Qualifying Exam Study Guide

Izak Oltman

April 12 2021

Contents

1 Partial Differential Equations

1.1 Distributions

Reference: (Oh 5). Basic definitions, basic operations, convolutions, fundamental solutions

1.1.1 Basic Definitions

Distributions are linear continuous functionals on the space of smooth compactly supported functions.

Definition 1.1 (Test Functions). *Let $U \subset \mathbb{R}^n$ be an open set. φ is a test function on U if it is smooth with compact support, and its support is contained in U. The space of such functions is denoted $\mathcal{D}(U)$ or $C_0^\infty(U)$*

A basic example of a test function is $\mathbb{1}_{|x|<1}\exp(-(1-|x|)^{-1})$. If we normalize this and call it $\varphi(x)$, we can get more test functions via convolution.

Proposition 1.1 (Density of Test Functions in C^k functions). *If $f \in C^k(\mathbb{R}^n)$, $\varphi_\delta = \delta^{-d}\varphi(\delta^{-1}x)$, then letting $f_\delta = f * \varphi_\delta$:*

1. *f_δ are smooth*
2. *$\partial^\alpha f_\delta$ converges uniformly on every compact set to $\partial^\alpha f$ for $|\alpha| \le k$ as $\delta \to 0$*
3. *if f has compact support, then f_δ has compact support.*

Proof.

1. $f * \varphi_\delta(x) = \int f(y)\varphi_\delta(x-y)dy$. The integrand is in L^1_{loc} so we can bring derivatives inside the integral which fall on φ_δ.
2. $\varphi_\delta * f - f = \int \varphi_\delta(y)(f(x-y) - f(x))dy$. Make δ small, use continuity to get this integral small. Uniform bounds come from this (easier to see with a change of variables). Derivative bound comes the exact same way.
3. this follows from $\text{supp}(f * g) \subset \text{supp} f + \text{supp}\, g$

□

Definition 1.2 (Distribution). *A distribution u on $\mathbb{R}^n$ is a linear functional on $C_0^\infty(\mathbb{R}^n)$ that is continuous with respect to the (very strong) topology on C_0^∞:*

$$\varphi_n \to \varphi \iff \text{supp}\,\varphi_n, \text{supp}\,\varphi \subset K \subset_{compact} \mathbb{R}^n \text{ and } \|\partial^\alpha(\varphi_n - \varphi)\|_{L^\infty(K)} \to 0 \ \forall \ \alpha \in \mathbb{N}^n$$

the space of distributions is denoted $\mathcal{D}'(\mathbb{R}^n)$

Proposition 1.2 (Alternate Distribution Definition based on Order). *A linear functional u on C_0^∞ is a distribution if and only if for all compact $K \subset \mathbb{R}^n$, there exists C_K and N such that for all $\varphi \in C_0^\infty$ with support in K, then:*

$$|u(\varphi)| \le C_K \|\varphi\|_{C_0^N(K)}$$

where $\|\varphi\|_{C_0^N(K)} = \sum_{|\alpha| \le N} \sup_{x \in K} |\partial^\alpha \varphi|$

Proof. The backwards definition is trivial, the backwards is a proof by contradiction, which is pretty easy. □

Definition 1.3 (Order of a Distribution). *If N in the above proposition is uniform and minimal, then we say u is a distribution of order N.*

Definition 1.4 (Support of a Distribution). *The support of a distribution u, denoted $\operatorname{supp} u$ is the compliment of the largest open set where u vanishes*[a]

We could we also define this as all $x \in \mathbb{R}^n$ such that for all $\varepsilon > 0$, there exists $\varphi \in C_0^\infty$ with $\operatorname{supp} \varphi \subset B_\varepsilon(x)$ and $u(\varphi) \neq 0$ Note that support is a closed set.

1.1.2 Basic Operations

1. **Adjoint**: if $\mathcal{A}, \mathcal{A}' : C_0^\infty(\mathbb{R}^n) \to C_0^\infty(\mathbb{R}^n)$ is such that $\int \mathcal{A}uv dx = \int u\mathcal{A}'v$, then for $u \in \mathcal{D}'(\mathbb{R}^n)$, define $\mathcal{A}u(\varphi) = u(\mathcal{A}'\varphi)$

2. **multiplication by smooth function** if $f \in C^\infty$, $u \in \mathcal{D}'(\mathbb{R}^n)$ define $(fu)(\varphi) = u(f\varphi)$

3. **Differentiation**: $(\partial_{x_i} u)(\varphi) = -u(\partial_{x_i}\varphi)$

4. **Convolution with compactly supported smooth function** $f \in C_0^\infty$, define $(f * u)(\varphi) = u(\int f(y - x)\varphi(y)dy)$

Convolution with C_0^∞ gives a smooth function. We can actually convolute with a compactly supported distribution – this will allow us to approximate distributions by test functions via mollification. However, we cannot convolute, in general, with a smooth function

Definition 1.5 (Convergence of Distributions). *$u_n \in \mathcal{D}'$ converges as a distribution to $u \in \mathcal{D}$ if for all $\varphi \in C_0^\infty$, $u_n(\varphi) \to u(\varphi)$ as $n \to \infty$.*

Theorem 1.1 (Sequential Convergence of Distributions). *If u_n are distributions such that for all $\varphi \in C_0^\infty$, $u_n(\varphi)$ converges, then there exists $u \in \mathcal{D}'$ such that $u_n \to u$ in the distributional sense. Furthermore, if $\varphi_n \to \varphi$, then $u_n(\varphi_n) \to u(\varphi)$ as $n \to \infty$ and on each compact set K, the order of u_n's are uniformly bounded.*

Proof.

1. Fix $K \subset_{compact} \mathbb{R}^n$ and $\varphi \in C_0^\infty(K)$, then $|u_n(\varphi)|$ is bounded uniformly in n (pointwise bound)

2. we can apply the uniform boundedness principal (because $C_0^\infty(K)$ is a Frechet space) to get $\|u_n\|_{C_0^\infty(K)\to\mathbb{C}} < C < \infty$ for all n

3. pass to a limit

□

This is nice. If u_n are nice functions, and we understand what they do to test functions, then we automatically get a distribution. And we can easily compute what this distribution does to things, or sequences of things.

Theorem 1.2 (Approximation of Distributions). *C_0^∞ is dense in the space of distributions.*

[a] u vanishes on an open set V if for all $\varphi \in C_0^\infty$ supported on V, $u(\varphi) = 0$

Proof. Let's approximate $u \in \mathcal{D}'$

1. let $\varphi \in C_0^\infty$, with $\int \varphi(x)dx = 1$ and let $\varphi_\delta = \delta^{-n}\varphi(\delta^{-1}x)$
2. $\varphi_\delta * u$ is smooth an converges in distribution to u
3. to see this use the definition of convolution, see that $\varphi_\delta \to \delta$ in distribution, then see all derivative go to the correct thing, then use continuity
4. for compact support, throw in cuttoff functions whose support expand to the entire space.

□

Proposition 1.3 (**Differentiation of Characteristic Function**). $\partial_j 1_U = -(\nu_{\partial U})_j dS_{\partial U}$, *where* ν_U *is the unit out normal vector, and* $dS_{\partial U}$ *is the distribution which is the surface measure on* ∂U

The first term direction makes sense, the function increase going towards the interior of U. The second term ensures the support is only on the boundary, and is properly scaled.

Definition 1.6 (Singular Support). *The singular support of a distribution is the compliment of the largest open set where the distribution locally coincides with a smooth function.*

Proposition 1.4 (**Multiplication of Distributions**). *We can multiply distributions if their singular supports are disjoint.*

The converse is false[a]. The prove this, show that if $u_n \to u$ and $v_n \to v$ are C_0^∞ approximations of said distributions, then $u_n v_n$ converges to a distribution. Use a cuttoff function, use linearity, and continuity of distributions.

Proposition 1.5 (**Convolution of Distributions**). *We can convolute distributions as long as at least one has compact support. And the usual support property holds*

Again this can be shown by approximation.

Corollary 1.1 (**Convolution of Distribution with Dirac**). *If* $u \in \mathcal{D}'$, *then* $\delta * u = u$

Remark 1.1. *For all these we use heavily the sequential convergence of distributions.*

1.1.3 Fundamental Solutions

Definition 1.7 (Fundamental Solution). *For a differential operator* P, *the fundamental solution* $E_y \in \mathcal{D}$ *at* y *is such that* $PE_y = \delta_y$

Not rigorous, but if we want to solve $Pu = f$ for a differential operator P with adjoint P' whose fundamental solution at x is denoted $(E')^x$, then:

$$u(x) = \langle u, \delta_x \rangle = \langle u, P(E')^x \rangle = \langle Pu, (E')^x \rangle = \langle f, (E')^x \rangle$$

[a] multiplication can't be made associative: $(\delta \cdot x)pv(1/x) \neq \delta(x \cdot pv(1/x))$

Proposition 1.6 (**Translation for Constant Coefficient Fundamental Solution**). *If P is a constant coefficient linear scalar PDE*[a] *and $PE_0(x) = \delta_0(x)$. Then (1) $P_x E_0(x-y) = \delta_y(x)$ and (2) $P'_y E_0(x-y) = \delta_y(x)$.*

In other words, $E_y = E_0(x-y)$ and $(E')^x(y) = E_0(x-y)$

Theorem 1.3 (**Representation via Fundamental Solutions**). *If $Pu = f$ with P a constant coefficient linear scalar PDE. Then if f has compact support, then $u = E_0 * f$. If u has compact support, then $u = E_0 * Pu$.*

Example 1.1. $\partial_x^k(\frac{x^k}{k!}H(x)) = \delta_0(x)$ *(where H is the Heaviside function).*

Using this and the representation formula, we get:

Theorem 1.4 (**Taylor's formula with integral remainder**).

$$u(x) = \sum_{j=0}^{N-1} \frac{\partial^j u(a)}{j!}(x-a)^j + \frac{1}{(N-1)!}\int_a^x \partial^N u(y)(x-y)^{N-1} dy$$

Note that the integral is $\partial^N u 1_{[a,b]} * x^N H$

1.1.4 Structure Theorems for Distributions

Theorem 1.5 (**Order of Compact Supported Distribution**). *If $u \in \mathcal{E}'(\mathbb{R}^n)$, then it has finite order.*

Proof. This is almost immediate from Proposition 1.2. □

Theorem 1.6 (**Distribution Supported at a point**). *If $u \in \mathcal{D}'(\mathbb{R})$ is supported only at x_0, then $u = \sum_{|\alpha|\le N} c_\alpha \partial^\alpha \delta_{x_0}$*

Proof.

1. By Theorem 1.5, u has order N,
2. for each $\varphi \in C_0^\infty$, Taylor expand to get: $\varphi = \sum_{|\alpha|\le N} a_\alpha x^\alpha + \varphi_N$, where $\partial^\alpha \varphi_N(0) = 0$ for all $|\alpha| \le N$
3. it can be shown that $\langle u, \varphi_N \rangle = 0$
4. so u will only detect a_α which depend on derivatives of φ at 0.

□

Theorem 1.7 (**Structure Theorem of Compactly Supported Distribution**). *IF u has compact support, then $u = \sum_{|\alpha|\le N} \partial^\alpha f_\alpha$ for $f_\alpha \in C^0$*

Theorem 1.8 (**Structure Theorem of Distributions**). *If $u \in \mathcal{D}'$, then there exist $f_\alpha \in C^0$ that are locally finite (I think) so that $u = \sum_\alpha \partial^\alpha f_\alpha$.*

[a] this gives existence of fundamental solutions by Malgrange–Ehrenpreis

1.1.5 Homogeneous Distribution

Definition 1.8 (Homogeneous Distribution)**.** *A distribution $u \in \mathcal{D}'(\mathbb{R}^n)$ has order a if for all $\varphi \in C_0^\infty(\mathbb{R}^n)$, $\lambda > 0$*

$$\lambda^{-a} \langle u, \varphi(\lambda \cdot) \rangle = \lambda^d \langle u, \varphi \rangle$$

Define u_λ as $\langle u_\lambda, \varphi \rangle = \lambda^{-d} \langle u, \varphi(\lambda^{-1} \cdot) \rangle$ (if u was a function, $u_\lambda(x) = u(\lambda x)$). Then an equivalent definition of homogeneous would be:

$$\langle u_\lambda, \varphi \rangle = \lambda^a \langle u, \varphi \rangle$$

Example 1.2. *$\delta \in \mathcal{D}'(\mathbb{R}^n)$ has degree $-n$*

Proposition 1.7 (**Properties of Homogeneous Distributions**)**.**

1. *if u is homogeneous of degree a, then $\partial^\alpha u$ is homogeneous of degree $a - |\alpha|$*
2. *if $u \in \mathbb{R}^n \setminus \{0\}$ is homogenous, than it has a unique extension to $\mathbb{R}^n$ we require $\mathbb{Z} \ni a > -d - 1$*
3. *$\lambda \frac{d}{d\lambda} \langle u_\lambda, \varphi \rangle = a \langle u_\lambda, \varphi \rangle$*

1.2 Four Important PDE

Reference: Evans 2.2-2.4. Laplace's equation: fundamental solution, mean value property, maximum principle, energy methods, Harnack inequality; the heat equation: fundamental solution, regularity/smoothing, maximum principle, energy methods; the wave equation: fundamental solution, finite propagation speed, Huygens' principle, energy methods

1.2.1 Laplace's Equation

To compute the fundamental solution of $-\Delta$, first note the Laplacian is rotationaly invariant (Exercise!), so we expect $E_0 = E_0(r)$ to be radial. We can take the Laplacian and get the delta function. The trick is to pair it with the correct thing, which is a ball of radius r:

$$1 = \left\langle \delta_0(x), 1_{B_r(0)} \right\rangle = \left\langle -\Delta E_0, 1_{B_r(0)} \right\rangle = \int \nabla E_0 \cdot \nabla 1_{B_r(0)} = \int \nabla E_0 \cdot \nu_{B_r(0)} dS_{B_r(0)} \tag{1}$$

$\nabla E_0 \cdot \nu_{B_r(0)} = \sum \partial_{x_j} E_0 \frac{x_j}{r}$. Chain rule: $\partial_{x_j} E_0(r(x)) = \partial_r E_0 \partial_{x_j} r = \partial_r E_0 \frac{x_j}{r}$. So $\nabla E_0 \cdot \nu_{B_r(0)} = \partial_r E_0(r) \frac{r^2}{r^2} = \partial_r E_0(r)$. So (1) becomes:

$$1 = \partial_r E_0(r) \int_{\partial B_r(0)} dS_{\partial B_r(0)} = \partial_r E_0(r) |\partial B_r(0)|$$

Therefore $E_0(r) = \int_0^r |\partial B_s(0)|^{-1} ds = c_d \int_0^r \frac{1}{s^{n-1}} ds = c_d \frac{1}{r^{n-2}}$

Theorem 1.9 (Fundamental Solution of Laplace Equation)**.**

$$E_0(x) = \begin{cases} c_2 \log |x| & n = 2 \\ c_n |x|^{2-n} & n > 2 \end{cases}$$

Note that $c\Delta \log|x| = \delta_0(x) = c\bar{\partial}_z\partial_z \log|z| = c\bar{\partial}_z|z|^{-1}$, so we have the fundamental solution for the dbar operator as well.

Note that E_0 is a distribution because it is locally integrable.[a]

Theorem 1.10 (Regularity of Harmonic Functions). *If $\Delta u = 0$, then $u \in C^\infty$*

Proof. Fix x, let $\chi \in C_0^\infty$ be a cuttoff, $\chi u = \chi u * \delta = \chi u * (-\Delta E_0) = (-\Delta(\chi u)) * E_0$. Note $\Delta(\chi u) = (\Delta\chi)u + 2\nabla\chi \cdot \nabla u$. So we get:

$$u(x) = -\int ((\Delta\chi)u + \nabla u \cdot \nabla\chi)E_0(x-y)dy$$

the integrand is supported away from the singular support of $E_0(x - \cdot)$ □

This can be generalized. If P is any differential operator whose fundamental solution has singular support at the origin, then $Pu = 0$ implies $u \in C^\infty$.

Theorem 1.11 (Derivative Control of Harmonic Functions). *If $\Delta u = 0$, then $|\partial^\alpha u| \le \frac{C}{r^{d+|\alpha|}} \int_{B_r(x)} |u(y)|dy$ for all $r > 0$.*

Proof.

1. use above representation to get $|\partial^\alpha u(x)| \le \int |(\Delta\chi u + 2\nabla\chi \cdot \nabla u)\partial_x^\alpha E_0(x-y)|dy$
2. let χ be 1 on $B_{r/2}(x)$ and supported on $B_r(x)$ with $|\partial^\alpha \chi| \le \frac{C}{r^{|\alpha|}}$
3. On support of integral, $|\partial_x^\alpha E_0(x-y)| \le \frac{C}{r^{d-2+|\alpha|}}$
4. move derivative off u, use control on χ to get get final result.

□

Theorem 1.12 (Mean Value Property of Harmonic Functions). *If $\Delta u = 0$, then for all $r > 0$, $u(x) = \fint_{\partial B_r(x)} u(y) dS_{\partial B_r(x)}$*

There are two tricks to this proof: (1) add a constant to fundamental solution so it vanishes where we want (2) remember what the derivative of the fundamental solution is.

Proof.

1. $u(x) = \int 1_{B_r(x)}(y)u(y)(-\Delta_y E_0(x-y))dy$
2. Integrate by parts, one term is $I_1 = -\int_{\partial B_r(x)} u(y)\nu_{\partial B_r(x)} \cdot \nabla E_0(x-y)dS_{B_r(x)}(y)$
3. The other term we integrate by parts, use $\Delta u = 0$, to get $I_2 = \int_{\partial B_r(x)} \nabla u(y) \cdot \nu_{B_r(x)}(y)E_0(x-y)dS_{\partial B_r(x)}$
4. Since the fundamental solution is radially symmetric, we can add a constant to get $E_0(x-y)|_{\partial B_r(x)} = 0$ so that $I_2 = 0$
5. The first term is $\int_{\partial B_r(x)} u(y)E_0'(r)dS = |\partial B_r(x)| \fint_{\partial B_r(x)} u(y)dS$

□

The same conclusion is true for solid balls, found by integrating this result.

[a] I got held up trying to compute $\int |x|^{-\alpha}$ using spherical coordinates. Computing the Jacobian is hard, but to remember the power of r in the volume element, just compute the volume of a sphere. We should have $dx = r^{n-1}$. Then we integrate $\int r^{n-1-\alpha}$, so we require $n - 1 - \alpha > -1$

1.2.1.1 Boundary Value Problem

Theorem 1.13 (Representation Formula for Laplace on Bounded Domain). *For $u \in C^\infty(\bar{U})$ with U an open bounded set in $\mathbb{R}^n$, then:*

$$u(x) = -\int_U \Delta u(y) E_0(x-y)dy + \int_{\partial U} (u(y)\nu_{\partial U} \cdot \nabla_y E_0(x-y) + \nabla u(y) \cdot \nu_{\partial U} E_0(x-y))\, dS(y)$$

It is important to remember that the derivative on E_0 is on the y variable.

Proof. $u = 1_U u * -\Delta E_0 = -1_U u * \partial_j \partial_j E_0 = -((\partial_j 1_U)u + 1_U(\partial_j u)) * \partial_j E_0$. First term we keep as $\int_{\partial U} u\nu \cdot \nabla E_0$. Second term, we expand to $-(\partial_j 1_U)(\partial_j u) * E_0 - 1_U(\partial_j \partial_j u) * E_0$. First one is $\int_{\partial U} \nu \cdot \nabla u E_0(x-y)$, second is $-\int_U \Delta u E_0(x-y)dy$ □

This gives us a mean-value theorem for functions which are only harmonic in a bounded region. Also we only need $u \in C^2$ (and probably even less) for this theorem to work.

Theorem 1.14 (Strong Maximal Property of Harmonic Functions). *Let u be Harmonic on U (a bounded, open, connected set). Then if there exists $x_0 \in U$ with $u(x_0) = \max_{\bar{U}} u(x)$, then u is constant on $\bar{U}$. This implies that harmonic functions achieve their maximum on the boundary (weak maximal principal).*

Proof. By mean-value property, $M = u(x_0) = ⨍_{\partial B} u(y)dy \leq M$ for all balls contained in U. This implies $u = M$ on the boundary. Use connectedness to get result. □

This gives uniqueness of the Dirichlet problem..

Theorem 1.15 (Harnack's Inequality). *If u is non-negative harmonic on an open set U. And V is an open, connected set with $\bar{V} \subset U$. Then there exists C (not depending on u) such that $\max_{x \in \bar{V}} u \leq C \min_{x \in \bar{V}} u$*

Proof. Idea: use balls, compactness, and mean value property.

1. let $r \ll dist(\bar{V}, \partial U)$
2. for $x, y \in V$, $|x-y| < r$, $u(x) = ⨍_{B_r(x)} u \leq |B_r(x)|^{-1} \int_{B_{2r}(y)} u = 2^d ⨍_{B_{2r}(y)} u = 2^d u(y)$
3. similarly, $u(x) = ⨍_{B_{2r}(x)} u \geq |B_{2r}(x)|^{-1} \int_{B_r(y)} u = 2^{-d} u(y)$
4. use compactness to cover $\bar{V}$ with N balls of radius r, to get $2^{-dN} u(y) \leq u(x) \leq 2^{dN} u(y)$ for all $x, y \in V$
5. take sup over x, and inf over y.

□

1.2.1.2 Green's Functions

To solve $\begin{cases} -\Delta u = f & U \\ u = g & \partial U \end{cases}$ split and solve $\begin{cases} -\Delta u = 0 & U \\ u = g & \partial U \end{cases}$ and $\begin{cases} -\Delta v = f & U \\ v = 0 & \partial U \end{cases}$. The first one is equivalent to solving $\begin{cases} -\Delta \tilde{u} = -\Delta \tilde{g} & U \\ \tilde{u} = 0 & \partial U \end{cases}$ with $\tilde{g}$ an extension of g. Then set $u = \tilde{u} + \tilde{g}$.

Conclusion: to solve inhomogenenous, nontrivial boundary condition Laplace equation, it suffices to solve inhomogeneous, trivial boundary condition.

Definition 1.9 (Green's Function). *A Green's function for a bounded set U is $G(\cdot, y) \in \mathcal{D}'(U) \cap C^1(\bar{U} \setminus \{y\})$ for all $y \in U$ that solves*

$$\begin{cases} -\Delta_x G(x,y) = \delta(x-y) & x, y \in U \\ G(x,y) = 0 & x \in \partial U \end{cases}$$

Theorem 1.16 (Basic Properties of Green's Functions). *1. $G \in C^\infty(U \times U \setminus \{x = y\})$*

2. $G(x,y) = G(y,x)$ for $x \neq y$, $x, y \in U$

3. G is unique.

A clean definition is $-\Delta_x G = \delta_y$, $-\Delta_y G = \delta_x$, and G vanishes if either variable is on the boundary.

Theorem 1.17 (Poisson Integral Formula). *If U is a C^1 domain and $u \in C^\infty(\bar{U})$, then:*

$$u(x) = \int_{\partial U} u(y) \nu \cdot \nabla_y G(x,y) dS(y) + \int_U (-\Delta u) G(x,y) dy$$

So to solve our inhomogeneous Dirichlet problem $-\Delta u = f$, then $u(x) = \int_U G(x,y) f(y) dy$. So $1_U G(x,y)$ is the kernel of the pseudodifferential operator $(-\Delta)^{-1}$. It makes sense it is smooth off the diagonal.

Example 1.3 (Green's Function for Half Space). *If $U = \{x \in \mathbb{R}^n : x_n > 0\}$, and $E_0(x)$ is the fundamental solution of the Laplacian. Then $G(x,y) = E_0(x-y) - E_0(x - \bar{y})$ is the Green's function for U (where $\overline{(y_1, \ldots, y_n)} = (y_1, \ldots, y_{n-1}, -y_n)$.*

This is because if $y \in \partial U$, then $y = \bar{y}$, so $G(x,y) = 0$. And $-\Delta_x G(x,y) = \delta(x-y) - \delta(x - \bar{y})$. If the support is restricted to U, then this is $\delta(x-y)$.

We can compute $\nu \cdot \nabla_y G(x,y)$ for $y \in \partial U$ as $\frac{C x_n}{|x-y|^n}$, so:

$$u(x) = c x_n \int_{y_n = 0} \frac{g(y)}{|x-y|^n} dy$$

solve $\nabla u = 0$ and $u = g$ on $x_n = 0$.

This can be rewritten as the Poisson kernel, $P_t(x) := c_d \frac{t}{(t^2+|x|^2)^{(d+1)/2}}$*, then* $u(x) = P_t(x) * f$ *solves the equation:*

$$\begin{cases} (\partial_t^2 + \Delta_x)u = 0 & t > 0 \\ u(0,x) = f(x) \end{cases}$$

Example 1.4 (**Green's Function for unit ball**)**.** *The Green function for the unit ball is* $G(x,y) = \Phi(y-x) - \Phi(|x|(y-\tilde{x}))$ *where* $\tilde{x} = \frac{x}{|x|^2}$.

This can be used to Laplace equation with boundary data on the unit ball. It turns out that $\partial_n u^y G(x,y) = C\frac{1-|x|^2}{|x-y|^n}$ *for* $|y| = 1$*. Therefore:*

$$u(x) = \frac{1-|x|^2}{n\alpha(n)} \int_{\partial B_1(0)} \frac{g(y)}{|x-y|^n} dS(y)$$

where $u(x) = g(x) \in C^0$ *for* $|x| = 1$.

We can solve this using harmonic analysis as well (for $d = 2$*). Suppose* $u(e^{i\theta}) = g(\theta)$*, then a solutions is* $u(re^{i\theta}) = \sum_{n\in\mathbb{Z}} \widehat{g}(n) r^{|n|} e^{in\theta}$*. If we let* $P_r(x) = \sum r^{|n|} e^{inx}$*, then we see that* $u(re^{i\theta}) = P_r * f$*. We call* P_r *the* ***Poisson kernel****, and is:*

$$\mathcal{P}_r(\theta) = \frac{1-r^2}{1-2r\cos\theta + r^2}$$

1.2.2 Wave Equation

We want to solve:

$$\begin{cases} \Box u(t,x) = f & \mathbb{R}^{1+d}_+ \\ u(0,x) = g \\ \partial_t u(0,x) = h \end{cases}$$

with $\Box = (-\partial_t^2 + \Delta)$

Theorem 1.18 **(Fundamental Solution for Wave Equation in 1-dimension).** *The forward fundamental solution to the wave equation in 1-dimension is* $E_+ = -\frac{1}{2}H(t-x)H(t+x)$

Proof.

1. let $u = t - x$, $v = t + x$, so $(t,x) = (\frac{1}{2}(u+v), \frac{1}{2}(v-u))$, so $\partial_u = \frac{1}{2}(\partial_t - \partial_x), \partial_v = \frac{1}{2}(\partial_t + \partial_x)$
2. So if $\Box E_0(t,x) = \delta(x,y)$, then $-4\partial_u\partial_v E_0(u,v) = \delta(x,y)$
3. $\delta(x,y) = \lim_{\varepsilon\to 0} \varepsilon^{-2}\chi(\varepsilon^{-1}t, \varepsilon^{-1}x) = \lim_{\varepsilon\to 0} \varepsilon^2 \chi(\varepsilon^{-1}\frac{1}{2}(u+v), \varepsilon^{-1}\frac{1}{2}(v-u)) = \delta(u,v)\int \chi(\frac{1}{2}(u+v), \frac{1}{2}(v-u)) = 2\delta(u,v)$
4. so $\partial_u\partial_v E_0 = \frac{-1}{2}\delta(u)\delta(v)$, so $E_0(u,v) = \frac{-1}{2}(H(u)+c_1)(H(v)+c_2)$
5. For E_0 to be supported in positive time, we require $c_1 = c_2 = 0$

□

For uniqueness, and in order to convolve with distributions, we require the technical definition:

Definition 1.10 (Forward Fundamental Solution to Wave Equation). *E_+ is a forward fundamental solution to the wave equation is*

1. $\square E_+ = \delta(t,x)$
2. $\operatorname{supp} E_+ \subset \{t \geq 0\}$
3. *if $I \subset \mathbb{R}$ is compact, then $\operatorname{supp} E_+ \cap \{(t,x) : t \in I\}$ is compact*

Theorem 1.19 (Basic Forward Fundamental Solution Properties). *If E_+ is a forward fundamental solution then (1) E_+ is unique (2) if $u \in \mathcal{D}'$ with $\operatorname{supp} u \subset \{(t,x) : t \geq C\}$, then $u * E_+$ is well defined.*

Theorem 1.20 (Representation Formula with Forward Fundamental Solution). *For $\varphi \in C^\infty(\mathbb{R}_{\geq 0} \times \mathbb{R}^d)$:*

$$\varphi(t,x) = \square\varphi * E_+ - \partial_t(E_+ * \delta_{t=0}\varphi) - E_+ * (\delta_{t=0}\partial_t\varphi)$$

Theorem 1.21 (Wave Equation Representation Formula $1d$). *If $u \in C^\infty(\mathbb{R}^2_+)$ solves* $\begin{cases} \square u = f \\ u(0,x) = g \\ \partial_t u(0,x) = h \end{cases}$ *, then:*

$$u(t,x) = \frac{1}{2}\int_0^t \int_{t-s-x}^{t-s-x} f(s,y)dsdy - \frac{1}{2}(g(t+x) - g(t-x)) - \frac{1}{2}\int_{t-x}^{t+x} h(y)dy$$

Theorem 1.22 (Forward Fundamental Solution to Wave Equation). *The forward fundamental solution to the wave equation is:*

$$E_+(t,x) = -\frac{1}{2\pi^{\frac{d-1}{2}}} 1_{t\in[0,\infty)} \chi_+^{\frac{-d+1}{2}}(t^2 - |x|^2)$$

with $\chi_+^a = \Gamma(a+1)^{-1} 1_{x>0} x^a$ for $a > -1$, for $k \in \mathbb{Z}$, $\chi_+^{-k} = \delta_0^{(k-1)}$ and $\chi_+^{-\frac{1}{2}-k} = \frac{1}{\sqrt{\pi}}\frac{d^k}{dx^k}(H(x)x^{-1/2})$

Proof.

1. (a) by symmetries, E_+ is homogenous (b) $\square E_+ = \delta$ has degree $-d-1$, so E_+ has degree $-d+1$ (c) $\square$ is invariant under Lorentz boost $t^2 - |x|^2$ (degree 2) and E_+ supported in forward time so $E_+ = \chi(t^2 - |x|^2)1_{(0,\infty)}(t)$ with χ some homogeneous distribution of order $\frac{1}{2}(-d+1)$ supported in $[0,\infty)$ note by uniqueness of homogenous distributions, only need to define it on $\mathbb{R}^{1+n} \setminus \{0,0\}$.
2. Computing $\square E_+$ away from origin, gets crazy cancellation to zero (there is a trick with derivatives of homogenous distributions that is used), thus it is supported on the origin. Therefore it is a constant times the delta function.

3. For $a \in \mathbb{C}$, define $\chi_+^a = c_a 1_{x>0} x^a$ (this is L^1_{loc} for $\mathfrak{R}(a) > -1$) and homogeneous of degree a. For $a \in \mathbb{R}$, $(\chi_+^a)' = c_a 1_{x>0} a x^{a-1} = \frac{c_a a}{c_{a-1}} \chi_+^{a-1}$, use this to inductively define χ_+^a. for all $a \in \mathbb{C}$

4. letting $c_a a = c_{a-1}$ gets everything to be 1. This is satisfied by $c_a = 1/\Gamma(a+1)$, so $(\chi_+^a)' = \chi_+^{a-1}$

5. $\chi_+^0 = H(x)$, so $\chi_+^{-1} = \delta_0(x)$, so $\chi_+^{-k} = \delta_0^{(k-1)}$.

6. $\chi_+^{-1/2} = \pi^{-1/2} H(x) x^{-1/2}$, so $\chi_+^{-\frac{1}{2}-k} = \frac{1}{\sqrt{\pi}} \frac{d^k}{dx^k}(H(x)x^{-1/2})$

7. there is a computation to get the proper constant.

□

Dimension	$E_+(t,x)$
1	$c_1 1_{t\geq 0} H(t^2 - x^2)$
2	$c_2 1_{t\geq 0} \frac{H(t^2-\|x\|^2)}{\sqrt{t^2-\|x\|^2}}$
3	$c_3 1_{t\geq 0} \delta_0(t^2 - \|x\|^2)$

Theorem 1.23 (Huyghen's Principals).

1. *(Weak Huygen's principal / finite speed of propagation) If (t,x) is such that $u(0,y) = u_t(0,y) = 0$ for all $|x-y| \leq t$ and $\Box u(s,y) = 0$ for all $s \in (0,t)$ and $|t-x| < t-s$, then $u(t,x) = 0$.*

2. *(Strong Huygen's Principal) If $d \geq 3$ and is odd and (t,x) is such that $u(0,y) = u_t(0,y) = 0$ for all $|x-y| = t$ and $\Box u(s,y) = 0$ for all $s \in (0,t)$ and $|y-x| = t-s$, then $u(t,x) = 0$*

Proof. First follows from $suppE_+ \subset \{(t,x) : |x| \leq t\}$, second follows from the fundamental solutions for $d \geq 3$, having $\chi = \delta^{(k)}$, which are supported on $\{(t,x) : |x| = t\}$. □

Theorem 1.24 (d'Alembert's Formula). *In one dimension, we have the following solution to the 1-dimensional wave equation :*

$$u(t,x) = \frac{1}{2}(g(x+t) + g(x-t)) + \frac{1}{2}\int_{x-t}^{x+t} h(y)dy$$

Theorem 1.25 (Poisson's Formula). *In two dimensions, the solution to the wave equation is:*

$$u(x,t) = \frac{1}{2} \fint_{B(x,t)} \frac{tg(y) + t^2 h(y) + tDg(y)\cdot(y-x)}{(t^2 - |y-x|^2)^{1/2}} dy$$

Theorem 1.26 (Kirchoff's Formula). *In 3-dimensions, the solution to the wave equation is of the form:*

$$u(x,t) = \fint_{\partial B(x,t)} th(y) + g(y) + Dg(y)\cdot(y-x) dS(y)$$

Theorem 1.27 (Duhamel's Principal For Wave Equation). *To solve $\Box u = f$ with zero initial data, then $u = \int_0^t u(t,x;s)ds$ where $u(t,x;s)$ solves:*

$$\begin{cases} \Box u = 0 & (s,\infty)\times\mathbb{R}^n \\ u(x,s) = 0 \\ u_t(x,s) = f(x,s) \end{cases}$$

If $d = 1$, then:

$$u(t,x) = \frac{1}{2}\int_0^t \int_{x-t+s}^{x+t-s} f(y,s)dyds$$

if $d = 3$, then:

$$u(t,x) = \frac{1}{4\pi}\int_{B(x,t)} \frac{f(y,t-|y-x|)}{|y-x|}dy$$

1.2.2.1 Energy Methods

We can prove uniqueness to the wave equation using the energy functional $e(t) = \int_U u_t^2 + |\nabla_x u|^2 dx$

1.2.3 Heat Equation

The heat equation is:

$$\begin{cases} u_t - \Delta u = f \\ u(0,x) = g \end{cases}$$

Theorem 1.28 (Fundamental Solution to Heat Equation). *The forward fundamental solution to the heat equation is*

$$\Phi(t,x) = 1_{t>0}\frac{1}{(4\pi t)^{n/2}}e^{-\frac{|x|^2}{4t}}$$

Proof.
1. Fourier transform pde: $\widehat{u}' = |\xi|^2\widehat{u}$ so $\widehat{u} = \widehat{u}(0)e^{-t|\xi|^2}$
2. Fundamental solution is $\mathcal{F}^{-1}[e^{-t|\xi|^2}]$.
3. This is $(2\pi)^{-n}\int e^{-t|\xi|^2}e^{ix\xi}d\xi$. Let $u = \sqrt{2t}\xi$.
4. Integral becomes $(2\pi)^{-n}(2t)^{-n/2}\int e^{-u^2/2}e^{ix\frac{u}{\sqrt{2t}}}du$, integral is Gaussian, becomes $(2\pi)^{n/2}e^{-\frac{x^2}{2\cdot 2t}}$
5. after everything, get $e^{-x^2/(4t)}$ with constant $(2\pi)^{-n/2}(2t)^{-n/2}$

□

Theorem 1.29 (Existence and Uniqueness of Homogenous Heat Equation with L^2 data). *If $g \in L^2$, then there exists a solution to the heat equation $u \in C_t([0,\infty), L^2)$ that is unique and $\|u(t,x)\|_{L^2_x} \le \|g\|_{L^2}$*

Proof. 1. Taking the Fourier transform, $\widehat{u}(t,\xi) = e^{-t|\xi|^2}\widehat{g}(\xi)$. The Fourier inversion formula gives existence (RHS is $\mathcal{S}L^2 \subset L^2$)

2. $\|u(t,x)\|_{L^2_x} = \|\widehat{u}(t,\xi)\|_{L^2_\xi} \le \|\widehat{g}\|_{L^2_\xi} = \|g\|_{L^2}$

3. $\|u(t,x) - u(s,x)\|_{L^2} = \left\|(e^{(s-t)|\xi|^2} - 1)\widehat{g}(\xi)\right\|_{L^2} \to 0$ by the dominated convergence theorem.

4. if $u = v$ solve this, let $w = u - v$, then $(\partial_t + |\xi|^2)\widehat{w} = 0$ so $\partial_t(e^{t|\xi|^2}\widehat{w}) = 0$, so $e^{t|\xi|^2}\widehat{w}$ is constant, but it is zero if $t = 0$, so it is zero. By Fourier inversion, $w = 0$.

□

Note that $\Phi(t,x)$ is an approximate identity sequence with parameter t. So if $u(t,x) = \Phi(t,x) *_x g$, then we see many things:

1. if $g \in C_b^0$, then $u(t,x) \xrightarrow{t\to 0} g(x)$ uniformly on compact sets.

2. if $g \in L^p$, then $u(t,x) \xrightarrow{t\to 0} g(x)$ in L^p

3. for $t > 0$, $u(t,x)$ is smooth in space and time if the initial data is a tempered distribution.

Theorem 1.30 (Solution of Nonhomogeneous Heat Equation). *To solve the nonhomogeneous heat equation, $(\partial_t - \Delta)u = f$ for $t > 0$ and $u = 0$ for $t = 0$, the trick is to write $u(t,x) = \int_0^t u(t,x;s)ds$ where $u(t,x;s)$ solves the initial value heat equation:*

$$\begin{cases} (\partial_t - \Delta)u = 0 & t > s \\ u(s,x) = f(s,x) \end{cases}$$

which has solution $u(t,x;s) = \Phi(t-s,x) * f(s,x)$ (of course this is not justified, but it works), so:

$$u(t,x) = \int_0^t \int_{\mathbb{R}^d} f(s,y)\Phi(x-y,t-s)dxds$$

Theorem 1.31 (Heat Equation Strong Maximum Principal). *If $U \subset \mathbb{R}^n$ is open, $U_T = U \times (0,T]$, $\Gamma_T = \bar{U}_T \smallsetminus U_T$. If $u \in C^{2_x,1_t}(U_T) \cap C(\bar{U}_T)$ solves the heat equation, then $\max_{\bar{U}_T} u = \max_{\Gamma_T} u$. If U is connected and u attains a maximum in U_T: $u(t_0,x_0) = \max_{\bar{U}_T} u$, then u is constant on $\bar{U}_{t_0}$ NOTE: constant on **earlier** times only*

Remember: weak does not imply strong because we could have a function that looks like a w.

The proof requires the following fact that must be memorized:

Theorem 1.32 (Heat Equation Mean Value Property). *If u solves the heat equation, then:*

$$u(t,x) = \frac{1}{4r^d} \iint\limits_{E_r(t,x)} u(s,y)\frac{|x-y|^2}{(t-s)^2}dsdy$$

*where $E_r(t,x) : \{(s,y) : s \le t \ \Phi(t-s,x-y) \ge 1/r^d\}$ is the **heat ball** of radius r*

Proof. 1. wlog $(x,t)=(0,0)$, let $f(r)=r^{-n}\iint_{E_r(0)}\frac{|y|^2}{s^2}dyds$

2. use that $\frac{1}{4\pi r^d}\iint_{E_1(0)}\frac{|y|^2}{s^2}=1$ [a], then by DCT $|f(r)-u(0,0)|\to 0$, therefore $f(r)\to 0$ as $r\to 0$

3. compute $f'(r)=0$

□

The weak maximal theorem can be proven much more simply (and can be generalized to parabolic equations):

Proof. Suppose $u(t_0,x_0)=\max_{\bar{U}_T}$ with $(t_0,x_0)\in U_T$, then first pretend that $u_t-\Delta u<0$. First, we require $\partial_t u\geq 0$ otherwise we can just go backwards in time. But we also require $\Delta u\leq 0$, this is a contradiction.

Consider $v_\varepsilon=u-\varepsilon t$, then $(\partial_t-\Delta)u_\varepsilon<0$. So if u solve the heat equation and has a maximum attained inside U_T, then since v_ε goes to u, we can use the above to get a contradiction. □

Proof. Here is a proof of the strong maximal property:

1. let (t_0,x_0) be an interior maximum with value, by mean value property $u=M$ on a small heat ball around it (this uses the fact that $(4r)^{-d}\iint_{E_r(t,x)}\frac{|x-y|^2}{(t-s)^2}dsdy=1$

2. for any earlier point, connect a line segment between the two, u must be M on this line segment (if not use continuity, get largest time this fails, extend heat ball, get contradiction)

3. any previous point can be connected via finitely many line segments, get $u=M$ on the whole previous time.

□

1.2.3.1 Regularity

Regularity for unbounded domains are trivial via the Fourier transform. However, for bounded domains it's a little trickier

Theorem 1.33 (Heat Equation Regularity on Bounded Domains). *If u is a classical solution of the heat equation in the bounded domain U_T, then $u\in C^\infty(U_T)$*

Proof. 1. assume u is smooth, then repeat this argument with mollifiers

2. fix (x_0,t_0), consider the cylinders C,C',C'' with radius $r,3r/2,r/2$ and heights the square of these the radius squared. Let ξ be a cuttoff function supported on C, and identically 1 on C'.

[a] I wasted a lot of time trying to verify this, I'm not quite sure how to do it

3. let $v(t,x) = \xi(t,x)u(t,x)$, this is zero at time 0, and it can be computed that $(\partial_t - \Delta)v = \xi_t u - 2D\xi \cdot Du - u\Delta\xi := f$

4. now we solve the nonhomogeneous heat equation, and use uniqueness to see that $v(t,x) = \int_0^t \in_{\mathbb{R}^n} \Phi(t-s, x-y) f(y,s) dyds$

5. expand this, integrate by parts to avoid derivatives falling on u, then by support properties of ξ and singsupp Φ, we see that the resulting thing is smooth for $x \in C''$

□

1.2.3.2 Energy Estimates

The correct energy for a solution to the wave equation is $e(t) = \int_U u^2(x,t)dx$. This is because:

$$\dot{e}(t) = \int_U u u_t dx = \int_U u \Delta u dx = -\int_U |\nabla u|^2 \leq 0$$

1.3 Characteristic Equations

Reference: Evans 3.2

Derivation, boundary conditions, local solutions

1.3.1 Derivation

Suppose we are solving a first order, scalar nonlinear PDE

$$\begin{cases} F(x, u(x), \nabla u) = 0 & x \in U \subset \mathbb{R}^d \\ u(x) = g(x) & x \in \Gamma \subset \partial U \end{cases} \tag{2}$$

with F and g smooth functions. Let $(x, u, \nabla u) = (x, z, p)$

Theorem 1.34 (Characteristic ODEs). *If u is a smooth solution to (2), then on a curve $x = x(s)$ such that $\dot{x}_i = \partial_{p_i} F$ $(i = 1, \dots, n)$, then $z(s) = u(x(s))$, $p(s) = (\partial_i u)(x(s))$ will satisfy:*

$$\dot{p}_i = -\partial_{x_i} F - (\partial_z F) p_i$$
$$\dot{z}(s) = \sum_{j=1}^n p_i \partial_{p_i} F$$

To remember: p_i is just negative the first two terms of $\partial_{x_i} F$, $\dot{x}_i$ has to just be memorized, $\dot{z}$ can be easily derived from the other two.

1. differentiate F with respect to x_i:

$$0 = \partial_{x_i} F = (\partial_{x_i} F) + (\partial_z F) p_i + \sum_{j=1}^n (\partial_{p_j} F) \partial_{x_i} \partial_{x_j} u \tag{3}$$

2. to rewrite the third term, suppose $x = x(s)$ is a curve, then:

$$\partial_s p_i(s) = \partial_s(\partial_{x_i} u(x(s))) = \sum_{j=1}^{n} (\partial_{x_i}\partial_{x_j} u)\dot{x}_j$$

3. If $\dot{x}_j = \partial_{p_j} F$, then (3) becomes:

$$0 = (\partial_{x_i} F) + (\partial_z F)p_i + \dot{p}_i$$

4. lastly:

$$\dot{z}(s) = \sum_{j=1}^{n} p_i \dot{x}_i = \sum_{j=1}^{n} p_i \partial_{p_i} F$$

Idea: differentiate with respect to x_i, use the chain rule. Get rid of the second derivative falling on u by cleverly letting $x = x(s)$ satisfy $\partial_{p_i} F = \dot{x}_i$

1.3.2 Boundary Conditions

Definition 1.11 **(C^k Boundary).** *$U \subset \mathbb{R}^n$ open and bounded is said to have C^k boundary if for all $x_0 \in \partial U$, there exists $R > 0$ and $\gamma \in C^k(\mathbb{R}^{n-1};\mathbb{R})$ such that $B_R(x_0) \cap U = \{x \in B_R(x_0) : x_n > \gamma(x_1, \ldots, x_{n-1})\}$*

Example 1.5. *Consider $\mathcal{S}^3$ and $x_0 = (0,0,-1)$, then we can have $R = 1/3$ and $\gamma = -\sqrt{1 - x^2 - z^2}$.*

Theorem 1.35 **(Straightening The Boundary).** *If U is open, bounded, with C^k boundary, then there exist $\Phi : \mathbb{R}^n \to \mathbb{R}^n$ and $\Psi = \Phi^{-1}$, both C^k such that $\det D\Phi = \det D\Psi$ and for all $x_0 \in U$, there exists $R > 0$ such that*

$$\Phi(U \cap B_R(x_0)) = \{y \in B_R(x_0) : y_n > 0\}$$

It now suffices to prove existence for characteristic equations that have a flat boundary. To see this:

1. let u solve (2). Fix a point on the boundary, get Φ and Ψ from Theorem 1.35. Define $v(y) = u(\Psi(x))$ so $u(x) = v(\Phi(x))$
2. $\partial_{x_i} u = \sum \partial_{y_j} v \partial_{x_i} \Phi_j = (\nabla v \cdot D\Phi)_i$, therefore $\nabla u = \nabla v \cdot D\Phi$
3. So $0 = F(x, u(x), \nabla u) = F(\Psi(y), v(y)\nabla v \cdot D\Phi) := G(y, v, \nabla v)$

Theorem 1.36 **(Noncharacteristic Boundary Conditions).** *If $D_p F(x_0, z_0, p_0) \cdot \nu(x_0) \neq 0$, then there exists a function $q : \mathbb{R}^n \to \mathbb{R}^n$ defined locally around x_0 such that $p(x) = q(x)$ satisfies:*

$$\begin{cases} u(x) = g(x) & x \in \Gamma \\ p(x) = \partial_\mu g(x) & x \in \Gamma,\ \mu \perp \nu \\ F(x, u, p) = 0 \end{cases}$$

Proof. (for flat boundary)

1. $u(x) = g(x)$ along Γ. And $p_i(x) = \partial_{x_i} u(x) = \partial_{x_i} g(x)$ for $i = 1, \dots, n-1$
2. Solve for p_n using $F(x, u, p) = 0$, using the intermediate value theorem ($\partial_{p_n} F \neq 0$).
3. now we need this to work in a neighborhood of x_0
   (a) Let $G(x,p) : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}^n$ be: $G_i(x,p) = p_i - \partial_{x_i} g(x)$ and $G_n(x,p) = F(x, g(x), p)$.
   (b) $G(x_0, p_0) = 0$,

$$D_p G(x_0, p_0) = \begin{pmatrix} 1 & 0 & \cdots & 0 & 0 \\ 0 & 1 & \cdots & \vdots & 0 \\ 0 & \ddots & \ddots & 0 & \vdots \\ 0 & 0 & \cdots & 1 & 0 \\ \partial_{p_1} F & \partial_{p_2} F & \cdots & \partial_{p_{n-1}} F & \partial_{p_n} F \end{pmatrix}$$

   the determinant is $\partial_{p_n} F \neq 0$, therefore we can locally solve for p continuously in terms of x.

□

Basically, we can locally solve our PDE if $D_p F \cdot \nu(x_0) \neq 0$, because this ensures we can find initial velocities that work.

1.3.3 Local Solutions

Theorem 1.37 (Characteristic Equations Local Existence). *Given the PDE*

$$\begin{cases} F(x, u(x), \nabla u) = 0 & x \in U \subset \mathbb{R}^n \\ u(x) = g(x) & x \in \Gamma \subset \partial U \end{cases}$$

With Γ flat. If there is $x_0 \in \Gamma$ such that $(\partial_{p_n} F)(x_0, g(x_0), \nabla g(x_0)) \neq 0$ (to get $\partial_n g(x_0)$, we have to solve $F(x, u(x), Du(x)) = 0$ (this is somewhat circular?)) then:

There exists $I \subset \mathbb{R}$ containing 0, a neighborhood of x_0 in $W \subset \Gamma$ and $V \subset \mathbb{R}^n$ such that for each $x \in V$ there is a unique $s \in I$, $y \in W$ such that the curve $x(s)$ solving $\partial_s x = \partial_p F$, so that $x(y, s) = x$. Inverting this, we have $y(x)$ and $s(x)$ are C^2.[a]

Now for each $x \in V$, we get $y(x) \in \Gamma$ and $s(x)$. Then define $u(x) = z(y(x), s(x))$, $p(x) = p(y(x), s(x))$ which come from existence of ODEs. The values of p in the normal direction are uniquely determined near x_0 via the implicit function theorem.

Then it is a matter of calculus to ensure that $u(x)$ defined this way actually solves our PDE.

[a]this is an easy application of the inverse function theorem.

1.4 Sobolev Spaces

Reference: Evans 5.2 - 5.8

Basic definitions, approximation, extensions, traces, Gagliardo-Nirenberg-Sobolev inequality, Morrey's inequality, general Sobolev inequality, compactness, Poincaré inequality

1.4.1 Basic Definitions

Definition 1.12 (Sobolev Space). *Given U and open subset of $\mathbb{R}^n$, we define:*

$$W^{k,p}(U) = \{f \in L^p(U) : \partial^\alpha f \in L^p \ \forall \ |\alpha| \le k\}$$

where derivatives are taken in the sense of distributions.

The norm we put on $W^{k,p}$ is:

$$\|f\|_{W^{k,p}} = \sum_{|\alpha| \le k} \|\partial^\alpha f\|_p$$

This norm is makes $W^{k,p}$ complete and hence a Banach space.

Definition 1.13 ($W_0^{k,p}(U)$). *For U and open set, we define $W_0^{k,p}(U)$ as the completion of $C_0^\infty(U)$ with respect to the metric $\|\cdot\|_{W^{k,p}(U)}$*

This is strictly smaller than $W^{k,p}$. When $p = 2$, we define $W^{k,p} = H^k$. An equivalent norm we can apply to H^k is:

$$\|f\|_{H^k} = \left\| \langle \xi \rangle^k \widehat{f} \right\|_{L^2}$$

and this works for $k \in \mathbb{R}$. $\langle \xi \rangle$ can mean $(1 + |\xi|)^k$ or $(1 + |\xi|^2)^{k/2}$

An equivalent definition of $H^s(\mathbb{R}^n)$ is:

$$H^s(\mathbb{R}^n) = \left\{ u \in \mathcal{S}'(\mathbb{R}^n) : \langle D \rangle^s u \in L^2 \right\}$$

where $\langle D \rangle^s = \mathcal{F}^{-1} \langle \xi \rangle^s \mathcal{F}$.

Definition 1.14 (Negative Sobolev Space). *$W^{-k,p}(U)$ is all $u \in \mathcal{D}'(U)$ such that there exist $g_\alpha \in L^p$ $(|\alpha| \le k)$ such that $u = \sum_{|\alpha| \le k} D^\alpha g_\alpha$. The norm is defined as:*

$$\|u\|^p_{W^{-k,p}(U)} = \inf_{g_\alpha} \sum_{|\alpha| \le N} \|g_\alpha\|_{L^p}$$

Theorem 1.38 (Duality of Negative Sobolev Spaces). *For a domain U, and $p \in (1, \infty)$, $(W_0^{k,p}(U))^* = W^{-k,p'}(U)$. Moreover if U has C^k domain, then $(W^{k,p}(U))^* = W_{\bar{U}}^{-k,p'}(\mathbb{R}^d)$*

The advantage of this is that, for instance:

$$\|u\|_{H^{-1}(U)} = \sup_{\|v\|_{H_0^1} = 1} \langle u, v \rangle_{L^2}$$

(For example if $f \in H^{-1}(U)$, then there exist $g_0, g_\alpha \in L^2$, such that $f = g_0 + \sum_{|\alpha| = 1} \partial^\alpha g_\alpha$).

1.4.2 Approximation

Theorem 1.39 (Approximation of $W^{k,p}$). *If U is an open set, then $C_0^\infty(\mathbb{R}^n)$ is dense in $W^{k,p}(U)$.*

If U is bounded with C^1 domain, then $C^\infty(\bar{U})$ is dense in $W^{k,p}(U)$.

The idea behind proving these is as follows:

1. By mollification can get convergence in $W^{k,p}_{loc}$ (the key thing to show is that $\partial^\alpha(\eta_\varepsilon * u) = \eta * \partial^\alpha u$)
2. By choosing a partition of unity, can approximate elements of $W^{k,p}(U)$ by elements in $C^\infty(U) \cap W^{k,p}$
3. by smoothing the boundary, get best result.

1.4.3 Extensions

Theorem 1.40 (Extension of $W^{k,p}(U)$). *If U is an open, bounded domain of $\mathbb{R}^n$ with C^k boundary and V an open set containing $\bar{U}$, then there exists $\mathcal{E} : W^{k,p}(U) \to W^{k,p}(\mathbb{R}^n)$ such that for $u \in W^{k,p}(U)$,:*

1. $\mathcal{E}(u)|_U = u$
2. *supp*$\mathcal{E}(u) \subset V$
3. $\|\mathcal{E}(u)\|_{W^{k,p}(\mathbb{R}^n)} \lesssim \|u\|_{W^{k,p}(U)}$

Proof. reduce to flat boundary, then reflect and brute force it.

1. since what we want is bounded, it suffices to show this for a dense domain, $C^\infty(\bar{U})$
2. cover the boundary in balls to smooth out the boundary. Use compactness to reduce to a finite subcover. Then we can use a partition of unity to consider extensions on each of these pieces
3. it suffices to extend $u \in C^k(B_1(0), \bar{\mathbb{R}}^d_+)$
   (a) set $\mathcal{E}(u) = u$ for $x_d > 0$ and for $x_d < 0$: $\mathcal{E}(u) = \sum_0^k \alpha_j u(x', -\beta_j x_d)$
   (b) for C^k, need left and right derivatives to agree: $1 = \sum_0^n (-\beta_j)^l \alpha_j$ for $l = 0, 1, \dots, l$.
   (c) this system of equations is solvable if β_j are distinct (vandermonde), and if they are less than 1, we get the correct support property.

□

1.4.4 Traces

Theorem 1.41 (Existence of Trace of Sobolev Functions). *If U is a open domain with C^1 boundary, then there exists a unique bounded linear map $Tr : W^{1,p}(U) \to L^p(\partial U)$ such that $Tr(u)| = u|_{\partial U}$ for all $u \in W^{1,p}(U) \cap C(\bar{U})$*

The fact that the norm of the boundary data can be controlled by the norm on the inside is non-trivial and relies heavily on the divergence theorem.

Proof. Suffices to show for $u \in C^\infty(\bar{U})$. After straightening the boundary with finitely many neighborhoods and using a partition of unity it suffices to control $\|u\chi\|_{L^p(\partial\Omega)}$ with $\Omega = \{x_n > 0\}$ and χ a smooth cutoff function.

1. $\|u\chi\|^p_{L^p(\partial\Omega)} = \int_{x_n=0} |u\chi|^p dx' = -\int |u\chi|^p(-1)dx'$, then apply the Gauss-Green theorem to get $-\int_\Omega \partial_{x_n}(|u\chi|^p)dx$
2. integrand is bounded by a constant times $|u|^p + |u|^{p-1}|\partial_{x_n} u|$
3. for the second term, use Young's inequality: $|u|^{p-1}|\partial_{x_n} u| \leq \frac{|u|^p(p-1)}{p} + \frac{|\partial_{x_n} u|^p}{p}$.

□

Trick: for Sobolev inequalities, the following useful identity was used:

$$a^{p-1}b \leq \frac{(p-1)a^p}{p} + \frac{b^p}{p}$$

for $p > 1, a, b \geq 0$

Theorem 1.42 (Trace Zero). *If U has C^1 boundary, and $u \in W^{1,p}(U)$. Then $Tu = 0$ if and only if $u \in W_0^{1,p}$*

1.4.5 Gagliardo-Nirenberg-Sobolev Inequality

In all these inequalities, we have $1 \leq p < n$

Theorem 1.43 (GNS $C_0^\infty(\mathbb{R}^n)$ inequality). $\|u\|_{L^{d/(d-1)}(\mathbb{R}^d)} \leq C\,\||\nabla u|\|_{L^1(\mathbb{R}^d)}$ *for all* $u \in C_0^\infty(\mathbb{R}^d)$

Theorem 1.44 (Dimensional Scaling for L^p functions). *If $u_\lambda = u(\lambda^{-1}x)$, then*

$$\|\partial^\alpha u_\lambda\|_{L^p(\mathbb{R}^d)} = \lambda^{\frac{d}{p}-|\alpha|}$$

Proof. (of GNS inequality)Use fundamental theorem of calculus and Holder's inequality.

1. Let $f_i = \sup_{x_i} |u(x_1, \ldots, x_n)|$, note $u(x) = \int_{-\infty}^{x_i} \partial_t u(x_1, \ldots, t, \ldots x_n)dt$, so that $f_i \leq \|Du\|_{L^1_{x_i}}$.[a]
2. since $u \leq f_i$ for all i, we have $\int |u|^{\frac{d}{d-1}} dx \leq \int \prod f_i^{(d-1)^{-1}} dx$.
3. Apply Looms-Whitney Inequality, to get $\int |u|^{\frac{d}{d-1}} dx \leq \prod \left\| f_i^{1/(d-1)} \right\|_{L^{d-1}}$

[a] here $Du = \sum_1^n \partial_i u$

4. note $\left\|f_i^{1/(d-1)}\right\|_{L^{d-1}} \le \|Du\|_{L^1}^{(d-1)^{-1}}$, so $\prod \left\|f_i^{1/(d-1)}\right\|_{L^{d-1}} \le \|Du\|_{L^1}^{\frac{d}{d-1}}$. Flip the exponents and we are done.

□

Theorem 1.45 (Loomis-Whitney Inequality). *Let f_i, $i = 1, \dots, n$, be nonnegative measurable functions on $\mathbb{R}^n$ such that f_i is independent of the ith coordinate. Then:*

$$\left\|\prod_1^n f_i\right\|_{L^1} \le \prod_1^n \|f_i\|_{L^{n-1}}$$

Proof. The key is that $\left\|\|f(x,y)\|_{L_x^p}\right\|_{L_y^p} = \|f(x,y)\|_{L_{x,y}^p}$

1. $\int f_1 f_2 \cdots f_n dx_1 = f_1 \int f_2 \cdots f_n dx_1 \le f_1 \prod_2^{n-1} \|f_i\|_{L_{x_1}^{n-1}}$ (iterated Holder inequality)
2. $\iint f_1 f_2 f_3 \cdots f_n dx_1 dx_2 \le \int f_1 \prod_2^{n-1} \|f_i\|_{L_{x_1}^{n-1}} dx_2 \le \|f_2\|_{L_{x_1}^{n-1}} \|f_1\|_{L_{x_2}^{n-1}} \|f_3\|_{L_{x_1,x_2}^{n-1}} \cdots \|f_n\|_{L_{x_1,x_2}^{n-1}}$
3. keep repeating this process.

□

Theorem 1.46 (GNS $W^{1,p}$ inequality). *Let $u \in C_0^\infty(\mathbb{R}^n)$, then $\|u\|_{L^{p^*}} \le C\|Du\|_{L^p}$.*

where $\frac{d}{p^*} = \frac{d}{p} - 1$. Note that $p^* > p$. So maybe we can think of this as bounding an L^{p^*} norm by pulling off weight in the p^* and putting on the regularity.

Proof.
1. $\int |u|^{p^*} dx = \int |u|^{\gamma(\frac{d}{d-1})} \le C \left\|D|u|^\gamma\right\|_{L^1}^{\frac{d-1}{d}}$
2. by Holder: $\left\|D|u|^\gamma\right\|_{L^1} \le C \left\||u|^{\gamma-1}\right\|_{\frac{p}{p-1}} \|Du\|_p$
3. then just do algebra on constants to get the result.

□

Theorem 1.47 (GNS $W^{1,p}(U)$ inequality). *If U is an open and connected subset of $\mathbb{R}^n$ then for $u \in W_0^{1,p}(U)$: $\|u\|_{L^{p^*}(U)} \le C\|Du\|_{L^p(U)}$. If $u \in W^{1,p}(U)$ and U has C^1 boundary, then $\|u\|_{L^{p^*}(U)} \le C\|u\|_{W^{1,p}(U)}$*

Note: the last inequality involves the full $W^{1,p}$ norm, but this can be weakened if $Tr(u) = 0$

Proof. The first is trivial as we can approximate u by functions in $C_0^\infty(\mathbb{R}^n)$ converging in $W_0^{1,p}$, for which we can apply the above GNS inequality (the details are also covered in the next case).

For the other case

1. extend u to $\mathcal{E}u := \bar{u} \in W^{1,p}(\mathbb{R}^n)$
2. approximate $\bar{u}$ by $u_n \in C_0^\infty$ in $W^{1,p}$

3. $\|u_n - u_m\|_{L^{p*}} \to 0$ by GNS, therefore $u_n \to \bar{u}$ in L^{p*}[a]

4. apply GNS and fact that $\|\bar{u}\|_{W^{1,p}} \leq \|u\|_{W^{1,p}(U)}$

□

We can refer to the the $W_0^{1,p}(U)$ case as Poincare's inequality, and can easily generalize it as:

Theorem 1.48 (Poincare Inequality). *For U and open, bounded subset of $\mathbb{R}^n$, $u \in W_0^{1,p}(U)$, then:*

$$\|u\|_{L^q} \leq C \|Du\|_{L^p}$$

for all $q \in [1, p^*]$

1.4.6 Morrey's Inequality

This section deals with $p > n$, in this case we have

Theorem 1.49 (Morrey's Inequality). *For $u \in C^1(\mathbb{R}^n)$:*

$$\|u\|_{C^{0,\alpha}(\mathbb{R}^d)} \leq \|u\|_{W^{1,p}(\mathbb{R}^d)}$$

with $\alpha = 1 - \frac{d}{p}$

where we define the Holder norm as:

$$\|u\|_{C^{0,\alpha}(\mathbb{R}^d)} = \underbrace{\sup_{x \neq y} \frac{|u(x) - u(y)|}{|x-y|}}_{:=[u]_{C^{0,\alpha}}} + \|u\|_{C^0}$$

Proof. idea: for both things we need to estimate ($|u(x)|$ and $|u(x) - u(y)|$, use a potential estimate, then Holder's inequality, the rest is algebra.

1. fix x, then:

$$\begin{aligned} |u(x)| &= |B_r(x)|^{-1} \left| \int_{B_r(x)} u(x) - u(y) + u(y) dy \right| \\ &\leq |B_r(x)|^{-1} \int_{B_r(x)} |u(x) - u(y)| dy + |B_r(x)|^{-1} \int_{B_r(x)} |u(y)| dy \end{aligned}$$

2. bound second term by Holder: $r^{-n} \|B_r(x)\|_{L^{p'}} \|u\|_{L^p(B_r(x))} = C r^{n(p-1)p^{-1} - n} \|u\|_{L^p(B_r(x))}$

3. second term is bounded by $\int_{B_r(x)} \frac{|Du(y)|}{|x-y|^{n-1}} dy$

   (a) $\int_{\partial B_r(x)} |u(y) - u(x)| dy = r^{n-1} \int_{\partial B_1(0)} |u(zr + x) - u(x)| dz = r^{n-1} \int_{\partial B_1(0)} |\int_0^r \frac{d}{dt} u(tz + x)| dt dz \leq r^{n-1} \int_{\partial B_1(0)} \int_0^r |Du(zt + x)| dt dz$

[a] if $f_n \to f$ in L^p and $f_n \to g$ in L^q, and we are on a compact set and $p < q$. Then $f_n \to g$ in L^p, therefore $f = g$ almost everywhere.

(b) $\int_{\partial B_1(0)} \int_0^r |Du(zt+x)|dzdt = \int_{\partial B_r(x))} |Du(y)||x-y|^{n-1} dy$[a]

(c) get $\int_{\partial B_r(x)} |u(y)-u(x)|dy \leq r^{n-1} \int_{\partial B_r(x))} |Du(y)||x-y|^{n-1} dy$. Integrate both sides from $r=0$ to $r=R$, replace on the RHS $B_r(x)$ by $B_R(x)$ to get: $\int_{B_R(x)} |u(y)-u(x)|dy \leq CR^N \int_{\partial B_R(x)} |Du(y)||x-y|^{n-1} dy$

4. by Holder $\int_{B_r(x)} \frac{|Du(y)|}{|x-y|^{n-1}} dy \leq \|Du\|_{L^p(B_r(x))} \||x-y|^{n-1}\|_{L^{p'}(B_r(x))}$. The second factor is finite because $p > n \Rightarrow (n-1)p(p-1)^{-1} < n$. It's value, after some algebra, is Cr^α

5. so $|u(x)| \leq r^\alpha \|Du\|_{L^p(B_r(x)} + r^{-n/p} \|u\|_{L^p(B_r(x)}$. If $r=1$, we get $\|u\|_{C^0} \leq \|u\|_{W^{1,p}}$

6. Next, for $x \neq y$, let $r = |x-y|$, then:

$$|u(x)-u(y)| \leq \frac{1}{|B_r(x) \cap B_r(y)|} \left(\int_{B_r(x) \cap B_r(y)} |u(x)-u(z)|dz + \int_{B_r(x) \cap B_r(y)} |u(y)-u(z)|dz \right)$$

since $|B_r(x) \cap B_r(y)| \geq C2^{-n} r^n$, we can bound each of these terms by

$$\fint_{B_r(x)} |u(x)-u(z)|dz \leq C \int_{B_r(x)} |Du(z)||x-z|^{-(n-1)} dz \leq C \|Du\|_{L^p(B_r(x))} r^\alpha$$

(or with x replaced by y)

7. divide by r^α, get $|u(x)-u(y)|r^{-\alpha} \leq \|Du\|_{L^p(\mathbb{R}^n)}$

□

Theorem 1.50 (Potential Estimate for Morrey's Inequality). *If $u \in C^1(\overline{B_r(x)})$, then*

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

The proof of this involves switching back and forth from polar coordinates.

Another useful bound that is used is:

$$\left(\int_U f(x)dx \right) \leq \|f\|_{L^p(U)} |U|^{(p-1)p^{-1}}$$

Theorem 1.51 (Morrey's Inequality for $W^{1,p}$). *If U is a bounded open domain with C^1 boundary, then for all $u \in W^{1,p}(U)$, then u is almost everywhere equal to $u^* \in C^{0,\alpha}(\bar{U})$ ($\alpha = 1 - \frac{d}{p}$) in U and $\|u^*\|_{C^{0,\alpha}(U)} \leq C \|u\|_{W^{1,p}(U)}$*

Proof. Use extension, approximations, and the above to get a sequence of smooth compactly supported functions which is Cauchy in Holder norm, and therefore has a limit which is is Cauchy. □

[a] we are integrating $\int_{\partial B_1(0)} \int_0^r f(t,\omega)d\omega = \int_{\partial B_1(0)} \int_0^r f(t,\omega)t^{n-1}t^{1-n} d\omega = \int_{B_r(x)} f(y)t^{1-n}(y)dy$

$u \in C_0^\infty(\mathbb{R}^n),\ p < n$	$\lVert u\rVert_{L^{p'}(\mathbb{R}^n)} \leq \lVert Du\rVert_{L^1(\mathbb{R}^n)}$	GNS
$u \in C_0^\infty(\mathbb{R}^n),\ p < n$	$\lVert u\rVert_{L^{p'}(\mathbb{R}^n)} \leq \lVert Du\rVert_{L^p(\mathbb{R}^n)}$	GNS
$u \in W_0^{1,p}(U),\ p < n$	$\lVert u\rVert_{L^{p'}(U)} \leq \lVert Du\rVert_{L^p(U)}$	Poincare/GNS
$u \in W^{1,p}(U),\ p < n$ [a]	$\lVert u\rVert_{L^{p'}} \leq \lVert u\rVert_{W^{1,p}(U)}$	GNS
$u \in C^1(\mathbb{R}^n),\ p > n$	$\lVert u\rVert_{C^{0,\alpha}(\mathbb{R}^n)} \leq \lVert u\rVert_{W^{1,p}(\mathbb{R}^n)}$	Morrey
$u \in W^{1,p}(U),\ p > n$ [b]	$\lVert u^*\rVert_{C^{0,\alpha}(\bar{U})} \leq \lVert u\rVert_{W^{1,p}(U)}$	Morrey

1.4.7 General Sobolev Inequalities

For both the following theorems, $k \in \mathbb{Z}_{>0}$, $p \in [1,\infty)$, U is a bounded domain and either $u \in W_0^{k,p}(U)$ or $u \in W^{k,p}(U)$ with U having C^k boundary.

Theorem 1.52 (General Sobolev Inequality).

$$\lVert u\rVert_{W^{\ell,q}(U)} \leq C\,\lVert u\rVert_{W^{k,p}(U)}$$

for $\frac{n}{q} - \ell \geq \frac{n}{p} - k$ $(\ell \in \mathbb{Z}_{>0},\ \ell \leq k,\ q \in [1,\infty))$ *and*

$$\lVert u^*\rVert_{C^{\ell,\alpha}(U)} \leq C\,\lVert u\rVert_{W^{k,p}(U)}$$

for $-\ell - \alpha \geq \frac{n}{p} - k$ *for* $\alpha \in (0,1)$, $\ell \in \mathbb{Z}_{>0}$, *and* $\ell \leq k$

To remember constants: (1) regularity cannot increase (2) $q \neq \infty$ (3) degree of homogeneity of top order term on LHS must be $\geq$ RHS.

One thing to look at when considering $u \in W^{k,p}$ is the ratio n/p. If $k > n/p$ we are Holder continuous, if $k < n/p$, then we have some control on the L^q norm

1.4.8 Compactness

Definition 1.15 (Compact Embedding of Banach Spaces). *We say X compactly embeds into Y for Banach spaces X and Y (written $X \Subset Y$) if (1) $X \subset Y$, (2) $\lVert x\rVert_Y \leq C\,\lVert x\rVert_X$ for all $x \in X$ (3) every bounded sequence in X has a convergent subsequence in Y*

Equivalently, the inclusion $\iota : X \to Y$ is a bounded compact operator.

Theorem 1.53 (Rellich-Kondrachov Compactness Theorem). *If $U \subset \mathbb{R}^n$ is bounded and open with C^1 domain and $1 \leq p < n$, then $W^{1,p}(U) \Subset L^q(U)$ for all $q \in [1, p^*)$*

For $q < p^*$, this is stronger statement than GNS, however, something funny happens for $q = p^*$

Proof.

1. Inclusion and boundedness of inclusion is trivial by GNS
2. if $u \in W^{1,p}(U)$ is bounded, WLOG may assume by extension they are elements of $W_0^{1,p}(V)$. Let u_m^ε by mollifications, wlog assume they are all supported in V.
3. claim: $u_m^\varepsilon \xrightarrow{\varepsilon \to 0} u_m$ in L^q uniformly in m.

4. claim: for each $\varepsilon > 0$, u_m^ε are uniformly bounded and equicontinuous.

5. for each $\delta > 0$, pick ε so $\|u_m^\varepsilon - u_m\|_{L^q} \le \delta/2$ for all m, then use Arzela-Ascoli to get subsequence of u_m^ε which converge uniformly (and hence Cauchy in L^q).

6. by triangle inequality, get $\limsup \left\|u_{m_j} - u_{m_l}\right\|_{L^q} \le \delta$.

7. By diagonalization with $\delta_n = 1/n$, get result.

□

Here is a proof of step 3, the idea uses: mollification approximation, fundamental theorem of calculus, L^p interpolation.

Proof.

1. $u_m^\varepsilon(x) - u_m(x) = \int \eta_\varepsilon(y)(u(x-y) - u_m(x))dy = \int_{B_1(0)} \eta(y)(u_m(x-\varepsilon y) - u_m(x))dy = \int_{B_1(0)} \eta(y) \int_0^1 \frac{d}{dt} u_m(x - \varepsilon t y) dt dy$

2. $\frac{d}{dt} u_m(x - \varepsilon t y) = -\varepsilon \nabla u_m(x - \varepsilon t y) \cdot y$

3. Taking absolute values get $|u_m^\varepsilon(x) - u_m(x)| \le \varepsilon \int_{B_1(0)} \int_0^1 \eta(y)|Du_m(x - \varepsilon t y)| dt dy$. Integrating over V, get $\|u_m^\varepsilon - u_m\|_{L^1(V)} \le \varepsilon \|Du_m\|_{L^1(V)}$ (haven't verified this, but seems clear).

4. then this goes to zero since u_m are uniformly bounded in $W^{1,p}$.

5. interpolate: $\|u_m^\varepsilon - u_m\|_{L^q} \le \|u_m^\varepsilon - u_m\|_{L^1}^\theta \|u_m^\varepsilon - u_m\|_{L^{p^*}}^{1-\theta}$ [a]

□

Theorem 1.54 (Rellich-Kondrachov Compactness Theorem (any p)). *For $U \subset \mathbb{R}^n$ bounded, open, with C^1 boundary. Then $W^{1,p}(U) \Subset L^p(U)$ for all $p \in [1, \infty]$*

Proof. If $p < n$, then since $p^* > p$, we immediately have this. Now let $p \ge n$:

1. if w_n are bounded in $W^{1,p}$, then let $p' < n$ (so $p' < p$) be such that $(p')^* > p$ (this is easy to get).

2. Then $\|w_n\|_{W^{1,p'}} \le \|w_n\|_{W^{1,p}}$. So w_n are bounded in $W^{1,p'}$, so we apply compactness to get w_{n_k} converging in L^p

□

[a] here's the key thing: $1 < q < p^*$, so $1 > q^{-1} > (p^*)^{-1}$ (if $q = 1$, then we were already done). Then Holder's interpolation theorem allows us to control q norm by any convex combination of 1 and p^*

1.4.9 Poincaré Inequality

Theorem 1.55 (Poincaré Inequality for $W^{1,p}$). *If $U \subset \mathbb{R}^n$ is bounded, open, with C^1 boundary, then for all $p \in [1, \infty]$:*

$$\left\| u - \fint_U u(y)dy \right\|_{L^p(U)} \leq C \|Du\|_{L^p(U)}$$

for all $u \in W^{1,p}(U)$

Note that our previous Poincare inquality didn't subtract the average but worked for trace zero functions.

Proof. Argue by contradiction, get bounded sequence in $W^{1,p}$, use compactness, get L^p function that has average zero, derivative zero, but L^p norm 1, which is impossible

1. assume false, let u_k be such that $\|u_k - \fint u_k\|_{L^p} \geq k \|Du_k\|_{L^p}$
2. set $v_k = \frac{u_k - \fint u_k}{\|u_k - \fint u_k\|_{L^p}}$, so that $\|v_k\|_{L^p} = 1$, and $\|Dv_k\|_{L^p} \leq k^{-1}$
3. v_k are bounded in $W^{1,p}$, so by compactness, get subsequence converging in L^p to $v \in L^p$
4. $\|v\|_{L^p} = 1$ and $\fint v = 0$[a]
5. checking $\langle v, \partial_{x_i}\varphi \rangle = \lim_{k\to\infty} \langle v_{n_k}, \partial_{x_i}\varphi \rangle \to 0$, get v is constant, and zero but this is a contradiction.

□

1.5 Second-order Elliptic Equations

Reference: Evans 6.1-6.5

Weak solutions, Lax-Milgram Theorem, existence and uniqueness, elliptic regularity, maximum principles, eigenvalues and eigenfunctions

A second order elliptic differential equation can be written:

$$\begin{cases} Lu = f & x \in U \\ u = 0 & x \in \partial U \end{cases} \tag{4}$$

with U a bounded open set in $\mathbb{R}^n$, $f \in L^2(U)$, and L a partial differential operator written either as

$$Lu = -\sum_{i,j=1}^{n} \partial_j(a^{ij}(x)\partial_i u) + \sum_{i=1}^{n} b_i(x)\partial_i u + c(x)u(x) = -\nabla \cdot (\nabla u A(x)) + B(x) \cdot \nabla u + c(x)u(x) \tag{divergence form}$$

$$Lu = -\sum_{i,j=1}^{n} a^{ij}(x)\partial_i\partial_j u + \sum_{i=1}^{n} b_i(x)\partial_i u + c(x)u(x) \tag{non-divergence form}$$

[a] sometimes things aren't obvious to me: $\fint v = \fint v + v_n - v_n = \fint v - v_n$, absolute value is bounded by $\fint |v - v_n| \leq C \|v - v_n\|_{L^p} \to 0$

Definition 1.16 (Elliptic Operator). *A differential operator defined above is uniformly elliptic if there exists $\theta > 0$ such that for almost every $x \in U$, $\xi \in \mathbb{R}^n$*

$$\sum_{ij} a^{ij}(x)\xi_i\xi_j \geq \theta|\xi|^2$$

(equivalently (a^{ij}) is uniformly positive definite). We will assume a^{ij} is symmetric in everything below.

1.5.1 Weak Solutions

Definition 1.17 (Weak Solution to Elliptic Equation). *For the above PDE we define the bilinear form:*

$$B[u,v] = \int_U \sum_{ij} a^{ij}\partial_i u \partial_j v + \sum_j b^j \partial_i uv + cuvdx \tag{5}$$

*for $u, v \in H_0^1(U)$. We say that u is a **weak solution** if $B[u,v] = \langle v, f\rangle$ for all $v \in H_0^1(U)$*

1.5.2 Existence of Weak Solutions

Theorem 1.56 (Lax-Milgram Theorem). *Let $B : H \times H \to \mathbb{R}$ be a bilinear map on Hilbert spaces H that is (1) bounded: $|B(u,v)| \leq C\|u\|_H \|v\|_H$ and (2) coercive: $\|u\|_H^2 \leq CB(u,u)$. Then for each $f : H \to \mathbb{R}$ bounded linear functional, there exists $u \in H$ such that $B(u,v) = \langle f, v\rangle$ for all $v \in H$.*

Proof. linear algebra

1. fix u, $v \mapsto B(u,v)$ is bounded linear. By Riesz-representation theorem, get $Au \in H$ such that $\langle Au, v\rangle = B(u,v)$.
2. A is bounded and linear (linear is easy), bounded: $\|Au\|_H^2 = \langle Au, Au\rangle = B(u, Au) \leq C\|u\| \|Au\|$
3. A is one-to-one with closed range. $\|u\|^2 \leq B(u,u) = \langle Au, u\rangle \leq \|Au\| \|u\|$, so $\|Au\| \geq \|u\|$
4. the range of A is H. If not, get w such that $0 = \langle Aw, w\rangle = B(w,w) \geq C\|w\|^2$
5. by Riesz-representation, have $w \in H$ such that $\langle f, u\rangle = \langle w, v\rangle$ for all $v \in H$, let $u \in H$ be such that $Au = w$
6. $B(u,v) = \langle Au, v\rangle = \langle w, v\rangle = \langle f, v\rangle$. By coercivity, get uniqueness.

□

Theorem 1.57 (Energy Estimate of Bilinear form for Elliptic 2nd Order PDE). *For the bilinear form in (5), there exist positive constants such that*

$$|B(u,v)| \leq \alpha\|u\|_{H_0^1}\|v\|_{H_0^1}$$
$$\beta\|u\|_{H_0^1}^2 \leq B(u,u) + \gamma\|u\|_{L^2}^2$$

Proof. for coercive: start with Poincare, use ellipticity, be wasteful with bounds on coefficients, use Peter-Paul to get rid of negative derivative term

Boundedness is trivial (coefficients are uniformly bounded (we're on a compact set) then we use Holder's inequality).

Coercive bound:

1. by Poincare (since we have trace zero functions): $\|u\|_{H_0^1} \le C \|Du\|_{L^2}$.
2. $\|Du\|_{L^2}^2 \le C \int \sum a^{ij} \partial_i u \partial_j u dx$ (by ellipticity), and this term is $B(u,u) - \int (b^i \partial_i u u + cu^2) dx$
3. $|\int (b^i \partial_i u u + cu^2) dx| \le \|b\|_\infty (\varepsilon \|Du\|_2^2 + \varepsilon^{-1} \|u\|_2^2)$ (by Young's inequality).
4. just rearrange, make ε small enough so that the coefficient of $\|Du\|_2^2$ is positive.

□

Another way to write the second energy is:

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

This is because $\|Pu\|_{H^{-1}(U)}^2 = \sup_{\|v\|_{H_0^1(U)}=1} \langle Pu, v \rangle \ge \|u\|_{H_0^1(U)}^{-1} \langle Pu, u \rangle$

As a corollary, we get:

Theorem 1.58 (Existence of Weak Solution for Modified Elliptic PDE). *Let $\mu \ge \gamma$ (for γ in the previous problem), then there exists a unique weak solution to:*

$$\begin{cases} Lu + \mu u = f & \text{in } U \\ u = 0 & \text{in } \partial U \end{cases}$$

for $f \in L^2(U)$.

Proof.

1. The new bilinear form is $B_\mu(u,v) = B(u,v) + \mu \langle u, v \rangle_{L^2}$
2. this is clearly bounded, by the above, $\beta \|u\|_{H_0^1} \le B(u,u) + \gamma \|u\|_2^2 = B_\mu(u,u) - (\mu - \gamma) \|u\|_2^2 \le B_\mu(u,u)$
3. by Lax-Milgram, we get a unique weak solution.

□

Theorem 1.59 (Second Existence Theorem of Weak Solutions to Elliptic PDE (via Fredholm alternative)). *For (4), such that $b^i \in C^1(\bar{U})$, the null spaces (solving homogeneous problem $f = 0$) of L and L^* (formal adjoint) have the same (finite) dimension, call them N and N^*. And there exists a unique solution to (4) if and only if $\langle f, u \rangle = 0$ for all $u \in N^*$.*

Furthermore, these weak solutions are unique modulo $\ker L$

Basically, we can solve second order PDEs, except for a small slice of L^2 removed having finite dimension. This slice is characterized by solving the homogeneous adjoint problem $L^*u = 0$

Proof.

1. let $\mu = \gamma$ is above theorem, for $f \in L^2$, define $L_\gamma^{-1} f = u$ to solve $B_\gamma(u,v) = \langle f, v\rangle$ for all $v \in H_0^1$.
2. u solves actual PDE if and only if $B(u,v) = B_\gamma(u,v) - \gamma(u,v) = \langle f, v\rangle$. So $B_\gamma(u,v) = \langle f + \gamma u, v\rangle$, that is $u = L_\gamma^{-1}(f + \gamma u)$
3. rearrange as $u - \gamma L_\gamma^{-1} u = L_\gamma^{-1} f := (I-K)u = h$ with $K = \gamma L_\gamma^{-1}$ and $h = L_\gamma^{-1} f$. So **if** we $h \in R(I-K)$, we have a solution.
4. claim: K is compact
   (a) Bounded: $\beta \|u\|_{H_0^1}^2 \le B_\gamma(u,u) = \langle g, u\rangle \le \|g\|_2 \|u\|_2 \le \|g\|_2 \|u\|_{H_0^1}$ so $\left\|L_\gamma^{-1} g\right\|_{H_0^1} \le C\|g\|_2$
   (b) Compact: let $g_n \in L^2$ be bounded, then by above Kg_n are bounded in H_0^1, by Rellich compactness, there exists subsequence Kg_{n_k} convergent in L^2.
5. apply Fredholm alternative: $(I-K)u = h$ has a solution if and only if $f \perp N(I - K^*)$. There are two options:
   (a) $N(I-K^*) = 0$, in which case $R(I-K) = H$, and so for every $f \in L^2$, there is a unique weak solution.
   (b) $dim(N(I-K)^*) \neq 0$, there is a m dimensional subspace of solutions to $B(u,v) = 0$ for all $v \in H_0^1$ (called N) and $B^*(\tilde{u}, v) = 0$ for all $v \in H_0^1$ (called N^*). And we can solve the original for f if and only if $\langle f, v\rangle = 0$ for all $v \in N^*$

□

Theorem 1.60 (Fredholm Alternative). *If K is a compact linear operator, then:*

1. $N(I-K)$ *is finite dimensional*
2. $R(I-K)$ *is closed*
3. $R(I-K) = N(I-K^*)^\perp$
4. $N(I-K) = 0$ *if and only if* $R(I-K) = H$
5. $\dim N(I-K) = \dim R(I-K^*)$

To remember: K being compact means it is approximately finite dimensional[a], so if we let $T = I - K$ it behaves like a finite dimensional operator[b] $N(T)$ is just the range of K

[a]it is the limit of finite rank operators in operator norm
[b]it is a Freholm operator

(which is basically finite dimensional), if H was finite dimensional, then, $R(T) = N(T^*)^\perp$ (which is the statement of 3). (4) is redundant.

The existence and uniqueness of PDE can be restated in the language of **Fredholm operators**. It says that $P : H_0^1(U) \to H^{-1}(U)$ is a Fredholm operator of index 0, that is it is bounded and the dimension of it kernel is finite and equals the dimension of its cokernel. That is P^{-1} exists if and only if $\ker P = \varnothing$.

Theorem 1.61 (Local Solvability of Elliptic PDE). *For every $x_0 \in U$, there exists a unique weak solution $u \in H_0^1(B_\varepsilon(x_0))$ to $Pu = f$ in $B_\varepsilon(x_0)$*

Proof.

1. we have shown that $\|Du\|_{L^2} \le C(B(u,u) + \|u\|_{L^2})$
2. By Poincare, $\|u\|_{L^2(B_\varepsilon(x_0))} \le C_\varepsilon \|Du\|_{L^2(B_\varepsilon(x_0))}$. By dimensional analysis, $C_\varepsilon = C_1\varepsilon$
3. Therefore, there exists $\varepsilon > 0$ such that $1 - C\varepsilon C_1 > 0$, and so $\|Du\|_{L^2(B_\varepsilon(x_0))} \le CB(u,u)$

□

1.5.3 Regularity

Theorem 1.62 (Interior H^2 Regularity of Elliptic PDE). *If $u \in H^1(U)$ is a weak solution to $Pu = f \in L^2$ with $a \in C^1(U)$ and $b, c \in L^\infty$. Then $u \in H^2_{loc}$ and for each $V \Subset U$, $\|u\|_{H^2(V)} \le C(\|f\|_{L^2(U)} + \|u\|_{L^2(U)})$*

Note possibly $u \notin H_0^1(U)$. Also this proof is very long and technical so only the main ideas are important. This result can be obtained through a parametrix construction using microlocal analysis..

Proof.

1. Since $B(u,v) = \langle f, v\rangle$ for $v \in H_0^1$. Expand $B(u,v)$, move the terms involving b and c to the RHS, redefine that as $\tilde{f}$ to get $\langle A\nabla u, \nabla v\rangle = \langle \tilde{f}, v\rangle$
2. set $v = \partial_j(\chi^2 \partial_j u)$ with χ a cuttoff function, and we actually use difference quotients
3. expand LHS, integrate by parts, use ellipticity, Young's inequality, to ultimately get a lower bound of RHS as $\ge \|\chi D^2 u\|_{L^2}^2 - C\|Du\|_{L^2}^2$

□

Theorem 1.63 (High Interior H^2 Regularity for Elliptic PDE). *Given $a, b, c \in C^{m+1}(U)$, $f \in H^m(U)$, then if $u \in H^1(U)$ is a weak solution to $Pu = f$ in U, then $u \in H^{m+2}_{loc}$, and for all $V \Subset U$:*

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

Proof.

1. induct on m, assume true for m. Fix $W \Subset V \Subset U$
2. let $\alpha \in \mathbb{N}^n$ with $|\alpha| = m+1$, pick $\varphi \in C_0^\infty(W)$, let $v = D^\alpha \varphi \in H_0^1$, therefore $B(u,v) = \langle f, v \rangle_2$
3. integrate by parts to get $B(\tilde{u}, \varphi) = \langle \tilde{f}, \varphi \rangle$ with $\tilde{u} = D^\alpha u \in H^1$ and $\tilde{f}$ some complicated thing. But now $\tilde{u}$ is a weak solution to $P\tilde{u} = \tilde{f}$
4. by previous result, $\|\tilde{u}\|_{H^2(V)} \le C(\|f\|_{L^2} + \|\tilde{u}\|_{L^2})$
5. expanding f, and using previous induction we can control $\|\tilde{u}\|_{L^2}$ and $\|f\|_{L^2}$ to get $\|\tilde{u}\|_{H^2(V)} \le C(\|f\|_{H^{m+1}(U)} + \|u\|_{L^2(U)})$, and therefore $u \in H^{m+3}$

□

Theorem 1.64 (Boundary Regularity of Elliptic PDE). *Given $a, b, c \in C^{m+1}(\bar{U})$, $f \in H^m(U)$, U has C^{m+2} boundary, and $u \in H_0^1(U)$ a weak solution to $Pu = f$ in U, then $u \in H^{m+2}(U)$ and $\|u\|_{H^{m+2}(U)} \le C(\|f\|_{H^m(U)} + \|u\|_{L^2(U)})$*

If u is the unique solution, then $\|u\|_{H^{m+2}(U)} \le C \|f\|_{H^m(U)}$.

1.5.4 Maximum Principles

Theorem 1.65 (Weak Maximum Principal for Elliptic PDE). *If $u \in C^2(U) \cap C(\bar{U})$ satisfies $Pu \le 0$ in U, with $a, b \in C$ and $c \equiv 0$, then $\max_{\bar{U}} u = \max_{\partial U} u$*

Proof. trivial for $Pu = -\Delta u < 0$ (otherwise need to orthogonally diagonalize a). The regularization term is nontrival to come up with: $\varepsilon e^{\lambda x_1}$

1. first show true for $Pu < 0$, assume $x_0 \in U$ such that u has maximum (wlog $x_0 = 0$), then $\nabla u = 0$ and $\Delta u \le 0$ at this point.
2. Moreover $Pu(x_0) = -a^{ij}\partial_i\partial_j u$. Now claim: $a^{ij}\partial_i\partial_j u \le 0$
   (a) orthogonally decompose $A = O\Lambda O^t$ with Λ diagonal with positive entries. Change coordinates to $y = Ox$, in this case: $\partial_{x_i}\partial_{x_j} u(0) = \sum_{k,l}(\partial_{y_k}\partial_{y_l} u(0))O_{k,j}O_{l,i}$
   (b) Then in these coordinates, we get $\sum a^{ij} O_{k,j} O_{l,i} (\partial_{y_k}\partial_{y_l} u(0))$, this sum contracts to just $\sum \Lambda_j (\partial_{y_j}^2 u) \ge \varepsilon \Delta u$, with $\varepsilon = \min(\Lambda_j)$.
3. If $Pu \le 0$, then let $u_\varepsilon = u + \varepsilon e^{x_1 \lambda}$
   (a) $P\varepsilon e^{x_1\lambda} = \varepsilon \lambda e^{x_1 \lambda}(b^1 - a^{11}\lambda)$
   (b) $a^{11} \ge \theta$, so if $\lambda > \|b\|_\infty \theta^{-1}$, then $Pu_\varepsilon < 0$
   (c) use above, and fact that as $\varepsilon \to 0$, $u_\varepsilon \to 0$

□

Theorem 1.66 (Hopf's Lema). *Given $u \in C^2(U) \cap C^1(\bar{U})$ such that $Pu \le 0$ in u with $c \ge 0$, $x_0 \in \partial U$ a strict max in U with $u(x_0) \ge 0$ and $x \in \partial B_r(\tilde{x})$ for some $B_r(\tilde{x}) \subset U$, then $\partial_\nu u(x_0) > 0$ (where ν is the unit outer normal of $B_r(\tilde{x})$).*

Note $\partial_\nu(x_0) \geq 0$ is trivial.

Proof. the idea is adding an auxiliary function to still be a subsolution with a maximum at x_0 such that normal derivative of the subsolution is strictly negative. The construction of the auxilarly function is clever: defined it on an annulus, the inner boundary can be trivially controlled, the outer boundary controlled by having the function vanish, and interior controlled by weak maximal

1. wlog $\tilde{x} = 0$. Let $R = B_r(0) \setminus B_{r/2}(0)$. Construct a function v such that for some $\varepsilon > 0$, $g(x) := u(x) + \varepsilon v(x) - u(x_0) \leq 0$ in R and $g(x_0) = 0$
   (a) Let $v(x) = e^{-\lambda|x|^2} - e^{-\lambda r^2}$ (so $v(r) = 0$), so $Pv \leq 0$ in R for $\lambda \gg 0$
      i. $Pv = e^{-\lambda|x|^2}\lambda(-4a^{ij}x_i x_j \lambda + 2\operatorname{Tr} A + b^i x_i + c) - ce^{-\lambda r^2}$
      ii. $(-4a^{ij}x_i x_j \lambda + 2\operatorname{Tr} A + b^i x_i + c) \leq -4\theta \|x\|_2^2 \lambda + 2\operatorname{Tr} A + \|b\|_2 \|x\|_2 \leq -c_1\lambda r + c_2 + c_3 r$
   (b) since $u(x_0)$ is maximal, set $0 < \varepsilon \ll 1$ so $u(x_0) \geq u(x) + \varepsilon v(x)$ on $\partial B_{r/2}$
   (c) $Lg \leq -cu(x_0) \leq 0$ in R and $g \leq 0$ on ∂R (since for inner we chose ε small and on outer v vanishes)
2. by weak maximal principal, $g \leq 0$ in R, therefore $\partial_\nu g(x_0) \geq 0$
3. $\partial_\nu g(x_0) = \partial_\nu u(x_0) + \varepsilon \partial_\nu v(x_0)$. Now $\partial_\nu v(x_0) = -2\lambda|x|e^{-|x|^2\lambda} < 0$, therefore $\partial_\nu(x_0) > 0$

□

Theorem 1.67 (Strong Maximal Principal for Elliptic PDE). *Let $u \in C^2(U) \cap C(\bar{U})$ satisfy $Pu \leq 0$ with $c = 0$. If there is $x \in U$ such that $u(x) = \max_{\bar{U}} u(x)$, then u is constant. (the condition $c = 0$ can be replaced by $c \geq 0$ and $\max_{\bar{U}} u \geq 0$)*

Proof.
1. suppose u is nonconstant and attains a max, let $M = \max_{\bar{U}} u$, let $A = \{x \in U : u(x) = M\}$ and $B = U \setminus A$.
2. pick $y \in B$ such that $dist(y, A) < dist(y, \partial U)$, expand a maximally sized ball around y contained in B
3. this ball has a point of A in its boundary. Apply Hopf's lemma, get contradiction because there will be a larger maximum.

□

Theorem 1.68 (Harnack's Inequality for Elliptic PDE). *If $0 \leq u \in C^2(U)$ is a solution to $Pu = 0$ and $V \Subset U$, then there exists a constant C (depending on V and P) such that* $\sup_V u \leq C \inf_V u$

1.5.5 Eigenvalues of Elliptic PDE

Theorem 1.69 (Spectrum of Elliptic Operators). *The real spectrum of P is at most countable, with eigenvalues going to infinity.*

Proof. $\lambda \notin \operatorname{spec}(P)$ if and only if $(P-\lambda)u = f$ has a unique weak solution for each $f \in L^2$ if and only if (by Fredholm-alternative) $(P-\lambda)u = 0$ has **only** the trivial solution.

1. let $\gamma \gg 0$ such that $(P+\gamma)u = f$ has a solution for all $f \in L^2$, then the above can be written $(P+\gamma)u = (\gamma+\lambda)u$, so $u = (P+\gamma)^{-1}(\gamma+\lambda)u$
2. let $K = \gamma(P+\gamma)^{-1}$, this is a compact operator, so we have $u = \frac{\gamma+\lambda}{\gamma}Ku$, so $Ku = \frac{\gamma}{\gamma+\lambda}u$
3. we are interested when this has only the trivial solution, therefore when $\frac{\gamma}{\gamma+\lambda} \notin \operatorname{spec}(K)$
4. since the eigenvalues of compact operators are at most countable and go to zero, the spectrum of P goes to infinity.

□

Theorem 1.70 (Boundedness of Inverse of Elliptic Operator). *If $\lambda \notin \operatorname{spec}(P)$, then $(P-\lambda)^{-1} : L^2 \to L^2$ is bounded.*

Proof.

1. suppose not, get sequence $u_k \in H_0^1$ and $f_k \in L^2$ such that $(P-\lambda)u_k = f_k$ and $\|u_k\|_{L^2} \geq k\|f\|_2$ (wlog $\|u_k\|_2 = 1$)
2. therefore $\|f_k\|_2 \to 0$ and by energy estimate $\|u_k\|_{H_0^1} \leq C(\|f_k\|_2 + \|u_k\|_2)$, we see u_k are bounded in H_0^1
3. by Banach-Alaglou and Sobolev compactness, get $u \in H_0^1$ such that u_{k_j} converge weakly in H_0^1 and strongly in L^2
4. therefore $Pu = \lambda u$[a] and $\|u\|_2 = 1$, but if $\lambda \notin \operatorname{spec}(P)$, we require $u \equiv 0$, so we get a contradiction.

□

For the remainder of this section, assume $Pu = -\partial_j(a^{ij}\partial_i u)$ with a uniformly positive definite and symmetric.

Theorem 1.71 (Eigenvalues of Symmetric Elliptic Operator). *If P is as above, then the eigenvalues are $0 < \lambda_1 \leq \lambda_2 \leq \cdots \to \infty$ and there exists a orthonormal basis for $L^2(U)$ of eigenfunctions in H_0^1 solving the Dirichlet problem. (By elliptic regularity, if ∂U is C^k, then the eigenfunctions are in H^k)*

Proof. It suffices to show that P^{-1} is a symmetric, nonnengative, compact operator.

1. P^{-1} is bounded $L^2 \to H_0^1$ (this follows from Theorem 1.70) which by Sobolev imbedding is compact $L^2 \to L^2$.

[a] this is somewhat nontrivial, but a common trick. The general heuristic is if $u_n \to u$ weakly then $B(u_n, v) \to B(u, v)$ for a bilinear form B

2. P^{-1} is symmetric
   (a) pick $f, g \in L^2$, then $\langle P^{-1}f, g\rangle = \langle u, g\rangle = \langle g, u\rangle = B(v, u)$ where u and v are weak solutions of $Pu = f$ and $Pu = g$ respectively.
   (b) Similarly, $\langle f, P^{-1}g\rangle = \langle f, v\rangle = B(u, v)$, since B is symmetric we get that P^{-1} is symmetric.
3. $\langle P^{-1}f, f\rangle = \langle u, f\rangle = B(u, u) \geq 0$
4. therefore, we get an orthonormal basis of eigenfunctions in L^2. The Eigenvalues of P^{-1} got zero, therefore the eigenvalues of P go to infinity (must be positive infinity because $P + \lambda$ is always invertible for $\lambda \gg 0$).

□

Theorem 1.72 (Principal Eigenvalue of Elliptic Operator). *For P as above $\lambda_1 \leq B(u, u)$ for all $u \in H_0^1$ with $\|u\|_2 = 1$ with a minimum achieved by the weak solution to $Pu_1 = \lambda_1 u_1$. Moreover, the eigenvalue λ_1 is simple.*

Proof.

1. let $u \in H_0^1(U)$, let w_k be orthonormal basis of L^2, write $u = \sum_1^\infty d_k w_k$, this sequence actually converges in $H_0^1(U)$
   (a) $\frac{w_k}{\sqrt{\lambda_k}}$ is an orthonormal basis of $H_0^1(U)$ with respect to the innerproduct $B(\cdot, \cdot)$.
   (b) to see this, let $u \in H_0^1(U)$ such that $B(\frac{w_k}{\sqrt{\lambda_k}, u}) = 0$ for each k. But this says that $0 = \sqrt{\lambda_k}(w_k, u)_2$, but this implies $u \equiv 0$.
   (c) write $u = \sum \mu_k w_k$, with $\mu_k = B(u, \frac{w_k}{\sqrt{\lambda_k}})$, then we can write $d_k = \mu_k \lambda_k^{-1/2}$
2. therefore $B(u, u) = \sum_1^\infty d_k^2 \lambda_k$. If $\|u\|_2 = 1$, then $\sum d_k^2 = 1$, therefore $B(u, u) \geq \lambda_1$ (with equality if and only if $d_1 = 1$).
   (a) if $B(u, u) = \lambda_1$, then $\sum d_k^2 \lambda_1 = \lambda_1 = B(u, u) = \sum d_k^2 \lambda_k$, so $\sum(\lambda_k - \lambda_1)d_k^2 = 0$, so $d_k = 0$ for $k > 1$
3. if $u \in H_0^1(U)$ is a weak eigenfunction with eigenvalue λ_1, then it is either strictly positive or strictly negative
4. if u_1, u_2 are two weak solutions to $Pu = \lambda_1 u$, then there is $c \neq 0$ such that $\int u_1 - cu_2 dx = 0$. Since $u_1 - cu_2$ is a weak eigenfunction, by previous step, must be zero, therefore $u_1 = cu_2$, therefore λ_1 is a simple eigenvalue.

□

Theorem 1.73 (Principal Eigenvalue of nonsymmetric Elliptic Operators). *If $P = -a^{ij}\partial_i\partial_j + b^j\partial_j + c$ with $a, b, c \in C^\infty(\bar{U})$, $c \geq 0$, then*

1. *there exists a real eigenvalue λ_1 of P that is simple with an eigenfunction which is positive*
2. *if $\lambda \in \operatorname{spec}(P)$, then $\Re(\lambda) \geq \lambda_1$*

1.6 Second-order Parabolic Equations

Reference: Evans 7.1

Definition, existence of weak solutions, regularity, maximum principles

Summary of Results:

Data	Weak Solution Definition	Energy	Regularity
$f \in L^2(I; L^2), g \in H_0^1$	$u \in L^2(I; H_0^1), u' \in L^2(I; H^{-1})$	$\|u(t)\|_{L_x^2}$	$u \in L^\infty(I; H_0^1) \cap L^2(I; H^2), u' \in L^2(I; L^2)$
Strong Maximal:	subsolution s.t. $u(t_0, x_0) = \max_{\bar{U}_t} u$	u constant on U_{t_0}	proof by Harnack
Harnack:	$u \geq 0$	then	$\sup_{x \in V} u(t_1, x) \leq C \inf_{x \in V} u(t_2, x)$

Definition 1.18 (Parabolic PDE). *Given $U \subset \mathbb{R}^n$ open bounded connect. A parabolic pde is:*

$$\begin{cases} \partial_t u + Lu = f & (t, x) \in (0, T] \times U := U_T \\ u(t, x) = 0 & x \in \partial U \\ u(t, x) = g & t = 0 \end{cases}$$

with $Lu = -\sum_{i,j=1}^n \partial_i (a^{ij} \partial_j u) + \sum_{j=1}^n b^j \partial_j u + cu$ with a, b, c functions of some regularity, and $\sum_{i,j=1}^n a^{ij} \xi_i \xi_j \geq \theta |\xi|^2$ for all t, x, ξ with $\theta > 0$.

Definition 1.19 (Weak Solutions of Parabolic PDE). *A function $u \in L^2(0, T; H_0^1(U))$ with $u' \in L^2(0, T; H^{-1}(U))$ is a weak solution of the above parabolic PDE if $\langle u', v \rangle_{L_x^2} + B[u, v; t] = \langle f, v \rangle_{L_x^2}$ for almost every $t \in [0, T]$ and $u(0, x) = g$. Where $B[u, v; t]$ is the bilinear form from elliptic PDE but we fix t.*

1.6.1 Sobolev Spaces involving time

Here's a review of Sobolev spaces involving time. $u \in L^2(0, T; H_0^1(U))$ means $u(t, \cdot) \in H_0^1(U)$ for all $t \in [0, T]$ [a], and $\int \|u(t, x)\|_{H_0^1(U)}^2 \, dt < \infty$.

If $u \in L^1(0, T; X)$ (for X a Banach space) and there exists $u' \in L^1(0, T; X)$ such that $\int_0^T \varphi u' dt = -\int_0^T \varphi' u dt$ for all $\varphi \in C_0^\infty([0, T]; \mathbb{R})$, then we call u' the weak derivative.

If $u \in L^p[0, T; X]$ and $u' \in L^p[0, T; X]$, then we say $u \in W^{1,p}[0, T; X]$. In this case we get that for all $0 \leq s \leq t \leq T$:

$$u(t) = u(s) + \int_s^t u'(\tau) d\tau$$

which tells us that (1) $u \in C[0, T; X]$ and (2) $\|u\|_{C[0,T;X]} \leq C \|u\|_{W^{1,p}[0,T;X]}$.

[a] actually $u : [0, T] \to H_0^1(U)$ is strongly measurable, which means it is the almost everywhere pointwise limit (in t) of simple function that take values in $H_0^1(U)$

If $u \in L^2(0,T;H_0^1(U))$ and $u' \in L^2(0,T;H^{-1}(U))$ [a], then (1) $u \in C(0,T;L^2(U))$ with $\|u\|_{C(0,T;L^2(U))} \le C(\|u\|_{L^2(0,T;H_0^1(U))} + \|u'\|_{L^2(0,T;H^{-1}(U))})$ and (2) for almost every $t \in [0,T]$ $\frac{d}{dt}\|u(t)\|_{L^2(U)} = 2\langle u'(t), u(t)\rangle_{L^2(U)}$

Theorem 1.74 (Parabolic PDE Uniqueness). *Weak solutions to parabolic pde are unique.*

Proof. Show energy cannot grow via Gronwall It suffices to show that if u is a weal solution to the parabolic pde with $f = g = 0$, then $u = 0$.

1. By Sobolev properties $\frac{d}{dt}\|u(t)\|_{L^2(U)} = 2\langle u'(t), u\rangle_{L^2(U)} = -2B(u,u)$ by definition of weak solution
2. by same energy estimates as elliptic pde, $\|u(t)\|^2_{H_0^1(U)} \le C(B(u,u;t) + \|u(t)\|^2_{L^2(U)})$, so $-2B(u,u) \le -c\|u(t)\|^2_{H_0^1(U)} + 2\gamma\|u(t)\|^2_{L^2(U)} \le 2\gamma\|u(t)\|^2_{L^2(U)}$
3. By Gronwall's inequality, $\|u(t)\|^2_{L^2(U)} \le e^{2\gamma t}\|u(0)\|^2_{L^2(U)} = 0$

□

Theorem 1.75 (Gronwall's Inequality). *If $\eta'(t) \le \eta(t)\varphi(t) + \psi(t)$, with $\eta, \psi, \varphi \ge 0$, summable, and η absolutely continuous, then $\eta(t) \le e^{\int_0^t \varphi(s)ds}(\eta(0) + \int_0^t \psi(s)ds)$*

Proof. The idea is to pretend we have equality, solve it, then realize that most of equalities can be replaced by inequality.

1. $\frac{d}{ds}(e^{-\int_0^s \varphi(t)dt}\eta(s)) = e^{-\int_0^s \varphi(t)dt}(\eta'(s) - \varphi(s)\eta(s)) \le e^{-\int_0^s \varphi(t)dt}\psi(s)$
2. Integrate to get $e^{-\int_0^s \varphi(t)dt}\eta(s) \le \eta(0) + \int_0^s e^{-\int_0^x \varphi(t)dt}\psi(x)dx \le \eta(0) + \int_0^s \varphi(t)dt$
3. rearrange

□

1.6.2 Existence of Weak Solutions

Step 1. Find a weak solution to a finite dimensional approximation to our PDE.

Let w_k be an orthogonal basis in $H_0^1(U)$ and orthonormal in $L^2(U)$. For each $m \in \mathbb{N}$, we want to find $u_m \in L^2(0,T;H_0^1(U))$ such that

1. $u_m(t) = \sum_{k=1}^m d_k^m(t)u_m$ for some $d_k^m(t)$
2. $u_m(0) = g$
3. $(u_m', w_k)_{L^2} + B(u_m, w_k; t) = (f, w_k)_{L^2}$ for a.e. t and every $k = 1, \ldots, m$

[a] basically, taking a derivative in time loses two derivatives in space, this makes sense for parabolic pde as $\partial_t u = \Delta_x u$

This is easily shown, as we get a first order system of linear ode for which we can solve for $d_k^m(t)$.

Step 2. Provide bounds on approximations to pass to subsequence that weakly converges by Banach-Alaglou

Lemma 1.1. *u_m are bounded in $L^2(0,T;H_0^1(U))$.*

Proof. All the identities come from looking at $E(t) = \|u_m(t)\|_{L^2(U)}$, and computing $\frac{d}{dt}\|E(t)\|_{L^2(U)}$

1. $E(t) = 2(u_m', u)$. Since u_m is a weak solution to finite dimensional problem, multiply this expression by $d_k^m(t)$ and sum to get $(u_m', u_m) + B(u_m, u_m; t) = (f, u_m)$. So $E(t) = 2(f, u_m) - B(u_m, u_m; t)$

2. since $C_1\|u(t)\|^2_{H_0^1(U)} \leq B(u_m, u_m; t) + C_2\|u\|^2_{L^2(U)}$ and Young on other term, get $E(t) \leq \|f\|^2_{L^2} + \|u_m\|^2_{L^2} + C_2\|u_m\|^2_{L^2} - \|u_m\|^2_{H_0^1}$, rearrange to get:

$$\frac{d}{dt}\|u_m(t)\|_{L^2(U)} + C_1\|u_m(t)\|^2_{H_0^1(U)} \leq \|f(t)\|_{L^2(U)} + C_2\|u_m(t)\|_{L^2(U)}$$

3. by Gronwall, (throw away second term on RHS for now) get $\|u_m(t)\|^2_{L^2} \leq C(\|g\|_{L^2(U)} + \|f\|^2_{L^2(0,T;L^2(U)})$ (this is a uniform bound in t).

4. next use the second term, integrate over time, use uniform bound on $\|u_m(t)\|_{L^2}$ to get $\|u_m\|^2_{L^2(0,T;H_0^1)} \leq C(\|g\|^2_{L^2(U)} + \|f\|^2_{L^2(0,T;L^2(U)})$

□

Lemma 1.2. *u_m' are bounded in $L^2(0,T;H^{-1}(U))$*

1. let $v \in H_0^1(U)$ with $v = v_1 + v_2$, $\|v\|_{H_0^1} \leq 1$, and v_2 orthogonal to the span of $\{w_k\}$. (note $\|v_1\|_{H_0^1(U)} \leq 1$

2. $(u_m', v) = (u_m', v_1) = (f, v_1) - B(u_m, v_1; t)$, the modulus of which is bounded by $\|f(t)\|_{L^2(U)} + C\|u_m\|_{H_0^1}$

3. therefore $\|u_m'\|_{H^{-1}} \leq C(\|f(t)\|_{L^2(U)} + \|u_m\|_{H_0^1(U)})$. Integrate this, use above bound to get that $\|u_m'\|^2_{L^2(0,T;H^{-1}(U)} \leq C(\|f\|^2_{L^2(0,T;L^2(U))} + \|g\|^2_{L^2(U)})$

Step 3. Show that the weak limit converges to the correct thing.

1. relabel subsequences, u_m converges weakly to u in $L^2(0,T;H_0^1(U))$ and u_m' converges weakly to u' in $L^2(0,T;H^{-1}(U))$ (showing they are equal is a small exercise)

2. Let $v = \sum_{k=0}^N d_k(t)w_k(x) \in L^2(0,T;H_0^1)$ with $d_k(t)$ smooth (functions of this form are dense in $L^2(0,T;H_0^1)$, so it suffices to test u with v.

3. with $m > N$, we can sum the finite dimensional weak solution condition of u_m with $d_k(t)$ and integrate over time to get: $\int_0^T (u_m(t)', v(t))_2 + B(u_m(t), v(t); t) dt = \int_0^T (f(t), v(t))_2 dt$

4. by weak convergence, the same is true with u_m replaced by u[a]. But this implies that $(u'(t), v) + B(u, v; t) = (f(t), v)$ for all $v \in H_0^1(U)$, and a.e. $t \in [0, T]$[b]

5. constructing v s.t. $v(T) = 0$, then integrate by parts above to get $\int_0^T -(v', u) + B(u, v; t) dt = \int_0^T (f, v) dt + (u(0), v(0))$, but we can also integrate by parts the identity of u_m, send to infinity and match terms to get $(u(0), v(0)) = (g, v(0))$, therefore $u(0) = g$

1.6.3 Regularity

Theorem 1.76 (First Regularity Estimate for Parabolic PDE). *If the coefficients of L are smooth and don't depend on t, u is a weak solution to our parabolic pde, then*

$$ess \sup_{0 \le t \le T} \|u(t)\|_{H_0^1(U)} + \|u\|_{L^2(0,T;H^2(U))} + \|u'\|_{L^2(0,T;L^2(U))} \le C(\|f\|_{L^2(0,T;L^2(U))} + \|g\|_{H_0^1(U)})$$

Proof. test Galerikin against u_m' to control $\|u'\|_{L^2}$, integrate, to get estimates on $\|u'\|_{L^2(I,L^2)}$ and $\|u\|_{L^\infty(I,H_0^1)}$, send m to infinity. Get pointwise weak solution, test against v, move over to get elliptic pde, use regularity to get bound on $\|u\|_{H^2}$, integrate to get result.

1. Galerkin approximations satisfy $(u_m', u_m') + B(u_m, u_m'; t) = (f, u_m')$, expand B, write first term as $\frac{1}{2}\frac{d}{dt}A(u_m, u_m) = \int_U a^{ij} \partial_i u_m' \partial_j u_m$

2. Peter-Paul everything to get $\|u_m'\|_{L^2(U)}$ terms small, then integrate to get: $\int_0^T (u_m', u_m') + \sup_{0 \le t \le T} A(u_m(t), u_m(t)) \le C(A(u_m(0), u_m(0)) + \int_0^T \|f(t)\|_{L^2(U)}^2 + \|u_m\|_{H_0^1(U)}^2)$

3. $u_m(0) = g$, and A is elliptic, so we get: $\|u_m'\|_{L^2(0,T;L^2(U))} + \sup_{0 \le t \le T} \|u_m\|_{H_0^1(U)} \le C(\|g\|_{H_0^1(U)}^2 + \|f\|_{L^2(0,T;L^2(U))}^2 + \|u_m\|_{L^2(0,T;H_0^1(U))}^2)$. But u_m is bounded by exactly what we want (from step 2 of existence).

4. let $m \to \infty$

5. since $(u', v) + B(u, v; t) = (f, v)$ for all $v \in H_0^1(U)$, we have $B(u, v; t) = (f - u', v)$, that is u is a weak solution to an elliptic pde: $f - u' \in L^2$, therefore $u \in H^2$ with:

$$\|u(t)\|_{H^2(U)} \le C(\|f - u'\|_{L^2(U)} + \|u\|_{L^2(U)})$$

6. integrate, use above to get $\|u\|_{L^2(0,T;H^2(U))} \le C(\textit{what we want})$

□

[a] it is nontrivial that $B(u_m, v; t) \to B(u, v; t)$

[b] this is nontrivial to me, my sense is that if this was not true, we can construct $d(t)$ a bump function and get a contradiction

Theorem 1.77 (Higher Regularity of Parabolic PDE). *Given $g \in H^{2m+1}(U)$ and $\partial_t^k f \in L^2(0,T;H^{2m-2k})$ for $k = 0,\dots,m$ (with compatibility conditions), then $\partial_t^k u \in L^2(0,T;H^{2m+2-2k}(U))$ for $k = 0,\dots,m+1$ (with the expected control on norms).*

The compatability conditions are: $g_0 := g \in H_0^1$, $g_1 := f(0) - Lg_0 \in H_0^1(U)$, . . . , $g_m := \partial_t^{m-1} f(0) - Lg_{m-1} \in H_0^1(U)$.

1.6.4 Maximum Principals

Theorem 1.78 (Weak Maximal Principal for Parabolic PDE). *If $u \in C_1^2(U_T) \cap C(\bar{U}_T)$ is a subsolution*[a] *of the operator $(\partial_t + L)$ with $c \equiv 0$, then $\max_{\bar{U}_T} u = \max_{\Gamma_T} u$*[b]

Proof.

1. first if $(\partial_t + L)u < 0$: if there is an interior maximum with $t < T$, then $u_t = 0$ and by same argument as elliptic pde, $Lu \geq 0$, so we have a contradiction. If $t = T$, then $u_t \geq 0$, and we still get a contradiction.

2. if $(\partial_t + L)u \leq 0$, let $u^\varepsilon = u - \varepsilon t$, then $(\partial_t + L)u^\varepsilon = (\partial_t + L)u - \varepsilon < 0$, so $\max_{\bar{U}_T} u^\varepsilon = \max_{\Gamma_T} u^\varepsilon$. Send $\varepsilon \to \infty$.[c]

□

Another variation involves $c \geq 0$, in this case subsolutions satisfy $\max_{\bar{U}_T} u \leq \max_{\Gamma_T} u^+$ and supersolutions satisfy $\min_{\bar{U}_T} u \geq -\max_{\Gamma_T} u^-$. The way to prove the first is assume false, then there exists interior point whose value is $> \max_{\Gamma_T} u^+ = \max(\max_{\Gamma_T} u, 0)$. It is therefore maximum, and positive. Then we do the same argument.

Theorem 1.79 (Harnack's Inequality for Parabolic PDE). *If $u \geq 0$ solves $(\partial_t + L)u = 0$ in U_T, $V \Subset U$, $0 < t_1 < t_2 \leq T$, then there exists $C > 0$ such that*

$$\sup_{x \in V} u(t_1, x) \leq C \inf_{x \in V} u(t_2, x)$$

To remember order: heat dissipates, so $\sup u(x,t_1) \geq \sup u(x,t_2)$. Harnack deals with controlling variation, so we want to bound early peak with later (inverse) peak.

The proof is very long.

Theorem 1.80 (Strong Maximal Principal for Parabolic PDE). *If $(\partial_t + L)u \leq 0$ in U_T, and $u(t_0, x_0) = \max_{\bar{U}_T} u := M$ for $(x_0, t_0) \in U_T$, then u is constant on U_{t_0}.*

Proof. idea is to show u is zero on parabolic boundaries of sets containing our maximum. This requires setting up Harnacks inequality to conclude $M - u = 0$ on the boundary.

1. Let $(x_0, t_0) \in W \Subset U$, let v solve $(\partial_t + L)v = 0$ in W_T and $v = u$ on the parabolic boundary of W_T (call it $\tilde{\Gamma}_T$)

2. Claim: $v(t_0, x_0) = M$

[a] subsolution of P means $Pu \leq 0$

[b] recall $U_T = (0,T] \times U$, $\Gamma_T = \bar{U}_T \setminus U_T = U \times \{t = 0\} \cup \partial U \times (0,T]$ (it's a cup)

[c] u^ε converges uniformly to u. So if $\max_{\bar{U}_T} \neq \max_{\Gamma_T}$, then we get $u(t_0, x_0) > \max_{\Gamma_T} u$. Let ε be small such that $|u^\varepsilon - u| < \frac{1}{2}(u(t_0, x_0) - \max_{\Gamma_T} u)$

(a) Note $(\partial_t + P)(u - v) \le 0$ and $u - v$ is zero on the boundary, so the weak maximal principal says that $u \le v$.

(b) also the weak maximal principal says $v \le M$, therefore $M = u(t_0, x_0) \le v(t_0, x_0) \le M$

3. $\tilde{v} := M - v \ge 0$ and solves $(\partial_t + L)\tilde{v} = 0$ with zero boundary condition. By Harnack, for each $x_0 \in V \Subset W$, $0 < t < t_0$:

$$0 \le \sup_V \tilde{v}(t, x) \le C \inf_V \tilde{v}(t_0, x) = 0$$

4. therefore $\tilde{v} = 0$ on W_{t_0}, so $v = M$ on W_{t_0}, therefore $u = M$ on $\partial W \times [0, t_0]$.

□

1.7 Hyperbolic Equations

Reference: Evans 7.2-7.3

Second-order hyperbolic equations: definitions, energy estimates, energy-momentum tensor, finite speed of propagation, regularity; hyperbolic systems of first-order equations: definitions, existence and uniqueness of weak solution

Summary of Results:

Data	Weak Solution Definition	Regularity
$f \in L^2(I : L^2), g \in H_0^1, h \in L^2$	$u \in L^2(I; H_0^1), u' \in L^2(I; L^2), u'' \in L^2(I; H^{-1})$	$u \in L^\infty(I; H_0^1), u' \in L^\infty(I, L^2)$

1.7.1 Definitions

The hyperbolic PDE I will consider will look for solutions $u : \bar{U}_T \to \mathbb{R}$ ($U \subset \mathbb{R}^d$ open, $U_T = (0, T] \times U$) satisfying:

$$\begin{cases} (\partial_{tt} + L)u = f & (t, x) \in U_T \\ u = 0 & (t, x) \in [0, T] \times \partial U \\ u(0, x) = g \\ u_t(0, x) = h \end{cases}$$

with $Lu = -\sum \partial_{x_j}(a^{ij}\partial_i u) + \sum b^j \partial_j u + cu$, uniformly elliptic.

Definition 1.20 (Weak Solution of Hyperbolic PDE). *For $f \in L^2(U_T), g \in H_0^1(U)$, $h \in L^2(U)$, u is a weak solution to our hyperbolic pde if $u \in L^2(0, T; H_0^1(U))$, $u' \in L^2(0, T; L^2(U))$, $u'' \in L^2(0, T; H^{-1}(U))$ and satisfies $(u''(t), v) + B(u, v; t) = (f(t), v)$ for all $v \in H_0^1(U)$ and almost every $t \in [0, T]$. And $u(0) = g(0)$, $u'(0) = h$.*

To remember spaces: bilinear form suggests testing with H_0^1, this requires $u''(t) \in H^{-1}$. We need to take weak derivatives of u, so $u(t) \in H_0^1$ is also needed. In Fourier space, with $L = -\Delta$, $\widehat{u}(t, \xi) = \cos(|\xi|t)\widehat{g} + |\xi|^{-1}\sin(|\xi|t)h$. For $u \in H_0^1$, we need $g \in H_0^1$ and $h \in L^2$ (looking at decay as $\xi \to \pm\infty$). Morever, differentiating this in time brings a factor of $|\xi|$, so that $u' \in L^2$

1.7.2 Existence, Uniqueness, and Regularity

It turns out that if u lives in the above space, then $u \in C(0,T;L^2)$ and $u' \in C(0,T;H^{-1})$

One way to prove existence of weak solutions is via the Galerkin approximation. The construction of such functions is identical to parabolic pde.

We ultimately get the energy estimate on these approximations $u_m(t,x)$

$$\max_{t\in[0,T]}(\|u_m\|_{H_0^1} + \|u_m'\|_{L^2}) + \|u_m''\|_{L^2(0,T;H^{-1})} \le C(\|f\|_{L^2(0,T;L^2)} + \|g\|_{H_0^1} + \|h\|_{L^2})$$

Note: this is slightly stronger than we would expect, as u_m, u_m' are actually bounded in their respective Banach spaces

Proof. 1. test against u_m': $(u_m'', u_m') + B(u_m, u_m'; t) = (f, u_m')$ □

We then pass to a weak limit to get $u \in L^\infty(0,T;H_0^1), u' \in L^\infty(0,T;L^2), u'' \in L^2(0,T;H^{-1})$. After some work it turns out that this weak solution is unique.

We furthermore have regularity results. If $\frac{d^k}{dt^2} f \in L^2(0,T;H^{m-k})$ for $k = 0,\dots,m$, $g \cap H^{1+m}$, $h \in H^m(U)$, with certain compatability conditions, then $\frac{d^k}{dt^k}u \in L^\infty(0,T;H^{1+m-k}(U))$ for $k = 0,\dots,m+1$. And we can control these norms of u by exactly what we expect.

1.7.3 Finite Speed of Propagation

Theorem 1.81 (Finite Speed of Propagation for Hyperbolic PDE). *If u is a smooth solution to $Lu = -a^{ij}\partial_{x_i}\partial_{x_j}u = -\partial_t u$ (a independent of time), (x_0,t_0) fixed, $q > 0$ a smooth solution (on $\mathbb{R}^n \setminus \{x_0\}$) of*

$$\begin{cases} a^{ij}q_{x_i}q_{x_j} = 1 \\ q(x_0) = 0 \end{cases}$$

and $C_t = \{x : q(x) < t_0 - t\}$ $(0 \le t < t_0)$, $C = \cup_{0\le t<t} C_t$.

Then if $u \equiv u_t \equiv 0$ on C_0, then $u \equiv 0$ on C.

Proof. 1. let $e(t) = \frac{1}{2}\int_{C_t} u_t^2 + a^{ij}u_{x_j}u_{x_i}dx$, then $\dot{e}(t)$ will have one term where derivative goes in the integral, and the second where the derivative hits the boundary, call these terms I_1 and I_2.

2. $I_1 = \int_{C_t} u_t u_{tt} + a^{ij}u_{x_j}u_{x_i,t}dx = \int_{C_t} u_t u_{tt} - \partial_{x_i}(a^{ij}u_{x_j})u_t dx + \int_{\partial C_t} a^{ij}u_{x_j}\nu_j u_t dS$

   (a) first two terms give just: $-\int_{C_t}(\partial_{x_i}a^{ij})u_{x_j}u_t dx$

   (b) third term: $\partial C_t = \{q(x) + t - t_0 = 0\}$, so $\nu_j = |\nabla q|^{-1}\partial_{x_j}q$, get:

   $$|\int_{\partial C_t} |\nabla q|^{-1}\langle A\nabla u, \nabla q\rangle u_t dS| \le \int_{\partial C_t} |\nabla q|^{-1}(|\langle A\nabla u, \nabla u\rangle|^{1/2}|\langle A\nabla q, \nabla q\rangle|^{1/2})u_t dS$$

   $|\langle A\nabla q, \nabla q\rangle|^{1/2} = 1$, so the integral is majorized by $\int_{\partial C_t} |\nabla q|^{-1}(\frac{u_t^2}{2} + \frac{a^{ij}\partial_{x_i}u\partial_{x_j}u}{2})$

(c) ultimately: $|I_1| \le C \int_{C_t} |\nabla u||u_t|dx + \frac{1}{2}(u_t^2 + a^{ij}\partial_{x_i} u \partial_{x_j} u)\frac{dS}{|\nabla q|}$ (by Young's inequality and ellipticity, the first term is bounded by $Ce(t)$.

3. use co-area formula[a], $I_2 = -\frac{1}{2}\int_{\partial C_t}(u_t^2 + a^{ij}u_{x_j}u_{x_i})\frac{dS}{|\nabla Q|}$

4. Therefore $\dot{e}(t) \le Ce(t)$. Apply Gronwall, get $e(t) = 0$ for all t.

□

1.7.4 Energy Momentum Tensor

Definition 1.21 (Energy Momentum Tensor). *For $u \in C^\infty(\mathbb{R}^{n+1})$, define the energy momentum tensor as $T^{\alpha\beta} = \partial^\alpha u \partial^\beta u - \frac{1}{2}m^{\alpha\beta}\partial^\gamma u \partial_\gamma u$. Where $m_{\alpha\beta} = diag(-1, 1, \ldots, 1)$, $m^{\alpha\beta} = m^{-1}_{\alpha\beta}$ $(= m_{\alpha\beta})$, $\partial^\alpha = m^{\alpha\beta}\partial_\beta$.*

Theorem 1.82 (Divergence Free Energy Momentum Tensor). *If $\Box u = 0$, then $\partial_\alpha T^{\alpha\beta} = 0$*

This is an easy computation, just remember $\Box u = -\partial_\alpha \partial^\alpha u$.

From here, we can see that $\partial_0 T^{0\beta} = -\partial_j T^{j\beta}$, so we if we integrate over space, and apply the divergence theorem (assuming compact support of u), then we see: $\partial_t \int_{\mathbb{R}^n} T^{0\beta}dx = 0$ for all $\beta = 0, \ldots, n+1$.

If $\beta = 0$, we have $T^{00} = \frac{1}{2}(u_t^2 + |Du|^2)$, which is our usual **energy**. If $\beta = j$, we have $T^{0j} = u_t \partial_j u$ which is sometimes called **momentum**.

We also have, for every $x \in \mathbb{R}^{n+1}$, $\partial_\alpha T^{\alpha\beta}x_\beta = 0$, if x is constant, then we have $\partial_\alpha(T^{\alpha\beta}x_\beta) = 0$, so by the same argument as above, we get conserved (over time) quantities $\int_{\mathbb{R}^n} T^{0\beta}x_\beta dx$.

It can be shown that $T^{0\beta}x_\beta > 0$ if and only if[b] $x_0 > 0$ and $x_1^2 + \cdots x_n^2 < x_0^2$ (a vector satisfying this with respect to our (Minkowski metric) is called **forward time-like**).

Integrating $\partial_\alpha T^{\alpha\beta}x_\beta$ over the region $\{(t,x) \in \mathbb{R}^{d+1} : t \in [t_0, t_1]\}$ and applying the divergence theorem gives the relation:

$$\int_{t=t_0} T^{0\beta}x_\beta dx = \int_{t=t_1} T^{0\beta}x_\beta dx$$

More generally, if Σ_0 and Σ_1 are two surfaces in $\mathbb{R}^{d+1}$, then we can again integrate $\partial_\alpha T^{\alpha\beta}x_\beta$ over the region between these two surfaces and apply the divergence theorem to get:

$$\int_{\Sigma_0} N_\alpha T^{\alpha\beta}x_\beta dx = \int_{\Sigma_1} N_\alpha T^{\alpha\beta}x_\beta dx$$

With N the normal vector of the surface. These quantities are positive definite if and only if x and N are both forward time-like or both backward time-like.

[a] $\partial_r \int_{\{q(x)\le r\}} f(y)|Dq(x)|dy = \int_{\{q(x)=r\}} f(y)dS(y)$

[b] backwards direction is pretty easy, forwards may be harder

1.7.5 Systems of Hyperbolic PDE

Definition 1.22 (System of Hyperbolic PDE). *A hyperbolic system of PDEs is of the form:*

$$\begin{cases} u_t + \sum B^l \partial_l u = f & (x,t) \in \mathbb{R}^n \times (0,T] \\ u = g & (x,t) \in \mathbb{R}^n \times \{t = 0\} \end{cases}$$

where $u = (u_1, \ldots, u_m) : \mathbb{R}^n \to \mathbb{R}^m$, $f : \mathbb{R}^n \to \mathbb{R}^m$, $g : \mathbb{R}^n \to \mathbb{R}^m$, $B^l(x,t) \in M^{m\times m}(\mathbb{R})$ *for* $l = 1, \ldots, n$*. Such that for all* $y \in \mathbb{R}^n$, $\sum y_j B^j$ *is diagonalizable with real eigenvalues*[a]*.*

Furthermore if:

1. B_j *are all symmetric, our system is called symmetric*
2. $\sum y_j B^j$ *have* n *distinct real eigenvalues, then our system is strictly hyperbolic.*

Definition 1.23 (Hyperbolic System Weak Solution). *A weak solution to the above system of hyperbolic PDEs*[b] *is* $u \in L^2([0,T]; H^1(\mathbb{R}^n; \mathbb{R}^m))$, $u' \in L^2([0,T]; L^2(\mathbb{R}^n; \mathbb{R}^m))$*, such that for all* $v \in H^1(\mathbb{R}^n; \mathbb{R}^m)$*:*

$$\langle u', v \rangle + B(u, v; t) = \langle f, v \rangle$$

for a.e. $t \in [0,T]$*, and* $u(0) = g$*. Where* $B(u,v;t) = \int_{\mathbb{R}^n} \sum B^j \partial_{y_j} u \cdot v dx$

The method to prove the existence of weak solutions is to first solve a related PDE with an added Laplacian term with a parameter ε. These solutions are called viscosity solutions. Then we establish energy estimates on the viscosity solutions, which allows us to pass to a weak subsequence, and show that the weak limit is a weak solution.

Step 1. Prove existence of viscosity solutions via fixed point argument.

1. wish to solve $\partial_t u - \varepsilon \Delta u + B^j \partial_j u = f$ with $u(0,x) = \eta_\varepsilon * g := g^\varepsilon$. For each $v_1, v_2 \in L^\infty([0,T]; H^1)$ we can solve $\partial_t u - \varepsilon \Delta u = f - B^j \partial_j v_{1,2}$ via fundamental solutions of the heat equation, denote the solutions u_1^ε and u_2^ε.
2. Let $\widehat{u^\varepsilon} = u_1^\varepsilon - u_2^\varepsilon$, then this solves the ε-heat equation with forcing term $-B^j(v_1 - v_2)$, by energy estimates, can get $\|u^\varepsilon(t)\|_{H^1} \le C(\varepsilon) T^{1/2} \|v_1 - v_2\|_{L^\infty([0,T];H^1)}$
3. force ε small enough to get $C(\varepsilon)T^{1/2} \le 1/2$, so that $v \mapsto u$ has a fixed point, denote it u^ε (which solves our pde).
4. these viscosity solutions u^ε are unique and $u^\varepsilon \in L^2 H^3$, $(u^\varepsilon)' \in L^2 H^1$

Step 2. Establish energy estimates on viscosity solutions

[a] To remember this, consider constant coefficient case with $f = 0$, then taking the Fourier transform, we see that $\widehat{u}(t,\xi) = \widehat{g}(\xi) \exp(-itB^l \xi_l)$. It would be nice if $B^l \xi_l$ was diagonalizable with real coefficients.

[b] here we require B_j symmetric, in C^2. $f \in H^1(\mathbb{R}^n \times (0,1); \mathbb{R}^m)$, $g \in H^1(\mathbb{R}^n; \mathbb{R}^m$

1. wish to prove $\|u^\varepsilon(t)\|_{L^\infty H^1} + \|(u^\varepsilon)'\|_{L^2L^2} \lesssim \|g\|_{H^1} + \|f\|_{L^2H^1} + \|f'\|_{L^2L^2}$

Step 3. Prove the weak limit of viscosity solutions is a solution.

Theorem 1.83 (Solution of Constant Coefficient Hyperbolic System). *There exists a unique solution to $u_t + B^l\partial_l u = 0$, $u(0) = g$ for $g \in H^s(\mathbb{R}^n;\mathbb{R}^m)$, $s > \frac{n}{2} + m$, B constant coefficients. And $u \in C^1([0,\infty);\mathbb{R}^m)$*

Proof.

1. we can see that $\widehat{u}(t,\xi) = e^{-itB(\xi)}\widehat{g}(\xi)$ with $B(\xi) = \sum A^l\xi_l$
2. Claim: $e^{-itB(\xi)} = \frac{1}{2\pi i}\oint_\gamma e^{-itz}(zI - B(\xi))^{-1}dz$ for γ a contour in $\mathbb{C}$ containing the eigenvalues of B
3. Claim: if the contour is circles of radius 1 around eigenvalues of $B(\xi)$, then $\|z - B(\xi)\| \le C\langle\xi\rangle^{m-1}$
4. Therefore $\|e^{-itB(\xi)}\| \le Ce^t\langle\xi\rangle^{m-1}$, and so the representation of u (taking the inverse Fourier transform) converges.

□

1.8 Pseudodifferential Operators

Reference: Grigis and Sjöstrand 1,3,4

Oscillatory integrals, basic calculus of pseudodifferential operators, parametrix construction

1.8.1 Symbols and Oscillatory Integrals

Definition 1.24 (Symbols). *The space of symbols, $S^m_{\rho,\delta}(X,\mathbb{R}^N)$ ($X \subset_{open} \mathbb{R}^n$, $\rho,\delta \in [0,1]$, $m \in \mathbb{R}$) are functions $a \in C^\infty(X,\mathbb{R}^N)$ such that for all $K \Subset X$, α,β multiindices:*

$$\sup_{(x,\xi)\in K\times\mathbb{R}^N} |\partial_x^\alpha\partial_\xi^\beta a(x,\xi)| \le C\langle\xi\rangle^{m-\rho|\beta|+\delta|\alpha|}$$

The best constants form semi-norms[a] make $S^m_{\rho,\delta}$ a Frechet space.

To work with symbols, we often need to approximate symbols by rapidly decaying symbols living in $S^{-\infty}$

Theorem 1.84 (Density of $S^{-\infty}$ in $S^m_{\rho,\delta}$). *For all $a \in S^m_{\rho,\delta}$, there exist $a_j \in S^{-\infty}$ such that $a_j \to a$ with respect to all semi-norms of $S^{m+\varepsilon}_{\rho,\delta}$ for all $\varepsilon > 0$.*

Proof. The idea is to let $a_j = \chi(\theta\varepsilon)a$ with $\chi \in C_0^\infty$, 1 near 0. Then a nontrivial fact is that if (1) $a_j \in S^m$ are bounded in S^m and (2) converge pointwise to some a (to some arbitrary! function), then $a \in S^m$ and $a_j \to a$ in $S^{m+\varepsilon}$. □

[a] $\|a\|_{\alpha,\beta,k} := \sup_{(x,\xi)\in X\times\mathbb{R}^N}\langle\xi\rangle^{-m+\rho|\beta|-\delta|\alpha|}|\partial_x^\alpha\partial_\xi^\beta a(x,\xi)|$

Definition 1.25 (Phase Function). *A phase function is some* $\varphi \in C^\infty(X \times \dot{\mathbb{R}}^N)$ [a] *such that for all* $(x, \theta) \in X \times \dot{\mathbb{R}}^N$

1. $\varphi(x, \lambda\theta) = \lambda\varphi(x, \theta)$ *for all* $\lambda > 0$
2. $d\varphi \neq 0$
3. $\Im\varphi \geq 0$

motivation is to generalize $\varphi(x, \theta) = (x_1 - x_2) \cdot \theta$, where $x = (x_1, x_2)$. Homogeneity is natural, and gives symbol control[b]. Nonvanishing of the differential allows integrating by parts. And positivity of the imaginary part avoids exponential growth of $e^{i\varphi}$.

Definition 1.26 (Oscillatory Integral). *For a symbol* $a \in S^m_{\rho,\delta}(X, \mathbb{R}^N)$ *and phase function* φ*, we get the oscillatory integral* $I(a, \varphi)$ *which can be formally written* $\int e^{i\varphi(x,\theta)} a(x, \theta) d\theta$

Note these integrals don't make sense, as they do not converge. One motivation is to generalize fourier transforms on tempered distributions. In the sense of distributions, the inverse fourier transform of 1 is the dirac delta function, so we could write, although the integral doesn't make sense:

$$\delta_0(x) = \frac{1}{(2\pi)^n} \int e^{ix\cdot\xi} d\xi$$

Theorem 1.85 (Existence of Oscillatory Integrals). *We can define* $I(a, \varphi) \in \mathcal{D}'(X)$ *in such a way that agrees with when the integral converges* ($m < -N$) *and* $a \mapsto I(a, \varphi)$ *is continuous.*[c]

Explicitly, $a \mapsto I(a, \varphi)$ is continuous if for all $\varepsilon > 0$ and $u \in C_0^\infty$, there exist A, B finite sets of multiindicies, $K \subset X$, and $\delta > 0$ such that $\sum_{\alpha\in A, \beta\in B} \|a - \tilde{a}\|_{\alpha,\beta,K} < \delta$ implies that $|\langle I(a, \varphi), u\rangle| < \varepsilon$

Proof. Integrate by parts using an operator with enough regularity to lower the symbol class. Use density of $S^{-\infty}$ and joint continuity of integration by parts to show the result is well-defined.

1. There exists $L = \sum a_j \partial_{\theta_j} + \sum b_j \partial_{x_j} + c$ with $a_j \in S^0_{1,0}$, $b_j, c \in S^{-1}_{1,0}$ such that $L^t(e^{i\varphi}) = e^{i\varphi}$ (the real transpose of L)
   (a) set $\Phi = |\theta|^2 \nabla_\theta \bar{\varphi} \cdot \nabla_\theta \varphi + \nabla_x \varphi \cdot \nabla_x \varphi$, this is smooth, nonzero for $\theta \neq 0$, positively homogeneous of degree 2 in θ, and therefore in $S^2_{1,0}$[d]
   (b) let $\chi(\theta) \in C_0^\infty(\mathbb{R}^N)$ be identically 1 near 0, then define $L^t = \frac{(1-\chi)}{i\Phi}(\sum_{j=1}^n (|\theta|^2 \partial_{\theta_j} \bar{\varphi} \partial_{\theta_j} + \partial_{x_j} \bar{\varphi} \partial_{x_j})) + \chi(\theta) = \sum a_j \partial_{\theta_j} + b_j \partial_{x_j} + c$.
   (c) Φ is nonzero away from zero, so $\Phi^{-1} \in S^{-2}$, $|\theta|^2 \in S^2$, $\partial_{\theta_j} \varphi \in S^0$, therefore $a_j \in S^0$. Similarly, $\partial_{x_j} \varphi \in S^1$, so $b_j \in S^{1-2} = S^{-1}$ and $\chi \in S^{-\infty} \subset S^{-1}$

[a] $\dot{R}^n := \mathbb{R}^n \setminus \{0\}$
[b] if $a(x, \theta)$ is homogeneous of degree m in θ, then $a \in S^m$
[c] Moreover, the order of the distribution is the smallest k such that $m - tk < -N$ where $t = \min(\rho, 1 - \delta)$
[d] differentiation in θ lowers degree by 1, that's why we need $|\theta|^2$

2. Note $L^k(au)$ (for $u \in C_0^\infty(X)$, $a \in S^m_{\rho,\delta}$) belongs to $S^{m-kt}_{\rho,\delta}$ (and is in fact continuous) (for $t = \min(\rho, 1-\delta)$
   (a) $L(au) = aL(u) + L(a)u$, the terms of $L(u)$ either vanish or are in $S^{-1}_{1,0} \cdot S^0_{1,0} = S^{-1}_{1,0}$[a].
   (b) $L(a) = a^j \partial_{\theta_j} a + b^j \partial_{x_j} a + ca$. $\partial_{\theta_j} a \in S^{m-\rho}_{\rho,\delta}$, so $a^j \partial_{\theta_j} a \in S^{m-\rho}_{\rho,\delta}$. Similarly, $L(a) \in S^{m+\delta-1}_{\rho,\delta}$, $ca \in S^{m-a}_{\rho,\delta}$. So just iterate this.
3. for $a \in S^m_{\rho,\delta}$, let k be such that $m - kt < -N$, and define $I_k(a,\varphi) \in \mathcal{D}'(X)$ as $\langle I_k(a,\varphi), u\rangle = \iint e^{i\varphi} L^k(au)dxd\theta$. (This is a distribution as it has finite order on every compact set, continuity of L^k shows that $a \mapsto I(a,\varphi)$ is continuous).
4. This is well-defined (no matter which k or L we choose[b] and continuity of L), so we just define $I(a,\varphi) = I_k(a,\varphi)$

□

Proposition 1.8. *We can compute $I(a,\varphi)$ as:*

$$I(a,\varphi) = \lim_{\substack{\varepsilon\to 0\\ \mathcal{D}'}} \int e^{i\varphi(x,\xi)} a(x,\xi)\chi(\xi\varepsilon)d\xi$$

for $\chi \in \mathcal{S}$ with $\chi(0) = 1$.

Proof. This can be used as alternative definition of oscillatory integrals, but note that $a(x,\xi)\chi(\xi\varepsilon) \in S^{-\infty}$ and approximate a, so this is essentially the same proof as above, but more concrete with the approximation.

1. pick $u \in C_0^\infty(\mathbb{R}^n)$, for each $\varepsilon > 0$, $\int e^{i\varphi} a\chi(x\varepsilon)udxd\xi$ converges
2. use the same operator L as above, and let k be such that $m - kt < -N$, then integrate by parts with L^k. When no derivatives fall on χ, we get $\int e^{i\varphi} L^k(au)\chi(\varepsilon\xi)dxd\xi$, this converges by DCT as $\varepsilon \to 0$, and is exactly what we want.
3. the rest of the terms are of the form $\varepsilon^j d(\varepsilon\xi)f(x,\xi)$ for $j = 1,\dots,k$ where $d \in C_0^\infty$ is supported away from zero (let's say $B_2(0) \setminus B_1(0)$), $f \in S^{m-(k-j)t}$, and $\Pi_x \operatorname{supp} f \subset \operatorname{supp} u$, then we can compute the magnitude of the integral of each of these terms:

$$\varepsilon^j \left| \int e^{i\varphi} d(\varepsilon\xi) f(x,\xi)dxd\xi \right| \le \varepsilon^j \sup_{\substack{x\in\operatorname{supp} u\\ \frac{1}{\varepsilon}\le|\xi|\le\frac{2}{\varepsilon}}} |f(x,\xi)| \le C\varepsilon^j \left(1 + |\frac{2}{\varepsilon}|\right)^{m-(k-j)t}$$

   collecting the exponents, and recalling that $m - kt < -N$, we see that the exponent of ε is positive, so as $\varepsilon \to 0$, all these terms converge to zero.

□

[a] A function with no dependence on θ is in $S^0_{1,0}$

[b] this is by density of $S^{-\infty}$

Definition 1.27 (Critical Set of Phase). *The critical set of the phase* φ *is*

$$C_\varphi := \left\{(x,\theta) \in X \times \dot{\mathbb{R}}^N : \nabla_\theta \varphi = 0\right\}$$

Theorem 1.86 (Singular Support of Oscillatory Integral). *If* $a \in S^m_{\rho,\delta}(X, \dot{\mathbb{R}}^N)$ *vanishes on* C_φ*, then* $I(a,\varphi)$ *is smooth. This implies that* singsupp $I(a,\varphi) \subset \Pi_X C_\varphi$

Proof. 1. Let $\chi(x,\theta) \in C_c^\infty$ have support contained in $C_\varphi \cap \operatorname{supp} A^c$

2. Let $L = \frac{(1-\chi)}{i|\nabla_\theta \varphi|^2} \sum \partial_{\theta_j} \bar{\varphi} \partial_{\theta_j}$, then $Le^{i\varphi} = (1-\chi)e^{i\varphi}$ and the vector field of L has components in $S^0_{1,0}(X, \dot{\mathbb{R}}^N)$ (note $a\chi = 0$)

3. Then formally (we need to pair) $\int e^{i\varphi} a d\theta = \int e^{i\varphi} a(1-\chi)^k d\varphi = \int (L^k e^{i\varphi}) a d\theta = \int e^{i\varphi}(L^k a) d\theta$. If k is large enough, this converges and is in $C^m(X)$ for all m, and is therefore smooth.

□

When $\varphi = (x-y)\cdot\theta$, we see that the singular support of $I(a,\varphi) = \Delta := \{(x,y) \in \mathbb{R}^n : x = y\}$. Hueristically: the 'action' of the oscillatory integral occurs on the diagonal. So we can decompose an oscillatory integral into a part near the diagonal and a part away from the diagonal. The part away from the diagonal is just a smooth function. Another hueristic, is the only nonsingular action is for θ near ∞

Theorem 1.87 (Schwartz Kernel Theorem). *There is a bijection between* $K \in \mathcal{D}'(X \times X)$ *and continuous maps* $A : C_0^\infty(Y) \to \mathcal{D}'(X)$*:* $\langle Au, v\rangle_X = \langle K, v \otimes u\rangle_{X\times Y}$.

So given a symbol $a \in S^m_{\rho,\delta}(X \times Y \times \mathbb{R}^N)$ and phase φ, we have above that $I(a,\varphi) \in \mathcal{D}'(X \times Y)$. This is the **Distributional Kernel** of the operator A, defined as $Au(x) = \langle \int e^{i\varphi} u(x) a, \cdot \rangle$ for $u \in C_0^\infty(Y)$.

1.8.2 Pseudo-differential Operators

Definition 1.28 (Pseudodifferential Operator). *Given a symbol* $a \in S^m_{\rho,\delta}(X \times \mathbb{R}^N)$ *with* $X = \mathbb{R}^N \times \mathbb{R}^N$*, and* $\varphi = \langle \theta, x - y\rangle$*, then if* $K = I(a,\varphi) \in \mathcal{D}'(\mathbb{R}^n \times \mathbb{R}^n)$ *is the Schwartz Kernel of* A*, then* A *is a psuedodifferential operator.*

Explicitly, for each $u \in C_0^\infty(\mathbb{R}^n_y)$ *and* $v \in C_0^\infty(\mathbb{R}^n_x)$*, then (after normalizing)*

$$\langle Au, v\rangle = \frac{1}{(2\pi)^n} \int e^{i(x-y)\cdot\theta} a(x,y,\theta) u(y) v(x) dy d\theta dx$$

The space of such operators is denoted $L^m_{\rho,\delta}(\mathbb{R}^n)$

Theorem 1.88 (Mapping Property of PDOs on C_0^∞). $L^m_{\rho,\delta}(\mathbb{R}^n) \ni A : C_0^\infty(\mathbb{R}^n_y) \to C^\infty(\mathbb{R}^n)$

Proof. 1. Let $\chi(\theta)$ be supported near 0, let

$$L = \chi(\theta) + \frac{1-\chi(\theta)}{i|\theta|^2(1+|x-y|^2)} \sum_{j=1}^n [|\theta|^2 (x_j - y_j)\partial_{\theta_j} + \theta_j \partial_{y_j}]$$

then $Le^{i\varphi} = e^{i\varphi}$ and $L = S^0_{1,0}\partial_\theta \oplus S^{-1}_{1,0}\partial_{y_j} \oplus S^{-1}_{1,0}$

2. for $u, v \in C_0^\infty(\mathbb{R}^n)$: $\langle Au, v\rangle = \langle K, u \otimes v\rangle = \int L^k(e^{i\varphi})auv = \int e^{i\varphi}L^k(au)v$.

3. Therefore Au agrees with $\int e^{i\varphi}L^k(a(x,y,\theta)u(y))d\theta dy$ (this integral exists if k is large enough, and is smooth in x (again by increasing k).

□

(This can be generalized to when if $(y,\theta) \mapsto \varphi(x,y,\theta)$ is a phase function for all x).

Theorem 1.89 (Mapping Property of PDOs on $\mathcal{E}'$). $L^m_{\rho,\delta}(\mathbb{R}^n) \ni A : \mathcal{E}'(\mathbb{R}^n) \to \mathcal{D}'(\mathbb{R}^n)$ *(really they have a unique continuous extension with this property).*

Proof. 1. Define for $u \in \mathcal{E}'(\mathbb{R}^n)$, $v \in C_0^\infty(\mathbb{R}^n)$: $\langle Au, v\rangle = \langle u, A^t v\rangle$

2. but the distributional kernel of A^t is $K(y,x)$ [a], so get that $A^t : C_0^\infty(\mathbb{R}^n) \to C^\infty(R^n)$. Therefore $\langle u, A^t v\rangle$ is the pairing of $\mathcal{E}'$ and C^∞, which works.

□

By the Schwartz Kernel theorem, if $A \in L^m_{\rho,\delta}(X)$ with symbol a, then A has kernel $K_A = \int e^{i(x-y)\theta}a(x,y,\theta) \in \mathcal{D}'(X \times X)$. By integrating by parts in θ, we can get that $\operatorname{singsupp} K \subset \{x = y\} := \Delta(X \times X)$. This gives us that $\operatorname{singsupp} Au \subset \operatorname{singsupp} u$ for $u \in \mathcal{E}'(X)$

Definition 1.29 (Smoothing pseudo-differential operators). $A \in L^m_{\rho,\delta}(X)$ *is smoothing (denoted $L^{-\infty}$) if any of the following hold: (1) A is continuous $\mathcal{E}'(X) \to C^\infty(X)$ (2) $K_A \in C^\infty(X \times X)$ (3) $A = Op(a)$ with $a \in S^{-\infty}$*

To get better mapping properties (ie mapping on functions without compact support) it makes sense to consider compactly supported Kernels. This is a little too restrictive, instead we just require compact support of kernels on cylinders. For example if $\varphi \in C^\infty$, for $\int K_A(x,y)\varphi(x)$ to converge, we would like for all $y \in B_R(y_0)$, $\operatorname{supp} K_A \cap B_R(y_0) \times X$ to be compact. Similar thinking gives the following definition:

Definition 1.30 (Properly Supported pseudo-differential operators). $A \in L^m_{\rho,\delta}(X)$ *is properly supported if K_A is properly supported, that is for all compact K: $\Pi_y(K \times X \cap \operatorname{supp}(K_A))$ and $\Pi_x(X \times K \cap \operatorname{supp}(K_A))$ are compact*[b]

The picture should be that $\operatorname{supp} K_A$ is contained (locally) in a diagonal strip containing $\Delta(X \times X)$

If A is properly supported, it maps all the following spaces to themselves: $C_0^\infty, C^\infty, \mathcal{D}', \mathcal{E}'$.

By constructing $\chi(x,y) \in C^\infty(X \times X)$ which is identically 1 in a neighborhood of $\Delta(X \times X)$, then if $A \in L^m_{\rho,\delta}$, we can decompose K_A as $K_A\chi + K_A(1-\chi)$. The first (distribution) is properly supported, and the second (smooth function) is smoothing.

Theorem 1.90 (Dependence of Symbol of pseudo-differential operator on y). *Given $A \in L^m_{\rho,\delta}(X)$ (properly supported), there exists $\sigma(A) := b \in S^m_{\rho,\delta}(X)$, such that:*

$$Au(x) = \frac{1}{(2\pi)^n}\int e^{ix\theta}b(x,\theta)\widehat{u}(\theta)d\theta dx$$

[a] (probably should have complex conjugate?)

[b] where $\Pi_x : X \times X \ni (x,y) \mapsto x \in X$

for $u \in C_0^\infty(X)$.

b is called the complete symbol of A. If A has symbol a, then:

$$b(x,\xi) = e^{-ix\xi}A(e^{i(\cdot)\xi}) \sim \sum_{\alpha\in\mathbb{N}^n} \frac{1}{i^{|\alpha|}\alpha!}(\partial_\xi^\alpha \partial_y^\alpha a(x,y,\xi)|_{y=x})$$

Definition 1.31 (Asymptotic Sum of Symbols). *We write* $a \sim \sum_{j=0}^\infty a_j \in S^{m_0}$ *if* $a_j \in S^{m_j}_{\rho,\delta}$ *where* $m_j \searrow -\infty$ *and* $a - \sum_{j=0}^{k-1} a_j \in S^{m_k}_{\rho,\delta}$.

It turns out that given any sequence of a_j, there exists a unique a satisfying the above, modulo an element in $S^{-\infty}$. The idea is as follows:

1. let $\{\mu_{n,k}\}_{n\in\mathbb{N}}$ a collection of semi-norms for S^{m_k}, for each j, construct $b_j \in S^{-\infty}$ such that $\mu_{n,k}(a_j - b_j) < 2^{-j}$ for $n, k \le j-1$
   (a) This follows from the fact that every element of S^m can be approximated by elements of $S^{-\infty}$ in the topology of $S^{m+\varepsilon}$ for all $\varepsilon > 0$
      i. for $a \in S^m_{\rho,\delta}$, let $\chi_j(\theta) = \chi(\theta/j)$ where $\chi \in C_0^\infty$. Then $a \in S^{-\infty}$, then the claim is that $\chi a \in S^{-\infty}$ and converges to a pointwise and in the topology of $S^{m+\varepsilon}_{\rho,\delta}$ for all $\varepsilon > 0$
      ii. this follows from: if $a_j \in S^m$ converge pointwise everywhere to a, and a_j are bounded in S^m, then $a \in S^m$ and converges in $S^{m'}$ for all $m' > m$.
2. Then for each j, $\sum_{k\ge j} a_k - b_k$ converges in $S^{m_j}_{\rho,\delta}$, so $a = \sum_0^\infty a_j - b_j$ works

Here is a way to prove that a smooth function is an asymptotic sum of symbols:

Theorem 1.91 (Converse of Asymptotic Sum of Symbols). *If* $a_j \in S^{m_j}$, $m_j \to -\infty$, *and* $a \in C^\infty(X \times \mathbb{R}^N)$ *is such that*

1. *for* α, β, K, *there exist* $C, M > 0$ *such that* $\sup_{(x,\theta)\in K\times\mathbb{R}^N} |\partial_x^\alpha \partial_\theta^\beta a| \le C\langle\theta\rangle^M$
2. *there exists* $m_k \in \mathbb{N}$ *going to* $-\infty$, *such that for all* k, K:

$$|a - \sum_{j=0}^{k-1} a_j| \le C\langle\theta\rangle^{m_k}$$

   for $(x,\theta) \in K \times \mathbb{R}^N$

then $a \sim \sum a_j$

The proof of Theorem 1.90 can be broken down into steps. The main idea is to first show that b is a symbol with the above asymptotic expansion. Then to show that it is the correct symbol.

Step 1. Show that $b(x,\xi)$ decays rapidly away from $\theta = \xi$.

1. note A is properly supported, $e^{iy\xi} \in C^\infty$, therefore $b(x,\xi)$ is smooth. We can write (dropping constants):

$$b(x,\xi) = \iint e^{i(x-y)(\theta-\xi)} a(x,y,\theta)\, d\theta\, dy$$

2. we want rapid decay in ξ, so it makes sense to integrate by parts in y to get factors of $(\theta_j - \xi_j)^{-1}$, so it is natural to consider $L = -i|\theta - \xi|^{-1}\sum_{j=1}^{n}(\xi_j - \theta_j)\partial_{y_j}$. We need a cuttoff function around $\theta = \xi$. It would make sense to choose $(1 - \chi(|\xi - \theta|))$, BUT, we will end up needing $(1 - \chi(\frac{|\xi-\theta|}{|\xi|}))$

3. we split the integral in to the part near ξ and the part away from ξ, away from ξ, we integrate by parts k times to get the integrand:

$$|\xi - \theta|^{-k}(L^k a)(1 - \chi) \leq C|\xi - \theta|^{-k}(1 + |\theta|)^{m+\delta k}$$

4. now we use that if $|\xi| > 2$ and $\operatorname{supp}\chi \subset B_{1/2}(0)$, then $|\xi - \theta| \sim 1 + |\xi| + |\theta|$. Then multiply and divide by $(1 + |\theta|)^{n+1}$ to get the integral to converge as long as k is large enough. This tells us that this part of the integral is in $S^{-\infty}$

Step 2 Control the second integral using the method of stationary phase.

1. Our integral is $\int \chi(\frac{|\xi-\theta|}{|\xi|})e^{i(x-y)(\theta-\xi)}a(x,y,\theta)dyd\theta$.

2. change variables to have phase: $e^{-|\xi|is\sigma}$, this gives us: $\lambda^n \iint a(x, x+s, \lambda(\omega+\sigma))\chi(|\sigma|)e^{-i\lambda s\sigma}ds\sigma$ (where $|\xi| = \lambda$ and $\xi = \lambda\omega$).

3. the integral can be written $\int e^{(i/2)\langle Qy,y\rangle}\varphi(y)dy$ with $Q = \begin{pmatrix} 0 & -I \\ -I & 0 \end{pmatrix}$, $y = (s, \sigma)$.

4. by below computation, this becomes:

$$\sum_{|\alpha|\leq N-1} \frac{1}{i^{|\alpha|}\alpha!}\partial_y^\alpha\partial_\theta^\alpha a(x,y,\theta)|_{x=y} + R_N a$$

5. it can be shown that the remainder has the correct decay for the asymptotic sum to make sense.

Theorem 1.92 (Method of Stationary Phase). *Given Q a symmetric (non-degenerate) $n \times n$ matrix, $u \in C_0^\infty(\mathbb{R}^n)$, $\lambda \in \mathbb{R}_+$ then:*

$$\int e^{i\lambda\langle Qx,x\rangle}u(x)dx = \sum_{k=0}^{N-1}\frac{(2\pi)^{n/2}e^{\frac{i\pi sgn(Q)}{4}}}{k!|\det Q|^{1/2}\lambda^{k+n/2}}(\frac{1}{2i}\langle D_x, Q^{-1}D_x\rangle)^k u(0) + S_N(u,\lambda)$$

with $|S_n(u,\lambda)| \leq C(N!)^{-1}\lambda^{-N-n/2}\left\|(\frac{1}{2}\langle D, Q^{-1}D\rangle)^N u\right\|_{H^{n/2+\varepsilon}(\mathbb{R}^n)}$ for any $\varepsilon > 0$.

I will prove the simpler case where $Q = \begin{pmatrix} 0 & -I \\ -I & 0 \end{pmatrix}$

Proof.

1. Want to compute $\int e^{-i\lambda x\cdot y}u(x,y)dxdy$, by Plancheral, this is $c_d\int \widehat{e^{-i\lambda x\cdot y}}\widehat{u}(\xi,\eta)d\xi d\eta$

2. $\widehat{e^{-i\lambda x\cdot y}} = (\frac{2\pi}{\lambda})^d e^{i\xi\cdot\eta/\lambda}$

(a) $\widehat{e^{-i\lambda x\cdot y}} = \prod_{j=1}^{d} \widehat{e^{-i\lambda x_j y_j}}$

(b) $\widehat{e^{-i\lambda xy}} = \int e^{-i(x\xi+y\eta)-ixy}dxdy$. Let $x = u+v$, $y = u-v$.

(c) Use fact that $\widehat{e^{ix^2\lambda/2}} = (2\pi/\lambda)^{d/2}e^{-i\pi/4}e^{-i\xi^2/(2\lambda)}$

3. we get $(\frac{2\pi}{\lambda})^d \int e^{i\xi\cdot\eta/\lambda}\widehat{u}(\xi,\eta)d\xi d\eta$, expand exponential, use fact that $\int \frac{i^k(\xi\cdot\eta)^k}{\lambda^k k!}\widehat{u} = \frac{1}{i^k\lambda^k k!}(\partial_x\cdot\partial_y)^k u(0,0)$, then multinomial expansion: $(\partial_x\cdot\partial_y)^k = \sum_{|\alpha|=k}\binom{k}{\alpha}\partial_x^\alpha\partial_y^\alpha$

4. so we get $(\frac{2\pi}{\lambda})^d \sum_{|\alpha|\le N} \frac{1}{i^{|\alpha|}\lambda^{|\alpha|}\alpha!}\partial_x^\alpha\partial_y^\alpha u(0,0) + R_N u$

Control on remainder

1. using $|e^{it} - \sum_0^{n-1}\frac{(it)^k}{k!}| \le \frac{t^n}{n!}$, we get that the magnitude of $R_N u$ is bounded by:

$$(\frac{2\pi}{\lambda})^d \frac{1}{\lambda^{N+1}(N+1)!}\int (\xi\cdot\eta)^{N+1}\widehat{u}d\xi d\eta$$

2. the integrand (modulo constants) is $\mathcal{F}[\partial_x\cdot\partial_y)^{N+1}u]$. Sufficient conditions for the $\widehat{f}\in L^1(\mathbb{R}^d)$ is $\partial_x^\alpha f\in L^1$ for all $|\alpha|\le d+1$

3. Therefore the integrand is bounded by:

$$\frac{(2\pi)^d}{\lambda^{d+N+1}(N+1)!}\sum_{|\alpha+\beta|\le 2d+1}\left\|\partial_x^\alpha\partial_y^\beta(\partial_x\cdot\partial_y)^N u\right\|_{L^1}$$

□

Theorem 1.93 (Product of Pseudodifferential Operators). *If $A\in L^{m_1}_{\rho,\delta}, B\in L^{m_2}_{\rho,\delta}$ (at least one properly supported), then $AB\in L^{m_1+m_2}_{\rho,\delta}$ with full symbol:*

$$\sigma_{AB}(x,\xi) \sim \sum_{\alpha\in\mathbb{N}^n}\frac{\partial_\xi^\alpha\sigma_A(x,\xi)D_x^\alpha\sigma_A(x,\xi)}{\alpha!}$$

Definition 1.32 ($a\#b$). *If $a\in S^{m_1}_{\rho,\delta}, b\in S^{m_2}_{\rho,\delta}$ define $a\#b := \sum_\alpha(\alpha!)^{-1}\partial_\xi^\alpha a D_x^\alpha b$*

1.8.3 Elliptic Operators and L^2 Continuity

Definition 1.33 (Elliptic PDO). *$A\in L^m_{\rho,\delta}$ with symbol a is elliptic at $(x_0,\xi_0)\in X\times\dot{\mathbb{R}}^N$, if $a(x,\xi)\ge C\langle\xi\rangle^m$ for x,ξ in a conical neighborhood of x_0,ξ_0. That is for the set:*

$$\left\{(x,\xi) : |\xi| > c_1, \left|\frac{\xi}{|\xi|}-\frac{\xi_0}{|\xi_0|}\right| < c_2, |x-x_0| < c_3\right\}$$

Furthermore, we say:

1. *A is elliptic at x_0 if A is elliptic for all $(x_0,\xi_0)\in X\times\dot{\mathbb{R}}^N$*

2. *A is elliptic on $Y\subset X$ if A is elliptic for all $x\in A$*

3. *A is elliptic of A is elliptic on X.*

Definition 1.34 (Classical PDO Operator). *$a \in S^m_{1,0}$ is called a classical symbol (denote S^m_{cl}) if $a \sim \sum_{j=0}^{\infty}(1-\chi(\theta))a_j$, where each $a_j \in S^{m-j}$ is positively homogeneous in θ of degree $m-j$, and $\chi(\theta) = 1$ near 0*[a] *.*

If $A \in L^m$ with $\sigma(A) \in S^m_{cl}$, we call A a classical pseudo-differential operators, which is denoted L^m_{cl}

Since the values of a for fixed x are uniquely determined by θ such that $|\theta| = 1$. We see that if $a(x_0, \xi_0) \neq 0$, then a is elliptic at x_0, ξ_0.

Theorem 1.94 (Parametrix Construction). *If $P \in L^m_{\rho,\delta}$ is properly supported and elliptic (and $\rho > \delta$), then there exists a unique (modulo smoothing operators) $Q \in L^{-m}_{\rho,\delta}$ (properly supported) such that $PQ = QP = I$ mod $L^{-\infty}$.*

Definition 1.35 (Parametrix). *The above Q is called that **parametrix** of P.*

Proof.

1. Using a partition of unity, can construct $Q_0 \in C^\infty(X, \mathbb{R}^n)$ so that for each compact set K $Q_0(x,\xi) = P(x,\xi)^{-1}$ for $x \in K$, $|\xi| > C_K$

2. Claim: $Q_0 \in S^{-m}_{\rho,\delta}$

   (a) use induction and the identity $Q_0 P = 1$.

3. $Q_0 \# P = 1 + \sum_{|\alpha|\geq 1}(\alpha!)^{-1}\partial_\xi^\alpha Q_0 D_x^\alpha P$, the second term is in $S^{-\rho+\delta}$, call it $-T$, similarly, define $R = 1 - P\#Q_0 \in S^{-\rho+\delta}$ (everything modulo $L^{-\infty}$)

4. Now define $Q_r := Q_0\#(1 + R + R\#R + R\#R\#R + \cdots) \in S^{-m}_{\rho,\delta}$ then $P\#Q_r = I$. Similarly, if $Q_l := (I + T + T\#T + \cdots)\#Q_0$, then $Q_l \# P = I$

5. $Q_l = Q_l\#I = Q_l\#(P\#Q_r) = Q_r$. Let $Q = Q_r$ modulo $S^{-\infty}$

□

Theorem 1.95 (Adjoint of PDO). *If $A \in L^m_{\rho,\delta}$, then the adjoint $A^* : C_0^\infty \to \mathcal{D}'$ belongs to $L^m_{\rho,\delta}$, has Schwartz kernel: $\overline{K_A(y,x)}$ and has symbol $\sigma_{A^*} \sim \sum(\alpha!)^{-1}\partial_\xi^\alpha D_x^\alpha \overline{\sigma_A(x,\xi)}$*

Theorem 1.96 (L^2 Boundedness of PDOs). *If $A \in L^0_{\rho,\delta}$ and $K_A \in \mathcal{E}'(\mathbb{R}^n \times \mathbb{R}^n)$, then A is bounded $L^2 \to L^2$.*

Proof.

1. Write $A^*A = MI - B^*B - K$ with $M \gg 0$, $B \in L^0_{\rho,\delta}$, $K \in L^{-\infty}$

   (a) For large M, the principal symbol of $C := M - A^*A$ is $c_0 := M - |a_0|^2 > 0$ (where a_0 is the principal symbol of A). c_0 is uniformly bounded above and below.

[a] the cuttoff allows us to the expression to be smooth in θ for all θ (otherwise we might have terms like $1/\theta$). Again it is only the behavior near ∞ that is important in this business.

(b) Let $b_0 = c_0^{1/2} \in S^0_{\rho,\delta}$, let $B_0 = Op(b_0)$, then by looking at principal symbols, $C_1 := M - A^*A - B_0^*B_0 \in L^{-(\rho-\delta)}_{\rho,\delta}$

(c) goal: find $B_1 \in L^{-(\rho-\delta)}_{\rho,\delta}$ such that $M - A^*A - (B_0+B_1)^*(B_0+B_1) \in L^{-2(\rho-\delta)}_{\rho,\delta}$

i. this is $C_1 - B_1^*B_0 + B_0B_1^* + B_1^*B_1^*$, so want $C_1 - B_1^*B_0 \in L^{-2(\rho-\delta)}_{\rho,\delta}$

ii. let B_1 have symbol $\frac{1}{2}C_1B_0^{-1}$

2. if $u \in C_0^\infty$, then $\|Au\|_{L^2}^2 = \langle A^*Au, u\rangle \le M\|u\|_2^2 - \|Bu\|_2^2 + \|K\|_{L^2\to L^2}\|u\|_2^2 \le C\|u\|_2^2$

(a) Boundedness of K on L^2 comes from Shur's lemma.

□

The above proof is really designed to establish a precise bound, which I didn't do. But this proof can also be modified to prove the same thing if we know the seminorms of the symbol are bounded uniformly in x. There is a much easier proof of this fact for $S^0_{1,0}$ with symbol compactly supported in x:

Proof. 1. for $u \in C_0^\infty$, we have $\widehat{Au}(\theta) = \int \widehat{u}(\xi)\int a(x,\xi)e^{-ix(\theta-\xi)}dxd\xi$

2. this is just Fourier transforming the x component of a: $\int \widehat{u}(\xi)\widehat{a}(\theta-\xi,\xi)d\xi$, it now suffices to show that $\widehat{a}(\theta-\xi,\xi) \in L^2$

3. note since $\partial_x^\alpha a \in L^\infty \cap L^1$ (compact support), so that $\langle\eta\rangle^m\widehat{a}(\eta,\xi) \in L^\infty$ for all $m \in \mathbb{Z}$.

□

Theorem 1.97 (PDO mapping on Sobolev Spaces). *If $A \in L^m_{\rho,\delta}$ is properly supported, then $A : H^s_{loc} \to H^{s-m}_{loc}$ is continuous.*

There are several variants: if the semi-norm bounds are uniform in x, then we have $H^s \to H^{s-m}$. This implies the same thing if A's Schwartz kernel is compactly supported.

Proof. (idea, in reality we have to use a bunch of cuttoffs)

1. $\langle D\rangle^l \in L^l_{1,0}$ and maps $H^s \to H^{s-l}$

2. $B := \langle D\rangle^{s-m} A \langle D\rangle^{-s} \in L^0_{\rho,\delta}$ is bounded $L^2 \to L^2$ by previous theorem.

3. $A = \langle D\rangle^{m-s} B \langle D\rangle^s : H^s \to L^2 \to L^2 \to H^{m-s}$

□

1.8.4 Change of Coordinates

We can change coordinates and have our symbol stay in the same class, which is useful for defining pseudo-differential operators on manifolds.

Theorem 1.98 (Change of Coordinates for PDOs). *If $A_2 \in L^m_{\rho,\delta}(\mathbb{R}^n_2)$ with $\rho > \delta$ and $\rho + \delta = 1$ and $\kappa : \mathbb{R}^n_1 \to \mathbb{R}^n_2$ is a diffeomorphism*[a]*, then we can define $A = \kappa^* \circ A \circ (\kappa^{-1})^* \in L^m(\mathbb{R}^n_1)_{\rho,\delta}$ with symbol:*

$$a(x,y,\theta) = \tilde{a}(\kappa(x), \kappa(y), G^{-1}(x,y)\theta)\frac{|\det \kappa'|}{|\det G|}$$

where $G(x,y) = F(x,y)^t$, and $F(x,y) = \int_0^1 (\partial_x \kappa)(tx + (1-t)y)dy$. From this, we can see that the principal symbol of A is just:

$$\sigma^0_A(x,\xi) = \sigma^0_{\tilde{A}}(\kappa(x), (((k')^t)^{-1})\xi)$$

Proof.

1. if $u \in C_0^\infty(\mathbb{R}^n_1)$, then $Au(x) = c_d \int e^{i(\kappa(x)-\tilde{y})\tilde{\theta}} a(\kappa(x), \tilde{y}, \tilde{\theta}) u(\kappa^{-1}(\tilde{y}))d\tilde{y}d\tilde{\theta}$
2. let $y = \kappa^{-1}(\tilde{y})$, get a Jacobian, now the only issue is the phase, which is now $(\kappa(x) - \kappa(y))\tilde{\theta}$
3. idea: Taylor expand κ to write $(\kappa(x) - \kappa(y) \cdot \tilde{\theta} = \left\langle F(x,y)(x-y), \tilde{\theta} \right\rangle = \langle (x-y), F^t(x-y) \rangle$. Then change variables of θ. This requires shrinking to a neighborhood of the diagonal so that F^t is invertible.

□

1.9 Important Tricks For Exercises

1.9.1 Things Involving Japanese Brackets

Proposition 1.9 (Integrating on one component). *If $\eta \in \mathbb{R}^n$, $\xi \in \mathbb{R}$, $a > 1$, then $\int \langle \xi, \eta \rangle^{-a} d\xi = C \langle \eta \rangle^{-a+1}$*

Proof. $\langle \xi, \eta \rangle^{-2s} = (\xi^2 + \langle \eta \rangle)^{-s}$, let $c = \langle \eta \rangle$, then we compute:

$$\int_{-\infty}^{\infty} (x^2 + c^2)^{-s} dx = c^{-2s} 2 \int_0^\infty \left(\frac{x^2}{c^2} + 1\right)^{-s} dx$$

let $u = \frac{x^2}{c^2} + 1$, so $\frac{du}{dx} = \frac{2x}{c^2}$, so we get:

$$c^{-2s+2} \int_1^\infty u^{-s}(c^2(u-1))^{-1/2} du = c^{-2s+1} \int_1^\infty \frac{1}{\sqrt{u-1}u^s} du$$

This is fine near 1, far away, the integral is asymptotically $u^{-s-1/2}$, which has finite integral if $s + 1/2 > 1$. Therefore $\int_{-\infty}^\infty \langle \xi, \eta \rangle^{-2s} d\xi = C \langle \eta \rangle^{-2s-1}$ □

[a] this works, and is more meaningul, if the domain and range are arbitrary open sets

This is the crucial ingredient in the following exercise:

Example 1.6. *Let* $Tu(x,y) = u(0,y)$ *for* $x \in \mathbb{R}$, $y \in \mathbb{R}^{n-1}$, *then* $T : H^s(\mathbb{R}^n) \to H^{s-1/2}(\mathbb{R}^{n-1})$

Proposition 1.10 (**Triangle Like Inequality**). *For* $s \geq 0$, *there exists a constant* C *such that* $\langle x \rangle^s \leq C(\langle x-y \rangle^s + \langle y \rangle^s)$

2 Harmonic Analysis

2.1 Fourier Inversion, Plancherels's Theorem, and Other Basics

Reference: Christ 1.1-1.8

Definitions of Fourier transforms, L^p facts, convolution, approximate identities, Plancherel's theorem, tempered distributions, Poisson summation formula

2.1.1 Fourier Series

Let $\mathbb{T}^d = [0, 2\pi]^d$.

Definition 2.1 (Fourier Series). *For $f \in L^1(\mathbb{T}^d)$, $\widehat{f}(n) = (2\pi)^{-d} \int_{\mathbb{T}^d} f(x)e^{-in\cdot x}$ for $n \in \mathbb{N}$.*

This is funny convention for our class, here is a table

Space	Fourier transform	Inverse Fourier Transform	Plancheral Constant
$\mathbb{R}^d$	$\widehat{f}(\xi) := \int_{\mathbb{R}^d} e^{-ix\cdot\xi} f(x)dx$	$f(x) = (2\pi)^{-d} \int_{\mathbb{R}^d} e^{ix\xi} \widehat{f}(\xi)d\xi$	$\\|f\\|_{L^2}^2 = (2\pi)^{-d} \\|\widehat{f}\\|_2^2$
$\mathbb{T}^d$	$\widehat{f}(n) := (2\pi)^{-d} \int_{\mathbb{T}^d} e^{-in\cdot x} f(x)dx$	$f(x) = \sum_{n\in\mathbb{Z}^d} \widehat{f}(n)e^{in\cdot x}$	$\\|\widehat{f}\\|_{\ell^2}^2 = (2\pi)^{-d} \\|f\\|_{L^2}^2$

The only thing to remember is in $\mathbb{R}$, $\mathcal{F}[e^{-x^2/2}] = \sqrt{2\pi}e^{-\xi^2/2}$, so $\widehat{e^{-x^2/2}} = (2\pi)^{d/2}e^{-\xi^2/2}$. Therefore the constant for the inverse Fourier transform must have $(2\pi)^{-d}$. Since $\|e^{-x^2/2}\|_{L^2(\mathbb{R})} = \sqrt{\pi}$, we can quickly recover the Plancheral constant. For Fourier series, just let $f(x) = e^{inx}$, then constants are easy to recover.

Theorem 2.1 (Plancherel's Theorem). *For $f \in L^2(\mathbb{T}^d)$, $\|\widehat{f}\|_{\ell^2}^2 = (2\pi)^{-d}\|f\|_{L^2}$ and $f = \sum_{n\in\mathbb{Z}^d} \widehat{f}(n)e^{in\cdot x}$ (in the sense that $\sum_n \widehat{f}e^{in\cdot x}$ converges in L^2 norm to f)*

Proof. By computation $e_n := (2\pi)^{-d/2}e^{in\cdot x}$ is an orthonormal set. By the Stone-Weierstrauss theorem, it's span is dense in $C^0(\mathbb{T}^d)$, and by measure theory results, its span is dense in L^2. Then by results about orthonormal dense sets in Hilbert spaces, these results follow. □

2.1.2 Fourier Transform

Definition 2.2 (Fourier Transform on L^1). *For $f \in L^1(\mathbb{R}^d)$, define $\widehat{f}(\xi) = \int e^{-ix\cdot\xi} f(x)dx$*

It is easy to see that $\widehat{\ }: L^1 \to L^\infty \cap C^0$. The inverse Fourier transform is $\mathcal{F}^{-1}g(\xi) = (2\pi)^{-n} \int g(\xi)e^{ix\cdot\xi}d\xi$.

Another way to remember all these constants is to remember the unitary Fourier transform $\mathcal{F}_u[f](\xi) := \int e^{-2\pi ix\cdot\xi} f(\xi) = \widehat{f}(2\pi\xi)$. In which case all constants are 1.

Lemma 2.1 (Planchärel's Theorem for $L^1 \cap L^2$). *For $f \in L^1 \cap L^2$, $\|f\|_2^2 = (2\pi)^{-d} \|\widehat{f}\|_2^2$*

Proof.

1. prove for dense subspace V.

2. for $f \in V$, let $f_t = f(tx)$ (will send $t \to \infty$)

3. $\|f\|_2^2 = t^d \|f_t\|_2^2$

4. if t large enough , use Plancheral for Torus: $\|f_t\|_2^2 = c_d \sum_n |\mathcal{F}_{\mathbb{T}} f_t(n)|^2$

5. $\mathcal{F}_{\mathbb{T}} f_t(n) = (2\pi)^{-d} \int e^{-in\cdot x} f(tx)dx = (2\pi t)^{-d} \mathcal{F}_{\mathbb{R}} f(n/t)$

6. therefore $\|f\|_2^2 = c_d t^{-d} \sum_n |\mathcal{F}_{\mathbb{R}} f(n/t)|^2$

7. the term on the right is the integral of a simple function approximating $|\widehat{f}(\xi)|^2$, and we can apply the dominated convergence theorem if $\widehat{f}(\xi) \in L^2$. If we assume $\widehat{f} \lesssim \langle \xi \rangle^{-d}$, then we are done.

8. the dense subspace is $C_0^\infty(\mathbb{R}^d)$ (this seems like overkill, but oh well).

□

the idea is to use a scaling argument to reduce to the case of Plancheral on the torus. This immediately requires compact support. Following our nose, we get a Riemman sum, which may not converge, we can force it to as long as it is dominated by something decaying sufficiently.

Theorem 2.2 (Density of $C_0^\infty(\mathbb{R}^d)$ in L^p). *$C_0^\infty(\mathbb{R}^d)$ is dense in L^p.*

Proof. 1. For $f \in L^p$, by a dyadic argument, get a sequence of simple functions that pointwise increase to f (wlog $f \geq 0$).

2. By the dominated convergence theorem, they converge in L^p to f.

3. Next approximate the characteristic functions by C_0^0 functions. This is by regularity of the measure (get a compact set below and open set above), then use Urysohn's lemma.

4. Next approximate these continuous compactly supported functions via mollification by smooth compactly supported functions, apply Young's convolution inequality to get final result.

□

We therefore have a continuous linear operator on a dense subspace of L^2, so we extend it to L^2 (this is a general fact about densely defined bounded operators on complete spaces).

Theorem 2.3 (Approximation of Identity). *An approximate identity sequence is $\varphi_n \in L^1$ such that (1) $\int \varphi_n = 1$ for all n (2) $\|\varphi_n\|_1 \leq C$ (3) $\int_{|x|>\varepsilon} \varphi_n(x) \to 0$ as $n \to \infty$ for all $\varepsilon > 0$.*

Then (1) if $f \in C_0^0$, then $\varphi_n * f$ converges uniformly to f (2) if $f \in C_b^0$, then $\varphi_n * f$ converges uniformly on compact sets to f (3) if $f \in L^p$ ($p \in [1, \infty)$) then $\varphi_n * f$ converges to f in L^p

This greatly generalizes the usual mollification family.

Theorem 2.4 (Plancherel's Theorem for L^2). *There exists a surjective bounded linear map $\mathcal{F} : L^2 \to L^2$ (which has the explicit formula above for L^1) such that $\|f\|_2 = (2\pi)^{-d/2} \|\mathcal{F} f\|_2$ and*

$$\left\| f - \frac{1}{(2\pi)^d} \int_{|\xi|\leq R} e^{ix\cdot\xi} \mathcal{F} f d\xi \right\|_{L^2} \xrightarrow{R\to\infty} 0$$

Proof. This operator comes from the fact that it densely defined, so it has a unique extension. To show it is onto, it suffice to show that if $\langle \mathcal{F} f, g \rangle = 0$ for all f, then $g = 0$, use adjoint of the Fourier transform to get this. To show last thing, compute this norm explicitly by writing $\|f - g\|_2^2 = \|f\|_2^2 + \|g\|_2^2 - 2\langle f, g \rangle$. □

2.1.3 Convolution

Theorem 2.5 (**L^p convolution bounds**). *If $f \in L^1$, $g \in L^p$, then $\|f * g\|_{L^p} \le \|f\|_1 \|g\|_p$*

Proof. Use duality: $\|f * g\|_p \le \langle f * g, h \rangle$ for $h \in L^{p'}$. Expand the integral, use Holder's inequality. □

This is just a special case of:

Theorem 2.6 (**Young's Convolution Identity**). *If $f \in L^p$ $g \in L^q$ and $\frac{1}{p} + \frac{1}{q} = 1 + \frac{1}{r}$, then $\|f * g\|_r \le \|f\|_p \|g\|_q$*

Theorem 2.7 (**Fubini-Toneli**). *If $f(x, y)$ is measurable with respect to $\sigma-$finite measure spaces, then taking absolute values we can integrate in any order. If any are finite, they all are, and we can drop the absolute values.*

I always forget that in absolute value, everything exists.

Definition 2.3 (**Radon Measure**). *A complex Radon measure is such that $|\mu|$ is (1) finite on compact sets (2) outer regular on Borel sets and (3) inner regular on open sets.*

Definition 2.4 (**Convolution with Radon Measure**). *Let μ be a complex radon measure and $f \in C_b^0$, define $\mu * f(x) = \int f(x - y) d\mu(y)$*

Theorem 2.8 (**Characterization of Convolution Operators With Radon Measures**). *If T is a bounded, linear map $C_{\to 0}^0(\mathbb{R}^n) \to C_b^0(\mathbb{R}^n)$ that is invariant under translations, then $Tf = f * \mu$ for some complex Radon measure.*

The proof relies on the Riesz representation theorem.

Theorem 2.9 (**Riesz Representation Theorems**).

1. (***Hilbert version***) *For Hilbert space H, for each $u \in H^*$, there exists a unique $f \in H$ such that for all $\varphi \in H$, $u(\varphi) = \langle u, f \rangle$ (and $\|f\|_H = \|u\|_{H^*}$).*

2. (C_0^0 ***dual on LCH space***) *If $\lambda \in (C_0^0(X))^*$ is positive for X a locally compact Hausdorff space, then there exists a unique regular Borel measure μ on X such that for all $\varphi \in C_0^(X)$, then $\lambda(f) = \int f d\mu(x)$.*

3. ($C_{\to 0}^0$ ***dual on LCH***) *If $\lambda \in (C_{\to 0}^0(X))^*$ is continuous, then there exists a unique complex countably additive Borel measure μ such that $\lambda(f) = \int f d\mu$ for all $f \in C_{\to 0}^0(X)$*

Convolution will add additional regularity.

Proposition 2.1 (**Covolution adds regularity**). *If $f \in C^1$ and $g \in L^1$, and $f, \nabla f \in L^\infty$, then $f * g \in C^1$, this can be iterated.*

If $f \in L^1$ and $g \in L^\infty$, then $f * g \in C_b^0$ (this uses that translation is continuous on L^1)

Theorem 2.10 (Basic Properties of Fourier Transform). *These are all obvious, the only one worth remembering and not deriving every time:*

1. $L \in GL(n)$, $f \in L^1(\mathbb{R}^n)$, *then* $\mathcal{F}[f \circ L] = (\det L)^{-1}\mathcal{F}[f]((L^*)^{-1} \circ \xi)$

Theorem 2.11 (Fourier Transform of Gaussian). *For* $f_z(x) = e^{-\frac{z|x|^2}{2}}$ *(*$\Re(z) > 0$*), then* $\mathcal{F}_x[f_z(x)] = (2\pi)^{d/2} z^{-d/2} e^{-\frac{|\xi|^2}{2z}}$

Proof.

∎

To remember, the constant can be recovered quickly by computing $\widehat{f}(0)$. The reason z goes to the denominator is that as z increases (assuming it is real), f_z becomes more localized → less regular → slower decay in Fourier space → z in denominator.

2.1.4 Tempered Distributions

Definition 2.5 (Schwartz Function). *f is a Schwartz function ($\mathcal{S}(\mathbb{R}^n)$) if f is smooth, complex valued and for all $\alpha, \beta \in \mathbb{N}^n$, $\langle x \rangle^\alpha \partial^\beta f$ is bounded.*

Theorem 2.12 (Fourier Transform on Schwartz Functions). *The Fourier transform is a bijective homeomorphism on Schwartz functions, mapping them to themselves.*

Definition 2.6 (Tempered Distributions). *A tempered distribution is a complex valued, continuous, linear map on $\mathcal{S}(\mathbb{R}^n)$.*

Note that this is strictly contained in distributions. Tempered distributions are 'nicer' than distributions.

The Fourier transform maps tempered distributions to themselves and it is a homeomorphism. Where $\langle \mathcal{F}u, \varphi \rangle \equiv \langle u, \mathcal{F}\varphi \rangle$ for $u \in \mathcal{S}'$ and $\varphi \in \mathcal{S}$.

2.1.5 Poisson Summation Formula

Theorem 2.13 (Poisson Summation Formula). *For* $f \in \mathcal{S}(\mathbb{R})$*:*

$$\sum_{k\in\mathbb{Z}^n} \widehat{f}(k) = \sum_{k\in\mathbb{Z}^n} f(k)$$

(where $\widehat{f}(\xi) = \int e^{-2\pi i x\cdot\xi} f(x)dx$*)*

As a consequence (not rigorous), letting $f(y) = \delta(x-y)$, we have:

$$\sum_{n\in\mathbb{Z}^n} e^{2\pi i n\cdot x} = \sum_{k\in\mathbb{Z}^n} \delta(x-k)$$

Proof. Fix $f \in \mathcal{S}$, define $g : [0,1]^d \to \mathbb{C}$ by $g(x) = \sum_{n\in\mathbb{Z}^d} f(x+n)$ (this is a periodic function. By the inversion formula: $g(x) = \sum_{k\in\mathbb{Z}^d} e^{2\pi i x\cdot k} \int_{[0,1]^d} g(y)e^{-2\pi i y k}dy$. $g(0) = \sum_{n\in\mathbb{Z}^d} f(n)$. And:

$$g(0) = \sum_{k\in\mathbb{Z}^d} \int_{[0,1]^d} \sum_{n\in\mathbb{Z}^d} f(y+n)e^{-2\pi i y\cdot k}dy = \sum_{k\in\mathbb{Z}^d} \int_{\mathbb{R}^d} f(y)e^{-2\pi i y\cdot k}dy = \sum_{k\in\mathbb{Z}^d} \widehat{f}(k)$$

(second to last equality requires justification) □

trick: define periodic function from f (not scaled but repeating). Evaluate periodic function at zero, use fourier inversion, and clever tiling to get integration over entire space.

Using the usual Fourier transform, Poisson summation becomes: $\sum_{n\in\mathbb{Z}^d} f(2\pi n) = (2\pi)^{-d} \sum_{n\in\mathbb{Z}^d} \widehat{f}(n)$.

The Poisson summation formula can also be written functionally (back to unitary ft) as $\sum_{n\in\mathbb{Z}^d} f(n+x) = \sum_{n\in\mathbb{Z}^d} e^{2\pi i n x}\widehat{f}(n)$. In our usual fourier transform, the Poisson summation is $\sum_{n\in\mathbb{Z}^d} f(x+2\pi n) = (2\pi)^{-d} \sum_{k\in\mathbb{R}^d} e^{ikx}\widehat{f}(k)$

2.1.6 List of Useful Fourier Transforms

Here are useful, or nontrivial, fourier transforms:

1. Gaussian: $e^{-x^2 t/2}$ has FT $(\frac{2\pi}{t})^{d/2} e^{-\xi^2/(2t)}$
2. Poisson Kernel: $e^{-t|\xi|}$ has IFT $\pi^{\frac{-d+1}{2}}\Gamma(\frac{d+1}{2})(t^2+|x|^2)^{\frac{-d+1}{2}}$ (where $\xi \in \mathbb{R}^d$)
   (a) this is used to solve $(\partial_t^2 + \nabla)u(t,x) = 0$ in $\mathbb{R}^{d+1}$ and $u(0,x) = f$.
3. $1 \mapsto (2\pi)^d \delta(\xi)$
4. $pv(1/x) \mapsto -i\pi\,\mathrm{sgn}(\xi)$ $(d=1)$
5. $\log|x| \mapsto \frac{-\pi}{|\xi|} - 2\pi\gamma\delta(\nu)$ (care is need to understand $1/|\xi|$ as a distribution)

2.2 Convergence of Fourier Series

Reference: Christ 3.1-3.5, 3.7-3.8

Decay of Fourier coefficients, Rademacher functions, Khinchine's inequlaity, uniform and pointwise convergence of Fourier series, almost everywhere divergence (Kolmogorov theorem), L^p norm convergence, almost everywhere convergence, Wiener's Tauberian theorem, Riesz-Thorin Theorem

Theorem 2.14 (Summary of Basic Fourier Mapping Properties).

L^1	$\to$	$L^\infty \cap C^0_{\to 0}$
L^2	$\to$	L^2
$\mathcal{S}$	$\to$	$\mathcal{S}$
$\mathcal{S}'$	$\to$	$\mathcal{S}'$
$(Lip)_{comp}$	$\to$	$\langle\xi\rangle^{-1} L^2 \cap \langle\xi\rangle^{-1} L^\infty$
$(\Lambda_\alpha)_{comp}$	$\to$	$\langle\xi\rangle^{-\alpha} L^\infty$
C^k_0	$\to$	$\langle\xi\rangle^{-k} (L^2 \cap L^\infty)$
C^k	$\leftarrow$	$\langle\xi\rangle^{n+1+k} L^\infty$
L^p $(p \in [1,2])$	$\to$	$L^{p'}$
$\Lambda_\alpha(\mathbb{T})$ $(\alpha > 1/2)$	$\to$	$\ell^1(\mathbb{Z})$

It is hard to keep everything together, but here is yet another attempt at a summary:

- regularity in base space leads to decay in Fourier space, but the details are funny.
- the preserved spaces are $\mathcal{S}$, $\mathcal{S}'$, L^2 (once you uniquely extend the Fourier transform)
- For $p \in [1,2]$, L^p is mapped boundedly (Hausdorff-Young) to $L^{p'}$ (but it is not surjective).
- For $p \in (2,\infty)$, the Fourier transform may not even live in $L^{p'}$
- L^1 not only goes to L^∞, but also C^0 and decays (Riemann-Lebesgue). But it can decay arbitrarily slowly ☹ and is not surjective onto this smaller function space.
- On the torus, if $p > 2$, then $L^p(\mathbb{T}^d) \subset L^2(\mathbb{T}^d) \subset L^1(\mathbb{T}^d)$. So, by the above, $L^p(\mathbb{T}^d)$ will be mapped boundedly to $\ell^\infty(\mathbb{Z}^d) \cap \ell^2(\mathbb{Z}^d) = \ell^2(\mathbb{Z}^d)$. If p increases, we can still only say that L^p functions are mapped to ℓ^2 sequences (Kahane).

2.2.1 Decay of Fourier Coefficients

Theorem 2.15 (Fourier Decay for C^k_0). *If $f \in C^k_0(\mathbb{R}^n)$, then $\langle\xi\rangle^k \widehat{f} \in L^2 \cap L^\infty$. This implies $\widehat{f} = \mathcal{O}(|\xi|^{-k})$*

Proof. $|\xi|^{2k}|\widehat{f}(\xi)|^2 \le \sum_{|\alpha|=k} |\xi^\alpha|^2 |\widehat{f}(\xi)|^2$. Each term on the right is $|(i\xi)^\alpha \widehat{f}(\xi)|^2 = |\mathcal{F}[\partial^\alpha f](\xi)|^2$. $\partial^\alpha f \in C_0 \subset L^1 \cap L^2$, therefore the Fourier transform is in $L^\infty \cap L^2$, so integrate both sides to get the result. □

Theorem 2.16 **(Regularity of Fourier Decaying Function).** *If* $\widehat{f} = \mathcal{O}(|\xi|^{-k-d-1})$, *then* $f \in C^k$. *A stronger statement is that* $\langle\xi\rangle^k \widehat{f} \in L^1$ *implies* $f \in C^k$

Proof. For each $|\alpha| \le k$, $(i\xi)^\alpha \widehat{f} \in L^1$, therefore $L^\infty \ni \mathcal{F}^{-1}[(i\xi)^k \widehat{f}] = \partial^\alpha f$. Note that $\partial^\alpha f$ always exist as weak derivatives, but then we see that they are in fact bounded functions. □

Definition 2.7 **(Lipschitz Continuous).** f *is Lipschitz if* $\sup_{x \neq y} |f(x) - f(y)||x-y|^{-1} < \infty$

Note that this is a uniform bound. Also Lipschitz functions are L^1_{loc}, so they are tempered distributions

Theorem 2.17 **(Fourier Decay for Lipschitz Functions).** *If* f *is Lipschitz with compact support, then* $\widehat{f} = \mathcal{O}(|\xi|^{-1})$ *and* $\langle\xi\rangle f \in L^2$

So $Lip \subset H^1$

Proof. (Proof of the second statement) Assume $d = 1$ Idea: f is basically C^1, so $\mathcal{F}[\partial_x f] = i\xi\widehat{f} \in L^2$. But derivative doesn't exist, so need to consider distributional derivative.

1. For $\varphi \in \mathcal{S}$, $\int f\varphi' = \lim_{h\to 0} \int f \frac{\varphi(x+h)-\varphi(x)}{h} dx = \lim_{h\to 0} \int \frac{f(x)-f(x-h)}{h}\varphi(x)dx$
2. let $f_h = \frac{f(x)-f(x-h)}{h}$, $\|f_h\|_\infty \le \|f\|_{Lip}$, by the Banach-Alagou theorem, get subsequence to converge weakly to $g \in (L^1)^* = L^\infty$. Since g has compact support, $g \in L^2 \cap L^1 \cap L^\infty$
3. therefore $f' = g$ (in distributional sense). Let $\eta \in C_0^\infty$ be 1 on the support of f, then $\int g\eta e^{-ix\xi}dx = \int f(x)(i\xi)e^{-ix\xi}dx = i\xi\widehat{f}$
4. since $g \in L^2$, $\widehat{g} \in L^2$, so $i\xi\widehat{f} \in L^2$

□

This is the edge case of Banach-Alaglou, bounded in L^∞ allows weak subsequence. But bounded in L^1 probably doesn't? Also note that we proved that every Lipshitz continuous function has a weak derivative in L^1

Theorem 2.18 **(Banach-Alaglou).** *If* X *is a normed vector space, then the closed unit ball in* X^* *is weak⋆ compact. If* X *is separable, then the closed unit ball is sequentially weak⋆ compact.*

A conclusion of this is that if $f_n \in L^p$ with $1 < p < \infty$, is bounded, then there exists $f \in L^p$ and a subsequence such that $\langle f_{n_k}, g\rangle \to \langle f, g\rangle$ for all $g \in L^{p'}$

Definition 2.8 **(Hölder Continuous).** Λ_α *consists of all* f *such that* $\sup_{x\neq y} |f(x)-f(y)||x-y|^{-\alpha} < \infty$ *with* $0 < \alpha < 1$

Theorem 2.19 **(Fourier Decay for Hölder functions).** *if* $f \in \Lambda_\alpha$ *with compact support, then* $\widehat{f} = \mathcal{O}(|\xi|^{-\alpha})$

Trick is break integral translation – translating a function only scales the Fourier transform by a modulus 1 complex number.

Proof. (Same proof for Lipschitz functions).

1. $f(\xi) = \int e^{-ix\xi} f(x)dx = \frac{1}{2}\int e^{-ix\xi}f(x)dx + \frac{1}{2}\int e^{-i(x+\frac{\pi\xi}{|\xi|^2})\xi} f(x + \frac{\pi\xi}{|\xi|^2})dx$

2. $e^{-i\pi} = -1$, so $|f(\xi)| \leq \int |f(x) - f(x + \frac{\pi\xi}{|\xi|^2})| dx$

3. by compact support of f, and $|\xi| > \pi$, we can bound this by $C \|f\|_{\Lambda_\alpha} |\xi|^{-\alpha}$

□

This bound is sharp:

Example 2.1. *If $f(x) = \sum_{i=0}^{\infty} 2^{-\alpha k} e^{2^k i\pi x}$, then $f \in \Lambda_\alpha$ but $\langle k \rangle^\alpha \widehat{f}(k) = 1$ for infinitely many k.*

Theorem 2.20 (Riemman-Lebesuge Lemma). *If $f \in L^1(\mathbb{R}^d)$, then $|\widehat{f}(\xi)| \to 0$ as $|\xi| \to \infty$*

Proof. Same idea as Λ_α Fourier decay, but use continuity of translation with respect to L^1 norm[a].

1. $f(\xi) = \int e^{ix\xi} f(x) dx = \frac{1}{2} \int e^{ix\xi} f(x) dx + \frac{1}{2} \int e^{i(x + \frac{\pi\xi}{|\xi|^2})\xi} f(x + \frac{\pi\xi}{|\xi|^2})$

2. combine, take absolute value, get: $|f(\xi)| \leq \|f(x) - f(x + \theta t)\|_{L^1(\mathbb{R}^d_x)}$, with $\theta = \pi\xi/|\xi|$ and $t = |\xi|^{-1}$.

3. This goes to zero.

□

Proposition 2.2. *The Fourier transform does not map $L^1 \to C^0_{\to 0}$ surjectively.*

Proof. If it was, then it is open (by the open mapping theorem) and injective (by fourier inversion theorem), so it's inverse is bounded. So there is $C > 0$ such that $\|f\|_{L^1} \leq C \|\widehat{f}\|_{C_0}$. But this is impossible, let $f_t = e^{-(1+it)|x|^2/2}$ and send $t \to \infty$. □

This is a trick to disprove surjectivity for injective bounded maps: if $T : X \to Y$ is a bounded, injective map, then surjectivity implies existence of $C > 0$ such that $\|f\|_X \leq C \|Tf\|_Y$ for all $f \in X$

Also the coefficients of an L^1 function can decay arbitrarily slowly:

Theorem 2.21 (Sharpness of Riemman-Lebesgue Lemma). *If $g(\xi) \to 0$ as $|\xi| \to \infty$ and be continuous and positive, then there exists $f \in L^1$ with $|\widehat{f}| \geq g$*

Theorem 2.22 (Hausdorff-Young Inequality). *For $p \in [1, 2]$, the Fourier transform is a bounded operator $L^p \to L^{p'}$*

This follows immediately from the Riesz-Thorin theorem.

Remark 2.1 (**Some Remarks About Hausdorff-Young**).

1. *If $\mathcal{F} : L^r \to L^p$ is bounded, then $r = p'$ (follows by a scaling argument)*

2. *It is enough to show $\mathcal{F} : L^{p_1} \to L^{p_2}$ (don't need to show bounded), and apply closed graph theorem*[b].

[a] this follows easily by approximation by C_0^∞ functions

[b] although it is easy to show the end points are bounded

3. *if $p > 2$, then $\mathcal{F}(L^p) \notin L^{p'}$ (H-Y is optimal on L^p spaces)*

4. *$p \in [1,2)$, $\mathcal{F} : L^p \to L^{p'}$ is not surjective.*

There are two tricks to showing counter examples to Fourier mapping properties (1) scaling (2) Gaussian with a scaled parameter

Proof.
(3 boundedness failure) Let $f_t(x) = e^{(1+it)|x|^2/2}$, then $|f_t|$ is independent of t, so $\|f_t\|_p$ is constant. While $\widehat{f_t}(\xi)$ has a term $(1+it)^{-d/2}$. Taking absolute values, we get $\left\|\widehat{f_t}\right\|_q^q \geq C\langle t\rangle^{d-\frac{qd}{2}}$. If $p > 2$, then $q < 2$, and this quantity grows with t, so $\left\|\widehat{f_t}\right\|_q \leq C\|f_t\|_p$ cannot be true.

(3 mapping failure) assume true, use duality:

1. pick $f \in L^p$, $\|f\|_p \leq 1$, $p > 2$, $q = p'$, $\left\|\widehat{f}\right\|_p = \sup_{\|g\|_q=1} \int \widehat{f}g$. So we want to prove the functional $\ell_f \in (L^q)^*$, $g \mapsto \int \widehat{f}g$ is bounded.
2. for each fixed g, $|\ell_f(g)| = |\int \widehat{f}g| = |\int f\widehat{g}| \leq \|\widehat{g}\|_q < \infty$
3. by uniform boundedness principal, we get $\|\ell_f\|$ are uniformly bounded, but this is $\left\|\widehat{f}\right\|_{L^q}$, and so we get that $\mathcal{F}$ is bounded $L^p \to L^q$, which is a contradiction

(4) Assume false, use the open mapping theorem to get a constant C such that $\|f\|_p \leq C\left\|\widehat{f}\right\|_{p'}$. Let $f_t = e^{-(1+it)|x|^2/2}$ to see that no such C can exist. □

Theorem 2.23 (Uniform Boundedness Principal). *If T is a family of continuous maps $X \to Y$ on Banach spaces*[a] *that is pointwise bounded ($\sup_{F\in T}|F(x)| < \infty$ for each $x \in X$), then it is uniformly bounded: $\sup_{F\in T}\|T\|_{X\to Y} < \infty$*

2.2.2 Rademacher Functions and Khinchine's Inequality

Definition 2.9 (Rademacher Functions). *Define $r_n(x) = \sum_{k=1}^{2^n} 1_{D_{n_k}}(x)(-1)^{(k+1)}$, with D_{n_k} the k^{th} dyadic interval of length 2^{-n}*

Proposition 2.3 (Basic Properties of Rademacher Functions). *r_n are orthonormal in L^2, but are not complete. This is because $r_1 r_2 \neq 0$ but is orthogonal to all r_i.*

If you index on $n \geq 1$, these are i.i.d. mean zero random variables on the measure space $[0,1]$ with Lebesgue measure.

Theorem 2.24 (Khinchine's Inequality). *For $c \in \ell^2$, for all $q \in (0,\infty)$ there exist $C_q > 0$ such that:*

$$C_q^{-1}\left\|\sum c_n r_n\right\|_{L^q} \leq \|c\|_{\ell^2} \leq C_q\left\|\sum c_n r_n\right\|_{L^q}$$

If $f = \sum c_n r_n$, then $\|f\|_{L^2} = \|c\|_{\ell^2}$, and this theorem says $\|f\|_{L^q} \sim \|f\|_{L^2}$

[a] the range can just be a normed vector space

Proof. Idea: bound above by even L^p, expand the sum, use independence of Rademacher functions

1. choose an even integer, $2p$, greater than q, since $\|f\|_q \leq \|f\|_{2p}$, it suffices to bound $\|f\|_{2p}$
2. $\|f\|_{2p}^{2p} = \int \bar{f}^p f^p dx = \int (\sum \bar{c}_n r_n)^p (\sum c_n r_n)^p$, recall how to expand this to get:

$$\|f\|_{2p}^{2p} = \sum_{\substack{n_1,\ldots,n_p=1 \\ m_1,\ldots,m_p=1}}^{\infty} \prod_{i,j=1}^{p} \bar{c}_{n_i} c_{m_j} \int_0^1 \prod_{i,j=1}^{p} r_{n_i} r_{m_j} dx$$

3. since r_i are mean zero independent random variables, for each n_i, m_i, the integral is:

$$\prod_{i,j=1}^{p} \mathbb{E}[r_{n_i} r_{m_j}] = \prod_{i,j=1}^{p} \delta_{n_i = m_i}$$

   this is zero unless $n_i = m_i$ for $i = 1, \ldots, q$. If we fix n_i, then we can rearragne the q entries of m_i, so we have $p!$ terms which are nonzero.

4. so $\|f\|_{2p}^{2p} = p! \sum_{n_1,\ldots,n_p=1}^{\infty} \prod_{j=1}^{p} |c_{n_j}|^2 = p! \, \|c\|_{\ell^2}^{2p}$

□

Theorem 2.25 (Kahane's Theorem). *If $a \in \ell^2$, there exists $f \in L^\infty(\mathbb{T})$ such that $|\widehat{f}(n)| \geq |a_n|$ for all n. Furthermore there exists $f \in \bigcap_{p<\infty} L^P$ with $|\widehat{f}(n)| = |a_n|$*

Since $L^2 \to \ell^2$ and $\ell^2 \to L^2$ by forward and backwards Fourier transform, it would be nice if $L^1 \to \ell^\infty$

Proof. Idea: let f be the inverse Fourier transform of a but with signs that are i.i.d. mean zero random variables. The L^p norm has finite expectation, so is almost surely finite. (of weaker statement).

1. Fix p, consider $f_\omega(x) = f(x) = \sum a_n r_n(\omega) e^{ix \cdot n}$, with $\omega \in [0,1]$. $|\widehat{f}(n)| = |a_n|$.
2. $\int \|f\|_{L^p}^p d\omega = \iint |f|^p d\omega dx \leq C \int \|a\|_{\ell^2}^p < \infty$ by Kitchine's inequality
3. So for almost every $\omega \in [0,1]$, $f_\omega \in L^p$.
4. take $p_n \to \infty$, take intersections of these full measure sets

□

2.2.3 Uniform and Pointwise Convergence of Fourier Series (Dirichlet Kernels, Cesaro means)

Here is a brief summary:

- Fourier series need not converge, they are funny. This is largely because Fourier series can be written as convolutions with certain functions (Dirichelt kernels) which just miss being an approximate identity sequence (unbounded L^1 norm).

- pointwise Fourier series convergence can fail for even C^0 functions
  - If we add up our series in a funny way (Cesaro means), then we can get uniform convergence in C^0 of trigonometric polynomials.
- there exists an L^1 function whose Fourier series diverges everywhere (Kolmogorov)
- pointwise Fourier series convergence holds at x_0 if $\int \frac{|f(x_0)-a|}{|x_0-x|}dx < \infty$ for some a ($f(x_0)$ if f is C^0).
- Fourier series uniformly converge for $f \in \Lambda_\alpha$ (so there is a transition from C^0 to C^1).
- Fourier series of functions in L^p ($p \in (1,\infty)$) converge in L^p.
  - In fact, Fourier series converge almost everywhere for $f \in L^p$, $p \in (1,\infty]$ (Carleson), but this is very hard to show.

Definition 2.10 (Dirichlet Kernel). *The Dirichlet Kernels are defined* $D_N(x) \coloneqq \sum_{n=-N}^{N} e^{inx} = \frac{\sin((N+\frac{1}{2})x)}{\sin(x/2)}$.[a]

Proposition 2.4 (Purpose of Dirichlet Kernels). *For* $f \in L^1(\mathbb{T}^d)$*, then* $D_N * f = S_N(x) = \sum_{n=-N}^{N} \widehat{f}(n)e^{inx}$ *(where the convolution has a normalizing factor* $(2\pi)^{-1}$

Note that in Fourier space, $\widehat{D}_N = 1_{[-N,N]}$. By the Poisson summation formula, we have that $\sum_{n\in\mathbb{Z}} e^{inx} = 2\pi \sum_{n\in\mathbb{Z}} \delta_{2\pi n}(x)$. So the Dirichlet kernel converges, in some sense, to a periodic delta distribution.

While $\int D_N = 2\pi$, $\|D_N\|_1 \to \infty$ so it is not an approximate identity sequence. (It also fails to have a support that is shrinking).

Theorem 2.26 (L^1 norm of Dirichlet Kernel). $\|D_N\|_{L^1} \geq c \log N$

Theorem 2.27 (Failure of pointwise convergence of Fourier Series). *There exists* $f \in C^0(\mathbb{T})$ *such that* $S_N(0)$ *doesn't converge.*

Proof. idea: use UBP, get contradiction by duality argument and unbounded L^1 norm of D_N

1. let ℓ_N be the functional $\ell_N(g) = S_n(0) = \int g(0-y)D_N(y)dy$.
2. assuming statement is true, ℓ_N are pointwise bounded, therefore by UBP, they are bounded uniformly
3. But $\|\ell_N\| = \sup_{g\in C^0:\|g\|_{C^0}=1} \geq |\ell(1)| = \|D_N\|_1 \to \infty$

□

Theorem 2.28 (Pointwise convergence of Fourier series). *If* $f \in L^1(\mathbb{T}^d)$*,* $x_0 \in \mathbb{T}$*,* $a \in \mathbb{C}$ *and* $\int |f(x_0) - a|/|x - x_0|dx < \infty$*, then* $S_N f(x_0) \to a$

[a] really they are defined this for $x \neq 0$ and $2N+1$ for $x = 0$

Proof. idea: act like normal approximation of identity, rewrite integral as fourier transform of L^1 function, apply Riemman-Lebesgue Theorem

1. WLOG, $x_0 = 0$, $S_N(0) - a = (2\pi)^{-1}\int D_N(y)(f(0-y) - a)$ (since $⨍ D_n = 1$)

2. $$\int D_N(y)(f(-y) - a) = \int_{-\pi}^{\pi} \frac{\sin((\frac{N}{2}+1)y)}{\sin(y/2)}(f(-y)-a)dy = \int_{-\pi}^{\pi} \frac{e^{i(\frac{N}{2}+1)y} - e^{-i(\frac{N}{2}+1)y}}{2i\sin(y/2)}(f(-y)-a)dy$$
$$= \frac{1}{2i}\left(\int_{\mathbb{R}^n} \frac{e^{i(\frac{N}{2}+1)y}(f(-y)-a)1_{[-\pi,\pi]}}{\sin(y/2)} - \int_{\mathbb{R}^n} \frac{e^{-i(\frac{N}{2}+1)y}(f(-y)-a)1_{[-\pi,\pi]}}{\sin(y/2)}\right)$$
$$= \frac{1}{2i}(\widehat{h}(\frac{N}{2}+1) + \widehat{h}(-\frac{N}{2}-1))$$
with $h(x) = 1_{[-\pi,\pi]}(f(-y) - a)(\sin(y/2))^{-1}$.

3. by hypothesis, $h \in L^1$, so $\widehat{h}(N) \to 0$ (by the Riemann-Lebesgue lemma).

□

Theorem 2.29 (Uniform Convergence of Fourier Series). *If $f \in \Lambda_\alpha(\mathbb{T})$, then $\|S_N f - f\|_{C^0} \le C_\alpha \|f\|_{\lambda_\alpha} N^{-\alpha}\log(N)$*

This is very important, the proof is rather lengthy, but the summary is short. This is improved with Cesaro means.

Proof.

1. wlog, show $|S_n(0)| \to 0$.

2. $S_n(0) = \frac{1}{2\pi}\int_{-\pi}^{\pi} f(x)\frac{\sin(Mx)}{\sin(x/2)}dx$ with $M = \frac{N}{2}+1$.

3. split integral into two parts, the first: $\int_{|x|\le\delta} f(x)\frac{\sin(Mx)}{\sin(x/2)}dx$, this is bounded by $C\int_{|x|\le\delta}|f(x)||x|^{-1}dx \le C\int_{|x|\le\delta}|x|^{\alpha-1}\|f\|_{\Lambda_\alpha}dx = C\|f\|_{\Lambda_\alpha}\delta^\alpha$

4. second term: $\int_{|x|\ge\delta} g(x)\sin(Mx)dx$. Split $\sin(Mx)$ into exponential terms, it suffices to look at only one (where we use the shifting trick):
$$\int_\delta^\pi g(x)e^{iMx}dx = \frac{1}{2}\int_\delta^\pi g(x)e^{iMx}dx - \frac{1}{2}\int_{\delta-\frac{\pi}{M}}^{\pi-\frac{\pi}{M}} g(x+\frac{\pi}{M})e^{iMx}dx$$

5. we get a lot of integrals, each can be handled either trivially, or using the Holder continuity condition.

□

Definition 2.11 (Cesaro means). *For a function f on $\mathbb{T}$, let $S_N(x) = \sum_{n=-N}^{N} \widehat{f}(n)e^{inx}$, define the Cesaro mean $\sigma_N f(x) := \frac{1}{N+1}\sum_{n=0}^{N} S_N(x)$*

Theorem 2.30 (Facts about Cesaro Sums).

1. *$\sigma_N f(x) = K_N * f$, with K_N the **Fejer Kernel** $K_N(x) = \frac{1}{N+1}\sum_{n=0}^{N} D_n = (N+1)^{-1}\frac{\sin^2((N+1)x/2)}{\sin^2(x/2)}$*
2. *In Fourier space, we have:*

$$\widehat{K}_N(n) = \begin{cases} 1 - \frac{|n|}{N+1} & |n| \le N \\ 0 & else \end{cases}$$

3. *The Fejer Kernels form an approximate identity sequence, therefore:*
   (a) *$f \in C^0$ implies that $\sigma_N f$ converges uniformly to f*
   (b) *$f \in L^p$ implies that $\sigma_N f$ converges in L^p to f*
4. *$f \in \Lambda_\alpha(\mathbb{T}^1)$ implies that $\|\sigma_N f - f\|_{C^0} \le C_\alpha \|f\|_{\Lambda_\alpha} N^{-\alpha}$ (here the Holder norm includes the supremum norm)*

Proof. To show $\sigma_N(x)$ is an approximate identity sequence, the only nontrivial thing to show is that $\int_{|x|>\varepsilon} \sigma_N(x)dx \to 0$. Replace N with $n-1$, use that $\sin(x/2) \le cx$:

$$\int_{|x|>\varepsilon} \sigma_N(x)dx \le Cn^{-1}\int_\delta^\pi x^{-2}dx = n^{-1}(\varepsilon^{-1} - \pi^{-1}) \to 0$$

□

Theorem 2.31 (Fourier Series Convergence of Bounded Variation Functions). *If $f \in C^0(\mathbb{T})$ and has bounded variation, then $S_N f \to f$ uniformly.*

This can be weakened to $f \in C^0(\mathbb{T})$ and $\widehat{f}(n) = \mathcal{O}(|n|^{-1})$ (which bounded variation functions satisfy).

2.2.4 Almost Everywhere Divergence

The construction of an L^1 function whose partial series diverge everywhere relies on the following:

Theorem 2.32 (Kronecker's Theorem). *If $t_1, \ldots, t_m \in \mathbb{R}$ are such that $\cup_i t_i \cup 1$ are linearly independent over the rationals, then for all ε and $z_j \in \mathbb{C}$ with $|z_j| = 1$, there exists $n \in \mathbb{N}$ such that:*

$$|e^{2\pi i n t_j} - z_j| < \varepsilon$$

for all $j = 1, \ldots, mj$

(This seems related (and probably provable by) to the Poincare recurrence theorem).

Proof. if not, get f supported on a ball that avoids integer periods. Average over it, taking Fourier transform, apply dominated convergence theorem to get contradiction (not super intuitive).

1. assume false, get $z_0 \in \mathbb{T}^m$ and $\varepsilon_0 > 0$ such that $e^{2\pi i n t}$ never lands in $B_\varepsilon(z_0)$ for $n \in \mathbb{N}$ (where $t = (t_1, \ldots, t_m) \in \mathbb{R}^m$)
2. let $f \in C^\infty$ be supported in $B_\varepsilon(z_0)$ with $f \geq 0$, $\int f = 1$.
3. $f(nt) = \sum_{k \in \mathbb{N}^n} \widehat{f}(k) e^{2\pi i n t \cdot k}$, take average, interchange sums:
$$N^{-1} \sum_{n=1}^{N} f(nt) = \sum_{k \in \mathbb{N}^n} \widehat{f}(k) N^{-1} \sum_{n=1}^{N} e^{2\pi i n t \cdot k}$$
4. apply dominate convergence theorem (for sums) to the RHS
   (a) it is clearly ℓ^1
   (b) for fixed k, have $N^{-1} \sum_1^N e^{2\pi i n \cdot k} \leq N^{-1} \frac{2}{|1 - e^{2\pi t \cdot k}|}$ (the denomoniator is never zero by hypothesis[a]) which goes pointwise to 0 for $k \neq 0$.
   (c) by DCT, get $\widehat{f}(0) = c_d \int f$
5. LHS is zero by hypothesis and support of f, so we get $0 = \int f$, which is a contradiction.

□

Another Kernel is used to get a variant of partial sums:

Definition 2.12 (Vallée Poussin Kernel). *Define* $V_N = 2K_{2N+1} - K_N$ *(*K_N *are Fejer Kernels). These have the properties: (1)* $\widehat{V_N}(n) = 1$ *for* $|n| \leq N+1$ *and* $\widehat{V_n}(n) = 0$ *for* $|N| \geq 2N+2$ *(2)* V_N *is an approximate identity sequence (*$\|V_N\|_1 \leq 3$*)*

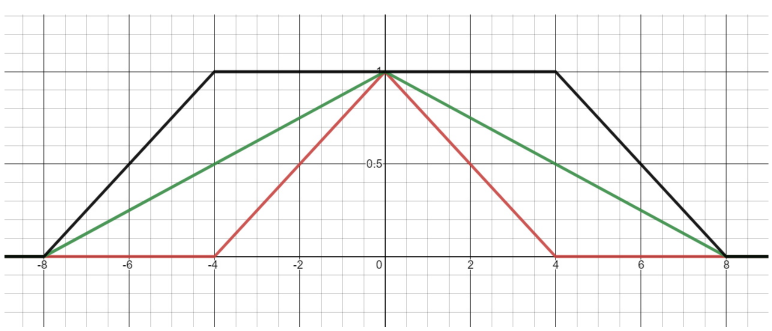

Figure 1: Fourier Coefficients for $n = 3$. Green and Red are Fejer kernels, and Black (the weighted sum) is the Vallee Poussin kernel.

[a] it is zero if and only $e^{2\pi t \cdot k} = 1$ if and only if $t \cdot k \in \mathbb{Z}$ if and only if t_i are linearly **dependent** with respect to 1

Theorem 2.33 (Kolmogorov's Divergence Theorem). *There exist $f \in L^1(\mathbb{T})$ such that $S_n f$ diverges almost everywhere.*

Proof. (main ideas)

1. For M, choose $\{y_j\}_{j=1}^M$ elements of $[-\pi, \pi]$ which are linearly independent over $\mathbb{Q}$ (along with π) and that are approximately evenly spaced.

2. let $\mu = M^{-1} \sum_{j=1}^M \delta_{y_j}$, then by writing out the Dirichlet Kernels and using **Kronecker's theorem**, we get for each x, a $N = N(x)$, such that $S_N \mu(x) \geq c \log M$.

3. this implies that for all $A > 0, \varepsilon > 0$, there exist $K < \infty$ such that $\sup_{N \leq K} |S_N \mu(x)| \geq A$ for all $x \in \mathcal{T} \setminus E$ with $|E| < \varepsilon$ [a]

4. the same is true with μ replaced by a trigonometric polynomial g with $\|g\|_1 = 1$

   (a) this by $g = \mu * V_k$

   (b) then $\widehat{g}(n) = \widehat{\mu}(n)$ for $|n| \leq k$, therefore $S_N g = S_N \mu$ for $N \leq k$.

   (c) $\|g\|_1 = \|\mu * V_K\|_1$ which is convex combination of translations of V_k, which has norm bounded by 3 (so we normalize to get $\|g\|_1 = 1$).

5. recursively choose g_j such that their fourier transforms have disjoint support and their partial sums are massive on sets whose measures approach 1. Then take $\sum 2^{-j} g_j$

□

Theorem 2.34 (Extra Fourier Decay of Holder Continuous Functions). *If $f \in \Lambda_\alpha(\mathbb{T})$ with $\alpha > 1/2$, then $\widehat{f} \in \ell^1(\mathbb{Z})$*

Proof.

1. let $f_n = V_{2^n} * f$, let $g_n = f_n - f_{n-1}$ so $f = f_0 + \sum_1^\infty g_n$

2. $g_n = V_{2^n} * f - f - (V_{2^{n-1}} * f - f)$, it can be shown that $\|V_n * f - f\|_{C_0} \leq C N^{-\alpha} \|f\|_\alpha$, therefore g_n is of order $2^{-n\alpha}$

3. use support of $\widehat{V}$ and Holder: $\|\widehat{g}_n\|_{\ell_1} \leq C 2^{n/2} \|\widehat{g}\|_{\ell^2} \leq C 2^{n/2} \|g_n\|_{L^2} \leq C 2^{n/2} \|g_n\|_{C_0} \leq C 2^{n/2} 2^{-n\alpha}$

4. therefore $\|f\|_{\ell_1} \leq C + C \sum_1^\infty 2^{\frac{k}{2} - \alpha k}$, this converges if $\alpha > 1/2$

□

This is actually sharp.

[a] this is true for a.e. x if $K \to \infty$, therefore by measure theory, we can get this statement

2.2.5 L^p norm convergence

The main ingredient is:

Theorem 2.35 **(Reisz-Thorin Theorem).** *If $T : S(X) \to (L^1 + L^\infty)(Y)$ is a linear map from simple functions on a measure space X to functions on a measure space Y such that $\|T\|_{L^{p_i} \to L^{q_i}} = A_i$ for $i = 1, 2$ and $p_i, q_i \in [1, \infty]$, then for all $\theta \in [0, 1]$, we have $\|T\|_{L^{p_\theta} \to L^{q_\theta}} \le A_1^\theta A_2^{(1-\theta)}$.*

Where $p_\theta^{-1} = \theta p_1^{-1} + (1-\theta) p_2^{-1}$ and $q_\theta^{-1} = \theta q_1^{-1} + (1-\theta) q_2^{-1}$

Proof. (idea)

1. for simple functions $f = \sum a_j 1_{E_j}$ and $g = \sum b_k 1_{F_k}$, define $f_z = \sum |a_j|^{L(z)} e^{i\varphi_j} 1_{E_j}$ and $g_z = \sum |b_k|^{K(z)} e^{i\psi_k} 1_{F_k}$. With L and K affine holomorphic functions defined on $\Re(z) \in [0, 1]$.

2. Let $\mathcal{F}(z) = (A_1^z A_2^{(1-z)}) - 1 \int T(f_z) g_z dy$. It can be shown that $\mathcal{F}(\theta) = (A_1^\theta A_2^{(1-\theta)})^{-1} \int T(f) g dy$

3. We use modified maximum principal to bound $\mathcal{F}(z)$ on the boundary of the strip to get $|\mathcal{F}(\theta)|$ is bounded by what we want.

□

Here is a Hueristic proof to remember the affine transformation

Proof. 1. take $f \in L^{p_\theta}$ and $g \in L^{q'_\theta}$ (both with norm 1). In the above, we are basically considering $H(s) := \int T f^{L(s)} \cdot g^{K(s)} dy$ for $s \in [0, 1]$

2. we want control on the actual f and g, so we would like $H(\theta) = \int T f \cdot g dy$. So we would like $L(\theta) = K(\theta) = 1$. But at the end of the day we want a bound like $|H(\theta)| \le A_1^\theta A_2^{(1-\theta)}$, so we should redefine H as $H(s) := (A_1^s A_2^{1-s})^{-1} \int T f^{L(s)} \cdot g^{K(s)} dy$

3. we want endpoint control, when $s = 0$, we have $H(0) = \frac{1}{A_2} \int T f^{L(0)} \cdot g^{K(0)} dy$. This suggests using the $L^{p_2} \to L^{q_2}$ bound:

$$\left| \int T f^{L(0)} \cdot g^{K(0)} dy \right| \le A_2 \left\| f^{L(0)} \right\|_{p_2} \left\| g^{K(0)} \right\|_{q'_2}$$

   for these terms on the right to be 1, we simply let $L(0) = \frac{p_\theta}{p_2}$ and $K(0) = \frac{q'_\theta}{q'_2}$

4. similarly, we want $L(1) = \frac{p_\theta}{p_1}$ and $K(1) = \frac{q'_\theta}{q'_1}$. It then magically works out that the affine map

□

Lemma 2.2 (**Modified Maximum Principal for Holomorphic functions**)**.** *If f is holomorphic on the strip $S = \{\Re(z) \in (0, 1)\}$, continuous on $\bar{S}$, $|f| \le 1$ on $\bar{S}$, and $|f| \le B < \infty$ on $\bar{S}$, then $|f| \le 1$ on S*

Proof. 1. let $f_\varepsilon = f(z) e^{\varepsilon z^2}$

2. apply usual maximum principal for $\{\Re z \in (0,1), \Im(z) \in (-M,M)\}$, get for large enough M that $f_\varepsilon \le 1$ on this set.

3. send $\varepsilon \to 0$.

□

The fact we can uniformly bound $F(z)$ on the strip is because we have finitely many terms, and they all only depend on the real part.

2.2.5.1 Interpolation results

The order of proof goes Young's Inequality → Hölder's inequality → interpolation on L^p → Riesz-Thorin.

First note $0 \le (a-b)^2 = a^2 + b^2 - 2ab$, so $ab \le \frac{a^2}{2} + \frac{b^2}{2}$. This is a useful bound, and is a baby version of Young's inequality:

Theorem 2.36 (Young's Inequality). *If $a, b \ge 0$, $1 < p, q$ with $\frac{1}{p} + \frac{1}{q} = 1$, then*

$$ab \le \frac{a^p}{p} + \frac{b^q}{q} \tag{6}$$

Proof. Let $t = 1/p$, so $1/q = (1-t)$ and $t \in (0,1)$. Then the logarithm right-hand-side of (6) is $\ln(ta^p + (1-t)b^q) \ge t\ln(a^p) + (1-t)\ln(b^q)$ (by concavity). And this is $\ln(a) + \ln(b) = \ln(ab)$. Take exponential to get result. □

Trick: take log, use concavity.

Theorem 2.37 (Hölder's Inequality). *If $f \in L^p$, $g \in L^q$, $p^{-1} + q^{-1} = 1$, then $\|fg\|_1 \le \|f\|_p \|g\|_q$*

Proof. Define $\bar{f} = \|f\|_p^{-1} f$ and $\bar{g}$ similarly. Then $|\bar{f}\bar{g}| \le \frac{|\bar{f}|^p}{p} + \frac{|\bar{g}|^q}{q}$ by Young. Integrate both sides: $\|\bar{f}\bar{g}\|_1 \le \frac{1}{p} + \frac{1}{q} = 1$. Multiply to get $\|fg\|_1 \le \|f\|_p \|g\|_q$ □

Theorem 2.38 (Hölder's Inequality Interpolation). *If $f \in L^p \cap L^q$, and $\theta \in (0,1)$ and $p_\theta^{-1} = p^{-1}\theta q^{-1}(1-\theta)$, then $\|f\|_{p_\theta} \le \|f\|_p^\theta \|f\|_q^{(1-\theta)}$*

Proof. Compute $\|f\|_{p_\theta}^{p_\theta}$, apply Holder's inequality to $\||f|^{p_\theta \theta}|f|^{p_\theta(1-\theta)}\|_1$ with conjugate exponents: $\frac{p_1}{p_\theta \theta}$ and $\frac{p_2}{(1-\theta)p_\theta}$. Everything cancels. □

2.2.5.2 Reisz Theorem

Theorem 2.39 (Reisz Theorem of L^p convergence of Fourier Series). *For $p \in (1,\infty)$, $S_N f \xrightarrow{L^p} f$ for all $f \in L^p$.*

Proof. decompose S_N into sum of compositions of simpler things. Bound those things by rewriting one as identity plus something. That thing can be bound using complex analysis.

1. suffices to show $\|S_n f\|_p \le C\|f\|_p$ for all trig functions p

(a) if $f \in L^p$, let $g = K_M * f$, such that $\|g - f\|_{L^p} < \varepsilon$. Then $\|S_N f - f\|_p \le \|S_N g - g\|_p + \|f - g\|_{L^p} + \|S_N(g - f)\|_p$. First term is zero since g has finite fourier coefficients, second is small by approximation, third term is what we want.

2. let $E_N f = e^{inx} f$, $\widehat{Pf} = \widehat{f} 1_{n \ge 0}$, then $S_N f = E_{-N} \circ P \circ E_N - E_{N+1} \circ P \circ E_{-N-1}$. So boundedness of S_N relies on boundedness of P.

3. let $\widehat{Hf} = -i \operatorname{sgn}(n) \widehat{f}$, then $\widehat{Pf}(n) = \frac{1}{2}(I + iH)f + \frac{1}{2}\widehat{f}(0)$. So boundedness of P relies on boundedness of H

4. By interpolation and duality[a], suffices to show bound for $p = 2q$ for $q \in \mathbb{N}$

5. H is bounded (it suffices to show this for trig polynomials u):

(a) Let f be an analytic function on $\overline{B_1(0)}$ such that $f(0) = \fint_{\mathbb{T}} u$ and $f(e^{i\theta}) = (u + iHu)(\theta)$[b]

(b) f^p is also holomorphic, so by mean value property of holomorphic functions $f^p(0) = \fint f^p$, therefore (losing constants) $(\int u)^p = \int (u + iH)^p$

(c) expand and rearrange, get $\|H\|_p^p \le C(\int u)^p + C \sum_1^p \int u^k H^{p-k} \le C \sum_1^p \int u^k H^{p-k}$

(d) $\int u^k H^{p-k} \le \|u^k\|_{p/k} \|H^{p-k}\|_{\frac{p}{p-k}} = \|u\|_p^k \|H\|_p^{p-k} \le \delta^{-1} \|u\|_p^p + \delta \|H\|_p^p$

□

Theorem 2.40 (Wiener's Tauberian Theorem). *If $a \in \ell^1(\mathbb{Z})$ and $\check{a}$ (which is an element of $C^0(\mathbb{T})$) vanishes nowhere, then $b := \mathcal{F}_{\mathbb{T}}((\check{a})^{-1}) \in \ell^1(\mathbb{Z})$*

Note that this implies $a * b = \delta_0$, i.e. a is invertible in the Banach Algebra of $\ell^1(\mathbb{Z})$ with the operation of convolution. So a slight strengthening of the theorem is $a \in \ell^1(\mathbb{Z})$ is invertible if and only if $\check{a}$ vanishes nowhere. Moreover, this theorem is significant, because all we know is that $(\check{a})^{-1} \in C^0(\mathbb{T})$ (if it vanishes nowhere), and all we know is $\mathcal{F}_{\mathbb{T}} : C^0(\mathbb{T}) \to \ell^2 \cap \ell^\infty$

Proof. Define $A = \{f \in C^0(\mathbb{T}) : \widehat{f} \in \ell^1(\mathbb{Z})\}$. So I need to show that if $f \in A$ and vanishes nowhere, then $f^{-1} \in A$. Define $\|f\|_A = \|\widehat{f}\|_{\ell^1}$

1. for $f \in A$, note that $f^{-1} = \delta \bar{f}(\delta |f|^2)^{-1}$, since $\delta \bar{f} \in A$ and A is closed, it suffices to show that $(\delta |f|^2)^{-1} \in A$

2. let $\delta |f|^2 = 1 - g$, then if we invert $(1 - g)$ and stay in A, we are done, so it now suffices to show that if $a \in \ell^1$ with $\|\check{a}\|_{C^0(\mathbb{T})} < 1$, then $(1 - \check{a})^{-1} \in A$. (Note $\delta f|$ is bounded above and below) [c]

3. note $(1 - \check{a})^{-1} = \sum_0^\infty (\check{a})^n$, take Fourier transform of both sides, get $\sum_{j=0}^\infty a^{\circledast j}$ [d], I want this to be in ℓ^1

[a] H is skew-self-adjoint
[b] this is by $f(z) = \sum_{n=0}^\infty a_n z^n$ with $a_0 = \widehat{u}(0)$, $a_n = \widehat{u + iH}(n)$ (only finitely man terms are nonzero)
[c] here $g = a$
[d] $f^{\circledast j} := \underbrace{f * f * \cdots * f}_{j \ factors}$

4. decompose $a = \alpha + \beta$, where $\operatorname{supp}\alpha \in [-M, M]$ and $\|\beta\|_{\ell^1} < \varepsilon$, then $a^{\circledast j} = \sum_{k=0}^{j} \binom{j}{k} \alpha^{\circledast k} * \beta^{\circledast(j-k)}$
   (a) $\|\alpha^{\circledast k}\|_{\ell^1} \le (2Mk+1)^{1/2} \|\alpha^{\circledast k}\|_{\ell^2} \le CM^{1/2}k^{1/2} \|\check{\alpha}\|_{C_0}^{k}$
   (b) $\|\beta^{\circledast j}\|_{\ell^1} \le \|\beta\|_{\ell^1}^{j}$
   (c) therefore $\|a^{\circledast j}\|_{\ell^1} \le \sum_{k=0}^{j} CM^{1/2}j^{1/2} \|\check{\alpha}\|_{C^0}^{k} \|\beta\|_{\ell^1}^{k-j} \binom{j}{k} \le CM^{1/2}j^{1/2}(\|\check{\alpha}\|_{C^0} + \|\beta\|_{\ell^1})^j$
5. since $\|\check{\alpha}\|_{C^0} + \|\beta\|_{\ell^1} < 1$, we get that this sum converges to something in ℓ^1

□

The idea of the proof is, after nontrivially reducing it to functions with small supremum norm, is break up the Fourier coefficients into finite support and small norm. Expand the Nuemann series as convolutions, the finite support terms can be bounded via Holder, while the small norms can be trivially bounded ($\|f * f\|_{\ell^1} \le \|f\|_{\ell^1} \|f\|_{\ell^1}$)

2.2.6 Almost Everywhere Convergence

Theorem 2.41 (Carleson's Theorem). *For all $f \in L^p(\mathbb{T})$, $p \in (1, \infty]$, $S_n f$ converges almost everywhere.*

This theorem is very difficult to show. To prove this, it suffices to show that $S_N^{\star} f = \sup_{n \le N} |S_n f(x)|$ is bounded $L^p \to L^p$. A weaker version of this is the following:

Theorem 2.42 (Kolmogorov-Seliverstov-Plessner). *For $f \in L^2(\mathbb{T}^1)$, $\|S_N^{\star} f\|_{L^2} \le C \|f\|_{L^2} \sqrt{\log N}$*

Proof.

1. for $f \in L^2$, define $n(x) \in \{1, \dots, N\}$ such that $S_N^{\star} f(x) = S_{n(x)} f(x)$, and let $Tf = S_{n(x)} f(x)$.
2. Now $Th(x) = \fint_{\mathbb{T}} D_{n(x)}(x-y)h(y)dy$ with transpose $T^t g(y) = \fint_{\mathbb{T}} D_{n(x)}(x-y)g(x)dx$. It suffices to bound T^t (as it has the same operator norm)
3. $\|T^t g\|_2^2 = \langle g, TT^t g\rangle$, and $TT^t g(x) = \fint_{\mathbb{T}} D_{n(x)\wedge n(x')}(x-x')g(x')dx'$
4. $\|T^t g\|_2^2 \le C \iint D_N^{\star}(x-x')|g(x)g(x')|dxdx' \le \|g\|_2 \|D_N^{\star} * g\|_2 \le \|g\|_2^2 \|D_N^{\star}\|_1 \le \|g\|_2^2 \log N$

□

2.2.7 Examples

It's always good to have examples of things.

Example 2.2 (Strictly Holder Continuous). *If $\alpha \in (0,1)$, then $f(x) = \sum_{n=0}^{\infty} e^{2^n ix} 2^{-n\alpha} \in \Lambda_\alpha$ and $f \notin \Lambda_{\alpha+\varepsilon}$ for all $\varepsilon > 0$. The Fourier coefficients decay at exactly $n^{-\alpha}$.*

Here is a proof of why this function is in Λ_α

1. *pick x, y (wlog $|x-y| \ll 1$), let N be such that $2^{-N} \le |x-y| \le 2^{-N+1}$*
2. *Then $f(x) - f(y) = \sum_1^N 2^{-k\alpha}(e^{2^k ix} - e^{2^k iy}) + \sum_{N+1}^{\infty} 2^{-k\alpha}(e^{2^k ix} - e^{2^k iy})$*

3. *second term is bounded by* $2\sum_{N+1}^{\infty} 2^{-k\alpha} \le 2C_\alpha 2^{-(N+1)\alpha} \le C_\alpha |x-y|^{-\alpha}$

4. *first term, use mean value theorem to get bound:* $\sum_1^N 2^{-k\alpha}2^k|x-y| \le C_\alpha|x-y|2^{N(1-\alpha)} \le C_\alpha|x-y||x-y|^{1-\alpha} = C_\alpha|x-y|^{-\alpha}$

2.3 Hardy-Littlewood Maximal Function

Reference: Christ 4.1-4.3, 4.5, 4.7

Weak L^p, distribution functions, Hardy-Littlewood maximal function, Marcinkiewicz Interpolation theorem, Calderón-Zygumund decomposition, BMO functions, John-Nirenberg inequality

2.3.1 Weak L^p space

Definition 2.13 (Distribution Function). *For a measurable function f, define $\lambda_f(\alpha)\mu(|f| > \alpha)$*

By Chebyshev's inequality: $\lambda_f(\alpha) \le \alpha^{-p}\|f\|_p^p$ if $f \in L^p$. The natural question is: if $\lambda_f(\alpha) \le C^p\alpha^{-p}$, is $f \in L^p$? The answer is no, but we call this weak L^p:

Definition 2.14 (Weak L^p Space). *Define $L^{p,\infty}$ as all measurable functions f, such that there exists $C = C(f) \ge 0$ such that $\lambda_f(\alpha) \le \alpha^{-p}C^p$. The smallest C is the $L^{p,\infty}$ norm (although it is not a norm (it fails the triangle inequality).*

Example 2.3. $|x|^{-d/p} \in L^{p,\infty} \setminus L^p$. *To see this, note $\lambda_f(\alpha) = ||x|^{-d/p} > \alpha| = ||x| < \alpha^{-p/d}| = |B_{\alpha^{-p/d}}(0)| = \alpha^p|B_1(0)|$. Therefore $\lambda_f(\alpha)\alpha^{-p} \le (|B_1(0)|^{1/p})^p$*

A useful identity is:

$$\|f\|_p^p = p\int_0^\infty \alpha^{p-1}\lambda_f(\alpha)d\alpha \tag{7}$$

2.3.2 Hardy-Littlewood Maximal Function

Definition 2.15 (Hardy-Littlewood Maximal Function). *For $f \in L^1_{loc}(\mathbb{R}^n;\mathbb{C})$, the Hardy-Littlewood maximal function is defined $Mf(x) = \sup_{r>0} \fint_{B_r(x)} |f(y)|dy$*

Theorem 2.43 (Boundedness of HLMF). *The HLMF is bounded $L^p \to L^p$ $(p \in (1,\infty])$ and $L^1 \to L^{1,\infty}$*

This is proven via interpolation on the end points.

We require the following lemma:

Theorem 2.44 (Vitali Covering Lemma). *Given any open cover of a compact set K, $\{B_\alpha\}_{\alpha\in A}$, there exists a finite subcollection $B_1,\dots,B_n$ that are disjoint and $K \subset \bigcup_{i=1}^n B_n^*$ where B_i^* is the same ball but with 3 times the radius.*

Proof. The obvious thing works, showing it works is a little work (that I omitted)

1. get a finite subcover, order it from largest measure to smallest measure: $B_1,\dots,B_n$

2. select B_1, select B_i only if it doesn't intersect the previous selected balls.
3. to show this works, suffices to show that each of the original finite subcollection balls is contained in the enlarged chosen ones.

□

Step 1. $\|Mf\|_{L^{1,\infty}} \le C\|f\|_{L^1}$

Proof. Use inner-regularity and Vitali covering lemma

1. for each $\alpha > 0$ let $U_\alpha = \{Mf > \alpha\}$, by inner regularity, it suffices to control measures of compact $K \subset U_\alpha$
2. for each $x \in K$, there exists a ball of radius $r = r(x)$, $B_r(x)$, such that $⨍_{B_r(x)} |f| > \alpha$. This is a an open cover, reduce to a collection by the **Vitali Covering Lemma**
3. $|K| \le 3^n \sum_1^n B_i(x) \le 3^n\alpha^{-1}\sum_1^n \int |f|1_{B_i} \le 3^n\alpha^{-1}\|f\|_1$
4. Therefore $|U_\alpha| \le \alpha^{-1}3^n\|f\|_1$

□

Step 2. $\|Mf\|_{L^\infty} \le C\|f\|_\infty$ (this is trivial)

Step 3. Interpolate:

Theorem 2.45 (Marcinkiewicz Interpolation Theorem). *If T is a sublinear operator such that for $f \in L^{p_0} \cap L^{p_1}$, $\|Tf\|_{L^{q_k,\infty}} \le A_k\|f\|_{L^{p_k}}$ (for $k = 0, 1$), then for all $\theta \in (0,1)$: $\|Tf\|_{L^{q_\theta}} \le A_\theta\|f\|_{L^{p_\theta}}$ (where $p_\theta^{-1} = \theta p_0^{-1} + (1-\theta)p_1^{-1}$ (and same for q_θ).*

This is true for $p_k, q_k \in [1,\infty]$, but we require $q_k \ge p_k$ and $q_0 \ne q_1$. And we define $L^{\infty,\infty} = L^\infty$.

Here is a proof of the case when $p_0 = q_0 = 1$, $p_1 = q_1 = \infty$:

Proof. repeatedly use (7), first on Tf, control λ_{Tf} by by splitting into small and big part (the small part vanishes by L^∞ bound). The big part is controlled by L^1 bound, use (7) to control this by λ_f, put everything together, use (7) to get to $\|f\|_p$

1. let $f \in L^p$ ($p \in (1,\infty)$, let $f = g + h$ were $g = g_\alpha = f1_{|f|<\alpha/2}$ and $h = h_\alpha = f1_{|f|\ge\alpha/2}$
2. $\|Tf\|_p^p = p\int_0^\infty \alpha^{p-1}\lambda_{Tf}(\alpha)d\alpha$
3. $\lambda_{Tf}(\alpha) = |\{|Tf| > \alpha\}| = |\{|T(h+g)| > \alpha\}| \le |\{|Th| > \alpha/2\}| + |\{|Tg| > \alpha/2\}|$. (WLOG $\|T\|_{L^\infty\to L^\infty} \le 1$ (otherwise divide α by this norm) so $\|Tg\|_\alpha \le \alpha/2$, so $|\{|Tg| > \alpha/2\}| = 0$)
4. $|\{|Th| > \alpha/2\}| = \lambda_{Th}(\alpha/2) \le (\alpha/2)^{-1}\|Th\|_{1,\infty} \le C\alpha^{-1}\|h\|_1$ (because $\|T\|_{L^1\to L^{1,\infty}} < \infty$)
5. $\|h\|_1 = \int_0^\infty \lambda_{h_\alpha}(t)dt$. Now $\lambda_{h_\alpha} = \lambda_f(t)1_{t>\alpha/2} + \lambda_f(\alpha/2)1_{t\le\alpha/2}$
6. $\|Tf\|_p^p \le c\int_0^\infty \alpha^{p-1}\alpha^{-1}(\int_0^{\alpha/2}\lambda_f(\alpha/2) + \int_{\alpha/2}^\infty \lambda_f(t)dt)d\alpha \le C\|f\|_p^p$ (by Fubini).

□

A consequence of weak boundedness is:

Theorem 2.46 (Lebesgue Differentiation Theorem). *For $f \in L^1_{loc}(\mathbb{R}^n)$, then for almost every $x \in \mathbb{R}^n$, $\lim_{r\to 0^+} ⨍_{B_r(x)} |f - f(x)|dy = 0$*

Proof. Decompose into continuous and small L^1 part. L^1 part is controlled by Hardy-Littlewood, use weak boundedness of HL to to bound measure of the bad set

1. by cutting off f, we can assume $f \in L^1$, let $f = g + h$ with $g \in C^\infty$ and $\|h\|_1 < \varepsilon$
2. $F_r(x) := ⨍_{B_r(x)} |f(y) - f(x)|dy \le ⨍_{B_r(x)} |g(y) - g(x)|dy + ⨍_{B_r(x)} |h(y) - h(x)|dy$, the first term goes to zero as $r \to 0^+$ by continuity of g. The second term is bounded by $|h(x)| + ⨍_{B_r(x)} |h(y)|dy \le |h(x)| + Mh(x)$
3. for each $\delta > 0$, $|\{\limsup_{r\to 0} F_r(x) > \delta\}| \le |\{|h| + Mh(x) > \delta\}| \le |\{|h| > \delta/2\}| + |\{Mh > \delta/2\}|$.
   (a) $|\{|h| > \delta/2\}| \le \frac{2}{\delta}\|h\|_1 \le \frac{2\varepsilon}{\delta}$
   (b) $|\{Mh > \delta/2\}| \le \frac{2}{\delta}\|h\|_1 \|M\|_{L^1\to L^{1,\infty}} \le \frac{C\varepsilon}{\delta}$

□

This is a very common proof technique used throughout harmonic analysis. The proof hinges on proving the maximal operator: $f \mapsto \limsup_{r\to 0^+} ⨍_{B_r(x)} |f(y) - f(x)|dy$ is bounded $L^1 \to L^{1,\infty}$.

Definition 2.16 (Dyadic Maximal Function). *For $f \in L^1_{loc}$, define $M_{\mathcal{D}}f(x) = \sup_{\substack{Q\ni x\\ Q\in\mathcal{D}}} ⨍_Q |f(y)|dy$, where $\mathcal{D}$ is the set of dyadic intervals in $\mathbb{R}^n$*[a]

Theorem 2.47 (Dyadic Maximal Function Boundedness). *$M_{\mathcal{D}}$ is bounded $L^1 \to L^{1,\infty}$*

This actually follows from the fact that $M_{\mathcal{D}}f(x) \le Mf(x)$ for all x[b] (note $f \le g \Rightarrow \lambda_f \le \lambda_g$). Here is an independent proof.

Proof. easy: establish upper-bound on size of cubes, choose maximal cubes by disjointedness and containment properties of cubes

1. need to bound $\alpha\lambda_{M_{\mathcal{D}}f}(\alpha)$, fix α, if $M_{\mathcal{D}}f(x) > \alpha$, then $⨍_Q |f| > \alpha$ flipping things around we see $|Q| < \alpha^{-1}\|f\|_1$
2. Since we have an upper bound on the size of the dyadic cubes, for each $x \in \{M_{\mathcal{D}}f > \alpha\}$, choose Q_x maximally sized (in containment and size) that satisfies the desired inequality
3. then $\{M_{\mathcal{D}}f > \alpha\} = \cup_j Q_j$ with Q_j disjoint, therefore $\lambda_{M_{\mathcal{D}}f}(\alpha) = \sum_j |Q_j| < \alpha^{-1}\sum \int_{Q_j} |f| \le \alpha^{-1}\|f\|_1$

---

[a] Dyadic cubes are of the form $c2^k + 2^k[0,1)^d$ where $c \in \mathbb{Z}^d$ and $k \in \mathbb{Z}$ (we first pick a coordinate c, scale it by 2^k which can be big or small, then fill in the cube with side lengths 2^k)

[b] for any $Q \ni x$ dyadic, then $⨍_Q |f| \le |Q|^{-1}\int_B |f|$ with $Q \subset B$, the radius of B is the longest diagonal of the cube: $n^{1/2}2^k$, and therefore has volume $|B_1(0)|n^{n/2}2^{nk} = C_n 2^{nk} = C_n|Q|$, so $|Q|^{-1}\int_B |f| \le C_n ⨍_B |f| \le C_n Mf$

□

Theorem 2.48 (Pointwise Convolution Bound using HLMF). *If $f \geq 0$ is measurable on $\mathbb{R}^n$, $L^1 \ni K = k(|x|) \geq 0$ with $k(r)$ nondecreasing, then $f * K(x) \leq \|K\|_1 Mf(x)$*

Proof.

1. Let $k_n(x) = \sum_{j=1}^{\infty} a_j 1_{[0,\frac{j}{n}]}$ approximate $k(x)$ (pointwise monotically), and let $K_n(x) := k_n(|x|)$
2. $f * K_n(0) = \int f(y)K_n(y)dy = \sum_{j=1}^{\infty} a_j \int_{|y|\leq\frac{j}{n}} f(y)dy = \sum_{j=1}^{\infty} a_j |B_{j/n}(0)| ⨍_{B_{j/n}(0)} f(y)dy \leq \|K_n\|_1 Mf(0)$
3. by MCT, $f * K_n(0) \to f * K(0)$. By translation invariance, we get this for all x[a]

□

This has the following application to the Dirichlet problem:

$$\begin{cases} \Delta u(t,x) = 0 & t \geq 0 \\ u(0,x) = f \end{cases}$$

Under certain assumptions on f, we have a solution given by $u(t,x) = \mathcal{P}_t * f(x)$, where $\mathcal{P}_t(x)$ is the **Poisson kernel**

$$\mathcal{P}_t(x) = \frac{c_d t}{(t^2 + |x|^2)^{(d+1)/2}}$$

$\mathcal{P}_t$ satisfies the hypothesis of the previous theorem with $\|\mathcal{P}_t\|_1 = 1$. Therefore we get:

1. for $f \in L^1 + L^\infty$, $\sup_{t>0} |u(x,t)| \leq Mf(x)$ (we can refine this to a bound on cones)
2. (there are some more)

2.3.3 Calderon-Zygmund Decomposition

Theorem 2.49 (Calderon-Zygmund Decomposition). *For $f \in L^1(\mathbb{R}^d)$, $\alpha > 0$, there exists $b \in L^1(\mathbb{R}^d)$ such that $g : f - b \in L^\infty$ with $\|g\|_\infty \leq 2^d\alpha$ and $b = \sum b_j$ where b_j are supported on dyadic cubes Q_j with*

1. $\int b_j = 0$
2. $\|b_j\|_1 \leq 2^{d+1}\alpha|Q_j|$
3. $\sum |Q_j| \leq \alpha^{-1} \|f\|_1$

(we additionally have $\|g\|_1 \leq C_1 \|f\|_1$ and $\|b\|_1 \leq C_2 \|f\|_1$)

The idea is if we have a distribution of charges, f, then we may remove a threshold of charge density, g, and are left with dipoles

[a] I'm not that comfortable with this, but it is easy to replicate the proof for general x

Proof.

1. **Stopping time algorithm to select dyadic cubes**
   (a) We look for cubes Q such that $⨍_Q |f| > \alpha$, since $f \in L^1$, there is an upperbound for the size of the cubes this works for.
   (b) Start with the largest cubes, and select Q_j^1 such that $⨍_{Q_j^1} |f| > \alpha$ with index $j \in J_1$
   (c) Consider $\mathbb{R}^d$ with the chosen cubes removed, and tile the space with cubes of sidelength half the previous step. Select Q_j^2 such that $⨍_{Q_j^2} |f| > \alpha$ with index $j \in J_2$
   (d) repeat the previous step to get a collection of dyadic cubes: $\left\{Q_j^i : j \in J_i, i \in \mathbb{N}\right\}$
2. reindex the cubes as Q_j, $j \in \mathbb{N}$, let $b_j = 1_{Q_j} f - 1_{Q_j} ⨍_{Q_j} f$, we get the following properties of b_j
   (a) $\int b_j = 0$
   (b) $\|b_j\|_1 \leq 2 \int_{Q_j} |f|$. Let Q_j' be the smallest dyadic cube that contains Q_j, it wasn't selected, so $⨍_{Q_j'} |f| \leq \alpha$, so $\int_{Q_j'} |f| \leq \alpha |Q_j'| = \alpha |Q_j| 2^d$. Therefore $\|b_j\|_1 \leq 2 \int_{Q_j'} |f| \leq 2^{d+1} |Q_j| \alpha$
   (c) since $|Q_j| \leq \alpha^{-1} \int_{Q_j} |f|$, and Q_j are disjoint, we get $\sum |Q_j| \leq \alpha^{-1} \|f\|_1$
   (d) $\|b\|_1 \leq \sum 2^{d+1} \alpha |Q_j| \leq 2^{d+1} \|f\|_1$
3. properties of g:
   (a) let $E = \mathbb{R}^d \setminus \bigcup_j Q_j$. By the dyadic cube version of the Lebesgue differentiation theorem, for almost every $x \in E$, $|f| = \lim_{|Q| \to 0} ⨍_Q |f| \leq \alpha$, therefore $|f| \leq \alpha$, and since $g = f$ on E, the same is true on E for g.
   (b) for each $x \in E^c$, $g = f - b_j =$ for some j, so $g(x) = ⨍_{Q_j} f$, so $|g| \leq ⨍_{Q_j} |f| \leq \frac{|Q_j'|}{|Q_j|} \alpha = 2^d \alpha$
   (c) $\|g\|_1 \leq \|f\|_1 + \|b\|_1 \leq \|f\|_1 (1 + 2^{d+1})$

□

2.3.4 BMO Functions

Here we discuss a slight generalization of L^∞

Definition 2.17 (Bounded Mean Oscillation). *For $f \in L^1_{loc}(\mathbb{R}^n)$, let $\mathcal{B}$ be the collection of balls in $\mathbb{R}^n$, let $f_B = ⨍_B f$, for $B \subset \mathbb{R}^n$, let $\|f\|_{BMO} = \sup_{B \in \mathcal{B}} ⨍_B |f - f_B|$. We say say f is of bounded mean oscillation (BMO) if $\|f\|_{BMO} < \infty$.*

Remark 2.2. *(1) $\|1\|_{BMO} = 0$, so BMO is not a norm, (2) $\|f\|_{BMO} \leq 2 \|f\|_\infty$ so $L^\infty \subset BMO$, (3) $\|f(x/t)\|_{BMO} = \|f\|_{BMO}$ which is similar to L^∞, but not L^p.*

Theorem 2.50 (Equivalent BMO Norms). *An equivalent norm is taking the supremum of averages over dyadic cubes. Another is:*

$$\sup_{B \in \mathcal{B}} \inf_b ⨍_B |f - b| dx$$

Example 2.4 (**BMO is strictly larger than** L^∞)**.** $\log(|x|) \in BMO \setminus L^\infty$.

Theorem 2.51 (**John-Nirenberg Inequality**)**.** *There exists* $C_d, \delta > 0$ *such that for all* f *with* $0 \neq \|f\|_{BMO} < \infty$ *and all balls* B*:*

$$⨍_B \exp\left(\frac{\delta|f - f_B|}{\|f\|_{BMO}}\right) \leq C_d$$

Corollary 2.1. *If* $f \in BMO$*, then* $f \in L^p_{loc}$.[a]

Proof.

1. by dividing by $\|f\|_{BMO}$, we may assume $\|f\|_{BMO} = 1$, we will first prove the JNI for $B = Q_0$ for Q_0 a cube. By scaling and translating, we may assume $|Q_0| = 1$ is a dyadic cube with corner on the origin.
2. We now partition Q_0 via a stopping time construction. Note that $⨍_{Q_0} |f - f_{Q_0}| \leq 1$
   (a) Select Q_j^1, $j \in I_1$ dyadic cubes contained in Q_0 (of **any** size) such that $⨍_{Q_j^1} |f - f_{Q_0}| > 2$ for all j and $Q_j^1 \notin Q_i^1$ for any $i \neq j$.
   (b) Next select Q_j^2, $j \in I_2$ dyadic cubes contained in some Q_i^1 ($i \in I_1$) such that $⨍_{Q_j^2} |f - f_{Q_i^1}| > 2$. And continue selecting cubes this way.
3. Claim: $Q_0 = \bigcup_{n\geq 0}(\bigcup_i Q_i^n \setminus \bigcup_j Q_j^{n+1})$ modulo null sets.
   (a) if x isn't in this RHS set, then it is in some cube for **every** generation, that is $x \in \bigcap_n \bigcup_j Q_j^n$
   (b) for each n, $\sum_{Q_j^n \subset Q_i^{n+1}} |Q_j^n| \leq \sum \frac{1}{2} \int_{Q_j^n} |f - f_{Q_i^{n+1}}|$ (by selection rule), this is bounded by $\frac{|Q_i^{n+1}|}{2} ⨍_{Q_i^{n+1}} |f - f_{Q_i^{n+1}}| \leq (1/2)|Q_i^{n+1}| \, \|f\|_{BMO} \leq (1/2)|Q_i^{n+1}$. By induction $\sum_j |Q_j^n| \leq 2^{-n}$
   (c) we have nested sets whose measures are bounded by 2^{-n}, therefore the intersection is null.
4. Now our integral is $\sum_{n=0}^\infty \int_{\cup_i Q_i^n \setminus \cup_j Q_j^{n+1}} \exp(\delta|f - f_{Q_0}|)$. Claim: if $x \in Q_i^n \setminus \cup_j Q_j^{n+1}$, then $|f - f_{Q_0}| \leq (n+1)2^{d+1}$
   (a) (for $n = 3$) Suppose we have $Q_i^2 \subset Q_j^1 \subset Q_0$, then $|f(x) - f_{Q_0}| \leq |f(x) - f_{Q_i^2}| + |f_{Q_i^2} - f_{Q_j^1}| + |f_{Q_j^1} - f_{Q_0}|$.
   (b) by Lebesgue differentiation theorem, $|f(x) - f_{Q_i^2}| < 2$.
   (c) $|f_{Q_i^2} - f_{Q_j^1}| = |⨍_{Q_i^2} f - f_{Q_j^1}| \leq ⨍_{Q_i^2} |f - f_{Q_j^1}| \leq 2^d ⨍_{(Q_i^2)'} |f - f_{Q_j^1}| \leq 2^{d+1}$ (where prime denotes dyadic parent). So our original thing is $\leq 2 + 2^{d+1} + 2^{d+1} \leq 3(2^{d+1})$
5. then $\int_{\cup_i Q_i^n \setminus \cup_j Q_j^{n+1}} \exp(\delta|f - f_{Q_0}|) \leq \sum_i |Q_i|^n \exp(\delta(n+1)2^{d+1}) \leq 2^{-n} \exp(\delta(n+1)2^{d+1})$. This is summable over n if δ is small enough

□

[a] By translating and scaling can assume $f_B = 0$ and $\|f\|_{BMO} = 1$. Then $|f(x)|^p \leq e^{\delta|f(x)|}$ for $|f| \gg 0$. Split up the integral into a big and small part, the big part is integrable by above.

2.4 Singular Operators

Reference: Christ 5.1-5.5

Calderón-Zygmund theorem, homogeneous distributions, almost everywhere existence of principal-value integrals, almost everywhere differentiablity, singular integral operators on L^∞.

2.4.1 Calderón-Zygmund theorem for Convolution Operators

The goal is to prove L^p boundedness of convolution with a certain class of functions. These functions will generalize $K(x) = x_i x_j |x|^{-(d-2)}$. The Kernels that satisfy the conditions of the following theorem will be called **CZ kernels**. $K(x)$ as previously defined is a CZ kernel.

Theorem 2.52 **(Calderon-Zygmund Theorem for Convolution Operators).** *For $d \geq 1$, let $K : \mathbb{R}^d \to \mathbb{C}$ be s.t. (1) $|x|^{d+1}|\nabla K| \in L^\infty$ (2) $\widehat{K} \in L^\infty$ (3) $K \in C^1(\mathbb{R}^d \setminus \{0\})$*

Then for all $p \in (1, \infty)$, $f \in L^p \cap L^1$: $\|f * K\|_{L^p} \leq C_p \|f\|_{L^p}$

Lemma 2.3. *Let $Tf = K * f$, it suffices to show that $T : L^1 \to L^{1,\infty}$ is bounded.*

Proof. Given this, then:

1. if $f \in L^2 \cap L^1$, then $\|f * K\|_2 = C \left\|\widehat{f}\widehat{K}\right\|_2 \leq C \left\|\widehat{K}\right\|_\infty \|f\|_2$, so $T : L^2 \to L^2$ is bounded.
2. By Marcinkiewicz interpolation, T is bounded $L^p \to L^p$ for all $p \in (1, 2]$
3. for $p > 2$, $\|Tf\|_p = \sup_{\|g\|_q = 1} |\langle Tf, g\rangle|$ with $q = p'$. But $\langle Tf, g\rangle = \langle f, T^* g\rangle$ with $T^* g = \tilde{K} * g$ with $\tilde{K}(x) = \overline{K(-x)}$. This satisfies all properties, so $|\langle Tf, g\rangle| \leq C \|f\|_p \|g\|_q$

∎

Lemma 2.4. *$T : L^1 \to L^{1,\infty}$ is bounded.*

Proof. Idea: use CZ decomp, the bounded part is trivial. For dipoles, need to control terms away from support, this is done by exploiting dipole condition: away from a dipole, we don't really see anything We want $S := |\{|Tf| > \alpha\}| \leq C\alpha^{-1} \|f\|_1$. Fix $\alpha > 0$.

1. let $f = g + b$ by the CZ decomposition with parameter α, then $S \leq |\{|Tg| > \alpha/2\}| + |\{|Tb| > \alpha/2\}|$. The first term is bounded by $C\alpha^{-2} \|Tg\|_2^2 \leq C\alpha^{-2} \|g\|_2^2 \leq C\alpha^{-2} \|g\|_1 \|g\|_\infty \leq C\alpha^{-2}\alpha \|f\|_1$[a]
2. b_j are supported on Q_j, let Q_j^* be the double cube containing Q_j, then $\left|\left\{x \in \bigcup_j Q_j^* : |Tb| > \alpha/2\right\}\right| \leq \sum |Q_j^*| \leq C \sum |Q_j| \leq C\alpha^{-1} \|f\|_1$.
3. $\left|\left\{x \in \mathbb{R}^n \setminus \bigcup_j Q_j^* : |Tb| > \alpha/2\right\} \leq C\alpha^{-1} \sum_j \|Tb_j\|_{L^1(\mathbb{R}^n \setminus Q_j^*)}$

[a] using $T : L^2 \to L^2$ is bounded, $\|g\|_\infty \leq C\alpha$, $\|g\|_1 \leq C \|f\|_1$

4. for each j, for $x \notin Q_j^*$, $Tb_j(x) = \int_{Q_j} b_j(y)K(x-y)dy = \int_{Q_j}(K(x-y) - K(x-\bar{y}))b_j(y)dy$ (with $\bar{y}$ the center of Q_j^*). Then by MVT and hypothesis, $|Tb_j(x)| \le C\int_{Q_j}|y-\bar{y}||x-\bar{y}|^{-d-1}|b_j|dy \le C\ell(Q_j^*)\|b_j\|_1|x-\bar{y}|^{-d-1}$

5. integrate to get $\|Tb_j\|_{L^1(\mathbb{R}^n \setminus Q_j^*)} \le C\ell(Q_j^*)\|b_j\|_1 \int_{C\ell(Q_j)}^{\infty} r^{-d-1}r^{d-1}dr = C\|b_j\|_1$

6. with this, then $|\left\{x \in \mathbb{R}^n \setminus \bigcup_j Q_j^* : |Tb| > \alpha/2\right\}| \le C\alpha^{-1}\sum \|b_j\|_1 \le C\alpha^{-1}\|f\|_1$

□

2.4.2 Calderon-Zygmund Theorem

The following theorems are motivated by proving the following:

Theorem 2.53 (Second Derivative Controlled by Laplacian). $\left\|\partial^2_{x_j x_i} f\right\|_{L^p} \le C\|\Delta f\|_{L^p}$ *for all* $p \in (1, \infty)$.

1. by density it suffices to prove this for $f \in \mathcal{S}$. By the Fourier transform, can let $Tf = K * f$ with $K = \check{m}$ with $m(\xi) = \frac{\xi_i \xi_j}{|\xi|^2}$, it suffices to show that T is bounded $L^p \to L^p$[a]

2. To apply the Calderon-Zygmund theorem, it suffices to show that $\check{m}$ is homogeneous of degree zero and C^1 everywhere, except possibly the origin.

To prove (2), I will prove a more general result.

Theorem 2.54 (Fourier Transform Homogeneous Distribution Smooth Away From Origin). *If* $\varphi \in \mathcal{S}'(\mathbb{R}^d)$ *is homogeneous of degree* a*, and is in* $C^\infty(\mathbb{R}^d \setminus \{0\})$*, then* $\widehat{\varphi}$ *is in* $C^\infty(\mathbb{R}^d \setminus \{0\})$ *and is homogeneous of degree* $-a-d$.

Proof. 1. easy to see $\widehat{\varphi}$ is homogeneous of this degree. Therefore it suffices to show $\widehat{\varphi}(\xi) \in C^\infty$ near $|\xi| = 1$.

2. fix $k \in \mathbb{N}$, construct $\psi = \psi_k \in C_0^\infty(\mathbb{R}^n)$ radially symmetric supported in $B_1(0)$, with $\widehat{\psi}(\xi) \neq 0$ for $|\xi| = 1$ (and a moment condition that will be defined below). Since $\widehat{\psi} \in C^\infty$, it suffices to show that $\widehat{\psi}\widehat{\varphi} \in C^k$

3. for this, suffices to show $\langle x\rangle^k(\psi * \varphi) \in L^1$ (note $\psi * \varphi \in C^\infty$), for this suffices to understand for large x. Let $R \gg 0$ and consider $|x| = R$:

$$\begin{aligned}\psi * \varphi(x) &= \int_{y \in B_1(x)} \psi(x-y)\varphi(y)dy = \int_{v \in B_{R^{-1}(u)}} \psi(R(u-v))\varphi(Rv)R^d dv \\ &= R^a R^d \int_{|v-u| \le R^{-1}} \psi(R(u-v))\varphi(v)dv = R^a \int_{|y| \le 1} \psi(y)\varphi(u - \frac{y}{R})dx\end{aligned}$$

where $u = x/R$

4. Taylor expand φ about u to get $\varphi = P_N(u) + \mathcal{O}(R^{-N-1})$. The moment condition requires $\int \psi x^n = 0$ for all $n = 0, \ldots, N$, so we get $\psi * \varphi \le CR^a R^{-N-1}$.

[a] because $T\Delta f = K * \Delta f = \mathcal{F}^{-1}(m|\xi|^2\widehat{f}) = \mathcal{F}^{-1}(\xi_j \xi_i \widehat{f}) = \partial^2_{x_j x_i} f$

5. Therefore $\langle x\rangle^k \psi * \varphi = \mathcal{O}(|x|^{a-N-1-k})$, let N be large so that this exponent $< -d$

□

We bounded the norm of a convolution operator. If K is our convolution, and our operator is A, then we can write $A : C_0^\infty \to \mathcal{D}'$ by: $\langle Au, v\rangle = \iint u(x)v(y)K(x-y)dxdy$. We wish to generalize boundedness of A to more general A, defined in a similar way.

An operator $A : C_0^\infty \to \mathcal{D}'$ has **kernel** $K(x,y)$ if $\langle Au, v\rangle = \int u(x)v(y)K(x,y)dxdy$. Let's consider the following kernels:

Definition 2.18 (Calderon-Zygmund Kernel). *K is called a **CZ kernel** if (1) $K \in C^0(\mathbb{R}^d \times \mathbb{R}^d \smallsetminus \Delta) \to \mathbb{C}$ (2) $|K(x,y)| \le C|x-y|^{-d}$ (3) for all $|y-y'| \le \frac{1}{2}|x-y|$: $|K(x,y) - K(x,y')| + |K(y,x) - K(y',x)| \le C|y-y'|^\delta |x-y|^{-d-\delta}$ for some $\delta \in (0,1]$*

Although due to the singularities, we may only define A as above, if $\operatorname{supp} v \cap \operatorname{supp} u = \varnothing$

Example 2.5. *If $k \in C^\infty(\mathbb{R}^d \smallsetminus \{0\})$ is homogeneous of degree $-d$, then $K(x,y) = k(x-y)$ is a CZ kernel.*

Remark 2.3. *A sufficient condition for condition (3) is $\Delta_{x,y}K(x,y) = \mathcal{O}(|x-y|^{-d-1})$*

Theorem 2.55 (Calderon-Zygmund Theorem). *If A is an operator with associated CZ kernel K that is bounded on L^q for some q, then T extends to a bounded operator $L^p \to L^p$ for all $p \in (1,\infty)$*

Definition 2.19 (Principal Value). *If $k \in C^0(\mathbb{R}^d \smallsetminus \{0\})$ is homogeneous of degree $-d$, and $\int_{S^{d-1}} k(x)d\sigma(x) = 0$, then there exists a distribution $\operatorname{pv} k$ defined as:*

$$\langle \operatorname{pv} k, f\rangle = \lim_{\varepsilon \to 0^+} \int k(x)f(x)dx$$

This comes up naturally when looking at x^{-1} in $\mathbb{R}$. Note $x^{-1} \notin L^1_{loc}$, so it cannot naturally be considered a distribution.

Theorem 2.56 (Fourier Transform bijection of Homogeneous Degree Zero Distributions). *The fourier transform is a continuous bijection between homogeneous distributions of order zero that are smooth away from the origin and homogeneous distributions of degree $-d$ that can be written as $pv(k) + c\delta_0$ where $k \in C^\infty(\mathbb{R}^d \smallsetminus \{0\})$*

Proof.

1. let $\varphi \in \mathcal{D}'(\mathbb{R}^n) \cap C^\infty(\mathbb{R}^n \smallsetminus 0)$ homogeneous of degree 0. Write $\varphi = h + c$ with c a constant, and $\int_{S^{n-1}} h = 0$
2. then $\widehat{\varphi} = \widehat{h} + \widehat{c}$. $\widehat{c} = c(2\pi)^n \delta_0$ and $\widehat{h} \in \mathcal{D}'(\mathbb{R}^n) \cap C^\infty(\mathbb{R}^n)$ homogeneous of degree $-d$.
3. claim: $\int_{S^{n-1}} \widehat{h} = 0$
4. then $\widehat{h}$ agrees with $pv\widehat{h}$ everywhere except possibly the origin. The (distributional) difference of the two is homogeneous of degree $-d$, supported at the origin, and is therefore a delta distribution.

□

Example 2.6. *$|x|^{-d}$ can not be expressed as a homogeneous distribution (ie no homogeneous distribution agrees with $|x|^{-d}$ outside $\{x = 0\}$.*

2.4.3 Almost Everywhere Existence Of Convolution

Theorem 2.57 (Almost Everywhere Existence of Principal Value Integrals). *For $k \in \mathcal{S}' \cap C^\infty(\mathbb{R}^n \setminus \{0\})$ (homogeneous of degree $-d$, mean zero on spheres), $pv(k) * f(x)$ exists almost everywhere for $f \in L^p$ ($p \in [1, \infty)$)*

Things are a little confusing. So here is what we know:

1. We consider the operator $T : L^2 \to L^2$ by $Tf = k * f$. Note $\widehat{k} \in L^\infty$ by Theorem 2.54, so this is well defined.
2. We then showed that $\|Tf\|_{L^p} \le C \|f\|_p$ for all $p \in (1, \infty)$ and $f \in \mathcal{S}$
3. Therefore there is a unique bounded extension of $T : L^p \to L^p$, but it is now unclear how to compute Tf for arbitrary f (other than approximating f by Schwartz functions)
4. The theorem says that $Tf(x) = pv(k) * f(x)$ almost everywhere.

Proof. 1. decompose $f = g + b$ with $g \in C_0^\infty$ and $b \in L^p$ with $\|b\|_p < \delta$

2. then $pv(k) * f = \lim_{\varepsilon \to 0^+} \int_{|y|>\varepsilon} (g(x-y) + b(x-y))k(y)dy$. The first term in the integral converges by continuity of g
3. second term is bounded by $\sup_{\varepsilon>0} |\int_{|y|>\varepsilon} b(x-y)k(y)dy| := T^*h$
4. claim: T^* is bounded $L^p \to L^p$ for $p \in (1, \infty)$
5. with this then $|T^*h| \le \|T^*\|_{p \to p} \delta$

□

The proof boils down to proving $T^* : L^p \to L^p$ is bounded.

1. let $f \in C_0^1$, $k_\varepsilon(x) = k(x)1_{|x|>\varepsilon}$, then $f * k_\varepsilon = f * k_\varepsilon + f * k * \varphi_\varepsilon - f * k * \varphi_\varepsilon = (f * k) * \varphi_\varepsilon + f * (k_\varepsilon - k * \varphi_\varepsilon)$ where φ_ε is a usual approximation of identity
2. $(f * k) * \varphi_\varepsilon \le CM(f * k) \le CM(Tf)$
3. Claim: $(k_\varepsilon - k * \varphi_\varepsilon) \le C\varepsilon^{-d} \langle \varepsilon^{-1}|y| \rangle^{-d-1}$
   (a) expand integral
4. Claim: if $h \in L^1$ is radial, nonincreasing, nonnegative, then $g * h \le cMg$ for $g \in L^1_{loc}$
   (a) approximate h below by simple functions supported on annuli, apply dominated convergence theorem.
5. therefore $T^*(x) \le C(M(Tf)(x) + Mf(x))$, take L^p norms of both sides, use that M and T are bounded $L^p \to L^p$

2.4.4 Almost Everywhere Differentiability

The basic idea is that Lipschitz functions are differentiable almost everywhere. A more general theorem will be proven about $L_1^p(\mathbb{R}^d)$ with $1 > d/p$ (so we are in the Morrey's inequality domain).

Definition 2.20 ($L_1^p(\mathbb{R}^d)$). *For $f \in L^p$, the weak gradient is a distribution $\nabla f := g : \mathbb{R}^d \to \mathbb{R}^d$ such that for all $\varphi \in C_0^\infty(\mathbb{R}^d)$*

$$\int -f(\sum_1^d \partial_j \varphi)dx = \int g(x) \cdot \varphi(x)dx$$

We say $f \in L_1^p$ if $\nabla f \in L^p$

Theorem 2.58 (Almost Everywhere Differentiability of L_1^p). *If $f \in L_1^p$ with $p > d$, then f is almost everywhere differentiable.*

The steps of the proof are outlined here:

1. Pick $f \in L_1^p$ with compact support, then for almost every $x \in \mathbb{R}^d$:

$$f(x) = c_d \int \frac{x-y}{|x-y|^d} \cdot \nabla f(y)dy$$

2. claim: if $g \in L^p(\mathbb{R}^d; \mathbb{C}^d)$ and $f(x) = \int \frac{x-y}{|x-y|^d} \cdot g(y)dy$ with x satisfying: (1) $g(x) = 0$, (2) x is a Lebesgue point of g[a] (3) $pv\nabla_x(\frac{x}{|x|^{-d}}) * g(x) := Tg(x)$ exists, then:

$$f(x+h) - f(x) = h \cdot Tg(x) + o(|h|)$$

3. now let $f \in L_1^p$, then get the good x. And let $g(x') = \nabla f(x') - \nabla f(x)\eta(x')$ with $\eta \in C_0^\infty$ with $\eta(x) = 1$.

An alternate proof relies on the following:

Lemma 2.5 (**Morrey's Estimate**). *If $v \in W^{1,p}$ with $p > n$, then:*

$$|v(x) - v(y)| \le Cr^{1-\frac{n}{p}} \|Dv\|_{L^p(B(x,2r))}$$

Now let x be a Lebesgue point of Du, then let $v(y) = u(y) - u(x) - Du(x) \cdot (y-x)$, then we can easily get:

$$|u(x) - u(y) - Du(x) \cdot (x-y)| \le Cr \|Du - Dv\|_{L^p(B_{2r}(x))} = o(r)$$

2.4.5 Singular Integral Operators on L^∞

Theorem 2.59 (CZ Operators on L^∞). *If T is a CZ operator in $\mathbb{R}^d$, then $\|Tf\|_{BMO} \le C\|f\|_\infty$ for all $f \in L^\infty$ that vanish outside a bounded set.*

[a] is in the full measure set given by the Lebesgue differentiation theorem

Proof. The goal is to, for each ball B, to find b such that $\fint_B |f-b|dx$ is bounded (uniformly). It can be reduced to showing this for a ball centered at the origin, let that be $B(0,r)$

1. Let $f = f_0 + f_\infty$ with $f_0 = f 1_{B_{4r}(0)}$.
2. let $b = \int_{|y|>4r} K(0,y)f(y)dy$, then for $x \in B_r(0)$:

$$|Tf_\infty(x) - b| \le C \|f\|_\infty \int_{|y|\ge 4r} |x||y|^{-d-1}dy \le C \|f\|_\infty r \int_{4r}^\infty \rho^{-2}d\rho \le C \|f\|_\infty$$

3. Then:

$$\fint_B |Tf - b| \le \fint_B |Tf_\infty - b| + \fint_B |Tf_0|$$

it suffices to bound the second term:

$$\int_B |Tf_0|dx \le |B|^{1/2} \|Tf_0\|_{L^2(B)} \le C|B|^{1/2} \|f_0\|_{L^2(B)} \le \|f\|_\infty |B|C$$

□

3 Probability

3.1 Basic Notions

Reference: Durret 1.4 - 1.7

Measure theory, $\pi-\lambda$ theorem, random variables, inequalities, change of variables, notions of convergence of random variables

3.1.1 Measure Theory

Definition 3.1 (Measure Space). *A measure space is a triple $X, \mathcal{M}, m$. X is the space. $\mathcal{M}$ is the $\sigma-$algebra of measurable sets, and m is the measure.*

Definition 3.2 (σ-algebra). *A $\sigma-$algebra $\mathcal{M}$ on X is a collection of sets such that:*

1. $\varnothing, X \in \mathcal{M}$
2. *$\mathcal{M}$ is closed under countable unions and compliments*

that definition is redundant, basically any countable collection of set operations is allowed

Definition 3.3 (Non-negative real measure). *m is a non-negative real measure on X with $\sigma-$algebra $\mathcal{M}$ if $m : \mathcal{M} \to \mathbb{R}_{\geq 0}$ such that:*

1. $m(\varnothing) = 0$
2. *if $\{A_i\}_{i\in\mathbb{N}} \subset \mathcal{M}$ are disjoint, then $m(\bigcup A_i) = \sum m(A_i)$*

again, there are lots of ways to define this. Just think it generalizes $m([a,b]) = b - a$.

Definition 3.4 (Probability Space). *A probability space is a measure space $P, \mathcal{M}, m$ such that $m(P) = 1$*

Definition 3.5 (Random variable). *A random variable is a measurable function $X : P \to \mathbb{R}$. That is $X^{-1}(B) \in \mathcal{M}$ for all open sets Borel sets $B \in \mathbb{R}$.*

Note, it is the measurable with respect to Borel sets, not just open sets. Recall the Borel $\sigma-$algebra is the $\sigma-$algebra generated by open sets. So it contains closed sets and countable intersections of open sets.

Basically everything is measurable and therefore a random variable, but here are useful things:

Theorem 3.1 (Measurable on Generating Set). *If $\{X^{-1}(A)\}$ is measurable for all $A \in \mathcal{A}$ and $\mathcal{A}$ generates $\mathcal{M}$, then X is measurable with respect to $\mathcal{M}$*

Theorem 3.2 (Combinations of Random Variables). *Since compositions of measurable functions are measurable, combinations (not rigorous) of random variables are random variables (like $X_1 + X_2$).*

Definition 3.6 (Almost-sure convergence). *Random variables X_n converge to X almost surely, if $P(\{\omega : X_n(\omega) \to X\}) = 1$*

Definition 3.7 **(Convergence in Probability).** *X_n converges in probability to X if for all $\varepsilon > 0$, $P(|X_n - X| > \varepsilon) \to 0$*

Definition 3.8 **(Convergence in L^p).** *X_n converges in L^p to X if $\mathbb{E}[|X - X_n|^p] \to 0$*

Theorem 3.3 **(Relations of Notions of Convergence).**

1. $X_n \xrightarrow{a.s} X \Rightarrow X_n \xrightarrow{P} X$ *(use dominated convergence theorem)*
2. $X_n \xrightarrow{L^p} X \Rightarrow X_n \xrightarrow{P} X$ *(use Markov's inequality)*
3. $X_n \xrightarrow{P} X \Rightarrow X_{n_k} \xrightarrow{a.s.} X$

Proposition 3.1 **(Push-forward via random variable).** *A random variable X induces a probability measure via the pushforward of the probability measure: $\mu(A) = P(X \in A)$ for $A \in \mathcal{B}(\mathbb{R})$*

Definition 3.9 **(Cumulative Distribution Function (CDF)).** *The CDF of a random variable X is defined as $F(x) = P(X \in (-\infty, x])$*

I believe it is convention to be $(-\infty, x]$, so this needs to be memorized.[a]

Theorem 3.4 **(Characterizations of CDF).** *A function F satisfies (1) non-decreasing (2) $F(-\infty) = 0$, $F(\infty) = 1$ (3) is right continuous if and only if it is a CDF for a random variable X.*

Proof. ($\Rightarrow$) Let $\Omega = (0,1)$, $\mathcal{M}$ the Borel $\sigma-$algebra, m the Lebesgue measure. Then define $X(\omega) = \sup\{y : F(y) < \omega\}$. idea: if $F \in C^0$, then $X(\omega) = F^{-1}(\omega)$ works easily, otherwise consider X s.t. $P(X = -1) = P(X = 1) = 1/2$, think what F is, then reconstruct to get the correct inequality. □

3.1.2 Inequalities

Definition 3.10 **(Expected Value).** *For a random variable X, define $\mathbb{E}[X] = \int X dP$*

Theorem 3.5 **(Jensen's Inequality).** *If φ is convex, then $\varphi(\mathbb{E}[X]) \leq \mathbb{E}[\varphi(X)]$*

A way to remember this is that the absolute value is convex, and we know that $|\int f| \leq \int |f|$

Theorem 3.6 **(Markov's Inequality).** *If $X \geq 0$, then $cP(X \geq c) \leq \mathbb{E}[X]$*

Proof. $cP(X \geq c) = \int c1_{X \geq c} dP \leq \int X dP = \mathbb{E}[X]$ □

To remember, write $P(X \geq c)$ as an integral, and see that if $c = 1$, we can bound this integral by X, so we just need to scale.

Theorem 3.7 **(Chebyshev's Inequality).** *If $X \geq 0$ and $\varphi \geq 0$ is measurable, and $\iota_A = \min\{\varphi(y) : y \in A\}$ with A a Borel set, then $\iota_A P(X \in A) \leq \mathbb{E}[X 1_{X \in A}]$*

(same proof as Markov).

Corollary 3.1. $P(|X| \geq k) \leq k^{-2}\mathbb{E}[|X|^2]$

[a] mnemonic: Hungarians use $F(x) = P(X < x)$...maybe because they don't believe in equality (that's a joke, please don't attack me)

Theorem 3.8 (Fatou). $\mathbb{E}[\liminf X_n] \leq \liminf \mathbb{E}[X_n]$

Think $1_{[n,n+1]}$

Theorem 3.9 (Change of Variables). *If $f \in L^1$, then $\mathbb{E}[f(X)] = \int f(x)d\mu(x)$ where X has probability density μ.*

3.2 Law of Large Numbers

Reference: Durret 1.4-1.7

Independence, weak law of large numbers, Borel-Cantelli lemmas, strong law of large numbers, Kolmogorov 0–1 law, Kolmogorov maximal inequality

3.2.1 Independence

Definition 3.11 (Independence).

1. *(sets) $A, B \in \mathcal{R}$ are independent if $P(AB) = P(A)P(B)$.*
2. *(random variables) X and Y are independent if $P(X \in A, Y \in B) = P(X \in A)P(Y \in B)$ for all $A, B \in \mathcal{R}$.*
3. *(σ−algebras) $\mathcal{M}$ and $\mathcal{N}$ are independent if $P(AB) = P(A)P(B)$ for all $A \in \mathcal{M}$ and $B \in \mathcal{N}$. (Finite collections of these objects are also independent by a similar definition)*

Note that X and Y are independent if and only if $\sigma(X)$ and $\sigma(Y)$ are independent.[a]

Definition 3.12 (π-system). *A π−system is a collection of sets that is closed under finite intersection.*

Definition 3.13 (λ-system). *A λ−system is a collection of sets $\mathcal{L}$ of Ω such that (1) $\Omega \in \mathcal{L}$, (2) if $A \subset B$ are in $\mathcal{L}$, then $B \setminus A \in \mathcal{L}$, (3) if $A_i \in \mathcal{L}$ $A_i \subset A_{i+1}$, then $\bigcup_{n\in\mathbb{N}} A_n \in \mathcal{L}$*

Theorem 3.10 ($\pi - \lambda$ theorem). *If $\mathcal{P}$ is a π-system and $\mathcal{L}$ is a λ-system such that $\mathcal{P} \subset \mathcal{L}$, then $\sigma(P) \subset \mathcal{L}$*

Theorem 3.11 (Criterion for Independence). *$X_1, \ldots, X_n$ are independent random variables if and only if $P(X_1 \leq x_1, \ldots X_n \leq x_n) = \prod P(X_i \leq x_i)$ for all $x_i \in \mathbb{R}$*

Proof. (for X, Y random variables, backwards direction)

1. Let $\mathcal{A} = \{X \leq x : x \in \mathbb{R}\}$, $\mathcal{B} = \{Y \leq y : y \in \mathbb{R}\}$. These sets are closed under finite intersection, and are therefore π-systems.
2. Let $\mathcal{P} = \{A \in \mathcal{R} : P(AB) = P(A)P(B)\ \forall\ B \in \mathcal{B}\}$, this is a λ-system.
3. $\mathcal{A} \subset \mathcal{P}$, so by the π-λ theorem, $\sigma(\mathcal{A}) \subset \mathcal{P}$
4. But $\sigma(\mathcal{A}) = \sigma(X)$, therefore X and $\mathcal{B}$ are independent
5. Repeat this argument to get $\sigma(X)$ and $\sigma(Y)$ independent.

[a] To see this, note that $\sigma(X) = \{X^{-1}B : B \in \mathcal{R}\}$

□

This proof can be generalized to give:

Theorem 3.12 (Independent Collections of Sets). *If A_i is a finite collection of independent families that are closed under finite intersection. Then $\sigma(A_i)$ are independent.*

Corollary 3.2 (Independent Array of Sets). *If $\mathcal{F}_{i,j}$ are independent sets with $1 \le i \le n$ and $1 \le j \le m_i$, then $\mathcal{G}_i = \sigma(\bigcup_j \mathcal{F}_{ij})$ are independent.*

Proof. Let $M_i = \{\cap A_j : A_j \in \mathcal{F}_{i,j}\}$. Each M_i are a π−system and independent, therefore $\sigma(M_i)$ are independent. Now $\mathcal{G}_i \subset M_i$, so $\sigma(\mathcal{G}) \subset \sigma(M_i)$. And so these are independent. □

Theorem 3.13 (Expectation of Function of Two Independent Random Variables). *Let X, Y be two independent random variables with distribution μ and ν, and let $h \in L^1(\mathbb{R}^2)$, then $\mathbb{E}[h(X,Y)] = \iint h(x,y)d\mu(x)d\nu(y)$.*

Proof.

1. change of variables: $\mathbb{E}[h(X,Y)] = \int h(x,y)d\lambda$ with λ the unique measure on $\mathbb{R}^2$ that agrees with the induced measure of $X \times Y$ on rectangles.
2. $\lambda(A \times B) = P(X \in A, Y \in B) = P(X \in A)P(Y \in B) = \mu_X(A)\mu_Y(B)$.
3. use Fubini to get $\int h(x,y)d\lambda = \iint h(x,y)d\mu_X(x)d\mu_Y(y)$

□

3.2.2 Weak law of large numbers

Theorem 3.14 (L^2 LLN). *If X_i are uncorrelated with mean μ and variance uniformly bounded, then $n^{-1}S_n := \sum_1^n X_i \xrightarrow{L^2} \mu$*

Proof. $\mathbb{E}[n^{-1}S_n] = \frac{n\mu}{n} = \mu$, so $\mathbb{E}[|n^{-1}S_n - \mu|^2] = Var(n^{-1}S_n) = \frac{1}{n^2}\sum_1^n Var(X_i) \le \frac{C}{n} \to 0$ □

Theorem 3.15 (L^1 LLN). *If $X_i \in L^1$ are iid random variables, then $S_n \xrightarrow{L^1} \mathbb{E}[X]$*

Theorem 3.16 (Weak Law Of Large Numbers). *If X_i are i.i.d. L^1 random variables, then $\frac{1}{n}S_n \to \mathbb{E}[X_1]$ in probability.*

Proof. Let $\bar{X}_k = X_k 1_{|X_k| \le n}$, and $\mu_n = \mathbb{E}[\bar{X}_n]$. By dominated convergence theorem, $\mu_n \to \mu$. Let $\bar{S}_n = \sum_1^n \bar{X}_k$, it therefore suffices to show that $P(|n^{-1}S_n - \mu_n| > \varepsilon) \to 0$ for all ε.

1. $P(n^{-1}S_n - \mu_n| > \varepsilon) \le P(\bar{S}_n \ne S_n) + P(|n^{-1}\bar{S}_n - \mu_n| > \varepsilon)$
2. first term goes to zero: $P(\bar{S}_n \ne S_n) \le \sum_1^n P(|X_n| > n) = nP(|X_1| > n) \le \mathbb{E}[|X_1| 1_{|X_1| > n}] \to 0$ by DCT.
3. Second term bounded by Chebyshev:

$$\varepsilon^{-2}n^{-2}Var(\bar{S}_n) \le \varepsilon^{-2}n^{-2}\sum_1^n \mathbb{E}[\bar{X}_k^2] = \varepsilon^{-2}n^{-1}\int_0^n 2yP(|X_1| > y)dy$$

4. let $g(y) + 2yP(|X_1| > y)$, then $0 \le g \le 2y$ and goes to zero (by same argument as above). Then compute $\int_0^n g(y)dy$ via change of variables, it will go to zero by the dominated convergence theorem.

□

3.2.3 Borel-Cantelli Lemmas

Definition 3.14 (Liminf and Limsup of Events). $\limsup A_n = \lim_{j\to\infty} \bigcap_{n\ge j} \bigcup_{m\ge n} A_m$. $\liminf A_n = \lim_{j\to\infty} \bigcup_{n\ge m} \bigcap_{m\ge n}$.

Then $\omega \in \limsup A_n$ if ω is in infinitely many A_n (ω is in A_n i.o.). And $\omega \in \liminf A_n$ if ω is in all but finitely many A_n.

Theorem 3.17 (1st Borel-Cantelli Theorem). *If $\sum_n P(A_n) < \infty$, then $P(A_n \ i.o.) = 0$*

Proof. Let $N = \sum_n 1_{A_n}$. Then $\mathbb{E}[N] \le \sum_n P(A_n) < \infty$ (monotone convergence theorem). If $P(A_n \ i.o) > 0$, then $\mathbb{E}[N] \ge N(\omega : \omega \in A_n \ i.o)P(A_n \ i.o) = \cdot P(A_n \ i.o) = \infty$, □

Theorem 3.18 (L^p convergence implies subsequence converging a.s.). *If $X_n \xrightarrow{L^p} X$, then $X_{n_k} \to X$ almost surely.*

Proof.

1. $X_n \to X$ in probability. We therefore have a subsequence such that $P(|X_{n_k} - X| > 1/k) < 2^{-k}$.
2. Let $A_k = P(|X_{n_k} - X| > 1/k)$, so $P(A_k \ i.o) = 0$.
3. Relabel subsequence X_n. Let $B_k = \{|X_n - X| < 1/k$ for all but finite $n\}$. So $P(B_k) = 1$ for all k.
4. $P(X_n \to X) = P(\bigcap B) = 1$

□

The more general theorem is that is $P(|X_n - X| > \varepsilon)$ is summable for all ε, then $X_n \to X$ almost surely.

Theorem 3.19 (2nd Borel-Cantellil Theorem). *If A_n are independent non-summable events, then $P(A_n \ i.o) = 1$*

Proof. After playing around with limit continuity, it suffices to show that (for all n) $\lim_{m\to\infty} P(\bigcap_{k=n}^m A_k^c) = 0$. $P(\bigcap_n^k A_k) = \prod(1 - P(A_k)) = \exp(\sum \log(1 - P(A_k))) \le \exp(-\sum P(A_k)) \to 0$ (where we use $\log(1 - x) \le -x$) □

The trick is to use logarithm properties, and an inequality that I always forget how useful it is.

Theorem 3.20 (Distributional Formula for L^p norm). *If $Y \ge 0$ is a random variable and $p > 0$, then $\mathbb{E}[Y^p] = \int_0^\infty py^{p-1}P(Y > y)dy$*

Proof. Trick: write $Y^p = \int_0^Y py^{p-1}dy$

$$\mathbb{E}[Y^p] = \int_\Omega Y^p dP(\omega) = \int_\Omega \int_0^Y py^{p-1}dydP(\omega) = \int_\Omega \int_0^\infty 1_{Y>y}py^{p-1}dpdP(\omega)$$
$$= \int_0^\infty \int_\Omega py^{p-1}1_{Y>y}dP(\Omega)dp = \int_0^\infty py^{p-1}P(Y>y)dy$$

□

Theorem 3.21 (SSN for infinite mean iid random variables). *If X_i are iid random variables with $\mathbb{E}[|X_i|] = \infty$, then $\frac{1}{n}S_n$ almost surely doesn't converge to something finite.*

Proof. Really clever.

$$\sum_{n=0}^\infty P(|X_n| \geq n) = \sum_{n=0}^\infty \int_{x=n}^{n+1} P(|X_1| \geq n)dx \geq \sum_{n=0}^\infty \int_{x=n}^{n+1} P(|X_1| \geq x)dx$$
$$= \int_0^\infty P(|X_1| \geq x)dx = \mathbb{E}[|X|] = \infty$$

therefore by the 2nd Borel-Cantelli theorem, $P(|X_n| \geq n \ i.o.) = 1$, therefore the tails of S_n/n cannot converge. □

Theorem 3.22 (Strong Law of Large Numbers). *Let X_i be iid L^1 random variables, then $\frac{S_n}{n} \to \mathbb{E}[X_i]$ a.s.*

Proof.

1. Let $Y_n = X_n 1_{|X_n|\leq n}$ and $T_n = \sum_1^n Y_i$. By a Borel-Cantelli argument, it suffices to show $T_n/n \to \mu$ (T_n and S_n agree after finite terms).
2. let $k_n = \lfloor \alpha^n \rfloor$, it suffice to show that $T_{k_n}/k_n \to \mu$ for all $\alpha > 1$
   (a) letting n_m be such that $k_{n_m} \leq m \leq k_{n_m+1}$, $\frac{T_{k_m}}{k_{n_m+1}} \leq \frac{T_m}{m} \leq \frac{T_{k_{n_m}+1}}{k_{n_m}}$ implies $\frac{1}{\alpha}\mu \leq \frac{T_m}{m} \leq \mu\alpha$
3. want to show $\sum_{k=0}^\infty P(|T_{k_n} - \mathbb{E}[T_{k_n}]| > \varepsilon k_n) < \infty$ for all ε, this would imply $\frac{|T_{k_n} - \mathbb{E}[T_{k_n}]|}{k_n} \to 0$ almost surely, which gives $\frac{T_{k_n}}{n} \to \mu$
4. $P(|T_{k_n} - \mathbb{E}[T_{k_n}]| > \varepsilon k_n) \leq \frac{Var(T_{k_n}}{\varepsilon^2 k_n^2} = \varepsilon^{-2}k_n^{-2}\sum_{i=1}^{k_n} Var(Y_i)$
5. switch limits, want to show finiteness of $\sum_{m=1}^\infty Var(Y_m)\sum_{n:k_n\geq m} k_n^{-2}$.
6. for each m, $\sum_{n:k_n \geq m} k_n^{-2} \leq m^{-2}$
   (a) $\sum_{n:k_n\geq m} k_n^{-2} \leq c\sum_{n:n\geq \log_\alpha m} \alpha^{-2n} = c\frac{\alpha^{-2\log_\alpha m}}{1-\alpha^{-2}} = cm^{-2}$
7. $\sum_{k=1}^\infty Var(Y_k)k^{-2} < \infty$
   (a) $Var(Y_k) \leq \mathbb{E}[Y_k^2] = \int_0^\infty 2yP(|Y_k| > y)dy$, remember what X_k is: $\int_0^\infty 2yP(|X_k 1_{|X_k|\leq k}| > y)dy = \int_0^\infty 1_{y\leq k}(y)2yP(|X_k| > y)dy$
   (b) Fubini on sum is: $\int_0^\infty P(|X_1| > y)\sum_{k=1}^\infty 1_{y\leq k}(y)2yk^{-2}dy$

(c) fix y, $\sum_{k=1}^{\infty} 1_{y\le k} 2yk^{-2} \lesssim 2y\int_y^{\infty} k^{-2}dx = \frac{2y}{y} = 2$ details will give 4

(d) sum is bounded by $2\int_0^{\infty} P(|X_1| > y)dy = 2\mathbb{E}[|X_1|] < \infty$

□

Theorem 3.23 (Kolmogorov 0-1 Law). *If X_i are independent random variables, and $A \in \bigcap_{j\ge 1} \sigma(\bigcup_{m\ge j} X_m)$, then $P(A) \in \{0,1\}$*

We can interpret these events (called tail events) as events which don't depend on any finite collection of X_i.

Proof. (this is a good easy exercise in using the $\pi - \lambda$ theorem (or corollary).

1. $A \in \sigma(X_1, \dots X_k)$ and $B \in \sigma(X_{k+1}, X_{k+2}, \dots)$ are independent
   (a) If $B \in \sigma(X_{k+1}, \dots, X_{k+j})$, then apply Corollary 3.2 to see A and B are independent.
   (b) By (a), $\sigma(X_1, \dots, X_k)$ and $\bigcup_{j\ge k} \sigma(X_{k+1}, \dots X_{k+j})$ are independent, and are π-systems, therefore, $\sigma(\bigcup_{j\ge k} \sigma(X_{k+1}, \dots X_{k+j})) = \sigma(X_{k+1}, \dots)$ and $\sigma(X_1, \dots, X_k)$ are independent.
2. $A \in \sigma(X_1, \dots X_k)$ and $B \in \mathcal{T}$ are independent (because $B \in \sigma(X_{k+1}, \dots))$.
3. $\bigcup_k \sigma(X_1, \dots, X_k)$ and $\mathcal{T}$ are independent by (2) and are $\pi-$systems. Therefore the $\sigma-$algebras are. Since $\tau \subset \sigma(\cup_k \sigma(X_1, \dots, X_k))$, we are done.

□

Theorem 3.24 (Kolmogorov Maximal Inequality). *Given X_i independent mean zero random variables with finite variance, and $S_n = \sum_1^n X_i$, then $P(\max_{1\le k\le n} |S_k| > x) \le x^{-2} Var(S_n)$*

Proof. Idea: split the event into disjoint sets, bound the variance below by the second moment, split into integrals, use independence of partial sums to get rid on integral

1. let $A_k = P(|S_k| \ge x \; and \; |S_n| < x \; for \; n < k)$, these are disjoint over k
2. since $Var(S_n) = \mathbb{E}[S_n^2] - \mathbb{E}[S_n]^2 \ge \mathbb{E}[S_n^2]$:
$$Var(S_n) \ge \int S_n^2 dP \ge \sum_1^n \int_{A_k} S_n^2 dP = \sum_1^n \int_{A_k} (S_n - S_k + S_k)^2 dP$$
3. this gives three integrals. $(S_n - S_k)^2 \ge 0$, so we throw it away. For the other use indepedence: $(S_n - S_k) \perp S_k 1_{A_k}$ and $\mathbb{E}[S_n - S_k] = 0$.
4. The third integral is:
$$\sum_1^n \int S_k^2 1_{A_k} dP \ge \sum_1^n \int x^2 1_{A_k} dP = \sum_1^n x^2 P(A_k) = x^2 P(\max_{1\le k\le n} |S_k| > x)$$

□

3.3 Central Limit Theorem

Reference: Durret 2.2 - 2.4

Convergence in distribution, Helly's selection theorem, characteristic functions, Levy's continuity theorem, central limit theorem, Lindenberg-Feller Theorem

Definition 3.15 (Weak Convergence of Distributions). *Probability measure μ_n converge weakly to probability measure μ if $\int f d\mu_n \to \int f d\mu$ for all $f \in C^0 \cap L^\infty$ (this is written $\mu \Rightarrow \mu$).*

There are several equivalent statements of weak convergence in measure as outlined by the following:

Theorem 3.25 (Portmanteau Theorem). *If μ_n, μ are Borel probability measures then the following are equivalent:*

1. $\mu_n \Rightarrow \mu$
2. $\int f\mu_n \to \int f\mu$ *for all* $f \in Lip \cap L^\infty$
3. $\lim \mu_n(B) = \mu(B)$ *for all measurable* B *such that* $\mu(\partial B) = 0$
4. $\limsup \mu_n(F) \leq \mu(F)$ *for all closed measurable* F.

Definition 3.16 (Convergence in Distribution of Random Variables). *Random variables X_n converge in distribution to X if $\mathbb{E}[g(X_n)] \to \mathbb{E}[g(X)]$ for all $g \in L^\infty \cap C^0$*

Proposition 3.2 (Random Variable vs Measure weak convergence). *$X_n \Rightarrow X$ if and only if $\mu_n \Rightarrow \mu$.*

To prove this, the following helps:

Lemma 3.1 (measure vs CDF convergence). *If X_n, X are random variables with measure μ_n, μ and CDF F_n, F then $\mu_n \Rightarrow \mu$ if and only if $F_n(x) \to F(x)$ for all x where F is continuous.*

Proof. $\Rightarrow$ $F_n(x) = \mu_n((-\infty, x]) \to \mu((-\infty, x]) = F(x)$ because $\mu(\{x\}) = 0$.

$\Leftarrow$. Want to show if G is open, then $\liminf \mu_n(G) \geq \mu(G)$.

1. G is a disjoint union of open intervals $\bigcup(a_i, b_i)$
2. for each interval get x_i, y_i, continuity points of F such that $a_i \leq y_i \leq b_i$
3. $\liminf \mu_n(G) \geq \sum_1^m \mu_n(x_i, y_i) = \sum_1^m F_n(y_i) - F(x_i)$.
4. $n \to \infty$, get RHS as $\mu(\bigcup_1^m (x_i, y_i))$. Send x_i, y_i to a_i, b_i. Then send $m \to \infty$.

□

Weak convergence is the weakest form of convergence (almost sure convergence or convergence in probability implies weak convergence).

If $F_n \Rightarrow F$, then using the inversion formula for CDFs, we can construct Y_n, Y with the CDF F_n and F respectively such that $Y_n \xrightarrow{a.s.} Y$. This is very useful.

Theorem 3.26 (Continuous mapping theorem). *If $X_n \Rightarrow X$ and g is a measurable function such that $P(X \in D_g) = 0$ (where D_g is the set of discontinuities of g), then $g(X_n) \Rightarrow g(X)$*

Proof.

1. Let $Y_n \to Y$ a.s. have the same distribution as X_n, X
2. Let f be continuous and bounded, then $\mathbb{E}[f(g(X_n))] = \mathbb{E}[f(g(Y_n))]$ (b/c same distribution) $\to \mathbb{E}[f(g(Y))]$ (because it is continuous outside a measure zero set so we get almost sure convergence and by DCT we get this)
3. this is $\mathbb{E}[f(g(X))]$ because they are the same distribution.

□

Theorem 3.27 (Helly's Selection Theorem). *Given a sequence of cumulative distribution functions, there exists a subsequence the converges to a nondecreasing right-continuous function $F(x)$ at the continuity points of F. This convergence is called* ***vague convergence****.*

Proof. 1. Enumerate rationals, pick subsequences that converge on rational points to a sequence of nondecreasing values, call the result $F(x)$ (defined only for $x \in \mathbb{Q}$).

2. Define $F(x) = \inf\{F(q) : q > x \; q \in \mathbb{Q}\}$

□

Note that the result may not be a cumulative distribution function. Consider $F_n = 1_{x \geq n} \to 0$ or $F_n = 1_{x \leq -n} \to 1$.

Definition 3.17 (Tight CDFs). *CDFs F_i are tight if for all $\varepsilon > 0$, there exists $M = M(\varepsilon)$:*

$$\limsup_i (1 - F_i(M_\varepsilon) + F_i(-M_\varepsilon)) < \varepsilon$$

equivalently, the associated measure μ_n are such that for all $\varepsilon > 0$, there exists K_ε, compact, such that $1 - \mu_n(K_\varepsilon) < \varepsilon$ for all n

Theorem 3.28 (Vague Convergence to CDF). *CDFs F_n are tight if and only if every vague subsequential limit is a CDF.*

Proof. Easy proof. Both directions involve taking continuity points of F s and t, and know that a nondecreasing function taking values between 0 and 1, that is right-continuous is a cdf if and only if $\lim_{x\to\infty} 1 - F(x) + F(-x) \leq \varepsilon$ for all ε. □

Proposition 3.3 (Way to check tightness). *If X_n are random variables,, $\varphi \geq 0$ is goes to infinity as $|x| \to \infty$, and $\mathbb{E}[\varphi(X_n)] \leq C$, then the CDFs of X_n are tight.*

Proof. Apply Chebyshev: $P(|X_n| \geq M) \leq \frac{\mathbb{E}[\varphi(X_n)]}{\inf_{|y|>M} \varphi(y)} \to 0$ □

3.3.1 Characteristic Functions

Definition 3.18 (Characteristic Functions). *For a random variable X, define the characteristic function $\varphi(t) = \mathbb{E}[e^{itX}]$*

(This is the inverse Fourier transform of the probability density function)

Example 3.1 (Characteristic Function of Normal Distribution). *The standard normal distribution has density $f(x) = (2\pi)^{-1/2}e^{-x^2/2}$. The characteristic function is computed as:*

$$\varphi(t) = (2\pi)^{-1/2}\int_{-\infty}^{\infty} e^{-x^2/2+itx}dx = \frac{e^{-t^2/2}}{\sqrt{2\pi}}\int_{-\infty}^{\infty} e^{-(\frac{x}{\sqrt{2}}+\frac{it}{\sqrt{2}})}dx = e^{-t^2/2}$$

Theorem 3.29 (Characteristic Function Measure Inversion Formula). *If μ is a probability measure with characteristic function $\varphi(t)$, then for all $a < b$*

$$\frac{1}{2\pi}\lim_{T\to\infty}\int_{-T}^{T}\frac{1}{it}(e^{-ita}-e^{-itb})\varphi(t)dt = \mu(a,b) + \frac{1}{2}\mu(\{a,b\})$$

Proof. expand integral with Fubini, use Dirichlet Kernel knowledge

1. using Fubini, the LHS is (without constant and limit):

$$\int_{\mathbb{R}}\int_{-T}^{T}\frac{e^{it(x-a)}}{it} - \frac{e^{it(x-b)}}{it}dtd\mu(x)$$

2. since $\cos(t)/t$ is odd, each term in the integral is:

$$\int_{-\infty}^{\infty}\frac{\sin(t(x-a))}{t}dt = \pi\,\mathrm{sgn}(x-a)$$

(via contour integration)

3. then thinking about things, if $x \in (a,b)$, the integrand is 2π, if $x = a$ or $x = b$, it is π ($\mathrm{sgn}(0) = 0$), and otherwise it is zero. So we are done.

□

Here is some intuition: if X has a continuous PDF $f(x)$ with CDF $F(x)$, then $\varphi(t) = \widehat{f}(t) = \widehat{F'(t)} = it\widehat{F}$. Therefore $\int e^{-itx}\varphi(t)/(it)dt = \int e^{-itx}\widehat{F}(t)dt = F(x)$

Theorem 3.30 (Characteristic Function PDF Inversion Formula). *If $\varphi \in L^1$ is a characteristic function, then μ has bounded continuous density with pdf:*

$$f(x) = \frac{1}{2\pi}\int\varphi(t)e^{-itx}dt$$

Proof.

1. since $\varphi \in L^1$, then $|\int \frac{e^{-ita}-e^{-itb}}{it}\varphi(t)dt| \le (b-a)\|\varphi\|_1$, since $\frac{e^{-ita}-e^{-itb}}{it} = \int_a^b e^{it}dt$ (so the integral converges absolutely if $|b-a| < \infty$.

2. letting $a = b$ from the above formula, we see there cannot be singular points of μ.
3. $\mu(x, x+h) = \frac{1}{2\pi}\int_{-\infty}^{\infty}\int_{x}^{x+h} e^{-ity}\varphi(t)dydt = \int_{x}^{x+h}\frac{1}{2\pi}\int_{-\infty}^{\infty} e^{-ity}\varphi(t)dtdy$ (this is the definition of a pdf).

□

Theorem 3.31 (Levy's Continuity Theorem). *If μ_n are probability measures with characteristic functions φ_n then*

1. *if $\mu_n \Rightarrow \mu$ then $\varphi_n(t) \to \varphi(t)$ pointwise.*
2. *if $\varphi_n(t) \to \varphi(t)$ pointwise and φ is continuous at 0, then μ_n are tight and $\mu_n \Rightarrow \mu$*

Proof. The first statement is trivial: $\mathbb{E}[e^{itX_n}] \to \mathbb{E}[e^{itX}]$ by the dominated convergence theorem (where will let $X_n \to X$ pointwise having induced measures μ_n, μ.

For the second:

1. $\int_{-u}^{u} 1 - e^{itx}dt = 2(u - \frac{\sin(ux)}{x})$, therefore $u^{-1}\int_{-u}^{u}\int(1-e^{itx})\mu_n(dx) = 2\int(1 - \frac{\sin(ux)}{ux})\mu_n(dx)$
2. LHS is $u^{-1}\int_{-u}^{u}\int 1-\varphi_n(t)du$, RHS is bounded below by $2\int_{|x|>2u^{-1}}(1-\frac{1}{|xu|})\mu_n(x) \geq \mu_n(|x| > 2u^{-1})$
3. since $\varphi(0) = 1$ and φ is continuous at 0, $u^{-1}\int_{-u}^{u}(1-\varphi(t))dt \to 0$, since $\varphi_n(t) \to \varphi(t)$, then the LHS goes to zero as $u \to 0$ and $n \to \infty$. Therefore μ_n are tight.
4. get subsequence $\mu_{n_k} \Rightarrow \mu$ (by tightness) with characteristic function φ (by (1)). If μ_n didn't converge to μ, then for every subsequence it doesn't converge, but by the above, we can find a subsequence that does converge.

□

The second statement isn't straightforward. It suffices to show μ_n are tight. The key inequality to show is $\fint_{B_\varepsilon(0)}(1-\varphi_n(t))dt \geq \frac{1}{2}\mu_n(|x| > 2\varepsilon^{-1})$ (I wonder if there is a Harmonic analysis interpretation? Something involving distribution functions.).

Theorem 3.32 (Taylor Expansion of Characteristic Function). *If $X \in L^2$, then $\varphi_X(t) = 1 + it\mathbb{E}[X] - \frac{t^2}{2}\mathbb{E}[X^2] + o(t^2)$.*

Proof.

1. use calculus: $e^{itx} = 1 + itx - it^2x^2/2 + R(t,x)$ with $|R(t,x)| \leq \min(\frac{t^3x^3}{3!}, t^2x^2)$
2. so $\varphi(t) = \mathbb{E}[e^{itX}] = \int e^{itx}\mu(dx) = 1 + it\mathbb{E}[X] - it^2\mathbb{E}[X^2]/2 + \int R(t,x)\mu(dx)$
3. $\lim_{t\to 0} t^{-2}\int R(t,x) = 0$ because it is dominated by $x^2 \in L^1$ and converges pointwise because $tx^3/3 \to 0$ almost everywhere.

□

Theorem 3.33 (Central Limit Theorem). *If X_i are iid with $\mathbb{E}[X] = \mu$, $Var(X) = \sigma^2 \in (0, \infty)$, and $S_n = \sum_1^n X_i$, then $\frac{S_n - n\mu}{\sigma\sqrt{n}} \Rightarrow N(0,1)$*

Proof.

1. WLOG, by shifting $\mu = 0$
2. the characteristic function is: $\mathbb{E}[\exp(it\frac{S_n}{\sigma\sqrt{n}})] = (\mathbb{E}[\frac{itX_1}{\sigma\sqrt{n}}])^n = \varphi_X(\frac{t}{\sigma\sqrt{n}})^n$
3. by Theorem 3.32, this is $(1 + 0 - \frac{t^2\mathbb{E}[X^2]}{2\sigma^2 n} + o(n^{-1}))^n = (1 + t^2/(2n) + o(n^{-1}))^n$
4. this converges, for each fixed t, to $e^{-t^2/2}$ as $n \to \infty$ [a].
5. by Levy's Continuity Theorem (3.31) the original random variable converges in distribution to the standard normal.

□

Theorem 3.34 (Lindeberg-Feller Theorem). *Let $X_{n,m}$ ($m = 1, \ldots, n$, $n = 1, \ldots, \infty$) be a triangular array of mean zero independent random variables with variance $\sigma^2_{n,m}$, such that:*

1. $\lim_{n\to\infty} \sum_{m=1}^n \sigma^2_{n,m} = \sigma^2$
2. $\lim_{n\to\infty} \sum_{m=1}^n \mathbb{E}[|X_{n,m}|^2 1_{|X_{n,m}|\geq\varepsilon}] = 0$ *for all* $\varepsilon > 0$

then $S_n = \sum_{m=1}^n X_{n,m} \Rightarrow N(0, \sigma^2)$

Proof.

1. set $\varphi_{n,m}(t) = \mathbb{E}[e^{itX_{n,m}}]$, so $\varphi_{S_n}(t) = \prod_{m=1}^n \varphi_{n,m}(t)$, we need to show for each fixed t, $|\prod_1^n \varphi_{n,m}(t) - e^{-t^2\sigma^2/2}| \to 0$
2. Claim: $e^{-t^2\sigma^2/2} = \lim_{n\to\infty} \prod_{m=1}^n (1 - \frac{t^2\sigma^2_{n,m}}{2})$
   (a) take log, taylor expand: $\sum_{m=1}^n \log(1 - \frac{t^2\sigma^2_{n,m}}{2}) = \sum_{m=1}^n \frac{-t^2\sigma^2_{n,m}}{2} + \mathcal{O}(\sigma^4_{n,m})$
   (b) the sum is $-t^2\sigma^2/2$, the remainder goes to zero: $\sigma^2_{n,m} = \mathbb{E}[X^2_{n,m} 1_{|X_{n,m|\leq\varepsilon}}] + \mathbb{E}[X^2_{n,m} 1_{|X_{n,m}|>\varepsilon}] \leq \varepsilon^2 + \mathbb{E}[X^2_{n,m} 1_{|X_{n,m}|>\varepsilon}]$. The sum of the second term over m goes to zero, so all those terms must go to zero, therefore $\sup_m \sigma^2_{n,m} \to 0$ as $n \to \infty$
3. Claim: $|\prod_1^n \varphi_{n,m}(t) - \prod_{m=1}^n (1 - \frac{t^2\sigma^2_{n,m}}{2})| \leq \sum_1^n |\varphi_{n,m}(t) - (1 - \frac{t^2\sigma^2_{n,m}}{2})|$
   (a) for $n \gg 1$, the terms in the products are bounded in modulus by 1, then induct using the algebra:
   $$|z_1 z_2 - w_1 w_2| = |z_1 z_2 - z_1 w_2 + z_1 w_2 - w_1 w_2| \leq |z_1||z_2 - w_2| + |w_2||z_1 - w_1| \leq |z_2 - w_2| + |z_1 - w_1|$$
4. Claim $\sum_1^n |\varphi_{n,m}(t) - (1 - \frac{t^2\sigma^2_{n,m}}{2})| \xrightarrow{n\to\infty} 0$

[a] nontrivially since the remainder term has complex values

(a) $|\varphi_{n,m}(t) - (1 - \frac{t^2\sigma_{n,m}^2}{2}) \le c\mathbb{E}[|t^3 X_{n,m}|^3 \wedge t^2|X_{n,m}|^2] \le c\mathbb{E}[|t^3 X_{n,m}|^3 1_{|X_{n,m}| \le \varepsilon}] + c\mathbb{E}[|t^2 X_{n,m}|^2 1_{|X_{n,m}| > \varepsilon}]$

(b) the first term is bounded by $c|t|^3\varepsilon\mathbb{E}[|X_{n,m}|^2 1_{|X_{n,m}| \le \ e}] \le \sigma_{n,m}^2 c|t|^3\varepsilon$

(c) taking the sum to infinity (noting the second term goes to zero) gives: $\varepsilon|t|^3 c\sigma^2$, $\varepsilon \to 0$ gives result

5. finally, let $A_n(t) = \prod_{m=1}^n (1 - \frac{t^2\sigma_{n,m}^2}{2})$. From (3) and (4), $|\varphi_{S_n}(t) - A_n(t)| \to 0$ for all t. By (2) $A_n(t) \to e^{-t^2\sigma^2/2}$, so by the triangle inequality $|S_n - e^{-t^2\sigma^2/2}) \to 0$

□

4 List of Theorems

5 List of Definitions