Composing convex functions preserves convexity, and composing subsolutions gives a subsolution, as well. Therefore, we can apply 2 to w^β, we get

$$\|w\|^\beta_{L^{p*\beta}(B_{\theta R})} \lesssim \frac{\Lambda}{\lambda}\frac{1}{\theta R}\|w\|^\beta_{L^{2\beta}(B_R)}.$$

We can rewrite this as

$$\|w\|_{L^{p*\beta}(B_{\theta R})} \lesssim \left(\frac{\Lambda}{\lambda}\frac{1}{\theta R}\right)^{1/\beta}\|w\|^\beta_{L^{2\beta}(B_R)}.$$

If we denote $q = 2\beta$ and $\alpha = \frac{p_*}{2} > 1$, then this equation looks like

$$\|w\|_{L^{\alpha q}(B_{\theta R})} \lesssim \left(\frac{\Lambda}{\lambda}\frac{1}{\theta R}\right)^{2/q} \|w\|^{\beta}_{L^q(B_R)}.$$

We want to iterate this equation $(2q)$. Start with $q_0 = 2$, then apply this to $q_1 = 2\alpha$ and so on, so $q_n = 2\alpha^n$. What should our θs be so that the radius of the ball does not go to 0? The radii are $R_0 = 1$, $R_1 = \theta_1$, $R_2 = \theta_1\theta_2$, and so on, so $R_n = \theta_1 \cdots \theta_n$. The constants we get will be

$$C_1 = \left(\frac{\Lambda}{\lambda}\frac{1}{\theta_1 R_0}\right)^{2/q_0}, \cdots,$$

$$C_n = \left(\frac{\Lambda}{\lambda}\frac{1}{\theta_n R_{n-1}}\right)^{2/q_{n-1}} C_{n-1}\cdots C_1 = \left(\frac{\Lambda}{\lambda}\frac{1}{\theta_n\cdots\theta_1}\right)^{1/\alpha^{n-1}} \cdots \left(\frac{\Lambda}{\lambda}\frac{1}{\theta_{n-1}\cdots\theta_1}\right)^{1/\alpha^{n-2}}.$$

Our goal is to choose $\theta_1, \theta_2, \ldots$ so that $\theta_1\theta_2\cdots = R_\infty = 1/2$. So we want

$$\theta_n^{-\frac{1}{\alpha^{n-1}}}\theta_{n-1}^{-\frac{1}{\alpha^{n-1}}-\frac{1}{\alpha^{n-2}}}\cdots\theta_1^{-\frac{1}{\alpha^{n-1}}\cdots-\frac{1}{\alpha}} = C_\infty < \infty.$$

If we let $a_n = \log\theta_n$, then we want $\exp(-\sum a_n) = 1/2$ and

$$\exp\left(\frac{1}{\alpha}a_1 + \frac{1}{\alpha^2}(a_1+a_2) + \cdots + \frac{1}{\alpha^{n-1}}(a_1+\cdots+a_{n-1}) + \cdots\right) < \infty. \qquad \square$$

These ingredients are the same things that de Giorgi's proof used, but his argument used sub-level sets instead of this iteration, so it was much more geometric.

25.3.2 Hölder seminorm bound via the de Giorgi oscillation lemma

The remaining step of the proof of the de Giorgi-Nash-Moser theorem is the following.

Proposition 25.2. *Let $w \in H^1(B_1)$ satisfy $Pw = 0$. Then there exists an $\alpha > 0$ such that*

$$[w]_{C^{0,\alpha}(B_{1/4})} \lesssim_{d,\lambda,\Lambda} \|w\|_{L^2(B_1)}.$$

This uses an oscillation lemma.

Lemma 25.2 (de Giorgi oscillation lemma)**.** *There exists a $\gamma \in (0,1)$ such that for $w \in H^1(B_1)$ with $Pw = 0$,*

$$\operatorname{osc}_{B_{1/2}} w \leq \gamma \operatorname{osc}_{B_1} w,$$

where $\operatorname{osc}_U w := \sup_U w - \inf_U w$.

Here is how the lemma implies the proposition.

Proof. The idea is to let $D = |x-y|$ and apply the oscillation lemma iteratively to get

$$|w(x) - w(y)| \leq \operatorname{osc} B_{B_D(w)} w \leq \gamma^n \operatorname{osc}_{B_{2^n D(x)}} w$$

Now let $n = -\log_2 D + c$ so that $B_{2^n D} \subseteq B_1$. We get

$$\begin{aligned} &\lesssim \gamma^n \|w\|_{L^\infty(B_1)} \\ &\lesssim D^\alpha \|w\|_{L^2(B_2)} \end{aligned}$$

where $\alpha = -\log_2 \gamma > 0$, so $\gamma^n = \gamma^{-\log_2 D}$. □

25.3.3 The de Giorgi-Harnack inequality

The way to prove the de Giorgi oscillation lemma lemma is to see that w should satisfy a sort of Harnack inequality.

Lemma 25.3 (de Giorgi-Harnack inequality)**.** *Let $w \in H^1(B_1)$ with $1 > w > 0$ and $Lw = 0$. Assume that*

$$\left| \left\{ x \in B_{1/2} : w \geq \frac{1}{2} \right\} \right| \geq \frac{1}{2} |B_{1/2}|.$$

Then there exists a $\gamma > 0$ such that $w \geq \gamma$ in $B_{1/2}$.

Here is how the de Giorgi-Harnack inequality implies the oscillation lemma.

Proof. Without loss of generality, we may arrange for $\sup_{B_1} w = 1 - \varepsilon$ and $\inf_{B_1} w = \varepsilon$. On $B_{1/2}$, one of the following must hold:

1. $|\{x \in B_{1/2} : w \geq \frac{1}{2}\}| \geq \frac{1}{2}|B_{1/2}|$: In this case, apply the de Giorgi-Harnack inequality for w.

2. $|\{x \in B_{1/2} : w \geq \frac{1}{2}\}| \leq \frac{1}{2}|B_{1/2}|$: This this case, $1 - w$ is still a solution, so we can apply he de Giorgi-Harnack inequality for $1 - w$. □

Moser's approach actually proves the de Giorgi-Harnack inequality without the last assumption, but this needs PMO theory. Here is a quick proof of the inequality:

Proof. The key idea is to look at $v = -\log w$. (Exercise: For $-\Delta u = 0$ in U, show that $|\nabla \log u|_{L^\infty(V)} \lesssim 1$ for all $V \subseteq\subseteq U$. Then get that $\min_V u \geq \gamma \max_V u$.) There is an a priori bound for $\nabla \log w$:

Lemma 25.4. *Suppose $w \in H^1(B_1)$ with $Pw \geq 0$ and $w > 0$. Then*

$$\|\nabla \log w\|_{L^2(B_{1/2})} \lesssim \frac{\Lambda}{\lambda}.$$

Proof. Multiply $Pw \geq 0$ by w^{-1} and integrate over U. □

This is deficient in two ways: it is not an L^∞ bound, and it is only a bound on the gradient, not w itself. However, notice that w is also a subsolution, so $v = -\log w$ is a subsolution: $Pv \geq 0$. When $w < 1$, $v > 0$. So we have inequality of the form

$$\|v\|_{L^\infty(B_{1/4})} \lesssim \|v\|_{L^2(B_{1/2})}.$$

The last assumption in the statement of the de Giorgi-Harnack inequality tells us that

$$|\{x \in B_{1/2} : v \leq \log 2\}| \geq \frac{1}{2}|B_{1/2}|.$$

Now we use a Poincaré-type inequality:

Lemma 25.5. *If the above bound (*) holds, ad $v \in H^1(B_{1/2})$, then*

$$\|v\|_{L^2(B_{1/2})} \lesssim \|Dv\|_{L^2(B_{1/2})} + 1.$$

Proof. By the standard Poincaré inequality, there exists a c such that

$$\|v - c\|_{L^2(B_{1/2})} \lesssim \|Dv\|_{L^2(B_{1/2})}.$$

Now split into cases: If $c \leq 100 \log 2$, we are done. If $c \geq 100 \log 2$, then

$$\begin{aligned}
\|Dv\|_{L^2(B_{1/2})} \geq \|v - c\|_{L^2(B_{1/2})} & \\
&\geq \|v - c\|_{L^2(A)} \\
&\geq \frac{99}{100} c \|1\|_{L^2(A)} \\
&\gtrsim c,
\end{aligned}$$

where the last step uses the above bound (*). ☐

This completes the proof of the Giorgi-Harnack inequality. ☐