# Analysis of Functions *

Zhiyuan Bai

Compiled on March 15, 2022

This document serves as a set of revision materials for the Cambridge Mathematical Tripos Part II course *Analysis of Functions* in Lent 2022. However, despite its primary focus, readers should note that it is NOT a verbatim recall of the lectures, since the author might have made further amendments in the content. Therefore, there should always be provisions for errors and typos while this material is being used.

## Contents



*Based on the lectures under the same name taught by Dr. C. Warnick in Lent 2022.

# 1 Lebesgue Integral

In analysis of functions, we will use the tools from linear analysis and measure theory to deduce results about concrete spaces of functions. We start by reminding ourselves how Lebesgue integrals work.

## 1.1 Review of Measure Theory

**Definition 1.1.** Given a set $E$, a $\sigma$-algebra on $E$ is a collection $\mathcal{E}$ of subsets of $E$ such that:
1. $\varnothing \in \mathcal{E}$.
2. $A \in \mathcal{E} \implies A^c = E \setminus A \in \mathcal{E}$.
3. $(A_n)_{n\in\mathbb{N}} \in \mathcal{E} \implies \bigcup_{n\in\mathbb{N}} A_n \in \mathcal{E}$.
$(E, \mathcal{E})$ is called a measurable space. Elements of $\mathcal{E}$ are called measurable sets.

**Definition 1.2.** A measure on a measurable space $(E, \mathcal{E})$ is a function $\mu : \mathcal{E} \to [0, \infty]$ such that:
1. $\mu(\varnothing) = 0$.
2. If $(A_n)_{n\in\mathbb{N}} \in \mathcal{E}$ are disjoint, then

$$\mu\left(\bigcup_{n\in\mathbb{N}} A_n\right) = \sum_{n\in\mathbb{N}} \mu(A_n)$$

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

**Example 1.1.** For any set $E$, we can take $\mathcal{E} = 2^E$ and $\mu(A) = \text{card}(A)$.

**Definition 1.3.** Given any collection $\mathcal{A}$ of subsets of $E$, we define the $\sigma$-algebra generated by $\mathcal{A}$ by

$$\sigma(\mathcal{A}) = \bigcap_{\mathcal{E}\ \sigma\text{-algebra}, \mathcal{E}\supset\mathcal{A}} \mathcal{E}$$

When $E$ carries a topology $\mathcal{T}$, $\mathcal{B}(E) = \sigma(\mathcal{T})$ is known as the Borel $\sigma$-algebra on the topological space $(E, \mathcal{T})$.

In measure theory, you would have proved the existence of the following example:

**Example 1.2.** On $E = \mathbb{R}^n$, there is a $\sigma$-algebra $\mathcal{M}$ and a measure $\lambda$ (the Lebesgue measure) such that:
1. $\mathcal{B}(\mathbb{R}^n) \subset \mathcal{M}$.
2. If $A = (a_1, b_1] \times \cdots \times (a_n, b_n]$ (where $-\infty < a_i \le b_i < \infty$ for all $i$) is a rectangle, then $\lambda(A) = (b_1 - a_1) \cdots (b_n - a_n)$.
3. ("Borel regularity") $A \in \mathcal{M}$ iff for any $\epsilon > 0$, there is an open set $O$ and a closed set $C$ such that $C \subset A \subset O$ and $\mu(O \setminus C) < \epsilon$.
It follows immediately from these conditions that such $\mathcal{M}, \lambda$ are unique. We sometimes write $\mathrm{d}x = \mathrm{d}\lambda$ and $|A| = \lambda(A)$.

**Definition 1.4.** For measurable spaces $(E, \mathcal{E}), (G, \mathcal{G})$, a function $f : E \to G$ is called measurable if $f^{-1}(A) \in \mathcal{E}$ for all $A \in \mathcal{G}$.

When $(G, \mathcal{G}) = (\mathbb{R}, \mathcal{B}(\mathbb{R}))$, then we simply say that $f$ is measurable on $(E, \mathcal{E})$. When $(G, \mathcal{G}) = ([0, \infty], \mathcal{B}([0, \infty]))$, we say $f$ is a nonnegative measurable function on $(E, \mathcal{E})$. When both $E, G$ are topological spaces with Borel $\sigma$-algebras,

then we say $f$ is Borel measurable.
As seen in measure theory, the class of measurable functions on $E$ is closed under vector space operations and product and a.e. limits.

**Definition 1.5.** A simple function on $(E, \mathcal{E})$ is a function $f : E \to \mathbb{R}$ that has the form $f = \sum_{k=1}^{N} a_k 1_{A_k}$ where $a_k$ are scalars (which might reside in $\mathbb{R}, [0, \infty], \mathbb{C}$, etc.) and $A_k \in \mathcal{E}$.
For a non-negative simple function $f$ (which can always be written in a form with $a_k \in [0, \infty]$), we define its integral to be

$$\mu(f) = \int_E f \, \mathrm{d}\mu = \sum_{k=1}^{N} a_k \mu(A_k)$$

with the convention that $0 \cdot \infty = \infty \cdot 0 = 0$.
For a non-negative measurable $f$, we define

$$\mu(f) = \int_E f \, \mathrm{d}\mu = \sup \{\mu(g) : 0 \leq g \leq f, g \text{ simple.}\}$$

A measurable function $f$ on $E$ is integrable if $\mu(|f|) < \infty$, in which case we write $\mu(f^+), \mu(f^-) < \infty$ where $f = f^+ - f^-$ with $f^\pm$ nonnegative and we can define

$$\mu(f) = \int_E f \, \mathrm{d}\mu = \mu(f) = \mu(f^+) - \mu(f^-)$$

The integral satisfies all the expected properties (linearity, additivity, monotonicity, etc.) and more.

**Theorem 1.1** (Monotone Convergence Theorem)**.** *Suppose $0 \leq f_n \uparrow f$ a.e., then*

$$\int_E f_n \, \mathrm{d}\mu \uparrow \int_E f \, \mathrm{d}\mu$$

**Theorem 1.2** (Dominated Convergence Theorem)**.** *Suppose $f_n \to f$ a.e. and $|f_n| \leq g$ for some integrable $g$, then*

$$\int_E f_n \, \mathrm{d}\mu \to \int_E f \, \mathrm{d}\mu$$

## 1.2 $L^p$ Spaces

With the notion of an integral, we can always associate any measure space $(E, \mathcal{E}, \mu)$ with a family of Banach spaces.

**Definition 1.6.** For $1 \leq p < \infty$ and $f : E \to \mathbb{C}$ measurable, we define the $L^p$ norm of $f$ as

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

and we define

$$\|f\|_{L^\infty} = \operatorname*{ess\,sup}_E |f| = \inf\{K : |f| \leq K \text{ a.e.}\}$$

For $1 \leq p \leq \infty$, the $L^p$ space over $E$ is the set

$$L^p(E, \mu) = \{f : E \to \mathbb{C} \text{ measurable} : \|f\|_{L^p} < \infty\}/\text{a.e.}$$

Again we know from measure theory that

**Theorem 1.3** (Riesz-Fischer). *$(L^p(E,\mu), \|\cdot\|_{L^p})$ is a Banach space for all $1 \le p \le \infty$.*

When $(E, \mathcal{E}, \mu) = (\mathbb{R}^n, \mathcal{M}, \lambda)$, we simply just write $L^p(\mathbb{R}^n, \lambda) = L^p(\mathbb{R}^n)$ with the Lebesgue measure understood.
As always in analysis, life is easier if we can find a nice dense subset of a space we are interested in (e.g. $\mathbb{Q} \subset \mathbb{R}$). This is a process called "mollification".
It's pretty straightforward from definition that, the set of simple functions $s : E \to \mathbb{C}$ with $\mu(\{x : s(x) \neq 0\}) < \infty$ is dense in $L^p(E, \mu)$ for $1 \le p < \infty$. For $p = \infty$, the set of simple functions is still dense, but we don't get to impose the finite support condition.
As simple as they are, they are not usually that pleasant to work with. Sometimes, we can get a much nicer looking dense subset.

**Theorem 1.4.** *The set $C_c^\infty(\mathbb{R})$ of smooth functions on $\mathbb{R}^n$ with compact support is dense in $L^p(\mathbb{R}^n)$ for $1 \le p < \infty$.*

This is sadly not true for $p = \infty$, since no sequence of continuous functions $\mathbb{R} \to \mathbb{R}$ can converge to a discontinuous function under the $L^\infty$ norm.

**Definition 1.7.** For $f, g : \mathbb{R}^n \to \mathbb{C}$, we define their convolution $f * g : \mathbb{R}^n \to \mathbb{C}$ via

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

provided that the integral exists.

The integral exists when e.g. $f \in L^1(\mathbb{R}^n), g \in L^\infty(\mathbb{R}^n)$.

**Lemma 1.5.** *For $f, g, h \in C_c^\infty(\mathbb{R}^n)$, we have $f * g = g * f$ and $(f * g) * h = f * (g * h)$. Moreover,*

$$\int_{\mathbb{R}^n} f * g = \left(\int_{\mathbb{R}^n} f\right)\left(\int_{\mathbb{R}^n} g\right)$$

The condition $f, g, h \in C_c^\infty(\mathbb{R}^n)$ can of course be significantly weakened, but the current strength of the lemma is sufficient for our purpose.

*Proof.* Change of variables and what not. □

A more useful result concerns the differentiability of convolutions.

**Definition 1.8.** A ($n$-)multi-index is a member of $\mathbb{Z}^n_{\ge 0}$. For $\alpha = (\alpha_1, \dots, \alpha_n) \in \mathbb{Z}^n_{\ge 0}$, we define $|\alpha| = \alpha_1 + \cdots + \alpha_n, \alpha! = \alpha_1! \cdots \alpha_n!$.
For $x = (x_1, \dots, x_n) \in \mathbb{R}^n$, we write $x^\alpha = x_1^{\alpha_1} \cdots x_n^{\alpha_n}$.
For $f \in C^k(\mathbb{R}^n)$ for $k \ge |\alpha|$, we write

$$\mathrm{D}^\alpha f = \frac{\partial^{|\alpha|} f}{\partial x_1^{\alpha_1} \cdots \partial x_n^{\alpha_n}}$$

**Definition 1.9.** We say $f \in L^p_{\text{loc}}(\mathbb{R}^n)$ if $f 1_K \in L^p(\mathbb{R}^n)$ for all compact $K \subset \mathbb{R}^n$.

**Theorem 1.6.** *Suppose $f \in L^1_{\text{loc}}(\mathbb{R}^n), g \in C_c^k(\mathbb{R}^n)$ for some $k \ge 0$, then $f * g \in C^k(\mathbb{R}^n)$ and $\mathrm{D}^\alpha(f * g) = f * \mathrm{D}^\alpha g$ for all multi-index $\alpha$ with $|\alpha| \le k$.*

*Proof.* First suppose $k = 0$. For $\xi \in \mathbb{R}^n$, $f : \mathbb{R}^n \to \mathbb{C}$, we write $\tau_\xi f(x) = f(x-\xi)$. So $\tau_\xi(f * g) = f * \tau_\xi(g)$ by definition. Furthermore,$\tau_\xi g \to g$ pointwise as $|\xi| \to 0$ and $|\tau_\xi g| \leq (\sup_{\mathbb{R}^n} |g|) 1_{B_R(0)}$ for $|\xi| < 1$ and $R$ large enough so that the support of $g$ is contained in $B_R(0)$. We then have $\tau_\xi(f * g) \to f * g$ as $|\xi| \to 0$ by dominated convergence, which means that $f * g$ is continuous.
Next suppose $k = 1$. Let $\Delta_i^h g(x) = (g(x+he_i) - g(x))/h$, then $\Delta_i^h g(x) \to \mathrm{D}_i g(x)$ as $|h| \to 0$. Also, by mean value theorem, there is some $t \in (-\|h\|, \|h\|)$ such that $\Delta_i^h g(x) = \mathrm{D}_i g(x + te_i)$, so we have the bound $|\Delta_i^h g| \leq (\sup_{\mathbb{R}^n} |\mathrm{D}_i g|) 1_{B_R(0)}$ again for sufficiently large $R$. So $\Delta_i^h (f * g) \to f * \mathrm{D}_i g$ as $|h| \to 0$ by dominated convergence, so $f * g$ has continuous first partials, hence $f * g \in C^1(\mathbb{R}^n)$.
An induction argument for $k > 1$ finishes the proof. □

To get to our desired density result, we just need some lemmas.

**Lemma 1.7** (Minkowski's Integral Inequality). *If $F : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{C}$ is integrable and $1 \leq p < \infty$, then*

$$\left( \int_{\mathbb{R}^n} \left| \int_{\mathbb{R}^n} F(x,y) \, \mathrm{d}x \right|^p \mathrm{d}y \right)^{1/p} \leq \int_{\mathbb{R}^n} \left( \int_{\mathbb{R}^n} |F(x,y)|^p \, \mathrm{d}y \right)^{1/p} \mathrm{d}x$$

*Proof.* Example sheet. □

**Lemma 1.8** ("Translation is continuous"). *Let $1 \leq p < \infty$ and $g \in L^p(\mathbb{R}^n)$, then $\|\tau_z g - g\|_{L^p} \to 0$ as $z \to 0$.*

*Proof.* The result is clear if $g = 1_R$ where $R = (a_1, b_1] \times \cdots \times (a_n, b_n]$ is a rectangle, which means that it is also true for $g = 1_A$ for any finite disjoint union $A$ of rectangles. If $B$ is (Lebesgue) measurable with $|B| < \infty$, then for all $\epsilon > 0$ there is a finite disjoint union $A_\epsilon$ of rectangles such that $\|1_{A_\epsilon} - 1_B\|_{L^p} = |A_\epsilon \Delta B| < \epsilon$ by Borel regularity (or Dynkin's Lemma). So

$$\begin{aligned} \|\tau_z 1_B - 1_B\|_{L^p} &\leq \|\tau_z 1_B - \tau_z 1_{A_\epsilon}\|_{L^p} + \|\tau_z 1_{A_\epsilon} - 1_{A_\epsilon}\| + \|1_{A_\epsilon} - 1_B\| \\ &= \|\tau_z 1_{A_\epsilon} - 1_{A_\epsilon}\| + 2\|1_{A_\epsilon} - 1_B\| < 3\epsilon \end{aligned}$$

for sufficiently small $z$. Hence the result is true for $g = 1_B$ for any Lebesgue measurable $B$ with finite measure. Thus it also holds for any simple function with finite support. The density of simple functions with finite support in $L^p(\mathbb{R}^n)$ then gives the result. □

Sadly, this is false for $p = \infty$: Just take $g = 1_{\mathbb{R}_{\geq 0}}$.

**Theorem 1.9.** *Suppose $\phi \in C_c^\infty(\mathbb{R}^n)$ is nonnegative and*

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

*Let $\phi_\epsilon(y) = \epsilon^{-n} \phi(y/\epsilon)$. Then for any $g \in L^p(\mathbb{R}^n)$ with $1 \leq p < \infty$, $\phi_\epsilon * g$ is always smooth and $\phi_\epsilon * g \to g$ in $L^p(\mathbb{R}^n)$ as $\epsilon \to 0$.*

*Proof.* Smoothness follows from Theorem 1.6 and the fact that

$$\int_{\mathbb{R}^n} |f| 1_K \, \mathrm{d}x \leq \|f\|_{L^p} \|1_K\|_{L^q}$$

where $p^{-1} + q^{-1} = 1$.
For the convergence, let's do some estimations. Pointwise, we have

$$\begin{aligned}
|\phi_\epsilon * g(x) - g(x)| &= \left| \int_{\mathbb{R}^n} \phi_\epsilon(y) g(x-y)\, \mathrm{d}y - g(x) \right| \\
&= \left| \int_{\mathbb{R}^n} \phi(z) g(x - \epsilon z)\, \mathrm{d}z - \int_{\mathbb{R}^n} \phi(z) g(x)\, \mathrm{d}z \right| \\
&= \left| \int_{\mathbb{R}^n} \phi(z)(g(x - \epsilon z) - g(x))\, \mathrm{d}z \right| \\
&\le \int_{\mathbb{R}^n} \phi(z) |g(x - \epsilon z) - g(x)|\, \mathrm{d}z
\end{aligned}$$

So by Lemma 1.7,

$$\begin{aligned}
\|\phi_\epsilon * g - g\|_{L^p} &= \left( \int_{\mathbb{R}^n} |\phi_\epsilon * g(x) - g(x)|^p\, \mathrm{d}x \right)^{1/p} \\
&\le \left( \int_{\mathbb{R}^n} \left| \int_{\mathbb{R}^n} \phi(z) |\tau_{\epsilon z} g(x) - g(x)|\, \mathrm{d}z \right|^p \mathrm{d}x \right)^{1/p} \\
&\le \int_{\mathbb{R}^n} \left( \int_{\mathbb{R}^n} \phi(z)^p |\tau_{\epsilon z} g(x) - g(x)|^p\, \mathrm{d}x \right)^{1/p} \mathrm{d}z \\
&= \int_{\mathbb{R}^n} \phi(z) \|\tau_{\epsilon z} g - g\|_{L^p}\, \mathrm{d}z \to 0
\end{aligned}$$

as $\epsilon \to 0$ by dominated convergence. □

Theorem 1.4 follows.

## 1.3 Lebesgue Differentiation Theorem

**Definition 1.10.** Given an integrable $f : \mathbb{R}^n \to \mathbb{C}$, the Hardy-Littlewood maximal function $M_f : \mathbb{R}^n \to [0, \infty]$ is defined to be

$$M_f(x) = \sup_{r>0} \frac{1}{|B_r(x)|} \int_{B_r(x)} |f(y)|\, \mathrm{d}y$$

**Lemma 1.10.** *Suppose $f \in L^1(\mathbb{R}^n)$, then $M_f$ is measurable and a.e. finite. Moreover, there is a constant $C_n > 0$ such that $|\{M_f > \lambda\}| \le C_n \|f\|_{L^1}/\lambda$ for any $\lambda > 0$.*

*Proof.* Write $A_\lambda = \{M_f > \lambda\}$. For each $x \in A_\lambda$, there is some $r_x > 0$ such that

$$\frac{1}{|B_{r_x}(x)|} \int_{B_{r_x}(x)} |f(y)|\, \mathrm{d}y > \lambda$$

We shall first show that $A_\lambda$ is open, which shows the measurability of $M_f$. Indeed, suppose for the sake of contradiction that $(x_m)_m \in A_\lambda^{\mathrm{c}}$ converges to some $x \in A_\lambda$. By domianted convergence,

$$\begin{aligned}
\lambda \ge \frac{1}{|B_{r_x}(x_n)|} \int_{B_{r_x}(x_n)} |f(y)|\, \mathrm{d}y &= \frac{1}{|B_{r_x}(x)|} \int_{\mathbb{R}^n} 1_{B_{r_x}(x_n)} |f(y)|\, \mathrm{d}y \\
\to \frac{1}{|B_{r_x}(x)|} \int_{\mathbb{R}^n} 1_{B_{r_x}(x)} |f(y)|\, \mathrm{d}y &= \frac{1}{|B_{r_x}(x)|} \int_{B_{r_x}(x)} |f(y)|\, \mathrm{d}y > \lambda
\end{aligned}$$

Contradiction. So $A_\lambda^c$ is closed, hence $A_\lambda$ is open.
Suppose $K \subset A_\lambda$ is compact. Choose a finite subcover $\{B_i\}_{i=1}^N$ of the open cover $\{B_{r_x}(x) : x \in A_n\}$ of $K$. By Wiener's covering lemma (proved in example sheet), there is a disjoint subcollection $\{B_{i_j}\}_{j=1}^k$ such that

$$\begin{aligned}|K| &\leq \left|\bigcup_{i=1}^N B_i\right| \leq 3^n \sum_{j=1}^k |B_{i_j}| \leq \frac{3^n}{\lambda} \sum_{j=1}^k \int_{B_{i_j}} |f(y)| \,\mathrm{d}y \\ &\leq \frac{3^n}{\lambda} \int_{\mathbb{R}^n} |f(y)| \,\mathrm{d}y = \frac{3^n}{\lambda} \|f\|_{L^1}\end{aligned}$$

As this is true for all compact $K \subset A_\lambda$, the inner regularity of Lebesgue measure gives $|A_\lambda| \leq 3^n \|f\|_{L^1}/\lambda$. The fact that $\{M_f = \infty\} \subset \{M_f > \lambda\}$ for all $\lambda$ then shows that $M_f$ is a.e. finite. □

**Theorem 1.11** (Lebesgue Differentiation Theorem). *Suppose $f : \mathbb{R}^n \to \mathbb{C}$ is integrable, then*

$$\lim_{r\to 0} \frac{1}{|B_r(x)|} \int_{B_r(x)} |f(y) - f(x)| \,\mathrm{d}y \to 0$$

*for almost every $x \in \mathbb{R}^n$ (the "Lebesgue points of $f$").*

The case where $f$ is continuous is straightforward.

*Proof.* For $\lambda > 0$, consider

$$A_\lambda = \left\{ x \in \mathbb{R}^n : \limsup_{r\to\infty} \frac{1}{|B_r(x)|} \int_{B_r(x)} |f(y) - f(x)| \,\mathrm{d}y > 2\lambda \right\}$$

We shall show that $|A_\lambda| = 0$ for all $\lambda$.
For $\epsilon > 0$, we choose $g \in C_c^\infty(\mathbb{R}^n)$ such that $\|f - g\|_{L^1} < \epsilon$. We have

$$\begin{aligned}\frac{1}{|B_r(x)|} \int_{B_r(x)} |f(y) - f(x)| \,\mathrm{d}y &\leq \frac{1}{|B_r(x)|} \int_{B_r(x)} |f(y) - g(y)| \,\mathrm{d}y \\ &\quad + \frac{1}{|B_r(x)|} \int_{B_r(x)} |g(y) - g(x)| \,\mathrm{d}y \\ &\quad + |g(x) - f(x)|\end{aligned}$$

Note that the first term can be bounded by

$$\frac{1}{|B_r(x)|} \int_{B_r(x)} |f(y) - g(y)| \,\mathrm{d}y \leq \sup_{r>0} \frac{1}{|B_r(x)|} \int_{B_r(x)} |f(y) - g(y)| \,\mathrm{d}y = M_{f-g}(x)$$

Also, the second term goes to 0 as $r \to 0$ by continuity of $g$, so

$$\limsup_{r\to 0} \frac{1}{|B_r(x)|} \int_{B_r(x)} |f(y) - f(x)| \,\mathrm{d}y \leq M_{f-g}(x) + |g(x) - f(x)|$$

So if $x \in A_\lambda$, then either $M_{f-g}(x) > \lambda$ or $|f(x) - g(x)| > \lambda$. By the preceding lemma, $|\{M_{f-g}(x) > \lambda\}| \leq C_n \|f - g\|_{L^1}/\lambda \leq C_n\epsilon/\lambda$. And by Markov's inequality, $|\{|f(x) - g(x)| > \lambda\}| \leq \|f - g\|_{L^1}/\lambda \leq \epsilon/\lambda$. Consequently $|A_\lambda| \leq (1 + C_n)\epsilon/\lambda$, which means that $|A_\lambda| = 0$ as $\epsilon$ is arbitrary. □

**Corollary 1.12.** *$\phi_\epsilon * f \to f$ a.e. as $\epsilon \to 0$.*

**Corollary 1.13.** *If $g \in L^2(\mathbb{R})$ and*

$$G(x) = \int_{-\infty}^{x} g(t)\,\mathrm{d}t$$

*Then $G$ is differentiable at almost every $x$ and $G'(x) = g(x)$ there.*

**Example 1.3.** For $g = 1_{\mathbb{R}_{>0}}$, every $x \neq 0$ is a Lebesgue point.

$$G(x) = \int_{-\infty}^{x} g(x)\,\mathrm{d}x = \begin{cases} 0 & \text{for } x \leq 0 \\ t & \text{for } x > 0 \end{cases}$$

is differentiable at every $x \neq 0$ and has $G'(x) = g(x)$ for $x \neq 0$.

## 1.4 Littlewood's Principle

"Every set is nearly a finite sum of intervals; Every function is nearly continuous; Every convergent sequence is nearly uniformly convergent." (Littlewood, 1944)

**Proposition 1.14.** *Suppose $A \subset \mathbb{R}^n$ is measurable and $|A| < \infty$, then for any $\epsilon > 0$ there is a finite union $B$ of rectangles such that $|A \Delta B| < \epsilon$.*

*Proof.* Borel regularity. □

**Theorem 1.15** (Egorov)**.** *Let $(f_n)_n, f$ be measurable functions $E \to \mathbb{C}$ with $E \subset \mathbb{R}^d, |E| < \infty$. Suppose $f_n \to f$ a.e. in $E$. Then for all $\epsilon > 0$, we can find a closed subset $A_\epsilon \subset E$ such that $|E \setminus A_\epsilon| < \epsilon$ and $f_n \to f$ uniformly on $A_\epsilon$.*

*Proof.* WLOG $f_n \to f$ everywhere. Let $E_k^n = \{x \in E : \forall j > k, |f_j(x) - f(x)| < 1/n\}$. Then $E_k^n \subset E_{k'}^n$ for all $k' \geq k$ and $\bigcup_k E_k^n = E$ for all $n$, thus $|E_k^n| \uparrow |E|$ as $k \to \infty$. Pick $k = k_n$ such that $|E \setminus E_{k_n}^n| < 2^{-n}$.
By construction, $|f_j(x) - f(x)| < 1/n$ for all $n, j > k_n, x \in E_{k_n}^n$. Pick $N$ such that $\sum_{n=N}^{\infty} 2^{-n} < \epsilon/2$ and set $A'_\epsilon = \bigcap_{n=N}^{\infty} E_{k_n}^n$. Then

$$|E \setminus A'_\epsilon| = \left| \bigcup_{n=N}^{\infty} (E \setminus E_{k_n}^n) \right| \leq \sum_{n=N}^{\infty} |E \setminus E_{k_n}^n| \leq \sum_{n=N}^{\infty} 2^{-n} < \frac{\epsilon}{2}$$

For any $\delta > 0$, we can take $n > \max\{N, \delta^{-1}\}$, then $x \in A'_\epsilon \implies x \in E_{k_n}^n \implies |f_j(x) - f(x)| < n^{-1} < \delta$. So $f_j \to f$ uniformly on $A'_\epsilon$. Borel regularity then produces the desired closed $A_\epsilon$. □

**Theorem 1.16** (Lusin)**.** *Suppose $f$ is measurable and a.e. finite on $E \subset \mathbb{R}^d, |E| < \infty$. Then for any $\epsilon > 0$, there is a closed $F_\epsilon \subset E$ with $|E \setminus F_\epsilon| < \epsilon$ and $f|_{F_\epsilon}$ is continuous.*

*Proof.* Suppose first that $f = \sum_{n=1}^{m} a_n 1_{A_n}$ is a simple function with $(A_n)_n$ disjoint. WLOG $(A_n)_n$ is a partition of $E$. For $\epsilon > 0$, we pick (with Borel regularity) compact $K_n \subset A_n$ with $|A_n \setminus K_n| < \epsilon/m$. $B = \bigcup_{n=1}^{m} K_n$ would have $|E \setminus B| < \epsilon$. As $(K_n)_n$ are compact and disjoint, there is a positive distance between every pair of them, hence $f$ is continuous on $B$ since it's constant on each $K_n$.

For general $f$, let $f_n$ be a sequence of simple functions converging to $f$ a.e. and choose (by above discussion) $C_n$ such that $|C_n| < 2^{-n}$ and $f_n$ is continuous on $E \setminus C_n$. For $\epsilon > 0$, the preceding theorem shows that we can find some $A \subset E$ with $f_n \to f$ uniformly on $A$ and $|E \setminus A| < \epsilon/3$. Let $N$ be such that $\sum_{n=N}^{\infty} 2^{-n} < \epsilon/3$, then we can set $F'_\epsilon = A \setminus \bigcup_{n=N}^{\infty} C_n$ which has $|E \setminus F'_\epsilon| < 2\epsilon/3$. Also, $f|_{F'_\epsilon}$ is continuous by uniform convergence. The proof is again finished by Borel regularity. $\square$

A word of warning: To say $f|_F$ is continuous is only to say that $f|_F$ is continuous under the subspace topology of $F$, which is very not the same as to say $f$ is continuous at each point of $F$ as a function on $E$. If one take $f = 1_{\mathbb{Q}}$, then $f$ is nowhere continuous on $\mathbb{R}$ but $f|_{\mathbb{R}\setminus\mathbb{Q}} \equiv 0$ is continuous.

# 2 Banach and Hilbert Spaces

We now move on to consider the consequences of the linear structures on $L^p(\mathbb{R}^n)$ as a result of being complete. Things are especially nice for $p = 2$ as we can further give it the structure of a Hilbert space with the inner product

$$\langle f, g \rangle_{L^2} = \int_E \bar{f} g$$

## 2.1 Orthogonal Systems of Functions

**Definition 2.1.** A subset $S = \{u_j\}_{j \in J}$ of a Hilbert space $H$ is orthogonal if $\langle u_i, u_j \rangle = 0$ whenever $i \neq j$, orthonormal if in addition that $\langle u_i, u_i \rangle = 1$ for all $i$.
We say $S$ is complete if $\overline{\operatorname{Span} S} = H$. $S$ is called an orthonormal basis if it is orthonormal and complete.
$H$ is separable if it has a countable orthonormal basis.

If $\{e_n\}_{n \in \mathbb{N}}$ is a countable orthonormal basis, then any $x \in H$ would have the form $x = \sum_{n \in \mathbb{N}} \langle e_n, x \rangle e_n$.

**Example 2.1.** 1. $S = \{x \mapsto e^{-2\pi i n x} : n \in \mathbb{Z}\}$ is an orthonormal set in $L^1((0,1))$. It is in fact an orthonormal basis, which can be proved either with Stone-Weierstrass or Theorem 1.4.
2. Let

$$\psi(x) = \begin{cases} 1 & \text{for } 0 \leq x < 1/2 \\ -1 & \text{for } 1/2 \leq x < 1 \\ 0 & \text{otherwise.} \end{cases}$$

Define $\psi_{n,k}(x) = 2^{n/2}\psi(2^n x - k)$, then $\{\psi_{n,k} : n, k \in \mathbb{Z}\}$ is an orthonormal basis for $L^2(\mathbb{R})$. This is known as a Haar system.
3. Consider $H = L^2(\mathbb{R}, e^{-x^2}\,\mathrm{d}x)$, the Hilbert space with inner product

$$\langle f, g \rangle = \int_{\mathbb{R}} f(x) g(x) e^{-x^2}\,\mathrm{d}x$$

Applying Gram-Schmidt orthogonalisation to the linearly independent collection $1, x, x^2, x^3, \ldots \in H$ gives an orthogonal sequence of polynomials $(H_k(x))_{k=0}^{\infty}$

with $\deg H_k = k$, normalised in such a way that the coefficient of $x^k$ in $H_k$ is $2^k$. They are known as the Hermite polynomials, which is a complete orthogonal set in $H$.

Let's quote something from linear analysis since y'all are told it's a prerequisite.

**Theorem 2.1** (Riesz Representation Theorem)**.** *Suppose $H$ is a Hilbert space and $\Lambda : H \to \mathbb{C}$ is a bounded linear operator, then there is a unique $w \in H$ such that $\Lambda u = \langle w, u\rangle$ for all $u \in H$.*

You are also told that measure theory is a prerequisite, so let's use this to prove something in measure theory.

**Definition 2.2.** Suppose $(E, \mathcal{E})$ is a measurable space and $\mu, \nu$ are measures on it. We say $\nu$ is absolutely continuous with respect to $\mu$, written $\nu << \mu$, if for any $A \in \mathcal{E}$ we have $\mu(A) = 0 \implies \nu(A) = 0$. $\mu, \nu$ are said to be mutually singular, written $\mu \perp \nu$, if there exists a measurable $Z$ such that $0 = \mu(Z) = \nu(Z^c)$.

**Theorem 2.2** (Radon-Nikodým)**.** *Suppose $(E, \mathcal{E})$ is a measurable space and $\mu, \nu$ are finite measures on it with $\nu << \mu$, then there is some $w \in L^1(E, \mu)$, nonnegative $\mu, \nu$-a.e., such that*

$$\nu(A) = \int_A w \, \mathrm{d}\mu$$

*for all $A \in \mathcal{E}$. In other words,*

$$\int_E F \, \mathrm{d}\nu = \int_E F w \, \mathrm{d}\mu$$

*for all measurable $F : E \to [0, \infty]$.*

*Proof.* Let $\alpha = \mu + 2\nu, \beta = \nu + 2\mu$ (don't ask) which are finite measures. Let $H = L^2(E, \alpha) = L^2(E, \mu) \cap L^2(E, \nu) = L^2(E, \beta)$. Consider the linear operator $\Lambda : H \to \mathbb{R}$ via

$$\Lambda(f) = \int_E f \, \mathrm{d}\beta$$

This is bounded since

$$\begin{aligned} |\Lambda(f)| &\le \int_E |f| \, \mathrm{d}\beta = 2\int_E |f| \, \mathrm{d}\mu + \int_E |f| \, \mathrm{d}\nu \le 2\int_E |f| \, \mathrm{d}\mu + 4\int_E |f| \, \mathrm{d}\nu \\ &= 2\int_E |f| \, \mathrm{d}\alpha \le 2\sqrt{\alpha(E)} \|f\|_{L^2(E,\alpha)} \end{aligned}$$

by Cauchy-Schwartz. Theorem 2.1 then produces $g \in H$ such that any $f \in H$ has

$$\int_E f \, \mathrm{d}\beta = \Lambda(f) = \int_E g f \, \mathrm{d}\alpha$$

In other words,

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

Set $f = 1_{A_j}$ with $A_j = \{x \in E : g(x) < 1/2 - 1/j\}$, then

$$\frac{3}{2}\mu(A_j) \le \int_E (2g-1)f \,\mathrm{d}\nu = \int_E (2-g)f \,\mathrm{d}\mu \le -\frac{1}{j}\nu(A_j)$$

So $\mu(A_j) = \nu(A_j) = 0$, thus $g \ge 1/2$ $\mu, \nu$-a.e.. Using the same argument on $A_j = \{x \in E : g(x) > 2 + 1/j\}$ gives $g \le 2$ $\mu, \nu$-a.e..
Take $Z = \{g(x) = 1/2\}$, then setting $f = 1_Z$ gives $\mu(Z) = 0$, so $\nu(Z) = 0$ as $\nu << \mu$. For a measurable $F : E \to [0, \infty]$, set

$$f(x) = \frac{F(x)}{2g(x)-1}, w(x) = \frac{2-g(x)}{2g(x)-1}$$

on $Z^c$ and $f = w = 0$ on $Z$. Then

$$\begin{aligned}\int_E F \,\mathrm{d}\nu &= \int_{E\setminus Z} F \,\mathrm{d}\nu = \int_E (2g-1)f \,\mathrm{d}\nu = \int_E (2-g)f \,\mathrm{d}\mu \\ &= \int_{E\setminus Z} Fw \,\mathrm{d}\mu = \int_E Fw \,\mathrm{d}\mu\end{aligned}$$

As desired. □

## 2.2 Dual Spaces

**Definition 2.3.** Given a topological vector space $X$ over $\mathbb{F} = \mathbb{R}$ or $\mathbb{C}$, the topological dual (or simply dual) $X'$ of $X$ is the vector space of continuous linear forms $\Lambda : X \to \mathbb{F}$.
If $X$ is normed with $\|\cdot\|_X$, we norm $X'$ with

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

$X'$ is always a Banach space. Turns out, if in addition that $X$ is a Banach space, then $X'$ separates points:

**Lemma 2.3.** *Let $X$ be a Banach space and suppose $x, y \in X$ are distinct, then there is some $\Lambda \in X'$ with $\Lambda(x) \ne \Lambda(y)$.*

We defer the proof after we prove Hahn-Banach.
In the case when $X$ is Banach, for each $x \in X$ there is a natural map $f_x : X' \to \mathbb{C}$ via $\Lambda \mapsto \Lambda x$ which is bounded and linear, hence an element of $X'' = (X')'$. If $f_x(\Lambda) = f_y(\Lambda)$ for all $\Lambda$, then the preceding lemma tells us that $x = y$, so $x \mapsto f_x$ is an injection $X \to X''$. If this map were also surjective, we say $X$ is reflexive and we write $X = X''$. Theorem 2.1 tells us that for any Hilbert space $H$ we can canonically identify an isomorphism $H \cong H'$, which in particular shows that any Hilbert space is reflexive.
How about the dual of $L^p(\mathbb{R}^n)$. Suppose $f \in L^p(\mathbb{R})$ for some $1 \le p \le \infty$ and let $p^{-1} + q^{-1} = 1$, then Hölder's inequality gives

$$\left|\int_{\mathbb{R}^n} gf\right| \le \|g\|_{L^q}\|f\|_{L^p}$$

for any $g \in L^q(\mathbb{R}^n)$. Thus the map

$$\Lambda_g(f) = \int_{\mathbb{R}^n} gf$$

is bounded and linear, and a result from example sheet shows $\|\Lambda_g\|_{L^p(\mathbb{R}^n)'} = \|g\|_{L^q(\mathbb{R}^n)}$. Thus the map $\kappa : L^q(\mathbb{R}^n) \to L^p(\mathbb{R}^n)', g \mapsto \Lambda_g$ is linear and isometric (hence injective). Therefore $L^q(\mathbb{R}^n)$ can be identified with a subspace of $L^p(\mathbb{R}^n)'$. For $p = q = 2$, $L^2(\mathbb{R}^n)$ is a Hilbert space and we in fact have $L^2(\mathbb{R}^n) = L^2(\mathbb{R}^n)'$, but what if $p \neq 2$?

**Theorem 2.4.** *Let $1 \leq p < \infty$ and $p^{-1} + q^{-1} = 1$, then we have $L^q(\mathbb{R}^n) = L^p(\mathbb{R}^n)'$ via $\kappa$.*

Sadly, the theorem is not true for $p = \infty$.

*Proof.* We first reduce to the case where the base field is $\mathbb{R}$. Assuming the case for $\mathbb{R}$, then if the base field is $\mathbb{C}$, we let $L^p(\mathbb{R}^n;\mathbb{R})$ be the subspace of $L^p(\mathbb{R}^n)$ consisting of functions which are real a.e.. If $f \in L^p(\mathbb{R})$, then there are unique $f_r, f_i \in L^p(\mathbb{R}^n;\mathbb{R})$ such that $f = f_r + if_i$. Given a bounded linear $\Lambda : L^p(\mathbb{R}^n) \to \mathbb{C}$, we define two real-linear maps $\Lambda_r, \Lambda_i : L^p(\mathbb{R}^n;\mathbb{R}) \to \mathbb{R}$ via $\Lambda_r(f) = \mathrm{Re}(\Lambda f), \Lambda_i(f) = \mathrm{Im}(\Lambda f)$ which determine $\Lambda$ uniquely via $\Lambda(f_r + if_i) = \Lambda_r(f_r) - \Lambda_i(f_i) + i(\Lambda_r(f_r) + \Lambda_i(f_i))$.
We say a real linear map $u : L^p(\mathbb{R}^n;\mathbb{R}) \to \mathbb{R}$ is called positive if $u(f) \geq 0$ whenever $f \geq 0$. Any real-linear map $u$ can be written as $u = u_+ - u_-$ with $u_\pm$ positive (example sheet).
The reduction and the rest of the theorem both follow from the following claim: Let $1 \leq p < \infty$ and $p^{-1} + q^{-1} = 1$. Suppose $u : L^p(\mathbb{R}^n;\mathbb{R}) \to \mathbb{R}$ is a positive linear bounded map, then there exists a nonnegative $g \in L^q(\mathbb{R}^n;\mathbb{R})$ with $\|g\|_{L^q(\mathbb{R}^n)} = \|u\|_{L^p(\mathbb{R}^n)'}$ with

$$u(f) = \int_{\mathbb{R}^n} gf$$

To prove this, we consider the Gaussian measure $\mathrm{d}\mu = e^{-|x|^2}\,\mathrm{d}x$, then $\mu(\mathbb{R}^n) < \infty$. For a Lebesgue-measurable $A$, we consider $\nu(A) = u(e^{-|x|^2/p}1_A)$ which is a finite measure on $\mathbb{R}^n$. Indeed, it is nonnegative by the positivity of $u$ and clearly $\nu(\varnothing) = 0, \nu(\mathbb{R}^n) < \infty$ and that $\nu$ is finitely additive. For countable additivity, if $B = \bigcup_{n=1}^\infty A_n$ for disjoint measurable $A_n$, we set $B_k = \bigcup_{n=1}^k A_n$ which has

$$\|e^{-|x|^2/p}1_B - e^{-|x|^2/p}1_{B_k}\|_{L^p} = \left|\int_{B\setminus B_k} e^{-|x|^2}\,\mathrm{d}x\right|^{1/p} = \mu(B \setminus B_k)^{1/p} \to 0$$

as $k \to \infty$, thus $\nu(B \setminus B_k) = u(e^{-|x|^2/p}1_B - e^{-|x|^2/p}1_{B_k}) \to 0$ which (together with finite additivity) gives countable additivity.
We also have $\nu << \mu$ since $\|e^{-|x|^2/p}1_A\|_{L^p} = \mu(A)^{1/p}$. Theorem 2.2 produces a nonnegative $G \in L^1(\mathbb{R}^n, \mu)$ such that

$$\nu(A) = \int_A G\,\mathrm{d}\mu = \int_A G(x)e^{-|x|^2}\,\mathrm{d}x$$

Since $u$ is linear, if $f$ is of the form $e^{-|x|^2/p}F$ for some simple $F$, then

$$u(f) = \int_{\mathbb{R}^n} fg, g(x) = e^{-|x|^2/q}G(x)$$

Such functions are dense in $L^p(\mathbb{R}^n;\mathbb{R})$ and we also have the estimate $|u(f)| \leq \|u\|_{L^p(\mathbb{R}^n)'}\|f\|_{L^p(\mathbb{R}^n)}$. Thus

$$\sup\left\{\left|\int_{\mathbb{R}^n} fg\right| : f \in L^p(\mathbb{R};\mathbb{R}), \|f\|_{L^p(\mathbb{R}^n)} \leq 1\right\} \leq \|u\|_{L^p(\mathbb{R}^n)'}$$

$\|g\|_{L^q(\mathbb{R}^n)} \leq \|u\|_{L^p(\mathbb{R}^n)'}$ by example sheet, and we get $\|g\|_{L^q(\mathbb{R}^n)} \geq \|u\|_{L^p(\mathbb{R}^n)'}$ by Hölder. □

Hence $L^p(\mathbb{R}^n)$ is reflexive for $1 < p < \infty$, although neither $L^1(\mathbb{R}^n)$ nor $L^\infty(\mathbb{R}^n)$ is.
Another space whose dual can be conveniently characterised is $C_c^0(\mathbb{R}^n)$, a result known as Riesz representation theorem for spaces of continuous functions (Riesz-Markov theorem). We shall show that any positive functional on $C_c^0(\mathbb{R}^n;\mathbb{R})$ is essentially integration against a suitable measure.

**Definition 2.4.** Let $E$ be a topological space and $\mathcal{E}$ a $\sigma$-algebra on $E$ containing $\mathcal{B}(E)$. We say a measure $\mu$ on $(E,\mathcal{E})$ is regular (or Borel regular) if for any $A \in \mathcal{E}$ and any $\epsilon > 0$, there is a closed set $C$ and an open $O$ such that $C \subset A \subset O$ and $\mu(O \setminus C) < \epsilon$.

**Example 2.2.** The Lebesgue measure $\lambda$ on $(\mathbb{R}^n,\mathcal{M})$ is regular.

Suppose $\mu$ is a finite regular meaure on $(\mathbb{R}^n,\mathcal{M})$. Any $f \in C_c^0(\mathbb{R}^n;\mathbb{R})$ would be measurable since $\mathcal{M} \supset \mathcal{B}(\mathbb{R}^n)$. The map

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

is then a positive bounded linear operator $C_c^0(\mathbb{R}^n;\mathbb{R}) \to \mathbb{R}$.
We want to recover $\mu$ from $\Lambda$. We'd quite like to set $f = 1_A$ and call it a day, but indicators are not members of $C_c^0(\mathbb{R}^n;\mathbb{R})$. The obvious way to continue is then to approximate them with elements in $C_c^0(\mathbb{R}^n;\mathbb{R})$.
Suppose $O$ is open. For $k \in \mathbb{N}$, we set $O_k = O \cap \{|x| < k\}$. Define

$$\chi_k(x) = \begin{cases} 1 & \text{for } x \in O_k \text{ with } d(x, O_k^c) \geq k^{-1} \\ kd(x,O_k^c) & \text{for } x \in O_k \text{ with } d(x,O_k^c) < k^{-1} \\ 0 & \text{for } x \notin O_k \end{cases}$$

Then $\chi_k \in C_c^0(\mathbb{R}^n;\mathbb{R})$ and $\chi_k \uparrow 1_O$. Monotone convergence then gives $\Lambda(\chi_k) \to \mu(O)$ which recovers everything about $\mu$ by regularity. Hence we have

**Lemma 2.5.** *Given a $\sigma$-algebra $\mathcal{M}$ on $\mathbb{R}^n$ containing $\mathcal{B}(\mathbb{R}^n)$ and a finite regular measure $\mu$ on $(\mathbb{R}^n,\mathcal{M})$, the map*

$$\Lambda : f \mapsto \int_{\mathbb{R}^n} f \,\mathrm{d}\mu$$

*defines a positive bounded linear operator $C_c^0(\mathbb{R}^n;\mathbb{R}) \to \mathbb{R}$. Furthermore, $\Lambda$ uniquely determines $\mu$.*

**Theorem 2.6** (Riesz-Markov)**.** *Given a positive bounded linear operator $\Lambda : C_c^0(\mathbb{R}^n;\mathbb{R}) \to \mathbb{R}$, there is a $\sigma$-algebra $\mathcal{M}$ containing $\mathcal{B}(\mathbb{R})$ and a unique finite regular measure $\mu$ such that*

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

*Proof.* Too technical even for this course. □

**Example 2.3.** For $\Lambda(f) = f(0)$, we recover $\mu = \delta$, the point measure at 0.

## 2.3 Strong, Weak, and Weak-* Topologies

Banach spaces have a natural topology coming from the norm. In some circumstances, this topology is too strong for our liking. E.g. Bolzano-Weierstrass fails for the closed unit ball when the dimension is infinite.
In search for compactness, we resort to disqualify some sets from being open (i.e. trying to consider a weaker topology).

**Definition 2.5.** A seminorm $p$ on a vector space $X$ over a field $\mathbb{F} = \mathbb{R}$ or $\mathbb{C}$ is a map $p : X \to \mathbb{R}$ with:
1. $p(x + y) \leq p(x) + p(y)$.
2. $p(\lambda x) = |\lambda| p(x)$

It then follows that any seminorm is nonnegative. This is a weaker notion than a norm precisely in the sense that $p(x) = 0$ doesn't necessarily mean $x = 0$. What would you do when you don't have quality? Do something with quantity, I guess.

**Definition 2.6.** A family $\mathcal{P}$ of seminorms on $V$ is separating if, for every $x \in X \setminus \{0\}$, there is some $p \in \mathcal{P}$ with $p(x) \neq 0$.

A separating family $\mathcal{P}$ of seminorms induces a topology $\tau_{\mathcal{P}}$ on $X$.
For $p \in \mathcal{P}, n \in \mathbb{N}$, we set $V(p, n) = \{x \in X : p(x) < 1/n\}$. Let $\dot{\beta}$ be the collection of finite intersections of $V(p, n)$'s and $\beta = \{x + B : x \in X, B \in \dot{\beta}\}$.

**Theorem 2.7.** *$\beta$ is a base for a Hausdorff topology $\tau_{\mathcal{P}}$ on $X$ such that addition, scalar multiplication, and all $p \in \mathcal{P}$ are continuous.*

**Definition 2.7.** A space of the form $(X, \tau_{\mathcal{P}})$ is called a locally convex topological vector space (LCTVS).

If $\mathcal{P} = \{p_n\}_{n \in \mathbb{N}}$ is countable, then the corresponding topology $\tau_{\mathcal{P}}$ is metrizable by

$$d(x, y) = \sum_{n=1}^{\infty} 2^{-n} \frac{p_n(x - y)}{1 + p_n(x - y)}$$

If this metric is complete, we say $X$ is Fréchet.
In the case where $X$ is a Banach space, taking $\mathcal{P} = \{\|\cdot\|\}$ gives the exact same topology. This topology $\tau_{\mathrm{s}} = \tau_{\mathcal{P}}$ is usually known as the "strong" topology, only because we are about to define weak topologies. We know that $x_n \to x$ in $\tau_{\mathrm{s}}$ iff $\|x - x_n\| \to 0$ as $n \to \infty$. When this happens, we say $x_n$ converges strongly to $x$ and write $x_n \to x$.
For $\Lambda \in X'$, $p_\Lambda : x \mapsto |\Lambda(x)|$ satisfies the condition for a seminorm. The family $\mathcal{P} = \{p_\Lambda : \Lambda \in X'\}$ is separating, thus induces a Hausdorff topology $\tau_{\mathrm{w}} = \tau_{\mathcal{P}}$, known as the weak-topology on $X$. It is then immediate that $x_n \to x$ in $\tau_{\mathrm{w}}$ iff $\Lambda(x_n) \to \Lambda(x)$ for all $\Lambda \in X'$. If this happens, we say that $x_n$ converges weakly to $x$ and write $x_n \rightharpoonup x$.
$X'$ is also a Banach space, hence has associated strong and weak topologies. It also has another interesting topology $\tau_{\mathrm{w}*}$, the weak-* topology generated by

the separating family of seminorms $\{p_x : \Lambda \mapsto |\Lambda(x)| : x \in X\}$. So $\Lambda_n \to \Lambda$ in $\tau_{w*}$ iff $\Lambda_n(x) \to \Lambda(x)$ for all $x \in X$. If this happens, we say that $\Lambda_n$ converges weakly-* to $\Lambda$ and write $\Lambda_k \rightharpoonup^* \Lambda$. When $X$ is reflexive, the weak and weak-* topologies on $X'$ coincide. And it is always true that $\Lambda_k \rightharpoonup \Lambda$ implies $\Lambda_k \rightharpoonup^* \Lambda$.

**Example 2.4.** 1. Let $1 \le p < \infty$ and $f_i \in L^p(\mathbb{R}^n)$. Then $f_i \to f$ iff $\|f_i - f\|_{L^p} \to 0$, as per usual. $f_i \rightharpoonup f$ iff

$$\int_{\mathbb{R}^n} g f_i \to \int_{\mathbb{R}^n} g f$$

for all $g \in L^q(\mathbb{R}^n)$ with $p^{-1} + q^{-1} = 1$. When $1 < p < \infty$, the space is reflexive so this is further equivalent to $f_i \rightharpoonup^* f$ by viewing $L^p(\mathbb{R}^n) = L^q(\mathbb{R}^n)'$.
For $f_i \in L^\infty(\mathbb{R}^n) = L^1(\mathbb{R}^n)'$, $f_i \to f$ iff $\|f_i - f\|_{L^\infty} \to 0$ still. But now $f_i \rightharpoonup^* f$ is equivalent to

$$\int_{\mathbb{R}^n} f_i g \to \int_{\mathbb{R}^n} f g$$

for all $g \in L^1(\mathbb{R})$, which is NOT the same as $f_i \rightharpoonup f$.
2. Let $H$ be a separable Hilbert space and $u, u_i \in H$. Then $u_i \to u$ iff $\|u_i - u\| \to 0$ whereas $u_i \rightharpoonup u$ iff $\langle w, u_i\rangle \to \langle w, u\rangle$ for all $w \in H$, which in turn is equivalent to $\langle e_n, u_i\rangle \to \langle e_n, u\rangle$ for all $n$, where $(e_n)_n$ is a Hilbertian basis for $H$.

## 2.4 Compactness

We've all seen the celebrated Arzelà-Ascoli theorem, which says

**Theorem 2.8** (Arzelà-Ascoli, a slightly weaker version)**.** *Let $f_i : [0,1] \to \mathbb{C}$ be a bounded equicontinuous sequence of functions, then $(f_i)$ has a uniformly converging subsequence.*

We want to get an even nicer compactness result in a weaker topology.

**Theorem 2.9** (Banach-Alaoglu)**.** *Let $X$ be a normed space and $\bar{B}' = \{\Lambda \in X' : \|\Lambda\|_{X'} \le 1\}$ the closed unit ball in the dual space. Then $\bar{B}'$ is compact in the weak-* topology.*

This is too general for our liking (git gud), and its proof relies on Tychonoff's theorem (therefore Axiom of Choice), so let's prove a weaker version.

**Theorem 2.10** (Banach-Alaoglu, weaker version)**.** *Let $X$ be a separable Banach space. For any sequence $(\Lambda_j)_{j=1}^\infty \in \bar{B}'$, there exists a subsequence $(\Lambda_{j_k})_{k=1}^\infty$ and $\Lambda \in \bar{B}'$ with $\Lambda_{j_k} \rightharpoonup^* \Lambda$.*

*Proof.* Let $D = (x_l)_{l=1}^\infty$ be a countable dense subset of $X$. Consider $(\Lambda_j(x_1))_{j=1}^\infty$ which is a bounded sequence in $\mathbb{C}$ as $|\Lambda_j(x_1)| \le \|\Lambda_j\|_{X'}\|x_1\|_X \le \|x_1\|_X$. By Bolzano-Weierstrass there is some $\Lambda(x_1) \in \mathbb{C}$ and a converging subsequence $\Lambda_{j_k}(x_1) \to \Lambda(x_1)$. Moreover, $|\Lambda(x_1)| \le \|x_1\|_X$. Write $\Lambda_{1,k} = \Lambda_{j_k}$ for brevity.
By a similar argument, we can also find a subsequence $\Lambda_{1,k_i} \to \Lambda(x_2)$ and we set $\Lambda_{2,i} = \Lambda_{1,k_i}$. Proceeding iteratively gives a family $(\Lambda_{l,k})_{l,k}$ such that $(\Lambda_{l,k})_{k=1}^\infty$ is a subsequence of $(\Lambda_{l-1,k})_{k=1}^\infty$ for all $l > 1$, and $\Lambda_{l,k}(x_j) \to \Lambda(x_j)$ for all $j \le l$ for some $\Lambda(x_j) \in \mathbb{C}$ with $|\Lambda(x_j)| \le \|x_j\|_X$.
Now $\Lambda_{j,j}(x) \to \Lambda(x)$ for all $x \in D$ where $\Lambda(x) \in \mathbb{C}, |\Lambda(x)| \le \|x\|_X$. We first

claim that $\Lambda : D \to \mathbb{C}$ is uniformly continuous (in fact it's 3-Lipschitz). For $\epsilon > 0$, suppose $x, y \in D$ has $|x - y| < \epsilon/3$. Pick $k$ sufficiently large so that $|\Lambda_{k,k}(x) - \Lambda(x)| < \epsilon/3, |\Lambda_{k,k}(y) - \Lambda(y)| < \epsilon/3$, then

$$|\Lambda(x) - \Lambda(y)| \leq |\Lambda(x) - \Lambda_{k,k}(x)| + |\Lambda_{k,k}(y) - \Lambda(y)| + |\Lambda_{k,k}(x) - \Lambda_{k,k}(y)| < \epsilon$$

Thus $\Lambda$ extends uniquely to a continuous function $X = \bar{D} \to \mathbb{C}$.
$\Lambda$ is also linear: Suppose $x, y \in X, a \in \mathbb{C}$. We set $z = x + ay$. For $\epsilon > 0$, then $|\Lambda(z) - \Lambda(x) - a\Lambda(y)|$ is bounded by

$$\begin{aligned}&|\Lambda(z) - \Lambda(z')| + |\Lambda(x) - \Lambda(x')| + |a||\Lambda(y) - \Lambda(y')| \\ &+ |\Lambda(z') - \Lambda_{k,k}(z')| + |\Lambda(x') - \Lambda_{k,k}(x')| + |a||\Lambda(y') - \Lambda_{k,k}(y')| \\ &+ |\Lambda_{k,k}(z' - x' - ay')|\end{aligned}$$

By choice of $x', y', z' \in D$ we can bound the first and third lines both under $\epsilon/3$, and by choice of $k$ the second line too. Thus this can be bounded below $\epsilon$, which gives the linearity of $\Lambda$. Combining this with continuity shows $\Lambda \in X'$.
It remains to show that $\Lambda_{j,j} \rightharpoonup^* \Lambda$. For $x \in X$, we do the estimate

$$|\Lambda_{j,j}(x) - \Lambda(x)| \leq |\Lambda_{j,j}(x - x')| + |\Lambda_{j,j}(x') - \Lambda(x')| + |\Lambda(x') - \Lambda(x)|$$

Again the first and third term can be bounded by choice of $x'$ and the second term by choice of $j$, which gives the desied convergence. And we have $\|\Lambda\|_{X'} \leq 1$ by construction. □

**Corollary 2.11.** *Suppose $1 < p \leq \infty$ and let $(f_j)_{j=1}^\infty$ be a sequence in $L^p(\mathbb{R}^n)$ satisfying $\forall j, \|f_j\|_{L^p} \leq K$. Then there is a subsequence $(f_{j_k})_{k=1}^\infty$ and $f \in L^p(\mathbb{R}^n)$ such that $\|f\|_{L^p} \leq K$ and*

$$\int_{\mathbb{R}^n} f_{k_j} g \to \int_{\mathbb{R}^n} fg \, \mathrm{d}x$$

*as $j \to \infty$ for all $g \in L^q(\mathbb{R}^n)$ with $p^{-1} + q^{-1} = 1$.*

*Remark.* If $X$ is reflexive (e.g. if $X$ is a Hilbert space), then Theorem 2.10 shows that every bounded sequence has a weakly convergent subsequence.

## 2.5 Hahn-Banach Theorem

The Hahn-Banach theorem concerns the problem of extending a bounded linear map defined on a subspace $M \leq X$ to a bounded linear map on $X$.
We need only to consider the case where the base field is $\mathbb{R}$, as every vector space over $\mathbb{C}$ also has the natural structure of a vector space over $\mathbb{R}$. Furthermore, a complex linear form $\Lambda : X \to \mathbb{C}$ gives rise to a real linear form $\ell = \operatorname{Re} \Lambda : X \to \mathbb{R}$; Conversely, a real linear form $\ell$ induces a complex linear form $\Lambda(x) = l(x) - i\ell(ix)$. Such a correspondence translates almost all results we'll obtain.
From now on, we work over $\mathbb{R}$ as our base field. For the sake of glorious generality, we introduce a new notion of boundedness to state the theorem.

**Definition 2.8.** A sublinear functional on a (real) vector space $X$ is a map $p : X \to \mathbb{R}$ satisfying $p(x + y) \leq p(x) + p(y)$ and $p(tx) = tp(x)$ for all $x, y \in X, t \geq 0$.

**Example 2.5.** 1. If $\ell : X \to \mathbb{R}$ is linear, then $p = |\ell|$ is sublinear.
2. Indeed, all seminorms are sublinear (but not vice versa!).

For a linear map $\ell : X \to \mathbb{R}$, we will work with one-sided bounds of the form $l(x) \le p(x)$ which in fact implies a two-sided bound $-p(-x) \le l(x) \le p(x)$.

**Lemma 2.12** (Banach Extension Lemma). *Let $X$ be a real vector space and $p : X \to \mathbb{R}$ a sublinear map on $X$ and $M \le X$. Suppose $\ell : M \to \mathbb{R}$ is a linear form with $\ell(y) \le p(y)$ for all $y \in M$, then for any $x \in X \setminus M$, there exists a linear form $\tilde{\ell} : \tilde{M} = \mathrm{Span}(M \cup \{x\}) \to \mathbb{R}$ extending $\ell$ with $\tilde{\ell}(w) \le p(w)$ for all $w \in \tilde{M}$.*

*Proof.* Any $w \in \tilde{W}$ can be uniquely written in the form $w = \lambda x + y, y \in M, \lambda \in \mathbb{R}$. So $\tilde{\ell}$ is determined by $a = \tilde{\ell}(x)$.
Suppose $y, z \in M$, then $\ell(y) + \ell(z) = \ell(y+z) \le p(y+z) \le p(y-x) + p(z+x)$. This can be rewritten as $\ell(y) - p(y-x) \le p(z+x) - \ell(z)$, thus $a = \sup_{y \in M}(\ell(y) - p(y-x)) \le p(z+x) - \ell(z) < \infty$.
So $\ell(y') - a \le p(y'-x)$ and $\ell(z) + a \le p(z+x)$ for all $y', z \in M$. For $\lambda > 0$, the bound follows by letting $z = \lambda^{-1}y$; For $\lambda < 0$, the bound follows by letting $y' = (-\lambda^{-1})y$. □

So if $\dim(X/M)$ is finite, then we can extend any linear form on $\ell$ to a linear form $\ell : X \to \mathbb{R}$ whilst keeping the same (sublinear) bound. When $X$ is separable, we can also obtain the result by induction. The general case, however, requires a form of the Axiom of Choice, namely:

**Proposition 2.13** (Zorn's . . . Proposition?). *Let $(S, \le)$ be a nonempty partially ordered set in which every totally ordered subset has an upper bound, then $S$ contains at least one maximal element.*

**Theorem 2.14** (Hahn-Banach). *Let $X$ be a real vector space and $p : X \to \mathbb{R}$ a sublinear functional. Suppose $M \le X$ and $\ell : M \to \mathbb{R}$ is lienar and $\ell(y) \le p(y)$ for all $y \in M$, then there exists a linear $\tilde{\ell} : X \to \mathbb{R}$ extending $\ell$ with $\tilde{\ell}(x) \le p(x)$ for all $x \in X$.*

*Proof.* Consider the set $S$ of pairs $(N, \ell^*)$ with $M \le N \le X$, $\ell^* : N \to \mathbb{R}$ linear and extending $\ell$, and that $\ell^*(x) \le p(x)$ for all $x \in N$. $S$ is nonempty by Lemma 2.12. We partially order $S$ by saying that $(N_1, \ell_1^*) \le (N_2, \ell_2^*)$ iff $N_1 \le N_2$ and $\ell_2^*$ extends $\ell_1^*$.
Suppose $T$ is a totally ordered subset of $S$, then it has an upper bound given by $(N_T, \ell_T)$ where $N_T = \bigcup_{(N,\ell^*) \in T} N$ and $\ell_T(x) = \ell^*(x)$ whenever $(N, \ell^*) \in T$ and $x \in N$. Proposition 2.13 shows that $S$ admits a maximal element $(\tilde{N}, \tilde{\ell})$. But then necessarily $\tilde{N} = X$ by Lemma 2.12, and we are done. □

**Corollary 2.15.** *Let $X$ be a Banach space over $\mathbb{F} = \mathbb{R}$ or $\mathbb{C}$ and let $M \le X$. Suppose $\Lambda : M \to \mathbb{F}$ is a bounded linear form, then there exists a bounded linear form $\tilde{\Lambda} : X \to \mathbb{F}$ extending $\Lambda$ with $\|\tilde{\Lambda}\|_{X'} = \|\Lambda\|_{M'}$.*

*Proof.* If $\mathbb{F} = \mathbb{R}$, then we can just apply the preceding theorem to $p(x) = \|\Lambda\|_{M'}\|x\|$. If $\mathbb{F} = \mathbb{C}$, write $\Lambda(x) = \ell(x) - i\ell(ix)$ for $\ell : M \to \mathbb{R}$ a real linear operator. Then $|\Lambda(x)| = \ell(e^{i\theta}x)$ for a suitable $\theta$ not depending on $x$, consequently $\sup_{x \in M, \|x\| \le 1} |\Lambda(x)| = \sup_{x \in M, \|x\| \le 1} |\ell(x)|$. Applying the previous result to the real vector space structure on $X$ gives some $\tilde{\ell} : X \to \mathbb{R}$ with $\|\tilde{\ell}(x)\|_X \le \|\ell\|_{M'}\|x\|$. Then $\tilde{\Lambda}(x) = \tilde{\ell}(x) - i\tilde{\ell}(ix)$ gives the desired extension. □

Let's now formulate a more geometric form of Theroem 2.14. Recall that a subset of a vector space $V$ is called convex if for any $x, y \in A$, we have $tx + (1-t)y \in A$ for all $t \in [0,1]$. Recall also that in $\mathbb{R}^n$, we have the following theorem:

**Theorem 2.16** (Supporting Hyperplane Theorem in Finite Dimension)**.** *Suppose $A, B$ are disjoint, compact, convex, nonempty subsets of $\mathbb{R}^n$, then we can find a codimension 1 hyperplane such that $A, B$ are on the opposite sides of it.*

With the power of Theorem 2.14, we can generalise this to infinite dimensions.

**Theorem 2.17** (Geometric Hahn-Banach)**.** *Suppose $A, B$ are disjoint, convex, nonempty subsets in a (real or complex) Banach space $X$.*
*(a) If $A$ is open, then there exists $\Lambda \in X'$ and $\gamma \in \mathbb{R}$ such that $\operatorname{Re}(\Lambda x) < \gamma \leq \operatorname{Re}\Lambda(y)$ for all $x \in A, y \in B$. If in addition that $B$ is also open, then the second inequality is also strict.*
*(b) If $A$ is compact and $B$ closed, then there is some $\Lambda \in X'$ and $\gamma_1, \gamma_2 \in \mathbb{R}$ such that $\operatorname{Re}\Lambda(x) < \gamma_1 < \gamma_2 < \operatorname{Re}\Lambda(y)$ for all $x \in A, y \in B$.*

*Proof.* WLOG $X$ is real.
(a) Pick $a_0 \in A, b_0 \in B$ and set $x_0 = b_0 - a_0$. Set $C = A - B + x_0$, then $C$ is a convex open neighbourhood of 0. Since $A, B$ are disjoint, $x_0 \notin C$. Let $p(x) = \inf\{t > 0 : t^{-1}x \in C\}$ ("Minkowski gauge"). One can check (example sheet) that $p$ is sublinear, $p(x) \leq k\|x\|$ for some constant $k > 0$, and $p(y) < 1$ for all $y \in C$. Since $x_0 \notin C$, $p(x_0) \geq 1$.
Let $M = \mathbb{R}x_0 \leq X$, then $f : M \to \mathbb{R}, f(tx_0) = t$ is a linear form bounded above by $p$ ($f(tx_0) = t \leq tp(x_0) = p(tx_0)$ for any $t > 0$ and $f(tx_0) = t \leq 0 \leq p(tx_0)$ for any $t \leq 0$). Theorem 2.14 then tells us that we can extend $f$ to a linear form $\Lambda : X \to \mathbb{R}$ with $-k\|x\| \leq -p(-x) \leq \Lambda(x) \leq p(x) \leq k\|x\|$, consequently $\Lambda \in X'$.
For any $a \in A, b \in B$, we have $\Lambda(a) - \Lambda(b) + 1 = \Lambda(a - b + x_0) \leq p(a - b + x_0) < 1$, so $\Lambda(a) < \Lambda(b)$. This means that $\Lambda(A)$ and $\Lambda(B)$ are disjoint convex subsets of $\mathbb{R}$ with $\sup \Lambda(A) \leq \inf \Lambda(B)$. By open mapping theorem (example sheet), $\Lambda(A)$ is open and therefore $\gamma = \sup \Lambda(A)$ has $\Lambda(a) < \gamma \leq \Lambda(b)$ for all $a \in A, b \in B$. If $B$ is also open, $\Lambda(B)$ would be open and hence $\gamma \leq \inf \Lambda(B) < \lambda(b)$ for all $b \in B$.
(b) Since $A$ is compact and $B$ closed, $d = \inf_{a \in A, b \in B} \|a - b\| > 0$. Let $V = B_{d/2}(0)$ and consider $A + V$ which is open, convex, and disjoint from $B$. By (a) we can find $\Lambda \in X'$ such that $\Lambda(A + V), \Lambda(B)$ are disjoint and $\sup \Lambda(A + V) \leq \inf \Lambda(B)$. The result then follows from the fact that $\Lambda(A)$ is a compact subset of $\Lambda(A + V)$. □

**Corollary 2.18.** *For any distinct $x, y \in X$, there exists $\Lambda \in X'$ such that $\Lambda(x) \neq \Lambda(y)$.*

*Proof.* Set $A = \{x\}, B = \{y\}$ in Theorem 2.17(b). □

**Corollary 2.19.** *Suppose $M \leq X, \bar{M} \neq X$ and $x_0 \in X \setminus \bar{M}$, then there exists $\Lambda \in X'$ with $\Lambda(x_0) = 1$ and $\Lambda|_M = 0$.*

*Proof.* Set $A = \{x_0\}, B = \bar{M}$ in Theorem 2.17(b). □

*Remark.* This particular corollary is useful in proving the density of a certain subspace (which in turn is useful in the theory of PDEs), since it's usually easier to show the nonexistence of such a linear form $\Lambda$.

# 3 Distribution Theory

We often want to make use of seemingly ill-defined "functions" in the study of PDEs.

**Example 3.1.** Consider

$$G : \mathbb{R}^3 \setminus \{0\} \to \mathbb{R}, x \mapsto \frac{-1}{4\pi|x|}$$

which "morally" should have $\nabla^2 G = \delta$, as it helps so much in the study of Laplace and Poisson equations. But this doesn't make sense!

The theory of distributions provide a rigourous setting where such kind of ideas makes sense. The idea is to construct some spaces of "test functions".

## 3.1 Spaces of Test Functions

Let $\Omega \subset \mathbb{R}^n$ be open.
The set $C_c^\infty(\Omega)$ consisting of smooth functions $\phi : \Omega \to \mathbb{C}$ with compact $\operatorname{supp}\phi = \overline{\{x : \phi(x) \neq 0\}}$ has the natural structure of a vector space. It can be equipped with a topology making it a locally convex topological vector space, which we shall call $\mathscr{D}(\Omega)$. The details of the construction is omitted as it's more important to know its defining properties.

**Theorem 3.1.** *$C_c^\infty(\Omega)$ can be equipped with a topology $\tau$ such that:*
*(i) The topology makes it a locally convex topological vector space.*
*(ii) A sequence $(\phi_j)_j \in C_c^\infty(\Omega)$ converges to 0 in $\tau$ if there is a compact $K \subset \Omega$ such that $\operatorname{supp}\phi_j \subset K$ for all $j$ and for each multi-index $\alpha$ we have $\sup_{x\in K} |D^\alpha \phi_j(x)| \to 0$.*
*(iii) If $Y$ is a locally convex topological vector space and $\Lambda : C_c^\infty(\Omega) \to Y$ is linear, then $\Lambda$ is continuous iff it is sequentially continuous, i.e. $\Lambda\phi_j \to \Lambda\phi$ whenever $\phi_j \to \phi$.*

**Example 3.2.** For $\phi \in C_c^\infty(\mathbb{R})$ nonzero:
(a) If $\phi_j(x) = e^{-j}\phi(jx)$, then $\phi_j \to 0$ in $\mathscr{D}(\mathbb{R})$.
(b) If $\phi_j(x) = j^{-1000}\phi(jx)$, then $\phi_j$ does not converge to 0 in $\mathscr{D}(\mathbb{R})$.
(c) If $\phi_j(x) = e^{-j}\phi(x-j)$, then again $\phi_j$ does not converge to 0 in $\mathscr{D}(\mathbb{R})$.

The next space of test functions we will be interested in is $\mathscr{E}(\Omega)$, which is a structure based on the vector space $C^\infty(\Omega)$. We construct this by giving $C^\infty(\Omega)$ a Fréchet topology.
First recall that $\Omega$ admits an exhaustion by compact sets, i.e. there is a sequence of compact sets $K_i \subset \Omega$ with $K_i \subset K_{i+1}^\circ$ and $\Omega = \bigcup_i K_i$. For $\phi \in C^\infty(\Omega)$, let

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

Then $\{p_N : N\}$ is a separating family of seminorms, which induces a topology $\tau$ on $C^\infty(\Omega)$ making it a locally convex topological vector space. We call this structure $\mathscr{E}(\Omega)$. This can be induced from a metric which is complete, hence it's a Fréchet topology.

**Theorem 3.2.** *For $(\phi_j)_j \in C^\infty(\Omega)$, $\phi_j \to 0$ in $\mathscr{E}(\Omega)$ iff for each compact $K$ and multi-index $\alpha$, $\sup_{x\in K} |D^\alpha \phi_j(x)| \to 0$.*

**Example 3.3.** If $\phi \in C_c^\infty(\mathbb{R})$ is nonzero, then $\phi_j(x) = \phi(x-j)$ converges to 0 in $\mathscr{E}(\mathbb{R})$ as $j \to \infty$.

One can show (example sheet) that the inclusion $\mathscr{D}(\Omega) \hookrightarrow \mathscr{E}(\Omega)$ is continuous.
The last space we'll look at is the space $\mathscr{S}(\mathbb{R}^n)$.

**Definition 3.1.** A function $\phi \in C^\infty(\mathbb{R}^n)$ is rapidly decreasing if $\sup_{x\in\mathbb{R}^n} |(1+|x|)^N D^\alpha \phi(x)| < \infty$ for all multi-index $\alpha$ and $N \in \mathbb{N}$.

**Example 3.4.** $\phi(x) = e^{-|x|^2}$ is rapidly decreasing, but $\phi(x) = (1+|x|^2)^{-1001}$ is not.

For $\phi$ rapidly decreasing, we can consider the seminorms

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

$\{p_N : N\}$ is again a separating family of seminorms and they make the set of rapidly decreasing smooth functions into a locally convex topological vector space, which is in fact Fréchet. Unlike $\mathscr{D}(\Omega)$ and $\mathscr{E}(\Omega)$, $\mathscr{S}(\mathbb{R}^n)$ has a name and it's called the Schwarz space.
We can construct $\mathscr{S}(\mathbb{R}^n)$ alternatively with the seminorms $\sup_{x\in\mathbb{R}^n,|\alpha|\le N} |(1+|x|^2)^N D^\alpha \phi(x)|$ or $\sup_{x\in\mathbb{R}^n,|\alpha|\le N,|\beta|\le N} |x^\beta D^\alpha \phi(x)|$.

**Theorem 3.3.** *If $(\phi_j)_j \in \mathscr{S}(\mathbb{R}^n)$, then $\phi_j \to 0$ in $\mathscr{S}(\mathbb{R}^n)$ iff for all multi-index $\alpha$ and $N \in \mathbb{N}$ we have $\sup_{x\in\mathbb{R}^n} |(1+|x|)^N D^\alpha \phi_j(x)| \to 0$.*

We also have $\mathscr{D}(\mathbb{R}^n) \subset \mathscr{S}(\mathbb{R}^n) \subset \mathscr{E}(\mathbb{R}^n)$ with each inclusion continuous.

## 3.2 Distributions

For $\Omega \subset \mathbb{R}^n$ open, the space of distributions $\mathscr{D}'(\Omega)$ on $\Omega$ is the continuous dual space of $\mathscr{D}(\Omega)$. Note that a linear map $u : \mathscr{D}(\Omega) \to \mathbb{C}$ is continuous iff $u[\phi_j] \to 0$ for all $\phi_j \to 0$ in $\mathscr{D}(\Omega)$. $\mathscr{D}'(\Omega)$ is equipped with the weak-* topology, which as you recall is the one induced by the separating family of seminorms $\{p_\phi : \phi \in \mathscr{D}(\Omega)\}\}$ where $p_\phi[u] = |u[\phi]|$. Recall also that $u_j \to u$ weakly-* in $\mathscr{D}'(\Omega)$ iff $u_j[\phi] \to u[\phi]$ for all $\phi \in \mathscr{D}(\mathbb{R})$.

**Example 3.5.** (a) For $x \in \Omega$, we can consider $\delta_x : \mathscr{D}(\Omega) \to \mathbb{C}$ via $\phi \mapsto \phi(x)$ which is a member of $\mathscr{D}'(\Omega)$. It is known as the Dirac $\delta$.
(b) For $f \in L^1_{\text{loc}}(\Omega)$, we can consider $T_f : \mathscr{D}(\Omega) \to \mathbb{C}$ via

$$\phi \mapsto \int_\Omega f\phi$$

which is a member of $\mathscr{D}'(\Omega)$: Suppose $\phi_j \to 0$ in $\mathscr{D}(\Omega)$, then there is some $K \subset \Omega$ compact such that $\operatorname{supp} \phi_j \subset K$ for all $j$ and $\sup_{x\in K} |D^\alpha \phi_j(x)| \to 0$ for all multi-index $\alpha$. As $\phi_j$ then necessarily converges uniformly, they are in particular uniformly bounded by some constant $C$. This means that $|f 1_K \phi_j| \leq C|f|1_K$ and the latter is integrable. Applying dominated convergence then shows that

$$T_f \phi_j = \int_\Omega f\phi_j = \int_\Omega f 1_K \phi_j \to 0$$

as $j \to \infty$.

**Lemma 3.4.** *Suppose $u : \mathscr{D}(\Omega) \to \mathbb{C}$ is linear. Then $u$ is continuous iff for each compact $K \subset \Omega$, there exists $N_K \in \mathbb{N}$ and $C_K > 0$ such that*

$$|u[\phi]| \leq C_K \sup_{x\in K} \sum_{|\alpha|\leq N_K} |D^\alpha \phi(x)|$$

*for all $\phi \in C_c^\infty(K)$.*

*Proof.* For the "if" direction, suppose $\phi_j \to 0$ in $\mathscr{D}(\Omega)$, then there is some compact $K \subset \Omega$ with $\operatorname{supp} \phi \subset K$ for all $j$ and $D^\alpha \phi_j \to 0$ uniformly on $K$ for all $\alpha$. The condition then forces $|u[\phi_j]| \to 0$, so $u$ is continuous.
Conversely, suppose $u$ is continuous but there is some compact $K$ and $\phi_j \in C_c^\infty(K)$ such that for all $j$,

$$|u[\phi_j]| > j \sup_{x\in K} \sum_{|\alpha|\leq j} |D^\alpha \phi(x)|$$

Let $\psi_j = \phi_j/|u[\phi_j]|$, then $\psi_j \to 0$ in $\mathscr{D}(\Omega)$. For any multi-index $\beta$,

$$|D^\beta \psi_j(x)| = \frac{|D^\beta \phi_j(x)|}{|u[\phi_j]|} \leq \frac{|D^\beta \phi_j(x)|}{j \sup_{x\in K} \sum_{|\alpha|\leq j} |D^\alpha \phi_j(x)|} \leq \frac{1}{j} \to 0$$

eventually as $j \to \infty$. So $\psi_j \to 0$, but $|u[\psi_j]| = 1$, contradicting the continuity of $u$. □

If we can choose $N = N_K$ independently from $K$, we say $u$ has finite order with order the least such $N$.

**Example 3.6.** 1. $\delta_x, T_f$ have order 0.
2. $\phi \mapsto \phi'(x)$ has order 1.
3. $u : \mathscr{D}(\mathbb{R}) \to \mathbb{C}, \phi \mapsto \sum_{i=1}^\infty \phi^{(i)}(i)$ is continuous but does not have finite order.

If $f, g \in L^1_{\text{loc}}(\Omega)$ have $T_f = T_g$, then

$$\int_\Omega (f-g)\phi = 0$$

for all $\phi \in C_c^\infty(\Omega)$. Using the test functions as mollifiers (Theorem 1.4), we conclude that $f = g$ a.e.. Thus $T : f \mapsto T_f$ is an injection $L^1_{\text{loc}}(\Omega) \to \mathscr{D}'(\Omega)$, so it makes sense to abuse the notation and write $f$ in place of $T_f$ and write $L^1_{\text{loc}}(\Omega)$ in place of its image under this injection.
$\mathscr{D}'(\Omega)$ is a $\mathbb{C}$-vector space, so we can make sense of a distribution that is a $\mathbb{C}$-linear combinations of distributions. Suppose $a \in C^\infty(\Omega)$, then for any $\phi \in$

$\mathscr{D}(\Omega)$ we have $T_{af}[\phi] = T_f[a\phi]$ for any $f \in L^1_{\text{loc}}(\Omega)$. Thus suggests that for any distribution $u$ and $a \in C^\infty(\Omega)$ we might be able to consider a new distribution via $au[\phi] = u[a\phi]$, which works.
Similarly, if $f \in C^1(\mathbb{R}), \phi \in \mathscr{D}(\Omega)$, we have $T_{D_i f}[\phi] = -T_f[D_i\phi]$. Motivated by this, we can define the derivative $D^\alpha$ (with $\alpha$ a multi-index) of a distribution $u$ via $D^\alpha u[\phi] = (-1)^{|\alpha|}u[D^\alpha\phi]$. However, one should note that this operation doesn't necessarily stabilise the subspace $L^1_{\text{loc}}(\Omega) \subset \mathscr{D}'(\Omega)$.

**Example 3.7.** Let $H : \mathbb{R} \to \mathbb{R}$ be the Heaviside step function $H = 1_{\mathbb{R}_{\geq 0}}$, then essentially $DT_H = \delta_0$.

We can now make sense of (linear) distributional PDEs, i.e. equations of the form $\sum_{|\alpha|\leq k} a_\alpha D^\alpha u = w$ where $w \in \mathscr{D}'(\Omega), a_\alpha \in C^\infty(\Omega)$ are given and we want $u \in \mathscr{D}'(\Omega)$.
How about the space $\mathscr{E}'(\Omega)$, the continuous dual to $\mathscr{E}(\Omega)$? This is a Fréchet space, so a linear map $u : \mathscr{E}(\Omega) \to \mathbb{C}$ is continuous iff it's sequentially continuous.

**Lemma 3.5.** *Suppose $u : \mathscr{E}(\Omega) \to \mathbb{C}$ is linear. Then $u$ is continuous iff there exists compact $K \subset \Omega$, $N \in \mathbb{N}$ and $C > 0$ such that*

$$|u[\phi]| \leq C \sup_{x\in K,|\alpha|\leq N} |D^\alpha\phi(x)|$$

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

*Proof.* The "if" direction is again straightforward: Let $\phi_j \to 0$ in $\mathscr{E}(\Omega)$, then for any compact $\tilde{K} \subset \Omega$ and $\tilde{N} \in \mathbb{N}$ we have $\sup_{x\in\tilde{K},|\alpha|\leq\tilde{N}} |D^\alpha\phi_j(x)| \to 0$. The condition then shows that $|u[\phi_j]| \to 0$.
Conversely, suppose the condition is not true. Choose a compact exhausion $(K_i)_i$ of $\Omega$ (i.e. each $K_j$ compact in $\Omega$, $K_j \subset K^\circ_{j+1}$ and $\bigcup_j K_j = \Omega$). So for each $j$ we can find $\phi_j \in \mathscr{E}(\Omega)$ such that

$$|u[\phi_j]| \geq j \sup_{x\in K_j,|\alpha|\leq j} |D^\alpha\phi_j(x)|$$

Let $\psi_j = \phi_j/|u[\phi_j]|$. Then $\psi_j \to 0$ in $\mathscr{E}(\Omega)$ since for any compact $\tilde{K}$ and any $\tilde{N} \in \mathbb{N}$, there is some $J > \tilde{N}$ with $\tilde{K} \subset K_j$ for any $j \geq J$. This means that

$$\sup_{x\in\tilde{K},|\alpha|\leq\tilde{N}} |D^\alpha\psi_j(x)| \leq \frac{1}{|u[\phi_j]|} \sup_{x\in K_j,|\alpha|\leq j} |D^\alpha\phi_j(x)| \leq \frac{1}{j} \to 0$$

eventually as $j \to \infty$. However, $|u[\psi_j]| = 1$, so $u[\psi_j]$ does not converge to 0 as $j \to \infty$ which means that $u$ is not continuous. □

The continuous inclusion $\mathscr{D}(\Omega) \subset \mathscr{E}(\Omega)$ gives rise to a continuous linear map $\mathscr{E}'(\Omega) \subset \mathscr{D}'(\Omega)$. Turns out this map is injective, so we can identify $\mathscr{E}'(\Omega)$ with the distributions of compact support.

**Definition 3.2.** We say a distribution $u \in \mathscr{D}'(\Omega)$ has compact support if there is some compact $K \subset \Omega$ such that $u[\phi] = 0$ whenever $\operatorname{supp}\phi \subset \Omega \setminus K$.

The preceding lemma shows that any $u \in \mathscr{E}'(\Omega) \subset \mathscr{D}'(\Omega)$ has compact support. Conversely, if $u \in \mathscr{D}'(\Omega)$ has compact support inside some compact $K$, we can extend $u$ uniquely to $\tilde{u} \in \mathscr{E}'(\Omega)$ by setting $\tilde{u}[\phi] = u[\chi\phi]$ where $\chi \in C^\infty_c(\Omega)$ has $\chi|_K \equiv 1$. It's clear that such $\chi$ exists and the extension does not depend on the choice of $\chi$.

**Example 3.8.** (a) If $f \in L^1(\Omega)$, vanishes a.e. in $\Omega \setminus K$ for some compact $K \subset \Omega$, then $T_f \in \mathscr{E}'(\Omega)$.
(b) For any $x \in \Omega$, $\delta_x \in \mathscr{E}'(\Omega)$.
(c) $u \in \mathscr{D}'(\mathbb{R})$ given by $u[\phi] = \sum_{n=-\infty}^{\infty} \phi(n)$ does not have compact support.

The continuous dual $\mathscr{S}'(\mathbb{R}^n)$ to $\mathscr{S}(\mathbb{R}^n)$ is called the space of tempered distributions. The continuous inclusions $\mathscr{D}(\mathbb{R}^n) \subset \mathscr{S}(\mathbb{R}^n) \subset \mathscr{E}(\mathbb{R}^n)$ induces continuous maps $\mathscr{E}'(\mathbb{R}^n) \to \mathscr{S}'(\mathbb{R}^n) \to \mathscr{D}'(\mathbb{R}^n)$, both of which happen to be injective.

**Example 3.9.** 1. If $f \in L^1_{\text{loc}}(\mathbb{R}^n)$ satisfies

$$\int_{\mathbb{R}^n} (1 + |x|)^{-N} |f(x)| \, \mathrm{d}x < \infty$$

Then $T_f$ is a tempered distribution. Indeed, if $\phi \in \mathscr{S}(\mathbb{R}^n)$, then

$$\begin{aligned}|T_f[\phi]| &= \left| \int_{\mathbb{R}^n} f\phi \right| = \left| \int_{\mathbb{R}^n} (1 + |x|)^{-N} f(x)(1 + |x|)^N \phi(x) \, \mathrm{d}x \right| \\ &\le \left( \int_{\mathbb{R}^n} (1 + |x|)^{-N} |f(x)| \, \mathrm{d}x \right) \left( \sup_{x \in \mathbb{R}^n} |(1 + |x|)^N \phi(x)| \right)\end{aligned}$$

So $|T_f[\phi_j]| \to 0$ whenever $\phi_j \to 0$ in $\mathscr{S}(\mathbb{R}^n)$, which means that $T_f \in \mathscr{S}'(\mathbb{R}^n)$.
2. If $f(x) \in e^{|x|^2}$, then $T_f \in \mathscr{D}'(\mathbb{R}^n) \setminus \mathscr{S}'(\mathbb{R}^n)$.
3. The map $u[\phi] = \sum_{m \in \mathbb{Z}} m^{50} \phi(m)$ belongs to $\mathscr{S}'(\mathbb{R})$ but not $\mathscr{E}'(\mathbb{R})$.

## 3.3 Convolutions

Recall that if $\phi : \mathbb{R}^n \to \mathbb{C}$, we had defined $\tau_x$ via $\tau_x \phi(y) = \phi(y - x)$. We introduce the involute $\check{\phi}$ of $\phi$ via $\check{\phi}(y) = \phi(-y)$ and the convolute of $\phi$ via $\tau_x \check{\phi}(y) = \phi(x - y)$.
Suppose $f \in L^1_{\text{loc}}(\mathbb{R}^n), \phi \in \mathscr{D}(\mathbb{R}^n)$, then their convolution is defined by

$$(f * \phi)(x) = \int_{\mathbb{R}^n} f(y)\phi(x - y) \, \mathrm{d}y = T_f(\tau_x \check{\phi})$$

which suggests that we might be able to define convolution of a distribution $u \in \mathscr{D}'(\mathbb{R}^n)$ and $\phi \in \mathscr{D}(\mathbb{R}^n)$ by $u * \phi(x) = u(\tau_x \check{\phi})$. This is perfectly well-defined and bilinear. We also have $(u * \check{\phi})(0) = u[\phi]$, so $u$ is determined by all these $u * \phi, \phi \in \mathscr{D}(\mathbb{R}^n)$.

**Example 3.10.** $\delta_0 * \phi = \phi$.

**Lemma 3.6.** *Suppose $u \in \mathscr{D}'(\mathbb{R}^n), \phi \in \mathscr{D}(\mathbb{R}^n)$, then:*
*(i) $u * \phi \in C^\infty(\mathbb{R}^n)$ and $D^\alpha(u * \phi) = D^\alpha u * \phi = u * D^\alpha \phi$.*
*(ii) If $u \in \mathscr{E}'(\mathbb{R}^n)$, then $u * \phi \in \mathscr{D}(\mathbb{R}^n)$.*

*Proof.* (i) We have

$$\frac{u * \phi(x + he_i) - u * \phi(x)}{h} = u \left[ \frac{\tau_{x+he_i} \check{\phi} - \tau_x \check{\phi}}{h} \right]$$

But $h^{-1}(\tau_{x+he_i} \check{\phi} - \tau_x \check{\phi}) \to \tau_x (\check{D_i \phi})(y)$ as $h \to 0$ in $\mathscr{D}(\mathbb{R}^n)$ (example sheet), so indeed $h^{-1}(u * \phi(x + he_i) - u * \phi(x)) = u * D_i \phi(x)$.
Iterating this gives $u * \phi \in C^\infty(\mathbb{R}^n)$ and $D^\alpha(u * \phi) = u * D^\alpha \phi$. Furthermore,

$$D^\alpha \tau_x \check{\phi}(y) = \frac{\partial^{|\alpha|}}{\partial y^\alpha} \phi(x - y) = (-1)^{|\alpha|} D^\alpha \phi(x - y) = (-1)^{|\alpha|} \tau_x (\check{D^\alpha \phi})(y)$$

Hence $(D^\alpha u * \phi)(x) = D^\alpha u[\tau_x \check{\phi}] = (-1)^{|\alpha|} u[D^\alpha \tau_x \check{\phi}] = (-1)^{|\alpha|} u[D^\alpha \tau_x \check{\phi}] = (-1)^{|\alpha|} u[(-1)^{|\alpha|} \tau_x (\check{D^\alpha \phi})] = (u * D^\alpha \phi)(x)$.
(ii) Suppose $u[\phi] = 0$ for all $\phi \in \mathscr{D}(\mathbb{R}^n)$ supported outside $K$, then for any $\phi \in \mathscr{D}(\mathbb{R}^n)$ we have $K \cap \operatorname{supp} \tau_x \check{\phi} = \varnothing$ for $|x|$ sufficiently large. Thus $u[\tau_x \check{\phi}] = 0$ for $|x|$ large and therefore $u * \phi \in \mathscr{D}(\mathbb{R}^n)$. □

With this, we can define the convolution of distributions $u_1 \in \mathscr{D}'(\mathbb{R}^n), u_2 \in \mathcal{E}'(\mathbb{R}^n)$ by requiring $(u_1 * u_2) * \phi = u_1 * (u_2 * \phi)$. Recall that $u * \check{\phi}(0) = u[\phi]$, so this does determine a new distribution.

**Example 3.11.** If $u \in \mathscr{D}'(\mathbb{R}^n)$ and $\phi \in \mathscr{D}(\mathbb{R}^n)$, then $(u * \delta_0) * \phi = u * (\delta_0 * \phi) = u * \phi$, hence $u * \delta_0 = u$.

**Lemma 3.7.** *For $u_1 \in \mathscr{D}'(\mathbb{R}^n), u_2 \in \mathscr{E}'(\mathbb{R}^n)$, then $D^\alpha(u_1 * u_2) = (D^\alpha u_1) * u_2 = u_1 * (D^\alpha u_2)$.*

*Proof.* For any $\phi \in \mathscr{D}(\mathbb{R}^n)$, we have $(D^\alpha(u_1 * u_2)) * \phi = (u_1 * u_2) * D^\alpha \phi = u_1 * (u_2 * D^\alpha \phi) = u_1 * ((D^\alpha u_2) * \phi) = (u_1 * D^\alpha u_2) * \phi$, so $D^\alpha(u_1 * u_2) = u_1 * D^\alpha u_2$. The other case is an easy exercise. □

Convolution is a great tool for inverting partial differential operators. Suppose $L = \sum_{|\alpha| \leq N} a_\alpha D^\alpha, a_\alpha \in \mathbb{C}$ is a linear partial differential operator with constant coefficient. A fundamental solution of $L$ is a distribution $G$ satisfying $LG = \delta_0$. Note that $G$ might or might not be unique.

**Theorem 3.8.** *If $G \in \mathcal{D}'(\mathbb{R}^n)$ is a fundamental solution of $L$ and $u_0 = \mathcal{E}'(\mathbb{R}^n)$, then $u = G * u_0$ satisfies $Lu = u_0$.*

*Proof.* $L(G * u_0) = \sum_\alpha a_\alpha D^\alpha (G * u_0) = \sum_\alpha a_\alpha ((D^\alpha G) * u_0) = (\sum_\alpha a_\alpha D^\alpha G) * u_0 = (LG) * u_0 = \delta_0 * u_0 = u_0$. □

The same argument also shows that if $f \in \mathscr{D}(\mathbb{R}^n)$, then $u = G * f \in C^\infty(\mathbb{R}^n)$ satisfies $Lu = f$.

**Example 3.12.** Take $L = -\nabla^2 = -\sum_{i=1}^3 \partial^2/\partial x_i^2$ which is a differential operator on $\mathbb{R}^3$. Consider $g : \mathbb{R}^3 \to \mathbb{R}, 0 \neq x \mapsto 1/(4\pi|x|), 0 \mapsto 0$. Then $G = T_g$ is a fundamental solution for $L$. So if $f \in C_c^\infty(\mathbb{R}^3)$ then

$$u(x) = \frac{1}{4\pi} \int_{\mathbb{R}^3} \frac{f(y)}{|x - y|} \, \mathrm{d}x$$

is a smooth solution to $-\Delta u = f$.

# 4 Fourier Transforms

## 4.1 Definition and Examples

**Definition 4.1.** Given $f \in L^1(\mathbb{R}^n)$, we define its Fourier transform $\mathcal{F}f = \hat{f} : \mathbb{R}^n \to \mathbb{C}$ to be

$$\hat{f}(\xi) = \int_{\mathbb{R}^n} f(x) e^{-i(x \cdot \xi)} \, \mathrm{d}x$$

Since $|f(x)e^{-i(x \cdot \xi)}| \leq |f(x)| \in L^1(\mathbb{R}^n)$, the integral absolutely converges for all $\xi \in \mathbb{R}^n$.

**Lemma 4.1** (Riemann-Lebesgue Lemma)**.** *Suppose* $f \in L^1(\mathbb{R}^n)$*, then* $\hat{f} \in C^0(\mathbb{R}^n)$ *and* $\|\hat{f}\|_{L^\infty} \le \|f\|_{L^1}$*. Moreover,* $\hat{f}(\xi) \to 0$ *as* $|\xi| \to 0$*.*

*Proof.* If $\xi_n \to \xi$ in $\mathbb{R}^n$, then $f(x)e^{-i(x\cdot\xi_n)} \to f(x)e^{-i(x\cdot\xi)}$ and $|f(x)e^{-i(x\cdot\xi_n)}| \le |f(x)| \in L^1(\mathbb{R}^n)$, so by dominated convergence $\hat{f}(\xi_n) \to \hat{f}(\xi)$, hence $\hat{f}$ is continuous. Also, for any $\xi \in \mathbb{R}^n$, we have

$$|\hat{f}(\xi)| \le \int_{\mathbb{R}^n} |f(x)e^{-i(x\cdot\xi)}|\,\mathrm{d}x = \int_{\mathbb{R}^n} |f(x)|\,\mathrm{d}x = \|f\|_{L^1}$$

Given $f \in L^1(\mathbb{R}^n)$ and $\epsilon > 0$, we can find $f_\epsilon \in C_c^\infty(\mathbb{R}^n)$ with $\|f - f_\epsilon\|_{L^1} < \epsilon$. Integration by parts gives

$$\begin{aligned}\hat{f}_\epsilon(\xi) &= \int_{\mathbb{R}^n} f_\epsilon(x)e^{-i(x\cdot\xi)}\,\mathrm{d}x = \int_{\mathbb{R}^n} f_\epsilon(x)\,\mathrm{div}_x\left(\frac{\xi}{-i|\xi|^2}e^{-i(x\cdot\xi)}\right)\mathrm{d}x \\ &= \int_{\mathbb{R}^n} \frac{\xi\cdot Df_\epsilon(x)}{-i|\xi|^2}e^{-i(x\cdot\xi)}\,\mathrm{d}x\end{aligned}$$

Thus $|\hat{f}_\epsilon(x)| \le |\xi|^{-1}\|Df_\epsilon\|_{L^1} < \epsilon$ for large enough $|\xi|$. We then have $|\hat{f}(\xi)| < 2\epsilon$ for such $\xi$. ☐

**Example 4.1.** We'll take $n = 1$.
1. $f = 1_{(-1,1)}$ has $\hat{f}(\xi) = 2\xi^{-1}\sin\xi$ for $\xi \ne 0$ and $\hat{f}(0) = 1$.
2. $f(x) = e^{-\operatorname{sgn}(x)x}$ has $\hat{f}(\xi) = 2/(1+\xi^2)$.
3. $f(x) = 1/(1+x^2)$ has $\hat{f}(x) = \pi e^{-\operatorname{sgn}(x)x}$.
4. $f(x) = e^{-x^2/2}$ has $\hat{f}(\xi) = \sqrt{2\pi}e^{-\xi^2/2}$.

As you've probably seen before, Fourier transforms have some properties. We introduce $e_y(x) = e^{i(x\cdot y)}$.

**Lemma 4.2.** *(i) Suppose* $f \in L^1(\mathbb{R}^n)$*,* $x \in \mathbb{R}^n$ *and* $\lambda > 0$ *and let* $f_\lambda(x) = \lambda^{-n}f(\lambda^{-1}x)$*. Then* $\hat{f}_\lambda(\xi) = \hat{f}(\lambda\xi), \widehat{e_y f}(\xi) = \tau_y\hat{f}(\xi), \widehat{\tau_y f}(\xi) = e_{-y}\hat{f}(\xi)$*.*
*(ii) Suppose* $f, g \in L^1(\mathbb{R}^n)$*, then* $f * g \in L^1(\mathbb{R}^n)$ *and* $\widehat{f * g} = \hat{f}\hat{g}$*.*

**Theorem 4.3.** *(i) Suppose* $f \in C^1(\mathbb{R}^n)$ *and* $f, D_j f \in L^1(\mathbb{R}^n)$ *for all* $j$*, then* $\widehat{D_j f}(\xi) = i\xi_j\hat{f}(\xi)$*.*
*(ii) Suppose* $(1+|x|)f \in L^1(\mathbb{R}^n)$*, then* $\hat{f} \in C^1(\mathbb{R}^n)$ *and* $D_j\hat{f}(\xi) = -i\widehat{x_j f}(\xi)$*.*

*Proof.* (i) For $\epsilon > 0$, we can find $f_\epsilon \in C_c^\infty(\mathbb{R}^n)$ such that

$$\|f_\epsilon - f\|_{L^1} + \sum_j \|D_i f - D_j f_\epsilon\|_{L^1} < \epsilon$$

which can be done by a more precise mollification (e.g. choose a smooth cutoff function). Then

$$\widehat{D_j f_\epsilon}(\xi) = \int_{\mathbb{R}^n} e^{-i(x\cdot\xi)}D_j f_\epsilon(x)\,\mathrm{d}x = \int_{\mathbb{R}^n} i\xi_j e^{-i(x\cdot\xi)}f_\epsilon(x)\,\mathrm{d}x = i\xi_j\hat{f}_\epsilon(\xi)$$

and we have

$$\begin{aligned}|\widehat{D_j f}(\xi) - i\xi_j\hat{f}(\xi)| &\le |\widehat{D_j f}(\xi) - \widehat{D_j f_\epsilon}(\xi)| + |i\xi_j\hat{f}_\epsilon(\xi) - i\xi_j\hat{f}(\xi)| \\ &\le \|D_j f - D_j f_\epsilon\|_{L^1} + |\xi|\|f - f_\epsilon\|_{L^1} < (1+|\xi|)\epsilon\end{aligned}$$

(ii) By assumption $x_j f \in L^1(\mathbb{R}^n)$ for all $j$, so $-i\widehat{x_j f} \in C^0(\mathbb{R}^n)$. We have

$$\begin{aligned}\frac{\hat{f}(\xi + he_j) - \hat{f}(\xi)}{h} &= \frac{1}{h}\int_{\mathbb{R}^n} f(x)e^{-i(x\cdot(\xi+he_j))}\,\mathrm{d}x - \frac{1}{h}\int_{\mathbb{R}^n} f(x)e^{-i(x\cdot\xi)}\,\mathrm{d}x \\ &= \int_{\mathbb{R}^n} f(x)e^{-i(x\cdot\xi)}\frac{e^{-ix_jh}-1}{h}\,\mathrm{d}x\end{aligned}$$

which gives the result by dominated convergence (noting $|e^{i\theta}-1| = 2|\sin(\theta/2)| \leq |\theta|$). □

**Corollary 4.4.** *The Fourier transform maps $\mathscr{S}(\mathbb{R}^n)$ to itself continuously.*

*Proof.* First note that for any $f \in \mathscr{S}(\mathbb{R}^n)$, we have the estimate

$$\|f\|_{L^1} = \int_{\mathbb{R}^n} |f|(1+|x|)^{n+1}\frac{1}{(1+|x|)^{n+1}}\,\mathrm{d}x \leq C\sup_{x\in\mathbb{R}^n}\left(|f(x)|(1+|x|)^{n+1}\right)$$

where

$$C = \int_{\mathbb{R}^n}\frac{1}{(1+|x|)^{n+1}}\,\mathrm{d}x < \infty$$

In particular, if $f \in \mathscr{S}(\mathbb{R}^n)$, then $D^\alpha(x^\beta f(x)) \in L^1(\mathbb{R}^n)$ for all multi-indices $\alpha, \beta$. Applying the preceding theorem iteratively shows that $|\widehat{D^\alpha(x^\beta f)}(\xi)| = |\xi^\alpha D^\beta \hat{f}(\xi)|$, so

$$\sup_\xi |\xi^\alpha D^\beta \hat{f}(\xi)| \leq \|D^\alpha(x^\beta f)\|_{L^1} \leq C_{\alpha,\beta}\sup_{x\in\mathbb{R}^n, |\delta|\leq|\alpha|}(1+|x|)^{|\beta|+n+1}|D^\delta f(x)|$$

Thus $\hat{f} \in \mathscr{S}(\mathbb{R}^n)$. Now if $f_j \to 0$ in $\mathscr{S}(\mathbb{R}^n)$, then the same estimate shows that $\sup_\xi |\xi^\alpha D^\beta \hat{f}_j(\xi)| \to 0$, which means that $\hat{f}_j \to 0$ in $\mathscr{S}(\mathbb{R}^n)$. □

**Theorem 4.5** (Fourier Inversion Theorem)**.** *Suppose $f, \hat{f} \in L^1(\mathbb{R}^n)$, then*

$$f(x) = \frac{1}{(2\pi)^n}\int_{\mathbb{R}^n}\hat{f}(\xi)e^{i(x\cdot\xi)}\,\mathrm{d}\xi$$

*for almost every $x$.*

In particular, $\mathcal{F}^2 f = (2\pi)^n \check{f}$ and $\mathcal{F}^4 f = (2\pi)^{2n} f$. With Lemma 4.1, we know that if $f$ is continuous then the inversion holds for every $x$. It also follows that $\mathcal{F} : \mathscr{S}(\mathbb{R}^n) \to \mathscr{S}(\mathbb{R}^n)$ is an isomorphism of topological vector spaces.

## 4.2 Extensions of Fourier Transforms

We want to extend our definition of Fourier transforms to some larger spaces. Let's start with trying to define Fourier transforms on $L^2(\mathbb{R}^n)$.

**Theorem 4.6** (Fourier-Plancherel/Parseval)**.** *Suppose $f, g \in L^1(\mathbb{R}^n) \cap L^2(\mathbb{R}^n)$, then $\hat{f}, \hat{g} \in L^2(\mathbb{R}^n)$ and $(f,g)_{L^2} = (2\pi)^{-n}(\hat{f},\hat{g})_{L^2}$ where $(,)_{L^2}$ is the inner product on $L^2(\mathbb{R}^n)$.*

*Proof.* Omitted. □

Note that if $f \in L^1(\mathbb{R}^n) \cap L^2(\mathbb{R}^n)$, then $\|\hat{f}\|_{L^2} = (2\pi)^{n/2}\|f\|_{L^2}$, so $\mathcal{F} : L^1(\mathbb{R}^n) \cap L^2(\mathbb{R}^n) \to L^2(\mathbb{R}^n)$ is an isometry (up to a nonzero constant). $\mathcal{F}$ then extends uniquely to a linear map $\bar{\mathcal{F}}$ from the closure of $L^1(\mathbb{R}^n) \cap L^2(\mathbb{R}^n)$ in $L^2(\mathbb{R}^n)$.

**Definition 4.2.** $\bar{\mathcal{F}}$ is known as the Fourier-Plancherel transform.

How would we calculate $\bar{\mathcal{F}}$? Suppose $f \in L^2(\mathbb{R}^n)$, then $f_R = f1_{B_R(0)} \in L^1(\mathbb{R}^n) \cap L^2(\mathbb{R}^n)$, then $f_R \to f$ as $R \to \infty$ in $L^2$ and hence $\hat{f}_R \to \bar{\mathcal{F}}(f)$ in $L^2$ as $R \to \infty$. So $\hat{f} = \bar{\mathcal{F}}(f)$ is the $L^2$ limit of

$$\xi \mapsto \int_{B_R(0)} f(x)e^{-i(x\cdot\xi)}\,\mathrm{d}x$$

as $R \to \infty$.
Since $\bar{\mathcal{F}}$ is the extension of $\mathcal{F}$ anyways, we typically abuse the notation and write $\mathcal{F}$ in place of $\bar{\mathcal{F}}$.
Now let's try to make sense of Fourier transforms on $\mathscr{S}'(\mathbb{R}^n)$. Suppose $f \in L^1(\mathbb{R}^n)$ and $\phi \in \mathscr{S}(\mathbb{R}^n)$, we have

$$\begin{aligned} T_{\hat{f}}[\phi] &= \int_{\mathbb{R}^n} \hat{f}(x)\phi(x)\,\mathrm{d}x = \int_{\mathbb{R}^n}\int_{\mathbb{R}^n} \phi(x)f(y)e^{-i(y\cdot x)}\,\mathrm{d}y\,\mathrm{d}x \\ &= \int_{\mathbb{R}^n}\int_{\mathbb{R}^n} \phi(x)f(y)e^{-i(y\cdot x)}\,\mathrm{d}x\,\mathrm{d}y = \int_{\mathbb{R}^n} f(y)\hat{\phi}(y)\,\mathrm{d}y = T_f[\hat{\phi}] \end{aligned}$$

by Fubini's theorem. Note that $\phi \in \mathscr{S}(\mathbb{R}^n) \implies \hat{\phi} \in \mathscr{S}(\mathbb{R}^n)$, we are inspired to produce the following definition.

**Definition 4.3.** For $u \in \mathscr{S}'(\mathbb{R}^n)$, we define its Fourier transform to be $\hat{u} \in \mathscr{S}'(\mathbb{R}^n)$ given by $\hat{u}[\phi] = u[\hat{\phi}]$.

$u$ is well-defined by Corollary 4.4. Indeed, it's quite problematic if one do this for $u \in \mathscr{D}(\mathbb{R}^n)$ as $\mathcal{F}$ doesn't stabilise $\mathscr{D}(\mathbb{R}^n)$.

**Example 4.2.** Fix some $\xi \in \mathbb{R}^n$ and $u = \delta_\xi$. Let's compute $\hat{u}$. We have

$$\hat{u}[\phi] = u[\hat{\phi}] = \hat{\phi}(\xi) = \int_{\mathbb{R}^n} \phi(x)e^{-i(x\cdot\xi)}\,\mathrm{d}x = T_{e_{-\xi}}[\phi]$$

So we might write $\hat{\delta}_\xi = T_{e_{-\xi}}$, which corresponds to what you've seen in (quote) "Methods or something".
Correspondingly, for $x \in \mathbb{R}^n$, the Fourier transform of $T_{e_x}$ has

$$\hat{T}_{e_x}[\phi] = T_{e_x}[\hat{\phi}] = \int_{\mathbb{R}^n} e^{i(x\cdot y)}\hat{\phi}(y)\,\mathrm{d}y = (2\pi)^n\phi(x) = (2\pi)^n\delta_x[\phi]$$

by Theorem 4.5. So $\hat{T}_{e_x} = (2\pi)^n\delta_x$.

For a multi-index $\alpha$, we write $X^\alpha : \mathbb{R}^n \to \mathbb{C}, x \mapsto x^\alpha = x_1^{\alpha_1}\cdots x_n^{\alpha_n}$. For a distribution $u$, we write $\tau_x u[\phi] = u[\tau_{-x}\phi], \check{u}[\phi] = u[\check{\phi}]$.

**Lemma 4.7.** *Suppose $u \in \mathscr{S}'(\mathbb{R}^n)$, then $\widehat{e_x u} = \tau_x\hat{u}, \widehat{\tau_x u} = e_{-x}\hat{u}, \widehat{D^\alpha u} = i^{|\alpha|}X^\alpha\hat{u}, D^\alpha\hat{u} = (-i)^{|\alpha|}\widehat{X^\alpha u}$. Moreover, $\hat{\hat{u}} = (2\pi)^n\check{u}$.*

*Proof.* Pick $\phi \in \mathscr{S}(\mathbb{R}^n)$, then

$$\widehat{e_x u}[\phi] = e_x u[\hat{\phi}] = u[e_x \hat{\phi}] = u[\widehat{\tau_{-x}\phi}] = \hat{u}[\tau_{-x}\phi] = \tau_x \hat{u}[\phi]$$

Thus $\widehat{e_x u} = \tau_x \hat{u}$. As for the derivative business,

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

The rest are mostly similar, and we have $\hat{\hat{\phi}} = \hat{u}[\hat{\phi}] = u[\hat{\hat{\phi}}] = u[(2\pi)^n \check{\phi}] = (2\pi)^n \check{u}[\phi]$. ☐

**Lemma 4.8.** *$\mathcal{F} : \mathscr{S}'(\mathbb{R}^n) \to \mathscr{S}'(\mathbb{R}^n)$ is a linear homeomorphism, with $\mathscr{S}'(\mathbb{R}^n)$ given the weak-* topology.*

*Proof.* Linearity is straightforward. We'll show sequential continuity, which is what we're gonna use anyways although it is weaker than we claimed.
Suppose $u_j \rightharpoonup^* u$, then $\hat{u}_j[\phi] = u_j[\hat{\phi}] \to u[\hat{\phi}] = \hat{u}[\phi]$ for all $\phi \in \mathscr{S}(\mathbb{R}^n)$, which means that $\mathcal{F}u_j \rightharpoonup^* \mathcal{F}u$. As $\mathcal{F}^4 = (2\pi)^n \operatorname{id}$, we get a sequentially continuous inverse. ☐

## 4.3 Periodic Distributions

Recall that if $f \in L^2([0,1])$, then we can write

$$f(x) = \sum_{n\in\mathbb{Z}} f_n e^{-2\pi i n x}, f_n = \int_0^1 e^{2\pi i n x} f(x)\,\mathrm{d}x$$

with the sum converging in $L^2$. We want to find an analogue of this in the realm of distributions.

**Definition 4.4.** A distribution $u \in \mathscr{D}'(\mathbb{R}^n)$ is periodic if $\tau_y u = u$ for all $y \in \mathbb{Z}^n$.

**Example 4.3.** 1. If $f \in L^1_{\mathrm{loc}}(\mathbb{R}^n)$ is periodic, i.e. has $\forall y \in \mathbb{Z}^n, f(x+y) = f(x)$, then $T_f$ is periodic.
2. Suppose $v \in \mathscr{E}'(\mathbb{R}^n)$, then $u = \sum_{y\in\mathbb{Z}^n} \tau_y v$ is in $\mathscr{D}'(\mathbb{R}^n)$ (indeed $u[\phi]$ is always a finite sum) and is periodic.

If $f \in C^\infty(\mathbb{R}^n)$ is periodic, we often just need to restrict our attention to a fundamental cell of $\mathbb{Z}^n$. In particular, we are interested in averages over $q = \{x \in \mathbb{R}^n : -1/2 \le x_i < 1/2, i = 1, \dots, n\}$, e.g.

$$M[f] = \int_q f(x)\,\mathrm{d}x$$

We sort of want to say $M[f] = T_f[1_q]$ and generalise everything to distributions, which sadly doesn't work since $1_q$ isn't a test function.

**Lemma 4.9.** *Let $Q = \{x \in \mathbb{R}^n : -1 \le x_i < 1\}$. Then there exists $\psi \in C^\infty(\mathbb{R}^n)$ such that:*
*1.* $\psi \ge 0$.
*2.* $\operatorname{supp}\psi \subset Q^\circ$.
*3.* $\sum_{y\in\mathbb{Z}^n} \tau_y \psi = 1$.
*Such $\psi$ is known as a periodic partition of unity. Suppose $\psi, \psi'$ are two such functions, then $u[\psi] = u[\psi']$ whenever $u \in \mathscr{D}'(\mathbb{R}^n)$ is periodic.*

So we'll be able to define $M[u] = u[\psi]$ for any periodic partition of unity $\psi$.

*Proof.* Start by picking any $\psi_0 \in C^\infty(\mathbb{R}^n)$ supported inside $Q^\circ$ with $\psi_0(x) = 1$ for $x \in q$ and $\psi_0 \geq 0$. Let $s(x) = \sum_{y\in\mathbb{Z}} \psi_0(x-y)$. This sum is locally finite, so $s \in C^\infty(\mathbb{R}^n)$. For each $x \in \mathbb{R}^n$, there is some $y \in \mathbb{Z}^n$ with $x - y \in q$, so $s \geq 1$. $s$ is also periodic by construction. So we can just use $\psi(x) = \psi_0(x)/s(x)$.
Suppose now that $u$ is periodic and $\psi, \psi'$ are periodic partitions of unity, then we want to show that $u[\psi] = u[\psi']$. Indeed,

$$u[\psi] = u\left[\psi \sum_{y\in\mathbb{Z}} \tau_y\psi'\right] = \sum_{y\in\mathbb{Z}^n} u[\psi\tau_y\psi'] = \sum_{y\in\mathbb{Z}} [\tau_{-y}u](\psi'\tau_{-y}\psi) = \sum_{y\in\mathbb{Z}} u(\psi'\tau_{-y}\psi)$$
$$= \sum_{y\in\mathbb{Z}} u(\psi'\tau_y\psi) = u\left[\psi' \sum_{y\in\mathbb{Z}} \tau_y\psi\right] = u[\psi']$$

as desired. □

If $u = T_f$ for some periodic $f \in L^1_{\text{loc}}(\mathbb{R}^n)$, then by choosing a sequence of bounded periodic partitions of unity $\psi_n$ such that $\psi_n \to 1_q$ pointwise, we conclude that

$$M[T_f] = \int_q f(x)\,\mathrm{d}x$$

**Lemma 4.10.** *Suppose $v \in \mathscr{E}'(\mathbb{R}^n)$ then $u = \sum_{y\in\mathbb{Z}^n} \tau_y v$ converges weakly-* in $\mathscr{S}'(\mathbb{R}^n)$ (the ordering of the sum doesn't matter). Conversely, if $u$ is any periodic distribution, then there exists $v \in \mathscr{E}'(\mathbb{R}^n)$ with $u = \sum_{y\in\mathbb{Z}^n} \tau_y v$. In particular, every periodic distribution is tempered.*

*Proof.* Let $K$ be a compact set with $\sigma[\phi] = 0$ for all $\phi$ with $\operatorname{supp}\phi \subset \mathbb{R}^n \setminus K$. We know that there is $k \in \mathbb{N}$ and $C > 0$ such that $|\sigma[\phi]| \leq C\sup_{x\in K, |\alpha|\leq k} |D^\alpha\phi(x)|$ for all $\phi \in \mathscr{E}(\mathbb{R}^n)$. Suppose $\phi \in \mathscr{S}(\mathbb{R}^n) \subset \mathscr{E}(\mathbb{R}^n)$. Then $|\tau_y v[\phi]| = |v[\tau_{-y}\phi]| \leq C\sup_{x\in K,|\alpha|\leq k} |D^\alpha\phi(x+y)|$. As $K$ is bounded, $K \subset B_R(0)$ for some $R > 0$. For $x \in K$, we have $1+|y| = 1+|x+y-x| \leq 1+R+|x+y| \leq (1+R)(1+|x+y|)$, which means that $1 \leq (1+R)(1+|x+y|)/(1+|y|)$ for all $x \in K$. So for $m \geq 1$, we have the estimate

$$|\tau_y v[\phi]| \leq C\frac{(1+R)^m}{(1+|y|)^m} \sup_{x\in K,|\alpha|\leq k} |(1+|x+y|)^n D^\alpha\phi(x+y)|$$
$$\leq C\frac{(1+R)^m}{(1+|y|)^m} \sup_{x'\in\mathbb{R}^n,|\alpha|\leq k} |(1+|x'|)^n D^\alpha\phi(x')|$$

Taking $m \geq n+1$ gives $\sigma_y\tau[\phi] \leq \tilde{C}/(1+|y|^{n+1})$ for some constant $\tilde{C}$ since $\sum_{y\in\mathbb{Z}^n}(1+|y|)^{n+1} < \infty$. So we conclude that $u$ converges weakly-* in $\mathscr{S}'(\mathbb{R}^n)$.
Conversely, suppose $u \in \mathscr{D}'(\mathbb{R}^n)$ is a periodic distribution, we take $\psi$ to be a periodic partition of unity. For any $\phi \in \mathscr{D}(\mathbb{R}^n)$, we have

$$u[\phi] = \left(\sum_{y\in\mathbb{Z}^n} \tau_y\psi\right) u[\phi] = \sum_{y\in\mathbb{Z}^n} u\left[\phi\tau_y\psi\right]$$

Note that $u\left[\phi\tau_y\psi\right] = \tau_y u[\phi\tau_y\psi] = u[\psi\tau_{-y}\phi] = (\psi u)[\tau_{-y}\phi] = \tau_y(\psi u)[\phi]$. So we take $v = \psi u$ which shall give $u[\phi] = \sum_{y\in\mathbb{Z}^n} \tau_y v[\phi]$. $v$ has compact support since if $\operatorname{supp}\phi \cap \operatorname{supp}\psi = \varnothing$ then $v[\phi] = u[\psi\phi] = 0$. So $v$ extends uniquely to an element of $\mathscr{E}'(\mathbb{R}^n)$. $\square$

Thus every periodic distribution has a Fourier transform. In order to characterise the Fourier transform of a periodic distribution, we first need a technical lemma.

**Lemma 4.11.** *Suppose $u \in \mathscr{S}'(\mathbb{R}^n)$ satisfies $(e_{-y'} - 1)u = 0$ for all $y' \in \mathbb{Z}^n$. Then*

$$u = \sum_{y \in \mathbb{Z}^n} c_y \delta_{2\pi y}$$

*for some $c_y \in \mathbb{C}$ satisfying $|c_y| \leq C(1 + |y|)^N$ for some $C > 0, N \in \mathbb{N}$, and the sum converges in $\mathscr{S}'(\mathbb{R}^n)$.*

*Proof.* Let $\Lambda^* = \{2\pi y : y \in \mathbb{Z}^n\}$. Suppose $\phi \in \mathscr{D}(\mathbb{R}^n)$ has $(\operatorname{supp}\phi) \cap \Lambda^* = \varnothing$. Then for $y' \in \mathbb{Z}^n$, $(e_{-y'} - 1)$ is nonzero on $\operatorname{supp}\phi$, so $(e_{-y'} - 1)\phi \in \mathscr{D}'(\mathbb{R}^n) \subset \mathscr{S}(\mathbb{R}^n)$, hence $0 = (e_{-y'} - 1)u[(e_{-y'} - 1)^{-1}\phi] = u[\phi]$. Therefore $\operatorname{supp} u \subset \Lambda^*$.
Pick a periodic partition of unity $\psi$ and let $\tilde{\psi}(x) = \psi(x/(2\pi))$. So

$$\sum_{y \in \mathbb{Z}^n} \tau_{2\pi y}\tilde{\psi}(x) = 1, \tilde{\psi} \geq 0, \operatorname{supp}\tilde{\psi} \subset \{|x_i| < 2\pi\}$$

Consider $v_y = (\tau_{2\pi y}\tilde{\psi})u$ which is supported at $2\pi y$. We have $\sum_{y \in \mathbb{Z}^n} v_y = u$ and $(e_{-y'} - 1)v_y = 0$ for all $y' \in \mathbb{Z}^n$. Set $y' = l_j, j = 1, \ldots, n$ where $\{l_j\}_j$ is the standard basis on $\mathbb{R}^n$. Then $(e^{-i(x_j - 2\pi y_j)} - 1)v_y = 0$. Note that $e^{-i(x_j - 2\pi y_j)} - 1 = (x_j - 2\pi y_j)\kappa(x_j)$ for some smooth $\kappa(x_j)$ nonzero near $y_j$. Therefore $(x_j - 2\pi y_j)v_y = 0$.
Suppose we are given $\phi \in \mathscr{S}(\mathbb{R}^n)$, then $\phi(x) = \phi(2\pi y) + \sum_j (x_j - 2\pi y_j)\phi_j(x)$ for some $\phi_j \in C^\infty(\mathbb{R}^n)$. Now $v_y$ has compact support, so extends to act on $\mathscr{E}(\mathbb{R}^n) = C^\infty(\mathbb{R}^n)$. This means that

$$v_y[\phi] = v_y[\phi(2\pi y)] + \sum_{j=1}^{n} (x_j - 2\pi y_j)v_y[\phi_j] = v_y[\phi(2\pi y)] = \phi(2\pi y)v_y[1]$$

So $(\tau_{2\pi y}\tilde{\psi})u[\phi] = (\tau_{2\pi y}\tilde{\psi})u[\phi(2\pi y)] = u[\tau_{2\pi y}\tilde{\psi}\phi(2\pi y)] = u[\tau_{2\pi y}\tilde{\psi}]\delta_{2\pi y}[\phi]$, which means that $v_y = u[\tau_{2\pi y}\tilde{\psi}]\delta_{2\pi y}$, so we conclude by taking $c_y = u[\tau_{2\pi y}\tilde{\psi}]$.
To get the desired bound, note that by example sheet there are $N, k \in \mathbb{N}$ and $C > 0$ such that

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

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

$$\begin{aligned}
|c_y| &\leq C \sup_{x \in \mathbb{R}^n, |\alpha| \leq k} |(1 + |x|)^N D^\alpha \tilde{\psi}(x - 2\pi y)| \\
&\leq C \sup_{x \in \mathbb{R}^n, |\alpha| \leq k} |(1 + |x + 2\pi y|)^N D^\alpha \tilde{\psi}(x)| \\
&\leq C' \sup_{x \in \mathbb{R}^n, |\alpha| \leq k} |(1 + |x|)^N D^\alpha \tilde{\psi}(x)|(1 + |y|)^N
\end{aligned}$$

Therefore $|c_y| \leq C^N (1 + |y|)^N$. $\square$

**Theorem 4.12.** *Suppose $u \in \mathscr{D}'(\mathbb{R}^n)$ is periodic, then there exists $c_y \in \mathbb{C}$ such that we have an $\mathscr{S}'(\mathbb{R}^n)$-convergence*

$$u = \sum_{y \in \mathbb{Z}^n} c_y T_{e_{2\pi y}}$$

*where $c_y = M[e_{-2\pi y} u]$ satisfies the bound $|c_y| \leq C(1+|y|)^N$ for some $C > 0, N \in \mathbb{N}$.*

We say $(c_y)_{y \in \mathbb{Z}^n}$ are the Fourier coefficients for $u$.

*Proof.* As $u$ is periodic and periodic distributions are tempered, for $y' \in \mathbb{Z}^n$ we have $(e_{-y'} - 1)\hat{u} = 0$ since $\tau_{y'} u = u$. By the preceding lemma,

$$\hat{u} = (2\pi)^n \sum_{y \in \mathbb{Z}^n} c_y \delta_{2\pi y}, |c_y| \leq C(1+|y|)^N$$

in $\mathscr{S}'(\mathbb{R}^n)$. Theorem 4.5 then gives

$$u = \sum_{y \in \mathbb{Z}^n} c_y T_{e_{2\pi y}}$$

Also, since $e_{2\pi y} \in L^1_{\text{loc}}$, we can write $M[e_{-2\pi y'} T_{e_{2\pi y}}] = \delta_{yy'}$. The map $u \mapsto M(e_{2\pi y} u)$ is a continuous map $\mathscr{S}'(\mathbb{R}^n) \to \mathbb{C}$, so

$$M[e_{-2\pi y} u] = \sum_{y' \in \mathbb{Z}^n} c_{y'} M[e_{-2\pi y} T_{e_{2\pi y'}}] = c_y$$

as desired. □

We sometimes abuse notation and write $u(x) = \sum_{y \in \mathbb{Z}^n} c_y e^{2\pi i (y \cdot x)}$.

**Example 4.4.** Consider the periodic distribution $u = \sum_{y \in \mathbb{Z}^n} \delta_y$. Then $c_y = M[e_{-2\pi y} u] = u[\psi e_{2\pi y}]$ for a periodic partition of unity $\psi$. So

$$c_y = \sum_{y \in \mathbb{Z}^n} \psi(y') e^{-2\pi i (y \cdot y')} = \sum_{y \in \mathbb{Z}^n} \psi(y') = 1$$

The preceding theorem then gives Poisson's formula:

$$\sum_{y \in \mathbb{Z}^n} \delta_y = \sum_{y \in \mathbb{Z}^n} T_{e_{2\pi y}}$$

or, by abuse of notation, $\sum_{y \in \mathbb{Z}^n} \delta(x-y) = \sum_{y \in \mathbb{Z}^n} e^{2\pi i (y \cdot x)}$.

# 5 Sobolev Spaces

## 5.1 Definitions

Sobolev spaces give contexts where we make sense of differentiation in the integral sense. They consists of $L^p$ functions whose distributional derivatives are also in $L^p$.

**Definition 5.1.** Suppose $\Omega \subset \mathbb{R}^n$ is open and let $k \in \mathbb{Z}_{\geq 0}, 1 \leq p \leq \infty$. $f \in L^p(\Omega)$ belongs to the Sobolev space $W^{k,p}(\Omega)$ if for any multi-index $\alpha$ with $|\alpha| \leq k$, there exists $f^\alpha \in L^p(\Omega)$ with $D^\alpha T_f = T_{f^\alpha}$ in $D'(\Omega)$. We say $f^\alpha$ is the $\alpha^{th}$ weak derivative of $f$.

From now on, we abuse the notation and drop the notational distinction between $f \in L^1_{\text{loc}}$ and $T_f \in \mathscr{D}'$. In particular, we might sometimes write $D^\alpha f \in L^1(\mathbb{R})$ to mean $f^\alpha \in L^1(\mathbb{R}^n)$ with $D^\alpha T_f = T_{f^\alpha}$.

**Definition 5.2.** $W^{k,p}(\Omega)$ is a Banach space equipped with the norm

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

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

The condition $D^\alpha T_f = T_{f^\alpha}$ translates to the statement that for any $\phi \in \mathscr{D}(\Omega)$,

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

**Example 5.1.** 1. Let

$$f(x) = \begin{cases} -1 & \text{for } x < -1 \\ x & \text{for } -1 \leq x \leq 1 \\ 1 & \text{for } x > 1 \end{cases}$$

Then $f \in W^{1,\infty}(\mathbb{R})$ with $Df(x) = 1_{|x| \leq 1}$.
2. Let $H = 1_{\mathbb{R}_{\geq 0}}$, then $H \notin W^{1,p}$ for any $p$ since $DT_H = \delta_0 \neq T_f$ for any $f \in L^1_{\text{loc}}$.

As is the usual case, we can say very little about Sobolev spaces except in the case $p = 2, \Omega = \mathbb{R}^n$, so we'll do that. Suppose $f \in L^2(\mathbb{R}^n)$, $D^\alpha f \in L^2(\mathbb{R}^n)$ iff $\xi^\alpha \hat{f}(\xi) \in L^2(\mathbb{R}^n)$.

**Definition 5.3.** For $s \in \mathbb{R}$, we say $u \in \mathscr{S}'(\mathbb{R}^n)$ belongs to $H^s(\mathbb{R}^n)$ if $\hat{u} \in L^2_{\text{loc}}(\mathbb{R}^n)$ and

$$\int_{\mathbb{R}^n} (1 + |\xi|^2)^s |\hat{u}(\xi)|^2 \,\mathrm{d}\xi < \infty$$

Then $H^s(\mathbb{R}^n)$ is a Hilbert space with inner product

$$(u, v)_{H^s} = \int_{\mathbb{R}^n} \overline{\hat{u}(\xi)} \hat{v}(\xi) (1 + |\xi|^2)^s \,\mathrm{d}\xi$$

If $s \in \mathbb{Z}_{\geq 0}$, then $H^s(\mathbb{R}^n) = W^{s,2}(\mathbb{R}^n)$. One important property of Sobolev spaces is that for large enough $k$ (or $s$), they embed into classical differentiable spaces.

**Theorem 5.1** (Sobolev Embeddings)**.** *Suppose $f \in H^s(\mathbb{R}^n)$ for some $s > k + n/2, k \in \mathbb{Z}_{\geq 0}$, then there exists $f^* \in C^k(\mathbb{R}^n)$ with $f = f^*$ a.e..*

So it makes sense to write $H^s(\mathbb{R}^n) \subset C^k(\mathbb{R}^n)$.

*Proof.* First suppose that $f \in \mathscr{S}(\mathbb{R}^n)$, then by Theorem 4.5, we can write

$$D^\alpha f(x) = \frac{i^{|\alpha|}}{(2\pi)^n} \int_{\mathbb{R}^n} e^{i(x\cdot\xi)} \xi^\alpha \hat{f}(\xi) \, \mathrm{d}\xi$$

So by Cauchy-Schwarz

$$\begin{aligned} |D^\alpha f(x)| &= \frac{1}{(2\pi)^n} \left| \int_{\mathbb{R}^n} e^{i(x\cdot\xi)} \xi^\alpha \hat{f}(\xi) \, \mathrm{d}\xi \right| \\ &\le \frac{1}{(2\pi)^n} \left( \int_{\mathbb{R}^n} (1+|\xi|^2)^s |\hat{f}(\xi)|^2 \, \mathrm{d}\xi \right)^{1/2} \left( \int_{\mathbb{R}^n} \frac{|\xi^\alpha|^2}{(1+|\xi|^2)^s} \, \mathrm{d}\xi \right)^{1/2} \end{aligned}$$

As $|\alpha| \le k$, we have $|\xi^\alpha| \le c_k (1+|\xi|^2)^k$ for some $c_k > 0$, so

$$\int_{\mathbb{R}^n} \frac{|\xi|^\alpha}{(1+|\xi|^2)^s} \, \mathrm{d}\xi \le c_k \int_{\mathbb{R}^n} \frac{1}{(1+|\xi|^2)^{s-k}} \, \mathrm{d}\xi = c_{k,n,s}^2 < \infty$$

So we deduce that $\sup_{x\in\mathbb{R}^n, |\alpha|\le k} |D^\alpha f(x)| \le c_{n,k,s} \|f\|_{H^s}$ for all $f \in \mathscr{S}(\mathbb{R}^n)$. Now suppose $f \in H^s(\mathbb{R}^n)$, then we can approximate $f$ by a sequence $(f_n)_{n=1}^\infty$ with $f_n \in \mathscr{S}(\mathbb{R}^n), f_n \to f$ in $H^s$, and $f_n \to f$ a.e. (example sheet). In particular, $f_n$ is Cauchy in $H^s(\mathbb{R}^n)$, so by our estimate $f_n$ is Cauchy in $C^k(\mathbb{R}^n)$. So there exists $f^* \in C^k(\mathbb{R}^n)$ such that $f_n \to f^*$ uniformly. Hence $f = f^*$ a.e.. □

Why are Sobolev spaces worth considering?

**Example 5.2.** Consider the PDE $-\nabla^2 u + u = f$ on $\mathbb{R}^n$. We claim that if $f \in H^s(\mathbb{R}^n)$, then there is a unique $u \in H^{s+2}(\mathbb{R}^n)$ satisfying the PDE in $\mathscr{S}'(\mathbb{R}^n)$. To see this, we take a Fourier transform of the equation, which gives $(|\xi|^2+1)\hat{u}(\xi) = \hat{f}(\xi)$ for almost every $\xi$. This is to say that $\hat{u}(\xi) = \hat{f}(\xi)/(1+|\xi|^2)$ and we can invert this. We also have

$$\|u\|_{H^{s+2}}^2 = \int_{\mathbb{R}^n} (1+|\xi|^2)^{s+2} |\hat{u}(\xi)|^2 \, \mathrm{d}\xi = \int_{\mathbb{R}^n} (1+|\xi|^2)^s |\hat{f}(\xi)|^2 \, \mathrm{d}\xi = \|f\|_{H^s}^2$$

*Remark.* 1. $u$ is more regular than $f$, which is a property of the differential operator $(-\nabla^2+1)$ (known as elliptic regularity). We have $-\nabla^2 u + u \in L^2(\mathbb{R}^n) \implies u \in H^2(\mathbb{R}^n)$.
2. Suppose $s > n/2$, then (possibly after augmenting on a set of measure zero) $f \in C^0(\mathbb{R}^n), u \in C^2(\mathbb{R}^n)$ and the PDE $-\nabla^2 u + u = f$ holds classically.

## 5.2 Elliptic Boundary Problems

We often want to consider the restriction of $u \in H^s(\mathbb{R}^n)$ $(s > 0)$ to an $(n-1)$-dimensional hypersurface in $\mathbb{R}^n$. Since $u$ is only defined modulo changes on sets of measure zero, doing this restriction directly usually gives us nonsense. To resolve this, we show the following:

**Theorem 5.2** (Trace Theorem)**.** *Let $s > 1/2$. There exists a bounded surjective linear map $T : H^s(\mathbb{R}^n) \to H^{s-1/2}(\mathbb{R}^{n-1})$ such that $Tf = f|_{x_n=0}$ for all $f \in \mathscr{S}(\mathbb{R}^n)$.*

$Tf$ is called the trace of $f$ on $\{x_n = 0\}$.

*Proof.* Example sheet. □

By combining this with certain coordinates transformations, we can establish a similar result with $\mathbb{R}^{n-1} = \{x_n = 0\}$ replaced by suitably regular (e.g. Lipschitz regular) $(n-1)$-hypersurfaces.
Let $\Omega \subset \mathbb{R}^n$ be open and suppose $f \in C_c^\infty(\Omega)$. By extending $f$ by zero, we can realise $f$ as an element of $C_c^\infty(\mathbb{R}^n)$ and thus automatically $f \in H^1(\mathbb{R}^n)$. In other words, we can embed $C_c^\infty(\Omega) \subset H^1(\mathbb{R}^n)$. Denote by $H_0^1(\Omega)$ the closure of $C_c^\infty(\Omega)$ in $H^1(\mathbb{R}^n)$. Note that we have a particularly nice way to write the norm on $H^1$ for functions $f \in C_c^\infty(\Omega)$, i.e.

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

Then $H_0^1(\Omega)$ is a Hilbert space whose inner product has the form

$$(u,v)_{H^1} = \int_\Omega \overline{Du} \cdot Dv + \bar{u}v \,\mathrm{d}x$$

Suppose $\phi \in C_c^\infty((\Omega^c)^\circ)$, then

$$\Lambda_\phi(u) = \int_{\mathbb{R}^n} \phi u \,\mathrm{d}x = 0$$

for all $u \in C_c^\infty(\Omega)$. We know that $\Lambda_\phi : H^1(\mathbb{R}^n) \to \mathbb{C}$ is continuous, so if $u_n \to u$ in $H^1(\mathbb{R}^n)$ then $\Lambda_\phi u = 0$. This means that any $u \in H_0^1(\Omega)$ vanishes a.e. on $\Omega^c$.
For sufficiently nice $\partial\Omega$, the trace map $T : H_0^1(\Omega) \to H^{1/2}(\partial\Omega)$ is bounded and $Tu = 0$ for any $C_c^\infty(\Omega)$. So if $u \in H_0^1(\Omega)$ then $Tu = 0$.
Nothing breaks, basically.
$H_0^1(\Omega)$ can be viewed (informally) as "the set of $H^1$ functions which vanish on $\partial\Omega$ and outside $\Omega$".
Now consider the problem

$$\begin{cases} -\nabla^2 u + u = f & \text{in } \Omega \\ u = 0 & \text{on } \partial\Omega \end{cases}$$

We seek a distributional solution $u \in H_0^1(\Omega)$. $u$ satisfies the boundary condition by definition. Suppose $v \in \mathscr{D}(\Omega)$, then

$$(u,v)_{H^1} = \int_\Omega \overline{Du} \cdot Dv + \bar{u}v \,\mathrm{d}x = \int_\Omega -\bar{u}\nabla^2 v + \bar{u}v \,\mathrm{d}x = \int_\Omega \bar{f}v \,\mathrm{d}x = (f,v)_{L^2}$$

By continuity, this holds for every $v \in H_0^1(\Omega)$ provided $f \in L^2(\Omega)$. Conversely, if this holds for every such $v$, then by taking $v \in C_c^\infty(\Omega)$ and integrating by parts, we have $-\nabla^2 u + u = f$ holds in $\mathscr{D}'(\Omega)$.
Motivated by this,

**Definition 5.4.** We say $u \in H_0^1(\Omega)$ is a weak solution of the boundary value problem above (for $f \in L^2(\Omega)$) if $(u,v)_{H^1} = (f,v)_{L^2}$ for all $v \in H_0^1(\Omega)$.

Note that $\Lambda : v \mapsto (f,v)_{L^2}$ is a bounded linear map $H_0^1(\Omega) \to \mathbb{C}$ since $|(f,v)_{L^2}| \le \|f\|_{L^2}\|v\|_{L^2} \le \|f\|_{L^2}\|v\|_{H^1}$. By Riesz representation theorem, there is a unique $u \in H_0^1(\Omega)$ such that $(u,v)_{H^1} = \Lambda v = (f,v)_{L^2}$ for all $v \in H_0^1(\Omega)$. In other words,

**Lemma 5.3.** *Given $f \in L^2(\Omega)$, there exists a unique weak solution $u \in H_0^1(\Omega)$ with $\|u\|_{H^1} \leq \|f\|_{L^2}$.*

If we let $Af = u$, then $A$ is a bounded linear operator $L^2(\Omega) \to H_0^1(\Omega)$. Suppose $f, g \in L^2(\Omega), a \in \mathbb{C}$ and $u = Af, w = Ag$, then

$$(u + aw, v)_{H^1} = (u, v)_{H^1} + a(w, v)_{H^1} = (f, v)_{L^2} + a(g, v)_{L^2} = (f + ag, v)$$

for any $v \in H_0^1(\Omega)$. So $A(f+ag) = Af + aAg$, i.e. $A$ is linear. It's also bounded:

$$\|Af\|_{H^1}^2 = (Af, Ag)_{H^1} = (f, Af)_{L^2} \leq \|f\|_{L^2}\|Af\|_{L^2} \leq \|f\|\|Af\|_{H^1}$$

Thus $\|Af\|_{H^1} \leq \|f\|_{L^2}$. $A$ is in fact also Hermitian, i.e. $(Af, g)_{L^2} = (f, Ag)_{L^2}$. Indeed,

$$(f, Ag)_{L^2} = (f, w)_{L^2} = (u, w)_{H^1} = \overline{(w, u)}_{H^1} = \overline{(g, u)}_{L^2} = (u, g)_{L^2} = (Af, g)_{L^2}$$

**Definition 5.5.** For $s > 0$, we set $H_{\text{loc}}^s(\Omega) = \{u \in L_{\text{loc}}^2(\Omega) : \forall \chi \in C_c^\infty(\Omega), \chi u \in H^2(\mathbb{R}^n)\}$.

If $U$ is open and $\bar{U} \subset \Omega$. Then we can find $\chi \in C_c^\infty(\Omega)$ with $\chi = 1$ on $U$. If $s > (n/2) + k$, $u \in H_{\text{loc}}^s(\Omega)$, then $\chi u \in H^s(\mathbb{R}^n) \subset C^k(\mathbb{R}^n)$. But $\chi u = u$ on $U$, so $u \in C^k(\Omega)$ by varying $U$.
Returning to our elliptic boundary value problem, we fix $K \subset \Omega$ compact and $\chi_K \in C_c^\infty(\Omega)$ with $\chi_K|_K = 1$. Suppose $u$ is a weak solution, i.e.

$$\int_\Omega \overline{Du} \cdot Dv + \bar{u}v \, \mathrm{d}x = \int_\Omega \bar{f}v \, \mathrm{d}x$$

for any $v \in H_0^1(\Omega)$. Pick any $\phi \in \mathscr{S}(\mathbb{R}^n)$ and let $v(x) = \chi_K(x)\phi(x)$, then $v \in C_c^\infty(\Omega) \subset H_0^1(\Omega)$. Plugging this in, we conclude

$$\int_{\mathbb{R}^n} \bar{w}(-\nabla^2\phi + \phi) \, \mathrm{d}x = \int_{\mathbb{R}^n} \bar{g}\phi \, \mathrm{d}x$$

where $w = \chi_K u$ and $g = -2Du \cdot D\chi_K - u\nabla^2\chi_K + f\chi_K$. This tells us that $w \in H_0^1(\Omega) \subset H^1(\mathbb{R}^n)$ solves $-\nabla^2 w + w = g$ in $\mathscr{S}'(\Omega)$ with $g \in L^2(\Omega)$. We've shown previously that if $g \in H^s(\mathbb{R}^n)$, then the solution to $-\nabla^2 w + w = g$ satisfies $w \in H^{s+2}(\mathbb{R}^n)$. In this instance, we obtain $w \in H^2(\mathbb{R}^n)$. Let $\psi \in C_c^\infty(\Omega)$ be arbitrary, then by taking $K = \operatorname{supp}\phi$ we deduce that $\psi u = \psi\chi_K u \in H^2(\mathbb{R}^n)$, thus $u \in H_{\text{loc}}^2(\Omega)$.
Suppose $f \in H_{\text{loc}}^1(\Omega) \cap L^2(\Omega)$. The expression of $g$ reveals that $g \in H^1(\mathbb{R}^n)$, so the same argument shows that $u \in H_{\text{loc}}^3(\Omega)$. Iterating, we see that if $f \in L^2(\Omega) \cap H_{\text{loc}}^k(\Omega)$ then $u \in H_0^1(\Omega) \cap H_{\text{loc}}^{k+2}(\Omega)$. In particular, if $f \in L^2(\Omega) \cap C_c^\infty(\Omega)$ then $u \in H_0^1(\Omega) \cap C_c^\infty(\Omega)$ and $-\nabla^2 u + u = f$ holds classically.
This property of the operator $-\nabla^2 + 1$ of improving regularity of its solution is called elliptic regularity, which sadly (and expectedly) doesn't hold for every PDE. One counterexample happens when you try to solve the wave equation $u_{tt} - u_{xx} = 0$ which has solutions $u(x, t) = u_+(x + t) + u_-(x - t)$ that does not get smoothness for free.

## 5.3 Rellich-Kondrachov Theorem

We are interested in the compactness for Sobolev spaces. Suppose $\Omega \subset \mathbb{R}^n$ is open and bounded, and suppose $(u_n) \in H_0^1(\Omega)$ satisfy $\|u_n\|_{H^1} \le K$. Theorem 2.9 shows that after possibly extracting a subsequence we can assume $u_n \rightharpoonup u$ in $H_0^1(\Omega)$ for some $u \in H_0^1(\Omega), \|u\|_{H^1} \le K$ (we are in a Hilbert space so reflexivity is free). Moreover, if $w \in L^2(\Omega)$, we can estimate

$$|(w,v)_{L^2}| \le \|w\|_{L^2}\|v\|_{L^2} \le \|w\|_{L^2}\|v\|_{H^1}$$

So $v \mapsto (w,v)_{L^2}$ is a bounded linear map $H_0^1(\Omega) \to \mathbb{C}$, so $(w,u_n)_{L^2} \to (w,u)_{L^2}$, i.e. $u_n \rightharpoonup u$ in $L^2$. In fact, the convergence is strong.

**Theorem 5.4** (Rellich-Kondrachov)**.** *Suppose $\Omega \subset \mathbb{R}^n$ is open and bounded. Let $(u_n) \in H_0^1(\Omega)$ be such that $u_n \rightharpoonup u$ in $L^1$ for some $u \in H_0^1(\Omega)$ and $\|u_n\|_{H^1} \le K$, then $u_n \to u$ strongly in $L^2$.*

*Proof.* Fix $\epsilon > 0$, we have

$$\begin{aligned}\|u_n - u\|_{L^2}^2 &= \frac{1}{(2\pi)^n}\|\hat{u}_n - \hat{u}\|_{L^2}^2 \\ &= \frac{1}{(2\pi)^n}\int_{|\xi|<R} |\hat{u}_n(\xi) - \hat{u}(\xi)|^2 \,\mathrm{d}\xi + \frac{1}{(2\pi)^n}\int_{|\xi|>R} |\hat{u}_n(\xi) - \hat{u}(\xi)|^2 \,\mathrm{d}\xi\end{aligned}$$

For the second term, we have the bound

$$\begin{aligned}\int_{|\xi|>R} |\hat{u}_n(\xi) - \hat{u}(\xi)|^2 \,\mathrm{d}\xi &\le \frac{2}{R^2}\int_{|\xi|>R} (1+|\xi|^2)(|\hat{u}_n(\xi)|^2 + |\hat{u}(\xi)|^2)\,\mathrm{d}\xi \\ &\le \frac{2}{R^2}(\|u_n\|_{H^1}^2 + \|u\|_{H^1}^2) \le \frac{2K}{R^2}\end{aligned}$$

which can be made arbitrarily small for sufficiently large $R$.
As for the first term, recall that $\hat{u}(\xi) = (e_\xi, u_n)_{L^2(\Omega)}$. Since $|\Omega| < \infty, e_\xi \in L^2(\Omega)$. We deduce using $u_n \rightharpoonup_{L^2} u$ that $\hat{u}_n(\xi) \to \hat{u}(\xi)$ for any $\xi$. Further, for $|\xi| < R$, we have by Lemma 4.1 that

$$\begin{aligned}|\hat{u}_n(\xi) - \hat{u}(\xi)|^2 &\le 2(|\hat{u}_n(\xi)|^2 + |\hat{u}(\xi)|^2) \le 2(\|u_n\|_{L^1}^2 + \|u\|_{L^1}^2) \\ &\le 2|\Omega|(\|u_n\|_{L^2}^2 + \|u\|_{L^2}^2) \le 2|\Omega|(K^2 + \|u\|_{L^2}^2) \in L^1(B_R(0))\end{aligned}$$

So by dominated convergence,

$$I_n = \int_{|\xi|<R} |\hat{u}_n(\xi) - \hat{u}(\xi)|^2 \,\mathrm{d}\xi \to 0$$

as $n \to \infty$. So by picking sufficiently large $n$, $|I_n|$ can be made arbitrarily small, and thus $\|u_n - u\|_{L^2}^2$ can be made arbitrarily small. ☐

**Corollary 5.5.** *Suppose $\|u_n\|_{H^1} \le K$, then $(u_n)$ has a subsequence converging weakly in $H^1$ and strongly in $L^2$, i.e. after passing to a subsequence, there is some $u \in H_0^1(\Omega)$ with $\|u_n - u\|_{L^2} \to 0$ and $(Du_n, w)_{L^2} \to (Du, w)_{L^2}$ for all $w \in L^2(\Omega)$.*

**Corollary 5.6.** *Suppose $\Omega \subset \mathbb{R}^n$ is open and bounded and $A : L^2(\Omega) \to H_0^1(\Omega)$. We can think of $A$ as $L^2(\Omega) \to L^2(\Omega)$. If $A$ is linear and bounded, then $A : L^2(\Omega) \to L^2(\Omega)$ is compact. In particular, if $\|u_n\|_{L^2} \leq K$ is bounded, then $\|Au_n\|_{H^1} \leq C\|u_n\|_{L^2} \leq CK$, so $(Au_n)$ has a strongly convergent subsequence in $L^2$.*

For $\Omega$ open and bounded, we consider the problem

$$\begin{cases} -\nabla^2 u + Vu = f & \text{in } \Omega \\ u = 0 & \text{on } \partial\Omega \end{cases}$$

with $V : \Omega \to \mathbb{R}$ smooth and bounded. Similar to what we did before, we can reformulate the problem to get the notion of a weak solution. We say $u \in H_0^1(\Omega)$ is a weak solution to the problem if

$$\int_\Omega \overline{Du} \cdot Dv + V\bar{u}v \,\mathrm{d}x = \int_\Omega \bar{f}v \,\mathrm{d}x$$

for all $v \in H_0^1(\Omega)$. But we can't import the previous argument exactly, since the left hand side needs not be an inner product unless $V$ is nonnegative. Let's try to rewrite the equation to see if we can do anything.

$$\int_\Omega \overline{Du} \cdot Dv + \bar{u}v \,\mathrm{d}x = \int_\Omega (1 - V)\bar{u}v + \bar{f}v \,\mathrm{d}x$$

In other words, $u = A(f+(1-V)u)$, or $(I-K)u = Af$ where $Ku = A((1-V)u)$, where $A : L^2(\Omega) \to H_0^1(\Omega)$ is the map that takes $g \in L^2(\Omega)$ to the weak solution $w \in H_0^1(\Omega)$ to

$$\begin{cases} -\nabla^2 w + w = g & \text{in } \Omega \\ w = 0 & \text{on } \partial\Omega \end{cases}$$

Now $K : L^2(\Omega) \to H_0^1(\Omega)$ is bounded and linear, so $K : L^2(\Omega) \to L^2(\Omega)$ is compact by Corollary 5.6. Thus by the general theory of compact operators, either there is some $w \in L^2(\Omega)$ such that $(I - K)w = 0$ or $(I - K)u = g$ has a unique solution $u \in L^2(\Omega)$ for any $g \in L^2(\Omega)$.

In the first case, if we have $(I-K)w = 0$ then $w = A((1-V)w)$, thus $w \in H_0^1(\Omega)$. The usual regularity theory then shows that $w \in C^\infty(\Omega)$.

In the second case, we can set $g = Af$ which shows the existence of a weak solution $u \in H_0^1(\Omega)$ to the boundary value problem.

In conclusion, either there is some $w \in H_0^1 \cap C^\infty(\Omega)$ solving $-\nabla^2 w + Vw = 0$, or for any $f \in L^2(\Omega)$ there is a unique weak solution $u \in H_0^1(\Omega)$.

Let's now consider $A : L^2(\Omega) \to H_0^1(\Omega)$ in some more detail. We've shown previously that $A$ is linear and bounded. As a map $L^2(\Omega) \to L^2(\Omega)$, it is also compact and Hermitian. By the spectral theorem, we know that $\sigma(A) = \{0, \mu_1, \mu_2, \ldots\}$ for some $\mu_k \in \mathbb{R}$ which can only accumulate at 0. Furthermore, there is an orthonormal (Hilbertian) basis of $L^2(\Omega)$ consisting of eigenvectors of $A$.

Suppose $Aw = \mu w$ for $\mu \in \mathbb{R}, w \in L^2(\Omega) \setminus \{0\}$, then $w \in H_0^1(\Omega)$ and for any $v \in H_0^1(\Omega)$ we have $(w, v)_{L^2} = (Aw, v)_{H^1} = \mu(w, v)_{H^1}$. Setting $w = v$ shows that $\mu \neq 0$, hence $w \in H_0^1(\Omega)$ is a weak solution to

$$\begin{cases} -\nabla^2 w + w = w/\mu & \text{in } \Omega \\ w = 0 & \text{on } \partial\Omega \end{cases}$$

So $w \in H_0^1(\Omega) \subset H^1(\mathbb{R}^n)$. In particular, $w \in H^1_{\text{loc}}(\Omega)$. The usual regularity strikes again and leaves $w \in H^3_{\text{loc}}(\Omega), w \in H^5_{\text{loc}}(\Omega), w \in H^7_{\text{loc}}(\Omega), \ldots$, so in fact $w \in C^\infty(\Omega)$.
If we rewrite the system as

$$\begin{cases} -\nabla^2 w = \lambda w & \text{in } \Omega \\ w = 0 & \text{on } \partial\Omega \end{cases}, w = -1 + \frac{1}{\mu}$$

We can deduce

**Theorem 5.7.** *Let $\Omega \subset \mathbb{R}^n$ be open and bounded, then there exists an orthonormal (Hilbertian) basis $(w_n)_n$ for $L^2(\Omega)$ such that $w_n \in H_0^1(\Omega) \cap C^\infty(\Omega)$ satisfying $-\nabla^2 w_k = \lambda_k w_k$ in $\Omega$ where $0 \le \lambda_1 \le \lambda_2 \le \cdots$ with $\lambda_n \to \infty$.*

**Corollary 5.8.** $(\sqrt{2}\sin(n\pi x))_{n=1}^\infty$ *is an orthonormal basis for $L^2((0,1))$.*

## 5.4 The Direct Approach to Calculus of Variations

Given $u \in H_0^1(\Omega), f \in L^2(\Omega)$, consider

$$S[u] = \int_\Omega |Du|^2 + |u|^2 - \bar{f}u - f\bar{u}\,\mathrm{d}x$$

We have the estimation

$$\begin{aligned} S[u] &= \|u\|^2_{H^1} - 2\operatorname{Re}(f,u)_{L^2} \ge \|u\|^2_{H^1} - 2\|u\|_{L^2}\|f\|_{L^2} \\ &\ge \|u\|^2_{H^1} - \frac{1}{2}\|u\|^2_{L^2} - 2\|f\|^2_{L^2} \ge \frac{1}{2}\|u\|^2_{H^1} - 2\|f\|^2_{L^2} \end{aligned}$$

Thus $B = \{S[u] : u \in H_0^1(\Omega) : u \in H_0^1(\Omega)\}$ is bounded below. Let $\sigma = \inf B > -\infty$. We claim that this infimum is attainable.
Let $u_n \in H_0^1(\Omega)$ be such that $S[u_n] \to \sigma$. $(S[u_n])_{n=1}^\infty$ is convergent hence bounded, so $\|u_n\|^2_{H^1} \le 2S[u_n] + 4\|f\|^2_{L^2} \le K$ for some constant $K$ and therefore by Theorem 2.9 we can assume $u_n \rightharpoonup w$ in $H^1$ after passing to a subsequence. By example sheet, we have $\|w\|_{H^1} \le \liminf_{n\to\infty} \|u_n\|_{H^1}$. Since $u_n \rightharpoonup w$ in $H^1$, $\operatorname{Re}(f,w) = \lim_{n\to\infty} \operatorname{Re}(f,u_n) = \liminf_{n\to\infty} \operatorname{Re}(f,u_n)$. So

$$\begin{aligned} S[w] = \|w\|^2_{H^1} - 2\operatorname{Re}(f,w)_{L^2} &\le \liminf_{n\to\infty}(\|u_n\|^2_{H^1} - 2\operatorname{Re}(f,u_n)) \\ &\le \liminf_{n\to\infty} S[u_n] = \sigma \end{aligned}$$

But $S[w] \ge \sigma$ by the definition of $\sigma$, so $S[w] = \sigma$ and $w$ is the desired minimiser.
By considering $S[w + tv]$ for $t \in \mathbb{R}, v \in H_0^1(\Omega)$, one can show that $w$ is a weak solution of

$$\begin{cases} -\nabla^2 w + w = f & \text{in } \Omega \\ w = 0 & \text{on } \partial\Omega \end{cases}$$

which in particular shows that this minimiser is unique.
How would one generalise this idea to a wider class of functionals?

**Definition 5.6.** Let $X$ be a topological space (usually taken to be a Banach space or subspace thereof, equipped with various topologies). A functional $S : X \to \mathbb{R}$ is coercive if $\{u : S[u] \le K\}$ is sequentially relatively compact for all $K \in \mathbb{R}$, i.e. if $(u_n)_n \in X$ have $S[u_n] \le K$, then $u_n$ has a convergent subsequence. It's lower semicontinuous if for any sequence $(u_n)_n$ with $u_n \to u^*$ in $X$, we have $S[u^*] \le \liminf_{n\to\infty} S[u_n]$.

Continuity certainly implies lower semicontinuity, but not the other way around.

**Theorem 5.9.** *Suppose $S : X \to \mathbb{R}$ is coercive and lower semicontinuous, then $S$ achieves its infimum on $X$.*

*Proof.* Let $\sigma = \inf\{S[u] : u \in X\}$. Pick a sequence $(u_n)_n \in X$ with $S[u_n] \to \sigma$. Then in particular $(S[u_n])_n$ is bounded above. As $S$ is coercive, $u_n \to u^* \in X$ after passing to a subsequence.
By lower semicontinuity, $S[u^*] \leq \liminf_{n\to\infty} S[u_n] = \sigma$, and we also have $S[u^*] \geq \sigma$, so $S[u^*] = \sigma$. $\square$

Verifying coercivity and lower semicontinuity can be hard, depending on which space and functional we are interested in. From now on, let's take $X$ to be a reflexive separable Banach space.
Typically, functionals of interest are not coercive in the strong topology of $X$. Weak topology, however, is nice.

**Lemma 5.10.** *Suppose $S : X \to \mathbb{R}$ satisfies $S[n] \leq K \implies \|u\| \leq \tilde{K}$ for some $\tilde{K}$ depending only on $K$. Then $S$ is coercive in the weak topology of $X$.*

*Proof.* Theorem 2.9. $\square$

Lower semicontinuity, on the other hand, becomes hard in weaker topologies. We need some additional structure to do this, usually convexity.

**Lemma 5.11.** *If $U \subset X$ is convex and closed in the strong topology, then it's closed in the weak topology.*

*Proof.* Pick $x \in U^c$. By Theorem 2.17 (applied to the convex compact $\{x\}$ and the convex closed $U$), there is some $\Lambda \in X'$ such that there are $\gamma_1, \gamma_2 \in \mathbb{R}$ with $\operatorname{Re}\Lambda x < \gamma_1 < \gamma_2 < \operatorname{Re}\Lambda y$ for all $y \in U$. Now $\{z : \operatorname{Re}\Lambda z < \gamma_1\}$ is open, contains $x$, and is disjoint from $U$. Thus $U^c$ is weakly open, which means that $U$ is weakly closed. $\square$

**Theorem 5.12.** *If $S : X \to \mathbb{R}$ is convex and lower semicontinuous in the strong topology, then it's lower semicontinuous in the weak topology.*

*Proof.* For any $K$, the set $\{u : S[u] \leq K\}$ is convex by convexity of $S$ and sequentially closed (hence closed) in the strong topology. Therefore it's closed in the weak topology by the preceding lemma.
Suppose $u_n \rightharpoonup u^*$ with $S[u_n] \leq K$. After passing to a subsequence $(n_j)_j$, we can asume WLOG that $S[u_{n_j}] \to \liminf_{n\to\infty} S[u_n]$. For any $\epsilon > 0$, there is some $J$ such that $\forall j \geq J, S[u_{n_j}] \leq \tilde{K} = \liminf_{n\to\infty} S[u_n] + \epsilon$. But $\{S[u] \leq \tilde{K}\}$ is weakly closed, so $S[u^*] \leq \tilde{K}$. As $\epsilon > 0$ is arbitrary, $S[u^*] \leq \liminf_{n\to\infty} S[u_n]$. $\square$

**Example 5.3.** Let $\Omega \in \mathbb{R}^n$ be open and bounded. Suppose $L : \mathbb{C}^n \to \mathbb{R}$ is convex and satisfies $\forall z \in \mathbb{C}^n, L[z] \geq \gamma |z|^2 - C$ for some $\gamma, C > 0$. Then $S : H_0^1(\Omega) \to \mathbb{R}$ via

$$u \mapsto \int_\Omega L(Du)\,\mathrm{d}x$$

has a minimiser.

**Example 5.4.**

$$u \mapsto \int_\Omega (1 + |Du|^4)^{1/4} \, \mathrm{d}x$$

has a minimiser, despite having terrible Euler-Lagrange equations.