

# 1 Lebesgue Integration Theory

## 1.1 Review of measure theory

**Definition.** Given a set  $E$ , a  $\sigma$ -algebra on  $E$  is a collection  $\mathcal{E}$  of subsets of  $E$  such that:

- (i)  $E \in \mathcal{E}$ ;
  - (ii)  $A \in \mathcal{E} \Rightarrow A^c = E \setminus A \in \mathcal{A}$ ;
  - (iii)  $A_n \in \mathcal{E}, n \in \mathbb{N} \Rightarrow \bigcup_{n \in \mathbb{N}} A_n \in \mathcal{E}$ .
- $(E, \mathcal{E})$  is called a *measurable space*, and any  $A \in \mathcal{E}$  is called a *measurable set*.

Given a collection  $\mathcal{A}$  of subsets of  $E$ ,  $\sigma(\mathcal{A})$  is the smallest  $\sigma$ -algebra containing  $\mathcal{A}$ .

**Definition.** A *measure* on  $(E, \mathcal{E})$  is a function  $\mu : \mathcal{E} \rightarrow [0, \infty]$  such that

- (i)  $\mu(\emptyset) = 0$ ;
- (ii)  $A_n \in \mathcal{E}, n \in \mathbb{N} \text{ disjoint} \Rightarrow \mu(\bigcup_{n \in \mathbb{N}} A_n) = \sum_{n \in \mathbb{N}} \mu(A_n)$ .

$(E, \mathcal{E}, \mu)$  is called a *measure space*.

**Definition** (Borel measure). If  $(E, \tau)$  is a topological space, then  $\sigma(\tau)$  is called a *Borel algebra*, denoted  $\mathcal{B}(E)$ , and a measure on  $(E, \mathcal{B}(E))$  is called a *Borel measure*.

**Example.**  $E = \mathbb{R}^n$ ,  $\mu$  the Lebesgue measure satisfying  $\mu((a_1, b_1) \times \dots \times (a_n, b_n)) = (b_1 - a_1) \dots (b_n - a_n)$ .

**Notation:** we write  $\mu(dx) = dx$  and  $\mu(A) = |A|$  when  $\mu$  is the Lebesgue measure.

**Definition** (Measurable function). Let  $(E, \mathcal{E})$  and  $(F, \mathcal{F})$  be measurable spaces. Then  $f : E \rightarrow F$  is *measurable* if  $f^{-1}(A) \in \mathcal{E}$  whenever  $A \in \mathcal{F}$ . If  $(E, \mathcal{E})$  and  $(F, \mathcal{F})$  are Borel algebras, a measurable function is called a *Borel function*. Special case:  $(F, \mathcal{F}) = ([0, \infty], \mathcal{B}([0, \infty]))$ , then  $f : E \rightarrow F$  is called a *non-negative measurable function*.

**Fact.** The class of measurable functions is closed under addition, multiplication and taking (pointwise) limits.

**Definition.**  $f : E \rightarrow F$  ( $F = [0, \infty]$  or  $\mathbb{R}^n$  or  $\mathbb{C}^n$ ) is a *simple function* if  $f = \sum_{k=1}^K a_k \mathbb{1}_{A_k}$  for some  $K \in \mathbb{N}$ ,  $a_k \in F$ ,  $A_k \in \mathcal{E}$ . For a simple function, the integral is

$$\int f d\mu = \int f(x) d\mu(x) = \sum_{k=1}^K a_k \mu(A_k) \quad (0 \cdot \infty := 0).$$

For a non-negative measurable  $f$ , we define

$$\int f d\mu = \sup \left\{ \int g d\mu : g \text{ simple}, 0 \leq g \leq f \right\}.$$

**Definition.** A measurable function  $f : E \rightarrow \mathbb{R}$  is said to be *integrable* if  $\int |f| d\mu < \infty$ . Write  $f = f_+ - f_-$  with  $f_{\pm}$  non-negative, measurable,  $\int f_{\pm} d\mu < \infty$ , and then  $\int f d\mu = \int f_+ d\mu - \int f_- d\mu$ . For  $f : E \rightarrow \mathbb{R}^n$ , this is applied in each component.

**Theorem** (Monotone convergence theorem). *Let  $(E, \mathcal{E}, \mu)$  be a measure space, and let  $(f_n)$  be a (pointwise) increasing sequence of non-negative functions on  $E$  converging to  $f$ . Then*

$$\lim_{n \rightarrow \infty} \int_E f_n d\mu = \int_E f d\mu.$$

**Theorem** (Dominated convergence theorem). *Let  $(f_n)$  be a sequence of measurable functions on a measure space  $(E, \mathcal{E}, \mu)$  such that:*

- (i)  $f_n \rightarrow f$  pointwise almost everywhere;
- (ii)  $|f_n| \leq g$  almost everywhere for some integrable  $g$ .

*Then*

$$\lim_{n \rightarrow \infty} \int_E f_n d\mu = \int_E f d\mu.$$

## 1.2 $L^p$ spaces

**Definition.** Let  $(E, \mathcal{E}, \mu)$  be a measure space. For  $p \in [1, \infty)$  and  $f : E \rightarrow \mathbb{R}$  define

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

and

$$\|f\|_{L^\infty} = \text{esssup}|f| = \inf\{K : |f| \leq K \text{ a.e.}\}.$$

The space  $L^p$ ,  $p \in [1, \infty]$  is defined by

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

Where  $f \sim g$  if  $f = g$  a.e.

**Theorem** (Riesz-Fisher theorem).  *$L^p$  is a Banach space for all  $p \in [1, \infty]$ .*

**Notation:** when  $E = \mathbb{R}^n$ ,  $\mu$  the Lebesgue measure, write  $L^p(E, \mu) = L^p(\mathbb{R}^n)$ .

**Fact.** For  $p \in [1, \infty)$ , the simple functions  $f$  with  $\mu(\{x : f(x) \neq 0\}) < \infty$  are dense in  $L^p$ . For  $p = \infty$  we can drop the condition on the measure of the support.

**Definition.** For  $f, g : \mathbb{R}^n \rightarrow \mathbb{R}$ , the *convolution*  $f * g$  is defined by

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

provided the integral exists. Note that  $f * g = g * f$ , convolution is associative, and  $\mu(f * g) = \mu(f)\mu(g)$ .

**Theorem.**  $C_c^\infty(\mathbb{R}^n)$  is dense in  $L^p(\mathbb{R}^n)$  for  $p \in [1, \infty)$ .

Before we prove the theorem, we will need some preliminary results.

**Remark.** This theorem is false for  $p = \infty$ .

**Notation:** a multiindex is  $\alpha = (\alpha_1, \dots, \alpha_n) \in \mathbb{Z}_{\geq 0}^n$ . Set  $|\alpha| = \alpha_1 + \dots + \alpha_n$ ;  $\alpha! = \alpha_1! \dots \alpha_n!$ ;  $X^\alpha = X_1^{\alpha_1} \dots X_n^{\alpha_n}$  for  $X \in \mathbb{R}^n$ ;  $\nabla^\alpha f = D^\alpha f = \frac{\partial^\alpha f}{\partial x^\alpha} = \frac{\partial^{|\alpha|} f}{\partial x_1^{\alpha_1} \dots \partial x_n^{\alpha_n}}$ .

**Definition.** We say  $f \in L_{\text{loc}}^p(\mathbb{R}^n)$  if  $f \mathbb{1}_K \in L^p(\mathbb{R}^n)$  for any  $K \subseteq \mathbb{R}^n$  compact.

**Proposition.** Let  $f \in L_{\text{loc}}^1(\mathbb{R}^n)$ ,  $g \in C_c^k(\mathbb{R}^n)$ , some  $k \geq 0$ . Then  $f * g \in C^k(\mathbb{R}^n)$  and  $\nabla^\alpha(f * g) = f * (\nabla^\alpha g)$  for all  $|\alpha| \leq k$ .

*Proof.* First we check for  $k = 0$ . Set  $T_z f(x) = f(x - z)$ ,  $z \in \mathbb{R}^n$ . Then  $T_z(f * g) = f * (T_z g)$ . Also  $T_z g(x) \rightarrow g(x)$  for all  $x$  as  $z \rightarrow 0$  (continuity of  $g$ ). Furthermore  $|T_z g(x)| \leq \|g\|_{L^\infty} \mathbb{1}_{B_R(0)}(x)$  if  $|x| + 1 \leq R$ ,  $|z| < 1$  (we can just take  $R$  large enough so it holds everywhere since  $g$  has compact support). Then  $|f(y)T_z g(x - y)| \leq C|f(y)|\mathbb{1}_{B_R(0)}(x - y)$ , for  $C := \|g\|_{L^\infty}$ .

Since  $f \in L_{\text{loc}}^1(\mathbb{R}^n)$ ,  $|f(y)|\mathbb{1}_{B_R(0)}(x - y)$  is integrable in  $y$ , so by the dominated convergence theorem,

$$T_z(f * g) = (f * T_z g)(x) = \int_{\mathbb{R}^n} f(y)T_z g(x - y)dy \xrightarrow{z \rightarrow 0} \int_{\mathbb{R}^n} f(y)g(x - y)dy = (f * g)(x).$$

And so  $f * g \in C^0$ . Now let  $k = 1$ . Let  $\nabla_i^h g(x) = \frac{g(x + h e_i) - g(x)}{h}$ , where  $e_i$  is the  $i$ th unit vector. Then  $\nabla_i^h g(x) \rightarrow \nabla_i g(x)$  as  $h \rightarrow 0$ .

By the mean value theorem, there exists  $t \in [-h, h]$  such that

$$\nabla_i^h g(x) = \nabla_i g(x + t e_i) \Rightarrow |\nabla_i^h g(x)| \leq \|\nabla_i g\|_{L^\infty} \mathbb{1}_{B_R(0)}(x).$$

Again by the dominated convergence theorem,  $\nabla_i^h(f * g) = f * (\nabla_i^h g) \rightarrow f * \nabla_i g$ . Thus  $f * g \in C^1$ . The case  $k > 1$  is similar, with induction.  $\square$

**Proposition** (Minkowski's integral inequality). Let  $p \in [1, \infty)$  and  $f : \mathbb{R}^n \times \mathbb{R}^n \rightarrow \mathbb{R}$  Borel. Then

$$\left[ \int_{\mathbb{R}^n} \left| \int_{\mathbb{R}^n} f(x, y) dx \right|^p dy \right]^{1/p} \leq \int_{\mathbb{R}^n} \left[ \int_{\mathbb{R}^n} |f(x, y)|^p dy \right]^{1/p} dx.$$

*Proof.* Example sheet 1.  $\square$

**Proposition.** Let  $p \in [1, \infty)$ ,  $g \in L^p(\mathbb{R}^n)$ . Then

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

**Remark.** This is not true for  $p = \infty$ . Let  $\theta(x) = \mathbb{1}_{x \geq 0}$ . Then  $\|T_z \theta - \theta\|_{L^\infty} = 1$  if  $z \neq 0$ .

*Proof.* Consider first  $g = \mathbb{1}_R$ ,  $R$  a rectangle. Then the result is clear. Hence it also follows for any finite union of rectangles. If  $B$  is a Borel set,  $|B| < \infty$ , then for every  $\varepsilon > 0$ , there exists a finite union of rectangles  $R$  such that

$$\|\mathbb{1}_B - \mathbb{1}_R\|_{L^p} = |B \triangle R|^{1/p} < \varepsilon.$$

Hence

$$\|T_z \mathbb{1}_B - \mathbb{1}_B\|_{L^p} \leq \underbrace{\|T_z \mathbb{1}_B - T_z \mathbb{1}_R\|_{L^p}}_{=\|\mathbb{1}_B - \mathbb{1}_R\|_{L^p} < \varepsilon} + \underbrace{\|T_z \mathbb{1}_R - \mathbb{1}_R\|_{L^p}}_{< \varepsilon \text{ for } |z| \text{ small}} + \underbrace{\|\mathbb{1}_R - \mathbb{1}_B\|_{L^p}}_{< \varepsilon}.$$

Thus the result holds for  $g = \mathbb{1}_B$ ,  $B \in \mathcal{B}(\mathbb{R}^n)$ . Thus the result holds for simple functions  $g$ . Finally, for any  $g \in L^p$ , there is a  $\tilde{g}$  simple such that  $\|g - \tilde{g}\|_{L^p} < \varepsilon$ . Then

$$\|T_z g - g\|_{L^p} \leq \underbrace{\|T_z g - T_z \tilde{g}\|_{L^p}}_{=\|g - \tilde{g}\|_{L^p} < \varepsilon} + \underbrace{\|T_z \tilde{g} - \tilde{g}\|_{L^p}}_{< \varepsilon \text{ for } |z| \text{ small}} + \underbrace{\|g - \tilde{g}\|_{L^p}}_{< \varepsilon}.$$

□

**Theorem.** Let  $\varphi \in C_c^\infty(\mathbb{R}^n)$  be such that  $\varphi \geq 0$  and  $\int_{\mathbb{R}^n} \varphi dx = 1$  and set  $\varphi_\varepsilon(x) = \varepsilon^{-n} \varphi(x/\varepsilon)$ . Then for any  $g \in L^p$ ,  $p \in [1, \infty)$ , it follows that  $\varphi_\varepsilon * g \in C^\infty(\mathbb{R}^n)$  and  $\varphi_\varepsilon * g \rightarrow g$  in  $L^p$ .

*Proof.* We have

$$\begin{aligned} |\varphi_\varepsilon * g(x) - g(x)| &= \left| \int_{\mathbb{R}^n} [\varphi_\varepsilon(y)g(x-y) - g(x)] dy \right| \\ &\stackrel{z:=y/\varepsilon}{=} \left| \int_{\mathbb{R}^n} \varphi(z) [g(x-\varepsilon z) - g(x)] dz \right| \\ &\leq \int_{\mathbb{R}^n} \varphi(z) |T_{\varepsilon z}g(x) - g(x)| dz. \end{aligned}$$

Hence

$$\begin{aligned} \|\varphi_\varepsilon * g - g\|_{L^p} &= \left( \int_{\mathbb{R}^n} \underbrace{|\varphi_\varepsilon * g - g|^p}_{\int_{\mathbb{R}^n} \varphi(z) |T_{\varepsilon z}g - g|^p dz} dx \right)^{1/p} \\ &\leq \int_{\mathbb{R}^n} \left( \int_{\mathbb{R}^n} \varphi(z)^p |T_{\varepsilon z}g(x) - g(x)|^p dx \right)^{1/p} dz \\ &= \int_{\mathbb{R}^n} \varphi(z) \underbrace{\|T_{\varepsilon z}g - g\|_{L^p}}_{\rightarrow 0 \text{ as } \varepsilon \rightarrow 0} dz \end{aligned}$$

where the inequality comes from Minkowski's inequality. Also the final term tends to 0 as  $\varepsilon \rightarrow 0$  by the DCT since  $\varphi(z)\|T_{\varepsilon z}g - g\|_{L^p} \leq 2\varphi(z)\|g\|_{L^p}$  and  $\varphi$  is integrable.  $\square$

**Definition.**  $\varphi$  as above is called a (smooth) mollifier.

**Corollary.**  $C_c^\infty(\mathbb{R}^n)$  is dense in  $L^p(\mathbb{R}^n)$ ,  $p \in [1, \infty)$ .

*Proof.* The previous theorem implies  $C^\infty(\mathbb{R}^n) \cap L^p(\mathbb{R}^n)$  is dense in  $L^p$ . Since  $\|f - f\mathbb{1}_{B_R(0)}\|_{L^p} \rightarrow 0$  as  $R \rightarrow \infty$  by the DCT, for  $f \in L^p$ , applying the theorem with  $g = f\mathbb{1}_{B_R(0)}$  it follows that  $C_c^\infty(\mathbb{R}^n)$  is dense in  $L^p$ .  $\square$

### 1.3 Lebesgue Differentiation Theorem

Recall:

**Theorem** (Fundamental Theorem of Calculus). For  $f : \mathbb{R} \rightarrow \mathbb{R}$  continuous,  $F(x) := \int_0^x f(t)dt$  is differentiable with  $F'(x) = f(x)$ .

We actually have a stronger result:

**Theorem** (Lebesgue Differentiation Theorem). For  $f : \mathbb{R}^n \rightarrow \mathbb{R}$  integrable,

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

The  $x$  for which this holds are called Lebesgue points.

We will need a few preliminary results and definitions before we can prove this.

**Corollary.** If  $g \in L^1(\mathbb{R})$  and  $G(x) = \int_{-\infty}^x g(t)dt$ , then  $G$  is differentiable for almost every  $x$  with  $G'(x) = g(x)$ .

**Corollary.** If  $\varphi$  is a smooth mollifier and  $g \in L^p(\mathbb{R}^n)$ , then  $\varphi_\varepsilon * g \xrightarrow{\varepsilon \rightarrow 0} g$  almost everywhere.

**Definition.** For  $f : \mathbb{R}^n \rightarrow \mathbb{R}$  integrable, the *Hardy-Littlewood Maximal Function*  $Mf : \mathbb{R}^n \rightarrow [0, \infty]$  is defined by

$$Mf(x) = \sup_{r>0} \frac{1}{|B_r(x)|} \int_{B_r(x)} |f(y)|dy.$$

**Remark.** We sometimes write  $\int_{B_r(x)} |f(y)|dy$  for  $\frac{1}{|B_r(x)|} \int_{B_r(x)} |f(y)|dy$ .

**Lemma** (Wiener's covering lemma). If  $K$  is compact and  $K \subseteq \bigcup_{i=1}^N B_i$  for open balls  $(B_i)_{i=1}^N$ , there exists a subcollection  $(B_{i_k})_k$  of disjoint balls such that

$$\left| \bigcup_{i=1}^N B_i \right| \leq 3^n \sum_k |B_{i_k}|.$$

*Proof.* Example sheet. □

**Proposition.** Take  $f \in L^1(\mathbb{R}^n)$ . Then  $Mf$  is a Borel function, finite almost everywhere, and

$$|\underbrace{\{Mf > \lambda\}}_{:=A_\lambda}| \leq \frac{3^n}{\lambda} \|f\|_{L^1}.$$

*Proof.* For each  $x \in A_\lambda$ , there exists  $r_x > 0$  such that

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

We claim that  $A_\lambda$  is open. Then we will have shown  $Mf$  is Borel as the  $A_\lambda = (Mf)^{-1}((\lambda, \infty])$  are open, and the sets  $(\lambda, \infty]$  generate the Borel  $\sigma$ -algebra.

We'll actually show  $A_\lambda^c$  is closed. Suppose  $(x_k)_{k \geq 1}$  is a sequence in  $A_\lambda^c$  with  $x_k \rightarrow x$ . Suppose  $x \in A_\lambda$ . By the Dominated Convergence Theorem,

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

Since  $x_k \notin A_\lambda$ , the LHS is  $\leq \lambda$  for all  $k$ , but the RHS is  $> \lambda$  which is impossible. Hence  $x \in A_\lambda^c$  and  $A_\lambda^c$  is closed.

To prove the inequality, let  $K \subseteq A_\lambda$  be compact. Since  $\{B_{r_x}(x)\}_{x \in A_\lambda}$  is an open cover of  $K$ , there exists a finite subcover  $K \subseteq \bigcup_{i=1}^N B_i$ , where  $B_i = B_{r_x}(x)$  for

some  $x \in A_\lambda$ . Now take a subcollection  $(B_{i_k})_k$  of disjoint balls as in Wiener's covering lemma.

Since  $\frac{1}{|B_i|} \int_{B_i} |f(y)| dy > \lambda$ , it follows that  $|B_i| < \frac{1}{\lambda} \int_{B_i} |f(y)| dy$ . Hence

$$|K| \leq 3^n \sum_k |B_{i_k}| < \frac{3^n}{\lambda} \sum_k \int_{B_{i_k}} |f(y)| dy \leq \frac{3^n}{\lambda} \int_{\mathbb{R}^n} |f(y)| dy.$$

Since this holds for any  $K \subseteq A_\lambda$  compact, by regularity of the Lebesgue measure, it also holds for  $A_\lambda$ . In particular,  $|\{Mf = \infty\}| \leq |\{Mf > \lambda\}| \xrightarrow{\lambda \rightarrow \infty} 0$ , i.e.  $Mf < \infty$  almost everywhere.  $\square$



Now we are ready to prove:

**Theorem** (Lebesgue Differentiation Theorem). *For  $f : \mathbb{R}^n \rightarrow \mathbb{R}$  integrable,*

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

*The  $x$  for which this holds are called Lebesgue points.*

*Proof.* Let

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

Then it suffices to show  $|A_\lambda| = 0$  for any  $\lambda > 0$ . Indeed, the non-Lebesgue points are then  $\bigcup_n A_{1/n}$ , a countable union of sets of measure 0.

Given  $\varepsilon > 0$ , let  $g \in C_c^\infty(\mathbb{R}^n)$  be such that  $\|f - g\|_{L^1} < \varepsilon$ . Then

$$\begin{aligned} \int_{B_r(x)} |f(y) - f(x)| dy &\leq \underbrace{\int_{B_r(x)} |f(y) - g(y)| dy}_{\leq M(f-g)(x)} + |f(x) - g(x)| + \underbrace{\int_{B_r(x)} |g(x) - g(y)| dy}_{\rightarrow 0 \text{ since } g \in C^\infty} \\ \implies \limsup_{r \rightarrow 0} \int_{B_r(x)} |f(y) - f(x)| dy &\leq M(f-g)(x) + |f(x) - g(x)|. \end{aligned}$$

If  $x \in A_\lambda$ , then either  $M(f-g)(x) > \lambda$  or  $|f(x) - g(x)| > \lambda$ . The Hardy-Littlewood maximal inequality says  $|\{M(f-g) > \lambda\}| \leq \frac{3^n}{\lambda} \|f-g\|_{L^1}$ . Then by Markov's inequality  $|\{|f-g| > \lambda\}| \leq \frac{1}{\lambda} \|f-g\|_{L^1}$ . Hence

$$|A_\lambda| \leq \frac{3^n + 1}{\lambda} \|f - g\|_{L^1} < \frac{3^{n+1} + 1}{\lambda} \varepsilon.$$

Since  $\varepsilon > 0$  was arbitrary,  $|A_\lambda| = 0$ . □

## 1.4 Littlewood's Principles

**Theorem** (Egorov). *Let  $E \subseteq \mathbb{R}^n$ ,  $|E| < \infty$ , and  $f_k : E \rightarrow \mathbb{R}$ ,  $k \geq 1$  be a sequence of measurable functions such that  $f_k \rightarrow f$  almost everywhere. Then for every  $\varepsilon > 0$ , there is a closed subset  $A_\varepsilon \subseteq E$  such that  $|E \setminus A_\varepsilon| < \varepsilon$  and  $f_k \rightarrow f$  uniformly on  $A_\varepsilon$ .*

*Proof.* Without loss of generality,  $f_k(x) \rightarrow f(x)$  for all  $x \in E$  (otherwise restrict to a subset of  $E$  of full measure). Let

$$E_k^n = \left\{ x \in E : |f_j(x) - f(x)| < \frac{1}{n} \ \forall j > k \right\}.$$

Then  $E_{k+1}^n \supseteq E_k^n$ ,  $\bigcup_k E_k^n = E$ , hence  $|E_k^n| \uparrow |E|$  as  $k \rightarrow \infty$ . Let  $k_n$  be such that  $|E \setminus E_{k_n}^n| < 2^{-n}$  and for  $N \in \mathbb{N}$  set

$$A_N = \bigcap_{n \geq N} E_{k_n}^n \implies |E \setminus A_N| \leq \sum_{n \geq N} |E \setminus E_{k_n}^n| \leq 2^{-N+1} < \varepsilon \text{ for } N = N_\varepsilon.$$

Now it suffices to show  $f_j \rightarrow f$  uniformly on  $A_N$ . Indeed, for  $x \in A_N$  and any  $n \geq N$ ,  $|f_j(x) - f(x)| < \frac{1}{n}$  for all  $j > k_n$ . Hence  $\limsup_{j \rightarrow \infty} \sup_{A_N} |f_j - f| \leq \frac{1}{n}$  for all  $n \geq N$ , hence  $\lim_{j \rightarrow \infty} \sup_{A_N} |f_j - f| = 0$ .  $\square$

**Theorem** (Lusin). *Let  $f : E \rightarrow \mathbb{R}$  be a Borel function, where  $E \subseteq \mathbb{R}^n$  and  $|E| < \infty$ . Then for every  $\varepsilon > 0$ , there exists  $F_\varepsilon \subseteq E$  closed such that  $|E \setminus F_\varepsilon| < \varepsilon$  and  $f|_{F_\varepsilon}$  is continuous.*

**Remark.** Careful: this does not mean that  $f$  is continuous at  $x \in F_\varepsilon$  in the topology of  $\mathbb{R}^n$ .

*Proof.* First we show that the statement holds for simple functions  $f$ . Let  $f = \sum_{m=1}^M a_m \mathbb{1}_{A_m}$  with the  $A_m$  disjoint and  $\bigcup_m A_m = E$ . Then there are compact sets  $K_m \subseteq A_m$  with  $|A_m \setminus K_m| < \frac{\varepsilon}{M}$  by regularity of the Lebesgue measure. Then if  $F_\varepsilon = \bigcup_m K_m$ ,  $|E \setminus F_\varepsilon| < \varepsilon$ . Since  $f$  is constant on each  $K_m$ , and the distance between  $K_m$  and  $K_{m'}$  is strictly positive for  $m \neq m'$  (compactness), this implies  $f|_{F_\varepsilon}$  is continuous.

Now we show the statement holds for any measurable  $f$ . Let  $f_n$  be simple functions such that  $f_n \rightarrow f$  almost everywhere, and  $C_n \subseteq E$  be such that  $|C_n| < 2^{-n}$  and  $f_n|_{E \setminus C_n}$  is continuous for all  $n$ . By Egorov's Theorem, there exists  $A_\varepsilon$  such that  $f_n \rightarrow f$  uniformly on  $A_\varepsilon$  and  $|E \setminus A_\varepsilon| < \varepsilon$ . Set  $F'_\varepsilon = A_\varepsilon \setminus \bigcup_{n \geq N} C_n$  so  $|E \setminus F'_\varepsilon| < 2\varepsilon$  for  $N = N_\varepsilon$  sufficiently large. Since  $f_n|_{F'_\varepsilon}$ ,  $n \geq N$  is continuous  $f_n \rightarrow f$  uniformly on  $F'_\varepsilon$ ,  $f|_{F'_\varepsilon}$  is continuous.

By regularity of the Lebesgue measure, there exists  $F_\varepsilon \subseteq F'_\varepsilon$  closed with  $|F'_\varepsilon \setminus F_\varepsilon| < \varepsilon$  so  $|E \setminus F_\varepsilon| < 3\varepsilon$  and we are done.  $\square$

## 2 Banach and Hilbert space analysis

### 2.1 The Hilbert space $L^2$

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

$$\langle f, g \rangle_{L^2} = \int_E \overline{f} g d\mu.$$

**Definition.** A subset  $S = \{u_j\}_{j \in J} \subseteq H$  of a Hilbert space  $H$  is

- *Orthogonal* if  $\langle u_j, u_k \rangle = 0$  for all  $j \neq k$ ;
- *Orthonormal* if it is orthogonal and  $\langle u_j, u_j \rangle = 1$  for all  $j$ ;
- *Complete* if  $\overline{\text{span}\{u_j\}} = H$ .

A complete orthonormal set is called a *Hilbert basis*.

**Fact.** A Hilbert space is separable (i.e there is a countable dense subset) if and only if there is a countable orthonormal (Hilbert) basis.

**Examples.**

- (i)  $L^2([-\pi, \pi])$ ,  $S = \left\{ \frac{1}{\sqrt{2\pi}} e^{-inx} \right\}_{n \in \mathbb{Z}}$ . Then  $S$  is a Hilbert basis; the Fourier basis (completeness follows from the Stone-Weierstrass theorem & density of  $C^\infty$ ).
- (ii)  $L^2(\mathbb{R})$ ,  $S = \{\psi_{n,k}\}_{n,k \in \mathbb{Z}}$  where

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

$$\psi(x) = \begin{cases} 1 & x \in [0, 1/2) \\ -1 & x \in [1/2, 1) \\ 0 & \text{otherwise} \end{cases}.$$

$S$  is a Hilbert basis; the *Haar system*.

- (iii)  $L^2(\mathbb{R}, \mu(dx))$ , where  $\mu(dx) = (2\pi)^{-1/2} \exp(x^2/2) dx$ ; the Gauss measure. Then take  $\{H_n\}_{n \in \mathbb{Z}_{\geq 0}}$ , where the  $H_n$  are obtained by applying Gram-Schmidt to  $\{1, x, x^2, \dots\}$ ; the Hermite polynomials. Then  $\{H_n\}_{n \in \mathbb{Z}_{\geq 0}}$  is a Hilbert basis.

**Theorem** (Reisz representation theorem). *For any bounded linear functional  $\Lambda : H \rightarrow \mathbb{R}$  (respectively  $\mathbb{C}$ ), there is a unique  $w \in H$  such that  $\Lambda(u) = \langle w, u \rangle$  for all  $u \in H$ .*

## 2.2 Radon-Nikodym Theorems

**Definition.** Let  $(E, \mathcal{E})$  be a measurable space and let  $\mu, \nu$  be two measures on  $(E, \mathcal{E})$ . Then  $\nu$  is said to be *absolutely continuous* with respect to  $\mu$ , written  $\nu \ll \mu$ , if for all  $A \in \mathcal{E}$ ,  $\nu(A) = 0$  whenever  $\mu(A) = 0$ . Two measures  $\mu, \nu$  are *mutually singular*, written  $\mu \perp \nu$  if there is  $B \in \mathcal{E}$  such that  $\mu(B) = 0 = \nu(B^c)$ .

**Theorem (Radon-Nikodym).** Let  $\mu$  and  $\nu$  be finite measures on  $(E, \mathcal{E})$  with  $\nu \ll \mu$ . Then there exists  $\omega \in L^1(E, \mathcal{E})$  such that for all  $A \in \mathcal{E}$ ,

$$\nu(A) = \int_A \omega d\mu.$$

Equivalently, for all  $h : E \rightarrow [0, \infty]$  Borel,

$$\int h d\nu = \int h \omega d\mu.$$

*Proof.* Set  $\alpha = \mu + 2\nu$  and  $\beta = 2\mu + \nu$ . Define

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

Then

$$|\Lambda(f)| \leq \int_E |f| d\beta \leq 2 \int_E |f| d\alpha \leq 2\sqrt{\alpha(E)} \|f\|_{L^2(E, \alpha)}.$$

So  $\Lambda : L^2(E, \alpha) \rightarrow \mathbb{R}$  is bounded and linear. So by the Riesz representation theorem, there is  $g \in L^2(E, \alpha)$  such that  $\Lambda(f) = \langle g, f \rangle_{L^2(E, \alpha)}$  for all  $f \in L^2(E, \alpha)$ . Hence  $\int f d\beta = \int g f d\alpha$ , and

$$\int f(2d\mu + d\nu) = \int g f(d\mu + 2d\nu) \iff \int f(2 - g)d\mu = \int f(2g - 1)d\nu.$$

We claim that  $g$  takes values in  $[1/2, 2]$   $\mu$ -a.e and  $\nu$ -a.e, and that  $g \neq 1/2$   $\mu$ -a.e (this implies  $g \neq 1/2$   $\nu$ -a.e since  $\nu \ll \mu$ ). Assuming the claim, the proof is completed as follows; by the monotone convergence theorem,  $(*)$  can be extended to all  $f : E \rightarrow [0, \infty]$ . Given  $h : E \rightarrow [0, \infty]$  measurable, set

$$f(x) = \frac{h(x)}{2g(x) - 1}, \quad \omega(x) = \frac{2 - g(x)}{2g(x) - 1}, \quad x \in \{g \neq 1/2\}. \quad (*)$$

Then

$$\int h d\nu = \int f(2g - 1)d\nu = \int f(2g - 1)d\mu = \int h \omega d\mu.$$

In particular, taking  $h = 1$ , we see  $\omega \in L^1(E, \mu)$ .

Now we prove the claim: let  $f = \mathbb{1}_{A_j}$ , with  $A_j = \{x \in E : g(x) < \frac{1}{2} - \frac{1}{j}\}$ . Then we have

$$\int f(2g - 1)d\nu \leq -\frac{2}{j}\nu(A_j),$$

$$\int f(2 - g)d\mu \geq \frac{3}{2}\mu(A_j),$$
$$\implies \frac{3}{2}\mu(A_j) \leq -\frac{2}{j}\nu(A_j) \implies \mu(A_j) = \nu(A_j) = 0.$$

Implying  $g \geq 1/2$  both  $\mu$ -a.e and  $\nu$ -a.e. To show  $g \leq 2$   $\mu$ -a.e and  $\nu$ -a.e the proof is analogous, instead with  $A_j = \{x \in E : g(x) \geq 2 + 1/j\}$ . To show  $\mu(\{g = 1/2\}) = 0$ , set  $f = \mathbb{1}_Z$ ,  $Z = \{g = 1/2\}$  in (\*), giving

$$\frac{3}{2} \int \mathbb{1}_{\{g = 1/2\}} d\mu = 0.$$

□

### 2.3 The dual of $L^p$

**Definition.** A *topological vector space* (TVS)  $X$  is a vector space together with a topology in which  $(x, y) \mapsto x + y$  and  $(\lambda, x) \mapsto \lambda x$  are continuous. The *dual space*  $X'$  is the linear space of continuous linear maps  $\Lambda : X \rightarrow \mathbb{R}$  (or  $\mathbb{C}$ ).

If  $X$  is a normed vector space equipped with the topology induced by the norm, then linear maps on  $X$  are bounded if and only if they are continuous. We can define a norm on  $X'$  by

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

Then  $X'$  is a Banach space (even if  $X$  isn't).

We aim to identify  $L^p(\mathbb{R}^n)'$  with  $L^q(\mathbb{R}^n)$ , where  $\frac{1}{p} + \frac{1}{q} = 1$ , if  $p \in [1, \infty)$ .

**Proposition.** Let  $q \in [1, \infty]$ . For every  $g \in L^q(\mathbb{R}^n)$ ,

$$\Lambda_g(f) = \int \bar{f}g dx$$

defines  $\Lambda_g \in L^p(\mathbb{R}^n)'$  with  $\|\Lambda_g\| = \|g\|_{L^q}$ .

*Proof.* By Hölder's inequality,  $|\Lambda_g(f)| = \|f\|_{L^p} \|g\|_{L^q}$ . Hence  $\Lambda_g \in L^p(\mathbb{R}^n)'$  and  $\|\Lambda_g\| \leq \|g\|_{L^q}$ . Equality: see Example sheet 1.  $\square$

**Corollary.** The map  $J : L^q(\mathbb{R}^n) \rightarrow L^p(\mathbb{R}^n)'$ ,  $g \mapsto \Lambda_g$  is a linear isometry. Thus we can identify  $L^q(\mathbb{R}^n)$  as a subspace of  $L^p(\mathbb{R}^n)'$ .

**Remark.** When  $p = 2$  then  $L^2(\mathbb{R}^n)' = L^2(\mathbb{R}^n)$ , i.e  $J$  is surjective (Riesz representation theorem).

**Theorem.** Let  $p \in [1, \infty)$ ,  $\frac{1}{p} + \frac{1}{q} = 1$ . Then  $J$  is surjective, i.e  $L^p(\mathbb{R}^n)' = L^q(\mathbb{R}^n)$ .

**Remarks.**

1.  $L^1(\mathbb{R}^n)' = L^\infty(\mathbb{R}^n)$ , but  $L^\infty(\mathbb{R}^n)' \neq L^1(\mathbb{R}^n)$ ;
2. The same is true if  $\mathbb{R}^n$  is replaced by  $U \subseteq \mathbb{R}^n$  open.

**Definition.**  $\Lambda \in L^p(\mathbb{R}^n)'$  is *positive* if

$$\Lambda(f) \geq 0 \text{ for all } f \in L^p(\mathbb{R}^n) \text{ such that } f \geq 0 \text{ a.e.}$$

**Lemma.** Let  $\Lambda \in L^p(\mathbb{R}^n)'$  be positive. Then there is  $g \in L^q(\mathbb{R}^n)$  non-negative with

$$\Lambda(f) = \int_{\mathbb{R}^n} f g dx \text{ for all } f \in L^p(\mathbb{R}^n).$$

Furthermore  $\|g\|_{L^q} = \|\Lambda\|$ .

*Proof.* Let  $\mu(dx) = e^{-|x|^2} dx$ . Then  $\mu(\mathbb{R}^n) < \infty$ . Define

$$\nu(A) = \Lambda \left( e^{-|x|^2/p} \mathbb{1}_A \right) \text{ for } A \in \mathcal{B}(\mathbb{R}^n).$$

First we show that  $\nu$  is a finite measure on  $\mathbb{R}^n$ . Clearly  $\nu(\emptyset) = 0$  and  $\nu(A) \in [0, \infty)$  since  $\Lambda$  is positive. Let  $A_k \in \mathcal{B}(\mathbb{R}^n)$  be a sequence of disjoint sets and  $B_m = \bigcup_{k=1}^m A_k$ . Then

$$\begin{aligned} |\nu(B_\infty) - \nu(B_m)| &\leq \|\Lambda\| \left\| e^{-|x|^2/p} (\mathbb{1}_{B_\infty} - \mathbb{1}_{B_m}) \right\|_{L^p} \\ &= \|\Lambda\| \mu(B_\infty \setminus B_m)^{1/p} \rightarrow 0. \end{aligned}$$

So  $\nu$  is countably additive, and thus a measure. Now we claim  $\nu \ll \mu$ . Indeed if  $\mu(A) = 0$ ,  $\nu(A) \leq \|\Lambda\| \mu(A)^{1/p}$ . Thus by the Radon-Nikodym theorem, there is  $\omega \in L^1(\mathbb{R}^n, \mu)$  non-negative such that

$$\nu(A) = \int_A \omega d\mu = \int_A \omega e^{-|x|^2} dx \text{ for all } A \in \mathcal{B}(\mathbb{R}^n).$$

Now let  $f = e^{-|x|^2/p} \tilde{f}$  where  $\tilde{f}$  is simple. Then by linearity of  $\Lambda$ ,

$$\begin{aligned} \Lambda(f) &= \int \tilde{f} d\nu = \int \tilde{f} \omega e^{-|x|^2} dx \\ &= \int f \underbrace{\omega e^{-(1-\frac{1}{p})|x|^2}}_{\tilde{\omega} = \omega e^{-\frac{1}{q}|x|^2}} dx. \end{aligned}$$

Hence  $\Lambda(f) = \int f \tilde{\omega} dx$  for all  $f$  as above. Exercise: functions of the form  $f = e^{-|x|^2/p} \tilde{f}$  for  $\tilde{f}$  are dense in  $L^p(\mathbb{R}^n)$ . Then we have  $\Lambda(f) = \int f \tilde{\omega} dx$  for all  $f \in L^p(\mathbb{R}^n)$  since  $\Lambda$  is continuous.

Example sheet 1 gives that

$$\|\tilde{\omega}\|_{L^q} = \sup \left\{ \int |f \tilde{\omega}| dx : \|f\|_{L^p} \leq 1 \right\}.$$

Thus

$$\|\tilde{\omega}\|_{L^q} \leq \|\Lambda\| \text{ since } \int |f \tilde{\omega}| dx = \int |f| \tilde{\omega} dx = \Lambda(|f|) \leq \|\Lambda\| \|f\|_{L^p}.$$

Conversely,  $\Lambda(f) \leq \|f\|_{L^p} \|\tilde{\omega}\|_{L^q}$  by Hölder's inequality, so  $\|\Lambda\| \leq \|\tilde{\omega}\|_{L^q}$  and  $\|\Lambda\| = \|\tilde{\omega}\|_{L^q}$ .  $\square$

**Theorem.** Let  $p \in [1, \infty)$ . Then  $\int : L^q(\mathbb{R}^n) \rightarrow L^p(\mathbb{R}^n)'$ ,  $g \mapsto \Lambda_g$  where  $\Lambda_g(f) = \int fg$  is a linear isometry and surjective.

*Proof.* First consider the real case. In Example Sheet 2 it is shown that if  $\Lambda \in L^p(\mathbb{R}^n)'$  is real-valued, there are  $\Lambda_+$  and  $\Lambda_-$  both bounded and positive such that  $\Lambda = \Lambda_+ - \Lambda_-$ . The claim follows from the previous lemma.

In the complex case, if  $\Lambda \in L(\mathbb{R}^n, \mathbb{C})'$  then  $\Lambda_r(f) = \Re \Lambda(f)$  and  $\Lambda_i(f) = \Im \Lambda(f)$  define two  $\mathbb{R}$ -linear  $\Lambda \in L^p(\mathbb{R}^n, \mathbb{R})$  such that

$$\Lambda(f_r + if_i) = \Lambda_r(f_r) - \Lambda_i(f_i) + i\Lambda_r(f_i) + i\Lambda_i(f_r).$$

The claim then follows by the real-valued case.  $\square$

## 2.4 Riesz-Markov Theorem

**Fact.** For any finite (positive) regular Borel measure on  $\mathbb{R}^n$ ,  $\Lambda_\mu(f) = \int f d\mu$  defines a positive bounded linear functional on  $C_c(\mathbb{R}^n, \|\cdot\|_\infty)$ .

**Lemma.**  $\Lambda$  uniquely determines  $\mu$  and for any  $U \in \mathbb{R}^n$  open

$$\mu(U) = \sup\{\Lambda_\mu(g) : g \in C_c(\mathbb{R}^n), 0 \leq g \leq \mathbb{1}_U\}. \quad (*)$$

*Proof sketch.* We would like to take  $f = \mathbb{1}_A$  for  $A \in \mathcal{B}(\mathbb{R}^n)$ , but this is not continuous. So we approximate by continuous functions: assume  $U \in \mathbb{R}^n$  is open, set  $U_k = U \cap \{|x| < k\}$ , and

$$\chi_k(x) = \begin{cases} 1 & x \in U_k, d(x, U_k^c) \geq \frac{1}{k} \\ 0 & x \notin U_k \\ kd(x, U_k^c) & x \in U_k, d(x, U_k^c) < 1/k \end{cases}.$$

Then  $\chi_k \in C_c(\mathbb{R}^n)$  and  $\chi_k \uparrow \mathbb{1}_U$ . So by the Monotone Convergence Theorem,

$$\mu(U) = \lim_{k \rightarrow \infty} \int \chi_k d\mu = \lim_{k \rightarrow \infty} \Lambda(\chi_k).$$

And  $(*)$  also follows. Since  $\mu$  is regular, this determines  $\mu$  on all Borel sets.  $\square$

**Definition.** A *signed measure* is the difference of two mutually singular finite positive measures.

**Theorem (Riesz-Markov Theorem).** Given  $\Lambda : C_c(\mathbb{R}^n) \rightarrow \mathbb{R}$  linear positive and bounded, there is a unique finite Borel measure  $\mu$  on  $\mathbb{R}^n$  such that

$$\Lambda(f) = \int_{\mathbb{R}} f d\mu, \quad \forall f \in C_c(\mathbb{R}^n).$$

The dual space  $C_c(\mathbb{R}^n)$  is the space of signed measures.



## 2.5 Strong, weak & weak-\* topologies

Example Sheet 2: if  $X$  is a Banach space, then the closed unit ball is compact iff  $X$  is finite dimensional.

Goal: recover some form of compactness by considering a weaker topology.

**Definition.** A *seminorm*  $p$  on a vector space  $X$  (over  $\mathbb{R}$  or  $\mathbb{C}$ ) is a map  $p : X \rightarrow \mathbb{R}$  such that

- (i)  $p(x + y) \leq p(x) + p(y)$  for all  $x, y \in X$ ;
- (ii)  $p(\lambda x) = |\lambda|p(x)$  for all  $x \in X$ ;
- (iii)  $p(x) \geq 0$  for all  $x \in X$ .

(Note: it is not necessarily positive semidefinite)

**Definition.** A family  $\mathcal{P}$  of seminorms is *separating* if for every  $x \in X$  with  $x \neq 0$  there is  $p \in \mathcal{P}$  such that  $p(x) \neq 0$ .

**Definition.** The topology  $\tau_{\mathcal{P}}$  induced by a family of seminorms  $\mathcal{P}$  is generated by

$$\beta = \{x + B : x \in X, B \in \dot{\beta}\}.$$

Where  $\dot{\beta}$  consists of finite intersections of  $V(p, n) = \{x \in X : p(x) < 1/n\}$  for  $p \in \mathcal{P}, n \in \mathbb{N}$ .  $(X, \tau_{\mathcal{P}})$  is a locally convex topological vector space (LCTVS).

**Theorem.**  $\beta$  is a neighbourhood base for the topology  $\tau_{\mathcal{P}}$  (every open set  $U \in \tau_{\mathcal{P}}$  is a union of sets in  $\beta$ ), and the vector space operations  $(x, y) \mapsto x + y$  and  $(\lambda, x) \mapsto \lambda x$  are continuous, as is every seminorm  $p \in \mathcal{P}$ .

Example Sheet 2: for  $(x_k)_{k \geq 1}$  in  $X$ ,  $x_k \rightarrow x$  in  $\tau_{\mathcal{P}}$  if and only if  $p(x - x_k) \rightarrow 0$  for all  $p \in \mathcal{P}$ .

**Fact.** If  $\mathcal{P} = (p_k)_{k \in \mathbb{N}}$  is countable, then the topology is induced by the metric

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

**Definition.** If  $\mathcal{P}$  is as above with metric as above, if the metric  $d_{\mathcal{P}}$  is complete,  $(X, d_{\mathcal{P}})$  is called a *Fréchet space*.

**Examples.**

- (i)  $X$  a Banach space,  $\mathcal{P}_s = \{|\cdot|\}$ : the corresponding topology  $\tau_s = \tau_{\mathcal{P}_s}$  is called *norm* or *strong topology*. We have  $x_k \rightarrow x$  in  $\tau_s$  if and only if  $\|x_k - x\| \rightarrow 0$ .
- (ii)  $X$  a Banach space,  $\mathcal{P}_w = \{p_\Lambda : \Lambda \in X^1\}$  where  $p_\Lambda(x) = |\Lambda(x)|$ . Each  $p_\Lambda$  is a seminorm and the Hahn-Banach theorem implies  $\mathcal{P}_w$  is separating. (For  $X = L^p(\mathbb{R}^n)$  this can be verified directly.) The topology  $\tau_w = \tau_{\mathcal{P}_w}$  is called the *weak topology*. We have  $x_k \rightarrow x$  in  $\tau_w$  if and only if  $\Lambda(x_k) \rightarrow \Lambda(x)$  for all  $\Lambda \in X'$ . We write  $x_k \rightharpoonup^w x$ . Also  $x_k \rightarrow x$  implies  $x_k \rightharpoonup^w x$ .
- (iii)  $X$  a Banach space, then  $X'$  is also a Banach space. Hence we have a strong and weak topology on  $X'$ . The *weak-\** topology  $\tau_{w^*}$  is generated by  $\mathcal{P}_{w^*} = \{p_x : x \in X\}$  where  $p_x(\Lambda) = |\Lambda(x)|$ . Then  $\Lambda_k \rightarrow \Lambda$  in  $\tau_{w^*}$  if and only if  $\Lambda_k(x) \rightarrow \Lambda(x)$  for every  $x \in X$ . We write  $\Lambda_k \rightharpoonup^{w^*} \Lambda$ .

**Remark.** If  $X$  is reflexive, i.e  $X'' = X$ , then  $\tau_w = \tau_{w^*}$ .

**Example.** Let  $p \in [1, \infty)$  and  $(f_k)_{k \geq 1}$  be a sequence in  $L^p(\mathbb{R}^n)$ . Then

$$\begin{aligned} f_k \rightarrow f \text{ in } L^p &\iff \int |f_k - f|^p dx \rightarrow 0 \\ f_k \rightharpoonup^w f \text{ in } L^p &\iff \int g(f_k - f) dx \rightarrow 0 \text{ for all } g \in L^q \\ f_k \rightharpoonup^{w^*} f \text{ in } L^p &\iff f_k \rightharpoonup^w f \text{ in } L^p \end{aligned}$$

On the other hand, if  $(f_k)_{k \geq 1}$  is in  $L^\infty(\mathbb{R}^n)$ ,

$$\begin{aligned} f_k \rightarrow f \text{ in } L^\infty &\iff \text{esssup} |f_k - f| \rightarrow 0 \\ f_k \xrightarrow{w^*} \text{ in } L^\infty &\iff \int g(f_k - f) dx \rightarrow 0 \text{ for all } g \in L^1 \\ f_k \xrightarrow{w^*} \text{ in } L^\infty &\iff f_k \xrightarrow{w} \text{ in } L^\infty \end{aligned}$$

**2.6 Compactness**

**Theorem** (Arzela-Ascoli Theorem). *Let  $I = [0, 1]$  (or a compact Hausdorff space). Suppose a sequence of continuous functions  $f_k : I \rightarrow \mathbb{R}$  is*

- *Bounded:*  $\sup_k \sup_{x \in I} |f_k(x)| < \infty$
- *Equicontinuous:* for all  $\varepsilon > 0$  there exists  $\delta$  such that  $\sup_k \sup_{x \in I} \sup_{y \in B(x, \varepsilon)} |f_k(x) - f_k(y)| < \varepsilon$ .

*Then there is a subsequence  $(i_k)$  such that  $f_{i_k}$  converges to some continuous  $f$ .*

Application:  $C^{0,\alpha}(I)$  embeds compactly into  $C^0(I)$ , where  $C^{0,\alpha}(I) = \{f \in C^0(I) : \|f\|_{C^{0,\alpha}} < \infty\}$ ,

$$\|f\|_{C^{0,\alpha}} = \sup_{x \in I} |f'(x)| + \sup_{\substack{x,y \in I \\ x \neq y}} \frac{|f(x) - f(y)|}{|x - y|^\alpha}.$$

The identity map  $\text{id} : C^{0,\alpha}(I) \rightarrow C^0(I)$  is compact, i.e any sequence  $(f_i)_{i \geq 1}$  in  $C^{0,\alpha}$  that is bounded in  $C^{0,\alpha}$  has a convergent subsequence in  $C^0(I)$ .

**Theorem** (Banach-Alaoglu). *Let  $X$  be a separable Banach space, and let  $(\Lambda_j)_{j \geq 1}$  be a bounded sequence in  $X'$ , say  $\sup_j \|\Lambda_j\|_{X'} \leq 1$ . Then there is a subsequence  $(j_i)$  and  $\Lambda \in X'$  such that  $\Lambda_{j_i} \rightarrow^{w^*} \Lambda$ .*

**Example.** Let  $p \in (1, \infty]$  and  $(f_j)_{j \geq 1}$  be a sequence in  $L^p(\mathbb{R}^n)$  such that  $\|f_j\|_{L^p} \leq K$  for all  $j$ . Then there is  $f \in L^p$  with  $\|f\|_{L^p} \leq K$  and a subsequence  $(j_i)$  such that for every  $g \in L^q(\mathbb{R}^n)$ ,  $\int f_{j_i} g dx \rightarrow \int f g dx$ . (Just apply Banach-Alaoglu noting  $L^q(\mathbb{R}^n)' = L^p(\mathbb{R}^n)$  for  $p \in [1, \infty)$  and  $L^q$  is separable for such  $q$ .)

*Proof.* Step 1: construction. Let  $D = \{x_k\}_{k=1}^\infty \subseteq X$  be dense (can do this by separability). Since  $(\Lambda_j(x_1))_{j \geq 1}$  is a bounded sequence, there is a subsequence  $J_1 \subseteq \mathbb{N}$  and  $\Lambda(x_1) \in \mathbb{R}$  (or  $\mathbb{C}$ ) such that  $\Lambda_j(x_1) \rightarrow \Lambda(x_1)$  for  $j \in J_1, j \rightarrow \infty$ . Iterating, there are nested subsequences  $J_1 \supseteq J_2 \supseteq J_3 \supseteq \dots$  and  $\Lambda(x_k) \in \mathbb{R}$  (or  $\mathbb{C}$ ) such that  $\Lambda_j(x_k) \rightarrow \Lambda(x_k)$  for  $j \in J_l, l \geq k$ .

Now take the ‘diagonal subsequence’  $J$  of  $J_1 \supseteq J_2 \supseteq \dots$  defined by  $J = (j_n)_{n \geq 1}$  where  $j_n$  is the first element of  $J_n$ . i.e it has first element which is the first element of  $J_1$ , second element which is the first element of  $J_2$ , etc. Then  $\Lambda_j(x_k) \rightarrow \Lambda(x_k)$  for  $j \in J, j \rightarrow \infty$ .

Step 2: we’ll show  $\Lambda : D \rightarrow \mathbb{R}$  is uniformly continuous so can be extended uniquely to  $\Lambda : X \rightarrow \mathbb{R}$  continuous. For each  $x, y \in D$  such that  $\|x - y\| < \varepsilon$ , there is  $j \in J$  such that  $|\Lambda_j(x) - \Lambda(x)| < \varepsilon$ ,  $|\Lambda_j(y) - \Lambda(y)| < \varepsilon$ . Hence

$$|\Lambda(x) - \Lambda(y)| \leq |\Lambda(x) - \Lambda_j(x)| + |\Lambda_j(x) - \Lambda_j(y)| + |\Lambda_j(y) - \Lambda(y)| \leq 3\varepsilon.$$

Step 3: we show  $\Lambda : X \rightarrow \mathbb{R}$  (or  $\mathbb{C}$ ) is linear. For  $x, y \in X$ ,  $a \in \mathbb{R}$  (or  $\mathbb{C}$ ), let  $x', y', z' \in D$  be such that  $\|x - x'\| < \varepsilon$ ,  $\|y - y'\| < \varepsilon$ ,  $\|x + ay - z'\| < \varepsilon$ . Then take  $j \in J$  such that  $|\Lambda(x') - \Lambda_j(x')| < \varepsilon$ ,  $|\Lambda(y') - \Lambda_j(y')| < \varepsilon$ ,  $|\Lambda(z') - \Lambda_j(z')| < \varepsilon$ . Then

$$\begin{aligned} |\Lambda(x + ay) - \Lambda(x) - a\Lambda(y)| &\leq |\Lambda(x + ay) - \Lambda(z')| + |\Lambda(x) - \Lambda(x')| + |a||\Lambda(y) - \Lambda(y')| \\ &\quad + |\Lambda(z') - \Lambda_j(z')| + |\Lambda(x') - \Lambda_j(x')| + |a||\Lambda(y') - \Lambda_j(y')| \\ &\quad + |\Lambda_j(x') - \Lambda_j(x) - a\Lambda_j(y')| \\ &\leq C\varepsilon + \|\Lambda_j\| \|z' - x' - ay'\| \leq C'\varepsilon \end{aligned}$$

so  $\Lambda(x + ay) = \Lambda(x) + a\Lambda(y)$ .

Step 4:  $\|\Lambda\| \leq 1$ . We have

$$\|\Lambda\| = \sup_{\substack{x \in X \\ \|x\| \leq 1}} |\Lambda(x)| = \sup_{\substack{x \in D \\ \|x\| \leq 1}} |\Lambda(x)| \leq 1 \text{ by density.}$$

Step 5:  $\Lambda_j \xrightarrow{w^*} \Lambda$ . For  $x' \in D$  take  $x \in X$  with  $\|x - x'\|' < \varepsilon$ . Then we have

$$|\Lambda_j(x) - \Lambda(x)| \leq |\Lambda_j(x - x')| + |\Lambda_j(x') - \Lambda(x')| + |\Lambda(x - x')| < 3\varepsilon.$$

So  $\Lambda_j(x) \rightarrow \Lambda(x)$  for all  $x \in X$ . □

## 2.7 Hahn-Banach Theorem

Suppose  $\Lambda : M \rightarrow \mathbb{R}$  (or  $\mathbb{C}$ ) is a bounded linear functional on a subspace  $M \subseteq X$  of a Banach space. Goal: extend  $\Lambda$  to  $\tilde{\Lambda} : X \rightarrow \mathbb{R}$  (or  $\mathbb{C}$ ) with  $\|\tilde{\Lambda}\|_{X'} = \|\Lambda\|_{M'}$ .

**Definition.** Let  $X$  be a real vector space. Then  $p : X \rightarrow \mathbb{R}$  is *sublinear* if

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

**Examples.**

- $p(x) = |l(x)|$  for  $l : X \rightarrow \mathbb{R}$  linear.
- Any seminorm.

**Note.** If  $p$  is sublinear,  $l$  is linear,  $l(x) \leq p(x)$  for all  $x \in M$ , then

$$-p(-x) \leq l(x) \leq p(x).$$

**Lemma** (Bounded extension lemma). *Let  $X$  be a real vector space,  $p : X \rightarrow \mathbb{R}$  sublinear,  $M \subseteq X$  a subspace. Assume  $l : M \rightarrow \mathbb{R}$  is linear and  $l(y) \leq p(y)$  for all  $y \in M$ . For  $x \in X \setminus M$ , let  $\tilde{M} = \text{span}\{x, M\}$ . Then there is an extension  $\tilde{l} : \tilde{M} \rightarrow \mathbb{R}$  linear such that  $\tilde{l}(y) = l(y)$  for all  $y \in M$  and  $\tilde{l}(z) \leq p(z)$  for all  $z \in \tilde{M}$ .*

*Proof.* If  $z \in \tilde{M}$ , there are unique  $y \in M$  and  $\lambda \in \mathbb{R}$  such that  $z = y + \lambda x$ . Define  $\tilde{l}(x) = a$  for some  $a$  to be defined, and  $\tilde{l}(y) = l(y)$  for  $y \in M$  and then  $\tilde{l}(z)$  is defined by linearity.

Claim:  $a = \sup\{l(y) - p(y - x) : y \in M\}$  works. For each  $y, z \in M$ ,

$$l(y) + l(z) = l(y + z) \leq p(y + z) \leq p(y - x) + p(x + z).$$

Hence

$$l(y) - p(y - x) \leq p(z + x) - l(z). \quad (*)$$

Note this implies  $a < \infty$ . Also  $(*)$  implies

$$l(y) - a \leq p(y - x) \text{ for all } y \in M. \quad (*')$$

and

$$l(z) + a \leq p(z + x) - (l(y) - p(y - x)) + a \text{ for all } y \in M. \quad (**)$$

So taking the infimum of  $(**)$  over  $y \in M$ :

$$l(z) + a \leq p(z + x) - a + a = p(z + x).$$

Now

$$\tilde{l}(y + \lambda x) = l(y) + a\lambda \leq p(y + \lambda x) \text{ for all } y \in M, \lambda > 0$$

by taking  $z = \lambda^{-1}y$  in  $(**)$  and multiplying across by  $\lambda$ . Also

$$\tilde{l}(y + a\lambda) = l(y) + a\lambda \leq p(y + \lambda x) \text{ for all } y \in M, \lambda > 0$$

by replacing  $y$  with  $|\lambda|^{-1}y$  in  $(*)$  and multiplying across by  $|\lambda|$ . Hence  $\tilde{l}(z) \leq p(z)$  for all  $z \in \tilde{M}$ .  $\square$

**Corollary.** *If  $M$  has finite codimension in  $X$ , then any  $l : M \rightarrow \mathbb{R}$  satisfying  $l(y) \leq p(y)$  for all  $y \in M$  can be extended to  $\tilde{l} : X \rightarrow \mathbb{R}$  linear with  $\tilde{l}(x) \leq p(x)$  for all  $x \in X$ .*

*Proof.* Apply lemma repeatedly.  $\square$

**Definition.** Let  $S$  be a set. A *partial order* is a binary relation  $\leq$  on  $S$  such that

- (i)  $a \leq a$  for all  $a \in S$  (reflexive);
- (ii)  $a \leq b, b \leq c \Rightarrow a \leq c$  (transitive);
- (iii)  $a \leq b, b \leq a \Rightarrow a = b$  (antisymmetry).

A set  $S$  with a partial order is called a *poset*. If additionally  $a \leq b$  or  $b \leq a$  holds for all  $a, b \in S$ , then  $\leq$  is called a *total order*. A totally ordered subset  $T \subseteq S$  of a poset  $S$  is called a *chain*. An element  $u \in S$  is an *upper bound* for  $T \subseteq S$  if  $t \leq u$  for all  $t \in T$ . A *maximal element*  $m \in S$  is an element such that  $m \leq x$  implies  $m = x$ .

**Examples.**

- (i) If  $A$  is any set,  $S = 2^A$  is a poset partially ordered by inclusion of sets.
- (ii)  $\mathbb{R}$  (with the usual ordering) is a totally ordered set with no maximal element.
- (iii) The collection of open balls in  $\mathbb{R}^n$  is a poset ordered by inclusion. The subset  $T = \{B_r(0) : 0 < r \leq 1\}$  is a chain in  $S$ .  $B_1(0)$  is a maximal element of  $T$ .  $B_2(0)$  is an upper bound of  $T$ .

**Lemma (Zorn's Lemma).** *Let  $(S, \leq)$  be a poset in which every totally ordered subset has an upper bound. Then  $(S, \leq)$  contains at least one maximal element.*

We will treat Zorn's Lemma as an axiom.

**Theorem (Hahn-Banach).** *Let  $X$  be a real vector space,  $p : X \rightarrow \mathbb{R}$  sublinear,  $M \subseteq X$  a subspace. For any  $l : M \rightarrow \mathbb{R}$  linear such that  $l(x) \leq p(x)$  for all  $x \in M$ , there exists  $\tilde{l} : X \rightarrow \mathbb{R}$  linear such that  $\tilde{l}|_M = l$  and  $\tilde{l}(y) \leq p(y)$  for all  $y \in X$ .*

*Proof.* Let

$$S = \{(N, \tilde{l}) : X \supseteq N \supseteq M, \tilde{l} : N \rightarrow \mathbb{R} \text{ linear}, \tilde{x} \leq p(x) \forall x \in N, \tilde{x} = p(x) \forall x \in M\}$$

and define the partial order  $(N_1, \tilde{l}_1) \leq (N_2, \tilde{l}_2) \iff N_1 \subseteq N_2, \tilde{l}_2|_{N_1} = \tilde{l}_1$ . For every totally ordered subset  $T \subseteq S$ , we obtain an upper bound for  $T$  via

$$N_T := \bigcup_{(N, \tilde{l}) \in T} N, \quad l_T(x) = \tilde{l}(x) \text{ if } x \in N \text{ for some } (N, \tilde{l}) \in T$$

which is well-defined since where the  $\tilde{l}$  are defined (for  $(N, \tilde{l}) \in T$ ), they agree since  $T$  is a total order. Further,  $(N, \tilde{l}) \leq (N_T, l_T)$  for every  $(N, \tilde{l}) \in T$ . Thus  $(N_T, l_T)$  is an upper bound.

Applying Zorn's Lemma, there is a maximal element  $(\tilde{N}, \tilde{l})$  of  $S$ . It suffices to show  $\tilde{N} = X$ . Suppose not, then there is  $x \in X \setminus \tilde{N}$  and the bounded extension lemma gives an extension  $l^*$  to  $N^* = \text{span}\{x, \tilde{N}\}$  such that  $(\tilde{N}, \tilde{l}) \leq (N^*, l^*)$ , contradicting maximality of  $(\tilde{N}, \tilde{l})$ .  $\square$

**Corollary.** Let  $X$  be a normed vector space over  $\mathbb{K} = \mathbb{R}$  or  $\mathbb{C}$  and  $M \subseteq X$  a subspace. Then every bounded linear functional  $\Lambda : M \rightarrow \mathbb{K}$  can be extended to a bounded linear functional  $\tilde{\Lambda} : X \rightarrow \mathbb{K}$  such that  $\|\tilde{\Lambda}\|_{X'} = \|\Lambda\|_{M'}$  and  $\tilde{\Lambda}|_M = \Lambda$ .

*Proof.* If  $\mathbb{K} = \mathbb{R}$ , then  $p(x) = \|\Lambda\| \cdot \|x\|$  is sublinear and the result follows immediately from Hahn-Banach. If  $\mathbb{K} = \mathbb{C}$ , then  $\Lambda(x) = l(x) - il(ix)$  with  $l : X \rightarrow \mathbb{R}$ ,  $l(x) = \Re(\Lambda(x))$  a real linear function. Since  $|\Lambda(x)| = l(e^{i\theta}x)$  for suitable  $\theta \in [0, 2\pi]$ ,

$$\sup_{\substack{\|x\| \leq 1 \\ x \in \tilde{N}}} |\Lambda(x)| = \sup_{\substack{\|x\| \leq 1 \\ x \in \tilde{N}}} l(x), \quad N \subseteq X.$$

So apply Hahn-Banach to  $l$  and the result follows.  $\square$

**Corollary.** Let  $X$  be a normed vector space and  $x \in X$ . Then there is  $\Lambda_x \in X'$  such that  $\|\Lambda_x\| = 1$  and  $\Lambda_x(x) = \|x\|$ .  $\Lambda_x$  is called a support functional.

*Proof.* Let  $M = \text{span}\{x\}$  and define  $l \in M'$  by  $l(tx) = t\|x\|$ ,  $t \in \mathbb{K}$ . Clearly,  $\|l\| = 1$  and  $l(x) = \|x\|$ . Apply Hahn-Banach to get the result.  $\square$

**Remark.** For  $X = L^p(\mathbb{R}^n)$ ,  $p \in [1, \infty]$ , can construct a support functional by hand (Example Sheet 1).

**Corollary.** Let  $X$  be a normed vector space and  $x \in X$ . Then  $x = 0$  if and only if  $\Lambda(x) = 0$  for all  $\Lambda \in X'$ .

**Corollary.** Let  $X$  be a normed vector space and  $x, y \in X$  be distinct. Then there exists  $\Lambda \in X'$  such that  $\Lambda(x) \neq \Lambda(y)$ : i.e linear functionals separate points.

**Corollary.** The map  $\Phi : X \rightarrow X''$ ,  $\Phi(x) = \tilde{x}$  where  $\tilde{x}(\Lambda) = \Lambda(x)$  is an isometry.

*Proof.* By definition

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

By choosing  $\Lambda = \Lambda_x$  (the support functional), there is equality.  $\square$

**Definition.**  $X$  is said to be *reflexive* if  $\Phi$  is surjective, i.e  $X = X''$ .

**Example.**  $L^p(\mathbb{R}^n)$  is reflexive iff  $p \in [1, \infty)$ .

**Theorem.** Let  $A, B \subseteq X$  be disjoint, nonempty, convex subsets of a normed space  $X$  (real or complex). Then

(a) If  $A$  is open, there exists  $\Lambda \in X'$  such that and  $\gamma \in \mathbb{R}$  such that

$$\Re \Lambda(x) < \gamma \leq \Re \Lambda(y), \forall x \in A, \forall y \in B.$$

If  $B$  is also open the second inequality can be made strict.

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

$$\Re \Lambda(x) < \gamma_1 < \gamma_2 < \Re \Lambda(y), \forall x \in A, \forall y \in B.$$

*Proof.* Assume  $X$  is a vector space over  $\mathbb{R}$  (otherwise just apply to real part)

(a) Fix  $a_0 \in A$ ,  $b_0 \in B$  and set

$$x_0 = b_0 - a_0, \quad C = A - B + x_0 \ni 0.$$

Note  $C$  is convex (since  $A$  and  $B$  are),  $x_0 \notin C$ . (since  $A \cap B = \emptyset$ ). Thus  $C$  is a convex neighbourhood of 0. Let  $p(x) = \inf\{t > 0 : t^{-1}x \in C\}$ . Then  $p$  is sublinear with  $p(x) \leq k\|x\|$  for some  $k$ , and  $p(x) < 1$  if and only if  $x \in C$  (Example Sheet 2). Define  $M = \{tx_0 : t \in \mathbb{R}\}$ , and define  $l : M \rightarrow \mathbb{R}$  by  $l(tx_0) = t$ .

We claim that  $l(x) \leq p(x)$  for all  $x \in M$ . If  $t > 0$ ,  $l(tx_0) = t \leq tp(x_0)$  since  $x_0 \notin C$ . If  $t < 0$ ,  $l(tx_0) = t \leq 0 \leq p(tx_0)$ . By Hahn-Banach,  $l$  can be extended to  $\Lambda : X \rightarrow \mathbb{R}$  with  $\Lambda(x) \leq p(x)$  for all  $x \in X$ . Moreover,  $-k\|x\| \leq -p(-x) \leq \Lambda(x) \leq p(x)$  so  $|\Lambda(x)| \leq k\|x\|$  and  $\Lambda \in X'$ .

We claim that  $\Lambda(a) < \Lambda(b)$  for all  $a \in A$  and all  $b \in B$ . Indeed

$$\underbrace{\Lambda(a - b + x_0)}_{\Lambda(a) - \Lambda(b) + 1} \leq p(a - b + x_0) < 1.$$

Since non-zero elements of the dual are open maps (Example Sheet 2),  $\Lambda(A)$  is an open interval (since  $A$  is open). Take  $\gamma$  to be the right endpoint of  $\Lambda(A)$ . Then  $\Lambda(x) < \gamma \leq \Lambda(x)$ . If  $B$  is also open, the inequality is strict.



(b) Since  $A$  is compact,  $B$  is closed and  $A \cap B$ ,

$$d = \inf\{\|a - b\| : a \in A, b \in B\} > 0.$$

Let  $V = B_{1/2}(0)$ . Then  $A + V$  is open and disjoint from  $B$ . By (a), there is a  $\Lambda \in X'$  such that  $\Lambda(A + V)$  and  $\Lambda(B)$  are disjoint intervals of  $\mathbb{R}$ . These intervals are also a positive distance apart so there exist  $\gamma_1 < \gamma_2$  between them.

□

**Corollary.** *Let  $X$  be a Banach space,  $M \subseteq X$  a subspace and  $x_0 \in X$ . If  $x_0 \notin \overline{M}$  then there is  $\Lambda \in X'$  such that  $\Lambda(x_0) = 1$  and  $\Lambda(x) = 0$  for all  $x \in \overline{M}$ .*

*Proof.* Apply (b) of the previous theorem with  $A = \{x_0\}$ ,  $B = \overline{M}$ . Thus there exists  $\Lambda \in X'$  such that  $\Lambda(x_0) \notin \Lambda(\overline{M})$ . Thus  $\Lambda(\overline{M})$  must be a proper subspace of  $\mathbb{K}$ , so  $\{0\}$ . Also  $\Lambda(x_0) \neq 0$ , so  $\frac{\Lambda}{\Lambda(x_0)}$  is the required element of  $X'$ . □

### 3 Distributions

Distributions are generalised functions.

**Example.**  $G(x) = \frac{1}{4\pi|x|}$ ,  $x \in \mathbb{R}^3$  solves  $-\nabla^2 G = \delta$  as distributions. What this means is that for all sufficiently nice  $f : \mathbb{R}^3 \rightarrow \mathbb{R}$ ,  $\int (-\nabla^2 f) G dx = f(0)$ .

#### 3.1 Distributions, the space $\mathcal{D}(U)$ and $\mathcal{D}'(U)$

For  $U \subseteq \mathbb{R}^n$  open,  $C_c^\infty(U) = \{\phi : U \rightarrow \mathbb{R} \text{ smooth and } \text{supp } \phi \subseteq U \text{ is compact}\}$ .

**Theorem.** *There is a topology on  $C_c^\infty(U)$  such that*

- (i) *The vector space operations are continuous;*
- (ii) *A sequence  $(\phi_j)_{j \geq 1}$  in  $C_c^\infty(U)$  converges to 0 if and only if there is  $K \subseteq U$  compact such that  $\text{supp } \phi_j \subseteq K$  for all  $j$  and for all  $\alpha$ ,  $\sup_K |\nabla^\alpha \phi_j| \rightarrow 0$ ;*
- (iii) *If  $Y$  is a LCTVS (Locally Compact TVS) and  $\Lambda : C_c^\infty(U) \rightarrow Y$  is linear then  $\Lambda$  is sequentially continuous if and only if it is continuous.*

*Proof.* Not given. □

**Definition.**  $C_c^\infty(U)$  with the above topology is called the space of *test functions* and is denoted  $\mathcal{D}(U)$ .

**Examples.** Let  $\phi \in C_c^\infty(\mathbb{R})$ .

- (a) If  $\phi_j(x) = e^{-j}\phi(jx)$ , then  $\phi_j \rightarrow 0$  in  $\mathcal{D}(\mathbb{R})$ ;
- (b) If  $\phi_j(x) = j^{-100}\phi(jx)$ , then  $\phi_j$  does not necessarily converge to 0 in  $\mathcal{D}(\mathbb{R})$ ;
- (c) If  $\phi_j(x) = e^{-j}\phi(x-j)$  then  $\phi_j$  does not necessarily converge to 0 in  $\mathcal{D}(\mathbb{R})$ .

**Definition.** The *space of distributions*  $\mathcal{D}'(U)$  is the dual space of  $\mathcal{D}(U)$  with the weak-\* topology.

In practice,  $u \in \mathcal{D}'(U)$  if and only if  $u(\phi_j) \rightarrow u(\phi)$  whenever  $\phi_j \rightarrow \phi$  in  $\mathcal{D}(U)$ . Also,  $u_j \rightarrow u$  in  $\mathcal{D}'(U)$  if and only if  $u_j(\phi) \rightarrow u(\phi)$  for all  $\phi \in \mathcal{D}(U)$ .

**Examples.**

- (a) For any  $x \in U$ , define  $\delta_x : \mathcal{D}(U) \rightarrow \mathbb{R}$  by  $\delta_x(\phi) = \phi(x)$ . This is called the *Dirac* or  *$\delta$  distribution*.
- (b) If  $f \in L^1_{\text{loc}}(U)$  then  $T_f : \mathcal{D}(U) \rightarrow \mathbb{R}$ ,  $T_f(\phi) = \int_U f\phi dx$  defines a  $T_f \in \mathcal{D}'(U)$ .

**Fact.**  $T_f = T_g \iff \int_U (f-g)\phi dx = 0$  for all  $\phi \in C_c^\infty(U) \iff f = g$  almost everywhere. Hence the map  $T : L^1_{\text{loc}}(U) \rightarrow \mathcal{D}'(U)$ ,  $f \mapsto T_f$  is an injection.

**Example.** If  $\alpha \in C^\infty(U)$ , then  $T_{\alpha f}(\phi) = \int_U f\alpha\phi dx = T_f(\alpha\phi)$  for all  $\phi \in \mathcal{D}(U)$ .

**Definition.** If  $u \in \mathcal{D}'(U)$  is a distribution we define  $\alpha u \in \mathcal{D}'(U)$  by  $\alpha u(\phi) = u(\alpha\phi)$  for all  $\phi \in \mathcal{D}(U)$ .

**Example.** If  $f \in C^1(U)$  then

$$T_{\nabla_i f}(\phi) = \int_U (\nabla_i f) \phi dx = - \int_U f (\nabla_i \phi) dx = -T_f(\nabla_i \phi) \quad \forall \phi \in C_c^\infty.$$

**Definition.** If  $u \in \mathcal{D}'(U)$  define  $\nabla^\alpha u \in \mathcal{D}'(U)$  by

$$\nabla^\alpha u(\phi) = (-1)^{|\alpha|} u(\nabla^\alpha \phi) \quad \forall \phi \in C_c^\infty.$$

**Example.** Define  $H : \mathbb{R} \rightarrow \mathbb{R}$  by  $H(x) = 1$  for all  $x \geq 0$  and  $H(x) = 0$  for all  $x < 0$  (Heaviside function). Then for  $\phi \in \mathcal{D}(\mathbb{R})$ ,  $\nabla T_H(\phi) = - \int H \phi' dx = - \int_0^\infty \phi'(x) dx = \phi(0) = \delta_0(\phi)$ . Hence  $\nabla T_H = \delta_0$  or  $H' = \delta_0$  in the sense of distributions.

### 3.2 Compactly supported distributions: $\mathcal{E}(U)$ and $\mathcal{E}'(U)$

Now consider  $C^\infty(U) = \{\phi : U \rightarrow \mathbb{R} \text{ smooth}\}$ .

Let  $(K_i \subseteq U : i \in \mathbb{N})$  be compact sets such that  $K_i \subseteq \text{int}(K_{i+1})$ ,  $U = \bigcup_i K_i$ . For  $\phi \in C^\infty$  define  $p_N(\phi) = \sup_{x \in K_N} \sup_{|\alpha| \leq N} |\nabla^\alpha \phi(x)|$ . Then  $\mathcal{P} = \{p_N\}_{N \geq 1}$  is a separating family of seminorms.

**Definition.** The space  $C^\infty(U)$  with the locally convex topology induced by  $\mathcal{P}$  is denoted  $\mathcal{E}(U)$ .

**Remark.** Since  $\mathcal{P}$  is countable,  $\mathcal{E}(U)$  is a metric space. It is also complete, i.e a Fréchet space.

A sequence  $(\phi_j)_{j \geq 1}$  in  $\mathcal{E}(U)$  converges to 0 if and only if for all  $K \subseteq U$  compact and all  $\alpha$ ,  $\sup_{x \in K} |\nabla^\alpha \phi_j(x)| \rightarrow 0$ .

**Fact.**  $\mathcal{D}(U) \subseteq \mathcal{E}(U)$  so  $\mathcal{E}'(U) \subseteq \mathcal{D}'(U)$ .

**Example.** If  $\phi \in C_c^\infty$ , then  $\phi_j(x) = e^{-j} \phi(x-j)$  converges to 0 in  $\mathcal{E}(\mathbb{R})$  but not in  $\mathcal{D}(\mathbb{R})$ .

**Definition.**  $u \in \mathcal{D}'(U)$  has *support* in  $S \subseteq U$  if  $u(\phi) = 0$  for all  $\phi \in C_c^\infty(U \setminus S)$ . If  $S$  can be taken compact, say  $u$  has *compact support*.

**Theorem.**  $\mathcal{E}'(U) = \{u \in \mathcal{D}'(U) : u \text{ has compact support}\}$ .

**Lemma.** Let  $u : \mathcal{E}(U) \rightarrow \mathbb{R}$  be linear. Then  $u$  is continuous if and only if

$$\exists \text{ compact } K \subseteq \mathbb{R}^n, N \in \mathbb{N}, C > 0 \text{ such that } |u(\phi)| \leq C \sup_{\substack{x \in K \\ |\alpha| \leq N}} |\nabla^\alpha \phi(x)|. \quad (*)$$

*Proof.* Recall that  $u \in \mathcal{E}'(U)$  if and only if  $u(\phi_j) \rightarrow 0$  for all sequences  $(\phi_j)$  in  $\mathcal{E}(U)$ , i.e.  $\phi_j \xrightarrow{\mathcal{E}(U)} 0$ . Now assume  $(*)$  and let  $(\phi_j)$  be a sequence in  $\mathcal{E}(U)$  with  $\phi_j \xrightarrow{\mathcal{E}(U)} 0$ . This is equivalent to: for all  $\tilde{K} \subseteq U$  compact,  $\tilde{N} \in \mathbb{N}$ ,  $\sup_{\substack{x \in \tilde{K} \\ |\alpha| \leq \tilde{N}}} |\nabla^\alpha \phi_j(x)| \rightarrow 0$ . Thus taking  $\tilde{K} = K$  and  $\tilde{N} = N$ ,  $(*)$  implies  $u(\phi_j) \rightarrow 0$ .

Now suppose  $(*)$  does not hold. Let  $K_i \subseteq U$  be compact,  $K_j \subseteq \text{int}(K_{j+1})$ ,  $\bigcup_j K_j = U$ . Since  $(*)$  does not hold, for each  $j$  we have  $\phi_j \in \mathcal{E}(U)$  such that  $|u(\phi_j)| > j \sup_{x \in K_j} \sup_{|\alpha| \leq j} |\nabla^\alpha \phi_j(x)|$ . Set  $\psi_j = \frac{\phi_j}{|u(\phi_j)|}$ . We claim that  $\psi_j \rightarrow 0$  in  $\mathcal{E}(U)$ . For any  $\tilde{K} \subseteq U$  compact,  $\tilde{N} \in \mathbb{N}$ , there exists  $J > \tilde{N}$  such that  $\tilde{K} \subseteq K_j$  for all  $j > J$ , so

$$\sup_{\substack{x \in \tilde{K} \\ |\alpha| \leq \tilde{N}}} |\nabla^\alpha \psi_j(x)| \leq \sup_{\substack{x \in \tilde{K} \\ |\alpha| \leq \tilde{N}}} \frac{|\nabla^\alpha \phi_j(x)|}{|u(\phi_j)|} < \frac{1}{j}.$$

As claimed. But  $|u(\psi_j)| = 1$ , so  $|u(\psi_j)| \not\rightarrow 0$ , so  $u$  is not continuous.  $\square$

*Proof of Theorem.* If  $u \in \mathcal{E}'(U)$ , the lemma implies that  $u$  has support in  $K$ . Conversely, if  $u \in \mathcal{D}'(U)$  has support in  $K \subseteq U$  compact, define  $\tilde{u} \in \mathcal{E}'(U)$  by  $\tilde{u}(\phi) = u(\chi\phi)$  for all  $\phi \in \mathcal{E}(U)$ , where  $\chi \in C_c^\infty(U)$  satisfies  $\chi = 1$  on  $K$ . The extension does not depend on  $\chi$  since for any other such  $\tilde{\chi}$  one has  $\chi - \tilde{\chi} \in C_c^\infty(U \setminus K)$ .  $\square$

### Examples.

- (a) If  $f \in L^1(U)$  vanishes almost everywhere in  $U \setminus K$  for  $K$  compact, then  $T_f \in \mathcal{E}'(U)$ ;
- (b) For any  $x \in U$ ,  $\delta_x \in \mathcal{E}'(U)$ ;
- (c)  $u \in \mathcal{D}'(U)$  where  $u(\phi) = \sum_{m=-\infty}^{\infty} \phi(m) \notin \mathcal{E}(\mathbb{R})$ .

### Tempered distributions: the spaces $S(\mathbb{R}^n)$ and $S'(\mathbb{R}^n)$

**Definition.**  $\phi \in C^\infty(\mathbb{R}^n)$  is *rapidly decreasing* if

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

for all multi-indices  $\alpha$  and  $N \in \mathbb{N}$ .

### Examples.

- (a)  $\phi(x) = e^{-|x|^a}$  is rapidly decreasing;  
 (b)  $\phi(x) = |x|^{-2023}$  is not rapidly decreasing.

**Definition.** The *Schwartz space*  $S(\mathbb{R}^n)$  is the space of rapidly decreasing functions with the topology generated by the separating family of seminorms

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

**Remark.** There are other equivalent families of seminorms such as

$$\begin{aligned} & \sup_{x \in \mathbb{R}^n} \sup_{|\alpha| \leq N} |(1 + |x|^2)^N \nabla^\alpha \phi(x)| \\ & \sup_{x \in \mathbb{R}^n} \sup_{|\alpha| \leq N} |\nabla^\alpha (1 + |x|^2)^N \phi(x)|. \end{aligned}$$

**Fact.**  $S(\mathbb{R}^n)$  is a Fréchet space,  $\mathcal{D}(\mathbb{R}^n) \subseteq S(\mathbb{R}^n) \subseteq \mathcal{E}(\mathbb{R}^n)$  continuously and  $\mathcal{E}'(\mathbb{R}^n) \subseteq S'(\mathbb{R}^n) \subseteq \mathcal{D}'(\mathbb{R}^n)$ .

**Definition.**  $S'(\mathbb{R}^n)$  is called the space of *tempered distributions* or *Schwartz distributions*.

**Examples.**

- (a) If  $f \in L^1_{\text{loc}}(\mathbb{R}^n)$  satisfies  $\int_{\mathbb{R}^n} (1 + |x|)^{-N} |f(x)| dx < \infty$  for some  $N \in \mathbb{N}$ , then  $T_f \in S'(\mathbb{R}^n)$ . Indeed, if  $\phi \in S(\mathbb{R}^n)$ , then

$$\begin{aligned} |T_f(\phi)| &= \left| \int f(x) \phi(x) dx \right| \\ &= \underbrace{\left( \int (1 + |x|)^{-N} |f(x)| dx \right)}_{\leq C} \underbrace{\sup_{x \in \mathbb{R}^n} (1 + |x|)^N |\phi(x)|}_{\rightarrow 0 \text{ if } \phi \xrightarrow{S(\mathbb{R}^n)} 0} \end{aligned}$$

so if  $\phi_j \xrightarrow{S(\mathbb{R}^n)} 0$  then  $T_f(\phi_j) \rightarrow 0$ .

- (b) If  $f(x) = e^{|x|^2}$  then  $T_f \in \mathcal{D}'(\mathbb{R}^n)$  but  $T_f \notin S'(\mathbb{R}^n)$ .  
 (c)  $u(\phi) = \sum_{m=-\infty}^{\infty} |m|^{2023} \phi(m)$  belongs to  $S'(\mathbb{R})$  but not  $\mathcal{E}'(\mathbb{R})$ .

## Convolution

**Example.** Let  $f \in L^1_{\text{loc}}(\mathbb{R}^n)$ ,  $\phi \in \mathcal{D}(\mathbb{R}^n)$ . Then

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

where  $\tau_x \check{\phi}(y) = \check{\phi}(y - x) = \phi(x - y)$ ,  $\check{\phi}(y) = \phi(-y)$ .

**Definition.** For  $u \in \mathcal{D}'(\mathbb{R}^n)$  and  $\phi \in \mathcal{D}(\mathbb{R}^n)$  define

$$u * \phi(x) = u(\tau_x \check{\phi}).$$

**Facts.**

- $(u_1 + au_2) * \phi = u_1 * \phi + au_2 * \phi$ ;
- $u * (\phi_1 + a\phi_2) = u * \phi_1 + au * \phi_2$ ;
- $u * \check{\phi}(0) = u(\phi)$  - thus  $u * \phi(0)$ ,  $\phi \in \mathcal{D}(\mathbb{R}^n)$  determines  $u \in \mathcal{D}'(\mathbb{R}^n)$ .

**Example.**  $\delta_0 * \phi(x) = \delta_0(\tau_x \check{\phi}) = \check{\phi}(-x) = \phi(x)$ . Thus  $\delta_0 * \phi = \phi$ .

**Proposition.** Let  $u \in \mathcal{D}'(\mathbb{R}^n)$  and  $\phi \in \mathcal{D}(\mathbb{R}^n)$ . Then

- (i)  $u * \phi \in C^\infty(\mathbb{R}^n)$  and  $\nabla^\alpha(u * \phi) = (\nabla^\alpha u) * \phi = u * \nabla^\alpha \phi$ ;
- (ii) If  $u \in \mathcal{E}'(\mathbb{R}^n)$  then  $u * \phi$  has compact support, i.e  $u * \phi \in \mathcal{D}(\mathbb{R}^n)$ .

*Proof.*

(i)

$$\frac{1}{h}(u * \phi(x + he_i) - u * \phi(x)) = u \left( \frac{1}{h}(\tau_{x+e_i h} \check{\phi} - \tau_x \check{\phi}) \right) \xrightarrow{h \rightarrow 0} u(\tau_x \widetilde{\nabla_i \phi}).$$

Where we used from Example Sheet 3:

$$\frac{1}{h}(\tau_{x+e_i h} \check{\phi} - \tau_x \check{\phi}) \xrightarrow{\mathcal{D}(\mathbb{R}^n)} \nabla_i \phi(x - \cdot) = \tau_x \widetilde{\nabla_i \phi}.$$

Hence  $\nabla_i(u * \phi)(x)$  exists and equals  $u(\tau_x \widetilde{\nabla_i \phi}) = u * \nabla_i \phi(x)$ . So by induction  $u * \phi \in C^\infty$  and  $\nabla^\alpha u * \phi = u * \nabla^\alpha \phi$  for all  $\alpha$ . Also;  $\nabla^\alpha \tau_x \check{\phi}(y) = \nabla_y^\alpha \phi(x - y) = (-1)^{|\alpha|} \nabla_x^\alpha \phi(x - y) = (-1)^{|\alpha|} \tau_x \widetilde{\nabla^\alpha \phi}(y)$ . Thus  $u * \nabla^\alpha \phi = \nabla^\alpha u * \phi$ .

- (ii) Assume  $u(\phi) = 0$  for all  $\phi \in C_c^\infty(\mathbb{R}^n \setminus K)$  for some  $K$  compact. Then for any  $\phi \in \mathcal{D}(\mathbb{R}^n)$ ,  $\text{supp } \tau_x \check{\phi} \cap K = \emptyset$  for  $|x|$  large enough, i.e  $u * \phi$  has compact support.

□

**Definition.** For  $u_1 \in \mathcal{D}'(\mathbb{R}^n)$  and  $u_2 \in \mathcal{E}'(\mathbb{R}^n)$ , define  $u_1 * u_2$  to be the unique distribution such that

$$(u_1 * u_2) * \phi = u_1 * (u_2 * \phi).$$

[Note that  $u_2 * \phi \in \mathcal{D}(\mathbb{R}^n)$  by the previous proposition so this makes sense.]

**Example.** Let  $u \in \mathcal{D}'(\mathbb{R}^n)$ . Then  $u * \delta_0 = u$ . Indeed,  $(u * \delta_0) * \phi = u * (\delta_0 * \phi) = u * \phi$ .

**Proposition.** Let  $u \in \mathcal{D}'(\mathbb{R}^n)$  and  $u_2 \in \mathcal{E}'(\mathbb{R}^n)$ . Then  $\nabla^\alpha(u_1 * u_2) = u_1 * (\nabla^\alpha u_2) = (\nabla^\alpha u_1) * u_2$ .

*Proof.* Let  $\phi \in \mathcal{D}(\mathbb{R}^n)$ . Then by the previous proposition

$$\begin{aligned} \underbrace{\nabla^\alpha(u_1 * u_2)}_{\in \mathcal{D}'} * \underbrace{\phi}_{\in \mathcal{D}} &= (u_1 * u_2) * (\nabla^\alpha \phi) \\ &= u_1 * (u_2 * (\nabla^\alpha \phi)) = (u_1 * \nabla^\alpha u_2) * \phi. \end{aligned}$$

□

**Definition.** Call  $L = \sum_{|\alpha| \leq k} a_\alpha \nabla^\alpha$ ,  $a_\alpha \in \mathbb{R}$ ,  $\nabla^\alpha u_2 * \phi$  a *constant coefficient partial differential operator* of order  $k$ . A *fundamental solution* of  $L$  is a distribution  $G$  such that  $LG = \delta_0$ .

**Theorem.** If  $G \in \mathcal{D}'(\mathbb{R}^n)$  is a fundamental solution of  $L$  and  $f \in \mathcal{E}'(\mathbb{R}^n)$  then  $u = G * f$  solves  $Lu = f$ . Moreover, if  $f \in \mathcal{D}(\mathbb{R}^n)$  then  $u = G * f \in C^\infty(\mathbb{R}^n)$  solves  $Lu = f$  in the classical sense.

*Proof.*

$$L(G * f) = \sum_{|\alpha| \leq k} a_\alpha \nabla^\alpha (G * f) = \left( \sum_{|\alpha| \leq k} a_\alpha \nabla^\alpha G \right) * f = \delta_0 * f = f.$$

□

**Example.**  $L = -\nabla^2 = -\sum_{i=1}^3 \frac{\partial^2}{\partial x_i^2}$  on  $\mathbb{R}^3$ . Define  $g(x) = \frac{1}{4\pi|x|} \in L^1_{\text{loc}}(\mathbb{R}^3)$ . Then  $G = T_g$  is a fundamental solution for  $L$ . In particular, if  $f \in C_c^\infty(\mathbb{R}^n)$  then

$$u(x) = \int_{\mathbb{R}^3} \frac{f(y)}{4\pi|x-y|} dy$$

solves  $Lu = f$ .

## Fourier Transform

**Definition.** If  $f \in L^1(\mathbb{R}^n)$  then the *Fourier transform* of  $f$  is  $\hat{f} = \mathcal{F}(f) : \mathbb{R}^n \rightarrow \mathbb{C}$ ,  $\hat{f}(\xi) = \int_{\mathbb{R}^n} f(x) e^{-i x \cdot \xi} dx$ .

**Example.** ( $n = 1$ )

(i)

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

The Fourier transform of  $f$  is  $\hat{f}(\xi) = 2 \frac{\sin \xi}{\xi}$ ;

(ii)  $f(x) = e^{-|x|}$  has Fourier transform  $\hat{f}(\xi) = \frac{2}{1+\xi^2}$ ;

(iii)  $f(x) = \frac{1}{1+x^2}$  has Fourier transform  $\hat{f}(\xi) = e^{-|\xi|} \pi$ ;

(iv)  $f(x) = \frac{1}{\sqrt{2\pi}} e^{-x^2/2}$  has Fourier transform  $\hat{f}(\xi) = e^{-|\xi|^2/2}$ .

Upshot:  $f$  regular  $\leftrightarrow \hat{f}$  decays.

**Theorem** (Riemann-Lebesgue Lemma). *Let  $f \in L^1(\mathbb{R}^n)$ . Then  $\hat{f} \in C^0(\mathbb{R}^n)$  and  $\sup_{\xi \in \mathbb{R}^n} |\hat{f}(\xi)| \leq \|f\|_{L^1}$ ,  $\hat{f}(\xi) \rightarrow 0$  as  $|\xi| \rightarrow \infty$ .*

*Proof.* Assume  $\xi_k \rightarrow \xi$ . Then for  $x \in \mathbb{R}^n$ ,  $f(x) e^{-i \xi_k \cdot x} \rightarrow f(x) e^{-i \xi \cdot x}$  pointwise and  $|f(x) e^{-i \xi_k \cdot x}| \in |f(x)| \in L^1$  so by the DCT,  $\hat{f}(\xi_k) \rightarrow \hat{f}(\xi)$ . Hence  $\hat{f} \in C^0$ . We also have

$$|\hat{f}(\xi)| = \left| \int f(x) e^{-i \xi \cdot x} dx \right| \leq \|f\|_{L^1}.$$

To show  $\hat{f}(\xi) \rightarrow 0$  as  $|\xi| \rightarrow \infty$ , let  $f_\varepsilon \in C_c^\infty(\mathbb{R}^n)$  be such that  $\|f - f_\varepsilon\|_{L^1} < \varepsilon$ . Then

$$\hat{f}_\varepsilon(\xi) = \int_{\mathbb{R}^n} f_\varepsilon(x) \underbrace{e^{-i \xi \cdot x}}_{-\frac{1}{|\xi|^2} \nabla_x^2 (e^{-i \xi \cdot x})} dx = -\frac{1}{|\xi|^2} \underbrace{\int_{\mathbb{R}^n} (\nabla^2 f_\varepsilon) e^{-i \xi \cdot x} dx}_{\leq \|\nabla^2 f_\varepsilon\|_{L^1}}.$$

Hence  $\limsup_{|\xi| \rightarrow \infty} |\hat{f}_\varepsilon(\xi)| = 0$ . Finally note

$$|\hat{f}(\xi)| \leq |\hat{f}_\varepsilon(\xi)| + \underbrace{|\hat{f}(\xi) - \hat{f}_\varepsilon(\xi)|}_{\leq \|f - f_\varepsilon\|_{L^1}} \leq |\hat{f}_\varepsilon(\xi)| + \varepsilon.$$

So  $\hat{f}(\xi) \rightarrow 0$  as  $|\xi| \rightarrow \infty$ . □

Notation:  $\tau_y f(x) = f(x - y)$  and  $e_y(x) = e^{i x \cdot y}$ .

**Proposition.**



- (i) Let  $f \in L^1(\mathbb{R}^n)$ ,  $x \in \mathbb{R}^n$ ,  $\lambda > 0$ , and set  $f_\lambda(x) = \lambda^{-n}f(x/\lambda)$ . Then  $\widehat{f_\lambda}(\xi) = \widehat{f}(\lambda\xi)$ ,  $\widehat{e_y f}(\xi) = \tau_y \widehat{f}(\xi)$ ,  $\widehat{\tau_y f}(\xi) = e_{-y}(\xi) \widehat{f}(\xi)$ ;
- (ii) Let  $f, g \in L^1(\mathbb{R}^n)$ . Then  $f * g \in L^1(\mathbb{R}^n)$  and  $\widehat{f * g}(\xi) = \widehat{f}(\xi) \widehat{g}(\xi)$ .

*Proof.* Change of variables and Fubini.  $\square$

**Proposition.**

- (i) If  $f \in C^1(\mathbb{R}^n)$  and  $f, \nabla_i f \in L^1(\mathbb{R}^n)$  for all  $1 \leq i \leq n$ , then

$$\widehat{\nabla_j f}(\xi) = i\xi_j \widehat{f}(\xi).$$

- (ii) Assume  $(1 + |x|)f \in L^1(\mathbb{R}^n)$ . Then  $\widehat{f} \in C^1(\mathbb{R}^n)$  and

$$\nabla_j \widehat{f}(\xi) = -i\widehat{x_j f}(\xi).$$

*Proof.*

- (i) Let  $f_\varepsilon \in C_c^\infty(\mathbb{R}^n)$  be such that  $\|f_\varepsilon - f\|_{L^1} + \sum_j \|\nabla_j f - \nabla_j f_\varepsilon\|_{L^1} < \varepsilon$  (Exercise: show we can do this). Then (IBP)

$$\widehat{\nabla_j f_\varepsilon}(\xi) = \int_{\mathbb{R}^n} e^{-ix \cdot \xi} \nabla_j f_\varepsilon(x) dx = i\xi_j \int_{\mathbb{R}^n} e^{-ix \cdot \xi} f_\varepsilon(x) dx = i\xi_j \widehat{f_\varepsilon}(\xi).$$

- (ii) Hence

$$|\widehat{\nabla_j f}(\xi) - i\xi_j \widehat{f}(\xi)| \leq \|\nabla_j f - \nabla_j f_\varepsilon\|_{L^1} + |\xi| \|f - f_\varepsilon\|_{L^1} \leq (1 + |\xi|)\varepsilon \xrightarrow{\varepsilon \rightarrow 0} 0.$$

- (iii) Since  $x_j f \in L^1$ ,  $-i\widehat{x_j f} \in C^0$ . Need to show  $\nabla_j \widehat{f}$  exists and equals  $-i\widehat{x_j f}$ .

$$\frac{\widehat{f}(\xi + he_j) - \widehat{f}(\xi)}{h} = \int_{\mathbb{R}^n} f(x) e^{-i\xi \cdot x} \underbrace{\left( \frac{e^{-ih \cdot x_j} - 1}{h} \right)}_{\substack{\rightarrow -ix_j \\ \leq |x_j|}} dx \xrightarrow{h \rightarrow 0} -i\widehat{x_j f}(\xi)$$

by the DCT, using  $|x_j|f \in L^1$ .  $\square$

**Corollary.** The Fourier Transform maps  $\mathcal{S}(\mathbb{R}^n)$  to  $\mathcal{S}(\mathbb{R}^n)$  continuously.

*Proof.* For any  $f : \mathbb{R}^n \rightarrow \mathbb{C}$ ,

$$\|f\|_{L^1} \leq \sup_{x \in \mathbb{R}^n} (1 + |x|)^{n+1} |f(x)| \underbrace{\int_{\mathbb{R}^n} \frac{dy}{(1 + |y|)^{n+1}}}_{< \infty}. \quad (*)$$

Hence if  $f \in \mathcal{S}(\mathbb{R}^n)$ , then  $\nabla^\alpha(x^\beta f(x)) \in L^1(\mathbb{R}^n)$  for any multi-indices  $\alpha, \beta$ . Thus by the previous proposition (applied repeatedly),

$$|\widehat{\nabla^\alpha(x^\beta f)}(\xi)| = |\xi^\alpha \nabla^\beta \hat{f}(\xi)|.$$

So in particular (using (\*)),

$$\sup_{\xi} |\xi^\alpha \nabla^\beta \hat{f}(\xi)| \leq C \sup_{\substack{x \in \mathbb{R}^n \\ |\gamma| \leq \alpha}} \left[ (1 + |x|)^{|\beta|+n+1} |\nabla^\gamma f(x)| \right] \rightarrow 0 \text{ if } f \rightarrow 0 \text{ in } \mathcal{S}(\mathbb{R}^n).$$

Therefore  $\hat{f} \rightarrow 0$  in  $\mathcal{S}(\mathbb{R}^n)$  if  $f \rightarrow 0$  in  $\mathcal{S}(\mathbb{R}^n)$ . Hence  $\mathcal{F} : \mathcal{S}(\mathbb{R}^n) \rightarrow \mathcal{S}(\mathbb{R}^n)$  is well-defined and continuous.  $\square$

**Theorem** (Fourier inversion). *Let  $f \in L^1(\mathbb{R}^n)$  and assume also  $\hat{f} \in L^1(\mathbb{R}^n)$ . Then*

$$f(x) = \frac{1}{(2\pi)^n} \int_{\mathbb{R}^n} \hat{f}(\xi) e^{i\xi \cdot x} d\xi \text{ for almost all } x.$$

*Thus writing  $\check{f}(x) = f(-x)$  we have  $\mathcal{F}^2(f) = (2\pi)^n \check{f}$ .*