Functional Analysis

MATH 510

Winter Term 2 2025/26

Foreword

This set of notes is transcribed from UBC's MATH 510 Functional Analysis course. If any errors are found, please feel free to email me at `nathan.cantafio@stat.ubc.ca`

Contents

1 Banach Spaces **2**

1.1 Normed Vector Spaces 2
1.2 Linear Maps 5
1.3 Linear Functionals 7
1.4 The Baire Category Theorem and its Consequences 9

2 Special cases: L^p and Hilbert Spaces **14**

2.1 L^p Spaces 14
2.2 The Dual of L^P 17
2.3 Hilbert Spaces 21

3 Topological Spaces **28**

3.1 Point Set Topology 28
3.2 Weak, Weak*, and Operator Topologies 31

4 Fourier Analysis **34**

4.1 The Fourier Transform 34
4.2 The Schwartz Class 37
4.3 The Inverse Fourier Transform 39
4.4 Fourier Series 43

5 Further Topics **47**

5.1 Van der Corput's Lemmas 47
5.2 Interpolation Theory 49
5.3 Distribution Theory 51

1 Banach Spaces

1.1 Normed Vector Spaces

Definition 1.1. X is a **vector space** over $\mathbb{C}$ if for all $x, y, z \in X$ and $a, b \in \mathbb{C}$ we have

- **Closure under addition:** $x + y \in X$
- **Commutativity of addition:** $x + y = y + x$
- **Associativity of addition:** $(x + y) + z = x + (y + z)$
- **Existence of 0 vector:** There is $0 \in X$ such that $x + 0 = x$
- **Existence of additive inverse:** For each $x \in X$, there exists $-x$ such that $x + (-x) = 0$
- **Closure under scalar multiplication:** $ax \in X$
- **Distributivity (scalar over vector sum):** $a(x + y) = ax + ay$
- **Distributivity (vector over scalar sum):** $(a + b)x = ax + bx$
- **Associativity of scalar multiplication:** $(ab)x = a(bx)$
- **Multiplicative identity:** $1x = x$

Definition 1.2. (X, d) is a **metric space** if $d : X \times X \to [0, \infty)$ satisfies

(i) $d(x, y) = d(y, x)$

(ii) $d(x, y) = 0 \iff x = y$

(iii) $d(x, z) \leq d(x, y) + d(y, z)$

Definition 1.3. A **norm** on a vector space X (over $\mathbb{C}$) is a function $\| \cdot \| : X \to [0, \infty)$ satisfying

(i) $\|x\| = 0 \iff x = 0$

(ii) $\|\lambda x\| = |\lambda| \|x\|$ for all $\lambda \in \mathbb{C}$ and $x \in X$

(iii) $\|x + y\| \leq \|x\| + \|y\|$

We call $(X, \| \cdot \|)$ a **normed vector space**.

Note 1.4.

- It is easy to show that any normed vector vector space $(X, \| \cdot \|)$ is also a metric space with metric given by $d(x, y) = \|x - y\|$.
- From the triangle inequality: $\|x\| - \|y\| \leq \|x - y\|$, symmetry then yields $|\|x\| - \|y\|| \leq \|x - y\|$. So norms are continuous!

Definition 1.5. Two norms $\| \cdot \|$ and $\| \cdot \|'$ are equivalent if there are constants $C, C' > 0$ such that

$$C\|x\| \leq \|x\|' \leq C'\|x\| \qquad \forall x \in X$$

Note 1.6. Any two norms on a finite dimensional vector space are equivalent.

Definition 1.7. A metric space (X, d) is **complete** if every sequence $\{x_m\}_{m=1}^{\infty}$ in X which is Cauchy (for all $\varepsilon > 0$ there is $N \in \mathbb{N}$ so that $m, n \geq N \implies d(x_m, x_n) < \varepsilon$) also converges in X (there is $x \in X$ so that $\lim_{n\to\infty} d(x, x_n) = 0$).

Proposition 1.8. A normed vector space $(X, \|\cdot\|)$ is complete if and only if every absolutely convergent series is convergent. Namely, a normed vector space is complete if and only if

$$\sum_{k=1}^{\infty} \|x_k\| < \infty \implies \exists x \in X : x = \sum_{k=1}^{\infty} x_k.$$

Proof. Suppose first that $(X, \|\cdot\|)$ is complete. Let $\sum_{k=1}^{\infty} \|x_k\| < \infty$ and define $S_N = \sum_{k=1}^{N} x_k$. Then for $N > M$ we have

$$\|S_N - S_M\| = \left\| \sum_{k=M+1}^{N} x_k \right\| \leq \sum_{k=M+1}^{N} \|x_k\| \to 0 \quad \text{as } M, N \to \infty.$$

So the sequence $\{S_N\}$ is Cauchy and hence convergent (since X is complete).
Conversely, suppose that every absolutely convergent sequence is convergent. Let $\{x_m\}_{m=1}^{\infty}$ be Cauchy. Take $m_1 < m_2 < m_3 < \cdots$ so that $\|x_m - x_n\| < 2^{-j}$ for all $m, n \geq m_j$. Then define $y_1 = x_{m_1}$ and $y_j = x_{m_j} - x_{m_{j-1}}$ for $j \geq 2$. Then $\sum_{j=1}^{k} y_j = x_{m_k}$. Moreover,

$$\sum_{j\in\mathbb{N}} \|y_j\| \leq \|x_{m_1}\| + \sum_{j\in\mathbb{N}} 2^{-j} < \infty.$$

So $\lim x_{m_k} = \sum_{j=1}^{\infty} y_j$ exists. Finally, we verify that $\{x_m\}$ and $\{x_{m_k}\}$ have the same limit. Since $\{x_m\}$ is Cauchy, given $\varepsilon > 0$ there exist $N_0, K_0 \in \mathbb{N}$ so that $\|x_m - x_{m_k}\| < \varepsilon/2$ for all $m, m_k \geq N_0$ and so that $\|x_{m_k} - x\| < \varepsilon/2$ for all $k > K_0$. Hence for $m \geq \max\{N_0, m_{K_0}\}$ we have

$$\|x_m - x\| \leq \|x_m - x_{m_k}\| + \|x_{m_k} - x\| < \varepsilon.$$

□

Definition 1.9. A **Banach space** is a complete, normed vector space.

Example 1.10.

(i) $\mathbb{C}^m$ over $\mathbb{C}$ or $\mathbb{R}^m$ over $\mathbb{R}$.

(ii) $\ell^\infty = \{x = (x_1, x_2, \dots) : x_j \in \mathbb{C} \text{ and } \|x\|_\infty < \infty\}$ where $\|x\|_\infty = \sup_j |x_j|$

(iii) $C([0,1]) = \{f : [0,1] \to \mathbb{C} : f \text{ is continuous}\}$ with $\|f\|_\infty = \sup_{x\in[0,1]} |f(x)|$

That (i) is a normed vector space is that (i) is complete are both standard exercises. That (ii) and (iii) are normed vector space is left as an exercise. Below we show that (ii) and (iii) are complete.

- Suppose that $\{x^k\}_{k=1}^{\infty} = \{(x_1^k, x_2^k, \dots)\}_{k=1}^{\infty}$ in ℓ^∞ is Cauchy. Then for each $m \in \mathbb{N}$ we have

  $$|x_m^k - x_m^j| \leq \|x^k - x^j\|$$

  so $\{x_m^k\}_{k=1}^{\infty}$ as a sequence in $\mathbb{C}$ is Cauchy, and since $\mathbb{C}$ is complete we get

  $$\lim_{k\to\infty} x_m^k = x_m$$

for some $x_m \in \mathbb{C}$. Define $x = (x_1, x_2, \dots)$. We need to show that $\|x^k - x\| \to 0$ as $k \to \infty$. Choose $K \in \mathbb{N}$ so that for all $j, k \geq K$ we have $\|x^k - x^j\| < \varepsilon$. In fact, for all $m \in \mathbb{N}$ we have

$$|x_m^k - x_m^j| \leq \|x^k - x^j\| < \varepsilon.$$

And furthermore that

$$|x_m^k - x_m| \leq |x_m^k - x_m^j| + |x_m^j - x_m| \leq \varepsilon + \underbrace{|x_m^j - x_m|}_{\to 0 \text{ as } j \to \infty}.$$

So we see that $|x_m^k - x_m| \leq \varepsilon$ for all $k \geq K$ independent of m. Hence

$$\|x^k - x\| = \sup_m |x_m^k - x_m| \leq \varepsilon$$

for all $k \geq K$ and so $\lim_{k\to\infty} \|x_k - x\| = 0$. It remains to show that $x \in \ell^\infty$. And indeed $\|x\| \leq \|x_k - x\| + \|x_k\| \leq 1 + \|x_k\| < \infty$ for k large enough.

- Suppose $\{f_k\}_{k=1}^\infty$ in $C([0,1])$ is Cauchy. Then for each $x \in [0,1]$ we have

$$\|f_k(x) - f_j(x)| \leq \|f_k - f_j\|$$

so $\{f_k(x)\}_{k=1}^\infty$ is Cauchy as a sequence in $\mathbb{C}$, and since $\mathbb{C}$ is complete we get that

$$\lim_{k\to\infty} f_k(x) = f(x)$$

for some $f(x) \in \mathbb{C}$. Define the function $f : x \mapsto f(x)$. Now choose $K \in \mathbb{N}$ so that for all $j, k \geq K$ we have $\|f_k - f_j\| < \varepsilon$. Then for all $x \in [0,1]$ we have

$$|f_k(x) - f_j(x)| \leq \|f_k - f_j\| < \varepsilon.$$

And furthermore that

$$|f_k(x) - f(x)| \leq |f_k(x) - f_j(x)| + |f_j(x) - f(x)| \leq \varepsilon + \underbrace{|f_j(x) - f(x)|}_{\to 0 \text{ as } j \to \infty}.$$

So we see that $|f_k(x) - f(x)| < \varepsilon$ for all $k \geq K$, independent of $x \in [0,1]$. Hence

$$\|f_k - f\| = \sup_{x\in[0,1]} |f_k(x) - f(x)| \leq \varepsilon$$

for all $k \geq K$ and so $\lim_{k\to\infty} \|f_k - f\| = 0$. It remains to show that $f \in C([0,1])$. Now fix $x \in [0,1]$. Let $\varepsilon > 0$ and choose $N \in \mathbb{N}$ so that $\|f_N - f\| < \varepsilon/3$. By continuity there is $\delta > 0$ such that $|y - x| < \delta \implies |f_N(x) - f_N(y)| < \varepsilon/3$. Hence for $|y - x| < \delta$ we have

$$|f(x) - f(y)| \leq |f(x) - f_N(x)| + |f_N(x) - f_N(y)| + |f_N(y) - f(y)| < \varepsilon.$$

Namely f is continuous at x, and this holds for all $x \in [0,1]$

1.2 Linear Maps

Definition 1.11. A **linear map** $T : X \to Y$ between vector spaces X and Y satisfies

$$T(\lambda_1 x + \lambda_2 y) = \lambda_1 T(x) + \lambda_2 T(y)$$

for all $x, y \in X$ and $\lambda_1, \lambda_2 \in \mathbb{C}$.

Definition 1.12. A linear map $T : X \to Y$ is **bounded** if there is $C > 0$ such that

$$\|Tx\|_Y \leq C\|x\|_X$$

for all $x \in X$.

Example 1.13.

(i) All linear maps from $\mathbb{C}^m$ to $\mathbb{C}^n$ can be written in the form $x \mapsto Ax$ where A is an $m \times n$ matrix with entries in $\mathbb{C}$. Now we can write

$$\|Ax\|_2^2 = \sum_j |(Ax)_j|^2 = \sum_j (Ax)_j \overline{(Ax)_j} = \sum_{j,k,\ell} \overline{A_{j,k}}\,\overline{x_k} A_{j,\ell} x_\ell.$$

Using the fact that $ab \leq \frac{a^2+b^2}{2}$ for $a, b \geq 0$ we can then bound

$$\|Ax\|_2^2 \leq \max_{s,t} |A_{s,t}|^2 \sum_{j,k,\ell} \frac{n}{2}(|x_k|^2 + |x_\ell|^2) \leq \underbrace{\max_{s,t} |A_{s,t}|^2 m^2}_{C} \|x\|_2^2.$$

Namely, all linear maps from $\mathbb{C}^m \to \mathbb{C}^n$ are bounded.

(ii) Let $X = C([0,1])$ with $\|\cdot\|_\infty$. Fix $g \in C([0,1])$ and define $T : X \to X$ which maps $f \mapsto gf$. It is clearly a well-defined map and it is linear. It is also bounded, since

$$\|Tf\|_\infty = \|gf\|_\infty = \sup_{x\in[0,1]} |g(x)f(x)| \leq \sup_{x\in[0,1]} |g(x)| \sup_{x\in[0,1]} |f(x)| \leq \underbrace{\|g\|_\infty}_{C} \|f\|_\infty.$$

(iii) Let $X = C^1([0,1])$ with $\|f\|_{C^1} = \|f\|_\infty + \|f'\|_\infty$ and let $Y = C([0,1])$ with $\|\cdot\|_\infty$. Then the map $\frac{\mathrm{d}}{\mathrm{d}x} : X \to Y$ is a bounded linear map. Indeed, we have

$$\|Tf\|_\infty = \|f'\|_\infty \leq \|f\|_\infty + \|f'\|_\infty = \|f\|_{C^1}.$$

(iv) Let $X = C^\infty([0,1])$ with $\|\cdot\|_\infty$. Then $\frac{\mathrm{d}}{\mathrm{d}x} : X \to X$ is unbounded. Indeed, consider the sequence $f_n(x) = \sin(2\pi n x)$. Then $\|f\|_\infty = 1$, however $\frac{\mathrm{d}}{\mathrm{d}x} f_n(x) = 2\pi n \cos(2\pi n x)$ which has $\|\frac{\mathrm{d}}{\mathrm{d}x} f_n\|_\infty = 2\pi n$. In particular, for any constant C, there is $n \in \mathbb{N}$ such that $f_n \in C^\infty([0,1])$ and $\|\frac{\mathrm{d}}{\mathrm{d}x} f_n\|_\infty > C = C\|f_n\|_\infty$. So $\frac{\mathrm{d}}{\mathrm{d}x}$ cannot be bounded.

Proposition 1.14. For any linear map $T : X \to Y$ between normed vector spaces X, Y. The following are equivalent:

(i) T is bounded

(ii) T is continuous

(iii) T is continuous at 0

Proof. (ii) $\implies$ (iii) is trivial. Now we show (i) $\implies$ (ii). There exists some constant C such that

$$\|Tx - Ty\| = \|T(x-y)\| \leq C\|x-y\|,$$

so T is continuous. Now we show (iii) $\implies$ (i). By linearity we have $T(0) = 0$. And since T is continuous at 0 there is $\delta > 0$ such that $\|Tx\| \leq 1$ whenever $\|x\| \leq \delta$. So

$$\|Tx\| = \left\| T\left(\frac{\|x\|}{\delta}\frac{x}{\|x\|}\delta\right)\right\| = \frac{\|x\|}{\delta}\left\|T\left(\frac{x}{\|x\|}\delta\right)\right\| \leq \frac{\|x\|}{\delta}.$$

Namely T is bounded. □

Definition 1.15. For normed vector spaces X, Y, we define $L(X,Y)$ to be the set of all bounded linear maps from X to Y. Note that $L(X,Y)$ is itself a vector space.

Definition 1.16. For $T \in L(X,Y)$, the following are equivalent definitions of the operator norm:

$$\begin{aligned}\|T\| &= \sup\{\|Tx\|_Y : \|x\|_X = 1\} \\ &= \sup_{x\neq 0}\frac{\|Tx\|_Y}{\|x\|_X} \\ &= \inf\{C \geq 0 : \|Tx\|_Y \leq C\|x\|_X, x \in X\}\end{aligned}$$

Note 1.17. Firstly,

$$T \text{ is bounded} \iff \text{there is } C > 0 \text{ such that } \|Tx\| \leq C\|x\| \text{ for all } x \in X \iff \|T\| < \infty.$$

Moreover, for $x \neq 0$, $\|Tx\| = \frac{\|Tx\|}{\|x\|}\|x\| \leq \|T\|\|x\|$ which implies the inf above is in fact a min. It is also worth explicitly mentioning that the operator norm as defined, is in fact a norm, and that $(L(X,Y), \|\cdot\|)$ is a normed vector space.

Proposition 1.18. Let X, Y be normed vector spaces. If Y is complete (Banach), then $L(X,Y)$ is Banach.

Proof. It suffices to check completeness. Let $\{T_n\}_{n\in\mathbb{N}}$ be Cauchy in $(L(X,Y), \|\cdot\|)$. Then for all $x \in X$ we have

$$\|T_n x - T_m x\| = \|(T_n - T_m)x\| \leq \underbrace{\|T_n - T_m\|}_{\to 0 \text{ as } n,m\to\infty}\|x\|.$$

In particular, $\{T_n x\}_{n\in\mathbb{N}}$ is Cauchy in Y. So by the completeness of Y there is $Tx \in Y$ such that $Tx = \lim_{n\to\infty} T_n x$. Define the linear map $T : x \mapsto Tx$. Then T is linear because limits are linear. Moreover, T is bounded. Indeed, there is $N \in \mathbb{N}$ such that $\|T_n - T_m\| \leq 1$ for $n, m \geq N$, hence

$$\|T_n\| = \|T_n - T_N + T_N\| \leq \|T_n - T_N\| + \|T_N\| \leq 1 + \|T_N\|$$

for all $n \geq N$. Then define

$$M := \max\{\|T_1\|, \|T_2\|, \ldots, \|T_N\|, 1 + \|T_N\|\}$$

so that $\|T_n\| \leq M$ for all $n \in \mathbb{N}$. Finally by the continuity of norms we write for any $x \in X$:

$$\|Tx\| = \|\lim_{n\to\infty} T_n x\| = \lim_{n\to\infty}\|T_n x\| \leq M\|x\|.$$

So T is bounded and hence $T \in L(X,Y)$. It remains to show that $\|T - T_n\| \to 0$ as $n \to \infty$. But given $\varepsilon > 0$, there is $N \in \mathbb{N}$ so that $n \geq N$

$$\|T_n - T\| = \|T_n - \lim_{m\to\infty} T_m\| = \lim_{m\to\infty}\|T_n - T_m\| \leq \varepsilon.$$

□

Note 1.19. If $S, T \in L(X, X)$, then $\|ST\| \leq \|S\|\|T\|$ where ST is the composition of S and T. This follows since

$$\|STx\| \leq \|S\|\|Tx\| \leq \|S\|\|T\|\|x\|,$$

for all $x \in X$. And $\|ST\|$ is the inf over all such constants. So $\|ST\| \leq \|S\|\|T\|$.

1.3 Linear Functionals

Definition 1.20. A linear map $f : X \to \mathbb{C}$ (or $\mathbb{R}$) is a **linear functional** on X

Definition 1.21. The **dual space** of X, denoted $X^\star$, is $X^\star = L(X, \mathbb{C})$ (or $L(X, \mathbb{R})$). Namely, $X^\star$ is the space of bounded linear functionals on X.

Note 1.22. By Prop. 1.18 $X^\star$ equipped with the operator norm $\|f\| = \sup_{\|x\|=1} |f(x)|$ is Banach.

Example 1.23. Let $X = C([0, 1])$ with $\|\cdot\|_\infty$.

(i) Given $f \in L^1([0, 1])$, define $I_f : X \to C$ via $g \mapsto \int_0^1 gf \, \mathrm{d}x$. This is linear. It is also bounded:

$$|I_f g| = \left| \int fg \, \mathrm{d}x \right| \leq \int |f||g| \, \mathrm{d}x \leq \|g\|_\infty \|f\|_{L^1}.$$

(ii) Given $x_0 \in [0, 1]$, define $E_{x_0} : X \to \mathbb{C}$ via $g \mapsto g(x_0)$. This is linear. It also bounded:

$$|E_{x_0} g| = |g(x_0)| \leq \|g\|_\infty.$$

Notice that both of the above are examples of integrating against some measure.

$$I_f g = \int_0^1 g(f \, \mathrm{d}x) \qquad E_{x_0} g = \int_0^1 g \, \mathrm{d}\delta_{x_0}.$$

Definition 1.24. Let X be a vector space over $\mathbb{R}$. A **sublinear functional** on X is a map $p : X \to \mathbb{R}$ such that

$$p(x + y) \leq p(x) + p(y) \text{ for all } x, y \in X \qquad \text{and} \qquad p(\lambda x) = \lambda x \text{ for all } \lambda > 0 \text{ and } x \in X.$$

Note: a norm or "seminorm" (same as a norm without requiring $\|x\| = 0 \implies x = 0$) are examples of sublinear functionals.

Theorem 1.25. *(Hahn-Banach, $\mathbb{R}$ version).* Let X be a vector space over $\mathbb{R}$, p a sublinear functional on X, $M \subset X$ a subspace, and $f : M \to \mathbb{R}$ a linear functional such that $f(x) \leq p(x)$ for all $x \in M$. Then there exists a linear functional $F : X \to \mathbb{R}$ such that $F \restriction_M = f$ and $F(x) \leq p(x)$ for all $x \in X$.

Proof. If $M = X$ we are done. So suppose $M \neq X$.

- There is $x \in X \setminus M$. Define $g : M + \mathbb{R}x \to \mathbb{R}$ so that $g(y + \lambda x) := f(y) + \lambda g(x)$ for all $y \in M$ and $\lambda \in \mathbb{R}$. Write $\alpha = g(x) \in \mathbb{R}$. We need to choose α. For $y_1, y_2 \in M$:

  $$f(y_1) + f(y_2) = f(y_1 + y_2) \leq p(y_1 + y_2) = p(y_1 - x + y_2 + x) \leq p(y_1 - x) + p(y_2 + x).$$

  This then implies

  $$f(y_1) - p(y_1 - x) \leq p(y_2 + x) - f(y_2) \qquad \text{for all } y_1, y_2 \in M.$$

Taking first the inf of the RHS and the sup of the LHS yields:

$$\sup_{y_1 \in M} f(y_1) - p(y_1 - x) \leq \inf_{y_2 \in M} p(y_2 + x) - f(y_2).$$

Namely there exists $\alpha \in [\sup \text{LHS}, \inf \text{RHS}]$. If $\lambda > 0$, then

$$g(y + \lambda x) = \lambda \left[f\left(\frac{y}{\lambda}\right) + \alpha \right] \leq \lambda \left[f\left(\frac{y}{\lambda}\right) + p\left(\frac{y}{\lambda} + x\right) - f\left(\frac{y}{\lambda}\right) \right] \leq p(y + \lambda x).$$

If $\lambda < 0$, a similar argument yields $g(y + \lambda x) \leq p(y + \lambda x)$, and if $\lambda = 0$ then

$$g(y + \lambda x) = g(y) = f(y) \leq p(y).$$

- Define $\mathcal{F}$ to be the family of linear extensions

$$\mathcal{F} = \left\{ \hat{f} : \hat{M} \to \mathbb{R} \,|\, \hat{M} \supset M, \hat{f} \restriction_M = f, \hat{f} \leq p \text{ on } \hat{M} \right\}.$$

Then $\mathcal{F}$ is partially ordered where

$$(f_1 : M_1 \to \mathbb{R}) \subset (f_2 : M_2 \to \mathbb{R})$$

means $M_1 \subset M_2$ and $f_2 \restriction_{M_1} = f_1$. Now, note that any linearly ordered (meaning for all $\alpha, \beta \in \mathcal{I}$ either $f_\alpha \subset f_\beta$ or $f_\beta \subset f_\alpha$) sub-family $\{f_\alpha\}_{\alpha \in \mathcal{I}} \subset \mathcal{F}$ has an upper bound. Indeed, set $\hat{M} = \bigcup_{\alpha \in \mathcal{I}} M_\alpha$ and define $\hat{f} : \hat{M} \to \mathbb{R}$ so that $\hat{f}(x) = f_\alpha(x)$ if $x \in M_\alpha$ (note that if $x \in \hat{M}$ then $x \in M_\alpha$ for some α and all M_β which are "bigger", then by the restriction property we have that $\hat{f}$ is well-defined). Defined this way, we see that $f_\alpha \subset \hat{f}$ for all $\alpha \in \mathcal{I}$. Finally, we apply Zorn's Lemma (actually an axiom) which says that if every linearly ordered subset of a partially ordered set S has an upper bound, then S has a maximal element. This yields a maximal extension F of f, such that $f \leq p$ and f is defined on X. (F must be defined on X, if not we apply the first bullet point to obtain a new extension larger than F, a contradiction).

□

Definition 1.26. A **seminorm** is a map $p : X \to [0, \infty)$ such that for all $x, y \in X$ and $\lambda \in \mathbb{C}$ we have $p(\lambda x) = |\lambda| p(x)$ and $p(y + x) \leq p(y) + p(x)$.

Theorem 1.27. *(Hahn-Banach, $\mathbb{C}$ version)* Let X be a complex vector space, p a seminorm on X, $M \subset X$ a subspace, and $f : M \to \mathbb{C}$ a linear functional such that $|f(x)| \leq p(x)$ for all $x \in M$. Then there is a linear functional $F : X \to \mathbb{C}$ so that $F \restriction_M = f$ and $|F(x)| \leq p(x)$ for all $x \in X$.

Proof. The proof is omitted. □

Theorem 1.28. If $M \subset X$ is a closed subspace and $x \in X \setminus M$. Then there is $f \in X^\star$ satisfying $f(x) = \delta := \inf_{y \in M} \|y - x\| > 0$, $f \restriction_M = 0$, and $\|f\| = 1$. [Note: M is assumed closed so that $\delta > 0$].

Proof. Define $f : M + \mathbb{C}x \to \mathbb{C}$ so that $f(y + \lambda x) = \lambda \delta$ for all $y \in M$ and $\lambda \in \mathbb{C}$. This satisfies $f \restriction_M = 0$, $f(x) = \delta$, and f linear. Now write

$$|f(y + \lambda x)| = |\lambda \delta| = |\lambda| \delta \leq |\lambda| \left\| \frac{y}{\lambda} + x \right\| = \|y + \lambda x\|.$$

Hence $\| \cdot \|$ works as p in the extension theorem. So we can extend f to X. It remains to check that $\|f\| = 1$. Since $|f(z)| \leq \|z\|$ we have that $\frac{|f(z)}{\|z\|} \leq 1$, and taking the supremum over all $z \in X$ yields $\|f\| \leq 1$. To check $\|f\| \geq 1$. Let $\varepsilon > 0$, by properties of infimum, there is $y \in M$ so that $\|y - x\| \leq \delta(1 + \varepsilon)$. Then $|f(y - x)| = \delta \geq \frac{\|y - x\|}{1 + \varepsilon}$, and so $\frac{|f(y-x)|}{\|y-x\|} \geq \frac{1}{1+\varepsilon}$. Since $\varepsilon > 0$ was arbitrary, we conclude that $\|f\| \geq 1$. □

Corollary 1.29. If $X \ni x \neq 0$, then there is $f \in X^\star$ such that $\|f\| = 1$ and $f(x) = \|x\|$.

Proof. Take $M = \{0\}$ and apply Hahn-Banach. □

Corollary 1.30. If $x, y \in X$ with $x \neq y$, then there is $f \in X^\star$ such that $f(x) \neq f(y)$ and $\|f\| = 1$

Proof. Apply Corollary 1.29 to $z = x - y$ and use linearity. □

Note 1.31. Recall we defined the dual space of Z, denoted $Z^\star = L(Z, \mathbb{C})$, as the space of bounded linear functions on Z. When $Z = X^\star$, we denote the dual space of $X^\star$ as $X^{\star\star} = Z^\star = L(X^\star, \mathbb{C})$.

Theorem 1.32.

(i) For $x \in X$, the map $\hat{x} : X^\star \to \mathbb{C}$ given by $\hat{x} : f \mapsto f(x)$ is a bounded linear functional on $X^\star$.

(ii) The map $X \to X^{\star\star}$ given by $x \mapsto \hat{x}$ is a linear isometry.

Proof.

(i) $\hat{x}$ is clearly linear. It is also bounded since $|\hat{x}(f)| = |f(x)| \leq \|f\|\|x\|$.

(ii) We need to show that $\|\hat{x}\| = \|x\|$. The previous point implies $\|\hat{x}\| \leq \|x\|$. Then by Corollary 1.29 there is $f \in X^\star$ such that $\|f\| = 1$ and $|f(x)| = \|x\|$. Hence

$$\|\hat{x}\| = \|\hat{x}\|\|f\| \geq |\hat{x}(f)| = |f(x)| = \|x\|.$$

□

Note 1.33. We can identify each $x \in X$ with an element of $X^{\star\star}$ via

$$X \approx \hat{X} := \{\hat{x} : x \in X\} \subset X^{\star\star}.$$

We know that $X^{\star\star}$ is complete by Prop. 1.18, but we don't know that $\hat{X}$ is closed in $X^{\star\star}$. We can study the closure $\overline{\hat{X}}$ called "the completion of X". A Banach space is called **reflexive** if $\hat{X} = X^{\star\star}$.

1.4 The Baire Category Theorem and its Consequences

Definition 1.34. A set D is dense in X if $\overline{D} = X$. Equivalently, a set D is dense in X if $O \cap D \neq \emptyset$ for all non-empty open $O \subset X$.

Theorem 1.35. *(The Baire Category Theorem)* Let X be a complete metric space. If $\{U_m\}_{m\in\mathbb{N}}$ are open and dense, then $\bigcap_{m\in\mathbb{N}} U_m$ is dense.

Proof. Let $W \neq \emptyset$ be open. Since U_1 is open and dense, we have that $U_1 \cap W$ is open and non-empty. Namely there is $x_1 \in U_1 \cap W$ and in fact by openness there is some $0 < r_1 < 1/2$ such that

$$\overline{B_{r_1}(x_1)} \subset U_1 \cap W.$$

Now $U_2 \cap B_{r_1}(x_1)$ is open and non-empty, so there is $\overline{B_{r_2}(x_2)} \subset U_2 \cap B_{r_1}(x_1)$ for some $0 < r_2 < 1/2^2$. Inductively we continue to construct

$$\overline{B_{r_m}(x_m)} \subset U_m \cap B_{r_{m-1}}(x_{m-1}) \qquad x_m \in X \text{ and } 0 < r_m < 1/2^m.$$

In particular, since the balls are nested, we have that $x_m, x_n \in B_{r_N}(x_N)$ for $n, m \geq N$. Then since $r_N \to 0$ as $N \to \infty$ we conclude that $\{x_m\}_{m\in\mathbb{N}}$ is Cauchy, and as X is complete there is $X \ni x = \lim_{m\to\infty} x_m$. Let $N \in \mathbb{N}$, then

$$\text{for all } m \geq N\text{: } x_m \in B_{r_N}(x_N) \implies x \in \overline{B_{r_N}(x_N)} \subset U_N \cap W.$$

Since this is true for all $N \in \mathbb{N}$ we conclude that $x \in \bigcap_{N\in\mathbb{N}} U_N \cap W$. In particular, the intersection is non-empty; and since W was arbitrary we have that $\bigcap_{N\in\mathbb{N}} U_n$ is dense. □

Definition 1.36. A set $E \subset X$ is **nowhere dense** if $\overline{E}$ does not contain any non-empty open set. That is, the interior of the closure is empty: $(\overline{E})^\circ = \emptyset$.

Definition 1.37. A set is **meagre** if it is a countable union of nowhere dense sets. The complement of a meagre set is called **residual**.

Corollary 1.38. If X is a complete metric space, a residual set is dense. In particular, the whole space cannot be meagre.

Proof. Let $R \subset X$ be a residual set. Namely $R^c = \bigcup_{m=1}^\infty E_m$ where E_m are nowhere dense. Then

$$R = \bigcap_{m=1}^\infty E_m^c \supset \bigcap_{m=1}^\infty (\overline{E_m})^c.$$

Since $(\overline{E_m})^c$ is open, it is enough to prove that $(\overline{E_m})^c$ is dense. Then indeed the Baire Category Theorem implies $\bigcap_{m=1}^\infty (\overline{E_m})^c$ is dense which then implies R is dense. Let U be any non-empty open set. Since E_m is nowhere dense, $U \not\subset \overline{E_m}$. Thus $U \cap (\overline{E_m})^c \neq \emptyset$ as desired. Moreover, $\emptyset$ is not dense and so cannot be residual. Hence $X = \emptyset^c$ is not meagre. □

Example 1.39. *The set of nowhere differentiable functions is dense in the space $C([0,1])$ equipped with $\|\cdot\|_\infty$.* Recall that $C([0,1])$ with $\|\cdot\|_\infty$ is Banach. Define

$$E_m = \{f \in C([0,1]) : \text{there exists } x_0 \text{ so that } |f(x) - f(x_0)| \leq m|x - x_0| \text{ for all } x \in [0,1]\}.$$

Claim: if f is differentiable at some point, then $f \in \bigcup_{m=1}^\infty E_m$. Namely $\mathcal{D} \subset \bigcup_{m=1}^\infty E_m$ where

$$\mathcal{D} = \{f \in C([0,1]) : f \text{ is differentiable at at least one point}\}.$$

Proof. If $\lim_{x\to y} \frac{|f(x)-f(y)|}{|x-y|}$ exists (i.e. $f'(y)$ exists), then there is $\delta > 0$ so that

$$|x - y| < \delta \implies \left|\frac{f(x) - f(y)}{x - y}\right| < |f'(y)| + 1.$$

On the other hand: if $|x - y| \geq \delta$, then $\left|\frac{f(x)-f(y)}{x-y}\right| \leq \frac{2\|f\|_\infty}{\delta}$. Namely there is some $y \in [0,1]$ (where $f'(y)$ exists) such that

$$|f(x) - f(y)| \leq \left(\underbrace{|f'(y)| + 1 + \frac{2\|f\|_\infty}{\delta}}_{C}\right)|x - y|$$

for all $x \in [0,1]$. And moreover there is $N \in \mathbb{N}$ so that $N \geq C$. Hence $f \in E_N$. □

So the set of nowhere differentiable functions $\mathcal{D}^c$ contains

$$\mathcal{D}^c \supset \left(\bigcup_{m=1}^{\infty} E_m\right)^c =: R.$$

Hence it is enough to show that R is dense, and by Cor. 1.38 it is enough to show that R is residual, or in other words that E_m are nowhere dense...

Definition 1.40. A map $T : X \to Y$ is **open** if $T(U)$ is open for all open $U \subset X$. [Contrast this to continuous which requires $T^{-1}(V)$ to be open for all open $V \subset Y$.]

Theorem 1.41. *(Open Mapping Theorem)* Let X and Y be Banach spaces. If $T \in L(X, Y)$ is surjective, then T is open.

Proof. It suffices to show that T maps open balls (in X) to open sets (in Y) since any open set can be written as a union of open balls. Moreover notice

$$T(B_r(x_0)) = T(x_0 + rB_1(0)) = T(x_0) + rT(B_1(0))$$

so it suffices to check that $T(B_1(0))$ is open.
Claim: it suffices to check that $T(B_1(0)) \supset B_\delta(0)$ for some $\delta > 0$. Indeed: if $T(B_1(0)) \supset B_\delta(0)$ for some $\delta > 0$, then for any $y \in T(B_1(0))$ we have $y = Tx$ with $x \in B_1(0)$. Choose $r := 1 - \|x\|$ so that $B_r(x) \subset B_1(0)$. Then

$$T(B_1(0)) \supset T(B_r(x)) = Tx + rT(B_1(0)) \supset y + rB_\delta(0) = B_{\delta r}(y).$$

Which implies $T(B_1(0))$ is open.
Now since T is surjective write:

$$\bigcup_{k\in\mathbb{N}} T(B_k(0)) = T\left(\bigcup_{k\in\mathbb{N}} B_k(0)\right) = T(X) = Y.$$

By corollary 1.38 we have that Y is not meagre. So there must be $k \in \mathbb{N}$ such that $T(B_k(0))$ is not nowhere dense. Note that one can check $\phi : x \mapsto kx$ is a homeomorphism (bijective, continuous, with continuous inverse) and it maps $\phi : T(B_1(0)) \mapsto T(B_k(0))$. Since nowhere denseness is preserved by homeomorphism [specifically $(\overline{\phi(A)})^\circ = \phi((\overline{A})^\circ)$], we have that

$$T(B_1(0)) \text{ is nowhere dense } \iff T(B_k(0)) \text{ is nowhere dense.}$$

So $T(B_1(0))$ is not nowhere dense. Namely there is $r > 0$ and $y_0 \in Y$ such that

$$\overline{T(B_1(0))} \supset B_{4r}(y_0).$$

Now since $T(B_1(0))$ is dense in its closure, and $y_0 \in \overline{T(B_1(0))}$, there is $y_1 \in T(B_1(0))$ such that $\|y_1 - y_0\| < 2r$ [moreover, $y_1 = Tx_1$ for some $x_1 \in B_1(0)$]. Now if $\|y\| < 2r$, then

$$y + y_1 \in B_{2r}(y_1) \subset B_{4r}(y_0) \subset \overline{T(B_1(0))}.$$

And then:

$$y = -Tx_1 + (y + y_1) \in -Tx_1 + \overline{T(B_1(0))} = \overline{T(-x_1 + B_1(0))} \subset \overline{T(B_2(0))}.$$

This then implies that for any $m \in \mathbb{N}$:

$$\|y\| < 2^{-m}r \implies y \in \overline{T(B_{2^{-m}}(0))}.$$

And in particular that $B_r(0) \subset \overline{T(B_1(0))}$. We now show that $B_{r/2}(0) \subset T(B_1(0))$.
Fix $\|y\| < r/2$. We inductively construct $x_1, x_2 \dots$ with $x_j \in B_{2^{-j}}(0)$ so that

$$\left\| y - \sum_{j=1}^{m} Tx_j \right\| < 2^{-(m+1)}r$$

for any $m \in \mathbb{N}$. For $m = 1$, since $\|y\| < 2^{-1}r$ we have $y \in \overline{T(B_{2^{-1}}(0))}$ which implies the existence of $x_1 \in B_{2^{-1}}(0)$ so that $\|y - Tx_1\| < r/4$. Now the inductive step: suppose $x_1, \dots, x_m$ are such that $x_j \in B_{2^{-j}}(0)$ and $\|y - \sum_{j=1}^{m} Tx_j\| < 2^{-(m+1)}r$. We have that $y - \sum_{j=1}^{m} Tx_j \in \overline{T(B_{2^{-(m+1)}}(0))}$ which implies the existence of $x_{m+1} \in B_{2^{-(m+1)}}(0)$ such that $\|y - \sum_{j=1}^{m+1} Tx_j\| < 2^{-(m+2)}r$. Now since $\sum_{j\in\mathbb{N}} \|x_j\| < \sum_{j\in\mathbb{N}} 2^{-j} = 1$ and X is Banach, by Prop. 1.8 implies that $\sum_{j\in\mathbb{N}} x_j = x \in X$. Finally,

$$y = \lim_{m\to\infty} \sum_{j=1}^{m} Tx_j = T\left(\lim_{m\to\infty} \sum_{j=1}^{m} x_j \right) = Tx \in T(B_1(0)).$$

Since this holds for any $y \in B_{r/2}(0)$ we see that $B_{r/2}(0) \subset T(B_1(0))$ as desired. □

Corollary 1.42. Let X and Y be Banach spaces. If $T \in L(X, Y)$ is bijective, then $T^{-1} \in L(Y, X)$.

Proof. By the Open Mapping Theorem T is open. T^{-1} exists (since T is bijective) and is continuous (since T is open). Moreover T^{-1} is linear and so by Proposition 1.14 we have $T^{-1} \in L(Y, X)$. □

Definition 1.43. A linear map $T : X \to Y$ is said to be **closed** if its graph

$$\Gamma(T) = \{(x, y) \in X \times Y : Tx = y\}$$

is closed in $X \times Y$ with $\|(x, y)\| = \|x\|_X + \|y\|_Y$.
Note: to check T is closed it suffices to check that

$$(x_j \to x, Tx_j \to y) \implies Tx = y.$$

Furthermore: if T is bounded, then T is closed. Indeed, if $x_j \to x$ and $Tx_j \to y$, then

$$\|y - Tx\| \le \|y - Tx_j\| + \|Tx_j - Tx\| \le \|y - Tx_j\| + \|T\|\|x_j - x\| \to 0.$$

Which then implies $y = Tx$ as desired.

Theorem 1.44. *(Closed Graph Theorem)* Let X and Y be Banach spaces. Suppose $T : X \to Y$ is linear and closed. Then T is bounded.

Proof. Since X and Y are Banach, one can show that $X \times Y$ is Banach. Then since the graph $\Gamma(T) \subset X \times Y$ is a closed subset of a Banach space, $\Gamma(T)$ is itself Banach. Define the projection maps

$$\begin{matrix} \pi_1 : \Gamma(T) \to X \\ \pi_1 : (x, Tx) \mapsto x \end{matrix} \qquad \text{and} \qquad \begin{matrix} \pi_2 : \Gamma(T) \to Y \\ \pi_2 : (x, Tx) \mapsto Tx \end{matrix}.$$

Both π_i are trivially bounded. For example $\|\pi_2(x, Tx)\|_Y = \|Tx\|_Y \le \|x\|_X + \|Tx\|_Y = \|(x, Tx)\|$. Moreover, π_1 is a bijection. So by Corollary 1.42 we have that π_1^{-1} is bounded. Hence $T = \pi_2 \circ \pi_1^{-1}$ as a composition of bounded maps is bounded. □

Theorem 1.45. *(Uniform Bounded Principle)* Let X and Y be normed vector spaces. Let $A \subset L(X,Y)$ be a collection of bounded linear maps. Then

(i) If X is not meagre and $\sup_{T\in A} \|Tx\| < \infty$ for all $x \in X$, then $\sup_{T\in A} \|T\| < \infty$

(ii) If X is Banach and $\sup_{T\in A} \|Tx\| < \infty$ for all $x \in X$, then $\sup_{T\in A} \|T\| < \infty$

Namely, X being meagre (in fact Banach) is enough to turn a point-wise bound on $\|Tx\|$ into a uniform bound.

Proof. First of all, by Corollary 1.38 we have (i) $\implies$ (ii) since a Banach space cannot be meagre. So it suffices to prove (i). Define

$$E_m = \{x \in X : \sup_{T\in A} \|Tx\| \leq m\} = \bigcap_{T\in A} \{x \in X : \|Tx\| \leq m\}.$$

Notice that E_m is closed since for each $T \in A$ the set $\{x \in X : \|Tx\| \leq m\}$ is closed, and an arbitrary intersection of closed sets is closed. Now we write

$$X = \{x \in X : \sup_{T\in A} \|Tx\| < \infty\} = \bigcup_{m\in\mathbb{N}} E_m$$

and note that X is not meagre. So at least one of the E_m is not nowhere dense. Hence there is $\overline{B_r(x_0)} \subset E_m$ for some $x_0 \in E_m$ and $r > 0$. For any $T \in A$: if $\|x\| \leq r$, then

$$\|Tx\| \leq \|T(x + x_0)\| + \|Tx_0\| \leq 2m$$

since both $x + x_0, x_0 \in B_r(x_0)$. Namely $\overline{B_r(0)} \subset E_{2m}$. Finally, we may write

$$\|T\| = \sup_{\|x\|=1} \|Tx\| = \frac{1}{r} \sup_{\|x\|=1} \|T(rx)\| = \frac{1}{r} \sup_{\|x\|=r} \|Tx\| \leq \frac{2m}{r}.$$

Since this holds for any $T \in A$ we have thatsup$_{T\in A} \|T\| \leq \frac{2m}{r} < \infty$ as desired. □

2 Special cases: L^p and Hilbert Spaces

2.1 L^p Spaces

Definition 2.1. Let $1 \leq p \leq \infty$. Let $(X, \mathcal{M}, \mu)$ be a measure space.

- If $p < \infty$,

$$\mathcal{L}^p = \left\{ \psi : X \to \mathbb{C} \text{ such that } \psi \text{ is measurable and } \int |\psi|^p d\mu < \infty \right\}$$

and

$$\|\psi\|_p = \left(\int |\psi|^p d\mu \right)^{1/p}$$

- If $p = \infty$,

$$\mathcal{L}^\infty = \left\{ \psi : X \to \mathbb{C} \text{ such that } \psi \text{ is measurable and } \operatorname*{ess\,sup}_{x \in X} |\psi(x)| < \infty \right\}$$

and

$$\|\psi\|_\infty = \operatorname*{ess\,sup}_{x \in X} |\psi(x)| \qquad \text{where } \operatorname*{ess\,sup}_{x \in X} |\psi(x)| = \inf\{M \geq 0 : |\psi(x)| \leq M \mu\text{-a.e.}\}$$

Note that for any $n \in \mathbb{N}$, there is N_n with $\mu(N_n) = 0$ such for all $x \in N_n^c$

$$|\psi(x)| \leq \|\psi\|_\infty + \frac{1}{n}.$$

Let $N = \bigcup_{n \in \mathbb{N}} N_n$. Then N is null and $|\psi(x)| \leq \|\psi\|_\infty$ for all $x \in N^c$.

Example 2.2. For any $1 \leq p \leq \infty$, $\mathcal{L}^p$ is a vector space. Some examples:

(i) Let $n \in \mathbb{N}$, $X = \{1, 2, \ldots, n\}$, $\mathcal{M} = \mathcal{P}(X)$, and μ the counting measure. Then $\psi : X \to \mathbb{C}$ is identified with the vector $\begin{pmatrix} \psi(1) & \cdots & \psi(n) \end{pmatrix}^\top \in \mathbb{C}^n$ and

$$\int_X |\psi|^2 d\mu = \sum_{i=1}^n |\psi(i)|^2$$

is the squared Euclidean norm.

(ii) Let $X = \mathbb{N}$, $\mathcal{M} = \mathcal{P}(\mathbb{N})$, and μ the counting measure. We denote $\mathcal{L}^p(\mathbb{N}, \mathcal{P}(\mathbb{N}), \mu) = \ell^p$. Here $\psi : X \to \mathbb{C}$ is identified with the sequence $\{\psi_i\}_{i \in \mathbb{N}}$. And

$$\ell^p = \left\{ \{\psi_i\}_{i \in \mathbb{N}} : \sum_{i=1}^\infty |\psi_i|^p < \infty \right\} \qquad \text{and} \qquad \ell^\infty = \left\{ \{\psi_i\}_{i \in \mathbb{N}} : \sup_{i \in \mathbb{N}} |\psi_i| < \infty \right\}.$$

Note 2.3. In the definition of a norm, (i) fails for $\mathcal{L}^p$. The solution is to define the equivalence relation

$$\psi \sim \phi \text{ if } \psi = \phi \text{ a.e.}$$

Then $[\psi] = \{\phi \in \mathcal{L}^p : \phi \sim \psi\}$.

Definition 2.4. $L^p(X, \mathcal{M}, \mu)$ is the set $\{[\psi] : \psi \in \mathcal{L}^p(X, \mathcal{M}, \mu)\}$ equipped with the operations

- $[\psi] + [\phi] = [\psi + \phi]$
- $\alpha[\psi] = [\alpha\psi]$
- $\|[\psi]\|_p = \|\psi\|_p$

Lemma 2.5. The above operations are well-defined and $L^p(X, \mathcal{M}, \mu)$ is a normed vector space.

Proof.

- "+" is well defined: we need to show that

  $$(\psi_1 \sim \psi_2 \ \& \ \phi_1 \sim \phi_2) \implies \psi_1 + \phi_1 \sim \psi_2 + \phi_2.$$

  Let $N_\psi = \{x \in X : \psi_1(x) \leq \psi_2(x)\}$ and $N_\phi = \{x \in X : \phi_1(x) \neq \phi_2(x)\}$. Then $\mu(N_\psi) = \mu(N_\phi) = 0$, and

  $$\{x \in X : \psi_1(x) + \phi_1(x) \neq \psi_2(x) + \phi_2(x)\} \subset N_\psi \cup N_\phi.$$

  Finally, since $\mu(N_\psi \cup N_\phi) = 0$, we conclude $\psi_1 + \phi_1 \sim \psi_2 + \phi_2$ as desired.
- Similar for well-definedness of scalar multiplication
- The norm is well-defined since integrals of a.e. equal functions are equal
- The vector space axioms are immediate with $0 = [0]$ being the class of functions that are equal to 0 a.e.
- The first two norm axioms are quick to prove. The triangle inequality takes a bit of work and is called "Minkowski's Inequality" (see MATH507 HW10)

□

Proposition 2.6. Let $(X, \mathcal{M}, \mu)$ be a finite measure space. Let $\psi \in L^\infty$. Then $\psi \in L^p$ for all $1 \leq p \leq \infty$ and

$$\|\psi\|_\infty = \lim_{p\to\infty} \|\psi\|_p.$$

This is one of many results that essentially says: "the infinity norm is the limit of the p norm whenever it makes sense".

Proof. Let $X_r = \{x \in X : |\psi(x)| \geq r\}$. If $\mu(X_r) > 0$ then

$$\liminf_{p\to\infty} \|\psi\|_p \geq \liminf_{p\to\infty} \left(\int_{X_r} |\psi|^p d\mu\right)^{1/p} \geq r \liminf_{p\to\infty} \mu(X_r)^{1/p} = r$$

and

$$\limsup_{p\to\infty} \|\psi\|_p \leq \|\psi\|_\infty \limsup_{p\to\infty} \mu(X)^{1/p} = \|\psi\|_\infty.$$

Now pick $r = \|\psi\|_\infty - \varepsilon$ so that $\mu(X_r) > 0$ and so that $\liminf_{p\to\infty} \|\psi\|_p \geq \|\psi\|_\infty - \varepsilon$ for all $\varepsilon > 0$. Thus altogether

$$\|\psi\|_\infty \leq \liminf_{p\to\infty} \|\psi\|_p \leq \limsup_{p\to\infty} \|\psi\|_p \leq \|\psi\|_\infty.$$

So all inequalities are equalities and

$$\liminf_{p\to\infty} \|\psi\|_p = \limsup_{p\to\infty} \|\psi\|_p = \lim_{p\to\infty} \|\psi\|_p = \|\psi\|_\infty.$$

□

Theorem 2.7. *(Hölder's Inequality)* Let $1 \leq p, q \leq \infty$.

(i) If $\frac{1}{p} + \frac{1}{q} = 1$ and $\psi \in L^p$, $\phi \in L^q$, then $\psi\phi \in L^1$ and $\|\psi\phi\|_1 \leq \|\psi\|_p \|\phi\|_q$

(ii) If $1 \leq r < \infty$, $\frac{1}{p} + \frac{1}{q} = \frac{1}{r}$, and $\psi \in L^p$, $\phi \in L^q$, then $\psi\phi \in L^r$ and $\|\psi\phi\|_r \leq \|\psi\|_p \|\phi\|_q$

Proof.

(i) See HW 10

(ii) Follows from (i) applied to $|\psi|^r$, $|\phi|^r$ and $\frac{p}{r}$, $\frac{q}{r}$.

$$\|\psi\phi\|_r^r = \||\psi|^r |\phi|^r\|_1 \overset{(i)}{\leq} \||\psi|^r\|_{p/r} \||\phi|^r\|_{q/r} = \|\psi\|_p^r \|\phi\|_q^r.$$

□

Corollary 2.8. Let $1 \leq p < q \leq \infty$. If $\psi \in L^p \cap L^q$, then $\psi \in L^r$ for all $r \in [p, q]$ and

$$\|\psi\|_{p_\theta} \leq \|\psi\|_p^{1-\theta} \|\psi\|_q^\theta,$$

where $\frac{1}{p_\theta} = \frac{1-\theta}{p} + \frac{\theta}{q}$ for $\theta \in [0, 1]$.

Proof. $(q < \infty)$ By Hölder's Inequality:

$$\|\psi\|_{p_\theta}^{p_\theta} = \int_X |\psi|^{(1-\theta)p_\theta} |\psi|^{\theta p_\theta} d\mu = \||\psi|^{(1-\theta)p_\theta} |\psi|^{\theta p_\theta}\|_1 \leq \||\psi|^{(1-\theta)p_\theta}\|_{p/(p_\theta(1-\theta))} \||\psi|^{\theta p_\theta}\|_{q/\theta p_\theta}.$$

We should just check that $\frac{(1-\theta)p_\theta}{p} + \frac{\theta p_\theta}{q} = 1$ (true by assumption), and that

$$|\psi|^{(1-\theta)p_\theta} \in L^{p/(p_\theta(1-\theta))}, \qquad |\psi|^{\theta p_\theta} \in L^{q/\theta p_\theta}.$$

And indeed

$$\||\psi|^{(1-\theta)p_\theta}\|_{p/(p_\theta(1-\theta))} = \left(\int_X \left(|\psi|^{(1-\theta)p_\theta} \right)^{\frac{p}{(1-\theta)p_\theta}} \right)^{\frac{(1-\theta)p_\theta}{p}} = \|\psi\|_p^{(1-\theta)p_\theta},$$

and similarly

$$\||\psi|^{\theta p_\theta}\|_{q/\theta p_\theta} = \|\psi\|_q^{\theta p_\theta}$$

from which the claim follows. □

Note 2.9. One can prove a result of a similar flavour to corollary 2.8 which states: if $1 \leq p < q \leq \infty$ and $\mu(X) < \infty$, then $L^q \subset L^p$.

Theorem 2.10. *(Reisz-Fischer)* Let $(X, \mathcal{M}, \mu)$ be a measure space and $1 \leq p \leq \infty$. Then $L^p(X, \mathcal{M}, \mu)$ is a Banach space.

Proof. (Reisz-Fischer) From what we have already shown about $L^p(X, \mathcal{M}, \mu)$, it suffices to check completeness.

- Case $1 \leq p < \infty$: Let $\{\psi_i\}_{i\in\mathbb{N}}$ be absolutely convergent, namely $\sum_i \|\psi_i\|_p = M < \infty$. By Proposition 1.8 it suffices to show that $\sum_i \psi_i$ converges in L^p. Let $G_n = \sum_{i=1}^n |\psi_i(x)|$, this increases point-wise to $G = \sum_i |\psi_i|$ (it may be ∞). Now by the triangle inequality:

$$\|G_n\|_p \leq \sum_{i=1}^n \|\psi_i\|_p \leq M < \infty.$$

Then by the M.C.T., $G \in L^p$ and $\int_X G^p d\mu = \lim_{n\to\infty} \int |G_n|^p d\mu \leq M^p$. In particular $G(x)$ is finite μ-a.e. Hence there is a null set N such that the numerical series $\sum_i \psi_i(x)$ converges absolutely for $x \in N^c$. Now by completeness of $\mathbb{C}$:

$$S_n(x) = \sum_{i=1}^n \psi_i(x)\chi_{N^c}(x) \to S(x)$$

for all $x \in X$. Altogether $|S_n(x)-S(x)|^p \to 0$ as $N \to \infty$ and $|S_n(x)-S(x)|^p \leq (2G(x))^p \in L^1$, so the D.C.T. implies

$$\lim_{n\to\infty} \int_X |S_n(x) - S(x)|^p d\mu = 0.$$

Namely $S = \lim_{n\to\infty} S_n$ in L^p and so the series is convergent.

- Case $p = \infty$: Let $\{\psi_i\}_{i\in\mathbb{N}}$ be a Cauchy sequence in L^∞. Then, by definition of $\|\cdot\|_\infty$, for each $j, k \in \mathbb{N}$ there is a null set $N_{j,k}$ such that

$$|\psi_j(x) - \psi_k(x)| \leq \|\psi_j - \psi_k\|_\infty$$

for all $x \in N_{j,k}^c$. The set $N = \bigcup_{j,k\in\mathbb{N}} N_{j,k}$ is again a null set. Let $x \in N^c$. Then $\{\psi_i(x)\}_{i\in\mathbb{N}}$ is Cauchy and hence convergent (by the completeness of $\mathbb{C}$), say to $\psi(x)$. It follows that $\psi_j \to \psi$ uniformly on N^c. Namely $\psi_j \to \psi$ in L^∞.

□

2.2 The Dual of L^P

Unless otherwise stated, in this section we let $(X, \mathcal{M}, \mu)$ be a measure space. And we denote $L^p = L^p(X, \mathcal{M}, \mu)$, $\int f = \int_X f \, \mathrm{d}\mu$. Also we set $1 \leq p \leq \infty$ we define the **"conjugate exponent"** of p to be $1 \leq q \leq \infty$ such that $1/p + 1/q = 1$.

Proposition 2.11. Fix any $1 \leq q \leq \infty$ and let p be conjugate. For any $g \in L^q$ define $\phi_g : L^p \to \mathbb{C}$ which maps $f \mapsto \int gf$. Then $\phi_g \in (L^p)^*$ and $\|\phi_g\| \leq \|g\|_q$.

Proof. ϕ_g is linear because $\int$ is linear. ϕ_g is bounded by Hölder:

$$|\phi_g(f)| = \left|\int gf\right| \leq \int |gf| \leq \|g\|_q \|f\|_p.$$

In particular $\|\phi_g\| \leq \|g\|_q$. □

Definition 2.12. A measure μ on $(X, \mathcal{M})$ is **semi-finite** if for every $E \in \mathcal{M}$ with $\mu(E) = \infty$, there is $F \subset E$ in $\mathcal{M}$ such that $0 < \mu(F) < \infty$.

Proposition 2.13. Fix $1 \leq q \leq \infty$ and let p be conjugate. Let $g \in L^q$ and define ϕ_g as above. If either (i) $1 \leq q < \infty$ or (ii) $q = \infty$ and μ is semi-finite, then $\|\phi_g\| = \|g\|_q$.

Proof. We can assume that g is not 0 a.e. as this case is trivial. Otherwise it suffices to find an $f \in L^p$ with $\|f\|_p = 1$ such that $\left|\int fg\right| = \|g\|_q$ since recall that $\|\phi_g\| = \sup\left\{\left|\int fg\right| : \|f\|_p = 1\right\}$ and we have already shown that $\|\phi_g\| \leq \|g\|_q$.

- When $1 < q < \infty$, take
$$f(x) = \begin{cases} \frac{|g(x)|^{q-1}}{\|g\|_q^{q-1}} \frac{\overline{g(x)}}{|g(x)|} & g(x) \neq 0 \\ 0 & g(x) = 0 \end{cases}.$$
Since p and q are conjugate, we have that $p(q-1) = q$. Then
$$\|f\|_p^p = \int |f|^p = \frac{1}{\|g\|_q^{p(q-1)}} \int |g(x)|^{p(q-1)} \, \mathrm{d}\mu(x) = \frac{1}{\|g\|_q^q} \int |g(x)|^q \, \mathrm{d}\mu(x) = 1$$
and
$$\int fg = \frac{1}{\|g\|_q^{q-1}} |g(x)|^q \, \mathrm{d}\mu(x) = \|g\|_q.$$
- When $q = 1$ (i.e. $p = \infty$) take
$$f(x) = \begin{cases} \frac{\overline{g(x)}}{|g(x)|} & g(x) \neq 0 \\ 0 & g(x) = 0 \end{cases}.$$
Then $\|f\|_\infty = 1$ and $\int fg = \int |g| = \|g\|_1$.
- When $q = \infty$ (i.e. p=1) and μ is semi-finite, for every $\varepsilon > 0$ define
$$A_\varepsilon = A = \{x \in X : |g(x)| > \|g\|_\infty - \varepsilon\}.$$
Notice that $\mu(A) > 0$ (by the definition of $\|\cdot\|_\infty$) so by semi-finiteness there is $B \subset A$ such that $0 < \mu(B) < \infty$. Then define
$$f(x) = \begin{cases} \frac{\chi_B(x)}{\mu(B)} \frac{\overline{g(x)}}{|g(x)|} & g(x) \neq 0 \\ 0 & g(x) = 0 \end{cases}.$$
Then $\|f\|_1 = \frac{1}{\mu(B)} \int_B 1 = 1$ and
$$\int fg = \frac{1}{\mu(B)} \int_B |g| > \frac{1}{\mu(B)} \int_B \|g\|_\infty - \varepsilon = \|g\|_\infty - \varepsilon.$$
This shows that $\|\phi_g\| \geq \|g\|_\infty - \varepsilon$ for every $\varepsilon > 0$. Hence $\|\phi_g\| \geq \|g\|_\infty$ as desired.

□

Corollary 2.14. Let either (i) $1 \leq q < \infty$ or (ii) $q = \infty$ and μ be semi-finite. Take p to be the conjugate exponent. Then the map from $L^q \to (L^p)^*$ which sends $g \mapsto \phi_g$ is a linear isometry.

Proof. propositions 2.11 and 2.13 show exactly this. □

Theorem 2.15. Let either $1 < q < \infty$ or $q = \infty$ and μ be σ-finite. Take p to be the conjugate exponent. Then for each $\phi \in (L^p)^*$, there is a $g \in L^q$ such that $\phi = \phi_g$. Namely the above map is surjective, and hence is an isometric isomorphism (a bijective isometry) between L^q and $(L^p)^*$. [Note: being an isometry implies the map is injective for free.]

Proof. Let $\phi \in (L^p)^*$.

- **Case 1 ($\mu(X) < \infty$):** Notice that if there was such a g, then $\phi(\chi_E) = \int_E g \, \mathrm{d}\mu$. This looks like a Radon-Nikodym density $g = \frac{\mathrm{d}\nu}{\mathrm{d}\mu}$ for some complex measure ν which would require $\nu \ll \mu$. Following this idea, define $\nu(E) := \phi(\chi_E)$. We will show that ν as defined is a complex measure, and that $\nu \ll \mu$. Then we will show that $\frac{\mathrm{d}\nu}{\mathrm{d}\mu} =: g \in L^q$ and satisfies $\phi = \phi_g$.

  - ν is a complex measure: First of all $\nu(\emptyset) = \phi(0) = 0$ since ϕ is linear. Now let $\{E_m\}_{m\in\mathbb{N}}$ be disjoint and $E = \bigcup_{m=1}^{\infty} E_m$. We must show that $\nu(E) = \sum_{m=1}^{\infty} \nu(E_m)$ and that the series is absolutely convergent. Let

$$\alpha_k = \overline{\operatorname{sgn} \phi(\chi_{E_k})} \quad \text{where } \operatorname{sgn} z = \begin{cases} \frac{z}{|z|} & z \neq 0 \\ 0 & z = 0 \end{cases}$$

    and define $f_m = \sum_{k=1}^{m} \alpha_k \chi_{E_k}$ and $f = \sum_{k\in\mathbb{N}} \alpha_k \chi_{E_k}$. Note $f_m \in L^p$ since μ is finite. Now

$$\|f_m - f\|_p^p = \int \left| \sum_{k=m+1}^{\infty} \alpha_k \chi_{E_k} \right|^p \mathrm{d}\mu = \sum_{k=m+1}^{\infty} \int \chi_{E_k} \, \mathrm{d}\mu = \mu\left(\bigcup_{k=m+1}^{\infty} E_k \right) \quad (\star).$$

    This measure approaches 0 as $m \to \infty$ since $\mu(X) < \infty$. So $f_m \to f$ in L^p. Now notice that

$$\phi(f_m) = \phi\left(\sum_{k=1}^{m} \alpha_k \chi_{E_k} \right) = \sum_{m=1}^{k} \alpha_k \phi(\chi_{E_k}) = \sum_{k=1}^{m} |\phi(\chi_{E_k})|.$$

    So by continuity

$$\sum_{k\in\mathbb{N}} |\nu(E_k)| = \lim_{m\to\infty} \left| \sum_{k=1}^{m} |\phi(\chi_{E_k})| \right| = \lim_{m\to\infty} |\phi(f_m)| = |\phi(f)| \leq \|\phi\| \|f\|_p.$$

    Hence $\sum_{k=1}^{\infty} \nu(E_k)$ is absolutely convergent. Finally, using a similar argument as in $(\star)$ one can show that $\sum_{k=1}^{m} \chi_{E_k} \to \chi_E$ in L^p. Hence

$$\nu(E) = \phi(\chi_E) = \lim_{m\to\infty} \phi\left(\sum_{k=1}^{m} \chi_{E_k} \right) = \sum_{k=1}^{\infty} \nu(E_k)$$

    as desired.

  - $\nu \ll \mu$: If $\mu(E) = 0$, then $\chi_E = 0$ μ-a.e. Hence

$$|\nu(E)| = |\phi(\chi_E)| \leq \|\phi\| \underbrace{\|\chi_E\|_p}_{0}.$$

    Hence $\nu(E) = 0$.

  - $g := \frac{\mathrm{d}\nu}{\mathrm{d}\mu} \in L^q$: Radon-Nikodym guarantees the existence of $g := \frac{\mathrm{d}\nu}{\mathrm{d}\mu} \in L^1(\mu)$ so that

$$\phi(\chi_E) = \nu(E) = \int_E g \, \mathrm{d}\mu = \int \chi_E g \, \mathrm{d}\mu.$$

    Then by linearity, $\phi(f) = \int fg \, \mathrm{d}\mu$ for all simple functions f. Hence

$$\left| \int fg \, \mathrm{d}\mu \right| \leq \|\phi\| \|f\|_p \quad (\star\star)$$

    for all simple functions f. We state the following proposition without proof:

Proposition. Let p, q be conjugate. Let Σ be the set of all simple functions vanishing outside a set of finite measure. Suppose $g : X \to \mathbb{C}$ is measurable function such that: (i) $f \in \Sigma \implies gf \in L^1$, (ii)

$$M_q(g) = \sup\left\{\left|\int fg\right| : f \in \Sigma, \|f\|_p = 1\right\} < \infty,$$

and (iii) $\{x : g(x) \neq 0\}$ is σ-finite or μ is semi-finite. Then $g \in L^q$ and $\|g\|_q = M_q(g)$.

Since μ is finite, all (i)-(iii) are satisfied which implies $g \in L^q$ and when combined with $(\star\star)$ we have $\|g\|_q = M_q(g) \leq \|\phi\|$.

– $\phi = \phi_g$: For any $f \in L^p$ there is a sequence of simple functions $\{f_n\}_{n\in\mathbb{N}}$ in L^p such that $f_n \to f$ in L^p. Hence

$$\left|\phi(f) - \int fg\,\mathrm{d}\mu\right| \leq |\phi(f) - \phi(f_n)| + \left|\int (f_n - f)g\right| \leq \|\phi\|\|f - f_n\|_p + \|f_n - f\|_p\|g\|_q \to 0.$$

So $\phi(f) = \int fg\,\mathrm{d}\mu$ for all $f \in L^p$. That is $\phi = \phi_g$.

- **Case 2 (μ is σ-finite):** We can write $X = \bigcup_{n\in\mathbb{N}} E_n$ where $E_1 \subset E_2 \subset \cdots$ and $\mu(E_n) < \infty$. By case 1, there is $g_n \in L^q(E_n)$ such that $\phi(f) = \int_{E_n} fg_n\,\mathrm{d}\mu$ for all $f \in L^p(E_n)$. Then by uniqueness, $g_{n+1}|_{E_n} = g_n$ μ-a.e. So $g = \lim_{n\to\infty} g_n$ satisfies $g|_{E_n} = g_n$. We show that $g \in L^q$. If $q < \infty$ we apply the Monotone Convergence Theorem to $|g_n|^q$:

$$\|g\|_q^q = \int_X |g|^q\,\mathrm{d}\mu = \lim_{n\to\infty}\int_{E_n} |g_n|^q\,\mathrm{d}\mu \leq \|\phi\|^q < \infty.$$

If $q = \infty$, for any $n \in \mathbb{N}$ we have

$$\operatorname*{ess\,sup}_{x\in E_n} |g(x)| = \|g_n\|_\infty \leq \|\phi\|.$$

Since E_n is increasing and $\bigcup_n E_n = X$ we have that

$$\operatorname*{ess\,sup}_{x\in X} |g(x)| = \sup_n \operatorname*{ess\,sup}_{x\in E_n} |g(x)| \leq \|\phi\|.$$

Hence $g \in L^\infty$ and $\|g\|_\infty \leq \|\phi\|$. Now we show $\phi = \phi_g$. For any $f \in L^p$, $|f(1 - \chi_{E_n})|^p \to 0$ pointwise, and $|f(1 - \chi_{E_n})|^p \leq |f|^p \in L^1$. So by the Dominated Convergence Theorem:

$$\|f - f\chi_{E_n}\|_p^p = \int |f(1 - \chi_{E_n})|^p\,\mathrm{d}\mu \to 0.$$

So by continuity of ϕ and the Dominated Convergence Theorem applied to $|g_n f| \leq |gf| \in L^1$ (which follows from Hölder):

$$\phi(f) = \lim_{n\to\infty} \phi(f\chi_{E_n}) = \lim_{n\to\infty} \int g_n f\,\mathrm{d}\mu = \int gf\,\mathrm{d}\mu.$$

- **Case 3 (μ is arbitrary and $1 < q < \infty$):** See Folland.

□

2.3 Hilbert Spaces

"A Hilbert Space is a Banach Space with a norm induced by an inner product: $\|x\|^2 = \langle x, x\rangle$."

Definition 2.16. An **inner product** (on X, a vector space over $\mathbb{C}$) is a function

$$\langle \cdot, \cdot \rangle : X \times X \to \mathbb{C}$$

with the following properties:

(i) $\langle \alpha x + \beta y, z\rangle = \alpha\langle x, z\rangle + \beta\langle y, z\rangle$ for all $\alpha, \beta \in \mathbb{C}$, and $x, y, z \in X$

(ii) $\langle y, x\rangle = \overline{\langle x, y\rangle}$

(iii) $\langle x, x\rangle > 0$ whenever $x \neq 0$ and $\langle x, x\rangle = 0$ when $x = 0$

Notice:

$$\langle x, \alpha y + \beta z\rangle \overset{\text{(ii)}}{=} \overline{\langle \alpha y + \beta z, x\rangle} \overset{\text{(i)}}{=} \overline{\alpha\langle y, x\rangle + \beta\langle z, x\rangle} \overset{(ii)}{=} \overline{\alpha}\langle x, y\rangle + \overline{\beta}\langle x, z\rangle.$$

Example 2.17. *(Cauchy-Schwarz)* Given $x, z \in X$ with $z \neq 0$, define

$$\begin{aligned}
\phi(t) &= \|x - tz\|^2 \qquad t \in \mathbb{R} \\
&= \langle x - tz, x - tz\rangle \\
&= \langle x, x\rangle - t\langle x, z\rangle - t\langle z, x\rangle + \langle -tz, -tz\rangle \\
&= \|x\|^2 - t\left(\langle x, z\rangle + \overline{\langle z, x\rangle}\right) + t^2\|z\|^2
\end{aligned}$$

We can further write

$$\phi(t) = \|x\|^2 - 2t\mathrm{Re}(\langle x, z\rangle) + t^2\|z\|^2.$$

This is clearly non-negative as an upward facing parabola. Moreover $\phi(t) \geq \phi(t^\star)$ for all t where $t^\star$ is determined from

$$0 = \phi'(t) = -2\mathrm{Re}(\langle x, z\rangle) + 2t\|z\|^2 \implies t^\star = \frac{\mathrm{Re}(\langle x, z\rangle)}{\|z\|^2}.$$

And

$$\phi(t^\star) = \|x\|^2 - \frac{(\mathrm{Re}(\langle x, z\rangle))^2}{\|z\|^2}.$$

This ϕ is non-negative we obtain

$$|\mathrm{Re}(\langle x, z\rangle)| \leq \|x\|\|z\|.$$

This holds for all $x, z \in X$ (the case when $z = 0$ is trivial). Notice we have equality if and only if $x = t^\star z$, that is, x is a scalar multiple of z. We can upgrade this further. Let $y \in X$ be such that $\langle x, y\rangle \neq 0$. Take $z = \sigma y$ for some $\sigma \in \mathbb{C}$. Then

$$\langle x, z\rangle = \langle x, \sigma y\rangle = \overline{\sigma}\langle x, y\rangle.$$

Taking $\sigma = \frac{\langle x,y\rangle}{|\langle x,y\rangle|}$ yields

$$\langle x, \sigma y\rangle = \left(\frac{\overline{\langle x, y\rangle}}{|\langle x, y\rangle|}\right)\langle x, y\rangle = |\langle x, y\rangle|.$$

Hence

$$|\langle x, y\rangle| \leq \|x\|\|\sigma y\| = \|x\|\|y\|$$

since σ has unit norm. This is valid for arbitrary $x, y \in X$ (again the case when $\langle x, y\rangle = 0$ is trivial). Moreover, we have equality if and only if $x = t^\star(\sigma y) = (t^\star\sigma)y$, that is, x is a complex scalar multiple of y. This is the **Cauchy-Schwarz Inequality**.

Proposition 2.18. In this setup, $\|x\| = \sqrt{\langle x, x\rangle}$ is truly a norm.

Proof.

- $\|x\| \geq 0$ for all $x \in X$ and with equality if and only if $x = 0$
- $\|\alpha x\|^2 = \langle \alpha x, \alpha x\rangle = \alpha\langle x, \alpha x\rangle = \alpha\overline{\alpha}\langle x, x\rangle = |\alpha|^2\|x\|^2$
- We note that $\|x + y\|^2$ is ϕ from Example 2.17 evaluated at -1

$$\begin{aligned}\|x + y\|^2 &= \|x\|^2 + 2\mathrm{Re}(\langle x, y\rangle) + \|y\|^2 \\ &\leq \|x\|^2 + 2|\langle x, y\rangle| + \|y\|^2 \\ &\leq \|x\|^2 + 2\|x\|\|y\| + \|y\|^2 \\ &= (\|x\| + \|y\|)^2\end{aligned}$$

□

Definition 2.19. A **Hilbert Space** is a pair $(X, \langle\cdot,\cdot\rangle)$ where X is a complex vector space, $\langle\cdot,\cdot\rangle$ is an inner product on X, and the norm $\|x\| = \sqrt{\langle x, x\rangle}$ makes $(X, \|\cdot\|)$ a Banach Space.

Proposition 2.20. *(Parallelogram Laws)* For all $x, y \in X$ we have

$$\|x + y\|^2 + \|x - y\|^2 = 2(\|x\|^2 + \|y\|^2).$$

Proof. Evaluating ϕ from Example 2.17 at ± 1 yields

$$\|x + y\|^2 = \|x\|^2 + 2\mathrm{Re}(\langle x, y\rangle) + \|y\|^2 \qquad \text{and} \qquad \|x - y\|^2 = \|x\|^2 - 2\mathrm{Re}(\langle x, y\rangle) + \|y\|^2.$$

Adding these together yields

$$\|x + y\|^2 + \|x - y\|^2 = 2(\|x\|^2 + \|y\|^2)$$

as desired. □

Example 2.21.

(i) For any $n \geq 1$, let $\mathbb{C}^n = \{(z_1, z_2, \ldots, z_n) : z_k \in \mathbb{C}\}$ and define $\langle y, z\rangle = \sum_{k=1}^n y_k\overline{z_k}$.

(ii) Extend to $\mathbb{C}^{\mathbb{N}} = \{(z_1, z_2, \ldots) : z_k \in \mathbb{C}\}$ with $\langle y, z\rangle = \sum_{k=1}^\infty y_k\overline{z_k}$. But this series may diverge! So let $\ell^2 \subset \mathbb{C}^{\mathbb{N}}$ be the set of complex sequences z for which

$$\|z\|^2 = \sum_{k=1}^\infty z_k\overline{z_k} = \sum_{k=1}^\infty |z_k|^2$$

is convergent.

(iii) Let $(\Omega, \mathcal{M}, \mu)$ be a measure space, then $L^p(\Omega, \mathcal{M}, \mu)$ with inner product $\langle f, g\rangle = \int_\Omega f\overline{g}\,\mathrm{d}\mu$ works. [One can view ℓ^2 as a special case of this with $\Omega = \mathbb{N}$ and μ the counting measure.]

Proposition 2.22. Let $X = L^p(\Omega, \mathcal{M}, \mu)$ where $1 \leq p \leq \infty$, $\mu(\Omega) > 0$ and $\mu(\{x\}) = 0$ for all $x \in \Omega$. Then X endowed with $\|\cdot\|_p$ is a Hilbert space if and only if $p = 2$.

Proof. One way to detect if there exists an inner product compatible with a given norm $\|\cdot\|$ is to see if $\|\cdot\|$ satisfies the parallelogram law. Choose $E, F \in \mathcal{M}$ with $m := \mu(E) = \mu(F)$ and $E \cap F = \emptyset$. Such E and F always exist given our assumptions. Define $f = \chi_E$ and $g = \chi_F$. When $1 \leq p < \infty$:

$$\|f \pm g\|_p^p = \int_\Omega |f \pm g|^p \, \mathrm{d}\mu = \int_{E \cup F} |f \pm g|^p \, \mathrm{d}\mu = \mu(E \cup F) = 2m.$$

While on the other hand $\|f\|_p^p = \int_\Omega |f|^p \, \mathrm{d}\mu = \mu(E) = m$ and similarly $\|g\|_p^p = m$. Hence

$$\|f + g\|_p^2 + \|f - g\|^2 = (2m)^{2/p} + (2m)^{2/p} = 2(2)^{2/p}(m)^{2/p}$$

while

$$2(\|f\|_p^2 + \|g\|_p^2) = 2(m^{2/p} + m^{2/p}) = 2(2)(m)^{2/p}.$$

These two expressions are equal if and only if $p = 2$. In the case where $p = \infty$, note that $\|f \pm g\|_\infty = \|f\|_\infty = \|g\|_\infty = 1$ and that $1 + 1 \neq 2 + 2$. So the only L^p space that is also a Hilbert space is L^2 and the norm satisfies $\|f\|_2^2 = \int_\Omega |f|^2 \, \mathrm{d}\mu = \int_\Omega f\overline{f} \, \mathrm{d}\mu$. □

Definition 2.23. A set $S \subset X$ is **convex** if $s_t := s_0 + t(s_1 - s_0) \in S$ for all $s_0, s_1 \in S$ and $t \in [0, 1]$.

Theorem 2.24. Let X be a Hilbert space. Let $S \subset X$ be non-empty, closed, and convex. Then given $x \notin S$, there exists a unique $\hat{s} \in S$ where $\|x - \hat{s}\| \leq \|x - s\|$ for all $s \in S$. In other words, there exists a unique $\hat{s}$ which is the nearest point in S to the given x.

Proof. Denote $d_S(x) = \inf\{\|x - s\| : s \in S\}$. For each $k \in \mathbb{N}$, let $r_k = d_S(x) + \frac{1}{k}$. By properties of infimum there exists $s_k \in S$ such that $\|x - s_k\| < r_k$. Now for arbitrary $m, n \in \mathbb{N}$ apply the parallelogram law to $(x - s_m)$ and $(x - s_n)$:

$$\|(x - s_m) + (x - s_n)\|^2 + \|(x - s_m) - (x - s_n)\|^2 = 2\left(\|x - s_m)^2 + \|x - s_n\|^2\right)$$

which then becomes

$$\|s_n - s_m\|^2 = 2\|x - s_m\|^2 + 2\|x - s_n\|^2 - 4\left\|x - (s_m + s_n)/2\right\|^2.$$

By convexity, $(s_m + s_n)/2 \in S$ so $\|x - (s_m + s_n)/2\| \geq d_S(x)$. Altogether we have that

$$\|s_n - s_m\|^2 \leq 2r_m^2 + 2r_n^2 - 4d_S(x)^2.$$

Now $r_k \to d_S(x)$ as $k \to \infty$ so the RHS $\to 0$ as $m, n \to \infty$. Namely $\{s_k\}_{k \in \mathbb{N}}$ is Cauchy. In particular, as X is assumed Hilbert (hence complete), there exists $\hat{s} = \lim_{k \to \infty} s_k$. Firstly $\hat{s} \in S$ since S is closed. Moreover,

$$d_S(x) \leq \|x - \hat{s}\| = \lim_{k \to \infty} \|x - s_k\| \leq \lim_{k \to \infty} r_k = d_S(x),$$

so $\|x - \hat{s}\| = d_S(x)$ as desired. It remains to check uniqueness. Suppose that $\hat{s}_0, \hat{s}_1 \in S$ both satisfy $d_S(x) = \|x - \hat{s}_j\|$. We can again apply the parallelogram law to, this time to $(x - \hat{s}_0)$ and $(x - \hat{s}_i)$, to obtain that

$$\|\hat{s}_0 - \hat{s}_1\|^2 \leq 2\hat{r}_0^2 + 2\hat{r}_1^2 - 4d_S(x)^2$$

for any $\hat{r}_j \geq \|x - \hat{s}_j\|$. Choose $\hat{r}_j = d_S(x)$ to obtain $\|\hat{s}_0 - \hat{s}_1\| \leq 0$ so that $\hat{s}_1 = \hat{s}_0$. □

Theorem 2.25. Let X be a Hilbert space. Let $S \subset X$ be non-empty, closed, and convex. Let $\hat{s} \in S$ and $x \in X$. The following are equivalent.

(i) $\|x - \hat{s}\| \leq \|x - s\|$ for all $s \in S$

(ii) $\mathrm{Re}\langle x - \hat{s}, s - \hat{s}\rangle \leq 0$ for all $s \in S$

Proof. First a generic calculation (for any $s \in X$):

$$\|x - s\|^2 = \|x - \hat{s} + \hat{s} - s\|^2 = \|x - \hat{s}\|^2 + 2\underbrace{\mathrm{Re}\langle x - \hat{s}, \hat{s} - s\rangle}_{-\mathrm{Re}\langle x-\hat{s}, s-\hat{s}\rangle} + \|\hat{s} - s\|^2.$$

First (ii) $\implies$ (i). The generic calculation implies

$$\|x - s\|^2 \geq \|x - \hat{s}\|^2 + 0 + \|\hat{s} - s\|^2 \geq \|x - \hat{s}\|^2.$$

Now (i) $\implies$ (ii). Let $s \in S$ be arbitrary. By convexity, $s_t = \hat{s} + t(s - \hat{s}) \in S$ for all $t \in (0, 1]$. By applying (i) and then the generic calculation:

$$\|x - \hat{s}\|^2 \overset{\text{(i)}}{\leq} \|x - s_t\|^2 = \|x - \hat{s}\|^2 - 2\mathrm{Re}\langle x - \hat{s}, s_t - \hat{s}\rangle + \|\hat{s} - s_t\|^2.$$

Then since $s_t - \hat{s} = t(s - \hat{s})$ we can write

$$\|x - \hat{s}\|^2 \leq \|x - \hat{s}\|^2 - 2\mathrm{Re}\langle x - \hat{s}, t(s - \hat{s})\rangle + \|t(s - \hat{s})\|^2.$$

Since $t > 0$ we can then write

$$\mathrm{Re}\langle x - \hat{s}, s - \hat{s}\rangle \leq \frac{t}{2}\|s - \hat{s}\|^2.$$

Taking the limit as $t \to 0^+$ yields $\mathrm{Re}\langle x - \hat{s}, s - \hat{s}\rangle \leq 0$ as desired. □

As a special case: if S is a closed linear subspace of X we can improve (ii). Pick an arbitrary $s' \in S$ and let $\sigma = \langle x - \hat{s}, s'\rangle \in \mathbb{C}$. Apply (ii) with $s = \hat{s} + \sigma s'$. Then

$$0 \geq \mathrm{Re}\langle x - \hat{s}, \sigma s'\rangle = \mathrm{Re}\overline{\sigma}\langle x - \hat{s}, s'\rangle = \mathrm{Re}|\sigma|^2 = |\sigma|^2.$$

Since the complex modulus is non-negative, $\sigma = 0$. That is $\langle x - \hat{s}, s'\rangle = 0$ for any $s' \in S$.

Definition 2.26. We call $v, w \in X$ **orthogonal** when $\langle v, w\rangle = 0$.

Theorem 2.27. If E is a closed linear subspace of X with $\hat{e} \in E$ and $x \in X$. The following are equivalent.

(i) $\|x - \hat{e}\| \leq \|x - e\|$ for all $e \in E$

(ii) $\langle x - \hat{e}, e\rangle = 0$ for all $e \in E$

Moreover, for any $x \in X$, there exists a unique $\hat{e} \in E$ satisfying (i)-(ii).

Proof. See above. □

Definition 2.28. For any non-empty $S \subset X$ the **orthogonal complement** of S is defined to be

$$S^{\perp} = \{x \in X : \langle x, s\rangle = 0 \text{ for all } s \in S\}$$

($S^{\perp}$ is also pronounced "S perp").

One can show that $S^\perp$ is a closed linear subspace of X, and that $(S^\perp)^\perp \supset S$ and is the smallest closed linear subspace containing S.

Example 2.29. Let E be a closed subspace of X and $x \in E$. Let x_E be the nearest point in E to x. Define $x^\perp = x - x_E$. By Theorem 2.27 (ii) we have that $\langle x - x_E, e\rangle = 0$ for all $e \in E$ and so $x^\perp \in E^\perp$. For any $y \in E^\perp$ we have $y - x^\perp \in E^\perp$ since $E^\perp$ is a subspace. So we have $0 = \langle x_E, y - x^\perp\rangle = \langle x - x^\perp, y - x^\perp\rangle$. Hence again by Theorem 2.27 (ii) we have that $x^\perp$ is the nearest point of $E^\perp$ to x. In particular, we can uniquely split $x = x_E + x^\perp$ and write

$$X = E \oplus E^\perp = \{x_E + x^\perp : x_E \in E, x^\perp \in E^\perp\}.$$

Now we seek to describe the dual space X^* when X is a Hilbert space. First a useful lemma.

Lemma 2.30. For $\phi \in X^*$ and $x, z \in X$,

$$\phi(x)z - \phi(z)x \in \ker(\phi).$$

Proof. By linearity $\phi(\phi(x)z - \phi(z)x) = \phi(x)\phi(z) - \phi(z)\phi(x) = 0$. □

Example 2.31. First pick any $u \in X$ and let $f_u = \langle x, u\rangle$. Clearly f_u is linear. And if $u \neq 0$ then by Cauchy Schwarz

$$\frac{|f_u(x)|}{\|x\|} = \frac{|\langle x, u\rangle|}{\|x\|} \leq \frac{\|x\|\|u\|}{\|x\|} \qquad \text{for all } x \neq 0.$$

So $\|f_u\|_* \leq \|u\|$, hence $f_u \in X^*$. Moreover $\frac{f_u(u)}{\|u\|} = \frac{\langle u,u\rangle}{\|u\|} = \frac{\|u\|^2}{\|u\|}$ so in fact $\|f_u\|_* = \|u\|$. This shows that an entire copy of X lives in X^*. What about its other elements? Pick any $\phi \in X^\star \setminus \{0\}$. Let $E = \ker(\phi)$ (which is a closed proper subspace of X since ϕ is not the 0 map). In particular $E^\perp \neq \{0\}$ and we can pick $z \in E^\perp$ such that $\|z\| = 1$. By the above lemma any arbitrary $x \in X$ yields $\phi(x)z - \phi(z)x \in E$. And since $z \in E^\perp$ we have

$$0 = \langle \phi(x)z - \phi(z)x, z\rangle = \phi(x)\underbrace{\langle z, z\rangle}_{\|z\|^2=1} - \phi(z)\langle x, z\rangle.$$

In particular

$$\phi(x) = \phi(z)\langle x, z\rangle = \langle x, \overline{\phi(z)}z\rangle.$$

So ϕ has the form f_u with $u = \overline{\phi(z)}z$. Namely

$$X^* = \{f_u : u \in X\}$$

and $\|f_u\|_* = \|u\|$ for all $u \in X$. Moreover one can check that

$$\langle f_u, f_w\rangle = \langle w, u\rangle$$

is an inner product on X^*. This is the *Riesz Representation Theorem for Hilbert Spaces.*

Proposition 2.32. *(Pythagorean Theorem)* Let $x_1, \ldots, x_m \in \mathcal{H}$ a Hilbert space with $x_j \perp x_k$ for all $j \neq k$. Then $\|\sum_{k=1}^m\|^2 = \sum_{k=1}^m \|x_k\|^2$.

Proof. $\|\sum_{k=1}^m x_k\|^2 = \langle\sum_{k=1}^m x_k, \sum_{j=1}^m x_j\rangle = \sum_{k=1}^m\langle x_k, x_k\rangle = \sum_{k=1}^m \|x_k\|^2$. □

Definition 2.33. A subset $\{u_\alpha\}_{\alpha\in\mathcal{A}}$ of a Hilbert space $\mathcal{H}$ is said to be **orthonormal** if $\|u_\alpha\| = 1$ for all $\alpha \in \mathcal{A}$ and $\langle u_\alpha, u_\beta\rangle = 0$ for $\alpha \neq \beta$.

Proposition 2.34. *(Bessel Inequality)* Let $\{u_\alpha\}_{\alpha\in\mathcal{A}}$ in a Hilbert space $\mathcal{H}$ be orthonormal. Then for all $x \in \mathcal{H}$,

$$\sum_{\alpha\in\mathcal{A}} |\langle x, u_\alpha\rangle|^2 \leq \|x\|^2.$$

Note that for $\mathbb{R} \ni a_\alpha \geq 0$ we define

$$\sum_{\alpha\in\mathcal{A}} a_\alpha := \sup\left\{\sum_{\alpha\in F} a_\alpha : F \subset \mathcal{A} \text{ is finite}\right\}.$$

And if $\sum_{\alpha\in\mathcal{A}} a_\alpha < \infty$, then $\{\alpha \in \mathcal{A} : a_\alpha > 0\}$ is countable.

Proof. Let $F \subset \mathcal{A}$ be finite. Then

$$\begin{aligned}
0 &\leq \left\langle x - \sum_{\alpha\in F}\langle x, u_\alpha\rangle u_\alpha, x - \sum_{\alpha\in F}\langle x, u_\alpha\rangle u_\alpha \right\rangle \\
&= \|x\|^2 - \sum_{\alpha\in F}\langle x, u_\alpha\rangle\langle u_\alpha, x\rangle - \sum_{\alpha\in F}\overline{\langle x, u_\alpha\rangle}\langle x, u_\alpha\rangle + \left\|\sum_{\alpha\in F}\langle x, u_\alpha\rangle u_\alpha\right\|^2 \\
&= \|x\|^2 - 2\sum_{\alpha\in F}|\langle x, u_\alpha\rangle|^2 + \sum_{\alpha\in F}|\langle x, u_\alpha\rangle|^2 \underbrace{\|u_\alpha\|^2}_{1} \\
&= \|x\|^2 - \sum_{\alpha\in F}|\langle x, u_\alpha\rangle|^2
\end{aligned}$$

So $\sum_{\alpha\in F}|\langle x, u_\alpha\rangle|^2 \leq \|x\|^2$ for any finite $F \subset \mathcal{A}$. Taking sup over all such F completes the proof. □

Theorem 2.35. Let $\{u_\alpha\}_{\alpha\in\mathcal{A}}$ in a Hilbert space $\mathcal{H}$ be orthonormal. The following are equivalent.

(i) *Completness:* $\langle x, u_\alpha\rangle = 0$ for all $\alpha \in \mathcal{A}$ implies $x = 0$

(ii) *Parseval Identity:* $\|x\|^2 = \sum_{\alpha\in\mathcal{A}}|\langle x, u_\alpha\rangle|^2$ for all $x \in \mathcal{H}$

(iii) $x = \sum_{\alpha\in\mathcal{A}}\langle x, u_\alpha\rangle u_\alpha$ (and at most countably many $\langle x, u_\alpha\rangle$ are non-zero) for all $x \in \mathcal{H}$

If (i)-(iii) hold we say that $\{u_\alpha\}_{\alpha\in\mathcal{A}}$ is an orthonormal basis of $\mathcal{H}$.

Note that (i) means $\mathcal{H} = \overline{\text{span}\{u_\alpha : \alpha \in \mathcal{A}\}}$. Indeed, define $M := \overline{\text{span}\{u_\alpha : \alpha \in \mathcal{A}\}}$, so that M is a closed subspace of $\mathcal{H}$. Then $\mathcal{H}$ can be decomposed as $\mathcal{H} = M \oplus M^\perp$. Then one can use (i) to show that $M^\perp = \{0\}$.

Proof.

- (ii) $\implies$ (i): Suppose $\langle x, u_\alpha\rangle = 0$ for all $\alpha \in \mathcal{A}$. Then by (ii)

$$\|x\|^2 = \sum_{\alpha\in\mathcal{A}}|\langle x, u_\alpha\rangle|^2 = \sum_{\alpha\in\mathcal{A}} 0 \implies x = 0.$$

- (i) $\implies$ (iii): Fix $x \in \mathcal{H}$. Then by Bessel's Inequality $\sum_{\alpha\in\mathcal{A}}|\langle x, u_\alpha\rangle|^2 \leq \|x\|^2 < \infty$. So $\{\alpha \in \mathcal{A} : |\langle x, u_\alpha\rangle| \neq 0\}$ is countable and we can index them as $\{\alpha_j\}_{j\in\mathbb{N}}$. For $n < m$ we have by Pythagoras and $\|u_\alpha\| = 1$ that

$$\left\|\sum_{j=n}^{m}\langle x, u_{\alpha_j}\rangle u_{\alpha_j}\right\|^2 = \sum_{j=n}^{m}|\langle x, u_{\alpha_j}\rangle|^2 \to 0 \qquad \text{as } n, m \to \infty.$$

Namely the partial sums are Cauchy and therefore converge to $z \in \mathcal{H}$. Define

$$y = x - z = x - \sum_{j=1}^{\infty} \langle x, u_{\alpha_j} \rangle u_{\alpha_j}.$$

Then for any $\alpha \in \mathcal{A}$:

$$\langle y, u_\alpha \rangle = \langle x, u_\alpha \rangle - \sum_{j=1}^{\infty} \langle x, u_{\alpha_j} \rangle \langle u_{\alpha_j}, u_\alpha \rangle = \langle x, u_\alpha \rangle - \langle x, u_\alpha \rangle = 0.$$

And so by (i) we have $y = 0$, i.e.

$$x = \sum_{j=1}^{\infty} \langle x, u_{\alpha_j} \rangle u_{\alpha_j} = \sum_{\alpha \in \mathcal{A}} \langle x, u_\alpha \rangle u_\alpha.$$

- (iii) $\Longrightarrow$ (ii): Let $S_m = \sum_{j=1}^{m} \langle x, u_{\alpha_j} \rangle u_{\alpha_j}$. First of all

$$\begin{aligned} \|x - S_m\|^2 &= \|x\|^2 - 2\mathrm{Re}\langle x, S_m \rangle + \|S_m\|^2 \\ &= \|x\|^2 - 2\mathrm{Re} \sum_{j=1}^{m} \langle x, u_{\alpha_j} \rangle \overline{\langle x, u_{\alpha_j} \rangle} + \sum_{j=1}^{m} |\langle x, u_{\alpha_j} |^2 \\ &= \|x\|^2 = 2 \sum_{j=1}^{m} |\langle x, u_{\alpha_j} \rangle|^2 + \sum_{j=1}^{m} |\langle x, u_{\alpha_j} \rangle|^2 \\ &= \|x\|^2 - \sum_{j=1}^{m} |\langle x, u_{\alpha_j} \rangle|^2 \end{aligned}$$

And by (iii) we have

$$\Big\| x - \underbrace{\sum_{j=1}^{m} \langle x, u_{\alpha_j} \rangle u_{\alpha_j}}_{S_m} \Big\|^2 \to 0 \qquad \text{as } m \to \infty.$$

So we obtain (ii).

□

Note 2.36. A few remarks.

- Using Zorn's Lemma (see the proof of Hahn-Banach) one can show that every Hilbert space has an orthonormal basis.
- A Hilbert space is *separable* (i.e. it contains a countable dense subset) if and only if it has a countable orthonormal basis.
- Every separable Hilbert space is isomorphic to ℓ^2.

3 Topological Spaces

3.1 Point Set Topology

Definition 3.1. Let $X \neq \emptyset$, then $\mathbb{T} \subset \mathcal{P}(X)$ is called a **topology** if

- $\emptyset, X \in \mathbb{T}$
- $\mathbb{T}$ is closed under arbitrary unions
- $\mathbb{T}$ is closed under finite intersections

The elements of $\mathbb{T}$ are called **opens sets** and the pair $(X, \mathbb{T})$ is a **topological space**.

Example 3.2.

- $\{\emptyset, X\}$ is a topology on X
- $\mathcal{P}(X)$ is a topology on X
- When (X, d) is a metric space, and $B_r(x) := \{y \in X : d(y, x) < r\}$ then
$$\mathbb{T} = \{U \in \mathcal{P}(X) : \text{ for all } x \in U \text{ there is } r > 0 \text{ such that } B_r(x) \subset U\}$$
is a topology on X

Definition 3.3. Let $(X, \mathbb{T})$ be a topological space.

- A set $F \subset X$ is said to be **closed** if $F^c \in \mathbb{T}$. That is, if its complement is open.
- Let $Y \subset X$. The **relative topology on** Y is $\mathbb{T}_Y = \{U \cap Y : U \in \mathbb{T}\}$. It is easy to check that this defines a topology on Y.
- A **basis** of $\mathbb{T}$ is a family $\mathcal{B} \subset \mathbb{T}$ such that for any non-empty $U \in \mathbb{T}$ there is $\{B_\alpha\}_{\alpha \in \mathcal{A}}$ in $\mathcal{B}$ such that $U = \bigcup_{\alpha \in \mathcal{A}} B_\alpha$.
- $S \subset \mathbb{T}$ is called a **sub basis of** $\mathbb{T}$ if $\mathbb{T}$ is the smallest topology containing S. In this case: for all open sets $U \in \mathbb{T}$, U can be written as a union of finite intersections of sets in S.

Definition 3.4. Let $(X, \mathbb{T})$ be a topological space. Let $A \subset X$.

- The **interior** of A is
$$A^\circ = \bigcup_{\substack{U \in \mathbb{T} \\ U \subset A}} U$$
- The **closure** of A is
$$\overline{A} = \bigcap_{\substack{F \text{ closed} \\ F \supset A}} F$$
- The **boundary** of A is
$$\partial A = \overline{A} \setminus A^\circ$$

Definition 3.5. Let $(X, \mathbb{T})$ be a topological space.

- A neighbourhood of $x \in X$ is any $N \subset X$ such that $x \in N^\circ$. [Note N need not be open!]
- A family $\mathcal{N}_x$ of subsets of X is a **neighbourhood basis at** x if each $N \in \mathcal{N}_x$ is a neighbourhood of x, and for any M a neighbourhood of x, there is $N \in \mathcal{N}_x$ such that $x \in N \subset M$
- We say that $\{x_n\}_{n \in \mathbb{N}}$ **converges** to x if for any N a neighbourhood of x, there is $n_0 \in \mathbb{N}$ such that $x_n \in N$ for all $n \geq n_0$

Definition 3.6. Let $(X, \mathbb{T}_X)$ and $(Y, \mathbb{T}_Y)$ be topological spaces. We say $f : X \to Y$ is **continuous** if $f^{-1}(V) \in \mathbb{T}_X$ for all $V \in \mathbb{T}_Y$.

Note 3.7. Let $X \neq \emptyset$. Topologies on X are partially ordered by

$$\mathbb{T} \prec \mathbb{T}' \iff \mathbb{T} \subset \mathbb{T}'$$

in which case we say $\mathbb{T}'$ is **stronger** than $\mathbb{T}$ or that $\mathbb{T}$ is **weaker** than $\mathbb{T}'$. Notice that $\{\emptyset, X\}$ is the weakest topology and that $\mathcal{P}(X)$ is the strongest. Some topologies are not comparable. For example on $\mathbb{R}$ one can define $\mathbb{T} = \{\emptyset, \mathbb{R}, (-\infty, 0)\}$ and $\mathbb{T}' = \{\emptyset, \mathbb{R}, [1, \infty]\}$

Definition 3.8. The topology **generated** by $\mathcal{E} \subset \mathcal{P}(X)$ is the smallest (weakest) topology containing $\mathcal{E}$. That is,

$$\mathbb{T}(\mathcal{E}) = \bigcap_{\substack{\mathbb{T} \text{ topology} \\ \mathbb{T} \supset \mathcal{E}}} \mathbb{T}.$$

Equivalently $\mathbb{T}(\mathcal{E}) = \{\emptyset, X\} \cup \{\text{unions of finite intersections of elements of } \mathcal{E}\}$.

Definition 3.9. Let $(X, \mathbb{T}_X)$ and $\{(Y_\alpha, \mathbb{T}_\alpha)\}_{\alpha \in \mathcal{A}}$ be topological spaces and let $\{f_\alpha\}_{\alpha \in \mathcal{A}}$ be a collection of functions $f_\alpha : X \to Y_\alpha$. The **weak topology generated by** $\{f_\alpha\}_{\alpha \in \mathcal{A}}$ is the weakest topology on X so that all f_α are continuous; i.e. it is the topology generated by

$$\mathcal{E} = \{f_\alpha^{-1}(U_\alpha) : U_\alpha \in \mathbb{T}_\alpha, \alpha \in \mathcal{A}\}.$$

Definition 3.10. Let $\{(X_\alpha, \mathbb{T}_\alpha)\}_{\alpha \in \mathcal{A}}$ be a collection of topological spaces and let $X = \prod_{\alpha \in \mathcal{A}} X_\alpha$.

- The **box topology** on X is generated by the family

$$\mathcal{B}_{\text{BOX}} = \left\{ \prod_{\alpha \in \mathcal{A}} U_\alpha : U_\alpha \in \mathbb{T}_\alpha \right\}.$$

  But the box topology has some problems with $\mathcal{A}$ is infinite.

- We define the **product topology** on X as the weakest topology such that the projection maps $\pi_\alpha : X \to X_\alpha$ [which map $(x_\beta)_{\beta \in \mathcal{A}} \mapsto x_\alpha$] are continuous for all $\alpha \in \mathcal{A}$. Equivalently, it is the topology generated by

$$\mathcal{E} = \{\pi_\alpha^{-1}(U_\alpha) : U_\alpha \in \mathbb{T}_\alpha, \alpha \in \mathcal{A}\} = \left\{ \prod_{\alpha \in \mathcal{A}} U_\alpha : U'_\alpha \in \mathbb{T}'_\alpha \text{ and } U_\beta = X_\beta \text{ for all } \beta \neq \alpha' \right\}.$$

Example 3.11. *(Why the box topology is bad)*
Let $X = \mathbb{R}^{\mathbb{N}} = \prod_{n \in \mathbb{N}} \mathbb{R}$ with the usual topology on $\mathbb{R}$. An element $x \in X$ looks like $x = (x_1, x_2, \dots)$. Define the sequence $\{x^{(m)}\}_{m \in \mathbb{N}}$ in X with

$$x^{(m)} = (\underbrace{0, 0, \dots, 0}_{m \text{ zeroes}}, 1, 1, \dots), \quad \text{that is } x_k^{(m)} = \begin{cases} 0 & k \leq m \\ 1 \text{otherwise} \end{cases}.$$

We expect that $x^{(m)}$ should converge to the sequence 0. But under the box topology we have the open set $U = \prod_{k \in \mathbb{N}} (-1/2, 1/2)$ containing the sequence 0, and yet for any $m \in \mathbb{N}$ we have $x_k^{(m)} = 1 \notin (-1/2, 1/2)$ for $k \geq m$ so $x^{(m)} \notin U$. In particular $x^{(m)}$ does not converge to the sequence 0. In fact, one can show that $x^{(m)}$ does not converge to any $\{x_k\}_{k \in \mathbb{N}} = x \in X$ with the set $U = \prod_{k \in \mathbb{N}} (x_k - 1/2, x_k + 1/2)$.

Note 3.12. *(More remarks on the product topology)*

- A basis for the product topology is given by the "cylinders" $\prod_{\alpha\in\mathcal{A}} U_\alpha$ where $U_\alpha \subset X_\alpha$ is open and $U_\alpha = X_\alpha$ for all but finitely many α's
- If X is topological space and $\mathcal{A}$ a set of indices, consider

$$X^{\mathcal{A}} = \prod_{\alpha\in\mathcal{A}} X = \{f : \mathcal{A} \to X\}.$$

  For each α, the projection map is given by $\pi_\alpha(f) = f(\alpha)$. Then a sequence $\{f_n\}_{n\in\mathbb{N}}$ in $X^{\mathcal{A}}$ converges to $f \in X^{\mathcal{A}}$ with respect to the product topology if and only if $f_m(\alpha) \to f(\alpha)$ in X with respect to X's topology for all $\alpha \in \mathcal{A}$. [See Folland Proposition 4.12].

Definition 3.13. Let X be a topological space. A set $Y \subset X$ is **compact** if for any cover $Y \subset \bigcup_{\alpha\in\mathcal{A}} U_\alpha$ of open sets $U_\alpha \subset X$ there exist $\alpha_1, \ldots, \alpha_n \in \mathcal{A}$ such that $Y \subset \bigcup_{i=1}^n U_{\alpha_i}$.

Proposition 3.14. *(Tychonoff Theorem)* If $\{X_\alpha\}_{\alpha\in\mathcal{A}}$ is a family of compact topological spaces, then $\prod_{\alpha\in\mathcal{A}} X_\alpha$ with the product topology is compact. [This is false for a general topology, e.g. it is not true of the box topology]. [See Folland Theorem 4.42].

Definition 3.15. A topological space $(X, \mathbb{T})$ is **Hausdorff** (or T_2) if for all $x \neq y \in X$, there are disjoint $U, V \in \mathbb{T}$ such that $x \in U$ and $y \in V$.

Proposition 3.16. Let $(X, \mathbb{T})$ be a topological space.

(i) If $(X, \mathbb{T})$ is Hausdorff, $K \subset X$ is compact, and $x \in K^c$ then there exist disjoint $U, V \in \mathbb{T}$ such that $x \in U$ and $K \subset V$.

(ii) If X is compact and $C \subset X$ is closed, then C is compact.

(iii) If X is Hausdorff and $K \subset X$ is compact, then K is closed.

Proof. Exercise. □

Proposition 3.17. *(Heine-Borel)* A subset $F \subset \mathbb{R}^n$ is compact (under the usual topology) if and only if F is closed and bounded.

Proof. Exercise. □

Proposition 3.18. *(Bolzano-Weierstrass)* Let (X, d) be a metric space. $F \subset X$ is compact if and only if every sequence in F has a subsequence converging to a point in F (in which case we say F is "sequentially compact"). [Note that compactness and sequential compactness are not equivalent in general topological spaces and that either implication could fail].

Example 3.19. *(Closed & Bounded does not imply Compact in Infinite Dimensional Space)* Let $X = L^2(\mathbb{R})$ and $B := \{f \in L^2(\mathbb{R}) : \|f\|_2 \leq 1\}$, then B is closed and bounded but not compact.

Proof. It is immediate that B is closed (since $\|\cdot\| : X \to \mathbb{R}$ is continuous, B is the pre-image of the closed set $[0, 1]$, and the pre-image of a closed set remains closed.) and bounded (since B is a ball). As $L^2(\mathbb{R})$ is a metric space, compactness is equivalent to sequential compactness (i.e. K is compact if and only if every sequence in K has a convergent subsequence). Since $L^2(\mathbb{R})$ is complete, it suffices to find sequences $\{f_k\}_{k\in\mathbb{N}}$ in B with no Cauchy subsequences. Any of the following will work (where $\phi : \mathbb{R} \to \mathbb{C}$ is L^2 and satisfies $0 < \|\phi\|_2 \leq 1$):

(i) $f_k(x) = \phi(x-k)$

(ii) $f_k(x) = \sqrt{k}\phi(kx)$

(iii) $f_k(x) = \frac{1}{\sqrt{2}}\chi_{[0,1]}(x)\sin(k\pi x)$

One can verify that all of (i)-(iii) are in B, but that none have any Cauchy subsequences. The verification of (iii) is the easiest. Indeed

$$\|f_k\|_2^2 = \frac{1}{2}\int_0^1 \sin^2(k\pi x)\,\mathrm{d}x = \frac{1}{2}\int_0^1 \frac{1-\cos(k\pi x)}{2}\,\mathrm{d}x = \frac{1}{4} \leq 1$$

so $f_k \in B$ for all $k \in \mathbb{N}$. Moreover, for $j \neq k$ we have $\langle f_j, f_k\rangle = \frac{1}{2}\int_0^1 \sin(j\pi x)\sin(k\pi x)\,\mathrm{d}x = 0$ and

$$\|f_j - f_k\| = \|f_j\|_2^2 + \|f_k\|_2^2 - 2\mathrm{Re}\langle f_j, f_k\rangle = \|f_j\|_2^2 + \|f_k\|_2^2 = \frac{1}{2}.$$

So no subsequence is Cauchy.

□

3.2 Weak, Weak*, and Operator Topologies

Definition 3.20. A **directed set** $\mathcal{I}$ is non-empty equipped with a relationship $\preceq$ such that

- $i \preceq i$ for all $i \in \mathcal{I}$ (reflexive)
- $i \preceq j$ and $j \preceq k$ implies $i \preceq k$ (transitive)
- for all $i, j \in \mathcal{I}$, there is $k \in \mathcal{I}$ such that $i \preceq k$ and $j \preceq k$ (directed)

Note that not all indices need be comparable, but for any two indices we can find some other index which is larger than both.

Definition 3.21. A **net** on X is a function $x : \mathcal{I} \to X$ where $\mathcal{I}$ is a directed set. A net can also be denoted $\{x_i\}_{i\in\mathcal{I}}$. As an example, consider the directed set $(\mathbb{N}, \leq)$ with nets the sequences $\{x_n\}_{n\in\mathbb{N}}$.

Definition 3.22. Let X be a topological space. We say a net $\{x_i\}_{i\in\mathcal{I}}$ on X **converges to** $x \in X$ if for every open neighbourhood U of x, there is $i_0 \in \mathcal{I}$ such that $x_i \in U$ for all $i \succeq i_0$. In which case we write $x_i \to x$.

Proposition 3.23. Let X and Y be topological spaces.

(i) $F \subset X$ is closed if and only if for all nets $\{x_i\}_{i\in\mathcal{I}}$ in F converging to $x \in X$, we have $x \in F$

(ii) If X is Hausdorff, then every convergent net in X has a unique limit

(iii) X is compact if and only if every net in X has a convergent subnet

(iv) Let $f : X \to Y$ be continuous. If $\{x_i\}_{i\in\mathcal{I}}$ in X converges to $x \in X$, then $f(x_i) \to f(x)$ in Y

Definition 3.24. Let X be a normed vector space.
The topology on X generated by X^* (i.e. the smallest topology on X such that all linear and bounded $f : X \to \mathbb{C}$ is continuous) is the **weak topology** on X.
We say that x_i **converges weakly** to x, denoted $x_i \xrightarrow{w} x$ if and only if $f(x_i) \to f(x)$ for all $f \in X^*$.
Notice that $\mathbb{T}_{\text{weak}} \subset \mathbb{T}_{\|\cdot\|}$ since all $f \in X^*$ are continuous with respect to the norm topology. If X is infinite dimensional, then this is a strict subset (proof omitted).

Example 3.25. Let $\mathcal{H}$ be a Hilbert space. By the Riesz Representation Theorem we can write $f \in \mathcal{H}^*$ as $\langle \cdot, y \rangle$ for some $y \in \mathcal{H}$. Thus weak convergence in a Hilbert space is

$$x_i \xrightarrow{w} x \iff \langle x_i, y \rangle \to \langle x, y \rangle \qquad \forall\, y \in \mathcal{H}.$$

Example 3.26. Consider $X = L^p$ with $1 \leq p < \infty$ and q conjugate so that $\frac{1}{p} + \frac{1}{q} = 1$. Then we previously showed $(L^p)^* \simeq L^q$ via the map

$$L^q \ni g \mapsto \int \cdot\, g \in (L^p)^*.$$

So weak convergence in L^p is

$$f_i \xrightarrow{w} f \iff \int f_i g \to \int f g \qquad \forall\, g \in L^q$$

Example 3.27. Let $1 < p < \infty$. Then $f_k = \chi_{[k,k+1]}$ is a sequence in $L^p(\mathbb{R})$ that is weakly convergent to 0, and $\|f_k\|_p = 1$. It's clear that f_m has norm 1. Let $\varepsilon > 0$ and fix $g \in L^q$. Since $\int |g|^q < \infty$ we have that $\int_{|x| \geq R} |g|^q < \varepsilon$ for R large enough (one can use the D.C.T. to show this). Now we use Hölder to write:

$$\left| \int g f_k \right| = \left| \int_k^{k+1} g \right| \leq \int_k^{k+1} |g| 1 \leq \left(\int_k^{k+1} |g|^q \right)^{1/q} \underbrace{\left(\int_k^{k+1} 1^p \right)^{1/p}}_{1} < \varepsilon^{1/q}$$

for $k \geq R$. Namely $\int g f_k \to 0$ for any $g \in L^q$ and so $f_k \xrightarrow{w} 0$.

Definition 3.28. Recall we saw the map $X \to X^{**}$ which maps $x \mapsto \hat{x}$ where $\hat{x}$ denotes evaluation at x. That is $\hat{x} : X^* \to \mathbb{C}$ maps $\hat{x} : f \mapsto f(x)$. The topology on X^* generated by $\{\hat{x} : x \in X\}$ is called the **weak* topology**. (It is the smallest topology on X^* such $\hat{x}$ is continuous for all $x \in X$). We have weak* convergence, denoted $f_i \xrightarrow{w^*} f$ if and only if $f_i(x) \to f(x)$ for all $x \in X$.

Note 3.29. Notice that $\{\hat{x} : x \in X\} \subset X^{**}$. Now the weak topology on X^*, denoted $\mathbb{T}_{\text{weak}}(X^*)$, is generated by all functionals in X^{**}. The weak* topology on X^*, denoted $\mathbb{T}_{\text{weak}*}$, is generated by all the evaluations. Hence, when X is reflexive (i.e. $X \simeq X^*$), then all elements of X^{**} are evaluations and so $\mathbb{T}_{\text{weak}*} = \mathbb{T}_{\text{weak}}(X^*)$. Otherwise $\mathbb{T}_{\text{weak}*} \subset \mathbb{T}_{\text{weak}}(X^*)$ strictly.

Theorem 3.30. *(Alaoglu's Theorem)* Let $(X, \|\cdot\|)$ be a normed vector space. Then

$$B^* = \{f \in X^* : \|f\| \leq 1\}$$

is compact with respect to the weak* topology.

Proof.

- Embed B^* into a "compact product". For $x \in X$, define $D_x = \{z \in \mathbb{C} : |z| \leq \|x\|\}$ which is a compact subset of $\mathbb{C}$. Then by Tychonoff's Theorem,

  $$D = \prod_{x \in X} D_x = \{\phi : X \to \mathbb{C} \text{ so that } |\phi(x)| \leq \|x\|\}$$

  is also compact. Now define the map $J : B^* \to D$ as $J(f) = (f(x))_{x \in X}$ which is well-defined (if $f \in B^*$, then $\|f\| = 1$ and so $|J(f(x))| = |\phi(x)| \leq \|x\|$) and injective.

- We now show the weak* topology restricted to B^* is the same as the subspace topology on B^* induced by D via J. Notice that $\hat{x}(f) = \mathrm{ev}_x(f) = f(x) = \pi_x((f(y))_{y\in X}) = \pi_x(J(f))$. Indeed, the weak* topology on X^* restricted to B^* is by definition the smallest topology on B^* such that $\mathrm{ev}_x : B^* \to \mathbb{C}$ is continuous for all $x \in X$. And the subspace topology on B^* induced by D via J is the smallest topology on B^* such that $\pi_x \circ J$ is continuous for all $x \in X$.

- $J(B^*)$ is closed in D. Let $(J(f_\alpha))_\alpha$ be a net in $J(B^*)$ converging to $\phi \in D$. We want to show that $\phi \in J(B^*)$. We have that $f_\alpha(x) = \pi_x J f_\alpha \to \pi_x\phi = \phi(x)$ for each $x \in X$. Define $f : X \to \mathbb{C}$ as $f(x) = \phi(x)$. We need to show that $f \in B^*$. That is, $f \in X^*$ and $\|f\| \leq 1$. First it is linear since for $x, y \in X$ and $a, b \in \mathbb{C}$ we have $f_\alpha \in B^*$ so: $f_\alpha(ax+by) = af_\alpha(x)+bf_\alpha(y)$. The LHS converges to $f(ax + by)$ and the RHS converges to $af(x) + bf(y)$. Moreover, $|f(x)| = \lim |f_\alpha(x)| \leq \|x\|$ which implies $\|f\| \leq 1$. So $f \in B^*$ and $\phi \in J(B^*)$. Namely $J(B^*)$ is closed.

- We have that $J(B^*)$ is a closed subset of the compact D. Hence $J(B^*)$ is compact. Moreover, since J is a homeomorphism we have that B^* is compact.

□

Note 3.31.

(i) For $M > 0$, $MB^* := \{f \in X^* : \|f\| \leq M\}$ is compact with respect to the weak* topology.

(ii) The map $T : (X^*, \mathbb{T}_{w^*}) \to (X^*, \mathbb{T}_{w^*})$ given by $f \mapsto Mf$ is continuous with a continuous inverse (given by $g \mapsto \frac{1}{M}g$). So T is a homeomorphism.

Theorem 3.32. *(Banach-Alaoglu)* X is a separable normed vector space. Then any bounded sequence in X^* has a weak* convergent subsequence.

Proof. Omitted. □

4 Fourier Analysis

4.1 The Fourier Transform

Definition 4.1. For $z \in \mathbb{C}$, we define the **complex exponential** as

$$e^z = \sum_{m=0}^{\infty} \frac{z^m}{m!}.$$

Note this series is absolutely convergent for all $z \in \mathbb{C}$ (for example via the ratio test). One can show that $e^{z+w} = e^z e^w$ for any $z, w \in \mathbb{C}$. In particular

$$e^{a+bi} = e^a e^{bi}$$

for all $a, b \in \mathbb{R}$. Moreover by splitting the even and odd terms one can write:

$$e^{ib} = \sum_{m=0}^{\infty} \frac{(ib)^m}{m!} = \sum_{m=0}^{\infty} \frac{(-1)^m b^{2m}}{(2m)!} + i \sum_{m=0}^{\infty} \frac{(-1)^m b^{2m+1}}{(2m+1)!} = \cos b + i \sin b.$$

Through rearrangement one also obtains

$$\cos\theta = \frac{e^{i\theta} + e^{-i\theta}}{2} \qquad \sin\theta = \frac{e^{i\theta} - e^{-i\theta}}{2i}.$$

For $x, y \in \mathbb{R}^n$, with $x = (x_1, \ldots, x_n)$ and $y = (y_1, \ldots, y_n)$ we have the dot product

$$x \cdot y = x_1 y_1 + \cdots + x_n y_n = |x||y| \cos\theta,$$

where θ is angle between the vectors x and y. Also $|x|^2 = x \cdot x$.

Definition 4.2. Let $f \in L^1(\mathbb{R}^n)$. The **Fourier transform** of f, denoted $\hat{f} : \mathbb{R}^n \to \mathbb{C}$, is given by

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

It is sometimes also denoted $\mathscr{F}(f) = \hat{f}$.

Example 4.3.

(i) Define $f(x) = \chi_{[-1,1]}(x)$ on $\mathbb{R}$. Then for $\xi \neq 0$ we have

$$\hat{f}(\xi) = \int_{\mathbb{R}} \chi_{[-1,1]} e^{-2\pi i \xi x} \, \mathrm{d}x = \int_{-1}^{1} e^{-2\pi i \xi x} \, \mathrm{d}x = \cdots = \frac{e^{-2\pi i \xi} - e^{2\pi i \xi}}{-2\pi i \xi} = \frac{\sin(2\pi\xi)}{\pi\xi}.$$

If $\xi = 0$, then $\hat{f}(\xi) = \int \chi_{[-1,1]} \, \mathrm{d}x = 2$.

(ii) Define $f(x) = e^{-|x|}$ on $\mathbb{R}$. Notice that f is not differentiable at 0. Then

$$\hat{f}(\xi) = \int_{\mathbb{R}} e^{-|x|} e^{-2\pi i \xi x} \, \mathrm{d}x = \int_{-\infty}^{0} e^{x} e^{-2\pi i \xi x} \, \mathrm{d}x + \int_{0}^{\infty} e^{-x} e^{-2\pi i \xi x} \, \mathrm{d}x = \cdots = \frac{2}{1 + 4\pi^2 \xi^2}$$

which is C^∞.

Proposition 4.4. Let $a > 0$ and $x \in \mathbb{R}^n$. Then

$$\mathscr{F}\left(e^{-\pi a|x|^2}\right)(\xi) = a^{-n/2} e^{-\pi|\xi|^2/a}.$$

In particular $\mathscr{F}\left(e^{-\pi|x|^2}\right)(\xi) = e^{-\pi|\xi|^2}$.

Proof. First we show that for $x \in \mathbb{R}$: if $g(x) = e^{-\pi x^2}$, then $\hat{g}(\xi) = e^{-\pi\xi^2}$. Let

$$h(\xi) := \hat{g}(\xi) = \int_{\mathbb{R}} e^{-\pi x^2} e^{-2\pi i \xi x}\, \mathrm{d}x.$$

One can then differentiate under the integral sign (justifiable via the D.C.T.) to obtain

$$\begin{aligned}
h'(\xi) &= \int_{\mathbb{R}} -2\pi i x e^{-\pi x^2} e^{-2\pi i \xi x}\, \mathrm{d}x \\
&= i \int_{\mathbb{R}} \partial_x (e^{-\pi x^2}) e^{-2\pi i \xi x}\, \mathrm{d}x \\
&= i \underbrace{\left[e^{-\pi x^2} e^{-2\pi i \xi x}\right]_{-\infty}^{\infty}}_{0} - i \int_{-\infty}^{\infty} e^{-\pi x^2} \partial_x (e^{-2\pi i \xi x})\, \mathrm{d}x \\
&= -i \int_{-\infty}^{\infty} e^{-\pi x^2} (-2\pi i \xi) e^{-2\pi i \xi x}\, \mathrm{d}x \\
&= -2\pi \xi h(\xi)
\end{aligned}$$

Solving this differential equation then yields $h(\xi) = Ce^{-\pi\xi^2}$. And since

$$1 = \int e^{-\pi x^2}\, \mathrm{d}x = h(0) = C,$$

we get that $h(\xi) = e^{-\pi\xi^2}$. To extend the result to $\mathbb{R}^n$, notice that

$$e^{-\pi|x|^2} = \prod_{j=1}^{n} e^{-\pi x_j^2}.$$

The Fourier transform also factors (by Fubini). Hence $\mathscr{F}(e^{-\pi|x|^2})(\xi) = e^{-\pi|\xi|^2}$. For $a > 0$, define

$$f_a(x) = e^{-\pi a|x|^2} = f(\sqrt{a}x).$$

Then use the substitution $y = \sqrt{a}x$, $\mathrm{d}x = a^{-n/2}\, \mathrm{d}y$ to obtain

$$\hat{f}_a(\xi) = \int_{\mathbb{R}^n} e^{-\pi a|x|^2} e^{-2\pi i \xi x}\, \mathrm{d}x = a^{-n/2} \int_{\mathbb{R}^n} e^{-\pi|y|^2} e^{-2\pi i \left(\frac{\xi}{\sqrt{a}}\right)\cdot y}\, \mathrm{d}y = a^{-\frac{n}{2}} \hat{f}\left(\frac{\xi}{\sqrt{a}}\right) = a^{-\frac{n}{2}} e^{-\pi|\xi|^2/a}.$$

□

Definition 4.5. Let $f : \mathbb{R}^n \to \mathbb{C}$. We define

(i) the **translation:** $\tau_y f(x) = f(x - y)$

(ii) the **modulation:** $M_\eta f(x) = e^{2\pi i \eta \cdot x} f(x)$

(iii) for $f, g \in L^1$ the **convolution:** $(f * g)(x) = \int_{\mathbb{R}^n} f(x-y)g(y)\,\mathrm{d}y$

Definition 4.6. A **multi-index** in $\mathbb{R}^n$ is a vector $\alpha = (\alpha_1, \ldots, \alpha_n)$ with elements $\alpha_j \in \mathbb{N} \cup \{0\}$. If α is a multi-index and $x \in \mathbb{R}^n$ we set

$$|\alpha| = \sum_{j=1}^{n} \alpha_j \qquad \alpha! = \prod_{j=1}^{n} \alpha_j! \qquad \partial^\alpha = \frac{\partial^{\alpha_1}}{x_1} \cdots \frac{\partial^{\alpha_n}}{x_n} \qquad x^\alpha = \prod_{j=1}^{n} x_j^{\alpha_j}.$$

Theorem 4.7. Let $f, g \in L^1(\mathbb{R}^n)$.

(i) $\hat{f} \in C(\mathbb{R}^n)$ and $\|\hat{f}\|_\infty \leq \|f\|_1$

(ii) $\mathscr{F}(\tau_y f)(\xi) = e^{-2\pi i y\cdot\xi}\hat{f}(\xi)$ and $\mathscr{F}(M_\eta f)(\xi) = \hat{f}(\xi - \eta)$

(iii) $\mathscr{F}(f * g)(\xi) = \hat{f}(\xi)\hat{g}(\xi)$.

(iv) If $x^\alpha f \in L^1$ for all multi-indices $|\alpha| \leq k$, then $\hat{f} \in C^k$ and $\partial_\xi^\alpha \hat{f}(\xi) = \mathscr{F}((-2\pi i x)^\alpha f(x))(\xi)$

(v) If $f \in C^k$, $\partial^\alpha f \in L^1$ for all $|\alpha| \leq k$, and $\partial^\alpha f \in C_0$ for all $|\alpha| \leq k-1$, then

$$\mathscr{F}(\partial^\alpha f)(\xi) = (2\pi i\xi)^\alpha \hat{f}(\xi)$$

(vi) *(Riemann–Lebesgue Lemma)*: $\mathscr{F}(L^1(\mathbb{R}^n)) \subset C_0(\mathbb{R}^n)$

Here C_0 denotes the continuous functions vanishing at infinity.

Proof.

(i) For any $\xi \in \mathbb{R}^n$ we have

$$|\hat{f}(\xi)| = \left| \int_{\mathbb{R}^n} f(x) e^{-2\pi i x\cdot\xi}\,\mathrm{d}x \right| \leq \int |f(x)|\,\mathrm{d}x = \|f\|_1.$$

This implies $\|\hat{f}\|_\infty \leq \|f\|_1$. Now suppose $\xi_m \to \xi$. Then $e^{-2\pi i x\cdot\xi_m} \to e^{-2\pi i x\cdot\xi}$ for all $x \in \mathbb{R}^n$ pointwise. Combined with $|f(x)e^{-2\pi i x\cdot\xi_m}| \leq |f(x)|$ and $f \in L^1$, the D.C.T. yields

$$\hat{f}(\xi_m) = \int f(x) e^{-2\pi i x\cdot\xi_m}\,\mathrm{d}x \to \int f(x) e^{-2\pi i x\cdot\xi}\,\mathrm{d}x = \hat{f}(\xi).$$

(ii) Fix $y \in \mathbb{R}^n$. Then using a change of variables $z = x - y$ we obtain

$$\begin{aligned}
\mathscr{F}(\tau_y f)(\xi) &= \int_{\mathbb{R}^n} f(x-y) e^{-2\pi i x\cdot\xi}\,\mathrm{d}x \\
&= \int_{\mathbb{R}^n} f(z) e^{-2\pi i (z+y)\cdot\xi}\,\mathrm{d}z \\
&= e^{-2\pi i y} \int_{\mathbb{R}^n} f(z) e^{-2\pi i z\cdot\xi}\,\mathrm{d}z \\
&= e^{-2\pi i y}\hat{f}(\xi).
\end{aligned}$$

Now fix $\eta \in \mathbb{R}^n$. Simple algebra yields

$$\mathscr{F}(M_\eta f)(\xi) = \int e^{2\pi i \eta\cdot x} f(x) e^{-2\pi i \xi\cdot x}\,\mathrm{d}x = \int f(x) e^{-2\pi i (\xi-\eta)\cdot x}\,\mathrm{d}x = \hat{f}(\xi - \eta).$$

(iii) We first show that $f * g \in L^1$. Indeed,

$$\int |(f*g)(x)|\,\mathrm{d}x \le \int_{\mathbb{R}^n}\int_{\mathbb{R}^n} |f(x-y)||g(y)|\,\mathrm{d}x\,\mathrm{d}y = \int_{\mathbb{R}^n}\left(\int_{\mathbb{R}^n} |f(x-y)|\,\mathrm{d}x\right)|g(y)|\,\mathrm{d}y = \|f\|_1\|g\|_1 < \infty.$$

Now we take the Fourier transform (which via change of variables becomes):

$$\begin{aligned}\mathscr{F}(f*g)(\xi) &= \int_{\mathbb{R}^m}\left(\int_{\mathbb{R}^m} f(x-y)g(y)\,\mathrm{d}y\right)e^{-2\pi i\xi\cdot x}\,\mathrm{d}x \\ &= \int_{\mathbb{R}^m}\int_{\mathbb{R}^m} f(z)g(y)e^{-2\pi i\xi\cdot(y+z)}\,\mathrm{d}y\,\mathrm{d}z \\ &= \left(\int_{\mathbb{R}^m} f(z)e^{-2\pi i\xi\cdot z}\,\mathrm{d}z\right)\left(\int_{\mathbb{R}^m} g(y)e^{-2\pi i\xi\cdot y}\,\mathrm{d}y\right) \\ &= \hat{f}(\xi)\hat{g}(\xi)\end{aligned}$$

(iv) Suppose $x^\alpha f \in L^1(\mathbb{R}^n)$ for all multi-indices $|\alpha| \le k$. One can show that differentiating under the integral sign is allowable (and in the process show that $\hat{f} \in C^k$). Hence by properties of the complex exponential we get:

$$\partial_\xi^\alpha \hat{f}(\xi) = \int_{\mathbb{R}^n} f(x)\partial_\xi^\alpha\left(e^{-2\pi ix\cdot\xi}\right)\mathrm{d}x = \int_{\mathbb{R}^n} f(x)(-2\pi ix)^\alpha e^{-2\pi ix\cdot\xi}\,\mathrm{d}x = \mathscr{F}(f(x)(-2\pi ix)^\alpha)(\xi).$$

Proofs of (v) and (vi) are omitted.

□

4.2 The Schwartz Class

Definition 4.8. The **Schwartz class** is

$$\mathbb{S}(\mathbb{R}^n) = \left\{f \in C^\infty(\mathbb{R}^n) : \sup_{x\in\mathbb{R}^n} |x^\alpha\partial^\beta f(x)| < \infty \text{ for all multi-indices } \alpha,\beta\right\}$$

Note 4.9. It is clear that $\mathbb{S}$ is closed under taking derivatives. One can further show that $\mathbb{S}$ is closed under multiplication by a polynomial, i.e. $f \in \mathbb{S} \implies x^\alpha f \in \mathbb{S}$ for any multi-index α.

Lemma 4.10. Let $f \in C^\infty(\mathbb{R}^n)$. The following are equivalent.

(i) $f \in \mathbb{S}$, i.e.

$$\sup_{x\in\mathbb{R}^n} |x^\alpha\partial^\beta f(x)| < \infty \text{ for all multi-indices } \alpha,\ \beta.$$

(ii) For all multi-indices β and $N \ge 0$, there is a constant $C_{N,\beta}$ such that

$$|\partial^\beta f(x)| \le \frac{C_{N,\beta}}{(1+\|x\|_2)^N}$$

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

Proof. First (i) $\implies$ (ii). Fix β and N so that $\sup_{x\in\mathbb{R}^n} |x^\alpha\partial^\beta f(x)| < \infty$ for all α. Now we claim that there is a constant C_N such that:

$$(1+\|x\|_2)^N \le C_N \sum_{|\alpha|\le N} |x^\alpha| \text{ for all } x \in \mathbb{R}^n.$$

Indeed, notice that $\|x\|_2 \leq |x_1| + \cdots + |x_n|$ and so

$$(1 + \|x\|_2)^N \leq (1 + |x_1| + \cdots + |x_n|)^N = \sum_{|\alpha| \leq N} C_\alpha |x^\alpha| \leq C_N \sum_{|\alpha| \leq N} |x^\alpha|,$$

where $C_N = \max_{|\alpha| \leq N} C_{|\alpha|}$, proving the claim. Now we obtain

$$(1 + \|x\|_2)^N |\partial^\beta f(x)| \leq C_N \sum_{|\alpha| \leq N} |x^\alpha \partial^\beta f(x)| \leq C_N \sum_{|\alpha| \leq N} \sup_{x \in \mathbb{R}^n} |x^\alpha \partial^\beta f(x)| < \infty.$$

Call this attained supremum $C_{N,\beta}$, and so

$$|\partial^\beta f(x)| \leq \frac{C_{N,\beta}}{(1 + \|x\|_2)^N}$$

as desired. Now we show (ii) $\implies$ (i). Fix multi-indices α, β and take $N = |\alpha|$. Then there exists a constant $C_{N,\beta}$ such that $|\partial^\beta f(x)| \leq \frac{C_{N,\beta}}{(1+\|x\|_2)^N}$ for all $x \in \mathbb{R}^n$. Then notice that

$$|x^\alpha| = |x_1^{\alpha_1}| \cdots |x_n^{\alpha_n}| \leq \|x\|_2^{\alpha_1} \cdots \|x\|_2^{\alpha_n} = \|x\|_2^{|\alpha|} = \|x\|_2^N.$$

hence

$$|x^\alpha \partial^\beta f(x)| \leq \frac{\|x\|_2^N C_{N,\beta}}{(1 + \|x\|_2)^N} \leq C_{N,\beta} < \infty.$$

□

Note 4.11. $C_c^\infty \subset \mathbb{S} \subset C_0^\infty$, where C_c^∞ are the infinitely differentiable functions with compact support and

$$C_0^\infty = \left\{ f \in C^\infty : \lim_{\|x\|_2 \to \infty} f(x) = 0 \right\}.$$

Proposition 4.12. $\mathbb{S}(\mathbb{R}^n) \subset L^p(\mathbb{R}^n)$ for $1 \leq p \leq \infty$.

Proof. By Lemma 4.10 (ii), for all $N \geq 0$ there exists a constant C_N such that

$$|f(x)| \leq \frac{C_N}{(1 + \|x\|_2)^N}.$$

- For $p = \infty$, take $N = 0$ to see that f is bounded and hence L^∞.
- For $1 \leq p < \infty$, write

$$\|f\|_p^p = \int_{\mathbb{R}^n} |f(x)|^p \, \mathrm{d}x = \int_{\|x\|_2 \leq 1} |f(x)|^p \, \mathrm{d}x + \int_{\|x\|_2 > 1} |f(x)|^p \, \mathrm{d}x.$$

  The first term is the integral of a continuous function over a compact set, and is hence bounded. For the second term we can write

$$|f(x)|^p \leq \frac{C_N^p}{(1 + \|x\|_2)^{Np}} \leq \frac{C_N^p}{\|x\|_2^{Np}}.$$

  The second term can then be shown to be bounded via integration in spherical coordinates.

□

Note 4.13. Fix $1 \leq p < \infty$. Since $C_c^\infty \subset \mathbb{S} \subset L^p$, and we know that C_c^∞ is dense in L^p [given $f \in L^p$ we can approximate $f_R = f\chi_{B_R(0)}$ so that $\|f - f_R\|_p \to 0$ as $R \to \infty$ via the D.C.T. Then using a "mollifier" $\rho_\varepsilon \in C_c^\infty$ we can show that $\|\rho_\varepsilon * f_R - f_R\|_p \to 0$ as $\varepsilon \to 0$ and that $\rho_\varepsilon * f_R \in C_c^\infty$]. We therefore see that $\mathbb{S}$ is dense in L^p.

However $\mathbb{S}$ is not dense in L^∞. For example take $f \equiv 1 \in L^\infty$. For any $\varphi \in \mathbb{S}(\mathbb{R}^n)$ we have

$$\|f - \varphi\|_\infty = \|1 - \varphi\|_\infty \geq 1.$$

And so $f \notin \overline{\mathbb{S}(\mathbb{R}^n)}^{\|\cdot\|_\infty}$.

Proposition 4.14. $\mathscr{F}(\mathbb{S}) \subset \mathbb{S}$

Proof. Let $f \in \mathbb{S}$. We need to show that $\sup_{\xi \in \mathbb{R}^n} |\xi^\alpha \partial^\beta \hat{f}(\xi)| < \infty$. From Theorem 4.7 (iv) and (v) we know

- $\partial_\xi^\beta \hat{f}(\xi) = \mathscr{F}((-2\pi i x)^\beta f(x))(\xi)$
- $(2\pi i \xi)^\alpha \hat{g}(\xi) = \mathscr{F}(\partial^\alpha g)(\xi)$

Applying these one after the other with $g(x) = (-2\pi i x)^\beta f(x)$ we obtain

$$(2\pi i\xi)^\alpha \partial_\xi^\beta \hat{f}(\xi) = (2\pi i\xi)^\alpha \mathscr{F}((-2\pi i x)^\beta f(x))(\xi) = \mathscr{F}(\partial^\alpha(-2\pi i x)^\beta f(x))(\xi).$$

Taking the magnitude of both sides and rearranging yields.

$$|\xi^\alpha \partial_\xi^\beta \hat{f}(\xi)| = (2\pi)^{|\beta|-|\alpha|} \left|\mathscr{F}(\partial^\alpha(x^\beta f(x))(\xi)\right| \leq (2\pi)^{|\beta|-|\alpha|} \|\partial^\alpha x^\beta f(x)\|_{L^1} < \infty.$$

Where we use Theorem 4.7 (i) [$\|\hat{h}\|_\infty \leq \|h\|_{L^1}$] and the fact that $\partial^\alpha x^\beta f(x) \in \mathbb{S} \subset L^1$. □

4.3 The Inverse Fourier Transform

Definition 4.15. For $h \in L^1(\mathbb{R})$, the **inverse Fourier transform** of h is

$$\check{h}(x) := \int_{\mathbb{R}^n} h(\xi) e^{2\pi i x\cdot\xi} \, \mathrm{d}\xi.$$

It is sometimes also denoted $\mathscr{F}^{-1}(h) = \check{h}$.

Note 4.16. One can prove a version of the Riemann-Lebesgue Lemma which says

$$\mathscr{F}^{-1}(L^1(\mathbb{R}^n)) \subset C_0(\mathbb{R}^n).$$

Theorem 4.17. If $f \in L^1(\mathbb{R}^n)$ and $\hat{f} \in L^1(\mathbb{R}^n)$, then

$$\mathscr{F}^{-1}(\mathscr{F} f) = f(x) \qquad \text{for a.e. } x \in \mathbb{R}^n.$$

And moreover, $\mathscr{F}^{-1}(\mathscr{F} f) \in C_0(\mathbb{R}^n)$.

Proof. Fix $t > 0$ and define

$$\Phi_t(\xi) = e^{-\pi t^2 |\xi|^2}, \qquad I_t(x) = \int_{\mathbb{R}^n} \hat{f}(\xi) e^{2\pi i x\cdot\xi} \Phi_t(\xi) \, \mathrm{d}\xi.$$

First we show that $I_t(x) \to \mathscr{F}^{-1}(\hat{f})(x)$ for every $x \in \mathbb{R}^n$. Since $\Phi_t(\xi) \to 1$ for all ξ as $t \to 0$, and

$$|\hat{f}(\xi)e^{2\pi i x\cdot\xi}\Phi_t(\xi)| \le |\hat{f}(\xi)|,$$

with $\hat{f} \in L^1(\mathbb{R}^n)$, the D.C.T. implies

$$I_t(x) \to \int_{\mathbb{R}^n} \hat{f}(\xi)e^{2\pi i x\cdot\xi}\,\mathrm{d}\xi = \mathscr{F}^{-1}(\hat{f})(x).$$

Next we rewrite I_t as a convolution. Substituting the definition of $\hat{f}$ we obtain

$$\begin{aligned} I_t(x) &= \int_{\mathbb{R}^n}\left(\int_{\mathbb{R}^n} f(y)e^{-2\pi i y\cdot\xi}\,\mathrm{d}y\right)e^{2\pi i x\cdot\xi}\Phi_t(\xi)\,\mathrm{d}\xi \\ &= \int_{\mathbb{R}^n}\int_{\mathbb{R}^n} f(y)e^{-2\pi i y\cdot\xi}e^{2\pi i x\cdot\xi}\Phi_t(\xi)\,\mathrm{d}y\,\mathrm{d}\xi. \end{aligned}$$

We can apply Fubini's theorem since

$$\int_{\mathbb{R}^n}\int_{\mathbb{R}^n} |f(y)||\Phi_t(\xi)|\,\mathrm{d}y\,\mathrm{d}\xi = \|f\|_1\|\Phi_t\|_1 < \infty.$$

Thus

$$\begin{aligned} I_t(x) &= \int_{\mathbb{R}^n} f(y)\left(\int_{\mathbb{R}^n} e^{2\pi i(x-y)\cdot\xi}\Phi_t(\xi)\,\mathrm{d}\xi\right)\mathrm{d}y \\ &= \int_{\mathbb{R}^n} f(y)\mathscr{F}(\Phi_t)(x-y)\,\mathrm{d}y. \end{aligned}$$

By Proposition 4.4 we know the Fourier transform of a Gaussian. Hence

$$\mathscr{F}(\Phi_t)(x) = t^{-n}e^{-\pi|x|^2/t^2} =: g_t(x).$$

Therefore

$$I_t(x) = (f * g_t)(x).$$

The family $\{g_t\}_{t>0}$ is an approximate identity: $g_t \ge 0$, $\int_{\mathbb{R}^n} g_t(z)\,\mathrm{d}z = 1$, and the mass of g_t concentrates near the origin as $t \to 0$. Hence by the standard theorem on approximate identities we obtain

$$f * g_t \to f \qquad \text{in } L^1(\mathbb{R}^n).$$

From convergence in L^1 there exists a sequence $t_k \to 0$ such that

$$I_{t_k}(x) \to f(x) \qquad \text{for a.e. } x \in \mathbb{R}^n.$$

On the other hand we already showed that

$$I_{t_k}(x) \to \mathscr{F}^{-1}(\hat{f})(x) \qquad \text{for all } x \in \mathbb{R}^n.$$

Therefore

$$\mathscr{F}^{-1}(\hat{f})(x) = f(x) \qquad \text{for a.e. } x \in \mathbb{R}^n.$$

Finally we show that $\mathscr{F}^{-1}(\hat{f}) \in C_0(\mathbb{R}^n)$. For each $t > 0$ we have

$$I_t(x) = \mathscr{F}^{-1}(\hat{f}\,\Phi_t)(x).$$

Since $|\hat{f}(\xi)\Phi_t(\xi)| \leq |\hat{f}(\xi)| \in L^1$, we have $\hat{f}\,\Phi_t \in L^1$. By Note 4.16 we obtain

$$I_t \in C_0(\mathbb{R}^n).$$

Moreover

$$\begin{aligned} |I_t(x) - \mathscr{F}^{-1}(\hat{f})(x)| &= \left| \int_{\mathbb{R}^n} \hat{f}(\xi)(\Phi_t(\xi) - 1) e^{2\pi i x\cdot\xi} \, \mathrm{d}\xi \right| \\ &\leq \int_{\mathbb{R}^n} |\hat{f}(\xi)||\Phi_t(\xi) - 1| \, \mathrm{d}\xi. \end{aligned}$$

The right-hand side is independent of x and tends to 0 as $t \to 0$ by the D.C.T. since

$$|\hat{f}(\xi)||\Phi_t(\xi) - 1| \leq 2|\hat{f}(\xi)| \in L^1.$$

Hence $\|I_t - \mathscr{F}^{-1}(\hat{f})\|_\infty \to 0$. Thus $\mathscr{F}^{-1}(\hat{f})$ is the uniform limit of functions in $C_0(\mathbb{R}^n)$ and therefore

$$\mathscr{F}^{-1}(\hat{f}) \in C_0(\mathbb{R}^n).$$

□

Proposition 4.18. Let $g \in L^1(\mathbb{R}^n)$ and suppose that $\mathscr{F}^{-1}g \in L^1(\mathbb{R}^n)$. Then

$$\mathscr{F}(\mathscr{F}^{-1}g)(\xi) = g(\xi) \qquad \text{for a.e. } \xi \in \mathbb{R}^n.$$

An moreover, $\mathscr{F}(\mathscr{F}^{-1}g) \in C_0(\mathbb{R}^n)$ and also $\mathscr{F}^{-1}g \in C_0(\mathbb{R}^n)$.

Proof. Define $\tilde{g}(\xi) = g(-\xi)$. Since $g \in L^1(\mathbb{R}^n)$ we also have $\tilde{g} \in L^1(\mathbb{R}^n)$.
We claim that

$$\mathscr{F}(\tilde{g}) = \mathscr{F}^{-1}(g).$$

Indeed for any $x \in \mathbb{R}^n$,

$$\mathscr{F}(\tilde{g})(x) = \int_{\mathbb{R}^n} \tilde{g}(\xi) e^{-2\pi i x\cdot\xi} \, \mathrm{d}\xi = \int_{\mathbb{R}^n} g(-\xi) e^{-2\pi i x\cdot\xi} \, \mathrm{d}\xi.$$

With the change of variables $\eta = -\xi$ this becomes

$$\mathscr{F}(\tilde{g})(x) = \int_{\mathbb{R}^n} g(\eta) e^{2\pi i x\cdot\eta} \, \mathrm{d}\eta = \mathscr{F}^{-1}(g)(x).$$

Thus

$$\mathscr{F}(\tilde{g}) = \mathscr{F}^{-1}(g).$$

By assumption $\mathscr{F}^{-1}(g) \in L^1(\mathbb{R}^n)$. Hence $\mathscr{F}(\tilde{g}) \in L^1(\mathbb{R}^n)$. Since $\tilde{g} \in L^1(\mathbb{R}^n)$ we may apply Theorem 4.17 to obtain

$$\mathscr{F}^{-1}(\mathscr{F}(\tilde{g})) = \tilde{g} \qquad \text{for a.e. } x \in \mathbb{R}^n.$$

Using $\mathscr{F}(\tilde{g}) = \mathscr{F}^{-1}(g)$ we get

$$\mathscr{F}^{-1}(\mathscr{F}^{-1}(g)) = \tilde{g} \qquad \text{a.e.}$$

Now for any integrable function h one has

$$\mathscr{F}(h)(\xi) = \mathscr{F}^{-1}(h)(-\xi).$$

Applying this with $h = \mathscr{F}^{-1}(g)$ yields

$$\mathscr{F}(\mathscr{F}^{-1}(g))(\xi) = \mathscr{F}^{-1}(\mathscr{F}^{-1}(g))(-\xi) = \tilde{g}(-\xi) = g(\xi)$$

for almost every $\xi \in \mathbb{R}^n$.
Finally, since $\mathscr{F}^{-1}(g) \in L^1(\mathbb{R}^n)$, the Riemann–Lebesgue Lemma implies

$$\mathscr{F}(\mathscr{F}^{-1}(g)) \in C_0(\mathbb{R}^n).$$

Moreover, by Note 4.16 we also have

$$\mathscr{F}^{-1}(g) \in C_0(\mathbb{R}^n).$$

□

Theorem 4.19. *(Plancherel)* If $f \in L^1(\mathbb{R}^n) \cap L^2(\mathbb{R}^n)$, then $\hat{f} \in L^2(\mathbb{R}^n)$ and $\|\hat{f}\|_2 = \|f\|_2$. Moreover, the Fourier transform extends uniquely to an isometric isomorphism

$$\mathscr{F} : L^2(\mathbb{R}^n) \to L^2(\mathbb{R}^n).$$

Moreover: $\langle \mathscr{F}f, \mathscr{F}g \rangle = \langle f, g \rangle$ for all $f, g \in L^2$.

Proof. Let $X = \{f \in L^1 : \hat{f} \in L^1\}$. By Propositions 4.12 & 4.14: $\mathscr{F}(\mathbb{S}) \subset \mathbb{S} \subset L^1$, which implies that $\mathbb{S} \subset X$. Also by Note 4.13: $\mathbb{S}$ is dense in L^2. Together this implies that $X \cap L^2$ is dense in L^2.

- Now let us show that $\langle \mathscr{F}f, \mathscr{F}g \rangle = \langle f, g \rangle$ on $X \cap L^2$. Let $f, g \in X \cap L^2$ so that

  $$g(x) = \mathscr{F}^{-1}\mathscr{F}g(x) = \int \hat{g}(\xi) e^{2\pi i x \cdot \xi} \, \mathrm{d}\xi \qquad \text{and} \qquad \overline{g}(x) = \int \overline{\hat{g}}(\xi) e^{-2\pi i x \cdot \xi} \, \mathrm{d}\xi.$$

  Then

  $$\langle f, g \rangle = \int f(x)\overline{g}(x) \, \mathrm{d}x = \int f(x) \left(\int \overline{\hat{g}}(\xi) e^{-2\pi i x \cdot \xi} \, \mathrm{d}\xi \right) \mathrm{d}x = \int \overline{\hat{g}}(\xi) \hat{f}(\xi) \, \mathrm{d}\xi = \langle \mathscr{F}f, \mathscr{F}g \rangle,$$

  where the third equality follows from Fubini. Taking $g = f$ yields $\|f\|_2 = \|\hat{f}\|_2$.

- We now extend to L^2. Let $f \in L^2$ and choose $\{f_k\}_{k \in \mathbb{N}}$ in $X \cap L^2$ be so that $f_k \to f$ in L^2. Then by linearity

  $$\|\hat{f}_k - \hat{f}_m\|_2 = \|\widehat{f_k - f_m}\|_2 = \|f_k - f_m\|_2$$

  which can be made small for k, m large enough. In particular $\{f_k\}_{k \in \mathbb{N}}$ is Cauchy in L^2 and so there exists $h \in L^2$ such that $f_k \to h$ in L^2 by completeness. We **define** $\mathscr{F}f := h$. And indeed this is uniquely defined $\mathscr{F}$ since if $\{g_k\}_{k \in \mathbb{N}}$ is another such sequence in $X \cap L^2$ such that $g_k \to f$ in L^2, then

  $$\|\hat{f}_k - \hat{g}_k\|_2 = \|\widehat{f_k - g_k}\|_2 = \|f_k - g_k\|_2 \leq \|f_k - f\|_2 + \|g_k - f\|_2 \to 0.$$

  So h is unique.

- We now check that $\mathscr{F}$ defined in this way is an isometry in L^2 and preserves the inner product. Let $f \in L^2$ and $\{f_k\}_{k \in \mathbb{N}} \subset X \cap L^2$ be such that $f_k \to f$ in L^2. By definition we have $\hat{f}_k \to \mathscr{F}f$ in L^2. We also know that $\|\hat{f}_k\|_2 = \|f_k\|_2$ for all $k \in \mathbb{N}$. Taking limits and using continuity of the norm yields $\|\mathscr{F}f\|_2 = \|f\|_2$.

Now let $f, g \in L^2$ and let $\{f_k\}_{k\in\mathbb{N}}, \{g_k\}_{k\in\mathbb{N}} \subset X \cap L^2$ be such that $f_k \to f$ and $g_k \to g$ in L^2. Then $\langle \mathscr{F} f_k, \mathscr{F} g_k\rangle = \langle f_k, g_k\rangle$ for all $k \in \mathbb{N}$.

Since $\mathscr{F} f_k \to \mathscr{F} f$ and $\mathscr{F} g_k \to \mathscr{F} g$ in L^2, continuity of the inner product implies

$$\langle \mathscr{F} f_k, \mathscr{F} g_k\rangle \to \langle \mathscr{F} f, \mathscr{F} g\rangle.$$

Similarly, since $f_k \to f$ and $g_k \to g$ in L^2, we have

$$\langle f_k, g_k\rangle \to \langle f, g\rangle.$$

Passing to the limit in

$$\langle \mathscr{F} f_k, \mathscr{F} g_k\rangle = \langle f_k, g_k\rangle$$

yields

$$\langle \mathscr{F} f, \mathscr{F} g\rangle = \langle f, g\rangle.$$

- We now show that $\mathscr{F}$ is a bijection. By the same construction as above (replacing $\mathscr{F}$ with $\mathscr{F}^{-1}$), the inverse Fourier transform extends to an isometry

$$\mathscr{F}^{-1} : L^2(\mathbb{R}^n) \to L^2(\mathbb{R}^n).$$

Moreover, on $X \cap L^2$ we already know that $\mathscr{F}^{-1}$ and $\mathscr{F}$ are inverses.

Let $f \in L^2$ and choose $\{f_k\}_{k\in\mathbb{N}} \subset X \cap L^2$ such that $f_k \to f$ in L^2. Since $\mathscr{F}$ is an isometry, we have

$$\mathscr{F} f_k \to \mathscr{F} f \qquad \text{in } L^2.$$

Applying $\mathscr{F}^{-1}$ and using that it is also an isometry, we obtain

$$\mathscr{F}^{-1}(\mathscr{F} f_k) \to \mathscr{F}^{-1}(\mathscr{F} f) \qquad \text{in } L^2.$$

But for each k, $\mathscr{F}^{-1}(\mathscr{F} f_k) = f_k$. Hence the left-hand side converges to f in L^2. By uniqueness of limits in L^2, we conclude

$$\mathscr{F}^{-1}(\mathscr{F} f) = f.$$

A symmetric argument (starting with $g \in L^2$ and approximating by a sequence in $X \cap L^2$) shows that

$$\mathscr{F}(\mathscr{F}^{-1} g) = g.$$

Therefore $\mathscr{F}$ is bijective with inverse $\mathscr{F}^{-1}$, and hence is an isometric isomorphism of $L^2(\mathbb{R}^n)$.

□

4.4 Fourier Series

Definition 4.20. A function $f : \mathbb{R}^n \to \mathbb{C}$ is said to be $\mathbb{Z}^n$–**periodic** if

$$f(x + k) = f(x) \qquad \text{for all} \qquad x \in \mathbb{R}^n \text{and} \;\; k \in \mathbb{Z}^n.$$

Definition 4.21. The **torus** $\mathbb{T}^n = \mathbb{R}^n/\mathbb{Z}^n$ (where $x \sim x + k$ for all $x \in \mathbb{R}^n$ and $k \in \mathbb{Z}^n$) can be identified with a **representative** $Q = [-1/2, 1/2)^n$ in the sense that every point $y \in \mathbb{R}^n$ can be uniquely written as $y = x + k$ for some $x \in Q$ and $k \in \mathbb{Z}^n$. We define

$$\int_{\mathbb{T}^n} f(x)\,\mathrm{d}x = \int_Q f(x)\,\mathrm{d}x.$$

Example 4.22. The exponentials: $f_k(x) = e^{2\pi i k\cdot x}$ with $k \in \mathbb{Z}^n$ are $\mathbb{Z}^n$–periodic. Since

$$f_k(x+m) = e^{2\pi i k\cdot(x+m)} = e^{2\pi i k\cdot x}e^{2\pi i k\cdot m} = e^{2\pi i k\cdot x} = f_k(x).$$

Theorem 4.23. The exponentials, namely $\{e^{2\pi i k\cdot x}\}_{k\in\mathbb{Z}^n}$, form an orthonormal basis of $L^2(\mathbb{T}^n)$.

Proof.

- First, that they are indeed orthonormal. Let $k, \ell \in \mathbb{Z}^n$. Then

$$\langle f_k, f_\ell\rangle = \int_Q e^{2\pi i(k-\ell)\cdot x}\,\mathrm{d}x = \int_Q \prod_{j=1}^n e^{2\pi i(k_j-\ell_j)x_j}\,\mathrm{d}x_j = \prod_{j=1}^n \int_{-1/2}^{1/2} e^{2\pi i(k_j-\ell_j)x_j}\,\mathrm{d}x_j.$$

  When $k_j = \ell_j$ we have $\int_{-1/2}^{1/2} 1\,\mathrm{d}x = 1$. When $k_j \neq \ell_j$ we have

$$\int_{-1/2}^{1/2} e^{2\pi i(k_j-\ell_j)x_j}\,\mathrm{d}x_j = \frac{e^{2\pi i(k_j-\ell_j)x_j}}{2\pi i(k_j-\ell_j)}\bigg|_{-1/2}^{1/2} = \frac{e^{\pi i(k_j-\ell_j)} - e^{-\pi i(k_j-\ell_j)}}{2\pi i(k_j-\ell_j)} = \frac{(-1)^{k_j-\ell_j} - (-1)^{k_j-\ell_j}}{2\pi i(k_j-\ell_j)} = 0.$$

  Hence

$$\int_{-1/2}^{1/2} e^{2\pi i(k_j-\ell_j)x_j}\,\mathrm{d}x_j = \delta_{k_j\ell_j} \implies \langle f_k, f_\ell\rangle = \prod_{j=1}^n \delta_{k_j\ell_j} = \delta_{k\ell}.$$

- Now, that they span $L^2(\mathbb{T}^n)$. Finite linear combinations $\sum_k c_k e^{2\pi i k\cdot x}$ are called **trigonometric polynomials**. The Stone–Weierstrass theorem implies that these trig polynomials are dense in $C(\mathbb{T}^n)$. We also know that $C(\mathbb{T}^n) = C([-1/2,1/2)^n)$ is dense in $L^2([-1/2,1/2)^n) = L^2(\mathbb{T}^n)$. Hence the trig polynomials dense in $L^2(\mathbb{T}^n)$. Note that this is precisely what it means to span in this setting $[\overline{\{\text{finite sums } \sum_k c_k e^{2\pi i k\cdot x} : k\in\mathbb{Z}^n\}} = L^2(\mathbb{T}^n)]$.

□

Definition 4.24. For $f \in L^2(\mathbb{T}^n)$, we write the **Fourier Series of** f,

$$f(x) = \sum_{k\in\mathbb{Z}^n} \langle f, e^{2\pi i k\cdot x}\rangle e^{2\pi i k\cdot x} \qquad \text{in } L^2(\mathbb{T}^n).$$

We denote the **Fourier coefficients** as $\hat{f}(k) := \langle f, e^{2\pi i k\cdot x}\rangle = \int_Q f(x)e^{-2\pi i k\cdot x}\,\mathrm{d}x$. Note that this representation of f was not pulled out of thin air, but comes from Thm. 2.35 for general Hilbert spaces. The equality above is understood in the sense that

$$\left\| f(x) - \sum_{|k|\le N} \hat{f}(k)\, e^{2\pi i k\cdot x} \right\|_{L^2(\mathbb{T}^n)} \longrightarrow 0 \qquad \text{as} \quad N\to\infty$$

for some norm on $\mathbb{Z}^n$ (e.g. $|k| = \|k\|_\infty$). Moreover, Thm. 2.35 yields the Parseval identity for free

$$\sum_{k\in\mathbb{Z}^n} |\hat{f}(k)|^2 = \|f\|^2_{L^2(\mathbb{T}^n)}.$$

In this setting, the Fourier transform defines an isometric isomorphism

$$\mathscr{F} : L^2(\mathbb{T}^n) \to \ell^2(\mathbb{Z}^n), \qquad f \mapsto \{\hat{f}(k)\}_{k\in\mathbb{Z}^n}.$$

Moreover, the inverse is given by

$$\{\hat{f}(k)\}_{k\in\mathbb{Z}^n} \mapsto \sum_{k\in\mathbb{Z}^n} \hat{f}(k) e^{2\pi i k\cdot x},$$

with convergence in $L^2(\mathbb{T}^n)$.

Theorem 4.25. *(Periodization Theorem)* Let $f \in L^1(\mathbb{R}^n)$ and define $Pf : \mathbb{T}^n \to \mathbb{C}$ by

$$Pf(x) = \sum_{m \in \mathbb{Z}^n} f(x - m)$$

so that Pf is $\mathbb{Z}^n$–periodic. Then

- $\|Pf\|_{L^1(\mathbb{T}^n)} \leq \|f\|_{L^1(\mathbb{R}^n)}$ and Pf is well-defined a.e.
- Defining
$$P_N f(x) = \sum_{\substack{\|m\|_\infty \leq N \\ m \in \mathbb{Z}^n}} f(x - m)$$
we have $P_N f \to Pf$ a.e. and in $L^1(\mathbb{T}^n)$
- $\widehat{Pf}(k) = \hat{f}(k)$ for all $k \in \mathbb{Z}^n$

Proof.

- By definition we write:
$$\begin{aligned}\|Pf\|_{L^1(\mathbb{T}^n)} &= \int_Q |Pf(x)| \, \mathrm{d}x \leq \int_Q \sum_{m \in \mathbb{Z}^n} |f(x - m)| = \sum_{m \in \mathbb{Z}^n} \int_Q |f(x - m)| \, \mathrm{d}x \\ &= \sum_{m \in \mathbb{Z}^n} \int_{Q+m} |f(x)| \, \mathrm{d}x = \int_{\mathbb{R}^n} |f(x)| \, \mathrm{d}x = \|f\|_{L^1(\mathbb{R}^n)}.\end{aligned}$$
Where we have interchanged the order of summation and integration with an appeal to Tonelli. In particular, $g(x) := \sum_{m \in \mathbb{Z}^n} |f(x - m)| < \infty$ a.e. and so Pf is well-defined a.e.
- As in the previous bullet, $Pf(x)$ is absolutely convergent a.e., hence $P_N f(x) \to Pf(x)$ a.e. Moreover,
$$|P_N f(x) - Pf(x)| \leq |P_N f(x)| + |Pf(x)| \leq 2g(x) \in L^1(\mathbb{T}^n).$$
Via convergence a.e. and bounded by an L^1 function, the D.C.T. yields $P_N f \to Pf$ in $L^1(\mathbb{T}^n)$.
- Again we appeal to Fubini to write
$$\begin{aligned}\widehat{Pf}(k) &= \int_Q Pf(x) e^{-2\pi i k \cdot x} \, \mathrm{d}x = \int_Q \sum_{m \in \mathbb{Z}^n} f(x - m) e^{-2\pi i k \cdot x} \, \mathrm{d}x \\ &= \sum_{m \in \mathbb{Z}^b} \int_Q f(x - m) e^{-2\pi i k \cdot x} \, \mathrm{d}x = \sum_{m \in \mathbb{Z}^n} \int_{Q+m} f(x) e^{-2\pi i k \cdot x} \, \mathrm{d}x \\ &= \int_{\mathbb{R}^n} f(x) e^{-2\pi i k \cdot x} \, \mathrm{d}x = \hat{f}(k).\end{aligned}$$
Where we also use that $e^{2\pi i k m} = 1$.

□

Lemma 4.26. If $f, g \in L^1(\mathbb{T}^n)$ and $\hat{f}(k) = \hat{g}(k)$ for all $k \in \mathbb{Z}^n$, then $f = g$ a.e. on $\mathbb{T}^n$

Proof. Omitted. □

Theorem 4.27. *(Poisson Summation Formula)* Let $f \in L^1(\mathbb{R}^n)$ be such that $\sum_{k\in\mathbb{Z}^n} |\hat{f}(k)| < \infty$ and $Pf(x) = \sum_{m\in\mathbb{Z}^n} f(x-m)$ is continuous at $x = 0$. Then

$$\sum_{m\in\mathbb{Z}^n} f(m) = \sum_{k\in\mathbb{Z}^n} \hat{f}(k).$$

Proof. By the preceding theorem, Pf is well-defined a.e. and $Pf \in L^1(\mathbb{T}^n)$. Moreover

$$G(x) := \sum_{k\in\mathbb{Z}^n} \hat{f}(k) e^{2\pi i k\cdot x}$$

is absolutely and uniformly convergent on $\mathbb{T}^n$ (which follows from $\sum_{k\in\mathbb{Z}^n} |\hat{f}(k)| < \infty$). Looking at partial sums which are (continuous functions) on the torus (a compact set), uniformly converging to $G(x)$ we have that G is continuous. Now, G has Fourier coefficients $\hat{f}(k)$ and the Pf has Fourier coefficients $\widehat{Pf}(k) = \hat{f}(k)$. By Lemma 4.26 we have that $Pf = G$ a.e. on $\mathbb{T}^n$. This combined with the continuity of G and the fact that Pf is continuous at 0 implies that $G(0) = Pf(0)$. Finally,

$$\sum_{k\in\mathbb{Z}^b} \hat{f}(k) = G(0) = Pf(0) = \sum_{m\in\mathbb{Z}^n} f(0-m) = \sum_{m\in\mathbb{Z}^n} f(m).$$

□

Note 4.28.

(i) One can show that Schwartz functions satisfy the hypotheses of Theorem 4.27

(ii) One can prove an even more general result:
If $\sum_{m\in\mathbb{Z}^n} |f(m)| < \infty$ *and* $\sum_{k\in\mathbb{Z}^n} |\hat{f}(k)| < \infty$ *then* $\sum f(m) = \sum \hat{f}(k)$.

(iii) As an example of the Poisson Summation Formulas's utility, consider $f(x) = e^{-2\pi|x|}$ on $\mathbb{R}$. One can show that $\hat{f}(\xi) = \frac{1}{\pi(1+\xi^2)}$. Hence

$$\frac{1}{\pi}\sum_{k\in\mathbb{Z}} \frac{1}{1+k^2} = \sum_{k\in\mathbb{Z}} \hat{f}(k) = \sum_{m\in\mathbb{Z}} f(m) = \sum_{m\in\mathbb{Z}} e^{-2\pi|m|} = 1 + \sum_{m=1}^{\infty} (e^{-2\pi})^m = \frac{e^{2\pi}+1}{e^{2\pi}-1},$$

where one should first verify both the LHS and RHS are absolutely convergent (as they are).

5 Further Topics

5.1 Van der Corput's Lemmas

First some motivation. Write

$$F(R) = \int_a^b e^{iRf(x)}\,\mathrm{d}x$$

where $f : \mathbb{R} \to \mathbb{R}$, we are interested in finding conditions on f so that F is bounded. First of all notice that $\int_0^{2\pi} e^{ix}\,\mathrm{d}x = 0$ and that therefore $\int_0^N e^{ix}\,\mathrm{d}x$ is bounded. What if we ask for which f is the integral $\int_0^N e^{if(x)}\,\mathrm{d}x$ bounded? This can be answered by analyzing $F(R)$ as $R \to \infty$.

Example 5.1. Define the integral

$$I_N = \int_0^N e^{ix^2}\,\mathrm{d}x.$$

Can we bound $|I_N|$ independent of N?

- First we peal off
$$\left|\int_0^1 e^{ix^2}\,\mathrm{d}x\right| \le \int_0^1 1\,\mathrm{d}x = 1.$$
- Now for $N > 1$
$$\int_1^N e^{ix^2}\,\mathrm{d}x = \int_1^N \frac{1}{2ix}\left[\frac{\mathrm{d}}{\mathrm{d}x}\left(e^{ix^2}\right)\right]\mathrm{d}x \overset{\text{IBP}}{=} \underbrace{\frac{e^{ix^2}}{2ix}\bigg|_1^N}_{\mathrm{I}} - \underbrace{\frac{1}{2i}\int_1^N e^{ix^2}\frac{1}{x^2}\,\mathrm{d}x}_{\mathrm{II}}.$$

  We can then bound
$$|\,\mathrm{I}\,| \le \left|\frac{e^{iN^2}}{2iN}\right| + \left|\frac{e^i}{2i}\right| = \frac{1}{2N} + \frac{1}{2} \le 1$$

  and
$$|\,\mathrm{II}\,| \le \frac{1}{2}\int_1^N \frac{1}{x^2}\,\mathrm{d}x = \frac{1}{2}\left(-\frac{1}{x}\right)\bigg|_1^N = \frac{1}{2}\left(1 - \frac{1}{N}\right) \le \frac{1}{2}$$

Altogether we see that

$$\left|\int_0^N e^{ix^2}\,\mathrm{d}x\right| \le 2 + \frac{1}{2} \qquad \forall\, N \in \mathbb{N}.$$

Lemma 5.2. *(Van der Corput [1])* Let $a < b$, $f \in C^1([a,b])$, $f' \ge 1$ and f' be strictly monotonic. Then

$$\left|\int_a^b e^{iRf(x)}\,\mathrm{d}x\right| \le \frac{4}{R}.$$

Proof. Since f is C^1 we may write:

$$F(R) = \int_a^b e^{iRf(x)}\,\mathrm{d}x = \int_a^b \left[\frac{\mathrm{d}}{\mathrm{d}x}\left(e^{iRf(x)}\right)\right]\frac{1}{iRf'(x)}\,\mathrm{d}x.$$

Then integrating by parts yields:

$$\underbrace{\frac{e^{iRf(x)}}{iRf'(x)}\bigg|_a^b}_{\mathrm{I}} - \underbrace{\frac{1}{iR}\int_a^b e^{iRf(x)}\left[\frac{\mathrm{d}}{\mathrm{d}x}\left(\frac{1}{f'(x)}\right)\right]\mathrm{d}x}_{\mathrm{II}}$$

We may assume that $R \gg 1$. Now since $f' \geq 1$ we write

$$|\mathrm{I}| \leq \left| \frac{e^{iRf(b)}}{iRf'(b)} \right| + \left| \frac{e^{iRf(a)}}{iRf'(a)} \right| \leq \frac{1}{R} + \frac{1}{R} = \frac{2}{R}.$$

And since f' is strictly monotone we have that $\frac{\mathrm{d}}{\mathrm{d}x}\left(\frac{1}{f'(x)}\right)$ is always the same sign (therefore we can pull out the absolute values and appeal to the Fundamental Theorem of Calculus).

$$|\mathrm{II}| \leq \frac{1}{R}\int_a^b \left| \frac{\mathrm{d}}{\mathrm{d}x}\left(\frac{1}{f'(x)}\right)\right| \mathrm{d}x = \frac{1}{R}\left|\int_a^b \frac{\mathrm{d}}{\mathrm{d}x}\left(\frac{1}{f'(x)}\right)\mathrm{d}x\right| = \frac{1}{R}\left|\frac{1}{f'(x)}\right|_a^b\Bigg| = \frac{1}{R}\left|\frac{1}{f'(b)} - \frac{1}{f'(a)}\right| \leq \frac{2}{R}.$$

□

We can easily extend this to suppose that $f'(x) \geq \delta > 0$. In which case

$$\left|\int_a^b e^{iRf(x)}\,\mathrm{d}x\right| = \left|\int_a^b e^{i\delta R\frac{f(x)}{\delta}}\,\mathrm{d}x\right| \leq \frac{4}{\delta R}.$$

Lemma 5.3. *(Van der Corput [2])* Let $a < b$ and $f \in C^2([a,b])$ with $f'' \geq 1$, then

$$\left|\int_a^b e^{iRf(x)}\,\mathrm{d}x\right| \leq \frac{10}{\sqrt{R}}.$$

Proof. Notice that $f'' = (f')' \geq 1$ and hence f' is strictly increasing. Namely there can be at most one x_0 such that $f'(x_0) = 0$. Then, assuming x_0 is neither a nor b, we can decompose the integral

$$\int_a^b e^{iRf(x)}\,\mathrm{d}x = \underbrace{\int_a^{x_0-\delta} e^{iRf(x)}\,\mathrm{d}x}_{\mathrm{I}} + \underbrace{\int_{x_0-\delta}^{x_0+\delta} e^{iRf(x)}\,\mathrm{d}x}_{\mathrm{II}} + \underbrace{\int_{x_0+\delta}^{b} e^{iRf(x)}\,\mathrm{d}x}_{\mathrm{III}}$$

for any $\delta > 0$. Now, appealing to the MVT, for any $x \in (x_0+\delta, b)$ we may write

$$f'(x) = f'(x) - f'(x_0) = \underbrace{f''(\mu)}_{\geq 1}\underbrace{(x-x_0)}_{\geq \delta} \geq \delta.$$

Hence by Van der Corput [1] we have $|\mathrm{III}| \leq 4/R\delta$. Similarly one can show that $|\mathrm{I}| \leq 4/R\delta$. We also have the bound

$$|\mathrm{II}| \leq \int_{x_0-\delta}^{x_0+\delta} 1\,\mathrm{d}x = 2\delta.$$

Taken together we obtain

$$\left|\int_a^b e^{iRf(x)}\,\mathrm{d}x\right| \leq \frac{8}{R\delta} + 2\delta = \frac{10}{\sqrt{R}}$$

where we chose $\delta = R^{-1/2}$. [Note that if $x_0 = a$ or $x_0 = b$ a similar idea can be applied to decompose the integral that achieves an even better bound]. □

Note 5.4. One can prove the following "m–Van der Corput Lemma". Let $a < b$ and $f \in C^m([a,b])$ with $f^{(m)} \geq 1$. Then

$$\left|\int_a^b e^{iRf(x)}\,\mathrm{d}x\right| \leq \frac{C_m}{R^{1/m}}.$$

And moreover the decay rate $R^{-1/m}$ cannot be improved.

Example 5.5. Consider the unit circle $\mathbb{S}^1$ in $\mathbb{R}^2 \sim \mathbb{C}$. Let σ be the arc length measure on the unit circle. We define

$$\hat{\sigma}(\xi) = \int_{\mathbb{S}^1} e^{-2\pi i x\cdot\xi}\,\mathrm{d}\sigma(x) = \int_0^{2\pi} e^{-2\pi i(\cos\theta\xi_1+\sin\theta\xi_2)}\,\mathrm{d}\theta =: I.$$

We seek a bound of the form

$$|\hat{\sigma}(\xi)| \le C(1+|\xi|)^{-\gamma}$$

for some $\gamma > 0$. We write things in such a way as to apply Van der Corput's Lemmas:

$$I = \int_0^{2\pi} e^{i\xi_2(-2\pi)\left(\cos\theta\left(\frac{\xi_1}{\xi_2}\right)+\sin\theta\right)}\,\mathrm{d}\theta$$

where we identify ξ_2 with R and $-2\pi\left(\cos\theta\left(\frac{\xi_1}{\xi_2}\right)+\sin\theta\right)$ with $f(\theta)$. Now

$$f'(\theta) = -2\pi\left(-\sin\theta\left(\frac{\xi_1}{\xi_2}\right)+\cos\theta\right) \qquad \text{and} \qquad f''(\theta) = 2\pi\left(\cos\theta\left(\frac{\xi_1}{\xi_2}\right)+\sin\theta\right)$$

both of which may vanish, making the Van der Corput lemmas difficult to apply. However, one can show the following:

(i) there exists a $c > 0$ such that $\max\{|f'(\theta)|, |f''(\theta)|\} \ge c$ for all θ

(ii) using (i) and the Van der Corput lemmas, one can show that $|\hat{\sigma}(\xi)| \le C(1+|\xi|_2)^{-1/2}$.

If $\xi_1 = 0$ then (ii) yields are desired bound with $\gamma = 1/2$. We can appeal to rotational symmetry to finish the argument. Let A be any rotation and denote $\omega := (\cos\theta, \sin\theta)$. Then

$$\hat{\sigma}(A\xi) = \int_0^{2\pi} e^{-2\pi i\omega\cdot A\xi}\,\mathrm{d}\omega = \int_0^{2\pi} e^{-2\pi i\xi\cdot(A^\top\omega)}\,\mathrm{d}\omega = \int_0^{2\pi} e^{-2\pi i\xi\cdot\omega'}\,\mathrm{d}\omega' = \hat{\sigma}(\xi).$$

Here we use the fact that

$$\omega\cdot A\xi = \omega^\top A\xi = (A\xi)^\top\omega = \xi^\top A^\top\omega = \xi\cdot(A^\top\omega)$$

and that ω traces around a circle and so rotating ω does not affect the value of the integral. We can then choose the rotation A so that $A\xi = (0, |\xi|)$, and then by (ii) we have

$$|\hat{\sigma}(\xi)| \le C(1+|\xi|)^{-1/2}$$

as desired.

5.2 Interpolation Theory

Definition 5.6. Let $\Omega \subset \mathbb{C}$ be open. A function $f : \Omega \to \mathbb{C}$ is said to be **holomorphic** if at every point $z_0 \in \Omega$ the following limit exists

$$\lim_{h\to 0}\frac{f(z_0+h)-f(z_0)}{h}.$$

In which case we denote its value $f'(z_0)$.

Lemma 5.7. (*Hadamard three-line*) Let $F : \overline{S} = \{z \in \mathbb{C} : 0 \leq \mathrm{Re}(z) \leq 1\} \to \mathbb{C}$ be holomorphic on $S = \{z \in \mathbb{C} : 0 < \mathrm{Re}(z) < 1\}$, continuous and bounded and $\overline{S}$, and suppose there exist constants A_0, A_1 such that $|F(it)| \leq A_0$ and $|F(1+it)| \leq A_1$ for all $t \in \mathbb{R}$. Then

$$|F(\theta + it)| \leq A_0^{1-\theta} A_1^{\theta}$$

for all $\theta \in [0,1]$ and $t \in \mathbb{R}$. In particular, by letting $t = 0$, we have $|F(\theta)| \leq A_0^{1-\theta} A_1^{\theta}$ for all $\theta \in [0,1]$.

Theorem 5.8. *(Riesz-Thorin Interpolation Theorem)* Let (X, μ) and (Y, ν) be σ-finite measure spaces. Let $1 \leq p_0, p_1, q_0, q_1 \leq \infty$ and let $T : L^{p_0}(X) \cap L^{p_1}(X) \to L^{q_0}(Y) \cap L^{q_1}(Y)$ be a linear operator. Suppose there exist $M_0, M_1 \in \mathbb{R}$ such that

$$\|Tf\|_{L^{q_0}(Y)} \leq M_0 \|f\|_{L^{p_0}(X)} \qquad \text{and} \qquad \|Tf\|_{L^{q_1}(Y)} \leq M_1 \|f\|_{L^{p_1}(X)}$$

for all $f \in L^{p_0}(X) \cap L^{p_1}(X)$. Now let $\theta \in [0,1]$ and define p_θ and q_θ via

$$\frac{1}{p_\theta} = \frac{1-\theta}{p_0} + \frac{\theta}{p_1} \qquad \frac{1}{q_\theta} = \frac{1-\theta}{q_0} + \frac{\theta}{q_1}.$$

Then T extends to a bounded linear operator $T : L^{p_\theta}(X) \to L^{q_\theta}(Y)$ and

$$\|Tf\|_{L^{q_\theta}(Y)} \leq M_0^{1-\theta} M_1^{\theta} \|f\|_{L^{p_\theta}(X)}.$$

Proof. Let $f \in L^{p_\theta}(X)$ and $g \in L^{q_\theta'}(Y)$ where q_θ' denotes the conjugate exponent of q_θ. By duality we may write

$$\|Tf\|_{L^{q_\theta}(Y)} = \|\phi_{Tf}\| = \sup_{\|g\|_{L^{q_\theta'}(Y)}=1} \left| \int_Y T(f)\, g \, \mathrm{d}\nu \right|.$$

By density it suffices to assume that f and g are simple functions with finite measure support. Now write $f(x) = |f(x)| e^{i\alpha(x)}$ and $g(x) = |g(x)| e^{i\beta(x)}$ where α and β are appropriate rotations. For complex z such that $0 \leq \mathrm{Re}(z) \leq 1$ we define

$$f_z(x) := |f(x)|^{p_\theta / p(z)} e^{i\alpha(x)} \qquad \text{and} \qquad g_z(x) := |g(x)|^{q_\theta' / q'(z)} e^{i\beta(x)}$$

where $p(z)$ and $q(z)$ are defined via

$$\frac{1}{p(z)} = \frac{1 - \mathrm{Re}(z)}{p_0} + \frac{\mathrm{Re}(z)}{p_1} \qquad \text{and} \qquad \frac{1}{q(z)} = \frac{1 - \mathrm{Re}(z)}{q_0} + \frac{\mathrm{Re}(z)}{q_1}$$

and $q'(z)$ is the conjugate exponent of $q(z)$. Now define

$$F(z) = \int_Y T(f_z)\, g_z \, \mathrm{d}\nu$$

for which one can check the conditions of Hadamard. Indeed, since f_z and g_z depend analytically on z, and T is linear, it follows that $F(z)$ is holomorphic in the strip $0 < \mathrm{Re}(z) < 1$, and continuous and bounded on the closed strip. Now at the boundary. If $\mathrm{Re}(z) = 0$, then $p(z) = p_0$ and $q(z) = q_0$, and hence

$$|F(z)| = \left| \int_Y T(f_z)\, g_z \, \mathrm{d}\nu \right| \leq \|T(f_z)\|_{L^{q_0}(Y)} \|g_z\|_{L^{q_0'}(Y)} \leq M_0 \|f_z\|_{L^{p_0}(X)} \|g_z\|_{L^{q_0'}(Y)}.$$

Now by construction we have

$$\|f_z\|_{L^{p_0}(X)} = \|f\|_{L^{p_\theta}(X)}^{\frac{p_\theta}{p_0}} \qquad \text{and} \qquad \|g_z\|_{L^{q_0'}(Y)} = \|g\|_{L^{q_\theta'}(Y)}^{\frac{q_\theta'}{q_0'}}.$$

Thus

$$|F(z)| \le M_0 \|f\|_{L^{p_\theta}(X)}^{\frac{p_\theta}{p_0}} \|g\|_{L^{q_\theta'}(Y)}^{\frac{q_\theta'}{q_0'}}.$$

Similarly, if $\operatorname{Re}(z) = 1$, then $p(z) = p_1$ and $q(z) = q_1$, and hence

$$|F(z)| \le M_1 \|f\|_{L^{p_\theta}(X)}^{\frac{p_\theta}{p_1}} \|g\|_{L^{q_\theta'}(Y)}^{\frac{q_\theta'}{q_1'}}.$$

Thus, by the Hadamard three-line lemma, we obtain that for $\theta \in [0,1]$

$$\begin{aligned} |F(\theta)| &\le \left(M_0 \|f\|_{L^{p_\theta}(X)}^{\frac{p_\theta}{p_0}} \|g\|_{L^{q_\theta'}(Y)}^{\frac{q_\theta'}{q_0'}} \right)^{1-\theta} \left(M_1 \|f\|_{L^{p_\theta}(X)}^{\frac{p_\theta}{p_1}} \|g\|_{L^{q_\theta'}(Y)}^{\frac{q_\theta'}{q_1'}} \right)^{\theta} \\ &= M_0^{1-\theta} M_1^{\theta} \|f\|_{L^{p_\theta}(X)}^{(1-\theta)\frac{p_\theta}{p_0} + \theta \frac{p_\theta}{p_1}} \|g\|_{L^{q_\theta'}(Y)}^{(1-\theta)\frac{q_\theta'}{q_0'} + \theta \frac{q_\theta'}{q_1'}}. \end{aligned}$$

Now by the definition of p_θ and q_θ we have

$$(1-\theta)\frac{p_\theta}{p_0} + \theta \frac{p_\theta}{p_1} = 1 \qquad \text{and} \qquad (1-\theta)\frac{q_\theta'}{q_0'} + \theta \frac{q_\theta'}{q_1'} = 1,$$

and hence

$$|F(\theta)| \le M_0^{1-\theta} M_1^{\theta} \|f\|_{L^{p_\theta}(X)} \|g\|_{L^{q_\theta'}(Y)}.$$

Finally, observing that $F(\theta) = \int_Y T(f)\, g \, \mathrm{d}\nu$, we obtain

$$\left| \int_Y T(f)\, g \, \mathrm{d}\nu \right| \le M_0^{1-\theta} M_1^{\theta} \|f\|_{L^{p_\theta}(X)} \|g\|_{L^{q_\theta'}(Y)}.$$

Taking the supremum over all g such that $\|g\|_{L^{q_\theta'}(Y)} = 1$, we conclude that

$$\|Tf\|_{L^{q_\theta}(Y)} \le M_0^{1-\theta} M_1^{\theta} \|f\|_{L^{p_\theta}(X)}.$$

This completes the proof. □

5.3 Distribution Theory

Theorem 5.9. *(Riesz-Markov Representation Theorem)* Let X be locally compact (every point in X has a locally compact neighbourhood), and Hausdorff (points can be separated with disjoint open sets). Consider $C_0(X)$ (the set of complex valued functions on X which vanish at infinity) equipped with the supremum norm $\|f\|_\infty = \operatorname{ess\,sup}_{x \in X} |f(x)|$. If $L : C_0(X) \to \mathbb{C}$ is a continuous linear functional, then there is a unique complex Borel measure μ on X such that

$$L(\varphi) = \int_X \varphi \, \mathrm{d}\mu$$

for all $\varphi \in C_0(X)$ and $\|L\| = |\mu|(X)$. Moreover, if μ is absolutely continuous w.r.t. the Lebesgue measure, then by R-N we may write $L(\varphi) = \int_X \varphi \, \mathrm{d}\mu = \int_X \varphi f \, \mathrm{d}x$ for some $f \in L^1$.

Proof. Omitted. □

Note 5.10. Now some motivation for Distribution Theory: we can recover functions in L^1_{loc} (almost everywhere) by using convolutions with "bump" functions.

- Let $f \in L^1_{\text{loc}}(\mathbb{R}^n)$ and $\varphi \in C^\infty_C(\mathbb{R}^n)$ such that φ is "concentrated" near 0, $\varphi \geq 0$, and $\int \varphi(x)\,\mathrm{d}x = 1$. Then we can see the convolution

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

  as a weighted average of f near x.

- For $\varepsilon > 0$ we define the "bump functions"

$$\varphi_\varepsilon(x) = \frac{1}{\varepsilon^n}\varphi\left(\frac{x}{\varepsilon}\right)$$

  where the inside scaling factor $\frac{1}{\varepsilon}$ concentrates φ_ε in a ball of proportional to ε, and the outside scaling factor $\frac{1}{\varepsilon^n}$ ensures φ_ε still integrates to 1.

- Recall that a point x is Lebesgue for f if

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

  and that for $f \in L^1_{\text{loc}}$, almost every point is a Lesbesgue point. In particular, if x is a Lebesgue point we have

$$\begin{aligned}
\lim_{r\to\infty}\left|\int \frac{\chi_{B_r(x)}}{|B_r(x)|}f(y)\,\mathrm{d}y - f(x)\right| &= \lim_{r\to 0}\left|\frac{1}{|B_r(x)|}\int_{B_r(x)} f(y)\,\mathrm{d}y - f(x)\right| \\
&= \lim_{r\to 0}\left|\frac{1}{|B_r(x)|}\int_{B_r(x)} f(y) - f(x)\,\mathrm{d}y\right| \\
&\leq \lim_{r\to 0}\frac{1}{B_r(x)}\int_{B_r(x)} |f(y) - f(x)|\,\mathrm{d}y = 0
\end{aligned}$$

  And so, a.e., L^1_{loc} functions are asymptotically equal to the average of their nearby values.

- In particular we have

$$\lim_{\varepsilon\to 0}(f * \varphi_\varepsilon)(x) = f(x)$$

  for almost every x. And it is in this sense that $\{\varphi_\varepsilon\}_{\varepsilon>0}$ is called an "approximate identity". We can write

$$f(x) \overset{\text{a.e.}}{=} \lim_{\varepsilon\to\infty} f(y)\varphi_{\varepsilon,x}(y)\,\mathrm{d}y,$$

  where $\varphi_{\varepsilon,x}(y) := \varphi_\varepsilon(x-y)$ is C^∞_C.

- This motivates looking at functions associated with f

$$F : C^\infty_C \to \mathbb{C}$$

  given by

$$F(\varphi) = \int_{\mathbb{R}^n} f(x)\varphi(x)\,\mathrm{d}x$$

  instead of f evaluated at points.

Definition 5.11. Let $\Omega \subset \mathbb{R}^n$ be open. The space of **test functions** is $\mathcal{D} = \mathcal{D}(\Omega) = C_C^\infty(\Omega)$, the set of smooth functions $\varphi : \Omega \to \mathbb{C}$ with compact support contained in Ω. We say that $\varphi_m \to \varphi$ in $\mathcal{D}$ if

(i) there exists a compact $K \subset \Omega$ such that $\operatorname{supp}(\varphi_m) \subset K$ for all $m \in \mathbb{N}$

(ii) $\partial^\alpha \varphi_m \to \partial^\alpha \varphi$ uniformly on K as $m \to \infty$, for every multi-index α

A **distribution** on Ω is a linear functional

$$F : \mathcal{D}(\Omega) \to \mathbb{C}$$

such that $F(\varphi_m) \to F(\varphi)$ whenever $\varphi_m \to \varphi$ in $\mathcal{D}$. The space of all distributions on Ω is denoted $\mathcal{D}'(\Omega)$.

Note 5.12.

(i) We saw that any $f \in L^1_{\text{loc}}$ can be identified with a distribution.

(ii) The space of distributions on Ω is a linear space.

Example 5.13. *(A distribution not arising from $f \in L^1_{\text{loc}}$)* Let μ be a Borel measure that is finite on compact subsets of Ω. Then $F(\varphi) = \int \varphi(x)\,\mathrm{d}\mu(x)$ is a distribution on Ω. Clearly it is linear. Now suppose $\varphi_m \to \varphi$ in $\mathcal{D}$. Then $\operatorname{supp}(\varphi_m) \subset K$ compact and $\|\varphi_m - \varphi\|_{L^\infty(K)} \to 0$. So

$$|F(\varphi_m) - F(\varphi)| = \left| \int_K \varphi_m - \varphi \,\mathrm{d}\mu \right| \leq \|\varphi_m - \varphi\|_{L^\infty(K)} \mu(K) \to 0$$

as desired. Without loss of generality, assume that $0 \in \Omega$, and take $\mu = \delta_0$. Then

$$F(\varphi) = \int \varphi(x)\,\mathrm{d}\mu = \varphi(0).$$

By way of contradiction, suppose there exists some $f \in L^1_{\text{loc}}$ such that

$$\varphi(0) = F(\varphi) = \int f(x)\varphi(x)\,\mathrm{d}x$$

for all $\varphi \in \mathcal{D}(\Omega)$. Since Ω is open and contains 0, there is a ball $B_r(0) \subset \Omega$. Fix $\eta \in \mathcal{D}(\Omega)$ satisfying $\eta(0) = 1$, $0 \leq \eta \leq 1$ and $\operatorname{supp}(\eta) \subset B_1(0)$. Then

$$\varphi_m(x) := \eta\left(\frac{2m}{r}x\right) \in \mathcal{D}(\Omega) \qquad \forall\, m \in \mathbb{N},$$

$\varphi_m(0) = 1$ and $\operatorname{supp}(\varphi_m) \subset B_{\frac{r}{2m}}(0)$. But then

$$1 = |\varphi_m(0)| = \left| \int f(x)\varphi_m(x)\,\mathrm{d}x \right| \leq \int |f(x)| \cdot 1\,\mathrm{d}x \to 0$$

as $m \to \infty$. A contradiction.

Example 5.14. Let $\Omega \subset \mathbb{R}^n$ be open. Let $f \in C^\infty(\Omega)$, $\varphi \in C_C^\infty(\Omega)$, and let α be a multi-index. Define the distribution $F(\varphi) = \int f(x)\varphi(x)\,\mathrm{d}x$. How do we take the derivative of F? By integration by parts $|\alpha|$ times (and noting the boundary terms vanish since φ has compact support) we obtain

$$\int \left(\partial_x^\alpha f(x)\right) \varphi(x)\,\mathrm{d}x = (-1)^{|\alpha|} \int f(x) \left(\partial_x^\alpha \varphi(x)\right) \mathrm{d}x = (-1)^{|\alpha|} F(\partial^\alpha \varphi).$$

This motivates the following definition.

Definition 5.15. Let $\Omega \subset \mathbb{R}^n$ be open and let $F \in \mathcal{D}'(\Omega)$ For a multi-index α define the **distributional derivative of** F

$$(\partial^\alpha F)(\varphi) = (-1)^{|\alpha|} F(\partial^\alpha \varphi) \quad \forall\, \varphi \in \mathcal{D}(\Omega).$$

This is well-defined since $\varphi \in \mathcal{D}(\Omega)$, so is $\partial^\alpha \varphi$. One should also verify that $\partial^\alpha F$ is again linear and sequentially continuous, hence a distribution.

Example 5.16. We show that the distributional derivative is distinct from the derivative of the underlying function. As an example, consider the Heaviside function

$$h(x) = \begin{cases} 1 & x > 0 \\ 0 & x \leq 0 \end{cases}.$$

As a function $h'(x) = 0$ a.e. except at $x = 0$ where the derivative does not exist. But as a distribution:

$$H(\varphi) = \int h(x)\varphi(x)\,\mathrm{d}x = \int_0^\infty \varphi(x)\,\mathrm{d}x$$

and so

$$H'(\varphi) = -1H(\varphi') = -\int_0^\infty \varphi'(x) = -\varphi(x)|_0^\infty = 0 - (-\varphi(0)) = \varphi(0)$$

Proposition 5.17. Let $\Omega \subset \mathbb{R}^n$ be open, $\mathcal{D}(\Omega)$ the set of test functions on Ω and let F be a distribution on Ω arising from a function f. Let $\varphi, \psi \in \mathcal{D}(\Omega)$.

- Let τ_h denote translation by h, i.e. $(\tau_h f)(x) = f(x - h)$
- Let s_a denote dilation by a for $a > 0$, i.e. $(s_a f)(x) = f(ax)$
- Let T_L denote $(T_L f)(x) = f(Lx)$ where $L : \mathbb{R}^n \to \mathbb{R}^m$ is linear with $\det(L) \neq 0$
- Let T_A denote $(T_A f)(x) = f(Ax)$ where $A(x) = Lx + b$ with $\det(L) \neq 0$.
- Let M_ψ denote multiplication by $\psi \in \mathcal{D}(\Omega)$, i.e., $(M_\psi f)(x) = \psi(x) f(x)$.

For any transformation T, let $(TF)(\varphi) = \int (Tf)\varphi$. Then

(i) $(\tau_h F)(\varphi) = F(\tau_{-h}\varphi)$

(ii) $(s_a F)(\varphi) = \frac{1}{a^n} F(s_{1/a}\varphi)$

(iii) $(T_L F)(\varphi) = |\det L|^{-1} F(T_{L^{-1}}\varphi)$

(iv) $(T_A F)(\varphi) = |\det L|^{-1} F(T_{A^{-1}}\varphi)$

(v) $(\psi F)(\varphi) = F(\psi\varphi)$

Proof.

(i) By the change of variables $y = x - h \implies \mathrm{d}y = \mathrm{d}x$:

$$(\tau_h F)(\varphi) = \int f(x - h)\varphi(x)\,\mathrm{d}x = \int f(y)\varphi(x + h)\,\mathrm{d}x = F(\tau_{-h}\varphi).$$

(ii) By change of variables $y = ax \implies \mathrm{d}y = a^n \,\mathrm{d}x$:

$$(s_a F)(\varphi) = \int f(ax)\varphi(x)\,\mathrm{d}x = \frac{1}{a^n}\int f(y)\varphi(y/a)\,\mathrm{d}y = \frac{1}{a^n}F(s_{1/a}\varphi).$$

(iii) By change of variables $y = Lx \implies \mathrm{d}y = |\det L|\,\mathrm{d}x$:

$$(T_L F)(\varphi) = \int f(Lx)\varphi(x)\,\mathrm{d}x = |\det L|^{-1}\int f(y)\varphi(L^{-1}y)\,\mathrm{d}y = |\det L|^{-1}F(T_{L^{-1}}\varphi).$$

(iv) By change of variables $y = Ax = Lx + b \implies x = A^{-1}y = L^{-1}(y-b)$ and $dx = |\det L|^{-1}dy$:

$$(T_A F)(\varphi) = \int f(Ax)\varphi(x)dx = |\det L|^{-1}\int f(y)\varphi(A^{-1}y)dy = |\det L|^{-1}F(T_{A^{-1}}\varphi).$$

(v) By the associativity of multiplication:

$$(\psi F)(\varphi) = \int (\psi(x)f(x))\varphi(x)dx = \int f(x)(\psi(x)\varphi(x))dx = F(\psi\varphi).$$

□

This motivates the definition for general distributions F not necessarily arising from a function f.

Definition 5.18. Let $\Omega \subset \mathbb{R}^n$ be open, $\mathcal{D}(\Omega)$ the set of test functions on Ω and let F be a distribution on Ω. Then

(i) $(\tau_h F)(\varphi) = F(\tau_{-h}\varphi)$

(ii) $(s_a F)(\varphi) = \frac{1}{a^n}F(s_{1/a}\varphi)$

(iii) $(T_L F)(\varphi) = |\det L|^{-1}F(T_{L^{-1}}\varphi)$

(iv) $(T_A F)(\varphi) = |\det L|^{-1}F(T_{A^{-1}}\varphi)$

(v) $(\psi F)(\varphi) = F(\psi\varphi)$ for any $\psi \in \mathcal{D}(\Omega)$.

Definition 5.19. The **support of a function** $f : \Omega \to \mathbb{R}$, denoted $\mathrm{supp}(f)$, is the closure of the set of points where f is non-zero:

$$\mathrm{supp}(f) = \overline{\{x \in \Omega : f(x) \neq 0\}}.$$

Equivalently it is the smallest closed set outside of which f is identically zero.

Definition 5.20. The **support of a distribution** F is the complement of the largest open set U such that $F(\varphi) = 0$ for all test functions φ with $\mathrm{supp}(\varphi) \subset U$. Equivalently, $x \notin \mathrm{supp}(F)$ if there exists an open $U \ni x$ such that $F(\varphi) = 0$ for all test functions φ with $\mathrm{supp}(\varphi) \subset U$.

Let φ be a test function and with $\mathrm{supp}(\varphi) = K$ compact. Then define $\varphi_x(y) = \varphi(x - y)$ so that

$$\mathrm{supp}(\varphi_x) = x - K = \{x - k : k \in K\}$$

is compact also.

Definition 5.21. Let $\Omega \subset \mathbb{R}^n$ be open. Let F be a distribution on Ω and φ a test function on Ω. The **convolution** is a function $F * \varphi : \mathbb{R}^n \to \mathbb{R}$ given by $(F * \varphi)(x) = F(\varphi_x)$.

This definition is consistent with when F arrises from a function f. Indeed:

$$(F * \varphi)(x) = F(\varphi_x) = \int f(y)\varphi_x(y)\,\mathrm{d}y = \int f(y)\varphi(x-y)\,\mathrm{d}y = (f * \varphi)(x).$$

Theorem 5.22. Let $\Omega \subset \mathbb{R}^n$ be open, $F \in \mathcal{D}'(\Omega)$ a distribution, and $\varphi \in \mathcal{D}(\Omega)$ a test function. Then $F * \varphi : \mathbb{R}^n \to \mathbb{C}$ is smooth, and for every multi-index α,

$$\partial^\alpha (F * \varphi) = F * (\partial^\alpha \varphi).$$

Proof. First of all,

$$\partial_x^\alpha \varphi(x-y) = (\partial^\alpha \varphi)(x-y) = (\partial^\alpha \varphi)_x(y).$$

By induction, it suffices to show $\partial_i (F * \varphi) = F * (\partial_i \varphi)$. Let e_i be the i-th unit vector. For $h \neq 0$, the linearity of F gives:

$$\frac{(F * \varphi)(x + he_i) - (F * \varphi)(x)}{h} = F\left(\frac{\varphi_{x+he_i} - \varphi_x}{h}\right).$$

As $h \to 0$, the quotient $\frac{\varphi_{x+he_i} - \varphi_x}{h}$ converges to $(\partial_i \varphi)_x$ in the topology of $\mathcal{D}(\Omega)$. Since F is a distribution (and thus continuous on $\mathcal{D}(\Omega)$):

$$\partial_i (F * \varphi)(x) = \lim_{h \to 0} F\left(\frac{\varphi_{x+he_i} - \varphi_x}{h}\right) = F((\partial_i \varphi)_x) = (F * \partial_i \varphi)(x).$$

Repeating this for higher orders yields $\partial^\alpha (F * \varphi) = F * (\partial^\alpha \varphi)$. □

Theorem 5.23. Let $\Omega \subset \mathbb{R}^n$ be open, $F \in \mathcal{D}'(\Omega)$, and $\varphi \in \mathcal{D}(\Omega)$. Then

$$\operatorname{supp}(F * \varphi) \subset \operatorname{supp}(F) + \operatorname{supp}(\varphi),$$

where $A + B = \{a + b : a \in A,\, b \in B\}$ denotes the Minkowski sum.

Proof. Let $C = \operatorname{supp}(F)$ and $K = \operatorname{supp}(\varphi)$. It is equivalent to show that $(C+K)^c \subset (\operatorname{supp}(F * \varphi))^c$. Let $x_0 \notin C + K$. We need to show that $x_0 \notin \operatorname{supp}(F * \varphi)$. Since C is closed and K is compact, $C + K$ is closed. Hence there is an open neighbourhood W of x_0 such that $W \cap (C + K) = \emptyset$. It now suffices to show that $(F * \varphi)(x) = 0$ for all $x \in W$. Fix any $x \in W$, then $C \cap (x - K) = \emptyset$ (otherwise there would exist $c \in C$ such that $c = x - k$ with $k \in K$ which would imply $x = c + k \in C + K$, contradicting $W \cap (C + K) = \emptyset$). Since C is closed and $x - K$ is compact, there is an open $U \supset C$ such that $U \cap (x - K) = \emptyset$. Now $\operatorname{supp}(\varphi_x) = x - K$ is disjoint with U, namely $\operatorname{supp}(\varphi_x) \subset U^c \subset C^c$. And so $F(\varphi_x) = 0$ since C^c is by definition the largest open set such that $F(\psi) = 0$ for all test functions ψ satisfying $\operatorname{supp}(\psi) \subset C^c$. □

Definition 5.24. Let F be a distribution. We define its **reflection**, denoted $\tilde{F}$, by

$$\tilde{F}(\varphi) = F(\tilde{\varphi})$$

for any test function φ. Here $\tilde{\varphi}(x) = \varphi(-x)$ is the reflection of φ in the function sense. Note this can be shown compatible with the case where F arises from a function f via change of variables:

$$\int f(x)\varphi(-x)\,\mathrm{d}x = \int f(-x)\varphi(x)\,\mathrm{d}x.$$

Proposition 5.25. $\operatorname{supp}(\tilde{F}) = -\operatorname{supp}(F)$.

Proposition 5.26. It suffices to show that $\operatorname{supp}(\tilde{F}) \subset -\operatorname{supp}(F)$. Indeed, this would allow writing:

$$\tilde{\tilde{F}} = F \implies \operatorname{supp}(F) = \operatorname{supp}(\tilde{\tilde{F}}) \subset -\operatorname{supp}(\tilde{F}) \implies -\operatorname{supp}(F) \subset \operatorname{supp}(\tilde{F}).$$

To that end, it suffices to show that $(\operatorname{supp}(\tilde{F}))^c \supset (-\operatorname{supp}(F))^c$. Suppose $x \notin -\operatorname{supp}(F)$, that is, $-x \notin \operatorname{supp}(F)$. By definition there exists an open neighbourhood of $-x$, U, such that $F(\varphi) = 0$ for all test functions φ with $\operatorname{supp}(\varphi) \subset U$. Now, $-U$ is open and contains x. Moreover, if $\operatorname{supp}(\psi) \subset -U$, then $\operatorname{supp}(\tilde{\psi}) = -\operatorname{supp}(\psi) \subset U$. Hence $\tilde{F}(\psi) := F(\tilde{\psi}) = 0$. Therefore $x \notin \operatorname{supp}(\tilde{F})$ as we have found an open set (namely $-U$) such that $\tilde{F}(\psi) = 0$ for all ψ with $\operatorname{supp}(\psi) \subset -U$.

Definition 5.27. Let $\Omega \subset \mathbb{R}^n$ be open, $F \in \mathcal{D}'(\Omega)$, and $G \in \mathcal{D}'(\Omega)$ a distribution with compact support. The **convolution** $F * G \in \mathcal{D}'(\Omega)$ is defined by

$$(F * G)(\varphi) := F(\tilde{G} * \varphi) \quad \text{for all } \varphi \in \mathcal{D}(\Omega),$$

where $\tilde{G}(\psi) := G(\tilde{\psi})$ is the reflection of G, and $(\tilde{G} * \varphi)(x) := \tilde{G}(\varphi_x)$ with $\varphi_x(y) = \varphi(x - y)$. One can verify that $\tilde{G} * \varphi \in \mathcal{D}(\Omega)$, so the right-hand side is well-defined.

Note 5.28. We verify that the distributional convolution $F*G$ agrees with the classical convolution $f * g$ when F and G arise from L^1_{loc} functions f and g respectively, with g compactly supported. Let $\varphi \in \mathcal{D}(\Omega)$. First of all, for any $x \in \mathbb{R}^n$, using the definition of convolution of a distribution with a test function and then the definition of $\tilde{G}$:

$$(\tilde{G} * \varphi)(x) = \tilde{G}(\varphi_x) = G(\widetilde{\varphi_x}).$$

Now $\widetilde{\varphi_x}(y) = \varphi_x(-y) = \varphi(x + y)$, so

$$(\tilde{G} * \varphi)(x) = G(\varphi(x + \,\cdot\,)) = \int_{\mathbb{R}^n} g(y)\, \varphi(x + y)\, dy.$$

Now by definition of F as integration against f:

$$(F * G)(\varphi) = F(\tilde{G} * \varphi) = \int_{\mathbb{R}^n} f(x)\, (\tilde{G} * \varphi)(x)\, dx = \int_{\mathbb{R}^n} f(x) \int_{\mathbb{R}^n} g(y)\, \varphi(x + y)\, dy\, dx.$$

We now verify the conditions for and apply Fubini. To interchange the integrals, note that g has compact support $K_1 \subset \mathbb{R}^n$ (so $y \in K_1$) and φ has compact support $K_2 \subset \mathbb{R}^n$ (so $x + y \in K_2$, hence $x \in K_2 - K_1$, a compact set). Therefore, with $\|\varphi\|_\infty < \infty$:

$$\iint |f(x)|\, |g(y)|\, |\varphi(x + y)|\, dx\, dy \leq \|f\|_{L^1(K_2 - K_1)}\, \|g\|_{L^1(K_1)}\, \|\varphi\|_\infty < \infty,$$

so Fubini's theorem applies. Substituting $u = x + y$ (with x fixed) gives:

$$\int_{\mathbb{R}^n} g(y)\, \varphi(x + y)\, dy = \int_{\mathbb{R}^n} g(u - x)\, \varphi(u)\, du,$$

and therefore

$$(F * G)(\varphi) = \int_{\mathbb{R}^n} \int_{\mathbb{R}^n} f(x)\, g(u - x)\, dx\, \varphi(u)\, du = \int_{\mathbb{R}^n} (f * g)(u)\, \varphi(u)\, du,$$

which is exactly the action of the distribution associated to the classical convolution $f * g$ on φ.

The below two results are stated without proof.

- The convolution of two distributions cannot be defined in complete generality without assuming that one has compact support.
- If F and G are distributions that both have compact support, then $F * G = G * F$.