

# **II Analysis of Functions**

Ishan Nath, Lent 2024

Based on Lectures by Prof. Richard Nickl

March 11, 2024

## Contents

<b>0</b>	<b>Review</b>	<b>2</b>
<b>1</b>	<b>Actual Stuff</b>	<b>4</b>
1.1	$L^p$ spaces and their approximations . . . . .	4
1.2	Lebesgue's Differentiation Theorem . . . . .	8
<b>2</b>	<b><math>L^2</math></b>	<b>13</b>
2.1	Dual Spaces . . . . .	15
2.2	Weak and Weak-* Topologies . . . . .	18
2.3	Compact Subsets of Infinite-Dimensional Spaces . . . . .	20
2.4	Hahn-Banach Theorem . . . . .	23
<b>3</b>	<b>Generalised Functions and Distributions</b>	<b>28</b>
3.1	Operations on Distributions . . . . .	31
3.2	Convolutions of Distributions . . . . .	33
3.3	Fundamental Theorem of PDEs . . . . .	35
<b>4</b>	<b>Fourier Transforms of Distributions</b>	<b>37</b>
4.1	Periodic Distributions . . . . .	43
<b>5</b>	<b>Sobolev Spaces and Elliptic PDE's</b>	<b>50</b>
5.1	Weird Space . . . . .	53
5.2	Variational Problems and PDEs . . . . .	60
	<b>Index</b>	<b>63</b>

## 0 Review

We go over some basic concepts.

1. A measure space is a tuple  $(E, \mathcal{E}, \mu)$ , where  $E$  is any set,  $\mathcal{E}$  is a  $\sigma$ -algebra over  $E$ , and  $\mu$  is a measure from  $\mathcal{E} \rightarrow [0, \infty]$ .
2. For  $\mathcal{A}$  a collection of subsets of  $E$ , we let  $\sigma(\mathcal{A})$  be the smallest  $\sigma$ -algebra containing  $\mathcal{A}$ . For  $E$  a topological space, let  $\mathcal{B}(E) = \sigma(\mathcal{A})$ , where

$$\mathcal{A} = \{A \subseteq E \mid A \text{ open}\}.$$

3. On  $E = \mathbb{R}^n$  with the product topology, we set  $\mathcal{B} = \mathcal{B}(\mathbb{R}^n)$ . On  $\mathcal{B}$ , there exists Lebesgue measure  $\mu$  such that

$$\mu((a_1, b_1) \times \cdots \times (a_n, b_n)) = (b_1 - a_1) \cdots (b_n - a_n),$$

for  $a_i < b_i$ . We often write  $|A| = \mu(A)$  for  $A \in \mathcal{B}$ . The Lebesgue measure satisfies the following properties:

- (i)  $\mu$  is translation invariant, so  $|A + z| = |A|$  for all  $A \in \mathcal{B}$  and  $z \in \mathbb{R}^n$ .
- (ii)  $\mu$  is ‘inner-regular’:

$$|A| = \sup\{|K| \mid K \in A, K \text{ compact}\}.$$

$\mu$  actually extends to the  $\mu$ -completion  $\mathcal{M}$  on  $\mathcal{B}$ , which also includes subsets of Borel  $\mu$ -null sets.  $\mathcal{M}$  are the Lebesgue-measurable sets.

4. For  $(E, \mathcal{E})$  and  $(F, \mathcal{F})$  measurable spaces, we say  $f : E \rightarrow F$  is measurable if  $f^{-1}(A) \in \mathcal{E}$  for all  $A \in \mathcal{F}$ . For  $F = \mathbb{R}$  or  $\mathbb{C}$ , we always take  $\mathcal{F} = \mathcal{B}$ , and say that  $f$  is Borel measurable. We say that  $f : E \rightarrow \mathbb{R}$  is non-negative if it takes values in  $[0, \infty]$ , and write  $f \geq 0$ .
5. A simple function  $f : (E, \mathcal{E}, \mu) \rightarrow \mathbb{R}$ ,  $f \geq 0$  is of the form

$$f = \sum_{k=1}^n a_k \mathbb{1}_{A_k}, \quad a_k \geq 0, A_k \in \mathcal{E}, N \in \mathbb{N}.$$

The  $\mu$ -integral  $\mu(f)$  of a simple function is

$$\mu(f) = \sum_{k=1}^n a_k \mu(A_k).$$

For  $f : (E, \mathcal{E}, \mu) \rightarrow \mathbb{R}$ ,  $f \geq 0$ , we define

$$\mu(f) = \sup\{\mu(g) \mid g \leq f, g \text{ simple}\},$$

which may be  $\infty$ . A map  $f : (E, \mathcal{E}, \mu) \rightarrow \mathbb{R}$  or  $\mathbb{C}$  is called integrable if  $\mu(|f|) < \infty$ . In this case, we can write  $\mu(f) = \mu(f^+) - \mu(f^-)$ , and write

$$\mu(f) = \int_E f \, d\mu = \int_E f(x) \, d\mu(x).$$

**Theorem 0.1** (Monotone Convergence Theorem). *If  $f_n \geq 0$  measurable,  $f_n \uparrow f$  pointwise on  $E$ , then  $\mu(f_n) \uparrow \mu(f)$  as  $n \rightarrow \infty$ .*

**Theorem 0.2** (Dominated Convergence Theorem). *If  $f_n$  measurable,  $f_n \rightarrow f$  pointwise on  $E$  and  $|f_n| \leq g$  for some  $\mu$ -integrable  $g$ , then  $f$  is also integrable, and  $\mu(f_n) \rightarrow \mu(f)$  as  $n \rightarrow \infty$ .*

# 1 Actual Stuff

## 1.1 $L^p$ spaces and their approximations

**Definition 1.1.** For  $f : (E, \mathcal{E}, \mu) \rightarrow \mathbb{R}$  or  $\mathbb{C}$  measurable and  $1 \leq p < \infty$ , set

$$\|f\|_{L^p} = \left( \int_E |f(x)|^p d\mu(x) \right)^{1/p},$$

and

$$\|f\|_{L^\infty} = \mu \sup |f| = \inf \{ \lambda > 0 \mid |f| \leq \lambda \text{ } \mu\text{-almost everywhere} \}.$$

We reserve  $\|\cdot\|_\infty$  for the usual sup-norm. Then,

$$L^p = L^p(E, \mathcal{E}, \mu) = \{f : E \rightarrow \mathbb{R} \text{ measurable, } \|f\|_{L^p} < \infty\}.$$

For  $E = \mathbb{R}^n$ ,  $\mathcal{E} = \mathcal{B}$  and  $\mu = dx$ , we let  $L^p(\mathbb{R}^n) = L^p$ .

**Theorem 1.1** (Riesz-Fischer). *Identifying  $f \sim g$  if  $f = g$  almost-everywhere, the spaces  $L^p$ ,  $1 \leq p \leq \infty$  are Banach spaces.*

Recall the spaces  $C^k(\mathbb{R}^n)$  of functions with continuous partial derivatives up to order  $k$ . Further, set

$$C^\infty(\mathbb{R}^n) = \bigcap_{k \in \mathbb{N}} C^k(\mathbb{R}^n),$$

the smooth or infinitely differentiable functions. Further, write  $C_c^\infty(\mathbb{R}^n)$  and  $C_c^k(\mathbb{R}^n)$  for the subspaces of functions of compact support in  $\mathbb{R}^n$ . For  $k = 0$ , write  $C = C^0$ ,  $C_c = C_c^0$ .

*Remark.*  $C_c^\infty(\mathbb{R}^n)$  is non-empty, since it contains

$$\psi(x) = \begin{cases} e^{(|x|^2-1)^{-1}} & |x| < 1, \\ 0 & \text{else.} \end{cases}$$

**Theorem 1.2.** *For  $1 \leq p < \infty$ ,  $C_c^\infty(\mathbb{R}^n)$  is dense in  $L^p(\mathbb{R}^n)$ .*

Note this is false for  $p = \infty$ . This follows from the following:

**Proposition 1.1.** *For  $1 \leq p < \infty$ ,  $C_c(\mathbb{R}^n)$  is dense in  $L^p(\mathbb{R}^n)$ .*

The idea of the proof is to restrict to  $f \in L^p$  of compact support, since

$$\int_{\mathbb{R}^n} |f(x)|^p \mathbb{1}_{|x| \geq R} dx \rightarrow 0$$

as  $R \rightarrow \infty$ , by DCT. Then the rest is an example sheet problem.

**Definition 1.2.** For  $f, g : \mathbb{R}^n \rightarrow \mathbb{C}$  or  $\mathbb{R}$ , we define the *convolution*

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

whenever the integral exists.

**Proposition 1.2.** For  $f, g, h \in C_c(\mathbb{R}^n)$ , we have

- (i)  $f * g = g * f$ .
- (ii)  $(f * g) * h = f * (g * h)$ .
- (iii) We have

$$\int_{\mathbb{R}^n} f * g(x) \, dx = \int_{\mathbb{R}^n} f(x) \, dx \cdot \int_{\mathbb{R}^n} g(x) \, dx.$$

The proof is easy.

For  $\alpha = (\alpha_1, \dots, \alpha_n) \in \mathbb{Z}_+^n$  a multi-index, define  $|\alpha| = \alpha_1 + \dots + \alpha_n$ ,  $\alpha! = \alpha_1! \dots \alpha_n!$ , and for  $x = (x_1, \dots, x_n) \in \mathbb{R}^n$ , we set

$$x^\alpha = x_1^{\alpha_1} x_2^{\alpha_2} \dots x_n^{\alpha_n}.$$

Then we can define

$$D^\alpha f = \frac{\partial^{|\alpha|}}{\partial x^\alpha} = \frac{\partial^{|\alpha|}}{\partial x_1^{\alpha_1} \dots \partial x_n^{\alpha_n}}.$$

Let us also say that  $f$  measurable is *locally* in  $L^p(\mathbb{R}^n)$ , and write  $f \in L_{\text{loc}}^p(\mathbb{R}^n)$ , if  $f \mathbb{1}_K \in L^p(\mathbb{R}^n)$  for all compact sets  $K \in \mathbb{R}^n$ .

**Theorem 1.3.** Let  $f \in L_{\text{loc}}^1(\mathbb{R}^n)$ ,  $g \in C_c^k(\mathbb{R}^n)$ . Then  $f * g \in C^k(\mathbb{R}^n)$ , and for  $\alpha \leq 1$ ,  $|\alpha| \leq k$ , we have

$$D^\alpha(f * g) = f * D^\alpha g.$$

**Proof:** For  $\tau_u g = g \circ (-u)$  the translation operator for  $u \in \mathbb{R}^n$ , notice that

$$|\tau_u g(z)| = |g(z - u)| \leq \|g\|_\infty \mathbb{1}_{|z| \leq R},$$

for some  $R > 0$  and  $u \leq 1$ . Thus,

$$\begin{aligned} \tau_u(f * g)(x) &= f * \tau_u g(x) = \int_{\mathbb{R}^n} g(x - u - y)f(y) \, dy \\ &\rightarrow \int_{\mathbb{R}^n} g(x - y)f(y) \, dy = f * g(x), \end{aligned}$$

by DCT, since this function is at most  $\|g\|_\infty \mathbb{1}_K f$ . Since  $x$  was arbitrary,  $f * g \in C^0(\mathbb{R}^n)$ .

For  $k = 1$ , we introduce the difference quotients

$$\Delta_h^i g(x) = \frac{g(x + he_i) - g(x)}{h},$$

for  $h \in \mathbb{R} \setminus \{0\}$ , and  $e_i$  the  $i$ 'th basis vector. Then as  $h \rightarrow 0$ , this approaches  $D_i g(x)$ , since  $g \in C_c^1(\mathbb{R}^n)$ .

Also, by MVT,

$$|\Delta_h^i g(x)| = |D_i g(x + t_h e_i)|,$$

where  $t_h \in (-|h|, |h|)$ . Now since  $g \in C_c^1$ ,  $D_i g$  also has compact support and hence is bounded, so it is less than  $\|D_i g\|_\infty \mathbb{1}_{\{|x| \leq R\}}$  for some  $R > 0$ ,  $|h| \leq 1$ .

Therefore, by DCT as earlier,

$$\Delta_h^i (f * g) = f * (\Delta_h^i g) \rightarrow f * D_i g.$$

As the limit is continuous by the  $k = 0$  case, we see that  $f * g$  has all partial derivatives continuous on  $\mathbb{R}^n$ , the function is  $C^1$ . The rest is induction on  $k$ .

**Proposition 1.3.** *Let  $1 \leq p < \infty$ , and  $\tau_z g = g \circ (-z)$ . Then for all  $f \in L^p(\mathbb{R}^n)$ ,*

$$\|\tau_z f - f\|_{L^p} \rightarrow 0 \text{ as } z \rightarrow 0.$$

**Proof:** Let  $g \in C_c(\mathbb{R}^n)$ , then

$$\|\tau_z g - g\|_{L^p}^p = \int_{\mathbb{R}^n} |g(x - z) - g(x)|^p dx \rightarrow 0$$

as  $z \rightarrow 0$ , by DCT, since the thing in the middle is at most  $2^p \|g\|_\infty^p \mathbb{1}_{|x| \leq R}$ .

Next, let  $f \in L^p(\mathbb{R}^n)$ ,  $\varepsilon > 0$  be arbitrary. Since  $C_c(\mathbb{R}^n)$  is dense in  $L^p$ , we can find  $g \in C_c(\mathbb{R}^n)$  such that  $\|f - g\|_{L^p} < \varepsilon/3$ . Then by translation invariance of  $dx$ ,

$$\|\tau_z f - \tau_z g\|_{L^p} = \|\tau_z(f - g)\|_{L^p} = \int_{\mathbb{R}^n} |f(x - z) - g(x - z)|^p dx = \|f - g\|_{L^p} < \frac{\varepsilon}{3}.$$

Thus by the triangle inequality,

$$\|\tau_z f - f\|_{L^p} \leq \|\tau_z g - g\|_{L^p} + \|f - g\|_{L^p} + \|\tau_z f - \tau_z g\|_{L^p} = \varepsilon.$$

**Proposition 1.4** (Minkowski's Inequality). *Let  $F : \mathbb{R}^n \times \mathbb{R}^n \rightarrow \mathbb{C}$  or  $\mathbb{R}$  be*

measurable. Then for all  $1 \leq p < \infty$ ,

$$\left\| \int_{\mathbb{R}^n} F(x, \cdot) \, dx \right\|_{L^p} \leq \int_{\mathbb{R}^n} \|F(x, \cdot)\| \, dx.$$

This is on the example sheet.

**Theorem 1.4** (Approximation Identities or Mollification). *Let  $\phi \in C_c^\infty$ ,  $\phi \geq 0$  with*

$$\int_{\mathbb{R}^n} \phi(x) \, dx = 1,$$

*and set*

$$\phi_\varepsilon = \frac{1}{\varepsilon^n} \phi\left(\frac{\cdot}{\varepsilon}\right).$$

*Then for  $f \in L^p(\mathbb{R}^n)$ , the functions  $\phi_\varepsilon * f$ ,  $\varepsilon > 0$  lie in  $C^\infty(\mathbb{R}^n)$ , and*

$$\phi_\varepsilon * f \rightarrow f$$

*in  $L^p$ , and almost surely as  $\varepsilon \rightarrow 0$ .*

**Proof:** We have

$$\begin{aligned} |\phi_\varepsilon * f(x) - f(x)| &= \left| \int_{\mathbb{R}^n} \phi\left(\frac{y}{\varepsilon}\right) f(x - y) \, dy - f(x) \right| \\ &= \left| \int_{\mathbb{R}^n} \phi(u) (f(x - \varepsilon u) - f(x)) \, du \right| \\ &\leq \int_{\mathbb{R}^n} \phi(u) |\tau_{\varepsilon u} f(x) - f(x)| \, dx. \end{aligned}$$

Therefore, by Minkowski's inequality,

$$\|\phi_\varepsilon * f - f\|_{L^p} \leq \int_{\mathbb{R}^n} \phi(u) \|\tau_{\varepsilon u} f - f\|_{L^p} \, du \rightarrow 0,$$

by DCT, as  $\varepsilon \rightarrow 0$ . Note  $\|\tau_{\varepsilon u} f - f\|_{L^p} \rightarrow 0$  as  $\varepsilon \rightarrow 0$  by continuity of translation in  $L^p$ .

Since  $C_c(\mathbb{R}^n)$  is dense in  $L^p$ , for  $1 \leq p < \infty$ , and since  $C_c(\mathbb{R}^n) * \phi_\varepsilon = \{\phi_\varepsilon * f \mid f \in C_c(\mathbb{R}^n)\} \subseteq C_c^\infty(\mathbb{R}^n)$ , we have shown that  $C_c^\infty(\mathbb{R}^n)$  is dense in  $L^p$ .



## 1.2 Lebesgue's Differentiation Theorem

Recall that the fundamental theorem of calculus implies that, for  $f : \mathbb{R} \rightarrow \mathbb{R}$  continuous,

$$\lim_{h \rightarrow 0} \frac{1}{h} \int_x^{x+h} f(y) dy = f(x) \quad \forall x \in \mathbb{R}.$$

**Theorem 1.5.** *Let  $f \in L^1(\mathbb{R}^n)$ . Then for almost all  $x \in \mathbb{R}^n$ ,*

$$\lim_{r \rightarrow 0} \frac{1}{|B_r(x)|} \int_{B_r(x)} |f(y) - f(x)| dy = 0.$$

Our proof will be based on the following.

**Definition 1.3** (Hardy-Littlewood Maximal Function). Let  $f \in L^1(\mathbb{R}^n)$ , and set

$$Mf(x) = \sup_{r>0} \frac{1}{|B_r(x)|} \int_{B_r(x)} |f(y)| dy, \quad \forall x \in \mathbb{R}^n.$$

**Lemma 1.1.** *Let  $f \in L^1(\mathbb{R}^n)$ . Then  $Mf : \mathbb{R}^n \rightarrow \mathbb{R}$  is measurable,  $Mf < \infty$  almost everywhere, and for all  $\lambda > 0$ , there exists  $C_n > 0$  (depending only on  $n$ ) such that*

$$|\{x \mid Mf(x) > \lambda\}| \leq \frac{C_n}{\lambda} \|f\|_{L^1}.$$

**Proof:** Define  $A_\lambda = \{x \mid Mf(x) > \lambda\}$ . If  $x \in A_\lambda$ , then there exists  $r_x$  such that

$$\frac{1}{|B_{r_x}(x)|} \int_{B_{r_x}(x)} |f(y)| dy > \lambda.$$

Let now  $x_m \in A_\lambda^c$  such that  $x_m \rightarrow x$  in  $\mathbb{R}^n$ , and note

$$\begin{aligned} \frac{1}{|B_{r_x}(x_m)|} \int_{B_{r_x}(x_m)} |f(y)| dy &= \frac{1}{|B_{r_x}(x)|} \int_{\mathbb{R}^n} \mathbb{1}_{B_{r_x}(x_m)}(y) |f(y)| dy \\ &\rightarrow \frac{1}{|B_{r_x}(x)|} \int_{B_{r_x}(x)} |f(y)| dy, \end{aligned}$$

by DCT. But since  $x_m \in A_\lambda^c$ , the left hand side is less than  $\lambda$ , which is inherited by the limit. So  $x \in A_\lambda^c$ .

Thus  $A_\lambda = Mf^{-1}(\lambda, \infty)$  is open for all  $\lambda > 0$ , hence by a  $\pi$ -system argument,  $Mf$  is Borel measurable.

To prove the inequality, take  $K \subseteq A_\lambda$  compact. Then the family  $\{B_{r_x}(x) \mid$

$x \in A_\lambda\}$  is an open cover of  $K$ , so by compactness, there exists a finite subcover  $B_1, \dots, B_N$  such that

$$K \subseteq \bigcup_{i=1}^n B_i,$$

and  $B_i = B_{r_x}(x)$  for some  $x \in A_\lambda$ . By Wiener's covering lemma, there exists a sub-collection of disjoint  $B_1, \dots, B_k$  such that

$$|K| \leq 3^n \sum_{j=1}^k |B_j|.$$

Now each  $B_j = B_{r_x}(x)$ , so

$$\frac{1}{|B_j|} \int_{B_j} |f(y)| \, dy > \lambda.$$

Thus, we have

$$\begin{aligned} |K| &\leq 3^n \sum_{j=1}^n |B_j| = \frac{3^n}{\lambda} \sum_{j=1}^n \lambda |B_j| \leq \frac{3^n}{\lambda} \sum_{j=1}^n \int_{B_j} |f(y)| \, dy \\ &\leq \frac{3^n}{\lambda} \|f\|_{L^1}, \end{aligned}$$

so since, by the inner regularity of  $\mu$ ,

$$|A_\lambda| \leq \sup\{|K| \mid K \subseteq A_\lambda, K \text{ compact}\} \leq \frac{3^n}{\lambda} \|f\|_{L^1}.$$

Finally, since  $\{x \mid Mf(x) = \infty\} \subseteq A_\lambda$  for all  $\lambda > 0$ , we have

$$|\{Mf = \infty\}| \leq \frac{3^n}{\lambda} \|f\|_{L^1} \quad \forall \lambda > 0.$$

Before we prove Lebesgue's theorem, we state a few consequences.

- (i) Points  $x \in \mathbb{R}^n$  for which the conclusion of the theorem holds are called *Lebesgue points*. Then

$$|\{x \mid x \text{ is not a Lebesgue point}\}| = 0.$$

(ii) When  $n = 1$ , we deduce that

$$\lim_{h \rightarrow 0} \frac{1}{2h} \int_{x-h}^{x+h} f(y) \, dy = f(x)$$

for almost all  $x \in \mathbb{R}$  and  $f \in L^1$ .

(iii) In the setting of the approximate identity theorem, we actually have  $\phi_\varepsilon * f \rightarrow f$  almost everywhere as  $\varepsilon \rightarrow 0$ , for  $f \in L^1$ .

(iv) Lebesgue's theorem also holds for sets  $\tilde{B}_r(x)$  that “shrink nicely” towards, but not always.

Now let's get to the proof.

**Proof:** As before, define

$$A_\lambda = \left\{ x \in \mathbb{R}^n \mid \limsup_{r \rightarrow 0} \frac{1}{|B_r(x)|} \int_{B_r(x)} |f(y) - f(x)| \, dy > 2\lambda \right\}$$

for  $\lambda > 0$ . It suffices to prove  $|A_\lambda| = 0$  for all  $\lambda$ , as the exceptional set is  $A = \bigcup A_{1/n}$ . So then we will have

$$|A| \leq \sum_n |A_{1/n}| = 0.$$

Let  $r > 0$ , and take  $g \in C_c(\mathbb{R}^n)$  such that  $\|f - g\|_{L^1} < \varepsilon$ . Then,

$$\begin{aligned} \frac{1}{|B_r(x)|} \int_{B_r(x)} |f(y) - f(x)| \, dy &\leq \frac{1}{|B_r(x)|} \int_{B_r(x)} |f(y) - g(y)| \, dy \\ &\quad + \frac{1}{|B_r(x)|} \int_{B_r(x)} |g(y) - g(x)| \, dy \\ &\quad + |f(x) - g(x)| \\ &= a + b + c. \end{aligned}$$

Since  $g$  is uniformly continuous,  $b \rightarrow 0$  as  $r \rightarrow 0$ . Therefore, if  $x \in A_\lambda$ , we must have either  $M[f - g](x) > \lambda$ , or  $|f(x) - g(x)| > \lambda$ . But by the Hardy-Littlewood maximal inequality,

$$|\{x \mid M[f - g](x) > \lambda\}| \leq \frac{C_n}{\lambda} \|f - g\|_{L^1} \leq \frac{C_n}{\lambda} \varepsilon,$$

for all  $\lambda > 0$ . For the second, by Markov's inequality,

$$|\{x \mid |f(x) - g(x)| > \lambda\}| \leq \frac{\|f - g\|_{L^1}}{\lambda} \leq \frac{\varepsilon}{\lambda},$$

Hence overall we have

$$|A_\lambda| \leq \frac{C_n + 1}{\lambda} \varepsilon.$$

Therefore the result follows, as  $\varepsilon > 0$  was arbitrary.

**Theorem 1.6** (Egorov). *Let  $E \subseteq \mathbb{R}^n$  be measurable,  $|E| < \infty$ , and let  $f_k : E \rightarrow \mathbb{C}$  (or  $\mathbb{R}$ ) be measurable and such that  $f_k \rightarrow f$  almost everywhere on  $E$  as  $k \rightarrow \infty$ . Then for all  $\varepsilon > 0$ , there exists  $A_\varepsilon \subseteq E$  such that  $|E \setminus A_\varepsilon| < \varepsilon$ , and  $f_k \rightarrow f$  uniformly on  $A_\varepsilon$ .*

**Proof:** Without loss of generality, assume  $f_k \rightarrow f$  pointwise on  $E$ , by throwing away the measure zero set where it does not. Then, define

$$E_k^m = \{x \in E \mid |f_j(x) - f(x)| < 1/m \quad \forall j > k\}.$$

This is clearly monotone, in the sense that  $E_k^m \subseteq E_{k+1}^m$ , and moreover  $\bigcup_k E_k^m = E$ , since  $f_j \rightarrow f$  pointwise. So

$$|E_k^m| \uparrow |E| \implies |E \setminus E_k^m| \rightarrow 0$$

as  $k \rightarrow \infty$ . Pick a subsequence  $k_m = k_m(\varepsilon)$  such that  $|E \setminus E_{k_m}^m| \leq \varepsilon 2^{-m}$ . Hence if

$$A_\varepsilon = \bigcap_{m=1}^{\infty} E_{k_m}^m,$$

then we have

$$|E \setminus A_\varepsilon| \leq \varepsilon \sum_{m=1}^{\infty} 2^{-m} = \varepsilon.$$

Also, on  $A_\varepsilon$ ,  $f_k \rightarrow f$  uniformly, and the result follows.

*Remark.* By inner regularity, we can take  $A_\varepsilon$  closed.

**Theorem 1.7** (Lusin). *Let  $E \subseteq \mathbb{R}^n$ ,  $|E| < \infty$ , and  $f : E \rightarrow \mathbb{R}$  (or  $\mathbb{C}$ ) measurable. Then for all  $\varepsilon > 0$ , there exists  $F_\varepsilon \subseteq E$  such that  $|E \setminus F_\varepsilon| < \varepsilon$ , and  $f : F_\varepsilon \rightarrow \mathbb{R}$  is continuous.*

*Remark.* Note that  $f$  is not continuous at  $x \in F_\varepsilon$  regarded as a map on  $E$ .

**Proof:** Take  $f$  simple, say

$$f = \sum_{k=1}^m a_k \mathbb{1}_{A_k},$$

for  $A_k \in \mathcal{B}(E)$  disjoint, and without loss of generality we can say

$$E = \bigcup_{k=1}^m A_k.$$

Then take  $K_k \subseteq A_k$  compact, such that  $|A_k \setminus K_k| \leq \varepsilon/m$ , and note that the  $K_k$ 's are at positive distance from one another. Since  $f$  is constant on  $K_k$ , it is continuous there, and if we let

$$B = \bigcup_{k=1}^m K_k,$$

then  $|E \setminus B| \leq \varepsilon$ . Next, for general  $f$ , take  $f_m$  simple such that  $f_m \rightarrow f$  pointwise on  $E$ . Now, by what proceeds, there exists  $C_m$  such that  $|C_m| \leq \varepsilon 2^{-m-1}$ , and  $f_m$  is continuous on  $C_m^c$ . Next, by Egorov's theorem, there exists  $A_\varepsilon$  compact such that  $f_m \rightarrow f$  uniformly, and  $|E \setminus A_\varepsilon| < \varepsilon/2$ . Then set

$$F_\varepsilon = A_\varepsilon \cap \left( \bigcap_{m=1}^{\infty} C_m^c \right).$$

So on  $F_\varepsilon$ ,  $f$  is the uniform limit of continuous functions, hence it is continuous. Finally,

$$|E \setminus F_\varepsilon| \leq |E \setminus A_\varepsilon| + \sum_{m=1}^{\infty} |C_m| \leq \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon.$$

## 2 $L^2$

For any measure space  $(E, \mathcal{E}, \mu)$ , the space  $L^2(\mu)$  is a Hilbert space for the inner product

$$(f, g)_\mu = (f, g) = \int_E f \cdot \bar{g} \, d\mu.$$

In any Hilbert space  $H$ , a family  $(u_j)$  is called *orthogonal* if  $(u_i, u_j) = 0$  if  $i \neq j$ . If further  $(u_i, u_i) = \|u_i\|^2 = 1$ , we say  $(u_j)$  are *orthonormal*, and if  $\text{span}\langle u_1, u_2, \dots \rangle = H$ , we say  $(u_j)$  is an *orthonormal basis* of  $H$ .

### Example 2.1.

1. Take  $E = [0, 1]$  and  $\mu = dx$ , then  $\{e_n(x) = e^{-2\pi i n x} \mid n \in \mathbb{Z}\}$  is an orthonormal basis of  $L^2([0, 1])$ , by density of  $C[0, 1]$  in  $L^2$ , and the Stone-Weierstrass theorem.
2. The Haar basis of  $L^2(\mathbb{R})$ . Take

$$\{\psi_{nk}(x) = 2^{n/2} \psi(2^n x - k), n, k \in \mathbb{Z}\}$$

where  $\psi = \mathbb{1}_{[0, 1/2)} - \mathbb{1}_{[1/2, 1)}$ .

Recall the following.

**Theorem 2.1** (Riesz Representation Theorem). *Let  $\Lambda : H \rightarrow \mathbb{R}$  be a bounded linear map on a Hilbert space  $H$ . Then there exists a unique  $g \in H$  such that  $\Lambda(h) = \langle g, h \rangle_H$  for all  $h \in H$ .*

On a measurable space  $(E, \mathcal{E})$ , consider measure  $\mu, \nu : \mathcal{E} \rightarrow [0, \infty]$ . We say  $\nu$  is *absolutely continuous* with respect to  $\mu$ , and we write  $\nu \ll \mu$ , if  $\mu(A) = 0$  for  $A \in \mathcal{E} \implies \nu(A) = 0$ . If  $\mu \ll \nu$  and  $\nu \ll \mu$ , we say  $\mu, \nu$  are *mutually absolutely continuous*, and if there exists  $B \in \mathcal{E}$  such that  $\mu(B) = 0 = \nu(B^c)$ , we say  $\mu, \nu$  are *mutually singular*.

**Theorem 2.2** (Radon-Nikodym). *Let  $\mu, \nu$  be finite measures on  $(E, \mathcal{E})$  such that  $\nu \ll \mu$ . Then there exists function  $w \in L^1(\mu)$ ,  $w \geq 0$  such that*

$$\nu(A) = \int_A w \, d\mu \quad \forall A \in \mathcal{E}.$$

We often call

$$w = \frac{d\nu}{d\mu},$$

the Radon-Nikodym derivative of  $\nu$  with respect to  $\mu$ .

*Remark.*

1. One can show that  $w$  is unique, and that the result holds for  $\mu, \nu$   $\sigma$ -finite.
2. The proof actually shows that for all  $h \geq 0$  and measurable,

$$\int_E h \, d\nu = \int_E hw \, d\mu.$$

**Proof:** Define new measures  $\alpha = \mu + 2\nu$ ,  $\beta = 2\mu + \nu$ , and on  $L^2(\alpha)$  consider

$$\Lambda(f) = \int_E f \, d\beta,$$

which is linear and continuous since

$$\begin{aligned} |\Lambda(f)| &\leq \int_E |f| \, d\beta = 2 \int_E |f| \, d\mu + \int_E |f| \, d\nu \\ &= 2 \int_E |f| \, d\alpha \stackrel{C.S.}{\leq} 2|\alpha(E)|^{1/2} \|f\|_{L^2(\alpha)}. \end{aligned}$$

Therefore by Riesz, there exists  $g \in L^2(\alpha)$  such that  $\Lambda(f) = \langle g, f \rangle_{L^2(\alpha)}$ . This is equivalent to

$$\begin{aligned} \int_E f \, d\beta &= \int_E fg \, d\alpha \iff \int_E f(2 \, d\mu + d\nu) = \int_E f(d\mu + 2 \, d\nu) \\ &\iff \int_E f(2 - g) \, d\mu = \int_E f(2g - 1) \, d\nu, \quad (*) \end{aligned}$$

for all  $f \in L^2(\alpha)$ . Testing for  $f = \mathbb{1}_{A_j}$ , where  $A_j = \{x \in E \mid g(x) \leq \frac{1}{2} - \frac{1}{j}\}$ , we get

$$\begin{aligned} \int_E f(2g - 1) \, d\nu &\leq -\frac{1}{j} \nu(A_j), \\ \int_E f(2 - g) \, d\mu &\geq \frac{3}{2} \mu(A_j). \end{aligned}$$

This identity means that  $\mu(A_j) = \nu(A_j) = 0$ . Therefore  $g \geq \frac{1}{2} \mu$  and  $\nu$ -almost everywhere. Similarly, consider  $A'_j = \{x \mid g(x) \geq 2 + \frac{1}{j}\}$ , one shows that  $g \leq 2 \mu$  and  $\nu$ -almost everywhere.

Since the integrals in (\*) are non-negative, we can approximate and use the monotone convergence theorem to establish

$$\int_E f(2 - g) \, d\mu = \int_E f(2g - 1) \, d\nu,$$

for all  $f \geq 0$ . Next for  $\{x \mid g(x) = \frac{1}{2}\} \in \mathcal{E}$ , we have from (\*) with  $f = \mathbb{1}_{\{g=\frac{1}{2}\}}$ , that

$$\frac{3}{2}\mu\left(\left\{g = \frac{1}{2}\right\}\right) = 0 \implies \nu\left(\left\{g = \frac{1}{2}\right\}\right) = 0,$$

by absolute continuity (if  $\nu$  is not absolutely continuous, we can decompose  $\nu$  into  $\nu_{\text{ac}} + \nu_{\perp}$  by intersecting with  $\{x \mid g(x) = \frac{1}{2}\}$ ).

Take any  $h \geq 0$  measurable. Then define

$$f = \frac{h}{2g-1}, \quad w = \frac{2-g}{2g-1}, \quad f = w = 0 \text{ in } \left\{x \mid g(x) = \frac{1}{2}\right\}.$$

Then by (\*),

$$\int_E h \, d\nu = \int_E f(2g-1) \, d\nu = \int_E f(2-g) \, d\mu = \int_E fw(2g-1) \, d\mu = \int hw \, d\mu,$$

and  $w \in L^1(\mu)$  because

$$\|w\|_{L^1(\mu)} = \int_E w \, d\mu = \int_E d\nu = \nu(E) < \infty.$$

*Remark.* If  $\mathbb{P}$  is a probability measure on a subset  $E \subseteq \mathbb{R}^n$ , then it has a unique decomposition  $\mathbb{P} = \mathbb{P}_{\text{ac}} + \mathbb{P}_{\perp}$ , where  $\mathbb{P}_{\text{ac}} \ll \mu$ , the Lebesgue measure, and the Lebesgue probability density  $d\mathbb{P}_{\text{ac}}/d\Gamma$ , the corresponding Radon-Nikodym density.

## 2.1 Dual Spaces

Let  $X$  be a topological vector space over  $\mathbb{R}$  or  $\mathbb{C}$ . The *topological dual space*  $X'$  is defined as

$$X' = \{\Lambda : X \rightarrow \mathbb{R} \text{ (or } \mathbb{C}) \text{ continuous, linear}\}$$

which is itself a vector space of linear maps. If  $X$  carries a norm  $\|\cdot\|_X$ , then on  $X'$  we have a norm topology from the dual norm

$$\|\Lambda\|_{X'} = \sup_{\|x\| \leq 1} |\Lambda(x)|,$$

and  $(X', \|\cdot\|_{X'})$  is always a Banach space.

In this case, we define the *bidual space*  $X'' = (X')'$ .  $X$  is *reflexive* if  $X = X''$ .

The *canonical embedding*  $X \hookrightarrow X''$  is given by  $x \mapsto L_x$ , where  $L_x(\Lambda) = \Lambda(x)$ , for  $\Lambda \in X', x \in X$ .



**Example 2.2.**

Take  $X = L^p(\mathbb{R}^n)$ , with the usual measure. We already know from self-duality of Hilbert spaces that  $(L^2)' = L^2$ . For  $q$  conjugate to  $p$ , and  $g \in L^q(\mathbb{R}^n)$ , we can define a linear functional

$$\Lambda_g(f) = \int_{\mathbb{R}^n} f(x)g(x) \, dx,$$

for  $f \in L^p$ . By Hölder's inequality,

$$|\Lambda_g(f)| \leq \|f\|_{L^p} \|g\|_{L^q},$$

hence  $\Lambda_g \in (L^p)'$ . In this sense, we can say that  $L^q \subseteq (L^p)'$ . In fact, we have

$$\|\Lambda_g\|_{(L^p)'} = \|g\|_{L^q}.$$

In fact, we can show this inclusion is exact.

**Theorem 2.3.** *Let  $1 \leq p < \infty$ , and  $q$  be conjugate. Then*

$$(L^p(\mathbb{R}^n))' = L^q(\mathbb{R}^n).$$

**Corollary 2.1.**  *$L^p(\mathbb{R}^n)$  is reflexive if  $1 < p < \infty$ .*

We begin with the following lemma:

**Lemma 2.1.** *For  $1 \leq p < \infty$  and  $q$  conjugate, for all  $U \in (L^p)'$  positive, there exists  $g \geq 0$  with  $g \in L^q$  such that*

$$U(f) = \int_{\mathbb{R}^n} f(x)g(x) \, dx,$$

for all  $f \in L^p$ , with  $\|U\|_{(L^p)'} = \|g\|_{L^q}$ .

**Proof:** On  $\mathbb{R}^n$  consider the finite measure  $\mu$  with density  $d\mu(x) = e^{-|x|^2} dx$ . Further, for  $A \in \mathcal{B}$ , set

$$\nu(A) = U(e^{-|x|^2/p} \mathbb{1}_A) \in [0, \infty),$$

as  $U$  is positive. Further, if  $A_n \in \mathcal{B}$  such that  $A_n \downarrow \bigcap A_n = \emptyset$ , then

$$\nu(A_n) = U(e^{-|x|^2/p} \mathbb{1}_{A_n}) \leq \|U\|_{(L^p)'} \|e^{-|x|^2/p} \mathbb{1}_{A_n}\|_{L^p} \rightarrow 0,$$

by DCT. Therefore  $\nu$  is a measure on  $\mathcal{B}$ . If  $\mu(A) = 0$ , then

$$\nu(A) \leq \|U\|_{(L^p)'} \|e^{-|x|^2/p} \mathbb{1}_A\|_{L^p} \leq \|U\|_{(L^p)'} \mu(A)^{1/p} = 0,$$

so  $\nu \ll \mu$ . Thus by Radon-Nikodym theorem, there exists  $G \in L^1(\mu)$ ,  $G \geq 0$  such that

$$\nu(A) = U(e^{-|x|^2/p} \mathbb{1}_A) = \int_{\mathbb{R}^n} \mathbb{1}_A G \, d\mu.$$

Next set  $f = e^{-|x|^2/p} F$ , where  $F = \sum a_k \mathbb{1}_{A_k}$ . Then by the linearity of  $U$ ,

$$\begin{aligned} U(f) &= U(e^{-|x|^2/p} F) = \sum_{k=1}^n a_k U(e^{-|x|^2/p} \mathbb{1}_{A_k}) = \sum_{k=1}^n a_k \int_{\mathbb{R}^n} \mathbb{1}_{A_k} G(x) e^{-|x|^2} \, dx \\ &= \int_{\mathbb{R}^n} \sum_{k=1}^n a_k \mathbb{1}_{A_k}(x) e^{-|x|^2/p} G(x) e^{-|x|^2/q} \, dx = \int_{\mathbb{R}^n} f(x) g(x) \, dx, \end{aligned}$$

where  $g(x) = G(x) e^{-|x|^2/q}$ . Since  $\{f = e^{-|x|^2/p} F \mid F \text{ simple}\}$  is dense in  $L^p \cap \{f \geq 0\}$ , and since

$$\int_{\mathbb{R}^n} |f(x) g(x)| \, dx = \int_{\mathbb{R}^n} |f(x)| |g(x)| \, dx = U(|f|) \leq \|U\|_{(L^p)', L^p} \|f\|_{L^p},$$

by MCT, taking  $0 \leq f_n \uparrow f \in L^p$ , the identity extends to

$$U(f) = \int_{\mathbb{R}^n} f(x) g(x) \, dx,$$

for all  $f \in L^p$ . Equality of norms is given by Hölder's.

For general real-valued elements in  $(L^p)'$ , we may break into the positive and negative parts respectively. An element  $\Lambda \in (L^p)'$  is called *positive* if  $\Lambda(f) \geq 0$  for all  $f \in L^p$  such that  $f \geq 0$  a.e. One can show that any  $\Lambda \in (L^p)'$  decomposes uniquely into  $\Lambda = \Lambda_+ - \Lambda_-$ , where  $\Lambda_+$  and  $\Lambda_-$  are positive, and we thus restrict to  $\Lambda \in (L^p)'$  positive.

Note that  $(L^\infty)' \neq L^1$ , and  $L^1$  is not reflexive. One can describe, however  $(C_c(\mathbb{R}^n))'$ , by the action of finite measure.

Recall a measure on a topological space  $E$  is called *regular* on any  $\sigma$ -algebra  $\mathcal{M} \supseteq \mathcal{B}(E)$ , if for all  $A \in \mathcal{M}$ ,  $\varepsilon > 0$ , there exists  $C$  closed and  $O$  open such that  $C \subseteq A \subseteq O$  and  $\mu(O \setminus C) < \varepsilon$ .

**Theorem 2.4** (Riesz-Markov).

1. Let  $\mu$  be a finite regular measure on  $\mathcal{M} \supseteq \mathcal{B}(\mathbb{R}^n)$ . Then

$$\Lambda_\mu(f) = \int_{\mathbb{R}^n} f \, d\mu$$

defines a positive element of  $(C_c(\mathbb{R}^n))'$ .

2. Let  $\Lambda \in (C_c(\mathbb{R}^n))'$  be positive. Then there exists  $\mathcal{M} \subseteq \mathcal{B}(\mathbb{R}^n)$  and a finite regular measure  $\mu$  on  $\mathcal{M}$  such that

$$\Lambda(f) = \int_{\mathbb{R}^n} f \, d\mu,$$

for all  $f \in C_c$ .

**Proof:**

For 1, note that

$$|\Lambda_\mu(f)| \leq \|f\|_\infty \mu(\mathbb{R}^n).$$

2 is in Rudin.

## 2.2 Weak and Weak-\* Topologies

Let  $X$  be a Banach space with dual  $X'$ . Then on  $X, X'$ , the *strong topology* is the norm-topology for  $\|\cdot\|_X, \|\cdot\|_{X'}$ .

The unit balls  $\{x \mid \|x\| \leq 1\}$  are compact  $\iff X, X'$  are finite dimensional  $\iff$  any norm bounded sequence  $x_n$  has a convergent subsequence.

To consider weak topologies, we need some definitions.

**Definition 2.1.** On a real (or complex) vector space  $X$ , a map  $p : X \rightarrow [0, \infty)$  is called a *semi-norm* over  $X$  if:

- (i)  $p(x + y) \leq p(x) + p(y)$  for all  $x, y \in X$ .
- (ii)  $p(\lambda x) = |\lambda|p(x)$  for all  $x \in X, \lambda \in \mathbb{R}$ .

Note that we do not require that  $p(x) = 0 \iff x = 0$ . A collection  $\mathcal{P}$  of semi-norms is said to *separate* points if, for all  $0 \neq x \in X$ , there is  $p \in \mathcal{P}$  such that  $p(x) \neq 0$ .

The topology  $\tau_{\mathcal{P}}$  over  $X$  generated by the ‘open’ sets

$$V_x(p, n) = \{y \in X \mid p(x - y) < 1/n\},$$

for  $n \in \mathbb{N}, x \in X$  and  $p \in \mathcal{P}$ . This is called a *locally convex* topology over  $X$ .

One can show that  $x_n \in X$  converges to  $x$  in  $\tau_{\mathcal{P}}$  if and only if  $p(x_n - x) \rightarrow 0$  as  $n \rightarrow \infty$ , for all  $p \in \mathcal{P}$ .

If  $\mathcal{P}$  separates points, then  $\tau_{\mathcal{P}}$  is a Hausdorff topology, but not in general a metrizable topology, unless  $\mathcal{P}$  is countable, in which case a metric is given by

$$d_{\mathcal{P}}(x, y) = \sum_{i=1}^{\infty} 2^{-i} \frac{p_i(x - y)}{1 + p_i(x - y)}.$$

We call  $(X, \tau_{\mathcal{P}})$  a locally complete topological vector space, and if it is complete, it is called a *Fréchet space*.

### Example 2.3. (Weak Topology)

On  $X$ , consider seminorms

$$p_{\Lambda}(x) = |\Lambda(x)|,$$

for  $\Lambda \in X'$ . We define the *weak topology*  $\tau_{\omega} = \tau_{\mathcal{P}}$  for  $\mathcal{P} = \{p_{\Lambda} \mid \Lambda \in X'\}$ .

Then  $x_n \rightarrow x$  weakly if  $\Lambda(x_n) \rightarrow \Lambda(x)$ , for all  $\Lambda \in X'$ .

We will show later (using the Hahn-Banach theorem) that  $\tau_{\omega}$  separates points.

Next, if  $X'$  is a dual space, then it has a weak topology as a Banach space, but further it carries a weak-\* topology  $\tau_{\omega^*}$  induced by the semi-norms

$$\mathcal{P} = \{p_x(\Lambda) = |\Lambda(x)| \mid x \in X\}.$$

Here  $\Lambda_n \rightarrow \Lambda$  if  $\Lambda_n(x) \rightarrow \Lambda(x)$  for all  $x \in X$ . Now  $\mathcal{P}$  clearly separates points and is Hausdorff.

If  $X = X''$ , then  $\tau_{\omega} = \tau_{\omega^*}$ , up to the Hahn-Banach theorem.

### Example 2.4.

For  $1 \leq p < \infty$ ,  $(L^p(\mathbb{R}^n))' = L^q(\mathbb{R}^n)$ , and so  $f_k \in L^p$  converges to  $f \in L^p$  weakly if and only if

$$\int_{\mathbb{R}^n} f_k(x)g(x) \rightarrow \int_{\mathbb{R}^n} f(x)g(x) \, dx,$$

for all  $g \in L^q$ . Similarly, if  $1 < p \leq \infty$ ,  $\Lambda_{f_n} \in (L^p)'$  corresponds to  $f_n \in L^p$ , then  $f_n \rightarrow f$  weak-\* in  $(L^p)'$  if the same limit holds.

Let  $\mu_n$  be a sequence of Borel probability measures on  $\mathbb{R}^n$ , which by the Riesz-Markov theorem defines a sequence  $\Lambda_{\mu_n} \in (C_c(\mathbb{R}^n))'$ . It converges

weak-\* in  $C'_c$  if

$$\int_{\mathbb{R}^n} f \, d\mu_k \rightarrow \int_{\mathbb{R}^n} f \, d\mu,$$

for all  $f \in C_c$ . In general, for a metric space  $D$ , and  $\mu_n$  Borel probability measures on  $D$ , then  $\mu_n \in (C_c(D))'$ , via

$$f \mapsto \int_D f \, d\mu_n,$$

and  $\mu_n \rightarrow \mu$  weak-\* if and only if

$$\int Df \, d\mu_n \rightarrow \int_D f \, d\mu$$

for all  $f \in C_c(D)$ . If  $D$  is compact, so  $C_c(D) = C(D)$ , this coincides with the weak convergence of the laws  $\mu_n$ . Otherwise, it is called *vague convergence*.

If we have tightness, then vague convergence will imply weak convergence.

## 2.3 Compact Subsets of Infinite-Dimensional Spaces

Recall:

**Theorem 2.5** (Arzelá-Ascoli). *Let  $I = [0, 1]$  (or in general a compact metric space). Let  $f_n \in C(I)$  be a such that  $\|f_k\|_\infty \leq M < \infty$ , and equicontinuous, so for all  $\varepsilon > 0$  there exists  $\delta > 0$  such that for all  $k$ ,*

$$|x - y| < \delta \implies |f_k(x) - f_k(y)| < \varepsilon.$$

*Then there exists  $f \in C(I)$  such that  $f_{k_j} \rightarrow f$  uniformly on  $I$ , on some subsequence.*

### Example 2.5. (Hölder Continuous Function)

For  $f \in C(I)$ , we say  $f$  is  $\gamma$ -Hölder continuous if

$$\|f\|_{C^{0,\gamma}} = \|f\|_\infty + \sup_{x \neq y} \frac{|f(x) - f(y)|}{|x - y|^\gamma} < \infty.$$

Moreover,  $\{f \mid \|f\|_{C^{0,\gamma}} \leq 1\}$  is compact in  $C(I)$  for  $\gamma > 0$ .

**Theorem 2.6** (Banach-Alaoglu). *Let  $X$  be a normed space. Then the unit ball  $B'_1 = \{\Lambda \in X' \mid \|\Lambda\|_{X'} \leq 1\}$  is compact for the  $\tau_{\omega^*}$  topology.*

**Proof:** We will only prove for  $X$  separable; for general  $X$  this is a consequence of Tychonoff's theorem and the axiom of choice.

Consider  $\tau_{\tilde{\mathcal{P}}}$ , the topology generated by the seminorms

$$\tilde{\mathcal{P}} = \{p_{x_k}(\Lambda) = |\Lambda(x_k)| \mid x_k \in D\},$$

where  $D$  is a countable dense subset of  $X$ . Then  $\tau_{\tilde{\mathcal{P}}} = \tau_{\omega^*}$  on  $B'_1$  as trace topologies (i.e. restrictions of topologies to a subset), and hence  $\tau_{\omega^*}$  is metrizable by

$$d(\Lambda, \Lambda') = \sum_{k \in D} 2^{-k} \frac{|(\Lambda - \Lambda')(x_k)|}{1 + |(\Lambda - \Lambda')(x_k)|}.$$

To show this, the topology  $\tau_{\tilde{\mathcal{P}}}$  is generated by open sets

$$V(x_n, n) = \{\Lambda \mid |\Lambda(x_k)| < 1/n\},$$

for  $x_k \in D$ ,  $n \in \mathbb{N}$ . This is indeed open for  $\tau_{\omega^*}$ . Conversely, we need to show that  $V(x, n)$ , for  $x \in X$ ,  $n \in \mathbb{N}$ , contains  $V(x_k, m)$  for some  $x_k \in D$ ,  $m \in \mathbb{N}$ . But taking  $\|x_k - x\| < \varepsilon$ , and for  $\Lambda \in V(x_k, m)$ , we have

$$|\Lambda(x)| \leq |\Lambda(x - x_k)| + |\Lambda(x_k)| \leq \|x - x_k\| + \frac{1}{m} \leq \varepsilon + \frac{1}{m} < \frac{1}{n},$$

since  $\|\Lambda\|_{X'} \leq 1$ , and by taking an appropriate choice of  $\varepsilon, m$ .

Now to prove the Banach-Alaoglu theorem, we will show that for a sequence  $\Lambda_j$  in  $B'_1$ , that there is  $\Lambda \in B'_1$  such that  $\Lambda_{j_k} \rightarrow \Lambda$  weak-\* along a subsequence.

Indeed, let  $D = \{x_1, x_2, \dots\}$  be a countable dense subset of  $X$ . The sequence  $|\Lambda_j(x_1)| \leq \|x_1\|_X$ , and therefore there exists  $j_k$  such that  $\Lambda_{j_k}(x_1) \rightarrow \Lambda(x_1) \in \mathbb{R}$  as  $k \rightarrow \infty$ , by Bolzano-Weierstrass, such that  $|\Lambda(x_1)| \leq \|x_1\|_X$ .

Next  $|\Lambda_{j_k}(x_2)| \leq \|x_2\|_X$ , so pick a further subsequence  $\Lambda_{2,j}(x_2) \rightarrow \Lambda(x_2)$  such that  $|\Lambda(x_2)| \leq \|x_2\|_X$ . Iterating this procedure, we obtain a sequence  $\Lambda_{l,j} \in B'_1$  such that  $\Lambda_{l,j}(x_k) \rightarrow \Lambda(x_k)$  for all  $k \leq l$ , with  $|\Lambda(x_k)| \leq \|x_k\|_X$ .

Then choosing the diagonal sequence  $\Lambda_{j,j}$ , we have  $\Lambda_{j,j}(x_k) \rightarrow \Lambda(x_k)$  for all  $k$ . It remains to show that indeed  $\Lambda \in B'_1$ . We have three things to show

First, we show that  $\Lambda$  is uniformly continuous on  $D$ . Take  $x, y \in D$  such that  $\|x - y\| < \varepsilon/2$ , and choose  $J$  such that  $|\Lambda_{j,j}(x) - \Lambda(x)| + |\Lambda_{j,j}(y) - \Lambda(y)| < \varepsilon/2$  for all  $j \geq J$ . Then note

$$\begin{aligned} |\Lambda(x) - \Lambda(y)| &\leq |\Lambda(x) - \Lambda_{j,j}(x)| + |\Lambda(y) - \Lambda_{j,j}(y)| + |\Lambda_{j,j}(x - y)| \\ &< \frac{\varepsilon}{2} + \|\Lambda_{j,j}\|_{X'} \|x - y\| < \varepsilon. \end{aligned}$$

So  $\Lambda$  extends to  $\tilde{\Lambda} = \Lambda$ , continuous on  $X$  (since a uniformly continuous function on a dense set extends to the completion in a continuous way).

Now we show that  $\Lambda$  is linear. Let  $x, y \in X$  and  $a \in \mathbb{R}$ . Set  $z = x + ay$ . Choose  $x', y'$  and  $z'$  such that  $\|x - x'\|_X + |a|\|y - y'\|_X + \|z - z'\|_X < \varepsilon$ . Then

$$\begin{aligned} |\Lambda(z) - \Lambda(x) - a\Lambda(y)| &\leq |\Lambda(z) - \Lambda(z')| + |\Lambda(x) - \Lambda(x')| + |a||\Lambda(y) - \Lambda(y')| \\ &\quad + |\Lambda(z') - \Lambda_{j,j}(z')| + |\Lambda(x') - \Lambda_{j,j}(x')| + |a||\Lambda(y') - \Lambda_{j,j}(y')| \\ &\quad + |\Lambda_{j,j}(z') - \Lambda_{j,j}(x') - a\Lambda_{j,j}(y')|. \end{aligned}$$

The first line we can bound by  $\varepsilon$  as  $|\Lambda(x') - \Lambda(x)| \leq \|x - x'\|$ , the second we can also bound by  $\varepsilon$  by taking  $j$  large enough, and the third line is

$$|\Lambda_{j,j}(z') - \Lambda_{j,j}(x') - a\Lambda_{j,j}(y')| \leq |\Lambda_{j,j}(z - x - ay)| + \|z - z'\| + \|x - x'\| + |a|\|y - y'\|,$$

which is at most  $\varepsilon$  as  $\Lambda_{j,j}$  is linear so the first term disappears. So  $\Lambda \in X'$ .

Finally, for  $x \in X$ ,

$$|\Lambda(x)| \leq |\Lambda(x')| + |\Lambda(x - x')| \leq \|x'\|_X + \varepsilon \leq \|x\|_X + 2\varepsilon,$$

where we choose  $x' \in D$ . So If  $\|x\|_X \leq 1$ , then  $|\Lambda(x)| \leq 1$ , and so  $\|\Lambda\|_{X'} \leq 1$ . Finally, we prove  $\Lambda_{j,j} \rightarrow \Lambda$ . Indeed,

$$|\Lambda_{j,j}(x) - \Lambda(x)| \leq |\Lambda_{j,j}(x') - \Lambda(x')| + 2\|x - x'\|_X \rightarrow 0.$$

So  $\Lambda_{j,j} \rightarrow \Lambda$  weak-\* as  $j \rightarrow \infty$ .

### Example 2.6.

Let  $f_k \in L^p(\mathbb{R}^n)$  such that  $\|f_j\|_{L^p} \leq 1$ , for all  $j$ . For  $1 < p \leq \infty$ ,  $f_j \in (L^q)'$ , so there exists  $f_{j_k}$  such that

$$\int_{\mathbb{R}^n} f_{j_k}(x)g(x) \, dx \rightarrow \int_{\mathbb{R}^n} f(x)g(x) \, dx,$$

for all  $g \in L^q$ , for some  $f \in L^p$  with  $\|f\|_{L^p} \leq 1$ .

Likewise, if  $\mu_n$  are probability measures on a compact metric space  $D$ , so  $\mu_n \in (C(D))'$ , then there exists  $\mu_{n_k}$  such that

$$\int_D f \, d\mu_{n_k} \rightarrow \int_D f \, d\mu,$$

for all  $f \in C(D)$  and some probability measure  $\mu$ . This is *Prokhorov's theorem*.

## 2.4 Hahn-Banach Theorem

**Definition 2.2.** Let  $X$  be a real vector space. A map  $p : X \rightarrow \mathbb{R}$  is called *sub-linear* of:

- (i)  $p(x + y) \leq p(x) + p(y)$  for all  $x, y \in X$ .
- (ii)  $p(tx) = tp(x)$  for all  $t \geq 0, x \in X$ .

*Remark.* Clearly if  $\ell : X \rightarrow \mathbb{R}$  is linear, then  $\ell, |\ell|$  are sub-linear, and so is any seminorm. If  $\ell$  is linear,  $p$  is a seminorm, then  $\ell(x) \leq p(x) \iff |\ell(x)| \leq p(x)$  for all  $x$ .

**Lemma 2.2** (Bounded Extension). *Let  $p : X \rightarrow \mathbb{R}$  be sub-linear on a real vector space  $X$ . Further, let  $M \subset X$  be a proper vector subspace, and  $\ell : M \rightarrow \mathbb{R}$  linear and such that  $\ell(x) \leq p(x)$  for all  $x \in M$ .*

*For  $x \in X \setminus M$ , define  $\tilde{M} = \text{span}(M, x)$ . Then there exists  $\tilde{\ell} : \tilde{M} \rightarrow \mathbb{R}$  linear, such that  $\tilde{\ell}|_M = \ell$  and  $\ell(y) \leq p(y)$  for all  $y \in \tilde{M}$ .*

**Proof:** Let  $y_1, y_2 \in M$  and write  $\ell(y_1) + \ell(y_2) = \ell(y_1 + y_2) \leq p(y_1 + y_2) \leq p(y_1 - x) + p(y_2 + x)$ . So rearranging,

$$\ell(y_1) - p(y_1 - x) \leq p(y_2 + x) - \ell(y_2),$$

for all  $y_1, y_2 \in M$ . So taking supremums and infimums, there exists  $a \in \mathbb{R}$  such that

$$\sup\{\lambda\ell(y) - p(y - x) \mid y \in M\} \leq a \leq \inf\{p(y + x) - \ell(y) \mid y \in M\}.$$

We define

$$\tilde{\ell}(z) = \tilde{\ell}(y + \lambda x) = \ell(y) + \lambda a,$$

which is linear on  $\tilde{M}$ . Clearly  $\tilde{\ell}|_M = \ell$ , and  $\tilde{\ell}(y) \leq p(y)$  for  $y \in M$ . Finally for  $z = y + \lambda x$ ,  $\lambda > 0$ ,

$$\tilde{\ell}(y + \lambda x) = \lambda \left( \ell\left(\frac{y}{\lambda}\right) + a \right) \leq \lambda \left( \ell\left(\frac{y}{\lambda}\right) + p\left(\frac{y}{\lambda} + x\right) - \ell\left(\frac{y}{\lambda}\right) \right) = p(y + \lambda x).$$

If  $\lambda < 0$ , set  $-\mu = \lambda$ , and we have

$$\tilde{\ell}(y - \mu x) = \mu \left( \ell\left(\frac{y}{\mu}\right) - a \right) \leq \mu \left( \ell\left(\frac{y}{\mu}\right) + p\left(\frac{y}{\mu} - x\right) - \ell\left(\frac{y}{\mu}\right) \right) = p(y - \mu x).$$



Now to extend  $\ell$  to  $X$ , we need a Zorn's type of argument. Consider the sets

$$S = \{(N, \tilde{\ell}) \mid M \subseteq N \subseteq X \text{ vector subspaces, } \tilde{\ell}|_M = \ell, \tilde{\ell} \text{ linear, } \tilde{\ell} \leq p \text{ on } N\}.$$

We can define an order relation on  $S$  by  $(N_1, \ell_1) \leq (N_2, \ell_2)$  if  $N_1 \subseteq N_2$  and  $\ell_2|_{N_1} = \ell_1$ . If  $T$  is a totally ordered subset of  $S$ , then define

$$N_T = \bigcup_{(N, \tilde{\ell}) \in T} N, \quad L(x) = \tilde{\ell}(x).$$

Then this is well-defined, and also an upper bound. Therefore by Zorn's lemma,  $S$  contains a maximal element  $(\mathcal{N}, \ell^*)$ . Thus, if  $X \neq \mathcal{N}$ , then by an application of the bounded extension lemma. We contradict maximality. Hence, we have proven:

**Theorem 2.7** (Hahn-Banach). *Let  $X$  be a real vector space,  $p : X \rightarrow \mathbb{R}$  sub-linear. Let  $\ell : M \rightarrow \mathbb{R}$  be linear on some subspace  $M \subseteq X$  such that  $\ell(x) \leq p(x)$  for all  $x \in M$ . Then there exists  $\tilde{\ell} : X \rightarrow \mathbb{R}$  linear such that  $\tilde{\ell}|_M = \ell$  and  $\tilde{\ell}(x) \leq p(x)$  for all  $x \in X$ .*

*Remark.*

1.  $\tilde{\ell}$  is not generally unique.
2. If  $X$  is a separable Banach space, then there exists a constructive proof by extending  $\ell$  on a separable subset, and then to its completion.
3. The result extends to complex valued functionals.

**Corollary 2.2.** *Let  $X$  be a Banach space,  $M \subseteq X$  a closed linear space. Then if  $\Lambda \in (M, \|\cdot\|_X)'$ , there exists  $\Lambda \in X'$  such that  $\Lambda|_M = \Lambda$  and  $\|\Lambda\|_{X'} = \|\Lambda\|_{M'}$ .*

**Proof:** We apply the Hahn-Banach theorem with the sub-linear function

$$p(x) = \|\Lambda\|_{(M, \|\cdot\|_X)'} \|x\|.$$

**Corollary 2.3** (Norming Functionals). *Let  $(X, \|\cdot\|_X)$  be a normed space,  $x \in X$ . Then there exists  $\Lambda = \Lambda_x \in X'$  such that  $\|\Lambda\|_{X'} = 1$ ,  $\Lambda(x) = \|x\|_X$ .*

*In particular, if  $\Lambda(x) = \Lambda(y)$  for all  $\Lambda \in X'$ , then  $x = y$ .*

**Proof:** Set  $M = \{cx \mid c \in \mathbb{R}\}$ . Then  $\ell(cx) = c\|x\|_X$  is linear and

$$\|\ell\|_{M'} = \sup_{\|cx\|_X \leq 1} |c|\|x\|_X = 1.$$

By Hahn-Banach, we obtain  $\Lambda_x$  such that  $\|\Lambda_x\|_{X'} = 1$ , and  $\Lambda(x) = \ell(x) =$

$\|x\|_X$ . Applying this with  $x - y$ , we see that

$$\Lambda_{x-y}(x - y) = \|x - y\| = 0 \implies x = y.$$

This implies that any normed linear space supports sufficiently many continuous linear functionals to separate points, and  $\tau_\omega$  on  $X$  is Hausdorff.

**Corollary 2.4.** *The canonical injection  $i : X \rightarrow X''$  for any normed linear space, is an isometry.*

**Proof:** We have

$$\|i(x)\|_{X''} = \sup_{\substack{\Lambda \in X' \\ \|\Lambda\|_{X'} \leq 1}} |\Lambda(x)| \leq \|x\|_X \cdot \|\Lambda\|_{X'} \leq \|x\|_X,$$

but this is at least  $|\Lambda_x(x)| = \|x\|_X$ .

A few consequences:

1. If  $X$  is reflexive, then it is a Banach space, as  $X'' = (X')'$  is complete.
2. If  $X$  is not complete, we can complete it on a subset of  $X''$ .
3. If  $X$  is reflexive, then  $\tau_\omega = \tau_{\omega^*}$ , so Banach-Alaoglu applies to  $\tau_\omega$  as well.

### Example 2.7.

Let  $I = [0, 1]$ ,  $\mu$  be the Lebesgue measure. Then  $(C(I), \|\cdot\|_\infty)$  is a closed linear subspace of  $L^\infty(I)$ .

Let  $\ell(f) = f(0)$ , linear and continuous as  $|\ell(f)| \leq \|f\|_\infty$ . Then by the Hahn-Banach theorem, there exists  $\tilde{\ell} \in (L^\infty)'$ .

Next take  $f_n(x) = \max(1 - nx, 0)$ . Then  $\ell(f_n) = f_n(0) = 1$ . But suppose that

$$\tilde{\ell}(f) = \int_I f(x)g(x) \, dx$$

for some  $g \in L^1(\mu)$ . Then

$$\tilde{\ell}(f_n) = \int_I f_n g \rightarrow 0$$

as  $n \rightarrow \infty$  by DCT. Hence we have a contradiction, and so no such  $g \in L^1$  exists.

**Theorem 2.8** (Hyperplane Separation). *Let  $A, B$  be convex, non-empty, disjoint in a Banach space  $X$ .*

(i) *Assume  $A$  is open. Then there exists  $\Lambda \in X'$  and  $\gamma \in \mathbb{R}$  such that*

$$\Lambda(a) < \gamma \leq \Lambda(b)$$

*for all  $a \in A, b \in B$ . If  $B$  is also open, we may take  $\gamma < \Lambda(b)$ .*

(ii) *If  $A$  is compact,  $B$  is closed, then there exists  $\Lambda \in X'$ ,  $\gamma_1, \gamma_2 \in \mathbb{R}$  such that*

$$\Lambda(a) < \gamma_1 < \gamma_2 \leq \Lambda(b),$$

*for all  $a \in A, b \in B$ .*

**Proof:**

(i) Fix  $a_0 \in A, b_0 \in B$ , define  $x_0 = b_0 - a_0$  and set

$$C = A - B + x_0 = \{a - b + x_0 \mid a \in A, b \in B\},$$

which is convex, open and contains 0, since  $a_0 - b_0 + x_0 = 0$ . However  $x_0 \notin C$ , since  $a \neq b$  for  $A \cap B = \emptyset$ .

Consider the Minkowski functional associated to  $C$ , i.e.

$$p_C(x) = \inf \left\{ t > 0 \mid \frac{x}{t} \in C \right\}$$

Then  $p_C(x) \leq k\|x\|_X$  for some  $k > 0$ . Moreover  $p$  is sub-linear, and  $p_C(z) < 1$  if  $z \in C$ .

Now define  $M = \{tx_0 \mid t \in \mathbb{R}\}$ ,  $\ell(tx_0) = t$ . This is linear, and for  $t \geq 0$ ,

$$\ell(tx_0) = t \leq tp(x_0) = p(tx_0),$$

and for  $t < 0$ ,  $\ell(tx_0) = t \leq p(tx_0)$ . So  $\ell \leq p$  on  $M$ . By the Hahn-Banach theorem, there exists  $\Lambda : X \rightarrow \mathbb{R}$  linear such that (using our condition on both  $x$  and  $-x$ )

$$-k\|x\|_X \leq -p_C(x) \leq \Lambda(x) \leq p_C(x) \leq k\|x\|_X.$$

Hence  $\Lambda \in X'$ . Now let  $a \in A, b \in B$ . Then

$$\begin{aligned} \Lambda(a) - \Lambda(b) + 1 &= \Lambda(a) - \Lambda(b) + \Lambda(x_0) = \Lambda(a - b + x_0) \\ &\leq p_C(a - b + x_0) < 1. \end{aligned}$$

So we see that  $\Lambda(a) < \Lambda(b)$  for all  $a \in A, b \in B$ . Since  $\Lambda(A)$  is open in  $\mathbb{R}$ ,

$$\Lambda(a) < \sup \Lambda(A) = \gamma \leq \Lambda(b).$$

If  $B$  is open, we may repeat this argument.

(ii) If  $A$  is compact,  $B$  is closed, then  $\|A - B\|_X = d > 0$ . Now if  $B_{d/2} = \{z \mid \|z\|_X \leq d/2\}$ , then  $A + B_{d/2}$  is open, and still disjoint of  $B$ , so (i) applies.

Since  $\Lambda(A)$  is compact in  $\mathbb{R}$ , it is a strict subset of  $\Lambda(A + B_{d/2})$  open in  $\mathbb{R}$ , so the result follows.

### 3 Generalised Functions and Distributions

The idea of this section is as follows. We want to take a vector space

$$X \subseteq \bigcap_{q \geq 1} L^q(\Omega),$$

for  $\Omega \subseteq \mathbb{R}^n$  open. Then any  $f \in L^p(\Omega)$  defines an action

$$g \xrightarrow{T_f} \int_{\Omega} f(x)g(x) \, dx,$$

for  $g \in X$ , and  $T_f \in X'$  in an appropriate topology. If  $X$  contains  $C_c^\infty(\Omega)$ , then suppose  $T_f(g) = 0$  for all  $g \in C_c^\infty$ . We can pick  $g(y) = \phi_\varepsilon(x - y)$ , where  $\phi_\varepsilon(y) = \varepsilon^{-n} \phi(\frac{y}{\varepsilon})$ , so

$$T_f(\phi_\varepsilon) = f * \phi_\varepsilon(x) \approx f(x) = 0 \text{ almost everywhere.}$$

In this way we can inject  $L^p(\Omega)$  into  $X'$ , but restricting  $X$  will entail that  $X'$  contains much more than just  $L^p$ .

Recall  $C^\infty(\Omega)$  for  $\Omega \subseteq \mathbb{R}^k$  open is the set of infinitely differentiable maps  $f : \Omega \rightarrow \mathbb{R}$  (or  $\mathbb{C}$ ). On  $C^\infty(\Omega)$ , we can define seminorms

$$p_N(\phi) = \sup_{x \in K_N} \max_{|\alpha| \leq N} |D^\alpha \phi(x)|,$$

for  $N \in \mathbb{N}$  and  $K_i \subseteq K_{i+1} \subseteq \Omega$  compact such that  $K_i \uparrow \Omega$ . Then  $\tau_{\mathcal{P}}$  for  $\mathcal{P} = \{p_N \mid N \in \mathbb{N}\}$  defines a locally convex topology on  $C^\infty(\Omega)$ , and we denote

$$\mathcal{E}(\Omega) = (C^\infty(\Omega), \tau_{\mathcal{P}}),$$

a Fréchet space. Note that  $\mathcal{E}(\Omega)$  contains unbounded functions, so we now consider

$$C_c^\infty(\Omega) = \{f \in C^\infty(\Omega) \mid f \text{ has compact support}\}.$$

We will state but not prove the following:

**Theorem 3.1.** *There exists a topology  $\tau$  on  $C_c^\infty(\Omega)$  such that:*

- (i) *Vector space operations are continuous for  $\tau$ .*
- (ii)  *$\phi_j \in C_c^\infty(\Omega)$  such that  $\phi_j \rightarrow 0$  for  $\tau$  if there exists  $K \subseteq \Omega$  compact such that*

$$\sup_{x \in K} |D^\alpha \phi_j(x)| \rightarrow 0$$

*for all  $0 \leq |\alpha| < \infty$  and such that  $\text{supp}(\phi_j) \subseteq K$  for all  $j$ .*

- (iii) If  $T : C_c^\infty \rightarrow \mathbb{R}$  (or  $\mathbb{C}$  or any locally convex TVS) is linear, then  $T$  is continuous for  $\tau$  if and only if for all  $\phi_j \rightarrow 0$  for  $\tau$ , we have  $T(\phi_j) \rightarrow 0$  as  $j \rightarrow \infty$ .

We define

$$\mathcal{D}(\Omega) = (C_c^\infty(\Omega), \tau).$$

This is the space of *test functions*.

**Example 3.1.**

The functions  $e^{-j}\phi(jx)$  for  $x \in \Omega$  converge to 0 in  $\mathcal{D}(\Omega)$ , but  $j^{-2024}\phi(jx)$  does not converge to 0 in  $\mathcal{D}(\Omega)$ .

If  $\Omega = \mathbb{R}$ , then  $e^{-j}\phi(x-j)$  for  $x \in \Omega$  does not converge in  $\mathcal{D}(\Omega)$ .

When  $\Omega = \mathbb{R}^n$ , we say  $\phi \in C^\infty(\mathbb{R}^n)$  is *rapidly decreasing* if

$$\sup_{x \in \mathbb{R}^n} (1 + |x|)^N |D^\alpha \phi(x)| < \infty$$

for all  $\alpha$  with  $|\alpha| \geq 0$  and all  $N$ .

**Example 3.2.**

$(1 + |x|)^{-2024}$  is not rapidly decreasing, but  $e^{-|x|^2}$  is.

We define seminorms

$$p_N(\phi) = \sup_{x \in \mathbb{R}^n} \max_{|\alpha| \leq N} (1 + |x|)^N |D^\alpha \phi(x)|,$$

for  $N \in \mathbb{N}$ , and let

$$\mathcal{S}(\mathbb{R}^n) = (\{\phi \text{ rapidly decreasing}\}, \tau_{\tilde{\mathcal{P}}}),$$

where  $\tilde{\mathcal{P}} = \{p_N \mid N \in \mathbb{N}\}$ . This is the *Schwartz class*, another Fréchet space. The set inclusions

$$\mathcal{D}(\mathbb{R}^n) \subset \mathcal{S}(\mathbb{R}^n) \subset \mathcal{E}(\mathbb{R}^n)$$

are clear, and one shows further that the embeddings are continuous for the corresponding topologies. Similarly,  $\mathcal{D}(\Omega) \subset \mathcal{E}(\Omega)$ . We define

$$\mathcal{D}'(\Omega) = \{T : \mathcal{D}(\Omega) \rightarrow \mathbb{R} \text{ (or } \mathbb{C}) \text{ linear, continuous}\},$$

the space of *Schwartz distributions*. Moreover we set

$$\mathcal{S}'(\mathbb{R}^n) = \{T : \mathcal{S}(\mathbb{R}^n) \rightarrow \mathbb{R} \text{ (or } \mathbb{C}) \text{ linear, continuous}\}$$

the space of ‘tempered’ distributions. For instance, if  $f(x) = e^{|x|^2}$  on  $\Omega = \mathbb{R}$ , then  $T_f \in \mathcal{D}'(\mathbb{R})$ , but not in  $\mathcal{S}'(\mathbb{R})$ . Finally

$$\mathcal{E}'(\mathbb{R}^n) = \{T : \mathcal{E}(\mathbb{R}^n) \rightarrow \mathbb{R} \text{ (or } \mathbb{C}) \text{ linear, continuous}\},$$

which will be shown to equal the elements of  $\mathcal{D}'(\Omega)$  of “compact support”. Reversing the set inclusions, we obtain

$$\mathcal{E}'(\mathbb{R}^n) \subset \mathcal{S}'(\mathbb{R}^n) \subset \mathcal{D}'(\mathbb{R}^n),$$

and  $\mathcal{E}'(\Omega) \subset \mathcal{D}'(\Omega)$ , and these embeddings are continuous if we equip  $\mathcal{E}'$ ,  $\mathcal{S}'$  and  $\mathcal{D}'$  with their weak-\* topologies, where

$$T_j \rightarrow T \iff T_j(\phi) \rightarrow T(\phi) \quad \forall \phi \in \mathcal{E}, \mathcal{S}, \mathcal{D}.$$

### Example 3.3.

Consider the Dirac measure  $\delta_x$ . Clearly if  $\phi_j \rightarrow 0$  in  $\mathcal{D}(\Omega)$ , and  $x \in \Omega$ , then

$$\delta_x(\phi_j) = \phi_j(x) \rightarrow 0$$

for all  $x$ , so  $\delta_x \in \mathcal{D}'(\Omega)$ . Similarly, for any  $\alpha$ ,

$$D^\alpha \delta_x(\phi_j) = D^\alpha \phi_j(x) \rightarrow 0$$

as  $j \rightarrow \infty$ , if  $\phi \rightarrow 0$  in  $\mathcal{D}(\Omega)$ , so all derivatives of the Dirac measure lie in  $\mathcal{D}'(\Omega)$  too.

Moreover, for any locally integrable  $f \in L^1_{\text{loc}}(\Omega)$ , we get a distribution  $T_f \in \mathcal{D}'(\Omega)$  via

$$T_f(g) = \int_{\Omega} f(x)g(x) \, dx,$$

for  $g \in \mathcal{D}(\Omega)$ , since for any  $\phi_j \rightarrow 0$  in  $\mathcal{D}(\Omega)$ , we have

$$T_f(\phi_j) = \int_{\Omega} f(x)\phi_j(x) \, dx \rightarrow 0$$

by DCT with dominating function  $|f| \mathbb{1}_K \sup \|\phi_j\|_{\infty} \in L^1(\Omega)$ . Moreover  $L^1_{\text{loc}}(\Omega) \subset \mathcal{D}'(\Omega)$  is an injection since testing with  $g(y) = \phi_{\varepsilon}(x - y)$  for  $x \in \Omega$ , then

$$T_f(\phi_{\varepsilon}) = f * \phi_{\varepsilon}(x) \rightarrow f(x),$$

so if  $T_f(g) = 0$  for all  $g$ , then  $f = 0$  almost everywhere on  $\Omega$ .

In fact for  $\phi_\varepsilon$  as above and  $g \in \mathcal{D}(\Omega)$ , we have

$$T_{\phi_\varepsilon}(g) = \phi_\varepsilon * g(0) \rightarrow g(0)$$

for all  $g \in \mathcal{D}(\Omega)$ , if  $0 \in \Omega$ . So  $\phi_\varepsilon \rightarrow \delta_0$  as  $\varepsilon \rightarrow 0$  in  $\mathcal{D}'(\Omega)$  (weak-\* topology!)

### 3.1 Operations on Distributions

It is clear that for  $T_1, T_2 \in \mathcal{D}'(\Omega)$  and  $\alpha \in \mathbb{C}$  (or  $\mathbb{R}$ ), that  $\alpha T_1 + T_2 \in \mathcal{D}'(\Omega)$ .

If  $f \in L^1_{\text{loc}}(\Omega)$  and  $a \in C^\infty(\Omega)$ , then  $af \in L^1_{\text{loc}}(\Omega)$ , so we can define

$$T_{af}(\varphi) = \int_{\Omega} a(x)f(x)\varphi(x) \, dx = T_f(a\varphi),$$

for all  $\varphi \in \mathcal{D}(\Omega)$ . This motivates the definition  $(aT)(\varphi) = T(a\varphi)$ , for all  $T \in \mathcal{D}'$ ,  $a \in C^\infty$  and  $\varphi \in \mathcal{D}(\Omega)$ , the product of  $T$  with  $a \in \mathcal{D}'(\Omega)$ .

If  $f \in C^1(\Omega)$ , then  $D_i f \in L^1_{\text{loc}}(\Omega)$ , so

$$T_{D_i f}(\varphi) = \int_{\Omega} D_i f(x)\varphi(x) \, dx \stackrel{IBP}{=} - \int_{\Omega} f(x)D_i \varphi(x) \, dx = -T_f(D_i \varphi).$$

Therefore for an arbitrary  $T \in \mathcal{D}'(\Omega)$  and  $\alpha$  any multi-index, we define

$$D^\alpha T(\varphi) = (-1)^{|\alpha|} T(D^\alpha \varphi),$$

for  $\varphi \in \mathcal{D}(\Omega)$ . In this sense, a Schwartz distribution is infinitely differentiable. One has to be careful, however as these are not the classical derivatives we are used to. If  $T = T_f$  and  $D^\alpha T = T_g$  for some  $f, g \in L^1_{\text{loc}}$ , then we say that  $g = D_w^\alpha f$  is the *weak (or distributional) derivative* of  $f$ .

#### Example 3.4.

Consider the ramp function

$$f(x) = \begin{cases} 0 & x < 0 \\ x & x \geq 0 \end{cases} \in L^1_{\text{loc}}(\mathbb{R}).$$

Then we have

$$DT_f(\varphi) = -T_f(D\varphi) = - \int_0^\infty x D\varphi(x) \, dx = \int_0^\infty \varphi(x) \, dx = T_H(\varphi),$$

where  $H = \mathbb{1}_{[0, \infty)}$  is the Heaviside function. So  $H$  is the weak derivative of  $f$ .



Note that  $H$  is not the classical derivative of  $f$ . Moreover,

$$D^2T_f(\varphi) = DT_H(\varphi) = -T_H(D\varphi) = -\int_0^\infty D\varphi(x) \, dx = \phi(0) = \delta_0(\phi).$$

So  $DT_H = \delta_0$  in  $\mathcal{D}'(\mathbb{R})$ , but  $H'(x) = \delta_0(x)$  is nonsense!

We want to define the notion of convolution, as we have seen it is useful and correspond to pointwise products in the Fourier space. However it turns out we can only take convolutions with distributions which have compact support, which we will need to define.

**Proposition 3.1.** *A linear map  $T : \mathcal{E}(\Omega) \rightarrow \mathbb{R}$  or  $\mathbb{C}$  is continuous if and only if there exists  $K \subseteq \Omega$  compact and  $N \in \mathbb{N}$ ,  $C > 0$  such that*

$$|T(\phi)| \leq C \sup_{x \in K} \max_{|\alpha| \leq N} |D^\alpha \phi(x)|, \quad (\dagger)$$

for all  $\phi \in \mathcal{E}(\Omega)$ .

This is similar to the result we have for normed spaces, except here our topology is not induced from a norm.

**Proof:** First, if the above statement holds, take  $\phi_j \rightarrow 0$  in  $\mathcal{E}(\Omega)$ . Then we can take  $K_j \supseteq K$  such that  $D^\alpha \phi_j \rightarrow 0$  uniformly on  $K_j$  for all  $\alpha$  of  $|\alpha| \leq j$ . So the right hand side converges to zero as  $j \rightarrow \infty$ , and hence  $T(\phi_j) \rightarrow 0$  as  $j \rightarrow \infty$ , so  $T$  is continuous.

Conversely, let  $T$  be continuous, and assume that the inequality does not hold. Then for any exhaustion of compact sets  $K_j \subseteq K_{j+1}$  and  $K_j \uparrow \Omega$ , there exists  $\phi_j \in \mathcal{E}(\Omega)$  such that

$$|T(\phi_j)| \geq j \sup_{x \in K_j} \max_{|\alpha| \leq j} |D^\alpha \phi_j(x)|.$$

Define  $\psi_j = \phi_j / |T(\phi_j)| \in \mathcal{E}(\Omega)$ . Then for any multi-index  $\beta$ ,

$$|D^\beta \psi_j(x)| \leq \frac{|D^\beta \phi_j(x)|}{j \sup_{x \in K_j} \max_{|\alpha| \leq j} |D^\alpha \phi_j(x)|} \leq \frac{1}{j}$$

for  $j$  such that  $j \geq |\beta|$ . Thus  $\psi_j \rightarrow 0$  in  $\mathcal{E}(\Omega)$ , but

$$|T(\psi_j)| = \frac{|T(\phi_j)|}{|T(\phi_j)|} = 1 \not\rightarrow 0,$$

so  $T$  is not continuous.

**Definition 3.1.** A distribution  $T \in \mathcal{D}'(\Omega)$  is supported in a closed set  $C \subseteq \Omega$  if  $T(\phi) = 0$  for all  $\phi \in C_c^\infty(\Omega \setminus C)$ . We say  $T \in \mathcal{D}'(\Omega)$  is *compactly supported* if there exists  $K \subseteq \Omega$  compact such that  $T$  is supported in  $K$ .

**Example 3.5.**

1. If  $f \in L^1_{\text{loc}}(\Omega)$  vanishes outside a closed set  $C \subseteq \Omega$ , then  $T_f$  is supported in  $C$ .
2.  $\delta_x$  for  $x \in \Omega$  is compactly supported in  $\{x\}$ .
3. If  $T$  is supported in  $C$ , then so is  $D^\alpha T$  for all  $\alpha$ .

We now want to show the compactly supported functions are the dual space of some compact subset of  $\Omega$ .

**Theorem 3.2.** Any  $T \in \mathcal{E}'(\Omega)$  defines an element in  $\mathcal{D}'(\Omega)$  of compact support. Moreover, every  $T \in \mathcal{D}'(\Omega)$  that has compact support extends to  $T \in \mathcal{E}'(\Omega)$ .

**Proof:** By the previous proposition, if  $T \in \mathcal{E}'(\Omega)$ , there exists  $K \subseteq \Omega$  such that the right hand side of (†) vanishes for all  $\phi \in C_c^\infty(\Omega \setminus K)$ , so  $T(\phi) = 0$  for such  $\phi$ , and  $T$  has support in  $K$ .

Conversely, if  $T$  is supported in  $K \subseteq \Omega$  compact, take  $\xi \in C_c^\infty(\Omega)$  such that  $\xi = 1$  on  $K$ , and define

$$\tilde{T}(\phi) = T(\xi\phi),$$

for  $\phi \in C_c^\infty(\Omega)$ , which is well-defined, and  $\tilde{T} \in \mathcal{E}'(\Omega)$ .

### 3.2 Convolutions of Distributions

Recall that for  $f, g \in C_c(\mathbb{R}^n)$ ,

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

for  $x \in \mathbb{R}^n$ . If  $f \in L^1$ ,  $g \in L^p$ , this extends for almost all  $x \in \mathbb{R}^n$  by Minkowski's integral inequality. Recall  $\tau_x\phi(y) = \phi(y - x)$ , and set  $\widetilde{\phi}(y) = \phi(y)$ . Then  $\phi(x - y) = \widetilde{\tau_x\phi}(y)$ . (this latex is horrible). Then if  $f \in L^1_{\text{loc}}(\mathbb{R}^n)$  and  $\phi \in \mathcal{D}(\mathbb{R}^n)$ ,

$$\int f(y)\phi(x - y) dy = T_f(\widetilde{\tau_x\phi}),$$

so we define, for  $T \in \mathcal{D}'(\mathbb{R}^n)$  and  $\phi \in \mathcal{D}(\mathbb{R}^n)$ , the *convolution* as

$$T * \phi(x) = T(\widetilde{\tau_x\phi}),$$

for  $x \in \mathbb{R}^n$ . If  $T = \delta_0$ , then

$$\delta_0(\widetilde{\tau_x \phi}) = \int \phi(x - y) d\delta_0(y) = \phi(x),$$

so  $\delta_0 * \phi = \phi$  for all  $\phi \in \mathcal{D}(\mathbb{R}^n)$ .

**Theorem 3.3.** *Let  $T \in \mathcal{D}'(\mathbb{R}^n)$ ,  $\phi \in \mathcal{D}(\mathbb{R}^n)$ . Let  $\alpha$  be any multi-index. Then  $T * \phi \in C^\infty(\mathbb{R}^n)$ , and*

$$D^\alpha(T * \phi) = (D^\alpha T) * \phi = T * (D^\alpha \phi).$$

If  $T \in \mathcal{E}'(\mathbb{R}^n)$ , then  $T * \phi \in \mathcal{D}(\mathbb{R}^n)$ .

**Proof:** For  $e_i$  a basis vector in  $\mathbb{R}^n$  and  $h \rightarrow 0$ ,

$$\frac{1}{h}(T * \phi(x + he_i) - T * \phi(x)) = T \left( \frac{1}{h}(\widetilde{\tau_{x+he_i} \phi} - \widetilde{\tau_x \phi}) \right),$$

but this term is

$$\frac{\phi(x + he_i - \cdot) - \phi(x - \cdot)}{h} \rightarrow D_i \phi(x - \cdot)$$

in  $\mathcal{D}(\mathbb{R}^n)$ , so the right hand side converges to

$$T(\widetilde{\tau_x D_i \phi}) = T * D_i \phi(x).$$

Similarly,  $\tau_x$  is continuous for  $\mathcal{D}(\mathbb{R}^n)$ , and so by iterating  $T * D_i \phi$  exists and is continuous on  $\mathbb{R}^n$ , hence  $T * \phi \in C^\infty(\mathbb{R}^n)$ , and  $D^\alpha(T * \phi) = T * (D^\alpha \phi)$ . Note

$$D^\alpha(\widetilde{\tau_x \phi}) = D^\alpha(\phi(x - y)) = (-1)^{|\alpha|}(D^\alpha \phi)(x - y) = (-1)^{|\alpha|}\widetilde{\tau_x D^\alpha \phi}(y),$$

so

$$(D^\alpha T) * \phi(x) = D^\alpha T[\widetilde{\tau_x \phi}] = (-1)^{|\alpha|}T[D^\alpha \widetilde{\tau_x \phi}] = T[\widetilde{\tau_x D^\alpha \phi}] = T * (D^\alpha \phi),$$

which gives the identity. Finally if  $T \in \mathcal{E}'(\mathbb{R}^n)$  is supported in  $K$  compact, then  $\text{supp}(\widetilde{\tau_x \phi})$  is, for  $|x|$  large enough, disjoint from  $K$ , and so

$$T * \phi = T(\widetilde{\tau_x \phi}) = 0,$$

for such  $x$ , and so  $T * \phi \in \mathcal{D}(\mathbb{R}^n)$ .

**Definition 3.2.** For  $T_1 \in \mathcal{D}'(\mathbb{R}^n)$ ,  $T_2 \in \mathcal{E}'(\mathbb{R}^n)$ , we define  $T_1 * T_2$  as the distribution in  $\mathcal{D}'(\mathbb{R}^n)$  satisfying, for all  $\phi \in \mathcal{D}(\mathbb{R}^n)$ ,

$$(T_1 * T_2) * \phi = T_1 * (T_2 * \phi).$$

Recall from the previous theorem that  $T_2 * \phi \in \mathcal{D}(\mathbb{R}^n)$ . Note that taking  $x = 0$  in  $(T_1 * T_2)(\widetilde{\tau_x \phi})$  in particular defines  $T_1 * T_2(\phi)$  for all  $\phi \in \mathcal{D}(\mathbb{R}^n)$ . Note also that this definition is not commutative.

Moreover since  $\delta_0 \in \mathcal{E}'(\mathbb{R}^n)$ , we have

$$(T_1 * \delta_0) * \phi = T_1 * (\delta_0 * \phi) = T_1 * \phi,$$

so  $(\cdot) * \delta_0$  is an identity operation on  $\mathcal{D}'(\mathbb{R}^n)$ . Moreover we can show that if  $T_1, T_2 \in \mathcal{E}'(\mathbb{R}^n)$ , then  $T_1 * T_2 = T_2 * T_1$ .

**Theorem 3.4.** Let  $T_1 \in \mathcal{D}'(\mathbb{R}^n)$ ,  $T_2 \in \mathcal{E}'(\mathbb{R}^n)$ , and  $\alpha$  any multi-index. Then

$$D^\alpha(T_1 * T_2) = (D^\alpha T_1) * T_2 = T_1 * (D^\alpha T_2).$$

This last thing is well defined as  $D^\alpha T_2 \in \mathcal{E}'(\mathbb{R}^n)$ .

**Proof:** Using the previous theorem, for all  $\varphi \in \mathcal{D}(\mathbb{R}^n)$ ,

$$\begin{aligned} D^\alpha(T_1 * T_2) * \varphi &= (T_1 * T_2) * D^\alpha \varphi = T_1 * (T_2 * D^\alpha \varphi) \\ &= T_1 * (D^\alpha T_2 * \varphi) = (T_1 * D^\alpha T_2) * \varphi, \end{aligned}$$

and the second identity is proven similarly.

### 3.3 Fundamental Theorem of PDEs

Consider a partial differential operator

$$L = \sum_{|\alpha| \leq k} a_\alpha D^\alpha,$$

where  $a_\alpha \in C^\infty(\mathbb{R}^n)$ , and  $k \in \mathbb{N}$ . This acts on  $\mathcal{D}'(\mathbb{R}^n)$ , and we can thus consider solving a PDE  $Lu = u_0$ , for  $u, u_0 \in \mathcal{D}'(\mathbb{R}^n)$ . This means we solve  $Lu(\phi) = u_0(\phi)$  for all  $\phi \in \mathcal{D}(\mathbb{R}^n)$ , sometimes called a ‘weak solution’.

A distribution  $\mathcal{G} \in \mathcal{D}'(\mathbb{R}^n)$  is called a *fundamental solution* for  $L$  if  $L\mathcal{G} = \delta_0$  in  $\mathcal{D}'(\mathbb{R}^n)$ . If  $\mathcal{G} = T_g$  for  $g \in L^1_{\text{loc}}(\mathbb{R}^n)$ , we call  $g$  the *Green function* (or kernel) of  $\mathcal{G}$ .

**Theorem 3.5.** Suppose  $L$  has constant coefficients  $a_\alpha \in \mathbb{R}$  or  $\mathbb{C}$ , and that  $\mathcal{G} \in \mathcal{D}'(\mathbb{R}^n)$  is a fundamental solution of  $L$ .

Then for any  $u_0 \in \mathcal{E}'(\mathbb{R}^n)$ , a solution to the PDE  $Lu = u_0$  in  $\mathcal{D}'(\mathbb{R}^n)$  is given by

$$u = \mathcal{G} * u_0.$$

*Remark.* If  $u_0 \in \mathcal{D}(\mathbb{R}^n)$ , then we know  $G * u_0 \in C^\infty(\mathbb{R}^n)$  is a classical solution of the PDE.

**Proof:** We have from the previous theorem that

$$\begin{aligned} Lu = L(\mathcal{G} * u_0) &= \sum_{|\alpha| \leq k} a_\alpha D^\alpha (\mathcal{G} * u_0) = \sum_{|\alpha| \leq k} a_\alpha (D^\alpha \mathcal{G} * u_0) \\ &= \left( a_\alpha \sum_{|\alpha| \leq k} D^\alpha \mathcal{G} \right) * u_0 = (L\mathcal{G}) * u_0 \\ &= \delta_0 * u_0 = u_0 * \delta_0 = u_0. \end{aligned}$$

### Example 3.6.

On  $\mathbb{R}^3$ , consider the negative Laplacian

$$L = -\Delta = -\sum_{i=1}^3 \frac{\partial^2}{\partial x_i^2}.$$

One can show that a fundamental solution  $\mathcal{G}$  to  $-\Delta u = \delta_0$  is given by  $T_g$  with

$$g(x) = \frac{1}{4\pi|x|}.$$

Therefore the solutions  $u$  to the Poisson equation  $Lu = f$  for  $f \in \mathcal{D}(\mathbb{R}^n)$  are given by

$$u(x) = \mathcal{G} * f(x) = \frac{1}{4\pi} \int_{\mathbb{R}^n} \frac{f(y)}{|x-y|} dy \in C^\infty(\mathbb{R}^n).$$

This is explored more in Schauder theory.

## 4 Fourier Transforms of Distributions

For  $f \in L^1(\mathbb{R}^n)$ , we define

$$\hat{f}(u) = \int_{\mathbb{R}^n} f(x) e^{-ix \cdot u} dx.$$

So  $F[f](u) = \hat{f}(u)$  defines a map from  $\mathbb{R}^n \rightarrow \mathbb{C}$ , so all  $L^p(\mathbb{R}^n)$  are now complex valued.

**Lemma 4.1** (Riemann-Lebesgue). *Let  $f \in L^1(\mathbb{R}^n)$ . Then*

$$\hat{f} \in C_0(\mathbb{R}^n) = \{h : \mathbb{R}^n \rightarrow \mathbb{C} \text{ continuous} \mid \lim_{|u| \rightarrow \infty} |\hat{h}(u)| = 0\},$$

and  $\|\hat{f}\|_\infty \leq \|f\|_{L^1}$ .

**Proof:** Taking  $u_j \rightarrow u$  in  $\mathbb{R}^n$ , then DCT implies

$$\hat{f}(u_j) = \int_{\mathbb{R}^n} f(x) e^{-ix \cdot u_j} dx \rightarrow \int_{\mathbb{R}^n} f(x) e^{-ix \cdot u} dx = \hat{f}(u),$$

so  $\hat{f}$  is continuous, and  $\|\hat{f}\|_\infty \leq \|f\|_{L^1}$ . For  $f \in L^1(\mathbb{R}^n)$ , we take  $f_k \in C_c^\infty(\mathbb{R}^n)$  such that  $f_k \rightarrow f$  in  $L^1(\mathbb{R}^n)$  as  $k \rightarrow \infty$ . So by this inequality,  $\hat{f}_k \rightarrow \hat{f}$  in the supremum norm. Since  $(C_0(\mathbb{R}^n), \|\cdot\|_\infty)$  is a Banach space, it suffices to show that  $\hat{f}_k \in C_0(\mathbb{R}^n)$  for all  $k$ .

For this, recall (or see next lecture) that, for all  $j = 1, \dots, n$ , and  $D_j = \partial/\partial x_j$ ,

$$|u_j| |\hat{f}_k(u)| = |\widehat{D_j f_k}(u)| \leq \|D_j f_k\|_{L^1} < \infty,$$

so  $\hat{f}_k \in C_0(\mathbb{R}^n)$ .

**Lemma 4.2.** *Let  $f \in L^1(\mathbb{R}^n)$ , and  $f_\lambda(x) = \lambda^{-n} f(\frac{x}{\lambda})$ , for  $x \in \mathbb{R}^n$ ,  $\lambda > 0$ . Then*

$$\hat{f}_\lambda(u) = \hat{f}(u\lambda).$$

Also, for  $\tau_y f(x) = f(x - y)$ , we have

$$\widehat{\tau_y f}(u) = e^{-iu \cdot y} \hat{f}(u), \quad \widehat{e^{i\langle \cdot, y \rangle} f}(u) = \tau_y \hat{f}(u),$$

for  $u \in \mathbb{R}^n$ .

Moreover if  $f, g \in L^1(\mathbb{R}^n)$ , then  $f * g \in L^1(\mathbb{R}^n)$ , and

$$\widehat{f * g}(u) = \hat{f}(u) \hat{g}(u),$$

for all  $u \in \mathbb{R}^n$ .

**Proof:** This is done by an integral substitution, and also Fubini's theorem. For example,

$$\widehat{\tau_y f}(y) = \int_{\mathbb{R}^n} \underbrace{f(x-y)}_{=z} e^{-ix \cdot u} dx = e^{-iy \cdot u} \int_{\mathbb{R}^n} f(z) e^{-iz \cdot u} dz = e^{-iu \cdot y} \hat{f}(u).$$

The others are left as an exercise.

**Theorem 4.1.**

(i) Let  $f \in C^1(\mathbb{R}^n)$ , and  $f, D_j f \in L^1(\mathbb{R}^n)$  for  $j = 1, \dots, n$ . Then

$$\widehat{D_j f}(u) = i u_j \hat{f}(u),$$

for all  $u = (u_1, \dots, u_n) \in \mathbb{R}^n$ .

(ii) Assume  $f \in L^1(\mathbb{R}^n)$  such that

$$\int_{\mathbb{R}^n} (1 + |x|) |f(x)| dx < \infty.$$

Then  $\hat{f} \in C^1(\mathbb{R}^n)$ , and

$$D_j \hat{f}(u) = -i \widehat{x_j f(x)}(u),$$

for all  $j = 1, \dots, n$ , and  $u \in \mathbb{R}^n$ .

**Proof:**

(i) Take  $f_\varepsilon \in C_c^\infty(\mathbb{R}^n)$ , such that  $\|f - f_\varepsilon\|_{L^1} + \|D_j f - D_j f_\varepsilon\|_{L^1} < \varepsilon$ .

This is possible by first restricting to  $f, D_j f$  of compact support. Take  $f\chi$ , where  $\chi \in C_c^\infty(\mathbb{R}^n)$ , and  $\chi = 1$  on  $\{|x| \leq R\}$  for  $R$  large, then using an approximation identity  $\phi_\varepsilon * f$ , and note

$$D_j(\phi_\varepsilon * f) = \phi_\varepsilon * D_j f \rightarrow D_j f.$$

Now for such  $f_\varepsilon$ , we use integration by parts:

$$\begin{aligned} \widehat{D_j f_\varepsilon}(u) &= \int_{\mathbb{R}^n} D_j f_\varepsilon(x) e^{-ix \cdot u} dx = - \int_{\mathbb{R}^n} f_\varepsilon(x) D_j(e^{-ix \cdot u}) dx \\ &= \int_{\mathbb{R}^n} f_\varepsilon(x) (i u_j) e^{-ix \cdot u} dx = i u_j \hat{f}_\varepsilon(u). \end{aligned}$$

We conclude that

$$\begin{aligned} |\widehat{D_j f}(u) - i u_j \hat{f}(u)| &\leq |\widehat{D_j f}(u) - \widehat{D_j f_\varepsilon}(u)| + |u_j(\hat{f}_\varepsilon(u) - \hat{f}(u))| \\ &\leq \|D_j f - D_j f_\varepsilon\|_{L^1} + |u| \|f - f_\varepsilon\|_{L^1} \leq (1 + |u|)\varepsilon. \end{aligned}$$

(ii) Since  $x_j f(x) \in L^1(\mathbb{R}^n)$ , then  $\widehat{x_j f(x)}(u)$  is continuous for all  $u$ , so it suffices to prove the identity, as then  $\hat{f} \in C^1(\mathbb{R}^n)$ .

Take  $e_j$  to be the  $j$ 'th basis vector in  $\mathbb{R}^n$ , and consider for  $h \rightarrow 0$ ,

$$\begin{aligned} \frac{1}{h}(\hat{f}(u + h e_j) - \hat{f}(u)) &= \int_{\mathbb{R}^n} \frac{1}{h}(e^{-ix \cdot (u + h e_j)} - e^{-ix \cdot u}) f(x) dx \\ &= \int_{\mathbb{R}^n} e^{-ix \cdot u} \left( \frac{e^{-ix \cdot h e_j} - 1}{h} \right) f(x) dx \\ &\rightarrow \int_{\mathbb{R}^n} e^{-ix \cdot u} (-ix_j) f(x) dx \\ &= -ix_j \widehat{f(x)}(u), \end{aligned}$$

by using DCT. Indeed, the quotient term is equal to  $|e^{-ix \cdot \tilde{h}}| |x_j|$  by MVT, which is dominated.

Iterating this, we see that

$$\widehat{D^\alpha f}(u) = i^{|\alpha|} u^\alpha \hat{f}(u),$$

where we denote

$$u^\alpha = u_1^{\alpha_1} \cdots u_n^{\alpha_n}.$$

**Theorem 4.2** (Fourier Inversion). *Let  $f \in L^1(\mathbb{R}^n)$  be such that  $\hat{f} \in L^1(\mathbb{R}^n)$ . Then for almost all  $x \in \mathbb{R}^n$ ,*

$$f(x) = \frac{1}{(2\pi)^n} \int_{\mathbb{R}^n} \hat{f}(u) e^{iu \cdot x} du = F^{-1}[\hat{f}],$$

*and this holds for all  $x$  if  $f \in [f]$  is the unique continuous selection.*

*Moreover,  $f = F F^{-1}(f)$  almost everywhere.*

**Proof:** The same proof from Probability and Measure, albeit with  $u \cdot x$  replaced with  $-u \cdot x$ .

**Corollary 4.1.**  $F : L^1(\mathbb{R}^n) \rightarrow C_0(\mathbb{R}^n)$  is injective.



**Proof:** If  $\hat{f} = 0$ , then  $f = F^{-1}F(f) = F^{-1}[0] = 0$ .

**Theorem 4.3.**  $F$  realises a continuous homeomorphism of

$$\mathcal{S}(\mathbb{R}^n) = \{\phi \in C^\infty(\mathbb{R}^n) \mid \sup_{x \in \mathbb{R}^n} |(1 + |x|)^N D^\alpha \phi(x)| < \infty \text{ for all } \alpha, N\}.$$

**Proof:** First note for  $f \in \mathcal{S}(\mathbb{R}^n)$ , then

$$\|f\|_{L^1} \leq \sup_{x \in \mathbb{R}^n} (1 + |x|)^{N+1} |f(x)| \cdot \int_{\mathbb{R}^n} \frac{1}{(1 + |x|)^{N+1}} dx < \infty.$$

Similarly, for all  $\alpha, \beta$ ,  $D^\alpha(x^\beta f(x))$ , for  $f \in \mathcal{S}(\mathbb{R}^n)$ , lies in  $L^1$ , so its Fourier transform  $\hat{f}$  verifies, for all  $\alpha, \beta$  multi-derivatives,

$$|u|^\alpha |D^\beta \hat{f}(u)| = |D^\alpha \widehat{(x^\beta f(x))}(u)| \leq \|D^\alpha(x^\beta f)\|_{L^1} < \infty,$$

so  $\hat{f} \in \mathcal{S}(\mathbb{R}^n)$ . Now if  $\phi_j \rightarrow 0$  in  $\mathcal{S}(\mathbb{R}^n)$ , then

$$|u|^\alpha |D^\beta \hat{\phi}_j(u)| \leq p_N(\phi_j) \rightarrow 0,$$

as  $j \rightarrow \infty$ , so  $\hat{\phi}_j \rightarrow 0$  in  $\mathcal{S}(\mathbb{R}^n)$ .

Finally  $F : \mathcal{S}(\mathbb{R}^n) \rightarrow \mathcal{S}(\mathbb{R}^n)$  is injective by the corollary, and  $F$  is also surjective since any  $\phi \in \mathcal{S}(\mathbb{R}^n)$  equals

$$\phi = F(F^{-1}(\phi)) = F(F(\phi(-\cdot)))$$

is a Fourier transform of  $F(\phi(-\cdot)) \in \mathcal{S}(\mathbb{R}^n)$ .

A function  $f : \mathbb{R}^n \rightarrow \mathbb{R}$  is called *slowly increasing* if

$$\sup_{x \in \mathbb{R}^n} (1 + |x|)^{-N} |f(x)| < \infty.$$

One shows that  $T_f \in \mathcal{S}'(\mathbb{R}^n)$ , and the same is true for  $f \in L^p(\mathbb{R}^n)$ , for  $1 \leq p \leq \infty$ . Then for  $f \in L^1$  so that  $\hat{f} \in L^\infty$ , and

$$\begin{aligned} T_{\hat{f}}(\phi) &= \int_{\mathbb{R}^n} \hat{f}(u) \phi(u) du = \iint e^{-ix \cdot u} f(x) dx \phi(u) du \\ &= \int_{\mathbb{R}^n} \hat{\phi}(x) f(x) dx = T_f(\hat{\phi}), \end{aligned}$$

for  $\phi \in \mathcal{S}(\mathbb{R}^n)$  (by Fubini's). We thus define, for  $T \in \mathcal{S}'(\mathbb{R}^n)$ , the *distributional Fourier transform* via the action

$$FT(\phi) = \hat{T}(\phi) = T(\hat{\phi}),$$

for all  $\phi \in \mathcal{S}(\mathbb{R}^n)$ . Note that  $\hat{\phi} \in \mathcal{S}(\mathbb{R}^n)$ , and if  $\phi_j \rightarrow 0$  in  $\mathcal{S}(\mathbb{R}^n)$ , then by the last theorem  $\hat{\phi}_j \rightarrow 0$  in  $\mathcal{S}(\mathbb{R}^n)$ , so  $\hat{T}(\phi_j) = T(\hat{\phi}_j) \rightarrow 0$  as  $j \rightarrow \infty$ , showing  $\hat{T} \in \mathcal{S}'(\mathbb{R}^n)$ .

Note further that if  $T_j \rightarrow T$  in  $\mathcal{S}'(\mathbb{R}^n)$ , then

$$\hat{T}_j(\phi) = T_j(\hat{\phi}) \rightarrow T(\hat{\phi}) = \hat{T}(\phi)$$

for all  $\phi \in \mathcal{S}(\mathbb{R}^n)$ , so  $F : \mathcal{S}'(\mathbb{R}^n) \rightarrow \mathcal{S}'(\mathbb{R}^n)$  is continuous, for the weak-\* topology.

Define the *inverse Fourier transform* via

$$F^{-1}T(\phi) = T(F^{-1}\phi),$$

and we see that

$$F^{-1}FT(\phi) = FT(F^{-1}\phi) = T(FF^{-1}\phi) = T(\phi),$$

by Fourier inversion on  $\mathcal{S}$ . We conclude the following:

**Theorem 4.4.**  $F : \mathcal{S}'(\mathbb{R}^n) \rightarrow \mathcal{S}'(\mathbb{R}^n)$  is a homeomorphism with inverse  $F^{-1}$ .

Recall Plancherel's theorem: for  $f, g \in L^2 \cap L^1(\mathbb{R}^n)$ , then  $\hat{f}, \hat{g} \in L^2$  and

$$(f, g)_{L^2} = \frac{1}{(2\pi)^n} (\hat{f}, \hat{g})_{L^2}.$$

Then since  $L^1 \cap L^2$  is dense in  $L^2$ , we can extend  $F/(2\pi)^{n/2}$  to an isometry of  $L^2(\mathbb{R}^n)$ , say  $\bar{F}/(2\pi)^{n/2}$ . So if we set

$$\bar{F}T(\phi) = T(\bar{F}\phi) = T(\hat{\phi}) = FT(\phi),$$

we see that  $\bar{F} = F$  on  $\mathcal{S}'(\mathbb{R}^n)$ .

For  $T \in \mathcal{E}'(\mathbb{R}^n)$ , a natural definition of the Fourier transform would be

$$T(e^{-i\langle \cdot, u \rangle}) = \bar{t}(u),$$

for all  $u \in \mathbb{R}^n$ . One shows that  $\bar{t}$  is slowly increasing, and that  $T_{\bar{t}} = FT$  in  $\mathcal{S}'(\mathbb{R}^n)$ .

**Example 4.1.**

Take  $\delta_z$  for  $z \in \mathbb{R}^n$ . Then  $\delta_z \in \mathcal{E}'(\mathbb{R}^n) \subseteq \mathcal{S}'(\mathbb{R}^n)$ , and

$$\hat{\delta}_z(\phi) = \delta_z(\hat{\phi}) = \hat{\phi}(z) = \int_{\mathbb{R}^n} e^{-ix \cdot z} \phi(x) \, dx = T_{e^{-i\langle \cdot, z \rangle}}(\phi),$$

for all  $\phi \in \mathcal{S}$ . So in  $\mathcal{S}'(\mathbb{R}^n)$ ,

$$\hat{\delta}_z = e^{-i\langle \cdot, z \rangle},$$

and hence  $\hat{\delta}_0 = 1$ .

Similarly, one shows

$$\hat{T}_{e^{-i\langle \cdot, z \rangle}}(\phi) = T_{e^{-i\langle \cdot, z \rangle}}(\hat{\phi}) = \int_{\mathbb{R}^n} e^{-iu \cdot z} \hat{\phi}(u) \, du = (2\pi)^n \phi(z),$$

so we find

$$\int_{\mathbb{R}^n} e^{ix \cdot u} \, dx = \hat{1} = \delta_0$$

in  $\mathcal{S}'(\mathbb{R}^n)$ .

For  $a \in C^\infty(\mathbb{R}^n)$  slowly increasing,  $a\phi \in \mathcal{S}(\mathbb{R}^n)$  for all  $\phi \in \mathcal{S}(\mathbb{R}^n)$ , so

$$(aT)(\phi) = T(a\phi)$$

for  $T \in \mathcal{S}'(\mathbb{R}^n)$  is well-defined as an element of  $\mathcal{S}'(\mathbb{R}^n)$ . Also

$$(\tau_y T)(\phi) = T(\tau_{-y} \phi)$$

is defined.

**Lemma 4.3.** *Let  $T \in \mathcal{S}'(\mathbb{R}^n)$ . Then,*

(i) *The following holds:*

$$\widehat{e^{i\langle \cdot, y \rangle} T} = \tau_y \hat{T}, \quad \widehat{\tau_y T} = e^{-i\langle \cdot, y \rangle} \hat{T}.$$

(ii) *The following holds:*

$$\widehat{D^\alpha T} = i^{|\alpha|} u^\alpha \hat{T}, \quad D^\alpha \hat{T} = (-i)^{|\alpha|} x^\alpha \widehat{T}.$$

**Proof:** We just use the definitions and what we know for elements of  $\mathcal{S}$ .

(i) First,

$$\begin{aligned}\widehat{e^{i\langle \cdot, y \rangle} T}(\phi) &= T(e^{i\langle \cdot, y \rangle} \hat{\phi}) = T(\widehat{\tau_{-y} \phi}) \\ &= \tau_y \hat{T}(\phi).\end{aligned}$$

(ii) For the second, do a similar thing.

$$\begin{aligned}\widehat{D^\alpha T}(\phi) &= D^\alpha T(\hat{\phi}) = (-1)^{|\alpha|} T(D^\alpha \hat{\phi}) = (-1)^{|\alpha|} (-i)^{|\alpha|} T(\widehat{u^\alpha \phi}) \\ &= i^{|\alpha|} u^\alpha \hat{T}(\phi),\end{aligned}$$

as required. The other one is similar.

For any  $\alpha$ ,  $D^\alpha \delta_0 \in \mathcal{E}'(\mathbb{R}^n) \subseteq \mathcal{S}'(\mathbb{R}^n)$ , and by the lemma,

$$\widehat{D^\alpha \delta_0} = i^{|\alpha|} u^\alpha,$$

so the linear span of the Fourier transform of all derivatives of  $\delta_0$  gives the polynomials.

## 4.1 Periodic Distributions

**Definition 4.1.**  $T \in \mathcal{D}'(\mathbb{R}^n)$  is *periodic* if  $\tau_k T = T$  for all  $k \in \mathbb{Z}^n$ .

### Example 4.2.

If  $f \in L^1_{\text{loc}}(\mathbb{R}^n)$  is periodic, then

$$\begin{aligned}\tau_k T_f(\phi) &= \int_{\mathbb{R}^n} f(x) \phi(x+k) \, dx = \int_{\mathbb{R}^n} f(x) \phi(x+k) \, dx = \int_{\mathbb{R}^n} f(x) \phi(x) \, dx \\ &= T_f(\phi).\end{aligned}$$

In particular,  $e^{2\pi i n \cdot x} = e_{2\pi n}$  are all periodic for  $n \in \mathbb{Z}^n$ .

If  $T \in \mathcal{E}'(\mathbb{R}^n)$ , then its periodisation is

$$T_{\text{per}} = \sum_{k \in \mathbb{Z}^n} \tau_k T.$$

Note that for  $\phi \in \mathcal{D}'(\mathbb{R}^n)$ ,  $\tau_k T(\phi)$  is non-zero only for finitely many  $k$ , since  $\text{supp}(T) \cap \text{supp}(\phi(\cdot + k))$  is empty eventually, so  $T_{\text{per}} \in \mathcal{D}'(\mathbb{R}^n)$  is periodic and

makes sense. Next, define

$$q = \{x \in \mathbb{R}^n \mid -1/2 \leq x_i < 1/2 \text{ for all } i = 1, 2, \dots, n\},$$

the *fundamental cell* of the integer lattice. Then

$$\sum_{k \in \mathbb{Z}^n} \tau_k \mathbb{1}_q(x) = \sum_{k \in \mathbb{Z}^n} \mathbb{1}_q(x - k) = 1.$$

**Lemma 4.4.** *There exists  $\psi \in C_c^\infty(\mathbb{R}^n)$  such that:*

- (i)  $\psi \geq 0$ .
- (ii)  $\text{supp} \psi \subseteq \text{int} Q$ , where  $Q = \{x \in \mathbb{R}^n \mid -1 \leq x_i \leq 1 \text{ for all } i\}$ .
- (iii) We have

$$\sum_{k \in \mathbb{Z}^n} \tau_k \psi(x) = 1,$$

for all  $x \in \mathbb{R}^n$ .

We call such  $\psi$  a (smooth) periodic partition of unity, or *ppu*. If  $T \in \mathcal{D}'(\mathbb{R}^n)$  is periodic, and  $\psi, \psi'$  are *ppu*'s, then  $T(\psi) = T(\psi')$ .

**Proof:** Take  $\psi_0 \in C_c^\infty(\mathbb{R}^n)$  supported in  $\text{int} Q$ , such that  $\psi \geq 0$ , and  $\psi = 1$  on  $q$ . Then define

$$S(x) = \sum_{k \in \mathbb{Z}^n} \psi_0(x - k),$$

for  $x \in \mathbb{R}^n$ . We can verify that  $S(x) \geq 1$ , since for any  $x \in \mathbb{R}^n$ , there exists  $k \in \mathbb{Z}^n$  such that  $x - k \in q$ . Also,  $S(x)$  is a finite series for each  $x$ , so we can define

$$\psi(x) = \frac{\psi_0(x)}{S(x)},$$

which satisfies the first two conditions, and also the last since

$$\sum_{k \in \mathbb{Z}^n} \psi(x - k) = \sum_k \frac{\psi_0(x - k)}{S(x - k)} = \frac{S(x)}{S(x)} = 1,$$

for all  $x \in \mathbb{R}^n$ .

Finally, take  $T \in \mathcal{D}'(\mathbb{R}^n)$  periodic. Then for  $\psi, \psi'$  being ppu's,

$$\begin{aligned} T(\psi) &= T\left(\psi \sum_{k \in \mathbb{Z}^n} \tau_k \psi'\right) = \sum_{k \in \mathbb{Z}^n} T(\psi \tau_k \psi') = \sum_{k \in \mathbb{Z}^n} \tau_k T(\psi \tau_k \psi') \\ &= \sum_{k \in \mathbb{Z}^n} T(\tau_{-k} \psi \psi') = T\left(\left(\sum_{k \in \mathbb{Z}^n} \tau_{-k} \psi\right) \psi'\right) = T(\psi'), \end{aligned}$$

as required.

Taking  $\psi_{0,j} \rightarrow \mathbb{1}_q$  pointwise in the previous proof, we obtain ppu's  $\psi_j \rightarrow \mathbb{1}_q$  pointwise and with  $\sup \|\psi_j\|_\infty < \infty$ , and hence for  $T_f, f \in L^1_{\text{loc}}$  periodic, we see

$$T_f(\psi_j) = T_f(\psi_{j+1}) = \int_{\mathbb{R}^n} f(x) \psi_j(x) dx \rightarrow \int_q f(x) dx$$

by DCT, and hence

$$T_f(\psi) = \int_q f(x) dx,$$

for all ppu's  $\psi$ .

**Definition 4.2.** The *mean*  $M(T)$  of a periodic distribution  $T \in \mathcal{D}'(\mathbb{R}^n)$  is defined as

$$M(T) = T(\psi),$$

for any ppu  $\psi$ .

Let's clear up a hole from earlier.

**Theorem 4.5.** For  $T \in \mathcal{E}'(\mathbb{R}^n)$ , the series

$$T_{\text{per}} = \sum_{k \in \mathbb{Z}^n} \tau_k T$$

converges in  $\mathcal{S}'(\mathbb{R}^n)$ . Moreover, if  $T \in \mathcal{D}'(\mathbb{R}^n)$  is periodic, then there exists  $V \in \mathcal{E}'(\mathbb{R}^n)$  such that  $T = V_{\text{per}}$ .

**Proof:** Recall that for  $T \in \mathcal{E}'(\mathbb{R}^n)$  with support  $K \subseteq B_R(0)$ , where  $R > 0$ , then there exists  $N, C$  such that

$$|T(\phi)| \leq C \sup_{\substack{x \in K \\ |\alpha| \leq N}} |D^\alpha \phi(x)|,$$

for all  $\phi \in \mathcal{E}(\mathbb{R}^n)$ .

For  $k \in \mathbb{Z}^n$ , we have, for  $x \in K$ ,

$$1 + |k| = 1 + |k + x - x| \leq 1 + R + |k + x| \leq (1 + R)(1 + |k + x|),$$

so

$$1 \leq \frac{(1 + R)^M (1 + |k + x|)^m}{|k|^m},$$

for any  $M \geq 0$ . Thus for all  $k \in \mathbb{Z}^n$ ,

$$|\tau_k T(\phi)| \leq C \frac{(1 + R)^M}{1 + |k|^M} \sup_{\substack{x \in K \\ |\alpha| \leq N}} (1 + |k + x|)^M |D^\alpha \phi(x + k)|,$$

which implies, with  $M = n + 1$ , that

$$\sum_{k \in \mathbb{Z}^n} |\tau_k T(\phi)| \leq C_{n, \phi, R} \sum_{k \in \mathbb{Z}^n} \frac{1}{(1 + |k|)^{n+1}} < \infty,$$

so this is absolutely convergent, hence the sum converges in  $\mathcal{S}'(\mathbb{R}^n)$ .

Finally, let  $T \in \mathcal{D}'(\mathbb{R}^n)$  be periodic, and take  $\psi$  a ppu. Then

$$\begin{aligned} T(\phi) &= \sum_{k \in \mathbb{Z}^n} \tau_k \psi T(\phi) = \sum_{k \in \mathbb{Z}^n} T((\tau_k \psi) \phi) \\ &= \sum_{k \in \mathbb{Z}^n} \tau_k T((\tau_k \psi) \phi) = \sum_{k \in \mathbb{Z}^n} T(\psi \tau_{-k} \phi) = \sum_{k \in \mathbb{Z}^n} \tau_k (\psi T)(\phi) = V_{\text{per}}. \end{aligned}$$

Because this series converges in  $\mathcal{S}'$ , the Fourier transform can thus be applied to periodic distributions in  $\mathcal{D}'(\mathbb{R}^n)$ .

We have not actually shown that  $\mathcal{S}'$  is complete in the weak-\* topology. However we can actually prove this; we will note. But note the following fact:

**Proposition 4.1.**  *$T : \mathcal{S}(\mathbb{R}^n) \rightarrow \mathbb{C}$  linear is continuous if there exists  $C, N, k$  such that*

$$|T(\phi)| \leq C \sup_{\substack{x \in \mathbb{R}^n \\ |\alpha| \leq k}} (1 + |x|)^N |D^\alpha \phi(x)|,$$

for all  $\phi \in \mathcal{S}(\mathbb{R}^n)$ .

**Theorem 4.6.** *Let  $U \in \mathcal{D}'(\mathbb{R}^n)$  be periodic. Then in  $\mathcal{S}'(\mathbb{R}^n)$ ,*

$$U = \sum_{k \in \mathbb{Z}^n} u_k T_{e_{2\pi k}},$$

where as before  $e_{2\pi k} = e^{2\pi i \langle k, \cdot \rangle}$ , and  $u_k = M(e_{-2\pi k} T)$  are the generalized Fourier coefficients.

**Proof:** Since  $U$  is periodic, we have

$$\widehat{\tau_k U - U} = (e_{-k} - 1)\hat{U} = 0,$$

which will be exploited in the following. We will prove if  $T \in \mathcal{S}'(\mathbb{R}^n)$  with  $(e_{-k} - 1)T = 0$  for all  $k \in \mathbb{Z}^n$ , then

$$T = \sum_{g \in \mathbb{Z}^n} c_g \delta_{2\pi g}$$

in  $\mathcal{S}'(\mathbb{R}^n)$ , for some  $c_g \in \mathbb{C}$ . Also  $|c_g| \leq C(1|g|)^N$  for some  $N, C > 0$ .

Indeed, define  $\Lambda^* = \{2\pi g \mid g \in \mathbb{Z}^n\}$ , and let  $\phi \in \mathcal{D}(\mathbb{R}^n)$  be such that  $\text{supp } \phi \cap \Lambda^* = \emptyset$ . Then  $(e_{-k} - 1)^{-1}\phi \in \mathcal{D}(\mathbb{R}^n)$ , so we have

$$T(\phi) = (e_{-k} - 1)T((e_{-k} - 1)^{-1}\phi) = 0$$

for all such  $\phi$ , hence  $T$  is also supported in  $\Lambda^*$ . Take  $\psi$  to be a ppu, and scale it to  $\tilde{\psi}(1/2\pi)$  such that  $\text{supp}(\tilde{\psi}) \subseteq \{-2\pi < x_i < 2\pi \forall i\}$ , such that

$$\sum_{g \in \mathbb{Z}^n} \tau_{2\pi g} \tilde{\psi}(x) = 1.$$

Now define

$$T_g = \tau_{2\pi g} \tilde{\psi} T,$$

which is supported in  $\{2\pi g\}$ . Clearly

$$\sum_{g \in \mathbb{Z}^n} T_g = \sum_{g \in \mathbb{Z}^n} \tau_{2\pi g} \tilde{\psi} T = T,$$

and moreover

$$(e_{-k} - 1)T_g = \tau_{2\pi g} \tilde{\psi} (e_{-k} - 1)T = 0,$$

for all  $k \in \mathbb{Z}^n$ . Now choosing  $k = g_j = (0, \dots, 0, 1, 0, \dots, 0)$ , the last identity implies

$$0 = (e^{-ix_j} - 1)T_g = (e^{-i(x_j - 2\pi g_j)} - 1)T_g.$$

Now  $e^{-i(x_j - 2\pi g_j)} - 1 = (x_j - 2\pi g_j)K(x_j)$ , where  $K(x_j)$  does not vanish near  $2\pi g_j$ , and  $K \in C^\infty$ . So we deduce

$$(x_j - 2\pi g_j)T_g = 0,$$



for all  $j$ . Next if  $\phi \in \mathcal{S}(\mathbb{R}^n)$ , then again by Taylor's theorem,

$$\phi(x) + \phi(2\pi g) + \sum_{j=1}^n (x_j - 2\pi g_j) \phi_j(x),$$

for  $\phi_j \in C^\infty(\mathbb{R}^n)$ . Since  $T_g$  is compactly supported, it extends to act on  $\mathcal{E}(\mathbb{R}^n)$ , so

$$\begin{aligned} T_g(\phi) &= T_g(\phi(2\pi g)) + \sum_{j=1}^n (x_j - 2\pi g_j) T_g(\phi_j) \\ &= \phi(2\pi g) T_g(1) = \delta_{2\pi g}(\phi) T_g \left( \sum_{g' \in \mathbb{Z}^n} \tau_{2\pi g'} \tilde{\psi} \right) = c_g \delta_{2\pi}(\phi), \end{aligned}$$

where  $c_g = T_g(\tau_{2\pi g} \tilde{\psi})$ .

We conclude that indeed

$$T = \sum_g c_g \delta_{2\pi g},$$

and convergence holds in  $\mathcal{S}'(\mathbb{R}^n)$  by

$$\begin{aligned} |c_g| &= |T_g(\tau_{2\pi g} \tilde{\psi})| \leq C \sup_{\substack{x \in \mathbb{R}^n \\ |\alpha| \leq k}} (1 + |x|)^N |D^\alpha \tilde{\psi}(x - 2\pi g)| \\ &\leq \tilde{C} (1 + |g|)^N \sup_{\substack{y \in \mathbb{R}^n \\ |\alpha| \leq k}} (1 + |y|)^N |D^\alpha \tilde{\psi}(y)| < \infty. \end{aligned}$$

TO prove the main theorem, note that  $U \in \mathcal{S}'(\mathbb{R}^n)$  can be Fourier transformed, so

$$\hat{U} = (2\pi)^n \sum_{g \in \mathbb{Z}^n} u_g \delta_{2\pi g},$$

for  $u_g = c_g / (2\pi)^n$ . Fourier transforming back, we see

$$U = \sum_{g \in \mathbb{Z}^n} u_g T_{e_{2\pi g}}$$

in  $\mathcal{S}'(\mathbb{R}^n)$ . To find  $u_g$ , note that  $T \mapsto M(e_{-2\pi k} T) = T(e_{-2\pi k} \psi)$  is continuous from  $\mathcal{S}'(\mathbb{R}^n) \rightarrow \mathbb{C}$ , therefore

$$M(e_{-2\pi k} U) = \sum_{g \in \mathbb{Z}^n} u_g M(e_{-2\pi k} T_{e_{2\pi g}}) = u_k.$$

Let us apply this to

$$T = \sum_{k \in \mathbb{Z}^n} \delta_k = \sum_{k \in \mathbb{Z}^n} \tau_k \delta_0 \in \mathcal{S}'(\mathbb{R}^n).$$

This is a periodisation of an element of  $\mathcal{E}'(\mathbb{R}^n)$ . We can take without loss of generality a ppu  $\psi$  such that  $\psi(0) = 1$ , and compute

$$M(e_{-2\pi k} T) = T(e_{-2\pi k} \psi) = \delta_0(e_{-2\pi k} \psi) = \psi(0) \cdot 1 = 1.$$

Hence this gives us

$$\sum_{k \in \mathbb{Z}^n} \delta_k = \sum_{k \in \mathbb{Z}^n} T_{e_{2\pi k}}.$$

Testing this against  $\phi(x - \cdot)$ , for  $\phi \in \mathcal{S}(\mathbb{R}^n)$ , gives

$$\sum_{k \in \mathbb{Z}^n} \phi(x + k) = \sum_{k \in \mathbb{Z}^n} T_{e_{2\pi k}}(\phi(x + \cdot)) = \sum_{k \in \mathbb{Z}^n} e^{2\pi i k \cdot x} \hat{\phi}(2\pi k),$$

so the Fourier coefficients of  $\phi_{\text{per}}$  are precisely  $\hat{\phi}(2\pi k)$ .

## 5 Sobolev Spaces and Elliptic PDE's

Let us define Banach spaces  $C^k(\Omega)$ , for  $k \in \mathbb{N}$ , as

$$\{f : \Omega \rightarrow \mathbb{C} \mid D^\alpha f \text{ exists, } \|f\|_{C^k(\Omega)} < \infty\},$$

where  $\Omega \subseteq \mathbb{R}^n$  is open, and

$$\|f\|_{C^k(\Omega)} = \sum_{0 \leq |\alpha| \leq k} \|D^\alpha f\|_\infty,$$

where  $\|f\|_\infty$  is the supremum norm. Similarly, we define *Hölder spaces*  $C^{k,\eta}(\Omega)$  with norm, for  $0 < \eta \leq 1$ ,

$$\|f\|_{C^{k,\eta}} = \|f\|_{C^k} + \sum_{|\alpha|=k} \sup_{x \neq y} \frac{|D^\alpha f(x) - D^\alpha f(y)|}{|x - y|^\eta}.$$

We can reface in the definition of  $C^k(\Omega)$ , the infinity norms by  $L^p(\Omega)$ , norms, and  $D^\alpha$  by  $D_w^\alpha$ , where we recall that  $D_w^\alpha f \in L^p$  is the weak derivative of  $f \in L^p(\Omega)$  if  $D^\alpha T_f = T_{D_w^\alpha f}$  in  $\mathcal{D}'(\Omega)$ , which is equivalent to requiring

$$(-1)^{|\alpha|} \int_\Omega f D^\alpha \phi = \int_\Omega D_w^\alpha f \phi,$$

for all  $f \in C_c^\infty(\Omega)$ .

**Definition 5.1.** Let  $1 \leq p \leq \infty$ , and  $k \in \mathbb{N}$ . Then  $f \in L^p(\Omega)$  is said to belong to the *Sobolev space*  $W^{k,p}(\Omega)$  if  $D_w^\alpha f \in L^p(\Omega)$ , for all  $0 < |\alpha| \leq k$ .

These spaces are complete for the norm

$$\|f\|_{W^{k,p}(\Omega)} = \left( \sum_{0 \leq |\alpha| \leq k} \|D_w^\alpha f\|_{L^p}^p \right)^{1/p},$$

for  $p < \infty$ , and  $\max_{0 \leq |\alpha| \leq k} \|D_w^\alpha f\|_{L^\infty}$  if  $p = \infty$ .

*Remark.* For completeness, weak derivatives are essential. As an example, consider the ramp function.

Henceforth, we write  $D^\alpha$  for  $D_w^\alpha$ . A rich theory arises when  $p = 2$ ,  $\Omega = \mathbb{R}^n$ , in which case we define:

**Definition 5.2.** Let  $s \in \mathbb{R}$ . Then  $f \in \mathcal{S}'(\mathbb{R}^n)$  is said to belong to  $H^s(\mathbb{R}^n)$  if  $\hat{f} \in L^2_{\text{loc}}(\mathbb{R}^n)$  and

$$\int_{\mathbb{R}^n} |\hat{f}(u)|^2 (1 + |u|^2)^s du < \infty.$$

On  $H^s(\mathbb{R}^n)$ , we can define a Hilbert space inner product by

$$(f, g)_{H^s} = \int_{\mathbb{R}^n} \hat{f}(u) \overline{\hat{g}(u)} (1 + |u|^2)^s du,$$

which is just  $L^2(\mu_s)$ , for  $d\mu_s(u) = (1 + |u|^2)^s du$ .

**Proposition 5.1.** *For  $s \in \mathbb{N}$ , we have  $H^s(\mathbb{R}^n) = W^{s,2}(\mathbb{R}^n)$ .*

**Proof:** (Sketch). Use  $|D^\alpha f| = |u^\alpha \hat{f}|$  for  $0 \leq |\alpha| \leq s$ , and then use Plancherel's theorem.

One shows:

- (i)  $H^s(\mathbb{R}^n) \subseteq H^t(\mathbb{R}^n)$  for  $s \geq t$ ,
- (ii)  $\mathcal{S}(\mathbb{R}^n)$  is dense in  $H^s(\mathbb{R}^n)$ , but we have to be careful on  $\Omega$  with boundary  $\partial\Omega$ .
- (iii)  $\delta_x \in H^{-s}(\mathbb{R}^n)$  for  $s > \frac{n}{2}$ , for any  $x \in \mathbb{R}^n$ .
- (iv) One can show that  $(H^s(\mathbb{R}^n))' = H^{-s}(\mathbb{R}^n)$ , viewed in  $\mathcal{S}'(\mathbb{R}^n)$ .

**Theorem 5.1** (Sobolev Embedding). *Let  $f \in H^s(\mathbb{R}^n)$  for  $s > k + \frac{n}{2}$ ,  $k \in \mathbb{N} \cup \{0\}$ . Then there exists  $f^* \in C^k(\mathbb{R}^n)$  such that  $f = f^*$  almost everywhere, and*

$$\|f\|_{C^k} \leq C_{s,n,k} \|f\|_{H^s}.$$

**Proof:** Let  $f \in \mathcal{S}(\mathbb{R}^n)$ , then  $|D^\alpha f(x)| = |F^{-1}(u^\alpha \hat{f})|(x)$ , so by the Riemann-Lebesgue lemma, for  $|\alpha| \leq k$ ,

$$\begin{aligned} |D^\alpha f(x)| &\leq \frac{1}{(2\pi)^n} \int_{\mathbb{R}^n} |u|^{|\alpha|} |\hat{f}(u)| \frac{(1 + |u|^2)^{s/2}}{(1 + |u|^2)^{s/2}} du \\ &\leq \frac{1}{(2\pi)^n} \left( \int_{\mathbb{R}^n} \frac{|u|^{2k}}{(1 + |u|^2)^s} du \right)^{1/2} \left( \int_{\mathbb{R}^n} |\hat{f}(u)|^2 (1 + |u|^2)^s du \right)^{1/2}, \end{aligned}$$

and the right term gives  $\|f\|_{H^s}$ , and the left term is finite for  $s > k + \frac{n}{2}$ , so this entire thing is less than  $C_{s,k,n} \|f\|_{H^s}$ .

Now take  $f_n \in \mathcal{S}(\mathbb{R}^n)$  such that  $f_n \rightarrow f$  in  $H^s$ , and hence almost-everywhere. So  $f_n$  is a Cauchy sequence in  $H^s$ , hence applying this inequality to  $f_m - f_k$ , we see that  $f_n$  is Cauchy in  $C^k(\mathbb{R}^n)$ . So since  $C^k(\mathbb{R}^n)$  is complete, there exists  $f^*$  such that  $f_n \rightarrow f^*$  in  $C^k$ , where  $f = f^*$  almost everywhere. Taking limits in the inequality, the theorem is proved.

*Remark.* Identifying  $f$  with its unique continuous version  $f^*$ , we have

$$H^s(\mathbb{R}^n) \subseteq C^k(\mathbb{R}^n),$$

for  $s > k + \frac{n}{2}$ , and

$$\bigcap_{s>0} H^s(\mathbb{R}^n) \subseteq C^\infty(\mathbb{R}^n).$$

One shows further that  $H^s(\mathbb{R}^n) \subseteq C^{k,\eta}(\mathbb{R}^n)$  if  $s + k + \eta + \frac{n}{2}$ .

We look at an easy example: the PDE

$$-\Delta v + v = f \tag{*}$$

in  $\mathcal{S}'(\mathbb{R}^n)$ , where  $f \in H^s(\mathbb{R}^n)$ , and  $\Delta$  is the Laplacian,

$$\Delta = \sum_{j=1}^n \frac{\partial^2}{(\partial x_j)^2}.$$

**Theorem 5.2.** *There exists a unique solution  $v \in H^{s+2}(\mathbb{R}^n)$  to (\*), and*

$$\|v\|_{H^{s+2}(\mathbb{R}^n)} \leq \|f\|_{H^s(\mathbb{R}^n)},$$

for  $s \in \mathbb{R}$ .

This inequality is known as ‘elliptic regularity’.

**Proof:** We Fourier transform (\*), and note that

$$\hat{\Delta} = \sum_{j=1}^n \widehat{\frac{\partial^2}{(\partial x_j)^2}} = i^2 \sum_{j=1}^n u_j^2 = -|u|^2.$$

So (\*) is equivalent in  $\mathcal{S}'(\mathbb{R}^n)$  to the equation

$$(1 + |u|^2)\hat{v} = \hat{f}(u),$$

which has a unique solution

$$\hat{v}(u) = \frac{\hat{f}(u)}{1 + |u|^2},$$

so we can let  $v = F^{-1}(\hat{v})$ , which solve (\*) in  $\mathcal{S}'(\mathbb{R}^n)$ . Moreover,

$$\begin{aligned} \|v\|_{H^{s+2}}^2 &= \int_{\mathbb{R}^n} (1 + |u|^2)^{s+2} |\hat{v}(u)|^2 \, du = \int_{\mathbb{R}^n} (1 + |u|^2)^s |\hat{f}(u)|^2 \, du \\ &= \|f\|_{H^s}^2. \end{aligned}$$

*Remark.* If  $s > \frac{n}{2}$ , then

$$v \in H^{s+2}(\mathbb{R}^n) \subseteq C^2(\mathbb{R}^n),$$

so  $\Delta v$  is pointwise defined, and testing

$$\int_{\mathbb{R}^n} (-\Delta v + v)\phi \, dx = \int_{\mathbb{R}^n} f\phi \, dx,$$

for all  $\phi \in \mathcal{S}(\mathbb{R}^n)$ . Testing against appropriate  $\phi$ , we see that

$$-\Delta v + v = f$$

on  $\mathbb{R}^n$  pointwise, and  $\|v\|_{C^2} \leq C\|f\|_{H^s}$ . In particular, if  $f \in \mathcal{S}(\mathbb{R}^n)$ , then so is  $v$ .

To solve similar equations on subsets  $\Omega$  (open, bounded) of  $\mathbb{R}^n$ , we need to be able to define Sobolev functions on null sets.

**Theorem 5.3** (Trace Theorem). *Let  $s > \frac{1}{2}$ . Then there exists a bounded linear map  $T : H^s(\mathbb{R}^n) \rightarrow H^{s-1/2}(\mathbb{R}^{n-1})$  such that for all  $f \in \mathcal{S}(\mathbb{R}^n)$ ,*

$$Tf = f|_{\mathbb{R}^{n-1} \times \{0\}}.$$

We call  $T = T_\Sigma$  the trace (restriction) map to  $\Sigma = \mathbb{R}^{n-1} \times \{0\}$ .

**Proof:** On the example sheet.

*Remark.* By coordinate transformations, the trace map extends to describe restrictions to boundary manifolds  $\partial\Omega$  of  $\Omega$  if  $\Omega$  is sufficiently regular. The ball is easy.

## 5.1 Weird Space

Let  $\Omega \subset \mathbb{R}^n$  be open and bounded. We can regard  $f \in C_c^\infty(\Omega)$  as an element of  $H^1(\mathbb{R}^n)$  by zero extension outside its support. We thus define  $H_0^1(\Omega)$  as the closure of  $C_c^\infty(\Omega)$  for  $H^1(\mathbb{R}^n)$ -norm.

The Hilbert norm of  $H^1(\mathbb{R}^n)$  is then given by

$$\|f\|_{H^1}^2 = \int_{\mathbb{R}^n} (1 + |u|^2) |\hat{f}(u)|^2 \, du = (2\pi)^n \int_{\Omega} (|f(x)|^2 + |Df(x)|^2) \, dx,$$

where  $Df = \nabla f$  is the gradient, and  $f \in C_c^\infty(\Omega)$ . This identity extends to all  $f \in H_0^1(\Omega)$  by virtue of the following proposition:

**Proposition 5.2.** *Any  $f \in H_0^1(\Omega)$  vanishes for almost all  $x \in \Omega^c$ , and if  $\partial\Omega$  is sufficiently regular, then  $f$  vanishes at  $\partial\Omega$  in a trace sense, i.e.  $T_{\partial\Omega}f = 0$ .*

**Proof:** First, for  $\phi \in C_c^\infty(\text{int } \Omega^c)$ , the map

$$\Lambda_\phi(f) = \int_{\mathbb{R}^n} \phi f \, dx$$

is continuous on  $H_0^1$ , since  $|\Lambda_\phi(f)| \leq \|\phi\|_{L^2} \|f\|_{L^2} \leq C \|f\|_{H^1}$ , and so if  $f_n \in C_c^\infty(\Omega)$ ,  $f_n \rightarrow f$  in  $H^1$ , then

$$0 = \Lambda_\phi(f_n) \rightarrow \Lambda_\phi(f) = 0,$$

so  $f$  is supported in  $\overline{\Omega}$ . Similarly, if  $f_n \rightarrow f$  in  $H^1$  and  $f_n \in C_c^\infty$ , then  $0 = T_{\partial\Omega}(f_n) \rightarrow T_{\partial\Omega}(f)$ . So  $f$  vanishes in a trace sense.

We now show existence of unique solutions to the boundary value problem

$$\begin{aligned} -\Delta v + v &= f \text{ on } \Omega, \\ v &= 0 \text{ on } \partial\Omega, \end{aligned}$$

in  $\mathcal{D}'(\Omega)$ , where  $f \in L^2$ . So

$$\int_{\Omega} (-\Delta v + v) \phi \, dx = \int_{\Omega} f \phi \, dx,$$

for  $\phi \in C_c^\infty(\Omega)$ . Doing this by integration by parts,

$$\int_{\Omega} (Dv \cdot D\phi + v\phi) \, dx = \int_{\Omega} f \phi \, dx,$$

for all  $\phi$ . This is equivalent to solving

$$\langle v, \phi \rangle_{H^1} = \langle f, \phi \rangle_{L^2}, \tag{\dagger'}$$

for all  $\phi \in H_0^1(\Omega)$ . We call  $v \in H_0^1$  solving  $(\dagger')$  a weak solution of  $(\dagger)$ .

**Theorem 5.4.** *There exists a unique solution  $v \in H_0^1(\Omega)$  to  $(\dagger')$ , and  $\|v\|_{H^1} \leq \|f\|_{L^2}$ . Further, the solution operator  $S : f \mapsto v_f = v$  maps  $L^2(\Omega)$  into  $H_0^1(\Omega)$ , is linear, continuous and also self-adjoint for  $\langle \cdot, \cdot \rangle_{L^2(\Omega)}$ .*

**Proof:** Define linear operator

$$\Lambda_f(\phi) = \int_{\Omega} f \phi \, dx,$$

which is linear and continuous on the Hilbert space  $H_0^1(\Omega)$  since

$$|\Lambda_f(\phi)| \leq \|f\|_{L^2} \|\phi\|_{L^2} \leq \|f\|_{L^2} \|\phi\|_{H_0^1}.$$

Thus by the Riesz-representation theorem on  $H_0^1(\Omega)$ , there exists  $v \in H_0^1$  such that

$$\langle v, \phi \rangle_{H^1} = \Lambda_f(\phi) = \langle f, \phi \rangle_{L^2},$$

for all  $\phi \in H_0^1(\Omega)$ , and  $\|v\|_{H^1} \leq \|\Lambda_f\|_{(H^1)'} \leq \|f\|_{L^2}$ , by the earlier bound.

To see that  $S : L^2(\Omega) \rightarrow H^1(\Omega)$  is linear, take  $f_1, f_2 \in L^2(\Omega)$ , with  $v_i = S(f_i)$ , and define  $v = v_1 + \alpha v_2$ , for  $\alpha \in \mathbb{R}$ , then

$$\begin{aligned} \langle v, \phi \rangle_{H^1} &= \langle v_1 + \alpha v_2, \phi \rangle_{H^1} = \langle v_1, \phi \rangle_{H^1} + \alpha \langle v_2, \phi \rangle_{H^1} \\ &= \langle f_1, \phi \rangle_{L^2} + \alpha \langle f_2, \phi \rangle_{L^2} = \langle f_1 + \alpha f_2, \phi \rangle_{L^2}, \end{aligned}$$

so  $S(f_1 + \alpha f_2) = v = v_1 + \alpha v_2 = S(f_1) + \alpha S(f_2)$ .

$S$  is continuous since  $\|S(f)\|_{H^1} = \|v_f\|_{H^1} \leq \|f\|_{L^2}$ , and self-adjoint as

$$\begin{aligned} \langle S(f), g \rangle_{L^2(\Omega)} &= \langle g, S(f) \rangle_{L^2} = \langle S(g), S(f) \rangle_{H^1} = \langle S(f), S(g) \rangle_{H^1} \\ &= \langle f, S(g) \rangle_{L^2}. \end{aligned}$$

To obtain more ‘regular’ solutions we define

$$H_{\text{loc}}^s(\Omega) = \{f \in L_{\text{loc}}^2(\Omega) \mid f\chi \in H^s(\mathbb{R}^n) \text{ for all } \chi \in C_c^\infty(\Omega)\}.$$

**Proposition 5.3.** *If  $f \in H_{\text{loc}}^s(\Omega)$  for  $s > k + \frac{n}{2}$ ,  $k \in \mathbb{N}$ , then for any open set  $U$  such that  $\overline{U} \subseteq \Omega$ , we have  $f \in C^k(U)$ .*

**Proof:** Let  $\chi \in C_c^\infty(\Omega)$  such that  $\chi = 1$  on  $U$ . Then  $f\chi \in H^s(\mathbb{R}^n) \subseteq C^k(\mathbb{R}^n)$  by the Sobolev embedding and on  $U$ ,  $f\chi = f$ , so  $f \in C^k(U)$ .

We now look at properties of solutions to PDEs.

**Theorem 5.5** (Interior Regularity Theorem). *Suppose  $v \in H_0^1(\Omega)$  solves  $(\dagger')$ , so*

$$\int_{\Omega} (Dv \cdot D\phi + v\phi) \, dx = \int_{\Omega} f\phi \, dx, \quad (\dagger)$$

*for all  $\phi \in H_0^1(\Omega)$ , and some  $f \in L^2(\Omega)$ . Then  $f \in H_{\text{loc}}^2(\Omega)$ . If  $f \in H_{\text{loc}}^k(\Omega)$ , for some  $k \in \mathbb{N}$  then  $v \in H_{\text{loc}}^{k+2}(\Omega)$ .*



**Proof:** Take  $K \subseteq \Omega$  compact,  $\chi \in C_c^\infty(\Omega)$  such that  $\chi = 1$  on  $K$ , and  $\varphi \in \mathcal{S}(\mathbb{R}^n)$ . Take  $\phi = \chi\varphi \in H_0^1(\Omega)$  in  $(\dagger)$  to obtain

$$\int_{\Omega} (Dv \cdot D(\chi\varphi) + v\chi\varphi) dx = \int_{\Omega} f\chi\varphi dx.$$

Using the chain rule and integration by parts, this rearranges to

$$\int_{\Omega} (D(v\chi) \cdot D\varphi + v\chi\varphi) dx = \int_{\Omega} g\phi dx,$$

where

$$g = -Dv \cdot D\chi - v\Delta\chi + f\chi \in L^2(\Omega),$$

since  $v \in H_0^1(\Omega)$  by hypothesis. In fact, we can replace  $\Omega$  by  $\mathbb{R}^n$  since  $\chi$  and  $\phi$  are supported only on  $\Omega$ . Therefore  $w = v\chi$  solves

$$-\nabla w + w = g$$

in  $\mathcal{S}'(\mathbb{R}^n)$ , and so by the elliptic regularity estimate on  $\mathbb{R}^n$ , we deduce

$$\|\chi v\|_{H^2} = \|w\|_{H^2(\mathbb{R}^n)} \leq \|g\|_{L^2(\mathbb{R}^n)} = \|g\|_{L^2(\Omega)} < \infty.$$

Now if  $\psi \in C_c^\infty(\Omega)$  is given, take  $K = \text{supp}(\psi)$  and, using the basic inequality,

$$\|\psi \cdot f\|_{H^k} \leq C_{k,n} \|\psi\|_{C^k} \|f\|_{H^k}.$$

This is easy, as  $\|\psi f\|_{H^2} \leq \|\psi\|_{\infty} \|f\|_{L^2}$ , plus a derivative. Thus

$$\|\psi \cdot v\|_{H^2} = \|\psi \cdot (v\chi)\|_{H^2} \leq C \|v\chi\|_{H^2} < \infty.$$

Now we “bootstrap” this argument, as for  $f \in H_{\text{loc}}^1(\Omega)$ , we now know  $g \in H_{\text{loc}}^1(\Omega)$ , so again using the elliptic regularity estimate on  $\mathbb{R}^n$ , we get  $v \in H_{\text{loc}}^3(\Omega)$ . The rest is just induction on  $k$ .

**Corollary 5.1.** *If  $f \in L^2(\Omega) \cap C^\infty(\Omega)$ , then there exists  $v \in H_0^1(\Omega) \cap C^\infty(\Omega)$  such that*

$$-\nabla v + v = f$$

*on  $\Omega$ , pointwise.*

**Theorem 5.6** (Rellich-Kondrashov). *Let  $\Omega \subset \mathbb{R}^n$  be open and bounded, and  $u_j \in H_0^1(\Omega)$  be such that  $\|u_j\|_{H^1(\Omega)} \leq K$  for all  $j = 1, 2, \dots$  and some  $K > 0$ . Then there exists  $u \in H_0^1(\Omega)$  such that  $u_{j_k} \rightarrow u$  in  $L^2(\Omega)$  along a subsequence.*

**Proof:** By the Banach-Alaoglu theorem in  $H_0^1(\mathbb{R}^n)$ , we extract a subsequence  $u_{j_k} \rightarrow u$  weakly in  $H_0^1$ , and then also weakly in  $L^2(\Omega)$ , for some  $u \in H_0^1$ . Then by Plancherel's theorem,

$$\begin{aligned} \|u_{j_k} - u\|_{L^2(\Omega)}^2 &= \frac{1}{(2\pi)^n} \|\hat{u}_{j_k} - \hat{u}\|_{L^2(\mathbb{R}^n)}^2 \\ &= \frac{1}{(2\pi)^n} \int_{|z|>R} |\hat{u}_{j_k}(z) - \hat{u}(z)|^2 dz \\ &\quad + \frac{1}{(2\pi)^n} \int_{|z|\leq R} |\hat{u}_{j_k}(z) - \hat{u}(z)|^2 dz \\ &= \frac{1}{(2\pi)^n} (\text{I} + \text{II}). \end{aligned}$$

Note

$$\begin{aligned} \text{I} &\leq 2 \int_{|z|>R} \frac{1+|z|^2}{1+|z|^2} (|\hat{u}_{j_k}(z)|^2 + |\hat{u}(z)|^2) dz \leq \frac{2}{1+R^2} (\|u_{j_k}\|_{H^1}^2 + \|u\|_{H^1}^2) \\ &\leq \frac{4K^2}{1+R^2} < \varepsilon, \end{aligned}$$

for  $R$  large. Now for  $\text{II}$ , note

$$\begin{aligned} \hat{u}_{j_k}(z) &= \int_{\Omega} e^{-iz \cdot x} u_{j_k}(x) dx = \langle e_{-z}, u_{j_k} \rangle_{L^2(\Omega)} \\ &\rightarrow \langle e_{-z}, u \rangle_{L^2(\Omega)} = \hat{u}(z) \end{aligned}$$

as  $k \rightarrow \infty$ , and also

$$|\hat{u}_{j_k}(z)| + |\hat{u}(z)| \leq \|u_{j_k}\|_{L^1(\Omega)} + \|u\|_{L^1(\Omega)} \leq C_{\Omega} (\|u_{j_k}\|_{L^2(\Omega)} + \|u\|_{L^2(\Omega)}) \leq 2C_{\Omega} K,$$

whose square is certainly  $dz$  integrable on  $\{z \mid |z| \leq R\}$ , so by DCT,  $\text{II} \rightarrow 0$  as  $k \rightarrow \infty$ .

**Corollary 5.2.**  $H_0^1(\Omega)$  is compactly embedded into  $L^2(\Omega)$  for  $\Omega$  bounded, open.

As a consequence, the solution operator  $S$  from  $(\dagger')$  is a compact, linear and self-adjoint operator on  $L^2(\Omega)$ , so from the spectral theorem there exists an orthonormal basis  $\{w_k \mid k = 1, 2, \dots\}$  of  $L^2(\Omega)$  such that

$$Sw_k = \mu_k w_k,$$

where  $\mu_k \in \mathbb{R}$ ,  $\mu_k \rightarrow 0$  as  $k \rightarrow \infty$ . In particular,  $w_k \in H_0^1(\Omega)$ , so by  $(\dagger')$ ,

$$\langle w_k, v \rangle_{L^2} = \langle Sw_k, v \rangle_{H^1} = \mu_k \langle w_k, v \rangle_{H^1},$$

for all  $v \in H_0^1(\Omega)$ . Choosing  $w_k = v$  gives  $1 = \mu_k \|w_k\|_{H^1}$ , so  $\mu_k > 0$  for all  $k$ . This gives us

$$-\Delta w_k + w_k = \frac{1}{\mu_k} w_k,$$

weakly for all  $k$ . So applying the interior regularity theorem repeatedly, we see  $w_k \in C^\infty(\Omega)$ , so in particular this identity holds pointwise, and

$$-\Delta w_k = \left( \frac{1}{\mu_k} - 1 \right) w_k = \lambda_k w_k,$$

where now  $w_k, \lambda_k$  are the eigenfunctions and eigenvalues of  $\Delta$  on  $\Omega$  with Dirichlet boundary conditions;  $w_k = 0$  on  $\partial\Omega$  in a trace sense, with  $\lambda_k \rightarrow \infty$ .

Now consider the Dirichlet problem for the Laplace equation,

$$\begin{aligned} -\Delta v &= f \text{ on } \Omega, \\ v &= 0 \text{ on } \partial\Omega, \end{aligned} \tag{*}$$

whose weak formulation in  $\mathcal{D}'(\Omega)$  is

$$\langle -\Delta v, \phi \rangle_{L^2} = \langle f, \phi \rangle_{L^2},$$

for all  $\phi \in C_c^\infty(\Omega)$ .

**Theorem 5.7** (Poincaré's Inequality).  $\lambda_1 > 0$  and  $\lambda_k \geq \lambda_1 > 0$  for all  $k$ .

**Proof:** Example sheet.

Then define

$$v_J = \sum_{k=1}^J \lambda_k^{-1} w_k \langle w_k, f \rangle \in H_0^1 \cap C^\infty(\Omega).$$

Now  $v_J$  is a Cauchy sequence in  $H_0^1$  since, by Parseval's identity,

$$\begin{aligned} \|v_J - v_{J'}\|_{H^1}^2 &= \langle v_J - v_{J'}, v_J - v_{J'} \rangle_{L^2} + \langle D(v_J - v_{J'}), D(v_J - v_{J'}) \rangle_{L^2} \\ &= \sum_{k=J+1}^{J'} \lambda_k^{-2} \langle w_k, f \rangle_{L^2}^2 - \langle v_J - v_{J'}, \Delta(v_J - v_{J'}) \rangle_{L^2} \\ &= \sum_{k=J+1}^{J'} \lambda_k^{-2} \langle w_k, f \rangle_{L^2}^2 + \sum_{k=J+1}^{J'} \lambda_k^{-1} \langle f, w_k \rangle^2 \\ &\leq C(\lambda_1^{-1}) \sum_{k=J'+1}^J \langle f, w_k \rangle^2 \rightarrow 0 \end{aligned}$$

as  $J \rightarrow \infty$ , since  $f \in L^2(\Omega)$ , so since  $H_0^1(\Omega)$  is complete,

$$v = \sum_{k=1}^{\infty} \lambda_k^{-1} w_k \langle w_k, f \rangle_{L^2} \in H_0^1(\Omega),$$

which solves  $(*)$  weakly, since

$$\begin{aligned} \langle -\Delta v, \phi \rangle_{L^2} &= \langle v, -\Delta f \rangle_{L^2} = \sum_{k=1}^{\infty} \lambda_k^{-1} \langle w_k, f \rangle_{L^2} \langle -\Delta f, w_k \rangle_{L^2} \\ &= \sum_{k=1}^{\infty} \lambda_k^{-1} \langle w_k, f \rangle \langle -\Delta w_k, \phi \rangle = \langle f, \phi \rangle_{L^2}, \end{aligned}$$

for all  $\phi \in \mathcal{D}(\Omega)$ . For uniqueness, note that if  $w \in H_0^1(\Omega)$  is such that  $-\Delta w = 0$ , then using the Poincaré inequality once again,

$$\begin{aligned} \|w\|_{H^1}^2 &= \langle w, w \rangle_{L^2} + \langle Dw, Dw \rangle_{L^2} \leq \left( \frac{1}{\lambda_1} + 1 \right) \langle Dw, Dw \rangle_{L^2} \\ &= - \left( \frac{1}{\lambda_1} + 1 \right) \langle \Delta w, w \rangle = 0, \end{aligned}$$

so  $w = 0$  almost everywhere. Hence if  $v' \in H_0^1$  is another weak solution to  $(*)$ , then

$$\Delta(v - v') = f - f = 0,$$

but  $v - v' = w \in H_0^1$ , so  $v = v'$ . In conclusion, we have proven the following:

**Theorem 5.8.** *There exists a unique weak solution  $v$  to  $(*)$  given by*

$$v = \sum_{k=1}^{\infty} \lambda_k^{-1} w_k \langle w_k, f \rangle_{L^2}.$$

Poisson summation formula.

One can use spectral theory to solve other PDE's such as the heat equation:

$$\begin{aligned} \frac{\partial}{\partial t} v(t, x) &= \Delta v(t, x) & \forall (t, x) \in (0, T] \times \Omega, \\ v(t, x) &= 0 & \forall t, x \in \partial\Omega, \\ v(0, x) &= \phi & \forall x \in \Omega, \end{aligned}$$

for some initial condition  $\phi \in L^2(\Omega)$ . The solution  $v$  can be represented spectrally by

$$v(t, x) = \sum_{k=1}^{\infty} e^{-t\lambda_k} w_k \langle w_k, \phi \rangle_{L^2}.$$

Since

$$\frac{\partial}{\partial t} e^{-t\lambda_k} w_k = e^{-t\lambda_k} (-\lambda_k) w_k = e^{-t\lambda_k} \Delta w_k,$$

and since  $e^{-t\lambda_k} \rightarrow 1$  as  $t \rightarrow 0$ , this indeed solves the heat equation if we can differentiate under the summation. This can be justified by a result from probability and measure, and the following bounds:

- $\|\chi w_k\|_{H^s(\mathbb{R}^n)} \leq c \lambda_k^{c(s,n)}$ , for some  $c(s, n) > 0$ . This is proved as in the interior regularity theorem.
- (Weyl's law/asymptotics).  $\lambda_k \simeq k^{2/n}$ , where  $n = \dim(\Omega)$ .

Notice that

$$p_t(x, y) = \sum_{k=1}^{\infty} e^{-\lambda_k t} w_k(x) w_k(y)$$

is the *heat kernel*, or fundamental solution. This is the momentum density of Brownian motion. The time average of the heat flow then equals

$$\int_0^{\infty} v(t, x) dt = - \sum_{k=1}^{\infty} \frac{1}{\lambda_k} w_k \langle w_k, \phi \rangle,$$

which is the solution to the elliptic PDE  $-\Delta u = \phi$  on  $\Omega$ , such that  $u = 0$  on  $\partial\Omega$ .

## 5.2 Variational Problems and PDEs

Consider minimising a functional

$$S(u) = \|u\|_{H^1}^2 - 2\langle f, u \rangle_{L^2},$$

over  $H_0^1(\Omega)$ , for fixed  $f \in L^2(\Omega)$ .

**Theorem 5.9.** *We have*

$$\inf\{S(u) \mid u \in H_0^1(\Omega)\} = \sigma > -\infty,$$

*and there exists a unique minimizer  $w \in H_0^1$  which solves the PDE*

$$-\Delta w + w = f$$

*weakly on  $\Omega$ .*

**Proof:** Using Cauchy-Schwarz and  $ab \leq \frac{1}{2}a^2 + \frac{1}{2}b^2$ ,

$$\begin{aligned} S(u) &= \|u\|_{H^1}^2 - 2\|f\|_{L^2}\|u\|_{L^2} \geq \|u\|_{H^1}^2 - \frac{1}{2}\|u\|_{L^2}^2 - 2\|f\|_{L^2}^2 \\ &\geq \frac{1}{2}\|u\|_{H^1}^2 - 2\|f\|_{L^2}^2 \geq -2\|f\|_{L^2}^2 > -\infty, \end{aligned}$$

so  $\sigma > -\infty$ . Taking  $u_k \in H_0^1$  such that  $S(u_k) \rightarrow \sigma$  as  $k \rightarrow \infty$  and noting  $\sup_k |S(u_k)| \leq \bar{S} < \infty$ , we have

$$\|u_k\|_{H^1}^2 \leq S(u_k) + 2\|f\|_{L^2}\|u_k\|_{L^2} \leq \bar{S} + \frac{1}{2}\|u_k\|_{L^2}^2 + 2\|f\|_{L^2}^2,$$

which rearranged gives

$$\|u_k\|_{H^1}^2 \leq \bar{S} + 4\|f\|_{L^2}^2 < \infty.$$

Since  $u_k$  is bounded in  $H^1$ , by the Banach-Alaoglu theorem there exists a subsequence  $u_{k_j} \rightarrow w$  weakly in  $H^1$  (and  $L^2$ ), for some  $w \in H_0^1$ . From example sheet 2, we know

$$\|s\|_{H^1}^2 \leq \liminf_j \|u_{k_j}\|_{H^1}^2, \quad \lim_j \langle f, u_{k_j} \rangle_{L^2} = \langle f, w \rangle_{L^2}.$$

Then  $S(w) \geq \sigma$ , and also

$$\begin{aligned} S(w) &= \|w\|_{H^1}^2 - 2\langle f, w \rangle_{L^2} \leq \liminf_j (\|u_{k_j}\|_{H^1}^2 - 2\langle f, u_{k_j} \rangle_{L^2}) \\ &= \liminf_j S(u_{k_j}) = \sigma, \end{aligned}$$

so  $S(w) = \sigma$ . This property is lower semicontinuity, i.e. we can estimate the value of a functional by a limit inferior.

To show uniqueness, it suffices to show that  $w$  solves the PDE. Take  $v \in H_0^1$ ,  $t \in \mathbb{R}$ . Then

$$S(w + tv) \geq S(w) \implies \left. \frac{d}{dt} S(w + tv) \right|_{t=0} = 0,$$

so

$$\begin{aligned} S(w + tv) &= \langle w + tv, w + tv \rangle_{H^1} - 2\langle f, tw \rangle_{L^2} \\ &= \|w\|_{H^1}^2 + t^2\|v\|_{H^1}^2 - 2t(\langle w, v \rangle_{H^1} - \langle f, v \rangle_{L^2}) - 2\langle f, w \rangle_{L^2}, \end{aligned}$$

so this forces

$$\langle w, v \rangle_{H^1} = \langle f, v \rangle_{L^2}$$

for all  $v \in H_0^1$ , and thus solves  $(\dagger')$ .

*Remark.* All that is required in this proof is that  $S$  is lower semicontinuous. For instance, this works for

$$T(u) = \|u - f\|_{H^1}^2.$$

This is the direct method of the calculus of variations.

# Index

- absolutely continuous, 13
- bidual space, 15
- compactly supported, 33
- convolution, 33
- Dirac measure, 30
- distribution, 29
- distributional derivative, 31
- distributional Fourier transform, 41
- Fréchet space, 19
- fundamental cell, 44
- fundamental solution, 35
- Green function, 35
- Hölder spaces, 50
- Hölder-continuous, 20
- heat kernel, 60
- inverse Fourier transform, 41
- Lebesgue point, 9
- locally convex, 18
- locally in  $L^p$ , 5
- mean, 45
- orthogonal, 13
- orthonormal, 13
- orthonormal basis, 13
- periodic, 43
- periodic partition of unity, 44
- Radon-Nikodym derivative, 13
- rapidly decreasing, 29
- reflexive, 15
- regular, 17
- Schwartz class, 29
- Schwartz distribution, 29
- semi-norm, 18
- separating semi-norms, 18
- slowly increasing, 40
- Sobolev space, 50
- strong topology, 18
- sub-linear, 23
- support, 33
- tempered distribution, 30
- test function, 29
- topological dual space, 15
- vague convergence, 20
- weak derivative, 31
- weak topology, 19
- weak-\* topology, 19