80600 - Analysis I
Matania Ben-Artzi

Asaf Karagila*

February 8, 2016

I will not include all proofs here. Things I consider these to be relatively trivial and easy to understand. Sometimes only an intuition for the proof will be given.

*Any mistakes in this document can be assumed to be the author's responsibility. Please read cautiously.

Contents

Chapter 1

Preliminaries

1.1 Baire's Category Theorem

Definition 1. Let X be a topological space. $F \subseteq X$ is said to be *dense* if $\overline{F} = X$. Alternatively, for every non-empty open set $V \subseteq X$, $V \cap F \neq \varnothing$.

$E \subseteq X$ is said to be *nowhere dense*[1] if $X \setminus E$ is dense. Alternatively, E is nowhere dense if $\overline{E}$ does not contain any non-empty open set.

Lemma 2. *Let X be a complete metric space, and $U_1, U_n, \ldots$ are dense open sets in X. Then $\bigcap_{i=1}^{\infty} U_i$ is dense in X.*

Theorem 3 (Baire Category Theorem)**.** *Let X be a complete metric space and $\varnothing \neq Y \subseteq X$ open. Then Y is not the countable union of nowhere dense sets.*

Proof. Assume otherwise, E_n is the sequence of nowhere dense sets. Define for $n \in \mathbb{N}$, $U_n = X \setminus \overline{E_n}$, then U_n is dense open. Using the lemma we have that $\bigcap U_n \cap Y \neq \varnothing$, contradiction. $\square$

Definition 4. We say that a topological space is of *first category*[2] if it is the countable union of nowhere dense sets. Otherwise the space is called *second category*.

Theorem 5 (Baire Category Theorem (II))**.** *If X is a complete metric space, and $Y \subseteq X$ is a non-empty open set, then Y is a second category space.*

1.2 Uniform Boundedness Principle

Theorem 6. *Let X be a complete metric space, and $\mathcal{F} \subseteq C(X, \mathbb{C})$. Assume that for every $x \in X$, $M_x = \sup\{|f(x)| \mid f \in \mathcal{F}\} < \infty$. Then there exists an open set $\varnothing \neq U \subseteq X$ for which $\sup\{M_x \mid x \in U\} < \infty$.*

Proof. Let $m > 0$ some natural number. Define $E_m(f) = \{x \in X \mid |f(x)| \leq m\}$ for $f \in \mathcal{F}$. Note that $E_m(f)$ is closed for each $f \in \mathcal{F}$. Finally, define $E_m = \bigcap\{E_m(f) \mid f \in \mathcal{F}\}$. The assumption gives us that $\bigcup_{m=1}^{\infty} E_m = X$, and therefore by BCT there is some m for which E_m contains a nontrivial open set, as wanted. $\square$

[1] Also *sparse* in some texts.

[2] Also *meager* in some texts.

1.3 Hahn-Banach Theorems

Definition 7. Let X be a real vector space. $p\colon X \to \mathbb{R}$ a function such that:

1. $p(x+y) \leq p(x) + p(y)$, and
2. $p(\alpha x) = \alpha p(x)$ for all $\alpha \geq 0$.

Then p is called a positive homogeneous function.

Theorem 8. *Suppose that X is a real vector space and $M \subseteq X$ is a [linear] subspace. If p is positive homogeneous function on X and $f_0\colon M \to \mathbb{R}$ a linear functional, such that $f_0(x) \leq p(x)$. Then there exists a linear functional $f\colon X \to \mathbb{R}$ such that $f \upharpoonright M = f_0$ such that $f(x) \leq p(x)$ for all $x \in X$.*

Proof. Let $y \notin M$ and look at $M_y = \operatorname{span}(M, y) = \{x + \alpha y \mid x \in M, \alpha \in \mathbb{R}\}$. We want to extend f_0 to M_y. So if f is such extension, $f(x+\alpha y) = f_0(x) + \alpha f(y) = f_0(x) + \alpha c$. We need to find c such that the following is satisfies:

$$f_0(x) + \alpha c \leq p(x + \alpha y)$$

If $\alpha = 0$ then we have nothing to check.

If $\alpha > 0$, then we get that $f_0(x)+\alpha c = f_0(\frac{1}{\alpha}x)+c \leq p(\frac{1}{\alpha}x+y)$, so $c \leq p(\frac{1}{\alpha}x+y) - f_0(\frac{1}{\alpha}x)$. If $\alpha < 0$ by similar arguments we get that $c \geq f_0(-\frac{1}{\alpha}x) - p(-(\frac{1}{\alpha}x + y))$.

For $x_1, x_2 \in M$ we have that:

$$f_0(x_1) + f_0(x_2) \leq p(x_1 + x_2) = p(x_1 + y + x_2 - y) \leq p(x_1 + y) + p(x_2 - y)$$

Therefore it is always the case that $p(x_1 + y) - f_0(x_1) \geq f_0(x_2) - p(x_2 - y)$. So there exists some $c \in \mathbb{R}$ such that for every $x_1, x_2 \in M$,

$$\inf_{x_1 \in M} [p(x_1 + y - f_0(x_1))] \geq c \geq \sup_{x_2 \in M} [f_0(x_2) - p(x_2 - y)]$$

Taking such c will satisfy the two requirements for extending f_0 to M_y. Consider the collection of all extensions of f_0, ordered by inclusion, Zorn's lemma gives us a maximal element which has to be an extension of f_0 to the entire space X. $\square$

Definition 9. Let X be a complex vector space. $p\colon X \to \mathbb{R}$ is called a *semi-norm* if:

1. $p(x+y) \leq p(x) + p(y)$, and
2. $p(\lambda x) = |\lambda| p(x)$ for all $\lambda \in \mathbb{C}$.

Proposition 10. *If p is a semi-norm, then $p(0) = 0$, $p(x) \geq 0$, and $p(x - y) \geq |p(x) - p(y)|$.*

Theorem 11 (Hahn-Banach[3]). *Let X be a complex vector space and p a semi-norm on X. If M is a subspace and $f_0\colon M \to \mathbb{C}$ a linear functional such that $|f_0(x)| \leq p(x)$, then there exists a linear extension $f\colon X \to \mathbb{C}$ which is bounded by p.*

[3]This theorem is actually due to Bohnenblust and Sobczyk

Proof. We write $f_0(x) = g_0(x) + ih_0(x)$ where g_0, h_0 are real functionals. We have that $f_0(ix) = if_0(x) = i(g_0(x) + ih_0(x)) = ig_0(x) - h_0(x)$. And therefore $g_0(x) = h_0(ix)$ and $g_0(ix) = -h_0(x)$. So we can always write $f_0(x) = g_0(x) - ig_0(ix)$.

We know that for every $x \in M$, $|g_0(x)|, |h_0(x)| \leq f_0(x) \leq p(x)$. In particular $|g_0(x)| \leq p(x)$ so we can extend it as a real linear functional to some g such that $g(x) \leq p(x)$ for all $x \in X$. Define $f(x) = g(x) - ig(ix)$, then f is a linear extension of f_0 as a complex linear functional. To see that, we only need to check $f(ix) = if(x)$.

Finally if $f(x) = re^{i\theta}$, then $|f(x)| = r = f(e^{i\theta}x) = g(e^{i\theta}x)$ (the last equality is because g is the real part of f and $r \in \mathbb{R}$). But we know that $|g(e^{i\theta}x)| \leq p(e^{i\theta}x) = p(x)$. Therefore $|f(x)| = r \leq p(x)$ as wanted. $\square$

Chapter 2

Locally Convex Spaces

2.1 Motivating example

Let $C[a,b]$ be the space of all continuous functions from $[a,b] \subseteq \mathbb{R}$ to $\mathbb{C}$. This is a complete metric space with $d(f,g) = \max_{x\in[a,b]} |f(x) - g(x)|$ (this is the uniform convergence topology). Let $F_n \subseteq C[a,b]$ be the following set,

$$\{f \mid \exists x_0 \in [a, b - \tfrac{1}{n}] : |f(x) - f(x_0)| \leq n(x - x_0) \forall x \in [x_0, b]\}$$

We claim that F_n is a closed set. If $\{f_k\}_{n=1}^{\infty} \subseteq F_n$ is a convergent sequence with limit g. For each f_k there is some x_0^k, and by the compactness of $[a,b]$ we can assume without loss of generality that $x_0^k \to x_0$ for some $x_0 \in [a, b - \frac{1}{n}]$. We know that $|f_k(x) - f_k(x_0^k)| \leq n(x - x_0^k)$ and by the uniform convergence to g it follows that $|g(x) - g(x_0)| \leq n(x - x_0)$. Therefore $g \in F_n$ as well.

Next we claim that F_n is nowhere dense. Given $f \in F_n$, we will approximate it by piecewise linear functions (this can be done using the fact f is uniformly continuous), and each piecewise linear function will be approximated by saw-graph functions, that in every point has a slope of more than n. This will show that if $f \in F_n$ then it can be approximated by functions not in F_n.

It follows that $\bigcup F_n$ is a first category set, whereas $C[a,b]$ is a complete metric space so it is of the second category. Therefore most functions in $C[a,b]$ are not in any of the F_n's. Such function h satisfies that for every $x \in [a,b]$, and for every $y \in [x,b]$, $|h(x) - h(y)| > n(y - x)$. So h is nowhere differentiable.

Remark. *From now on, unless stated otherwise, X is a complex vector space.*

2.2 Convex, Balanced and Absorbing sets

Definition 12. Let $M \subseteq X$.

1. We say that M is *convex* if whenever $x, y \in M$ and $t \in [0,1]$, then $tx + (1-t)y \in M$.

2. We say that M is *balanced* if whenever $x \in M$ and $\lambda \in \mathbb{C}$ with $|\lambda| \leq 1$, $\lambda x \in M$.

3. We say that M is *absorbing* if whenever $y \in X$ there is some $\mu > 0$ for which $\mu^{-1} y \in M$.

Let p be a semi-norm on X, and let $M = \{x \mid p(x) \leq 1\}$.

Proposition 13. *M is convex, balanced, and absorbing.* $\square$

Proposition 14. *For every $y \in X$, $p(y) = \inf\{\mu > 0 \mid \mu^{-1}y \in M\}$.* $\square$

Proposition 15. *Suppose now that $M \subseteq X$ is convex, balanced and absorbing. Define $p_M(y) = \inf\{\mu > 0 \mid \mu^{-1}y \in N\}$. Then p_M is a semi-norm.*

Proof. To see linearity use absorption and balanced to show it works.

To see the triangle inequality, given $x, y \in X$ fix $\varepsilon > 0$ and take $x' = \frac{1}{p_M(x)+\varepsilon}x$ and $y' = \frac{1}{p_M(y)+\varepsilon}y$, then $t = \frac{p_M(x)+\varepsilon}{p_M(x)+p_M(y)+2\varepsilon}$ and use convexity, then $\varepsilon \to 0$ gives the wanted result. $\square$

Definition 16. Given M convex, balanced and absorbing, p_M as defined before is called the *Minkowski functional of M*.

Theorem 17. *If p is a semi-norm, then it is the Minkowski functional for $\{x \mid p(x) \leq 1\}$.* $\square$

2.3 Topology induced from semi-norms

Definition 18. Let $\{p_\gamma\}_{\gamma\in\Gamma}$ a family of semi-norms on X. We say that this family is *separating* if for every $x_0 \neq 0$ there is some $\gamma \in \Gamma$ such that $p(x_0) > 0$.

Definition 19. Given a separating family of semi-norms, we define a topology on X as follows:

1. For 0, define a basic neighborhood of 0 as $U^{\varepsilon_1,\dots,\varepsilon_n}_{\gamma_1,\dots,\gamma_N} = \{x \in X \mid p_{\gamma_i}(x) < \varepsilon_i\}$. Now $B_0 = \{U^{\varepsilon_1,\dots,\varepsilon_N}_{\gamma_1,\dots,\gamma_N} \mid N \geq 0, \gamma_i \in \Gamma, \varepsilon_i > 0\}$.
2. For every other point, we consider the basis of neighborhood obtains by shifting each open set in B_0 by y.

Proposition 20. *If $U^{\varepsilon_1,\dots,\varepsilon_n}_{\gamma_1,\dots,\gamma_N} \in B_0$, then it is the intersection of $U^{\varepsilon_1}_{\gamma_1} \cap \cdots \cap U^{\varepsilon_N}_{\gamma_N}$.* $\square$

Proposition 21. *If the family is separating, the topology is Hausdorff.*

Proof. Take $x \neq y \in X$. And without loss of generality $x = 0, y \neq 0$. There is some γ for which $p_\gamma(y) = \varepsilon > 0$. Now $U^{\varepsilon/4}_\gamma$ and $y + U^{\varepsilon/4}_\gamma$ are open and disjoint as wanted. $\square$

Remark. *Under this topology every p_γ is a continuous function.*

Proposition 22. *X is a topological vector space. Namely the addition and scalar multiplication are both continuous.* $\square$

Definition 23. Let X be a complex vector space which has a topology satisfying:

1. 0 has a basis of neighborhoods which are convex, balanced and absorbing.
2. X is a topological vector space.

Then we say that X is a *locally convex* vector space.

Corollary 24. *The topology defined on X from the family of semi-norms is locally convex.*

Proposition 25. *Suppose that X is topologized by a family of semi-norms $\{p_\gamma\}_{\gamma\in\Gamma}$. Let p be a semi-norm on X, then p is continuous if and only if there are finitely many $\gamma_1,\dots,\gamma_n \in \Gamma$ and $c \geq 0$, such that for all $x \in X$, $p(x) \leq c \cdot \bar{p}(x) = c \cdot \max\{p_{\gamma_i}(x) \mid i < n+1\}$.*

Proof. We begin by proving the following claim. p is continuous if and only if it is continuous at 0. One direction is trivial. For the other direction, suppose that p is continuous at 0. So for every $\varepsilon > 0$ there exists some open neighborhood $U = U^{\eta,\dots,\eta}_{\delta_1,\dots,\delta_k}$ such that $p(U) \subseteq [0,\varepsilon)$. For any given $x_0 \in X$, consider $x_0 + U$ as an open neighborhood of x_0, then for every $x \in x_0 + U$ we have that $p(x) \leq |p(x) - p(x_0)| < \varepsilon$ as wanted.

Suppose that $p(x) \leq c \cdot \bar{p}(x)$ for all x. Given some $\varepsilon > 0$ and $x \in U^{\varepsilon/c,\dots,\varepsilon/c}_{\gamma_1,\dots,\gamma_n}$, then

$$p(x) \leq c \cdot \bar{p}(x) = c \cdot \max_{i<n+1} p_{\gamma_i}(x) < c \cdot \frac{\varepsilon}{c} = \varepsilon.$$

In the other direction if p is continuous at 0, then for some neighborhood $U^{\eta,\dots,\eta}_{\delta_1,\dots,\delta_k}$ in which $p(x) < 1$. Given $y \in X$, take $x = \frac{\eta}{\max_{i<k+1} p_{\delta_i}(y)} y$. Then $x \in U^{\eta,\dots,\eta}_{\delta_1,\dots,\delta_k}$ and therefore $p(x) < 1$, so $p(y) \leq \frac{1}{\eta} \max_{i<k+1} p_{\gamma_i}(y)$ as wanted.[1] □

Remark. *Let $\varphi(x,y)$ be a non-negative function on $X \times X$ which satisfies $\varphi(x,y) = \varphi(y,x)$ and $\varphi(x,z) \leq \varphi(x,z) + \varphi(y,z)$. Then $\psi(x,y)$ defined as $\frac{\varphi(x,y)}{1+\varphi(x,y)}$ satisfies the triangle inequality.*

Proposition 26. *Suppose that $\Gamma = \{\gamma_n \mid n \in \mathbb{N}\}$ is a countable family of semi-norms. The topology it induces on X is metrizable by the metric:*

$$d(x,y) = \sum_{k=1}^{\infty} 2^{-k} \frac{p_{\gamma_k}(y-x)}{1 + p_{\gamma_k}(y-x)}$$

Theorem 27. *Suppose that $\{p_n \mid n \in \mathbb{N}\}$ is a family of semi-norms. Then the topology it induces on X is equivalent to the topology induced by the metric d they induce.*

Proof. Let $\varepsilon > 0$, and consider the open ball $B(0,\varepsilon)$. Take $N \gg 1$ such that $2^{-N} < \frac{\varepsilon}{2}$. Look at the neighborhood $U = U^{\varepsilon/2,\dots,\varepsilon/2}_{1,\dots,N}$. For $x \in U$ we calculate $d(x,0)$:

$$d(x,0) = \sum_{n=1}^{\infty} 2^{-n} \frac{p_n(x)}{1+p_n(x)} = \sum_{n=1}^{N} + \sum_{n=N+1}^{\infty} < \frac{\varepsilon}{2} + \sum_{n=1}^{N} 2^{-n} \frac{\varepsilon}{2} < \varepsilon.$$

Therefore $U \subseteq B(0,\varepsilon)$. In the other direction we need to show that given $U^{\eta,\dots,\eta}_{1,\dots,N}$ contains some $B(0,\delta)$ for some $\delta > 0$. Namely, we need to show that if $\sum_{n=1}^{\infty} 2^{-n} \frac{p_n(x)}{1+p_n(x)} < \delta$, then $\max\{p_i(x) \mid i \leq N\} < \eta$. Choose $\delta = 2^{-N} \frac{\eta}{1+\eta}$, and juggle semi-norms around. □

Remark. *d is continuous on $X \times X$.*

Definition 28. Suppose that X is locally convex by a countable family of semi-norms and d the induced metric. X is a *Fréchet space* if it is complete under this metric. Additionally, if the countable family is a singleton (so we have a normed space), we say that X is a *Banach space*.[2]

[1] Note that we divide by the maximum over several $p_{\delta_i}(y)$. What happens when all of those are 0?

[2] We will often write F-space and B-space to denote these, although this is not entirely standard in the literature.

Chapter 3

Examples For Fréchet Spaces

3.1 Some topology in $\mathbb{R}^n$

We say that $\Omega \subseteq \mathbb{R}^n$ is a *domain* if it is open and connected.

Proposition 29. *Suppose that Ω is a domain, then there is a sequence of compact sets K_n such that $K_n \subseteq \operatorname{int}(K_{n+1})$ and $\Omega = \bigcup K_n$ and for every $K \Subset \Omega$ there is some m for which $K \subseteq K_m$.*

Proof. Consider the collection of closed balls $C_1, C_2, \ldots \Subset \Omega$ with center and radius being rational. Define $K_1 = C_1$. If K_n was defined, let K_{n+1} be a compact set whose interior includes $K_n \cup C_{n+1}$. There exists such K_{n+1} by normality of $\mathbb{R}^n$ (separate K_n from $\mathbb{R}^n \setminus \Omega$; if you cannot then $\Omega = \mathbb{R}^n$ in which case define $K_n = B(0, n)$).

It is not hard to see why K_n is as wanted. But moreover we have that if $K \Subset \Omega$, there exists some m for which $K \subseteq \bigcup_{i=1}^m C_i$, and therefore $K \subseteq K_m$. $\square$

Such a sequence of compact sets is called an *exhausting sequence* of Ω.

Consider $C(\Omega)$, the continuous complex functions on Ω. Let $K \Subset \Omega$, let $p_K(f) = \max\{|f(x) \mid x \in K\}$. Then p_K is a semi-norm, and $\{p_K\}_{K \Subset \Omega}$ is a separating family. Therefore this family induces a locally convex topology on $C(\Omega)$, and when we consider $C(\Omega)$ as a topological space, we will always consider it with this topology.

Proposition 30. *The topology on $C(\Omega)$ is actually the induced only by $\{p_{K_m}\}_{m \in \mathbb{N}}$ where K_m is an exhausting sequence. In particular the topology on $C(\Omega)$ is metrizable.*

Proof. It suffices to show that given any compact set $K \Subset \Omega$, the semi-norm p_K is continuous with respect to the topology induced by the exhausting sequence. However, if $K \Subset \Omega$, there is some m for which $K \Subset K_m$ and therefore $p_K(f) \leq p_{K_m}(f)$ for all $f \in C(\Omega)$, which is the criterion for p_K to be continuous, as wanted. $\square$

Proposition 31. $C(\Omega)$ *is completely metrizable.*

Proof. Fix an exhausting sequence, and look at the metric it defines:

$$d(f, g) = \sum_{j=1}^{\infty} 2^{-j} \frac{p_{K_j}(f - g)}{1 + p_{K_j}(f - g)},$$

and suppose that f_m is a Cauchy sequence in the metric. So for every $\varepsilon > 0$ there is some N so whenever $r, s > N$, $d(f_r, f_s) < \varepsilon$. In particular for every j, the restrictions $f_m \upharpoonright K_j$ give us Cauchy sequences for each K_j. Therefore for each j, there is some g_j which is the (uniform convergence) limit of $f_m \upharpoonright K_j$. Because $K_j \subseteq K_{j+1}$ we have that $g_j = g_{j+1} \upharpoonright K_j$ for all j.

Therefore $g = \bigcup g_j$ is defined on Ω, and by uniform convergence on each K_n, ensures that g is continuous on Ω. It remains to show that g is the limit of f_m, which is the usual ε juggling arguments. □

3.2 Another Example

Take $X = C[0,1]$ and define two semi-norms:

1. $p_1(f) = \max_{0 \leq x \leq 1}\{|f(x)|\}$.
2. $p_2(f) = \int_0^1 |f(x)|dx$.

However $p_2(f) \leq p_1(f)$, and so p_2 is continuous with respect to p_1. The topology defined from p_1 is in fact completely metrizable, and so $C[0,1]$ is a B-space with this norm.

Consider now the subspace $P = \{f \in C[0,1] \mid f \text{ a polynomial}\}$, then P is a linear subspace of X and therefore it is a normed space. However by Weierstrass' theorem every function in X can be approximated by polynomials and therefore P is not closed, so it is not complete under this norm.

Compare the two topologies of $C[0,1]$, the $\max$-norm and the topology induced from an exhausting sequence. The former is a B-space, but the latter is a F-space and not a Banach space.

Now consider the topology on $C(0,1)$, defined by the pointwise convergence of sequences, this is the usual product topology $\mathbb{C}^{(0,1)}$. This topology is given by the family of semi-norms, $p_x(f) = |f(x)|$ for every $x \in (0,1)$. Therefore the topology is indeed locally convex. But is it metrizable, and does it give rise to an F-space?

3.3 Differentiable Functions

Let $\Omega \subseteq \mathbb{R}^n$ be a domain. We write $C^k(\Omega)$ for the space of all continuous functions which can be differentiated at least k times. $C(\Omega) = C^0(\Omega)$. Given $\alpha = \langle \alpha_1, \ldots, \alpha_n \rangle \in (\mathbb{N} \cup \{0\})^n$ we define,

$$\partial^\alpha f = \left(\frac{\partial}{\partial x_1}\right)^{\alpha_1} \cdots \left(\frac{\partial}{\partial x_n}\right)^{\alpha_n} f.$$

We define $|\alpha| = \sum_{i=1}^n \alpha_i$, and for $k \in \mathbb{N}$ we define $C^k(\Omega) = \{f \mid \partial^\alpha f \in C^0(\Omega), \forall |\alpha| \leq k\}$.

Proposition 32. *Let $p_{K_j,\ell}(f) = \max_{K_j,|\alpha|\ell} |\partial^\alpha f|$, where K_j is an exhausting set for Ω. For $\ell \leq k$, $C^k(\Omega)$ is an F-space in the topology induced by these semi-norms.*

This is the space of continuously differentiable functions of order k, in the uniform convergence topology on every derivative on every compact subset of Ω.

Proposition 33. *Suppose that* $\{f_m\}_{m=1}^{\infty} \subseteq C^k(\Omega)$ *a sequence of functions such that for every* α *such that* $|\alpha| \leq k$ *the following sequences converge uniformly on every compact set:*

$$\lim_{n\to\infty} f_m = g \in C(\Omega), \quad \lim_{n\to\infty} \partial^\alpha f_m = g_\alpha \in C(\Omega).$$

Then g *is continuously differentiable up to order* k *and* $g_\alpha = \partial^\alpha g$.

Chapter 4

Locally Convex Spaces (II)

4.1 Linear Operators on Locally Convex Spaces

Let X be a locally convex complex linear space.

Definition 34. Let $f: X \to \mathbb{C}$ be a function. We say that f is a linear functional if $f(x+y) = f(x) + f(y)$ and for all $\lambda \in \mathbb{C}$, $f(\lambda x) = \lambda f(x)$.

Proposition 35. *Let f be a linear functional, then f is continuous on X if and only if it is continuous at* 0. $\square$

Proposition 36. *Let f be a linear functional on X. Then f is continuous if and only if there exists a continuous semi-norm p and $c > 0$ such that $|f(x)| \leq c \cdot p(x)$ for all $x \in X$.*

Proof. If $|f(x)| \leq c \cdot p(x)$, then it is not hard to verify that f is continuous at 0. In the other direction, if f is continuous, then there is some U open neighborhood of 0, such that $f(U) \subseteq B(0,1)$ (where $B(0,1)$ is computed in $\mathbb{C}$). However there is some continuous semi-norm p and $\varepsilon > 0$ such that $\{x \mid p(x) \leq \varepsilon\} \subseteq U$. Take $y \neq 0$ so $\frac{\varepsilon \cdot y}{p(y)} \in U$ and therefore $|f(\frac{\varepsilon \cdot y}{p(y)})| \leq 1$ and therefore $|f(y)| \leq \frac{1}{\varepsilon} p(y)$. $\square$

Definition 37. Let X, Y be locally convex spaces. $T: X \to Y$ is called a *linear operator* if it is linear.

Proposition 38. *A linear operator $T: X \to Y$ is continuous if and only if T is continuous at* 0. $\square$

Proposition 39. *$T: X \to Y$ is continuous if and only if for every continuous semi-norm q on Y, there exists a continuous semi-norm p on X, and $c > 0$ such that $q(Tx) \leq c \cdot p(x)$.* $\square$

So for example, if X and Y are normed spaces, with $\|\cdot\|_X$ and $\|\cdot\|_Y$ the respective norms. Then $T: X \to Y$ is continuous if and only if $\|Tx\|_Y \leq c\|x\|_X$.

Proposition 40. *Let X be an F-space. If $A, B \subseteq X$ then $\overline{A} + \overline{B} \subseteq \overline{A + B}$. And if $W \subseteq X$ is open, then $W - W$ is an open environment of* 0.

Proof. Suppose that $x_n \to x$ where $x_n \in A$ and $x \in \overline{A}$ and $y_n \to y$ where $y_n \in B$ and $y \in \overline{B}$, then $x_n + y_n \to x + y$ since addition is continuous. Therefore $x + y \in \overline{A + B}$.

For the second part, note that $W - W = \bigcup_{y \in W} y - W$. Since moving an open set is open, $y - W$ is open for each y, so $W - W$ is open, and it is easy to see that $0 \in W - W$. $\square$

Proposition 41. *Let $T\colon X \to Y$ be a surjective linear operator. Then for every $0 \in U \subseteq X$ open, $Y = \bigcup_{k=1}^{\infty} kT(U)$.*

Proof. Since U is open, it contains a ball around 0, therefore every $x \in X$ has some $k \in \mathbb{N}$ such that $x \in kU$. By linearity it follows that $T(x) \in kT(U)$, and by surjectivity of T the claim follows. $\square$

Proposition 42. *Let $T\colon X \to Y$ be a surjective linear operator. If $0 \in U \subseteq X$ open, then $\overline{T(U)}$ contains an open neighborhood of 0 in Y.*

Proof. Since $Y = \bigcup_{k=1}^{\infty} k\overline{T(U)}$, it follows from Baire's Category Theorem that there is some k_0 such that $k_0\overline{T(U)}$ which contains an open set. However, scalar multiplication is a homeomorphism, so $\overline{T(U)}$ contains an open environment Y.

But now there is some $W \subseteq X$ which is an open neighborhood of 0 and $W - W \subseteq U$ (for example, a small enough ball around 0). Therefore $\overline{T(W)} - \overline{T(W)} \subseteq \overline{T(U)}$. But by the above argument $\overline{T(W)}$ contains an open set, and therefore $\overline{T(W - W)} \subseteq \overline{T(U)}$ and so we have that $\overline{T(U)}$ contains an open neighborhood of 0 as wanted. $\square$

4.2 Open Mapping and Closed Graph Theorems

Theorem 43 (Open Mapping Theorem)**.** *Let X, Y be F-spaces and $T\colon X \to Y$ a continuous linear operator between them. If $T(X) = Y$ (namely, T is surjective), then T is an open map. In other words, whenever $U \subseteq X$ is open, $T(U)$ is open in Y.*

Proof. Let $U \subseteq X$ be an open neighborhood of 0. Pick some $\varepsilon > 0$ such that $B_X(0, \varepsilon) \subseteq U$. Define $B_{X,i} = B(0, 2^{-i}\varepsilon)$ for $i \in \mathbb{N}$. So for each i there is some η_i such that $B_Y(0, \eta_i) \subseteq \overline{B_{X,i}}$. Without loss of generality $\eta_i \to 0$ monotonically as well.

Pick $y \in \overline{T(B_{X,1})}$. Then there is some $x_1 \in B_{X,1}$ such that $d_Y(y - Tx_1, 0) < \eta_2$, so $y - Tx_1 \in \overline{T(B_{X,2})}$ by the choice of η_i's. Continue by induction finding $x_n \in B_{X,n}$ such that $d_Y(y - T(\sum_{i \leq n} x_i, 0) < \eta_{n+1}$. Denote by $z_n = \sum_{i<n} x_i$, then z_n is a Cauchy sequence in X. So there is some $x = \lim_{n\to\infty} z_n$.

But now we have that $d_X(x, 0) \leq \sum_{i=1}^{\infty} d_X(x_i, 0) < \varepsilon$. So $x \in B_0 \subseteq U$.

On the other hand T is continuous, and $d_Y(y - T(z_n), 0) < \eta_{n+1}$, so $Tx = y$. And therefore $T(U)$ contains an open neighborhood of 0 since $\overline{T(B_{X,1})} \subseteq T(U)$.

And now if $V \subseteq X$ is any open set, and $x_0 \in V$, then there is some $\delta > 0$ such that $x_0 + B_X(0, \delta) \subseteq V$. But now using continuity of addition and the previous part, we have that $Tx_0 + T(B(0, \delta)) \subseteq T(V)$ and contains an open neighborhood of Tx_0 as wanted. $\square$

Corollary 44. *If $T\colon X \to Y$ is a continuous linear operator which is bijective, then it is a homeomorphism.* $\square$

Proposition 45. *Suppose that X is a Banach space under two norms, $\|\cdot\|_1$ and $\|\cdot\|_2$. And suppose that there is some $c > 0$ for which $\|x\|_2 \leq c\|x\|_1$ for all $x \in X$. Then the norms are equivalent.*

Proof. Simply take the identity map $x \mapsto x$ which is linear and bijective. By the assumption on $\|x\|_2 \leq c\|x\|_1$, this is a continuous operator as well. Therefore by the open mapping theorem, this is a homeomorphism, so the norms are equivalent. $\square$

Proposition 46. *If X and Y are F-spaces, then $X \times Y$ is an F-space as well.*

Proof. We define the product topology on $X \times Y$. If $\{p_n\}_{n\in\mathbb{N}}$ and $\{q_k\}_{k\in\mathbb{N}}$ are the family of semi-norms defining the topologies on X and Y respectively, then $t_{n,k}(x, y) = p_n(x) + q_k(y)$ is a countable collection of semi-norms which define the product topologies on $X \times Y$.

The space is completely metrizable since the product of two completely metrizable spaces is completely metrizable. □

Definition 47. Let $T\colon X \to Y$ be a linear operator. The *graph* of T is the subspace $G_T = \{(x, Tx) \mid x \in X\}$ of $X \times Y$.

Theorem 48 (Closed Graph Theorem)**.** *Suppose that $T\colon X \to Y$ is a linear operator between F-spaces, then T is continuous if and only if G_T is closed.*

Proof. Suppose that T is continuous, let $\{(x_n, Tx_n)\}_{n=1}^{\infty}$ a Cauchy sequence in $X \times Y$. Since X is an F-space, there is some $x \in X$ such that $\lim_{n\to\infty} x_n = x$, and by continuity $Tx_n \to Tx$ in T. Therefore $(x, Tx) \in G_T$, and therefore this is a closed set.

Suppose that G_T is closed, then by the virtue of being a closed subspace of an F-space, G_T is an F-space on its own. Consider the map $\Phi(x) = (x, Tx)$. Easily this is a bijection between X and G_T. Consider the projection map, $\Phi^{-1}(x, Tx) = x$. By the definition of the product topology we have that Φ^{-1} is continuous, and easily a linear operator, and surjective. By the Open Mapping Theorem we have that Φ is continuous and therefore T is continuous as the composition of Φ and a projection to Y. □

4.3 More On Linear Operators Between F-spaces

Here X and Y will always be two F-spaces.

Theorem 49 (Uniform Boundedness Theorem)**.** *Let $\{T_a\}_{a\in A}$ a family of continuous linear operators from X to Y. Moreover, assume that the following condition holds:*

$$\forall x \in X \ \sup_{a\in A} d_Y(T_a x, 0) < \infty,$$

then there exists an open ball $B_X(0, \varepsilon) \subseteq X$ for which $\sup_{a\in A} \sup_{x\in B_X(0,\varepsilon)} d_Y(T_a x, 0) < \infty$.

Proof. For every $a \in A$ define the function $f_a(x) = d_Y(T_a x, 0)$. Then $\{f_a\}_{a\in A}$ is a family of continuous functions from X to $\mathbb{C}$. Now the condition of the uniform boundedness principle holds, therefore there is an open set $U \subseteq X$ such that $\sup_{x\in U} \sup_{a\in A} f_a(x) < \infty$. Find some $B_X(x_0, \varepsilon) \subseteq U$, and now for $x \in B_X(x_0, \varepsilon)$ it holds that,

$$d_Y(T_a(x - x_0), 0) \le d_Y(T_a x, 0) + d_Y(T_a x_0, 0),$$

and therefore

$$\sup_{z\in B(0,\varepsilon)} \sup_{a\in A} d_Y(T_a z, 0) \le \sup_{a\in A}\{d_Y(T_a x_0, 0)\} + \sup_{x\in U} \sup_{a\in A}\{d_Y(T_a x, 0)\} < \infty$$

Therefore $B_X(0, \varepsilon)$ is as wanted. □

4.3.1 The Particular Case of Normed Spaces

Proposition 50. *Let $T\colon X \to Y$ a linear operator between two normed spaces. Then T is continuous if and only if there exists some $c > 0$ for which $\|Tx\|_Y \le c\|x\|_X$ for all $x \in X$.*

Proof. If $\|T_x\|_Y \le c\|x\|_X$, then it is clear that T is continuous (it is continuous at 0). In the other direction, if T is continuous, then we saw in general in the case of semi-norms that this is a condition guaranteeing continuity. □

Proposition 51. *If X and Y are normed spaces, and $T\colon X \to Y$ is a continuous linear operator, then:* [1]

$$\sup_{\|x\|_X \le 1} \|Tx\|_Y = \inf\{c > 0 \mid \|Tx\|_Y \le c\|x\|_X\}.$$

Proof. If $c > 0$ satisfies that $\|Tx\|_Y \le c\|x\|_X$, then in particular $\|Tx\|_Y \le c$ for all $\|x\|_X \le 1$. Therefore we have that $\sup \le \inf$.

In the other direction, if $c_1 < \sup$ then there is $\|x\|_X \le 1$ such that $\|Tx\|_Y > c_1, c_1\|x\|_X$. Therefore $c_1 < \inf$. Therefore $\sup \ge \inf$ and equality holds as wanted. □

We say that an operator is bounded if it is continuous, namely if the image of the unit ball is bounded.

Definition 52. If X and Y are normed spaces, $T\colon X \to Y$ is a continuous linear operator, we define the *operator norm* of T as

$$\|T\| = \sup_{\|x\|_X \le 1} \|Tx\|_Y.$$

If $Y = \mathbb{C}$, we denote by X^* the space of all continuous linear functionals from X to $\mathbb{C}$. This is called the *dual space of X*.

4.3.2 We Continue From the Digression... But Now Only For Normed Spaces

Theorem 53. *Suppose that $\{T_a\}_{a\in A}$ is a family of continuous linear operators from a B-space X to a normed space Y. Suppose that for all $x \in X$, $\sup_{a\in A} \|T_a x\|_Y < \infty$. Then $\sup_{a\in A} \|T_a\| < \infty$.*

We will write $B(X, Y)$ for the space of all bounded linear operators $T\colon X \to Y$. If $X = Y$ we just write $B(X)$.

Proposition 54. *If X and Y are normed spaces, then $B(X, Y)$ is a normed space using the operator norm as defined above.*

Proposition 55. *If Y is a Banach space, then $B(X, Y)$ is a Banach space.*

Proof. If $\{T_n\}$ is a Cauchy sequence in $B(X, Y)$, given $\varepsilon > 0$, $n, m > N(\varepsilon)$ satisfy:

$$\|(T_n - T_m)x\|_Y < \varepsilon\|x\|_X,\ \forall x \in X.$$

Use the completeness of Y to define $Tx = \lim_{n\to\infty} T_n x$, which is therefore well-defined. It follows from the definition that T is a bounded linear operator and that it is the limit of $\{T_n\}$. □

[1] In the following equality we can use $\|x\|_X = 1$. Why?

Corollary 56. *If X is a normed space, then $X^* = B(X, \mathbb{C})$ is a Banach space.* $\square$

Proposition 57. *Let X be a B-space and Y is a normed space, and let $\{T_n\}_{n=1}^{\infty} \subseteq B(X, Y)$ a sequence of bounded operators such that for every $x \in X$ the following limit exists in Y:*

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

Then the following things are true:

1. $T \in B(X, Y)$,
2. $\sup_{n\geq 1} \|T_n\| < \infty$,
3. $\|T\| = \liminf_{n\to\infty} \|T_n\|$.

Proof. First note that because the limit exists for every $x \in X$, it follows that for every $x \in X$, $\sup_n \|T_n x\| < \infty$, and therefore by the uniform boundedness theorem $\sup \|T_n\| < \infty$. Additionally, T is linear because $\lim$ is linear.

From the fact that norms are continuous on normed spaces, it follows that $\|Tx\|_Y = \lim_n \|T_n x\|_Y \leq \|T_n\|\|x\|_X$. Since x plays no particular role it follows that indeed $\|T\| = \liminf_n \|T_n\|$. $\square$

Remark. *If $T, S \in B(X)$ then $\|TS\| \leq \|T\|\|S\|$.*

Proposition 58. *Suppose that X is a B-space, and let $T \in B(X)$ such that $\|T\| < 1$. Then $I - T$ is an invertible and $(I-T)^{-1}$ is bounded and the following equality holds:*

$$(I-T)^{-1} = I + T + T^2 + \dots \tag{*}$$

Proof. We first need to show that the infinite series has meaning. Let $S_N = I + T + \dots + T^N$ for $N \geq 1$. Given any $M > N$ we have that:

$$\|S_M - S_N\| \leq \|T^{N+1}\| + \dots \|T^M\| \leq \|T\|^{N+1} + \dots \|T\|^M.$$

By the assumption that that $\|T\| < 1$ this is a Cauchy sequence and therefore it has a limit S. To see that $S = (I-S)^{-1}$ note that:

$$(I-T)S = \lim_{N\to\infty} (I-T)S_N = \lim_{N\to\infty} I - T^{N+1} = I.$$

This follows from the fact that $\|T^{N+1}\| \leq \|T\|^{N+1} \to 0$. Therefore $S = (I-T)^{-1}$. $\square$

The series $(I-T)^{-1} = I + T + T^2 + \dots$ is called the *Neumann series*.

Corollary 59. *The set of invertible operators in $B(X)$ is open (in the operator norm).*

Proof. Suppose that $T \in B(X)$ is invertible, then for $S \in B(X)$, $T - S = T(I - T^{-1}S)$. If $\|S\| < 1/\|T^{-1}\|$, namely S lies inside the open ball U of radius $1/\|T^{-1}\|$ around 0, then $\|T^{-1}S\| \leq \|T^{-1}\|\|S\| < 1$ and therefore $T(I - T^{-1}S)$ is invertible. $\square$

If we consider the algebraic structure of $B(X)$, then we have addition and scalar multiplication. But since $B(X)$ is closed under composition, it is in fact a $\mathbb{C}$-algebra. Moreover when we consider the norms:

$$\|TSx\|_X \leq \|T\|\|Sx\|_X \leq \|T\|\|S\|\|x\|_X.$$

Therefore $\|TS\| \leq \|T\|\|S\|$. This is a normed algebra.

Another example for a normed algebra is $C(K)$ where $K \Subset \mathbb{R}^n$, given by the usual norm $\|f\| = \sup\{|f(x)| \mid x \in K\}$ and the multiplication is given pointwise.

Chapter 5

Some Properties of Dual Spaces

Recall that if X is a normed space, its dual space $X^* = B(X, \mathbb{C})$ is a B-space. We will always assume X is a normed space, unless stated otherwise.

5.1 On the Existence of Functionals

Proposition 60. *Let $0 \neq x_0 \in X$. Then there is $f \in X^*$ such that $f(x_0) = \|x_0\|_X$ and $\|f\| = 1$.*

Proof. Consider the space $X_0 = \{\alpha x_0 \mid \alpha \in \mathbb{C}\}$. Now consider the linear functional defined on X_0, $f_0(\alpha x_0) = \alpha \|x_0\|_X$. It is easy to see that f_0 is a linear functional.

>From its definition, it follows that $|f_0(\alpha x_0)| \leq \|\alpha x_0\|_X$. Of course the norm is a semi-norm, so Hahn-Banach applies to extend f_0 to a linear functional on X which is bounded by the norm, and therefore continuous. Moreover for all $x \in X$, $|f(x)| \leq \|x\|_X$ and therefore $\|f\| \leq 1$.

On the other hand, $f(x_0) = f_0(x_0) = \|x_0\|_X$ and therefore taking $u = \frac{1}{\|x_0\|_X} x_0$ gives us that $f(u) = 1$ and so $\|f\| = 1$ as wanted. $\square$

Corollary 61.

$$\|x_0\|_X = \max_{\|f\| \leq 1,\ f \in X^*} |f(x_0)|$$

Proof. It is clear that if $\|f\| \leq 1$ then $|f(x_0)| \leq \|x_0\|_X$. So $\sup_{\|f\| \leq 1} |f(x_0)| \leq \|x_0\|_X$. By the previous proposition, there is some f giving us exactly $f(x_0) = \|x_0\|$ so $\sup$ is in fact $\max$. $\square$

Is the dual proposition true? Namely, $\|f\| = \max_{\|x\|_X \leq 1} |f(x)|$? If we replace $\max$ by $\sup$ we get the definition of $\|f\|$. So the question is whether or not we can realize this $\sup$ as an actual point on the closed unit ball. The answer is negative, as the following example shows.

Example 62. *Let $X = \{\varphi \in C[0,1] \mid \varphi(0) = \varphi(1) = 0\}$. It is not hard to show that X is a closed subspace of $C[0,1]$ and therefore a B-space. Consider $f(\varphi) = \int_0^1 \varphi(t)dt$, then $f \in X^*$, $\|f\| = 1$. But there is no $\varphi \in X$ such that $f(\varphi) = 1$.*

Proposition 63. *Let $M \subseteq X$ a closed subspace of X, and $y \in X \setminus M$ such that $d(y, M) = \delta$. Then there exists some $f \in X^*$ such that $f \restriction M = 0$, $f(y) = \delta$, and $\|f\| \leq 1$.*

Proof. Consider $M_y = \operatorname{span} M \cup \{y\} = \{m + \alpha \cdot y \mid m \in M, \alpha \in \mathbb{C}\}$. Define f_0 on M_y by $f_0(m + \alpha \cdot y) = \alpha\delta$. In order to apply the Hahn-Banach theorem, we need to check that $|f_0(m + \alpha \cdot y)| \leq \|m + \alpha \cdot y\|_X$. If $\alpha \neq 0$, then we have that:

$$\|m + \alpha \cdot y\|_X = |\alpha| \left\| \frac{m}{\alpha} + y \right\|_X \geq |\alpha|\delta = |\alpha\delta| = |f_0(m + \alpha \cdot y)|.$$

Use Hahn-Banach to obtain $f \in X^*$ defined on X such that $\|f\| \leq 1$, and it is easy to see that $f \restriction M = 0$ and $f(y) = \delta$ as wanted. □

Recall that a topological space X is *separable* when it has a countable dense subset. We make the following remark.

Remark. *Let X be a normed space, then X is separable if and only if $S = \{x \mid \|x\|_X = 1\}$ is separable.*

Proposition 64. *Let X be a B-space. If X^* is separable, then X is separable.*

Proof. Let $\{f_n\}_{n=1}^\infty$ a countable dense subset on the unit sphere of X^*. For every n there is $x_n \in X$ such that $|f_n(x_n)| \geq \frac{1}{2}$ and $\|x_n\|_X = 1$. We will show that $M = \operatorname{span}\{x_n \mid n \in \mathbb{N}\}$ is dense in X. Suppose not, then there is some $y \in X \setminus \overline{M}$, by the previous proposition we have some $f \in X^*$ such that $f \restriction \overline{M} = 0$, $\|f\| = 1$ and $f(y) = \beta \neq 0$.

We have now that:

$$\frac{1}{2} \leq |f_n(x_n)| \leq |(f_n - f)(x_n)| + |f(x_n)| = |(f_n - f)(x_n)| \implies \|f_n - f\| \geq \frac{1}{2}.$$

This is for every n, but this means that f is not an accumulation point of the $\{f_n\}_{n=1}^\infty$. □

Proposition 65. *In Theorem 63 we can require that $\|f\| = 1$.*

We first prove the following lemma.

Lemma 66. *Let X be a normed space, $M \subseteq X$ a closed subspace, then for every $1 > \varepsilon > 0$ there is some $x_\varepsilon \in X$ such that $d(x_\varepsilon, M) > 1 - \varepsilon$ and $\|x_\varepsilon\|_X = 1$.*

Proof. Let $u \in X \setminus M$ such that $d(u, M) = d > 0$, then there is some $v \in M$ such that $\|u - v\|_X < \frac{d}{1-\varepsilon}$. But now $d(u - v, M) = d$ as well, since $\inf\{\|(u - v) - m\|_X \mid m \in M\} = \inf\{\|u - (v + m)\|_X \mid m \in M\}$. Let $x_\varepsilon = \frac{u-v}{\|u-v\|_X}$. We have that $d(x_\varepsilon, M) = \frac{1}{\|u-v\|_X} \cdot d > \frac{1-\varepsilon}{d} \cdot d = 1 - \varepsilon$. □

Now we can prove the proposition.

Proof. Let M_y be as in the original proof, and $f_0(m + \alpha \cdot y) = \alpha\delta$, where $\delta = d(y, M)$. For every $\varepsilon > 0$ we can find by the lemma some $\alpha_\varepsilon \in \mathbb{C}$ such that $\|m + \alpha_\varepsilon \cdot y\|_X = 1$ and $|\alpha_\varepsilon|\delta = d(m + \alpha_\varepsilon \cdot y, M) > 1 - \varepsilon$.

Therefore the following holds:

$$f_0(m + \alpha_\varepsilon \cdot y) = \alpha_\varepsilon\delta \implies |f_0(m + \alpha_\varepsilon \cdot y)| = |\alpha_\varepsilon|\delta > 1 - \varepsilon.$$

And since $\|m + \alpha_\varepsilon \cdot y\|_X = 1$, it follows that $\|f_0\| \geq 1$ and therefore $\|f_0\| = 1$, and the extension to f also satisfies $\|f\| \leq 1$ and so $\|f\| = 1$ as wanted. □

Proposition 67 (Mazur's Theorem). *Let $K \subseteq X$ a closed set which is also convex, balanced and absorbing, and let $y \in X \setminus K$. Then there is some $f \in X^*$ such that $|f(x)| \leq 1$ for $x \in K$ and $f(y) > 1$.*

Proof. Since K is closed there is some open ball around y, $B(y, \varepsilon) \cap K = \varnothing$, but this means that $B(y, \frac{\varepsilon}{2}) \cap (K + B(0, \frac{\varepsilon}{2})) = \varnothing$ again.

However $V = K + B(0, \frac{\varepsilon}{2})$ is an open neighborhood of 0 which is convex, balanced and absorbing. Moreover $y \notin \overline{V}$, since the above shows there is an open neighborhood of y disjoint from V. Take p to be the Minkowski functional of $\overline{V}$. Then $x \in \overline{V} \iff p(x) \leq 1$, so $p(y) > 1$; and p is continuous as a semi-norm of such set.

Take now $H_y = \operatorname{span}\{y\} = \{\alpha \cdot y \mid \alpha \in \mathbb{C}\}$, and define on H_y the functional $f_0(\alpha \cdot y) = \alpha \cdot p(y)$. So we have that $|f_0(\alpha \cdot y)| = |\alpha| p(y) = p(\alpha \cdot y)$ and therefore $|f_0| \leq p$ for $x \in H_y$. Using Hahn-Banach there is a linear extension of f_0 to some f defined on all X such that $|f(x)| \leq p(x)$ for all $x \in X$.

Now f is continuous because it is bounded by a continuous semi-norm. Therefore $f \in X^*$ and it satisfies the wanted properties. □

Proposition 68. *Let X be a normed space, and $M \subseteq X$ a subspace. Given $f_0 \in M^*$, there is some $f \in X^*$ which extends it, and $\|f_0\| = \|f\|$.*

Proof. For every $m \in M$ we have $|f_0(m)| \leq \|f_0\| \cdot \|m\|_X$. Define a semi-norm on X, $p(x) = \|f_0\| \|x\|_X$. Then $|f_0(m)| \leq p(m)$ for all $m \in M$. By the simplest form of Hahn-Banach, there is some f extending f_0 and $|f(x)| \leq p(x)$ for all x.

Therefore $f \in X^*$, and $\|f\| \leq \|f_0\|$ as a result, so $\|f_0\| = \|f\|$. □

Definition 69. Suppose $M \subseteq X$ is a closed subspace. For $\varepsilon > 0$ there is some $y_\varepsilon \in X$ such that $\|y_\varepsilon\|_X = 1$ and $d(y_\varepsilon, M) > 1 - \varepsilon$. We say that y_ε is *almost orthogonal* to M.

Remark. *If X is a finite dimensional space, then the unit sphere S is compact, and therefore we can find for each $k \in \mathbb{N}$ some y_ε for $\varepsilon = \frac{1}{k}$, denote it by y_k. Then $\{y_k\}_{k=1}^{\infty}$ has a limit point on S, which is some y such that $\|y\|_X = 1$ and $d(y, M) = 1$ as well.*

Example 70. *Let P_k be the space of complex polynomials on $[0,1]$ with the $\max$-norm. Consider $X = P_{18}$ and $M = P_{17}$. Then there is some polynomial p of degree 18 such that $\max_{x\in[0,1]} |p(x)| = 1$, and for every $q \in M$ we have that $\max_{x\in[0,1]} |p(x) - q(x)| \geq 1$.*

5.2 Continuous Functionals on a Normed Space X

Suppose that X is a normed space, we know that X^* is a B-space. We denote by X^{**} the dual space of X^*. We define a linear operator Φ on X:

$$\Phi(x)(f) = f(x)$$

Namely $\Phi\colon X \to X^{**}$, because if $f \in X^*$ then evaluating f by a fixed x is a linear map from X^* to $\mathbb{C}$. We need that $\Phi(x)$ to be continuous on X^* in order to solidify our claim. But we have $|\Phi(x)(f)| = |f(x)| \leq \|f\| \cdot \|x\|_X$, but here $\|x\|_X$ is fixed, so we have the wanted continuity, and $\|\Phi(x)\| \leq \|x\|_X$.

On the other hand we also know that there is some $f_x \in X^*$ such that $\|f_x\| = 1$ and $f_x(x) = \|x\|_X$. Therefore $|\Phi(x)(f_x)| = \|x\|_X$, and therefore the inequality above is not strict. Namely, $\|\Phi(x)\| = \|x\|_X$.

In conclusion: The map $\Phi\colon X \to X^{**}$ is a linear isometry, namely $\|\Phi(x)\| = \|x\|_X$. In particular Φ is injective and its image is a subspace of X^{**} isometrically isomorphic to X.

Corollary 71. *Some corollaries from the above construction.*

1. *If* $\dim X = \infty$*, then* $\dim X^* = \infty$ *as well.*
2. *if* X *is a B-space, then* $\Phi(X)$ *is closed in* X^{**}*.*
3. *More generally, if* X *is a normed space, then* $\overline{\Phi(X)}$ *is a metric completion of* X*. In particular every normed spaces can be extended to a Banach space.*

Proof. If $\dim X^* = n$, then $\dim X^{**} = n$ as well. But $\Phi(X)$ has infinite dimensional, which is a contradiction.

If X is a B-space, taking a convergent sequence $y_n \to y$ in $\Phi(X)$, let $x_n = \Phi^{-1}(y_n)$. Then x_n is a convergent sequence in X, because X is a B-space. Let x be the limit in X of x_n, the by continuity of Φ, $\Phi(x) = y$ as wanted. $\square$

Definition 72. We say that a normed space X is *reflexive* if Φ is onto X^{**}.[1]

Proposition 73. *Suppose that* X *is a reflexive space, then* X *is separable if and only if* X^* *is separable.*

Proof. If X^* is separable, then X is separable as we proved before. If X is separable, then X^{**} is separable, and therefore X^* is separable. $\square$

5.2.1 Families of Functionals and Bilinear Forms

Recall the uniform boundedness theorem. If X is an F-space and $\{f_a\}_{a\in A} \subseteq X^*$, then if for every $x \in X$, $\sup_{a\in A} |f_a(x)| < \infty$, then there is some open ball U around 0 such that $\sup_{a\in A} \sup_{x\in U} |f_a(x)| < \infty$. In particular if X is a B-space it follows that $\sup_{a\in A} \|f_a\| < \infty$.

Corollary 74. *If* $\{f_n\}_{n=1}^{\infty} \subseteq X^*$ *and* X *is a B-space, and suppose* $f_n(x) \to f(x)$ *for all* $x \in X$*. Then:*

1. $f \in X^*$.
2. $\|f\| \leq \liminf_{n\to\infty} \|f_n\|$.

Definition 75. $Q\colon X \times Y \to \mathbb{C}$ is called a *bilinear form* if for every $x \in X$ and $y \in Y$, the functional $Q(x, \cdot)$ and $Q(\cdot, y)$ are linear functionals from X and Y respectively.

Theorem 76. *Let* X *and* Y *be B-spaces, and* Q *a bilinear form such that* $Q(x, \cdot) \in X^*$ *and* $Q(\cdot, y) \in Y^*$ *for all* $x \in X$ *and* $y \in Y$*. Then* Q *is continuous on* $X \times Y$*.*

Proof. It suffices to prove continuity at $(0, 0)$. Let $U = B_X(0, 1)$ the open unit ball in X. Consider the family of functionals on Y, $\{f_x\}_{x\in U}$ defined as $f_x(y) = Q(x, y)$. By the assumption each $f_x \in Y^*$.

[1]Note that a reflexive space is necessarily Banach.

For a fixed y, $\sup_{x\in U}|f_x(y)| < \infty$ by the continuity of $Q(\cdot, y)$. By the uniform boundedness theorem there is some $V = B_Y(0,r)$ such that $\sup_{x\in U}\sup_{y\in V}|f_x(y)| = M < \infty$. In other words, $\sup_{x\in U}\sup_{y\in V}|Q(x,y)| = M < \infty$. In particular for every $\varepsilon > 0$,

$$\sup_{x\in\varepsilon\cdot U}\sup_{y\in\varepsilon\cdot V}|Q(x,y)| = M\varepsilon^2.$$

Therefore in $\varepsilon\cdot U\times\varepsilon\cdot V \subseteq X\times Y$, for all (x,y) in that neighborhood $|Q(x,y)| \le M\varepsilon^2$. $\square$

Corollary 77. *In particular, if X and Y are B-spaces and Q is a bilinear form on $X\times Y$, then there is some $c > 0$ such that:*

$$|Q(x,y)| \le c\|x\|_X\|y\|_Y.$$

If we take U and V as the open unit balls of X and Y as in the previous proofs, $c = M$.

Example 78. *Let $X = C(0,1)$ be the complex-valued continuous functions on $(0,1)$. Let $\{x_k\}_{k=1}^{\infty} \subseteq (0,1)$ a sequence of points such that for all $f \in C(0,1)$, $\sup_k|f(x_k)| < \infty$. Then there is some compact set $K \Subset (0,1)$ such that $x_k \in K$ for all $k\in\mathbb{N}$.*

For each k, consider the evaluation functional, $\phi_x(f) = f(x_k)$. It is a continuous linear functional on X. By the assumption, $\sup_k|\phi_k(f)| < \infty$.

If X is a normed space, then $f \in X^*$ if and only if $|f(x)| \le \|f\|\cdot\|x\|_X$. In particular f is bounded on every open ball $B_X(0,r)$. Consider any F-space, e.g. $X = C(\Omega)$ where $\Omega\subseteq\mathbb{R}^n$ some open domain, with the usual semi-norms $p_K(f) = \max_K|f(x)|$ for $K\Subset\Omega$. Fix an exhausting sequence K_i and let p_i be p_{K_i}, these semi-norms induce a complete metric on X.

For every $y\in\Omega$ the point evaluation $\varphi_y(f) := f(y)$ is a continuous linear functional on X. Consider the open ball $B_X(0,\frac{3}{4})$, then the neighborhood $\{f \mid p_1(f) < \frac{1}{4}\} \subseteq B_X(0,\frac{3}{4})$ because for every f like that, $d(f,0) < \frac{3}{4}$ by definition of the metric. Take $\varphi_y(f) := f(y)$ for some $y\notin K_1$, then for every $\alpha \gg 1$ there is some $f\in B_X(0,\frac{3}{4})$ and $f(y) = \alpha$.

This means that φ_y is not bounded on $B_X(0,\frac{3}{4})$, since $\sup_{f\in B_X(0,\frac{3}{4})}|\varphi_y(f)| = \infty$. This means that for F-spaces the continuity of linear functionals is not the same as being bounded on every open ball. Note that this is not contradicting the fact that the continuity of $\varphi\in X^*$ implies that for every $\varepsilon > 0$ some $\delta > 0$ such that $\sup_{f\in B_X(0,\delta)}|\varphi(f)| < \varepsilon$.

We return to continuity of bilinear forms on $X\times Y$. The above shows that the proof of Theorem 76 cannot be salvaged for F-space, since we use the fact that for every y, $\sup_{x\in U}|f_x(y)| < \infty$. However we can use the metric of F-spaces to prove the theorem for F-spaces.

Proof of Theorem 76 for F-spaces. Since $X\times Y$ is metric, it is enough to check that whenever $(x_n,y_n)\to(0,0)$ we have that $Q(x_n,y_n)\to 0$ as well. Define for every n, $f_n(y) := Q(x_n,y)$ and $f(y) := Q(0,y)$, then $f, f_n \in Y^*$ by the assumption on continuity for each variable. Now for every $y\in Y$, $f_n(y)\to f(y)$ in Y^*. Then for every $y\in Y$ we have now that $\sup_n|f_n(y)| < \infty$, and by the completeness of Y there is some open ball $V = B_Y(0,\delta)$ such that $\sup_n\sup_{y\in V}|f_n(y)| < \infty$. Thus, we have that $|Q(x_n,y_n)| = |f_n(y_n)| < \varepsilon$ for δ small enough and $\varepsilon\to 0$ gives us the wanted result. $\square$

Chapter 6

Lebesgue Spaces

Let (Ω, Σ, μ) be a measure space, with μ being a non-negative σ-finite measure. We consider $L^p(\Omega, \Sigma, \mu)$ for $1 \leq p \leq \infty$, and when the measure and σ-algebra is clear from context (which is usually the case) we write $L^p(\Omega)$. We define:

- For $p < \infty$, $L^p(\Omega) = \{f : \Omega \to \mathbb{C} \mid f \text{ measurable and } \int_\Omega |f|^p \, \mathrm{d}\mu < \infty\}$. And we define $\|f\|_p = (\int_\Omega |f|^p \, \mathrm{d}\mu)^{1/p}$.
- For $p = \infty$, $L^\infty(\Omega) = \{f : \Omega \to \mathbb{C} \mid f \text{ measurable and } \operatorname{ess\,sup}_\Omega |f| < \infty\}$. Where $\operatorname{ess\,sup}$ is the infinimum of the supremum on co-null sets. Namely $L^\infty(\Omega)$ is the measurable functions which are bounded almost everywhere. Similarly, $\|f\|_\infty = \operatorname{ess\,sup}_\Omega |f|$.

We investigate the basic properties of L^p spaces, the case of $p = \infty$ is generally much easier than $p < \infty$, and we ignore it here almost entirely unless explicitly said otherwise.

Proposition 79 (Minkowski inequality). *For $p < \infty$, $\|\cdot\|_p$ satisfies the triangle inequality (and therefore a normed space).*

Proof. Take $f, g \in L^p$, define f_0 by $|f(x)| = \|f\|_p f_0(p)$ and similarly define g_0 (without loss of generality, $\|f\|_p, \|g\|_p > 0$). We have now that:

$$\begin{aligned} f(x) + g(x)|^p &= |\|f\|_p f_0(x) + \|g\|_p g_0(x)|^p \\ &= (\|f\|_p + \|g\|_p)^p \left| \frac{\|f\|_p}{\|f\|_p + \|g\|_p} f_0(x) + \frac{\|g\|_p}{\|f\|_p + \|g\|_p} g_0(x) \right|^p \\ &\leq (\|f\|_p + \|g\|_p)^p \left(\frac{\|f\|_p}{\|f\|_p + \|g\|_p} f_0(x)^p + \frac{\|g\|_p}{\|f\|_p + \|g\|_p} g_0(x)^p \right) \end{aligned}$$

Now integrating gives us that $\int_\Omega |f + g|^p \, \mathrm{d}\,\mu \leq (\|f\|_p + \|g\|_p)^p$ and therefore $\|f + g\|_p \leq \|f\|_p + \|g\|_p$ as wanted, and in particular $f + g \in L^p$ as well. $\square$

Proposition 80 (Riesz-Fischer Theorem). *For $p < \infty$, L^p is a complete normed space.*

Proof. Let $\{f_n\}_{n=1}^\infty \subseteq L^p$ be a Cauchy sequence. Without loss of generality we may assume that $\|f_n - f_{n+1}\|_p < 2^{-n}$, otherwise dilute the sequence. Consider for each $N \in \mathbb{N}$ the function $g_N(x) = \sum_{n=1}^N |f_{n+1}(x) - f_n(x)|$, for every N we have that $\|g_N\|_p \leq \sum_{n=1}^N \|f_{n+1} - f_n\|_p \leq 1$. Moreover it is easy to see that $g_N(x) \geq 0$ for any x.

Therefore for every $x \in \Omega$, $\{g_N(x)\}_{N=1}^{\infty}$ is increasing and $\sup_N \int |g_N(x)|^p \,\mathrm{d}\,\mu < \infty$. By the monotone convergence theorem we have some g such that $g_N(x) \uparrow g(x)$ and $g(x) < \infty$ almost everywhere, and also $\|g\|_p \leq 1$.

Consider now the series, for every fixed x, $\sum_{n=1}^{\infty} f_{n+1}(x) - f_n(x)$ is absolutely convergent almost everywhere. But this is a telescopic series, and so it converges to some $f(x)$ almost everywhere, and $|f(x)| \leq g(x) + |f_1(x)|$ for all $x \in \Omega$.

We have, if so, that $0 \xleftarrow{\text{a.e}} |f_n(x) - f(x)|^p \leq 2||f_1(x)| + g(x)|^p$ But the right hand side is an integrable function, and by the triangle inequality $f \in L^p$. Finally, using Lebesgue's dominated convergence theorem $\int_\Omega |f_n - f|^p \,\mathrm{d}\,\mu \to 0$, so $\|f_n - f\|_p \to 0$ as well, therefore the Cauchy sequence converges to f. $\square$

The above proof can be modified for $p = \infty$ as follows, by the Cauchy condition we have that $\operatorname{ess\,sup}_{x\in\Omega} |f_n(x) - f_m(x)| < \varepsilon$ for large enough $n > m$, define f as the pointwise limit, and then we can show that $\operatorname{ess\,sup} |f| < \infty$ and $\operatorname{ess\,sup} |f - f_n| \to 0$.

Definition 81. Define on $\mathbb{N}$ the counting measure μ such that for all $n \in \mathbb{N}$, $\mu(\{n\}) = 1$. We write $\ell^p(\mathbb{N}) := L^p(\mathbb{N}, \mathcal{P}(\mathbb{N}), \mu)$.

6.1 Hölder Inequality

Suppose that $1 < p < \infty$ and let $q = \frac{p-1}{p}$, namely $\frac{1}{p} + \frac{1}{q} = 1$. Then for every $a, b \geq 0$ we have that $ab \leq \frac{a^p}{p} + \frac{b^q}{q}$.

More generally, if $a_1, \ldots, a_n \geq 0$, and $p_1, \ldots, p_n$ such that $\sum_{i=1}^{n} \frac{1}{p_i} = 1$, then we have that $\prod_{i=1}^{n} a_i \leq \sum_{i=1}^{n} \frac{a_i^{p_i}}{p_i}$.

Proposition 82. *Suppose that $f \in L^p$ and $g \in L^q$, such that $\frac{1}{p} + \frac{1}{q} = 1$. Then we have that*

$$\int_\Omega |fg \,\mathrm{d}\mu| \leq \|f\|_p \|g\|_q.$$

Proof. The case for $p = \infty$ (or $p = 1$) is trivial: $\int_\Omega |fg| \,\mathrm{d}\,\mu \leq \|f\|_\infty \int_\Omega |g| \,\mathrm{d}\,\mu$. So we may assume $1 < p, q < \infty$. Take $\alpha > 0$ and have that

$$|f(x)g(x)| = \left| \frac{f(x)}{\alpha} \cdot \alpha g(x) \right| \leq \frac{|f(x)|^p}{p\alpha^p} + \frac{\alpha^q |g(x)|^q}{q}.$$

Therefore we have that,

$$\int_\Omega |fg| \,\mathrm{d}\,\mu \leq \frac{1}{p\alpha^p} \|f\|_p^p + \frac{\alpha^q}{q} \|g\|_q^q.$$

If we have that $\frac{\|f\|_p^p}{\alpha^p} = \alpha^q \|g\|_q^q$, then we have that $\alpha^{p+q} = \frac{\|f\|_p^p}{\|g\|_q^q}$. From this follows that

$$\alpha^p = (\alpha^{p+q})^{\frac{p}{p+q}} = \left(\frac{\|f\|_p^p}{\|g\|_q^q} \right)^{\frac{1}{q}} = \frac{\|f\|_p^{p/q}}{\|g\|_q}.$$

Returning to the integral and the inequality after choosing the appropriate α, we have that:

$$\frac{1}{p\alpha^p} \|f\|_p^p + \frac{\alpha^q}{q} \|g\|_q^q = \left(\frac{1}{p} + \frac{1}{q} \right) \frac{\|g\|_q}{\|f\|_p^{p/q}} \|f\|_p^{p - \frac{p}{q}} = \|f\|_p \|g\|_q. \qquad \square$$

Corollary 83. *If $g \in L^q$ then it defines a continuous linear functional on L^p by defining*

$$F_g(f) = \int_\Omega fg \,\mathrm{d}\,\mu.$$

Moreover $\|F_g\| = \|g\|_q$.

Proof. The only nontrivial part is showing that $\|F_g\| \geq \|g\|_q$. We take $h(x)$ such that $|h(x)| = 1$ and $h(x)g(x) = |g(x)|$. Define $f(x) = |g(x)|^{q-1}h(x)$. Then $f \in L^p$, since we have that:

$$|f(x)|^p = |g(x)|^{p(q-1)} = |g(x)|^q.$$

Moreover we have that $\|f\|_p = \|g\|_q^{q/p}$. Calculating $F_g(f)$ we can see that the result is in fact $\|g\|_q^q$, and therefore

$$\|F_g\| \geq \frac{\|g\|_q^q}{\|f\|_p} = \frac{\|g\|_q^q}{\|g\|_q^{q/p}} = \|g\|_q. \qquad \square$$

Corollary 84. *The above holds also for $p = 1$ and $q = \infty$, and for $p = \infty$ and $q = 1$.*

Proof. We first deal with $p = 1$. Again the only nontrivial part is showing $\|F_g\| = \|g\|_\infty = M$. Fix $\varepsilon > 0$, and E_ε such that $\mu(E_\varepsilon) > 0$ and $\{|g(x)| > M - \varepsilon \mid x \in E_\varepsilon\}$. Let $h(x) = \operatorname{sgn} g(x)$, so $g(x)h(x) = |g(x)|$. Define $f(x) = \chi_{E_\varepsilon} \cdot \frac{h(x)}{\mu(E_\varepsilon)}$. Then $\|f\|_1 = 1$ and we have that:

$$\int_\Omega fg \,\mathrm{d}\,\mu = \frac{1}{\mu(E_\varepsilon)} \int_{E_\varepsilon} |g(x)| \,\mathrm{d}\,\mu \geq \frac{1}{\mu(E_\varepsilon)} \cdot (M - \varepsilon) = M - \varepsilon.$$

It follows that $\|F_g\| \geq M - \varepsilon$ for all $\varepsilon > 0$, and thus $\|F_g\| = \|g\|_\infty = M$.

Next we deal with $p = \infty$ and $q = 1$. Now $g \in L^1$, and we want to show that $\|F_g\| = \|g\|_1$. Define $f(x) = \chi_E \cdot h(x)$ where $h = \operatorname{sgn} g$. It follows that $F_g(f) = \int_E |g| \,\mathrm{d}\,\mu$, and therefore $\|F_g\| \geq \sup_{E \in \Sigma} \int_E |g| \,\mathrm{d}\,\mu = \|g\|_1$. $\square$

It follows, if so, that the map taking $g \in L^q$ to $F_g \in (L^p)^*$ is a linear isometry. And in fact the following theorem shows that it is also surjective, at least when $p < \infty$.

6.2 Riesz Theorem

Theorem 85 (Riesz Theorem)**.** *Let $1 \leq p < \infty$. Then for every continuous functional F on L^p, there is some $g \in L^q$ such that $F = F_g$, as defined above.*

Proof. We begin by reducing the case where F only gives real values. If we know that for every real-valued functional on the real-valued version of L^p the theorem is true, and F is a complex-valued functional on L^p, then $F(f) = \operatorname{Re} F(f) + i \operatorname{Im} F(f)$. We take $F_1 = \operatorname{Re} F$ and $F_2 = \operatorname{Im} F$ and we get that $F = F_1 + iF_2$, and then there are g_1, g_2 such that $F(f) = \int_\Omega g_1 f \,\mathrm{d}\,\mu + i \int_\Omega g_2 f \,\mathrm{d}\,\mu$. And now $F(f) = F(\operatorname{Re} f) + iF(\operatorname{Im} f) = \int_\Omega gf \,\mathrm{d}\,\mu$ and the conclusion follows.

Lemma 86. *Suppose that $1 \leq p < \infty$, the simple functions are dense in L^p.*

Proof. Suppose that $f \geq 0$, then there is a sequence of simple functions $0 \leq \varphi_n \uparrow f$ almost everywhere. Therefore,

$$\forall n, \int_\Omega \varphi^p \leq \int_\Omega f^p < \infty.$$

Therefore $|\varphi_n - f|^p \leq 2^p f^p$ implies that $\int_\Omega |\varphi_n - f|^p \to 0$ (almost everywhere).

If f is an arbitrary real-valued function, we define f_+ and f_- in the usual way, and apply the previous part to each on separately. $\square$

Lemma 87. *Suppose that $\mu(\Omega) < \infty$ and g is a function such that for every simple function it holds that $|\int_\Omega g\varphi \,\mathrm{d}\,\mu| \leq M\|\varphi\|_p$, when $M > 0$ is some constant. Then $g \in L^q(\Omega)$ and $\|g\|_q \leq M$.*

Proof. First we prove for $p > 1$. Let $\{\psi_n\}_{n=1}^\infty$ a sequence of non-negative simple functions such that $\psi_n \uparrow |g|^q$. Define now $\varphi_n = \psi_n^{1/p} \operatorname{sgn} g$. Then these are also simple functions and we have that:

$$0 \leq \varphi_n g = \psi_n^{1/p}|g| \geq \psi_n^{1/p}\psi_n^{1/q} = \psi_n.$$

Therefore,

$$\int_\Omega \psi_n \,\mathrm{d}\,\mu \leq \int_\Omega \varphi_n g \,\mathrm{d}\,\mu \leq M\|\varphi_n\|_p.$$

But we also have that $\int_\Omega |\varphi_n|^p \,\mathrm{d}\,\mu = \int_\Omega \psi_n \,\mathrm{d}\,\mu$, so $\|\varphi_n\|_p = (\int_\Omega \psi_n \,\mathrm{d}\,\mu)^{1/p}$ and therefore,

$$\int_\Omega \psi_n \,\mathrm{d}\,\mu \leq M\left(\int_\Omega \psi_n \,\mathrm{d}\,\mu\right)^{1/p} \implies \left(\int_\Omega \psi_n \,\mathrm{d}\,\mu\right)^{1/q} \leq M.$$

Using the monotone convergence theorem we obtain that $|g| \in L^q$, so $g \in L^q$ as well and $\|g\|_q \leq M$ as wanted.

If $p = 1$, then $|\int_\Omega g\varphi \,\mathrm{d}\mu| \leq M\|\varphi\|_1$ for every simple φ. In particular if $\varphi = \chi_E \operatorname{sgn} g$ then $\int_E |g| \,\mathrm{d}\,\mu \leq M\mu(E)$. Now consider $E = \{x \mid \delta + M < |g(x)|\}$ for some fixed $\delta > 0$. Then we have that:

$$(M + \delta)\mu(E) \leq \int_E |g| \,\mathrm{d}\,\mu \leq M\mu(E) \implies \mu(E) = 0.$$

Therefore $g \in L^\infty$ and $\operatorname{ess\,sup} |g| = \|g\|_\infty \leq M$. $\square$

Lemma 88. *Let $E \in \Sigma$ and $E = \bigcup_{n=1}^\infty E_n$, such that $\{E_n \mid n \in \mathbb{N}\} \subseteq \Sigma$ is a family of pairwise disjoint sets. Let f be a function, we write $f_n = f\chi_{E_n}$. Then the following equivalence holds:*

$$f \in L^p(E) \iff \sum_{n=1}^\infty \|f_n\|_p^p < \infty,$$

and in that case $f = \sum_{n=1}^\infty f_n$ in $L^p(E)$.

Proof. If $f \in L^p(E)$ then for every $N \in \mathbb{N}$ we have that:

$$\infty > \int_E |f|^p \,\mathrm{d}\,\mu \geq \int_E \sum_{n=1}^N |f_n|^p \,\mathrm{d}\,\mu = \int_E \left|\sum_{n=1}^N f_n\right|^p \,\mathrm{d}\,\mu.$$

Therefore the conclusion follows. On the other hand, if $\sum_{n=1}^{\infty} \int_{E_n} |f_n|^p \,\mathrm{d}\,\mu < \infty$ then by the monotone convergence theorem, $\sum_{n=1}^{\infty} \|f_n\|_p^p < \infty$, then $\sum_{n=1}^{N} |f_n|^p \uparrow |f|^p$. Therefore,

$$\sum_{n=1}^{\infty} \int_{\Omega} |f_n|^p \,\mathrm{d}\,\mu = \sum_{n=1}^{\infty} \int_{E_n} |f_n|^p \,\mathrm{d}\,\mu = \int_{E} |f|^p \,\mathrm{d}\,\mu.$$

In particular $\sum_{n=1}^{N} f_n \to f$ in $L^p(E)$. $\square$

We can start the proof of Riesz' theorem now. We remark, again, that our functions (and spaces) are real. We also know that the map $g \mapsto F_g$ is an isometry, so every functional in $(L^p)^*$ has at most one $g \in L^q$ matching it.

We first consider the case where $\mu(\Omega) < \infty$. Let $F \in (L^p)^*$, then for every $E \in \Sigma$ define $\nu(E) := F(\chi_E)$. If $E = \bigcup_{n=1}^{\infty} E_n$, where the E_n's are pairwise disjointm define $\alpha_n = \operatorname{sgn}(\chi_{E_n})$, and consider the function f defined as $\sum_{n=1}^{\infty} \alpha_n \chi_{E_n}$. From the third lemma, we have that $\|f_n = \alpha_n \chi(E_n)\|_p = \mu(E_n)^{1/p}$. Therefore:

$$\sum_{n=1}^{\infty} \|f_n\|_p^p = \|f\|_p^p.$$

So we have that $\sum_{n=1}^{N} f_n \to f$ in L^p. By continuity of F we have that $F(\sum_{n=1}^{N} f_n) \to F(f)$, and so

$$F(f) \leftarrow \sum_{n=1}^{N} F(\chi_{E_n}) F(\chi_{E_n}) = \sum_{n=1}^{N} \alpha_n \nu(E_n) = \sum_{n=1}^{N} |\nu(E_n)|.$$

And therefore $\sum_{n=1}^{\infty} \nu(E_n)$ is absolutely convergent. On the other hand, $\sum_{n=1}^{\infty} \chi_{E_n} = \chi_E$ (as a sum in L^p). So $F(\sum_{n=1}^{N} \chi_{E_n}) \to F(\chi_E)$, so $\sum_{n=1}^{\infty} \nu(E_n) = \nu(E)$.

Therefore $\nu(E)$ is a finite measure (although perhaps not non-negative). We have that $|F(\alpha\chi_E)| \leq \|F\| \|\chi_E\|_p \leq \|F\| \mu(\Omega)^{1/p}$. Moreover if $\mu(E) = 0$ then $\chi_E = 0$ in L^p, so $F(\chi_E) = 0$ and in that case $\nu(E) = 0$ as well. We have that $\nu \ll \mu$ (ν is absolutely continuous with respect to μ[1]), and the Radon-Nikodym Theorem implies there is $g \in L^1(\Omega, \Sigma, \mu)$ such that $\nu(E) = \int_E g \,\mathrm{d}\,\mu = \int_\Omega g\chi_E \,\mathrm{d}\,\mu$.

Now by finite additivity, for every simple function φ we have that $F(\varphi) = \int_\Omega g\varphi \,\mathrm{d}\,\mu$. By the continuity of F we have that $\|F\| \|\varphi\|_p \geq |F(\varphi)|$.

Using the second lemma we have that $g \in L^q$ and $\|g\|_q \leq \|F\|$. Consider now $F_g \in (L^p)^*$, then for every simple function φ, $F(\varphi) = F_g(\varphi)$. By the first lemma, the simple functions are dense in L^p, and since F and F_g are continuous and agree on a dense subset they are equal.

Now suppose that $\mu(\Omega) = \infty$, by σ-finiteness there are $\Sigma_n \uparrow \Omega$ such that $\mu(\Sigma_n)$ is finite. We repeat the proof on each Ω_n, to obtain $g_n \in L^q(\Omega)$ such that $g_n = g_n \chi_{\Omega_n}$. If so, for every $f \in L^p(\Omega_n)$ we have that

$$F(f) = \int_{\Omega_n} fg \,\mathrm{d}\,\mu.$$

As we assume that $\Omega_n \subseteq \Omega_{n+1}$, this gives us a sequence g_n in L^q such that $g_{n+1} \restriction_{\Omega_n} = g_n$. So the function $g(x) = g_n(x)$ for some n such that $x \in \Omega_n$ is a well-defined function. Moreover if $f \in L^p(\Omega_n)$ we have $F(f) = \int_\Omega gf \,\mathrm{d}\,\mu$. And this completes the proof by convergence theorem. $\square$

[1]So every μ-null set is ν-null as well.

6.2.1 Some remarks

It is known that $L^\infty(\Omega) \supseteq L^1(\Omega)$, since if $g \in L^1$, then for $F_g(f) = \int_\Omega fg \,\mathrm{d}\,\mu$ we get from Hölder's inequality that $|F(f)| \leq \|f\|_\infty \|g\|_1 \implies \|F\| \leq \|g\|_1$. Moreover we actually get $\|F\| = \|g\|_1$.

Suppose that $L^1 = (L^\infty)^*$, take for example $\Omega = \mathbb{R}$ and μ the Borel measure. Then L^1 is separable, for example by simple functions with rational values. This means that L^∞ has to be separable as well. But this is not the case, since $\varphi(t) = \chi_{\leq t}$ is an uncountable family of functions in L^∞ such that for $t \neq s$, $\|\varphi_t - \varphi - s\|_\infty = 1$. So we cannot *prove* that equality holds regardless to the measure space.

Consider next ℓ^p spaces. Then $f \in \ell^p$ if and only if $\sum_{n=1}^\infty |f(n)|^p < \infty$. In this case we also have that $\ell^1 \neq (\ell^\infty)^*$ by considering the characteristics functions of subsets of $\mathbb{N}$. However, $\ell^1 = c^*$ and $\ell^1 = c_0^*$.[2]

[2]It is consistent, however, without the axiom of choice that $\ell^1 = (\ell^\infty)^*$.

Chapter 7

Weak Topologies

7.1 The Weak Topology on Normed Spaces

Let X be a normed space with $\|\cdot\|_X$ its norm. Recall that X^* is a Banach space and for $f \in X^*$ we have that $|f(x)| \leq \|f\|\|x\|_X$. Additionally there is a linear isometry $\Phi\colon X \to X^{**}$ defined by $\Phi(x)(f) = f(x)$. If Φ is onto X^{**} we say that X is reflexive.

An immediate conclusion from the previous chapter we have that L^p is reflexive for $1 < p < \infty$. Moreover for $p = 2$, $L^2 = (L^2)^*$.

Note that if $f \in X^*$ then $|f(x)|$ is a semi-norm on X.

Definition 89 (Weak topology)**.** Given a normed space X, the *weak topology* on X is the locally convex topology defined from $\{|f(x)| \mid f \in X^*\}$.

Recall that an open neighborhood U of 0 in the weak topology is given by $f_1, \dots, f_n \in X^*$ and $\varepsilon_1, \dots, \varepsilon_n > 0$ so $U = U^{\varepsilon_1,\dots,\varepsilon_n}_{f_1,\dots,f_n} = \{x \mid |f_i(x)| < \varepsilon_i\}$. Because there are uncountably many f's it is often the case the weak topology is not Frechét.

Proposition 90. *The weak topology is weaker than the normed topology. Namely, every open set in the weak topology is open in the normed topology.*

Proof. It suffices to consider neighborhoods of 0. If $U^{\varepsilon_1,\dots,\varepsilon_n}_{f_1,\dots,f_n}$ is an open neighborhood of 0, then it is the intersection of preimages $|f_i|^{-1}((-\varepsilon_i, \varepsilon_i))$. By the definition we have that f_i is continuous, so this is a finite intersection of open sets in the normed topology. $\square$

Example 91. *Suppose that $\{x_n\}_{n=1}^\infty$ weakly converges to x. Then for every $f \in X^*$ we have that $\lim_{n\to\infty} f(x_n) = f(x)$.*

This is because whenever U is an open neighborhood of x, then for every $\varepsilon > 0$ there is some N such that for all $n > N$, $|f(x - x_n)| < \varepsilon$. But it is clear now that in fact $\{x_n\}_{n=1}^\infty$ weakly converges to x if and only if for every $f \in X^*$, $f(x_n) \to f(x)$.

It is also clear that if $\|x_n - x\|_X \to 0$, then x_n weakly converges to x. We will write $x_n \xrightarrow{w} x$ to denote weak convergence.

Proposition 92. *Let $x_n \xrightarrow{w} x$, then the following is true:*

1. $\sup\{\|x_n\|_X \mid n \in \mathbb{N}\} < \infty$.

2. $\|x\|_X \leq \liminf \|x_n\|_X$.

Proof. Recall that for every $f \in X^*$, $f(x_n) \to f(x)$. Consider the canonical embedding $\Phi\colon X \to X^{**}$. Then $\Phi(x_n)(f) \to \Phi(x)(f)$, for every $f \in X^*$. Since X^* is complete we can apply the uniform boundedness theorem to have that $\sup_n \|\Phi(x_n)\| < \infty$ so $\sup \|x_n\| < \infty$ since Φ is norm-preserving.

The second point follows from the fact that $\|T\| \leq \liminf \|T_n\|$ when $T_n(y) \to T(y)$ for all $y \in Y$ whenever T is continuous from some Banach space X to a normed space. Take $Y = X^*$ and $T_n = \Phi(x_n)$, and we get again the second inequality by the norm-preservation of Φ. $\square$

Suppose that $\{f_n\}_{n=1}^{\infty} \subseteq L^p$ for $1 \leq p < \infty$, and suppose that there is some $f \in L^p$ so for every $g \in L^q$:

$$\lim_{n\to\infty} \int_\Omega f_n g \,\mathrm{d}\mu = \int_\Omega f g \,\mathrm{d}\mu.$$

In other words, $f_n \xrightarrow{w} f$ in L^p (endowed with the weak topology). Then we have by the previous proposition that:

1. $\sup_n \|f_n\|_p < \infty$.
2. $\|f\|_p \leq \liminf_{n\to\infty} \|f_n\|_p$.

Proposition 93. *Let $M \subseteq X$ be a closed subspace in the normed topology. Then M is also a closed subspace in the weak topology.*

Proof. Let M such subspace and $y \notin M$, then there is some $f \in X^*$ such that $f \restriction M = 0$ and $f(y) = 1$. Now the weak-open set $U = \{v \mid |f(y - v)| < \frac{1}{2}\}$ is disjoint from M, since whenever $m \in M$ we have that $f(y - m) = f(y) = 1$, and $y \in U$. So M is weak-closed as wanted. $\square$

This generalizes nicely to the following proposition.

Proposition 94. *Let $K \subseteq X$ a convex and balanced which is closed in the normed topology. Then K is closed in the weak topology as well.*

Proof. Let $y \in X \setminus K$. By the regularity of X we can extend K slightly to make it absorbing as well while not adding y. So we can apply Mazur's theorem, and find $f \in X^*$ such that $|f(x)| \leq 1$ for $x \in K$ and $f(y) = 1 + \delta > 1$. Consider the weak-open neighborhood of y, $U = \{v \mid |f(y - v)| < \delta\}$, by a similar argument as before U is open and disjoint from K as wanted. $\square$

It follows that $\overline{B_X(0,r)} = \{x \in X \mid \|x\|_X \leq r\}$ is weak-closed.

Proposition 95. *If $c > 0$ is a constant, and $\{x \in X \mid c < |f(x)|\}$ for some fixed $f \in X^*$ is weak-open.*

Proof. Fix $x_0 \in X$ such that $|f(x_0)| > c$, we want to find a weak-open set U such that $x_0 \in U$ and for all $x \in U$, $|f(x)| > c$ as well. Denote by $\delta = |f(x_0)| - c > 0$, and consider $V = \{x \mid |f(x)| < \delta/2\}$, and let $U = x_0 + V$. Then whenever $y \in U$ we have that $y - x_0 \in V$, so $|f(y - x_0)| < \frac{\delta}{2}$ and therefore $|f(y) - f(x_0)| < \frac{\delta}{2}$, so $|f(y)| > c + \frac{\delta}{2}$, so U is as wanted. $\square$

Proposition 96. *Let $\{x_n\}_{n=1}^{\infty} \subseteq X$ and $x \in X$ such that: $\sup_n \|x_n\|_X < \infty$, and there is a [normed-]dense set $Y \subseteq X^*$ such that for all $g \in Y$, $\lim_n g(x_n) \to g(x)$. Then $x_n \xrightarrow{w} x$.* $\square$

7.2 Weak Convergence in Lebesgue Spaces

Suppose that $1 \le p < \infty$, and q such that $\frac{1}{p} + \frac{1}{q} = 1$, we know that $L^q = (L^p)^*$. This means that $\{f_n\}_{n=1}^\infty \subseteq L^p$ is weakly convergent to f if for all $g \in L^q$,

$$\int_\Omega g f_n \,\mathrm{d}\mu \to \int_\Omega g f \,\mathrm{d}\mu.$$

In particular, $\sup_n \|f_n\|_p < \infty$ and $\|f\|_p \le \liminf_n \|f_n\|_p$. In general the uniform boundedness theorem tells us that if $\{\int g f_n\}_{n=1}^\infty$ is bounded for all $g \in L^q$, then $\sup \|f_n\|_p < \infty$.

What happens if $p = \infty$? Suppose that $\{f_n\}_{n=1}^\infty \subseteq L^\infty$, and for every $g \in L^1$ the sequence $\int g f_n$ is bounded. Does that mean that $\|f_n\|_\infty$ is bounded? Since $L^\infty = (L^1)^*$ we can consider $\{f_n\}_{n=1}^\infty$ as continuous functionals on L^1, and we are given that the sequences $\{\int_\Omega f_n g\}_{n=1}^\infty$ is bounded for every $g \in L^1$, so uniform boundedness gives us that $\|f_n\|_\infty$ is indeed bounded.

Let $\Omega = S^1$, the unit circle in $\mathbb{C}$, with μ being the usual Lebesgue measure on it, given by rotations. Consider now the sequence $\{e^{in\theta}\}_{n=0}^\infty$. Take $\varphi(\theta) \in C^1(S_1)$ and consider:

$$\int_{S^1} \varphi(\theta) e^{in\theta} \,\mathrm{d}\theta = \frac{1}{in} \int_{S^1} \varphi(\theta) \frac{\mathrm{d}}{\mathrm{d}\theta}(e^{in\theta}) \,\mathrm{d}\theta = -\frac{1}{in} \int_{S_1} \varphi'(\theta) e^{in\theta} \,\mathrm{d}\theta \xrightarrow{n \to \infty} 0,$$

Consider the $e^{in\theta}$ as elements of L^2 (as bounded functions they belong to every L^p). It follows that $\|e^{in\theta}\|_2 = \sqrt{2\pi}$. So from the last proposition, $X = L^2(S^1)$ and $\varphi \in C^1$ defines $g \in L^2$ (here $Y = C^1(S^1)$ is dense there). We get the following theorem as an easy corollary.

Theorem 97 (Riemann-Lebesgue Lemma)**.** *For every $g \in L^2(S^1)$ we have that*

$$\lim_{n \to \infty} \int g(\theta) e^{in\theta} \,\mathrm{d}\theta = 0.$$

In addition we get that for every $g \in L^\infty(S^1)$ the Riemann-Lebesgue lemma holds (since $L^\infty(S^1) \subseteq L^2(S^1)$). But now $e^{in\theta} \in L^1(S^1)$ is a functional on $L^\infty(S^1)$, and therefore $e^{in\theta} \xrightarrow{w} 0$ in $L^1(S^1)$. On the other hand, $\|e^{in\theta}\|_1 = 2\pi$, so it is certainly not the case that $e^{in\theta} \not\to 0$ in $L^1(S^1)$, so it follows that the sequence $e^{in\theta}$ cannot converge in norm to anything. Therefore in $L^1(S^1)$ weak convergence need not imply strong convergence.[1]

7.3 Back in the General Setting

We saw that a normed space X carries two topologies, strong (normed) and weak. Both topologies make it into a locally convex space. We can ask, if so, what are the continuous functionals on each of them.

Proposition 98. *If $f\colon X \to \mathbb{C}$ is a linear functional, then it is strongly continuous[2] if and only if it is weakly continuous.*

[1]This property is called the *Schur property*, and therefore $L^1(S^1)$ does *not* have the Schur property.

[2]Continuous in the normed topology.

Proof. Suppose that $f \in X^*$, namely f is strongly continuous. Fix $\varepsilon > 0$ and consider $f^{-1}(B_{\mathbb{C}}(0,\varepsilon)) = \{x \mid |f(x)| < \varepsilon\}$, and by definition of the weak topology it is an open neighborhood of 0. So f is weakly continuous at 0, and by linearity f is weakly continuous. The other direction is trivial. $\square$

In general the weak topology is the weakest topology such that every $f \in X^*$ is still continuous in that topology.

Theorem 99. *Let X and Y be normed spaces, then $T\colon X \to Y$ linear is continuous with respect to both normed topologies if and only if it is continuous with respect to both weak topologies. Namely $B(X,Y) = B(X_w, Y_w)$, where X_w and Y_w denote the weak topologies.*

Proof. Let $T \in B(X,Y)$ and we want to show that whenever V is weakly open in Y, $T^{-1}(V)$ is weakly open in X. We can assume that $V = V^{\varepsilon_1,\ldots,\varepsilon_n}_{g_1,\ldots,g_n}$, then we are looking for $U = U^{\delta_1,\ldots,\delta_m}_{f_1,\ldots,f_m}$ weakly open in X such that $T(U) \subseteq V$.

Let $f_i = g_i \circ T$ for $i < n$ (so $m = n$ in our case). By the fact that T and g_i are strongly continuous we have that $f_i \in X^*$. Taking $\delta_i = \varepsilon_i$ gives us that if $x \in U$ the $|g_i(Tx)| = |f_i(x)| < \varepsilon_i$, so $Tx \in V$ as wanted.

In the other direction assume that X and Y are Banach, if T is continuous with respect to the weak topologies, we will show that $G(T)$ is closed and by the closed graph theorem T is continuous.

Let $\langle x_n, Tx_n\rangle \to \langle x, y\rangle$, we want to show that $Tx = y$, but for every linear function $g \in Y^*$ we have that $g\colon T$ is continuous on X_w. Therefore $(g \circ T)(x_n) \to g(y)$. So $g(Tx) = g(y)$ for every $g \in Y^*$, and by the fact that functionals on Y separate points gives us that $Tx = y$ as wanted.

If X and Y are not assumed to be Banach spaces, we first prove the following lemma.

Lemma 100. *If $T\colon X \to Y$ is linear, and for every $g \in Y^*$, $g \circ T \in X^*$, then T is continuous.*

Proof. Define the following bilinear form on $X \times Y^*$, $Q(x,g) = g(Tx)$. It is not hard to see that Q is linear and that for every $x \in X$ we have that $Q(x,\cdot) = \Phi(Tx) \in Y^{**}$ and also the assumption tells us that for every $g \in Y^*$, $Q(\cdot, g) \in X^*$. By a theorem we proved it means that Q is a continuous bilinear form.

Take $x_n \to 0$ in X, then we have that $Q(x_n,\cdot) \to Q(0,\cdot) = \Phi(T(0)) = \Phi(0) = 0$. Since Φ is an isometry from Y into Y^{**} it follows that $Tx_n \to 0$ in Y as wanted. $\square$

Using the lemma , if T is continuous in the weak topologies the condition of the lemma holds and therefore T has to be continuous. $\square$

Recall that if X and Y are normed spaces, then $\overline{\Phi_X(X)}$ and $\overline{\Phi_Y(Y)}$ are Banach completions of X and Y respectively. If T is continuous between X and Y, then it is Lipschitz and therefore can be extended to the Banach completions. We can now ask whether or not a weakly continuous operator between X and Y can be extended to the weak topologies of their Banach completion. The theorem we proved above indeed tells us that the answer is positive.

7.4 The Weak-* Topology

We begin this section with a lemma.

Lemma 101. *Let X be a linear space, and suppose that $f_1, \dots, f_n$ are linear functions and f a linear functional with the following property: For fixed $\varepsilon_1, \dots, \varepsilon_n$, if $|f_i(x)| < \varepsilon_i$ then $|f(x)| < 1$. Then f is a linear combination of the f_i's.*

Proof. First, observe that $\bigcap_{i=1}^n \ker(f_i) \subseteq \ker(f)$. Define a linear operator $T\colon X \to \mathbb{C}^n$ by $Tx = \langle f_1(x), \dots, f_n(x)\rangle$, this is a linear operator, and let L be the image of T. We define $\varphi\colon L \to \mathbb{C}$ by $\varphi(x) = f(x)$, this is a linear functional on L. We can now extend φ to a functional on $\mathbb{C}^n$, so $\varphi(z_1, \dots, z_n) = \sum \alpha_i z_i$ so $f(x) = \sum \alpha_i f(x_i)$. □

Suppose that X is a linear space and $\mathcal{F}$ is a separating family of linear functionals on X. Then $\mathcal{F}$ defines a locally convex topology $\tau_{\mathcal{F}}$ on X given by the semi-norms, $|f(x)|$ for $f \in \mathcal{F}$. If f is a linear functional with respect to $\tau_{\mathcal{F}}$ then f is a linear combination of elements from $\mathcal{F}$.

Definition 102. Given a normed space X, the *weak-$*$ topology* is the topology on X^* given by the linear functionals in X^{**} obtained from the natural embedding $\Phi\colon X \to X^{**}$.

Note that $\mathcal{F} = \{\Phi(x) \mid x \in X\}$ is separating for X^* since whenever $f(x) = 0$ for all $x \in X$, $f = 0$, so if $\Phi(x)(f) = 0$ for all x, it follows that $f = 0$.

Theorem 103 (Banach-Alouglu Theorem)**.** *Let $B^* = \{f \in X^* \mid \|f\| \le 1\}$ the unit ball of X^*. Then B^* is compact in the weak-$*$ topology.*

Proof. Define a function on B^* mapping f to $\langle f(x) \mid x \in X\rangle$. This is a continuous embedding of B^* into $\prod_{x\in X} D(0, \|x\|_X)$, where $D(0, \|x\|_X)$ is the closed disc around 0 of radius $\|x\|_X$. The codomain is compact by Tychonoff's theorem, so it remains to show that the image of the embedding is closed. Let $g \in \overline{B^*}$ (the closure as computed in the product). If $x \in X$ and $\varepsilon > 0$, then there is some $f \in B^*$ such that $|f(x) - g(x)| < \varepsilon$. But by definition of $f \in X^*$ it follows that $|f(x)| \le \|x\|_X$. Therefore $|g(x)| \le \|x\|_X$ for all $x \in X$.

It remains to show that g is indeed linear. If $x, y, z \in X$ and $\varepsilon > 0$ then there is some $f \in B^*$ such that $|f(x) - g(x)|, |f(y) - g(y)|, |f(z) - g(z)| < \varepsilon$. But f is linear, so if $z = \alpha x + \beta y$, then $f(z) = \alpha f(x) + \beta f(z)$. By ε-manipulations we obtain that

$$|g(\alpha x + \beta y) - \alpha g(x) - \beta g(y)| < \varepsilon + (|\alpha| + |\beta|)\varepsilon,$$

so g is also linear as wanted. □

Corollary 104. *Let X be a normed space. Then X is linearly isometric to a subspace of $C(K)$ for a compact space K.*

Proof. Consider the map $\Phi\colon X \to X^{**}$ when considering $\Phi(x)$ restricted to B^*, the unit ball of X^*. This is a linear isometry. The weak-$*$ topology was defined as a topology on which $\Phi(x)$ are continuous. Take $K = B^*$ in the weak-$*$ topology, then we have the following equalities:

$$\|\Phi(x)\|_{C(K)} = \sup\{|\Phi(x)(f)| \mid f \in K\} = \sup\{|f(x)| \mid f \in B^*\} = \|x\|_X.$$

□

Theorem 105. *If X is a normed space then B^* is metric if and only if X is separable.*

Proof. Suppose that X is separable, let $\{x_n\}_{n=1}^\infty$ a dense countable subset. Then $\varphi_n = \Phi(x_n)$ is a separating family of semi-norms on X^*, so the locally convex topology $\{\varphi_n \mid n \in \mathbb{N}\}$ defines a metric topology. Denote by τ this topology and τ^* the weak-$*$ topology. The identity function is continuous between $(B^*, \tau^*) \to (B^*, \tau)$, but a continuous bijection from a compact space to a Hausdorff space is a homeomorphism. Therefore $\tau^* = \tau$ as wanted.

In the other direction, if B^* is metrizable in the weak-$*$ then 0 has a countable neighborhood basis $U_n = U^{\vec{\varepsilon}_n}_{\Phi(x_1^n),\ldots,\Phi(x^n_{k(n)})}$. Take $D = \{x_j^n \mid j \leq k(n), n \in \mathbb{N}\}$, this is a countable subset of X and we claim that $\overline{\operatorname{span}(D)}$ is dense in X. Otherwise there will be some $f \in X^*$ such that $f(x) = 0$ for all $x \in \overline{\operatorname{span}(D)}$ and $\|f\| = 1$, but then $f \in U_n$ for all n, therefore $f \in \bigcap U_n$ but this means $f = 0$ which is a contradiction. $\square$

>From this follows that every normed space X is linearly isometric to a subspace of $C(K)$ where K is a compact metric space, and not just a compact space. In particular every continuous linear functional on X we can match a continuous linear functional on $C(K)$. This means that every $f \in X^*$ has a Radon measure μ_f on K which defines its counterpart by $\int_K \mathrm{d}\,\mu_f(g)$ when $g \in C(K)$. On the other hand, $f(x) = \Phi(x)(f)$ with $\Phi(x) \in C(K)$ and therefore $f(x) = \int_K \mathrm{d}\,\mu_f(\Phi(x))$.

7.4.1 Some Useful Remarks

Proposition 106. *Suppose that X is a complex locally convex vector space. If $f\colon X \to \mathbb{R}$ is a continuous real-linear functional, then there is some continuous linear functional $\widetilde{f}\colon X \to \mathbb{C}$ such that $\operatorname{Re}\widetilde{f} = f$.*

Proof. We define $\widetilde{f} = f(x) - if(ix)$. It is enough to check scalar multiplication for i

$$\widetilde{f}(ix) = f(ix) - if(-x) = f(ix) + if(x) = i\widetilde{f}(x). \qquad \square$$

Proposition 107. *If $f\colon X \to \mathbb{C}$ is continuous and $\gamma \in \mathbb{R}$, then $A = \{x \mid \operatorname{Re} f(x) > \gamma\}$ is open in the weak topology.*

Proof. Take $x_0 \in A$, so $\operatorname{Re} f(x) > \gamma$. Take $\varepsilon = \frac{1}{2}(\operatorname{Re} f(x) - \gamma)$ and look at the weakly open neighborhood $U = \{z \mid |f(z)| < \varepsilon\}$. Then for $z \in U$ we have that $\operatorname{Re} f(x_0 + z) = \operatorname{Re} f(x_0) + \operatorname{Re} f(z) > \operatorname{Re} f(x_0) - |f(z)| > \gamma$. $\square$

Proposition 108. *The weak closure of a convex set is convex.*

Proof. If $\Lambda \subseteq X$ is convex, and look at its weak closure $\overline{\Lambda}$. If $\overline{\Lambda}$ is not convex then there are $x, y \in \overline{\Lambda}$ and some $t \in (0,1)$ such that $z_t = tx + (1-t)y \notin \overline{\Lambda}$. Then there is some open neighborhood V of z_t such that $V \cap \Lambda = \varnothing$. Look at the continuous bilinear form $F(u,v) = tu + (1-t)v$, and find U_x, U_y open neighborhoods of x, y respectively such that $F(U_x \times U_y) \subseteq V$. Now there are $u, v \in U_x \times U_y \cap \Lambda \times \Lambda$ but by the convexity of Λ, $F(u,v) \in V \cap \Lambda$ which is a contradiction to the assumption that $V \cap \Lambda = \varnothing$. $\square$

Remark. *Not that the weak-$*$ topology is **not** metrizable in general. While it might be that the closed unit ball is metrizable (and even compact), this metric is not canonical in any way and degenerates as the ball "grows".*

7.4.2 Some Concrete Examples

First Example

Suppose that $1 < p < \infty$, and consider L^p defined on some measure space Ω. We know that q such that $\frac{1}{p} + \frac{1}{q} = 1$ satisfies that $1 < q < \infty$ and $(L^p)^{**} = (L^q)^* = L^p$. So L^p is reflexive, and in particular L^p has the same weak topology and weak-$*$ topology.

As a corollary $B_p = \{f \in L^p \mid \|f\|_p \leq 1\}$ is weakly compact, and it is metrizable if L^p is separable. In other words, the weak topology of L^p is determined entirely by sequences. If $f_n \xrightarrow{w} f$ in B_p, it is convergent if for every $g \in L^q$, $\int_\Omega f_n g \,\mathrm{d}\mu \to \int_\Omega fg \,\mathrm{d}\mu$.

In particular if $E \subseteq \Omega$ such that $\mu(E) < \infty$, then $\chi_E \in L^q$, so taking $g = \chi_E$ we get immediately that $\int_E f_n \,\mathrm{d}\mu \to \int_E f \,\mathrm{d}\mu$.

Let $T\colon L^p \to L^p$ is a linear operator such that whenever $f_n \xrightarrow{w} f$, it follows that $Tf_n \xrightarrow{w} Tf$. Then T is weakly continuous, but therefore continuous.

Note that in the case that $p = \infty$ know that $L^\infty = (L^1)^*$, so if L^1 is separable, the unit ball of L^∞ is compact and metrizable in the weak-$*$ topology.

Second Example

Take $X = C(K)$ for some compact metric space K. We know that $X^* = \mathcal{M}_R(K)$ is the space of Radon measures on K. If X is reflexive then $X \sim C(K)$, then for every $x \in X$ we match $\Phi(x) \in X^{**}$. This means that for every $f \in X^*$ we can match a measure μ_f, so now $f(x) = \int_K \Phi(x)(g) \,\mathrm{d}\mu_f(g)$. And here $\mu_f = \delta_f$ is the atomic measure.

Take $K = [0,1]$ and consider $C[0,1]$. Suppose that we take $\mu \in \mathcal{M}_R([0,1])$ is a positive measure. Then there is some f_μ monotonous on $[0,1]$ which induces μ. More specifically, $f_\mu(x) = \mu([0,x))$. Then if $g \in C[0,1]$ μ is a continuous functional on $C[0,1]$ and we have that $\mu(g) = \int_0^1 g(x) \,\mathrm{d} f_\mu(x)$. This is the Riemann-Stieltjes integral.

Compactness now tells us that if $\{f_n\}_{n=1}^\infty$ is a sequence of monotonously increasing functions such that $f_n(0) = 0$ and $\sup_n f_n(1) \leq C$, then there is a monotonously increasing function f such that:[3]

$$\int_0^1 g(x) \,\mathrm{d} f_n(x) \to \int_0^1 g(x) \,\mathrm{d} f(x).$$

[3]This is known as Helly's Selection Theorem.

Chapter 8

Convexity and Compactness

We go back to locally convex spaces over $\mathbb{C}$. In this chapter, unless stated otherwise X is such a space.

Proposition 109. *Let $C \subseteq X$ a closed convex set and $0 \in C$ and $x_0 \notin C$. Then there exists a continuous real-valued linear functional f such that $f(x_0) > \sup_{y \in C} f(y)$.*

Proof. Since $x_0 \notin C$, there is a convex open neighborhood V of 0, such that $x_0 + V \cap C + V = \varnothing$, so in other words $x_0 + 2V \cap C = \varnothing$. Now $C + V$ is a convex open neighborhood of 0, therefore it has a Minkowski functional p. Recall that $p(x+y) \leq p(x) + p(y)$ and $\alpha p(x) = p(\alpha x)$ for all $\alpha \geq 0$.

Define $f_0(\beta x_0) = \beta p(x_0)$ for all $\beta \in \mathbb{R}$. Then $f_0 \leq p$ on $\operatorname{span}\{x_0\}$. Using the most basic form of Hahn-Banach we have an extension of f_0 to $f \colon X \to \mathbb{R}$ which is real-valued linear functional and $f(\beta x_0) = f_0(\beta x_0)$ and $|f| \leq p$ everywhere. This f is continuous as it is bounded by p, and it satisfies $f(x) \leq p(x) \leq 1$ for $x \in C$ and $f(x_0) = p(x_0) > 1$ since $x_0 \notin C + V$. Therefore the conclusion follows. $\square$

We can extend f obtained in the previous proof to $\widetilde{f} \colon X \to \mathbb{C}$ such that $\operatorname{Re} \widetilde{f} = f$. We can now reformulate the above proposition as the following corollary.

Corollary 110. *There exists $f \in X^*$ such that* $\operatorname{Re} f(x_0) > \sup_{y \in C} \operatorname{Re} f(y)$. $\square$

Corollary 111. *Every $x_0 \notin C$ has a weakly open neighborhood of x_0 disjoint from C.* $\square$

Corollary 112. *Every closed and convex set is weakly closed.* $\square$

8.1 Extremal Sets

Definition 113. Let $S \subseteq K \subseteq X$, we say that S is *extremal in K* if for every $x, y \in K$ if for $t \in (0,1)$, $tx + (1-t)y \in S$, then $x, y \in S$. We say that $y \in K$ is an *extremal point (in K)* if $\{y\}$ is extremal in K.

Theorem 114. *Let X be a locally convex complex vector space, $K \subseteq X$ is compact. Then K has an extremal point.*

Proof. Let $P = \{S \subseteq K \mid \varnothing \neq S \text{ compact and extremal}\}$, note that $K \in P$. Consider P as ordered by reverse inclusion, this partial order satisfies the condition for Zorn's lemma (note that extremal sets are closed under intersections).

If f a continuous linear functional, S is compact and extremal in K, let $M = \max_S \operatorname{Re} f$ then $S_M = \{x \in S \mid \operatorname{Re} f(x) = M\} \in P$. This is true because S_M is closed and thus compact, and if $x, y \in K$ and $tx + (1-t)y \in S_M$ then:

$$\operatorname{Re} f(tx + (1-t)y) = t \operatorname{Re} f(x) + (1-t) \operatorname{Re} f(y) = M.$$

But since $\operatorname{Re} f(x) \leq M$ and $\operatorname{Re} f(y) \leq M$ it has to be the case that $\operatorname{Re} f(x) = \operatorname{Re} f(y) = M$ so $x, y \in S_M$.

Let S be the maximal element of P obtained by Zorn's lemma. If $x, y \in S$ then for every continuous linear function f, $\operatorname{Re} f(x) = \operatorname{Re} f(y)$ and this implies that $x = y$. $\square$

Theorem 115 (Krein-Milman Theorem)**.** *Let $K \subseteq X$ compact and let Λ be the set of extremal points in K. Then $K \subseteq \overline{\operatorname{co}(\Lambda)}$.*[1]

Proof. Denote by H the set $\overline{\operatorname{co}(\Lambda)}$, it is a closed and convex set. Assume towards contradiction that there is some $x_0 \in K \setminus H$. Fix $c \in H$ and consider $C = H - \{c\}$, as a closed and convex set with $0 \in C$, by the lemma we proved there is some $f \in X^*$ such that $\operatorname{Re} f(x_0 - c) > \sup_{z \in H} \operatorname{Re} f(z - c)$ which means that $\operatorname{Re} f(x_0) > \sup_{z \in H} \operatorname{Re} f(z)$.

Let $M = \max_{y \in K} \operatorname{Re} f(y)$ and look at $K_M = \{z \in K \mid \operatorname{Re} f(z) = M\}$ which is extremal and compact in K. Our previous theorem tells us that K_M has an extremal point b. But this means that b is extremal in K as well so $b \in \Lambda \subseteq H$. But we chose f such that for all $z \in H$, $\operatorname{Re} f(z) < \operatorname{Re} f(x_0) \leq \operatorname{Re} f(b)$ which is a contradiction. $\square$

Corollary 116. *If $K \subseteq X$ is compact and convex then $K = \overline{\operatorname{co}(\Lambda)}$, where Λ is the set of extremal points in K.* $\square$

8.2 Examples!

The canonical example here is when X is a normed space and $K = B^*$ in the weak-$*$ topology.

Example 117. *Consider L^p for $1 < p \leq \infty$ with q such that $L^p = (L^q)^*$. Then the unit ball in L^p with the weak-$*$ topology the closed unit ball is closed in the weak-$*$ topology, and it is compact there. The extremal points are clearly on the unit sphere.*

Example 118. *Let $X = L^1[0,1]$ and B the closed unit ball. Take $f \in \partial B$, namely $\|f\|_1$. Take $a \in (0,1)$ for which $\int_0^a |f(x)| \,\mathrm{d}\, x = \frac{1}{2}$ and take g to be $2\chi_{[0,a]} \cdot f$ and $h = 2\chi_{[a,1]} \cdot f$. Then $\|g\|_1 = \|h\|_1 = 1$. On the other hand $f = \frac{1}{2}g + \frac{1}{2}h$ and $g \neq f, h \neq f$. Therefore f is not an extremal point of B. So B has no extremal points.*

Corollary 119. *There is no normed space X such that $X^* = L^1[0,1]$.* $\square$

We return to the canonical example of B^* in the weak-$*$ topology. So what does it mean? If $\varphi \in B^*$, namely $\|\varphi\| \leq 1$, then for every neighborhood V of 0 in the weak-$*$ topology

[1] The co denotes the convex closure.

there are extremal functionals $\psi_1, \dots, \psi_N$ with $\|\psi_i\| = 1$, and scalars $\alpha_1, \dots, \alpha_N \in [0,1]$ and $\sum_{i=1}^N \alpha_i = 1$ such that:

$$\varphi - \sum_{i=1}^N \alpha_i \psi_i \in V.$$

If X is also separable, then B^* is metrizable when considering the weak-$*$ topology. Then we can improve the statement and obtain that for every $\varphi \in B^*$ there is a sequence $\{\psi^k\}_{k=1}^\infty$ and $\psi^k = \sum_{i=1}^{N_k} \alpha_i^k \psi_i^k$, with ψ_i^k extremal in B^*, and $\varphi = \lim_{k\to\infty} \psi^k$.

Example 120. *Take $X = C[0,1]$ and X^* is the space of Radon measures in $[0,1]$. The extremal points of B^* are exactly δ_t for $t \in [0,1]$. This means that every measure can be approximated (in the weak-$*$ topology) by linear combinations of such δ_t's.*

But we are in context where B^* is indeed metrizable and compact. Consider the sequence $\sum_{i=1}^N \frac{1}{N}\delta_{j/N}$, it has a convergent subsequence, so the entire sequence converges, and its limit is in fact $\mu = \mathrm{d}x$ the Lebesgue measure. This is because this sequence is just the usual Riemann sum and on $C[0,1]$ Riemann integral is the same as Lebesgue integral.

Suppose that X is a normed space and $x_n \xrightarrow{w} x$. Consider $K = \overline{\mathrm{co}(\{x_n\}_{n=1}^\infty)}$ (weak closure is taken) then K is convex and weakly compact, and therefore it is convex and closed in the normed topology. In particular x is a limit point of $\mathrm{co}(\{x_n\}_{n=1}^\infty)$ in the strong topology.

This means that for every $\varepsilon > 0$ there is an eventually 0-sequence of $\alpha_j \in [0,1]$ such that

$$\left\| x - \sum_{j=1}^\infty \alpha_j x_j \right\|_X < \varepsilon.$$

In particular in $X = L^p$ for $1 < p < \infty$, the weak-$*$ topology is just the weak topology, if $f_j \xrightarrow{w} f \in L^p$, and there is a convex linear combination such that

$$\left\| \sum_{j=1}^\infty \alpha_j f_j - f \right\|_p < \varepsilon.$$

Chapter 9

Distributions in Finite Dimensions

9.1 Convolutions

Definition 121. If u, v are measurable complex valued functions on $\mathbb{R}^n$, the *convolution* of u and v is $u * v(x) = \int_{\mathbb{R}^n} u(x-y)v(y)\,\mathrm{d}y$.

Proposition 122. *Convolution is commutative, namely* $u * v = v * u$.

Proof. Simply observe the following equality

$$u * v = \int_{\mathbb{R}^n} u(x-y)v(y)\,\mathrm{d}y = \int_{\mathbb{R}^n} u(z)v(x-z)\,\mathrm{d}z = v * u.$$

This is true because over $\mathbb{R}^n$ the following holds, $\int_{\mathbb{R}^n} u(y)\,\mathrm{d}y = \int_{\mathbb{R}^n} u(-y)\,\mathrm{d}y$. □

Proposition 123. *For every* $1 \leq p \leq \infty$, $\|u * v\|_p \leq \|u\|_1 \|v\|_p$.

Proof.

$$|u * v(x)| \leq \int_{\mathbb{R}^n} |u(x-y)v(y)|\,\mathrm{d}y \leq \|u\|_1 \|v\|_\infty \implies \|u * v\|_\infty \leq \|u\|_1 \|v\|_\infty.$$

We also have, by Fubini's theorem:

$$\int_{\mathbb{R}^n} |u * v(x)|\,\mathrm{d}x = \int_{\mathbb{R}^n} \left(\int_{\mathbb{R}^n} |u(x-y)v(y)|\,\mathrm{d}y \right) \mathrm{d}x = \int_{\mathbb{R}^n} \left(\int_{\mathbb{R}^n} |u(x-y)|\,\mathrm{d}x \right) |v(y)|\,\mathrm{d}y.$$

Therefore $\|u * v\|_1 = \|u\|_1 \|v\|_1$.

More generally, if $1 < p < \infty$ then we want to approximate $\|v * u\|_p$, and the following holds:

$$\begin{aligned} |u * v(x)| &\leq \int_{\mathbb{R}^n} |u(x-y)v(y)|\,\mathrm{d}y \\ &= \int_{\mathbb{R}^n} |u(x-y)|^{1/q} |u(x-y)|^{1/p} |v(y)|\,\mathrm{d}y \\ &\leq \left(\int_{\mathbb{R}^n} |u(x-y)|\,\mathrm{d}y \right)^{1/q} \left(\int_{\mathbb{R}^n} |u(x-y)||v(y)|^p\,\mathrm{d}y \right)^{1/p} \end{aligned}$$

This implies that $|u * v(x)|^p \le \|u\|_1^{p/q} \int_{\mathbb{R}^n} |u(x-y)||v(y)|^p \,\mathrm{d}y$. We integrate again:

$$\begin{aligned}\int_{\mathbb{R}^n} |u * v(x)|^p \,\mathrm{d}x &\le \int_{\mathbb{R}^n} \left(\int_{\mathbb{R}^n} |u(x-y)||v(y)|^p \,\mathrm{d}y \right) \mathrm{d}x \cdot \|u\|_1^{p/q} \\ &= \|u\|_1^{1+p/q} \int_{\mathbb{R}^n} |v(y)|^p \,\mathrm{d}y = \|u\|_1^p \|v\|_p^p\end{aligned}$$

Therefore $\|u * v\|_p \le \|u\|_1 \|v\|_p$ whenever $1 \le p \le \infty$. $\square$

So if we take $p = 1$ there is a nice symmetry so $\|u * v\|_1 \le \|u\|_1 \|v\|_1$ and therefore convolution turns L^1 into an algebra over $\mathbb{C}$.

Definition 124. Let u be function on $\mathbb{R}^n$, we define $\operatorname{supp} u = \overline{\{x \in \mathbb{R}^n \mid u(x) \neq 0\}}$.

The idea of taking the closure is that we want $\mathbb{R}^n \setminus \operatorname{supp} u$ to be the maximal open set on which u is 0.

Proposition 125. $\operatorname{supp}(u * v) \subseteq \operatorname{supp} u + \operatorname{supp} v$.

Proof. Suppose not, and let x be a counterexample. So there is no $y \in \operatorname{supp} v$ such that $x - y \in \operatorname{supp} u$. So $u(x-y)v(y) \equiv 0$ for all $y \in \mathbb{R}^n$. $\square$

Consider the function $\widetilde{\varphi}$ defined as:

$$\widetilde{\varphi}(x) = \begin{cases} \exp\left(\frac{1}{|x|^2 - 1}\right) & |x| < 1 \\ 0 & |x| \ge 1 \end{cases}$$

This $\widetilde{\varphi}$ is in $C^\infty(\mathbb{R}^n)$ and $\operatorname{supp} \widetilde{\varphi} = \overline{B_1(0)}$. We are led to the following definition.

Definition 126. If $\Omega \subseteq \mathbb{R}^n$ is an open domain and $k \ge 0$ a natural number, $C_0^k(\Omega)$ is the set of all functions u with domain Ω such that u is continuously differentiable up to k, with a compact support.

Now $\widetilde{\varphi}$ as defined above belongs to $C_0^\infty(\mathbb{R}^n)$. Define $\varphi(x) = \frac{\widetilde{\varphi}(x)}{\int_{\mathbb{R}^n} \widetilde{\varphi}(y) \,\mathrm{d}y}$, we have that $\varphi \ge 0$, $\varphi \in C^\infty(\mathbb{R}^n)$, $\operatorname{supp} \varphi = \overline{B_1(0)}$ and $\int_{\mathbb{R}^n} \varphi(x) \,\mathrm{d}x = 1$.

Definition 127. Let $\Omega \subseteq \mathbb{R}^n$ an open domain, then $L^1_{loc}(\Omega)$ is the set of all measurable functions $u\colon \Omega \to \mathbb{C}$, such that for every $K \Subset \Omega$, $|u| \in L^1(K)$. Such u is said to be *locally integrable*.

Definition 128. With $\varphi \in C^\infty(\mathbb{R}^n)$ as defined above and $\varepsilon > 0$, we define $\varphi_\varepsilon(x) = \varepsilon^{-n} \varphi\left(\frac{x}{\varepsilon}\right)$.

Proposition 129. *For* $\varepsilon > 0$, $\operatorname{supp} \varphi_\varepsilon = \overline{B_\varepsilon(0)}$, *and* $\int_{\mathbb{R}^n} \varphi_\varepsilon = 1$. $\square$

Proposition 130. *Let* $u \in L^1_{loc}(\Omega)$ *such that* $\operatorname{supp} u = K \Subset \Omega$ *(namely,* $u \in L^1(\Omega)$*). Then for a sufficiently small* $\varepsilon > 0$, $u_\varepsilon = u * \varphi_\varepsilon$ *satisfies:*

1. $\operatorname{supp} u_\varepsilon \subseteq K_\varepsilon = K + \overline{B_\varepsilon(0)}$, *and*
2. $u_\varepsilon \in C^\infty(\Omega)$.

Proof. Take $0 < 2\varepsilon < d(K, \partial\Omega)$. Consider the extension of u by 0 to $\mathbb{R}^n \setminus K$. Then $u * \varphi_\varepsilon$ is well-defined and its support is a subset of $K_\varepsilon \Subset \Omega$.

To see that $u_\varepsilon \in C^\infty(\Omega)$, we pick a coordinate x_i for $x \in K_\varepsilon$. Consider the following:

$$\frac{u_\varepsilon(x + \Delta x \cdot e_i) - u_\varepsilon(x)}{\Delta x} = \int_{\mathbb{R}^n} u(y) \frac{\varphi_\varepsilon(x + \Delta x \cdot e_i - y) - \varphi_\varepsilon(x - y)}{\Delta x} \, \mathrm{d}y,$$

The integration is really on $K = \operatorname{supp} u$, and not the entire $\mathbb{R}^n$, and taking $|\Delta x| < \varepsilon$, then as $\Delta x \to 0$ we have that

$$\frac{\varphi_\varepsilon(x + \Delta x \cdot e_i - y) - \varphi_\varepsilon(x - y)}{\Delta x} \to \frac{\partial \varphi_\varepsilon}{\partial x_i}(x - y)$$

uniformly for all $y \in K$. This partial derivative is uniformly continuous on K_ε, and so:

$$\lim_{\Delta x \to 0} \frac{u_\varepsilon(x + \Delta x \cdot e_i) - u_\varepsilon(x)}{\Delta x} = \int_{\mathbb{R}^n} u(y) \frac{\partial \varphi_\varepsilon}{\partial x_i}(x - y) \, \mathrm{d}y$$

This implies $\frac{\partial u_\varepsilon}{\partial x_i} = \int_{\mathbb{R}^n} u(y) \frac{\partial \varphi_\varepsilon}{\partial x_i}(x - y) \, \mathrm{d}y$.

We can repeat this process by induction, and we obtain that $D^\alpha u_\varepsilon = u * D^\alpha \varphi_\varepsilon$ for every multi-index α. $\square$

Let us re-examine the definition u_ε. Denote by $z = \frac{x-y}{\varepsilon}$, then we have that:

$$u_\varepsilon(x) = \varepsilon^{-n} \int_{\mathbb{R}^n} u(y)\varphi(z) \, \mathrm{d}y = \varepsilon^{-n} \int_{\mathbb{R}^n} u(x - \varepsilon z)\varphi(z)\varepsilon^n \, \mathrm{d}z = \int_{\mathbb{R}^n} u(x - \varepsilon z)\varphi(z) \, \mathrm{d}z.$$

But we can restrict our integration to $\overline{B_1(0)}$, and so $u_\varepsilon \in C_0^\infty(\Omega)$.

To summarize the intuition, for $x \in K_\varepsilon$ we have that $u_\varepsilon(x)$ is a "rearrangement" of u by the "weights of $\varphi(z) \, \mathrm{d}z$".

Example 131. *Take $K \Subset \Omega$ with $K + \overline{B_{3\varepsilon}(0)} \subseteq \Omega$. Now consider $u = \chi_{K + \overline{B_\varepsilon(0)}}$. This function is indeed in $L^1_{loc}(\Omega)$ with compact support. We calculate:*

$$u_\varepsilon(x) = \int_{K + B_\varepsilon(0)} u(y)\varphi_\varepsilon(x - y) \, \mathrm{d}y = \int_{x - y \in B_\varepsilon(0)} \varphi_\varepsilon(x - y) \, \mathrm{d}y = 1.$$

This example leads us to the following corollary.

Corollary 132. *For every $K \Subset \Omega$ there is $\psi_{K,\varepsilon} \in C_0^\infty(\Omega)$ such that $\psi_{K,\varepsilon}(x) = 1$ for all $x \in K$, and generally $0 \leq \psi_{K,\varepsilon}(x) \leq 1$.* [1]

Proposition 133. *Suppose that u is continuous on Ω (so $u \in C_0(\Omega)$), then $u_\varepsilon(x) \to u(x)$ for all $x \in \Omega$.*

Proof. For every $x \in \Omega$ if ε is small enough, then we have that:

$$\begin{aligned}
u_\varepsilon(x) - u(\varepsilon) &= \int_{B_1(0)} u(x - \varepsilon z)\varphi(z) \, \mathrm{d}z - u(x) \\
&= \int_{B_1(0)} (u(x - \varepsilon z) - u(x))\varphi(z) \, \mathrm{d}z \\
\implies |u_\varepsilon(x) - u(x)| &\leq \int_{B_1(0)} |u(x - \varepsilon z) - u(x)|\varphi(z) \, \mathrm{d}z < \delta
\end{aligned}$$

For some $\delta > 0$ small enough, by uniform continuity of u on Ω. $\square$

[1] Such function is sometimes called a *cutoff function*.

Corollary 134. $C_0^\infty(\Omega)$ *is dense in* $C_0(\Omega)$ *in the topology induced by the* sup *norm.* $\square$

Note that on $C_0(\Omega)$ we can induce the topology from the F-space $C(\Omega)$ as topologized by the semi-norms p_K for $K \Subset \Omega$. These two topologies are not the same, to see why consider $\Omega = (0,1)$ and some $x_j \to 1$, let $\varphi_j(x) = \varphi_\varepsilon(x - x_j)$. In the topology of $C(\Omega)$ this sequence converges to the zero function, as on every compact set we will eventually be 0. But in the $\sup$ norm this is not the case, and the sequence is not convergent at all.

Recall that on $C_0^\infty(\Omega)$ we have a topology inherited from $C^\infty(\Omega)$ where it is defined by the semi-norms $p_{K,\alpha} = \max_{x\in K} |D^\alpha f(x)|$ for all $K \Subset \Omega$ and $\alpha \in \mathbb{N}^n$ for some $n \in \mathbb{N}$.

Proposition 135. $C_0^\infty(\Omega)$ *is dense in* $L^p(\Omega)$ *for all* $1 \le p < \infty$.

Proof. First note we can take $f \in L^p(\Omega)$ such that $\operatorname{supp} f = K \Subset \Omega$. Then we have that $f \in L^1(\Omega)$, so we can define $f_\varepsilon(x) = \int_\Omega f(y)\varphi_\varepsilon(x-y)\,\mathrm{d}\,y$. Now $\|f_\varepsilon\|_p \le \|f\|_p$ so the map $f \mapsto f_\varepsilon$ is a continuous linear operator on $L^p(K)$.

So to prove that $\lim_{\varepsilon\to 0} \|f_\varepsilon - f\|_p = 0$ it suffices to prove this for a dense subset of f's. So now we can assume that $f \in C_0(\Omega)$ in which case $f_\varepsilon \to f$ uniformly when $\varepsilon \to 0$. So certainly $\int_\Omega |f_\varepsilon - f|^p \to 0$. $\square$

9.2 Test Functions

Definition 136. Let $\mathcal{D}(\Omega)$ denote $C_0^\infty(\Omega)$ and we call it the *test functions* on Ω. If $K \Subset \Omega$ we denote by $\mathcal{D}_K(\Omega)$ the subspace of $\mathcal{D}_\Omega$ of all the test functions whose support is included in K.

On $\mathcal{D}_K(\Omega)$ there is one topology making it into an F-space, inherited from $C^\infty(\Omega)$ and defined by $p_{K,\alpha}$ as before.

We know that $C_0^\infty(\Omega)$ inherits a topology on $C^\infty(\Omega)$. But this is not the topology we are interested in. $\mathcal{D}(\Omega)$ is topologized in the following way: $U \subseteq \mathcal{D}(\Omega)$ which is convex, absorbing and balanced and $0 \in U$ is a basic open neighborhood of 0 if $U \cap \mathcal{D}_K(\Omega)$ for every $K \Subset \Omega$.

It is easy to see that this topology is Hausdorff, if $\varphi, \psi \in \mathcal{D}(\Omega)$, there is some $K \Subset \Omega$ such that $\operatorname{supp}\varphi, \operatorname{supp}\psi \subseteq K$ and then in $\mathcal{D}_K(\Omega)$ these two are separated.

9.2.1 Semi-norms

For every $K \Subset \Omega$ and a multi-index α, $p_{\alpha,K}(u) = \max_{x\in K} |D^\alpha u|$ is a continuous semi-norm, trivially. Another example is $p_{\max}(u) = \max_\Omega |u|$, and this one is continuous because if $U^\varepsilon = \{u \in \mathcal{D}(\Omega) \mid |u| < \varepsilon\}$ then $U^\varepsilon \cap \mathcal{D}_K(\Omega) = \{u \in C_0^\infty(\Omega) \mid \operatorname{supp} u \subseteq K, |u| < \varepsilon\}$ for every $K \Subset \Omega$, and this is by definition open in $\mathcal{D}_K(\Omega)$ so U^ε is open in $\mathcal{D}(\Omega)$.

A third example is when $x^k \to x \in \partial\Omega$, we define $p_{\{x^k\}}(u) = \sum_k c_k |u(x^k)|$ where $c_k > 0$ is a continuous semi-norm on $\mathcal{D}(\Omega)$, since when we intersect this with $\mathcal{D}_K(\Omega)$, for $K \Subset \Omega$, only a finite number of summands will appear.

If $f \in L^1_{loc}(\Omega)$ we can define $p_f(u) = \int_\Omega |f(x)u(x)|\,\mathrm{d}\,x$ is a continuous semi-norm. This is because when $u \in \mathcal{D}_K(\Omega)$ then $p_f(u) \le \max_K |u(x)| \int_K |f(x)|\,\mathrm{d}\,x$, which is a defined number since f is locally integrable. If μ is a non-negative measure giving compact subsets of Ω finite measure, then integrating with respect to μ, instead of integrating $f \cdot u$, gives a continuous semi-norm by similar considerations.

Proposition 137. *Let* $\{\varphi_j\}_{j=1}^{\infty} \subseteq \mathcal{D}(\Omega)$*, then* $\varphi_j \to 0$ *in* $\mathcal{D}(\Omega)$ *if and only if the following condition holds: There exists* $K \Subset \Omega$ *such that* $\operatorname{supp} \varphi_j \subseteq K$ *for all* j*, and* $\varphi_j \to 0$ *in* $\mathcal{D}_K(\Omega)$*.*

Proof. Suppose the condition holds, then for every $U \subseteq \mathcal{D}(\Omega)$ basic open set will satisfy $U \cap \mathcal{D}_K(\Omega)$ is open, and since $\varphi_j \to 0$ in $\mathcal{D}_K(\Omega)$ it means that for all but finitely many j, $\varphi_j \in U \cap \mathcal{D}_K(\Omega)$. But this means that almost all the φ_j's are in U.

Conversely assume that $\varphi_j \to 0$ in $\mathcal{D}(\Omega)$. Assume towards contradiction that there exists a subsequence φ_{j_k} and x^{j_k} such that $x^{j_k} \to x \in \partial\Omega$ (so there is no compact support for all the φ_j's), and $\varphi_{j_k}(x^{j_k}) \neq \varnothing$. Consider the continuous semi-norm:

$$p(u) = \sum_{k=1}^{\infty} \frac{|u(x^{j_k})|}{|\varphi_{j_k}(x^{j_k})|}$$

Then $p(\varphi_{j_k}) \geq 1$ which is impossible since we assumed $\varphi_j \to 0$. So there must be a compact K as wanted, but now $\varphi_j \to 0$ in $\mathcal{D}_K(\Omega)$ as wanted. □

Remark. *We can reformulate the condition from the previous proposition as follows:* $\varphi_j \to 0$ *in* $\mathcal{D}(\Omega)$ *if there is some* $K \Subset \Omega$ *such that* $\operatorname{supp} \varphi_j \subseteq K$ *and for every* $\alpha \in \mathbb{N}^n$, $\max_K |D^\alpha \varphi_j| \to 0$*.*

If in the above remark we require instead that for all $K \Subset \Omega$, $\max_K |D^\alpha \varphi_j| \to 0$ (without requiring that there is a uniform support), does that mean that $\varphi_j \to 0$ in $\mathcal{D}(\Omega)$? Not at all, because that would mean exactly that $\varphi_j \to 0$ in $C^\infty(\Omega)$ and we know that this is not the same as converging to 0 in $\mathcal{D}(\Omega)$.

9.3 Distributions (or Generalized Functions)

Definition 138. The space of continuous linear functions on $\mathcal{D}(\Omega)$ will be denoted by $\mathcal{D}'(\Omega)$ and its elements are called *distributions*[2]. The topology on $\mathcal{D}'(\Omega)$ is the weak-$*$ topology.

For example $\delta_{x_0}(\varphi) = \varphi(x_0)$ for $x_0 \in \Omega$ and $\varphi \in \mathcal{D}(\Omega)$. Functionals such as $u_f(\varphi) = \int_\Omega f\varphi \,\mathrm{d}\,x$ for $f \in L^1_{loc}(\Omega)$, or if $\mu \geq 0$ is a measure giving finite values to compact sets, then $u_\mu(\varphi) = \int_\Omega f \,\mathrm{d}\,\mu$ give us more examples.

Remark. *We will use Greek letters such as* φ, ψ *and so on for elements in* $\mathcal{D}(\Omega)$ *and Latin letters such as* u, v *will denote elements of* $\mathcal{D}'(\Omega)$ *and we define the notation:* $\langle u, \varphi \rangle := u(\varphi)$*.*[3]

Theorem 139. *Let* u *be a linear functional on* $\mathcal{D}(\Omega)$*, then* $u \in \mathcal{D}'(\Omega)$ *if and only if* $u \restriction \mathcal{D}_K(\Omega)$ *is continuous for all* $K \Subset \Omega$*.*

Proof. Denote by $B_\varepsilon = \{z \in \mathbb{C} \mid |z| < \varepsilon\}$. Suppose that $u \in \mathcal{D}'(\Omega)$, then for every $\varepsilon > 0$, $U_\varepsilon = u^{-1}(B_\varepsilon)$ is open in $\mathcal{D}(\Omega)$. By definition this means that for every $K \Subset \Omega$, $U_\varepsilon \cap \mathcal{D}_K(\Omega)$ is open. So $u^{-1}(B_\varepsilon) \cap \mathcal{D}_K(\Omega)$ is open for every compact K, which means that u is continuous on every $\mathcal{D}_K(\Omega)$.

In the other direction, assume that $u \restriction \mathcal{D}_K(\Omega)$ is continuous for every $K \Subset \Omega$, then the intersection $U_\varepsilon \cap \mathcal{D}_K(\Omega) = u^{-1}(B_\varepsilon) \cap \mathcal{D}_K(\Omega)$ is open by continuity, and therefore U_ε is open as wanted. □

[2]In some places distributions are called *generalized functions*

[3]This is called *pairing*.

Corollary 140. *$u \in \mathcal{D}'(\Omega)$ if and only if for every $K \Subset \Omega$ there are $c_K > 0$ and $k \in \mathbb{N}$ such that for every $\varphi \in \mathcal{D}_K(\Omega)$, $|u(\varphi)| \le c_K \max_{|\alpha|\le k} \max_{x\in K} |D^\alpha \varphi(x)|$.* □

Proposition 141. *Let u be a linear functional on $\mathcal{D}(\Omega)$. Then $u \in \mathcal{D}'(\Omega)$ if and only if for every $\varphi_j \to 0$ in $\mathcal{D}(\Omega)$, $u(\varphi_j) \to 0$.* □

Proposition 142. *Suppose that there is a sequence $\{u_j\} \subseteq \mathcal{D}'(\Omega)$ such that for all $\varphi \in \mathcal{D}(\Omega)$ the sequence $u_j(\varphi) \to u(\varphi)$ (so u_j converge pointwise to u). Then $u \in \mathcal{D}'(\Omega)$.*

Proof. It is clear that u is a linear functional. Now given $K \Subset \Omega$, $\mathcal{D}_K(\Omega)$ is a complete metric space, so by the uniform boundedness principle, $u_j \to u$ pointwise on $\mathcal{D}_K(\Omega)$ and u is continuous on $\mathcal{D}_K(\Omega)$ and by the previous theorem u is continuous as wanted. □

Example 143. *Take $\Omega = \mathbb{R}$ and consider $u(\varphi) = \sum_{j=1}^\infty e^{e^{e^j}} \varphi(j)$, then φ is a distribution.*

In slightly better generality, let $A\colon \mathcal{D}(\Omega) \to X$ a linear operator into a locally convex space X.

Proposition 144. *A is continuous if and only if $A \restriction \mathcal{D}_K(\Omega)$ is continuous for every $K \Subset \Omega$.*

Proof. To verify continuity it is enough to verify that if $V \subseteq X$ is an open neighborhood of 0, $A^{-1}(A)$ is an open neighborhood of 0 in $\mathcal{D}(\Omega)$. Namely, $A^{-1}(A) \cap \mathcal{D}_K(\Omega)$ is an open neighborhood of 0 for every $K \Subset \Omega$. □

The above is true in particular when taking $X = \mathcal{D}(\Omega)$ itself.

Example 145. *Let $\psi \in C^\infty(\Omega)$ and define $T_\psi(\varphi) = \psi\varphi$ as a linear operator from $\mathcal{D}(\Omega)$ to itself. We claim that T_ψ is continuous. So it is enough to check continuity on every $K \Subset \Omega$. But given any such K, and every $\alpha \in \mathbb{N}^n$:*

$$\max_{x\in K} |D^\alpha(\psi\varphi)| \le C \max_{x\in K, |\beta|\le|\alpha|} |D^\beta \varphi|.$$

Where we take $C = \max_{x\in K,|\beta|\le|\alpha|} |D^\beta \psi|$.

Example 146. *For $1 \le i \le n$ we define $T_i\colon \mathcal{D}(\Omega) \to \mathcal{D}(\Omega)$ by:*

$$T_i(\varphi) = \frac{\partial}{\partial x_i}\varphi.$$

This is continuous because given some $K \Subset \Omega$ we can find $C > 0$ and $k \in \mathbb{N}$ such that:

$$\max_{x\in K} \left|\frac{\partial \varphi}{\partial x_i}\right| \le C \max_{|\alpha|\le k} |D^\alpha \varphi(x)|.$$

We can define T_α for every multi-index $\alpha \in \mathbb{N}^n$, and by induction every T_α is continuous.

Example 147. *We know that the functional defined on $D(\Omega)$ by taking $x^j \to x \in \partial\Omega$ and defining:*

$$\varphi \mapsto \sum_{j=1}^\infty c_j D^{\alpha_j} \varphi(x^j).$$

Suppose now that $\Omega = \mathbb{R}^n$, and consider the operator: $T\varphi(x) = \sum_{k=0}^\infty \varphi(x+ka)$ where $a \in \mathbb{R}^n$ is nonzero. While $T\varphi$ does not have a compact support, it is true that $T\varphi \in C^\infty(\mathbb{R}^n)$. So we can ask whether or not $T\colon \mathcal{D}(\mathbb{R}^n) \to C^\infty(\mathbb{R}^n)$. And this is left as a riddle to the interested reader.

9.4 Operators on Distributions

Proposition 148. *Let $u \in D'(\Omega)$ and $\psi \in C^\infty(\Omega)$. Then ψu defined on $\mathcal{D}(\Omega)$ by:*

$$\psi u(\varphi) = u(\psi\varphi)$$

is continuous. Namely, $\psi u \in \mathcal{D}'(\Omega)$. Similarly, if $\alpha \in \mathbb{N}^n$ is a multi-index, then $D^\alpha u$ defined by $D^\alpha u(\varphi) = (-1)^{|\alpha|}u(D^\alpha\varphi)$ is also in $\mathcal{D}'(\Omega)$. □

The two operators are actually continuous operators on $\mathcal{D}'(\Omega)$.

9.4.1 Examples

Consider $\Omega = \mathbb{R}$ and the function $H(x) = \begin{cases} 0 & x < 0 \\ 1 & x > 0 \end{cases}$ which is in $L^1_{loc}(\mathbb{R})$. Then T_H is continuous, where T_H is defined as follows:

$$T_H(\varphi) = \int_{\mathbb{R}} H\varphi \,\mathrm{d}x = \int_0^\infty \varphi(x)\,\mathrm{d}x.$$

Now if we want to differentiate T_H, namely

$$\frac{\mathrm{d}}{\mathrm{d}x}T_H(\varphi) = -T_H\left(\frac{\mathrm{d}\varphi}{\mathrm{d}x}\right) = -\int_0^\infty \varphi'(x)\,\mathrm{d}x = -(\varphi(\infty) - \varphi(0)) = \varphi(0) = \delta_0(\varphi).$$

We differentiate again:

$$\frac{\mathrm{d}}{\mathrm{d}x}\delta_0(\varphi) = -\delta_0(\varphi') = -\varphi'(0).$$

Remark. *The keen eyed reader will notice that $(-1)^{|\alpha|}$ appears in the definition of $D^\alpha u$, and while it does not matter for the continuity, it does seem a bit odd. Why is it there?*

Suppose that $\psi \in C^\infty(\mathbb{R})$, then the definition for differentiating distributions gives us:

$$\frac{\mathrm{d}}{\mathrm{d}x}T_\psi(\varphi) = -T_\psi\left(\frac{\mathrm{d}\varphi}{\mathrm{d}x}\right) = -\int_{\mathbb{R}} \psi(x)\varphi'(x)\,\mathrm{d}x.$$

On the other hand, $\psi' \in C^\infty(\mathbb{R})$ as well, so we can talk about $T_{\psi'}$,

$$T_{\psi'}(\varphi) = \int_{\mathbb{R}} \psi'(x)\varphi(x)\,\mathrm{d}x.$$

We would like that there will be equality between $\frac{\mathrm{d}}{\mathrm{d}x}T_\psi = T_{\psi'}$. And we get that equality exactly because of that added sign.[4]

We remain in the context of $\Omega = \mathbb{R}$. Consider now $f(x) = x \cdot H(x)$ and we calculate:

$$\begin{aligned}\frac{\mathrm{d}}{\mathrm{d}x}T_f(\varphi) &= -\int_R Rf(x)\varphi'(x)\,\mathrm{d}x \\ &= -\int_0^\infty x\varphi'(x)\,\mathrm{d}x \\ &= -([x\varphi(x)]_0^\infty - \int_0^\infty \varphi(x)\,\mathrm{d}x \\ &= \int_0^\infty \varphi(x)\,\mathrm{d}x = T_H(\varphi).\end{aligned}$$

[4]This follows from integration by parts to calculate the equality.

Definition 149. If $H, f \in L^1_{loc}(\Omega)$ we say that H is the *weak derivative* of f when $\frac{\mathrm{d}}{\mathrm{d}x}T_f = T_H$.

9.4.2 Applications to Differential Equations

We work in $\Omega = \mathbb{R}^2_+ = \{\langle x, y\rangle \mid x \in \mathbb{R}, t > 0\}$. We have a function on the line $\Phi_0(x)$ and we are looking for a function $\Phi(x,t)$ defined on $\mathbb{R}^2_+$ such that $\Phi(x,0) = \Phi_0(x)$ and $\Phi_t + \Phi_x = 0$.[5]

We define lines of the form $x(t) = k + t$ for $k \in \mathbb{R}$. Suppose we had Φ and it is at least $C^1(\Omega)$. Then we have that:

$$\frac{\mathrm{d}}{\mathrm{d}t}\Phi(x(t),t) = \Phi_x \frac{\mathrm{d}x}{\mathrm{d}t} + \Phi_t = \Phi_x + \Phi_t = 0.$$

So Φ is constant on each line $x(t)$, so it only depends on the trace k. So we have that:

$$\Phi(x,t) = \widetilde{\Phi}(k) = \Phi_0(k) = \Phi_0(x-t).$$

So we get that $\Phi(x,t) = \Phi_0(x-t)$ is such a solution. To verify that, note that $\Phi_x = \Phi_0'$ and $\Phi_t = -\Phi_0'$ so $\Phi_t + \Phi_x = 0$ and of course that $\Phi(x,0) = \Phi_0(x)$.[6]

If, for example, we take $\Phi_0(x) = \sin x$ we will get $\Phi(x,t) = \sin(x-t)$. But what happens when we want $\Phi_0(x) = H(x)$? We could argue that H is not in C^1 so the question is meaningless. But we did not come here to give excuses, we came here to give solutions!

For the case $\Phi_0(x) = H(x)$ we still have $\Phi(x,t) = \Phi_0(x-t)$. But now we consider these as distributions. If $\Phi(x,t)$ is a distribution on $\mathbb{R}^2_+$ we know how to calculate Φ_x and Φ_t as distributions and we require that $\Phi_x + \Phi_t = 0$ as distributions. So we have that

$$\frac{\partial}{\partial t}\Phi_0(x-t) = -\frac{\mathrm{d}}{\mathrm{d}y}\Phi_0(y)|_{y=x-t}\,, \qquad \frac{\partial}{\partial x}\Phi_0(x-t) = \frac{\mathrm{d}}{\mathrm{d}y}\Phi_0(y)|_{y=x-t}.$$

So we claim that $\Phi(x,t) = H(x-t)$ is still a solution. We verify for $\varphi \in \mathcal{D}(\mathbb{R}^2_+)$:

$$\begin{aligned}
\Phi_x(x,t)(\varphi) &= -\int_\Omega \Phi(x,t)\varphi_x(x,t)\,\mathrm{d}x\,\mathrm{d}t \\
&= -\int_0^\infty \left(\int_t^\infty \varphi_x(x,t)\,\mathrm{d}x\right)\mathrm{d}t \\
&= -\int_0^\infty \varphi(t,t)\,\mathrm{d}t
\end{aligned}$$

Similarly we compute $\Phi_t(x,t)(\varphi)$ and we have that $\Phi_x + \Phi_t = 0$ as distributions.

[5]These are derivatives with respect to the two parameters.

[6]We also need to verify that the solution is unique, but this is beyond the scope of the course.

Chapter 10

Tempered Distributions

We begin by setting conventions for this chapter. We only work with functions in $C^\infty(\mathbb{R}^n)$, we will denote by ∂_j the operator $\frac{\partial}{\partial x_j}$ and D_j will be the operator $\frac{1}{i}\frac{\partial}{\partial x_j}$. Similarly, ∂^α and D^α will be used as before where $\alpha \in \mathbb{N}^n$ is a multi-index.

10.1 The Schwartz Space

Definition 150. *The Schwartz space* $S(\mathbb{R}^n) \subseteq C^\infty(\mathbb{R}^n)$ is the space of rapidly decreasing, smooth functions $f(x)$ such that for every $k, j \in \mathbb{N}$:

$$p_{k,j}(f) = \sup\{(1+|x|)^j |D^\alpha f(x)| \mid x \in \mathbb{R}^n, |\alpha| \leq k\} < \infty,$$

and the semi-norms $p_{k,j}$ define the topology.

The functions in the Schwartz space are those that can be multiplied by any polynomial, and be differentiated in any given order, and remain bounded.[1]

For example $C_0^\infty(\mathbb{R}^n) \subseteq S(\mathbb{R}^n)$, and $\exp(-|x|^2) \in S(\mathbb{R}^n)$ as well. Moreover, if $f \in S(\mathbb{R}^n)$ then for every α, $D^\alpha f \in S(\mathbb{R}^n)$ and for every polynomial p, $pf \in S(\mathbb{R}^n)$ as well.

Proposition 151. *If $r(x)$ is a polynomial, then $f \mapsto rf$ is a continuous operator on $S(\mathbb{R}^n)$.*

Proof. This is clearly a linear operator, so it is enough to verify continuity at 0. Let $p_{k,j}$ be a semi-norm, we want to find a semi-norm $p_{k',j'}$ such that $p_{k,j}(f) < \varepsilon \implies p_{k',j'}(rf) < 1$. But we have the following equality:

$$D^\alpha(rf) = \sum C_{\alpha,\beta} D^\beta r \cdot D^{\alpha-\beta} f.$$

Therefore multiplying by r and all its derivatives give us an increase by at most $|x|^{\deg r}$, so we take $j' = j + \deg r$. Now,

$$(1+|x|)^j |D^\alpha(rf)| \leq C_1(1+|x|)^{\deg r}(1+|x|)^j \sup\{|D^\beta f(x)| \mid x \in \mathbb{R}^n, \beta \leq \alpha\} = C_1 p_{k,j'}(f).$$

Taking supremum over $x \in \mathbb{R}^n$ gives us that $p_{k,j}(rf) \leq C_1 p_{k,j'}(f)$. Taking $\varepsilon = \frac{1}{C_1}$ finishes the proof. □

[1]This condition actually implies that as $x \to \infty$, $f(x) \to 0$.

Proposition 152. *For every multi-index α, $f \mapsto D^\alpha f$ is a continuous operator on $S(\mathbb{R}^n)$.* □

To improve readability (and save this author a lot of typing) we will simply write S for $S(\mathbb{R}^n)$ from here on end.

Proposition 153. *Let $f, g \in S$.*

1. *The pointwise multiplication $fg \in S$.*
2. *For every $h \in \mathbb{R}^n$, $f(x+h) \in S$.*
3. *For every $0 \neq \varepsilon \in \mathbb{R}$, $f(\varepsilon x) \in S$.*

Proposition 154. *S is a metric space, and in fact an F-space.*

Proof. Note that the convergence in S is uniform convergence on $\mathbb{R}^n$, for every order of differentiation and for every multiplication by a polynomial. In particular the semi-norms on $C^\infty(\mathbb{R}^n)$ are continuous on S. So given a Cauchy sequence in S, it has a limit in $C^\infty(\mathbb{R}^n)$, and this limit is in fact in S. □

10.2 Tempered Distributions and Fourier Transforms

Definition 155. The dual space of $S(\mathbb{R}^n)$ will be denoted by $S'(\mathbb{R}^n)$, and it is topologized by the weak-$*$ topology. We call the elements of S' *tempered distributions*.

For example δ_{x_0} is a tempered distribution. But the following proposition gives us a very different tempered distribution.

Proposition 156. *The functional I defined by $I(f) = \int_{\mathbb{R}^n} f(x)\,\mathrm{d}x$ is a tempered distribution.*

Proof. Note that $(1+|x|)^{-n-1}$ is integrable on $\mathbb{R}^n$ by transforming to spherical coordinates. So we can write,

$$|I(f)| = \left| \int_{\mathbb{R}^n} (1+|x|)^{n+1} f(x) (1+|x|)^{-n-1}\,\mathrm{d}x \right| \leq C \sup_{x \in \mathbb{R}^n} (1+|x|)^{n+1} |f(x)|.$$ □

Definition 157. We define the operator $\mathcal{F}$ on S as follows, for every $\xi \in \mathbb{R}^n$ we define:

$$(\mathcal{F}f)(\xi) = (2\pi)^{-n/2} \int_{\mathbb{R}^n} f(x) e^{-i\xi \cdot x}\,\mathrm{d}x,$$

where $\xi \cdot x$ is the usual dot product. We call $\mathcal{F}$ the *Fourier transform*.

It is clear why $\mathcal{F}f$ is well-defined for all $\xi \in \mathbb{R}^n$, and so for all $f \in S$. We will simply denote $\widehat{f}$ for $\mathcal{F}f$.

Proposition 158. *The operator $\mathcal{F}$ is a continuous linear operator on S.*

Proof. Given $1 \leq j \leq n$, consider $D_j\widehat{f}(\xi)$, we expand the definition to have that:

$$D_j\widehat{f}(\xi) = \frac{1}{i} \cdot \frac{\partial}{\partial \xi_j}\widehat{f}(\xi) = \frac{(2\pi)^{-n/2}}{i} \cdot \frac{\partial}{\partial \xi_j}\int_{\mathbb{R}^n} f(x)e^{-i\xi\cdot x}\,\mathrm{d}x,$$

We can move the differentiation inside the integration to obtain the following continuation of the equality,

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

Next we move to calculate $\widehat{D_jf}(\xi)$:

$$\begin{aligned}
\widehat{D_jf}(\xi) &= (2\pi)^{-n/2}\int_{\mathbb{R}^n} D_jf(x)e^{-i\xi\cdot x}\,\mathrm{d}x \\
&= (-i)(2\pi)^{-n/2}\int_{\mathbb{R}^n} \frac{\partial}{\partial x_j}f(x)e^{-i\xi\cdot x}\,\mathrm{d}x \\
&= (-i)(2\pi)^{-n/2}\int_{\mathbb{R}^n} f(x)\left(-\frac{\partial}{\partial x_j}e^{-i\xi\cdot x}\right)\,\mathrm{d}x \\
&= (-i)(2\pi)^{-n/2}\int_{\mathbb{R}^n} f(x)i\xi_je^{-i\xi\cdot x}\,\mathrm{d}x \\
&= (2\pi)^{-n/2}\xi_j\int_{\mathbb{R}^n} f(x)e^{-i\xi\cdot x}\,\mathrm{d}x = \xi_j\widehat{f}(\xi).
\end{aligned}$$

This gives us the following formulas, $D_j\widehat{f}(\xi) = \widehat{-x_jf}(\xi)$ and $\xi_j\widehat{f}(\xi) = \widehat{D_jf}(\xi)$. We can finish the proof now. It is clear that $\mathcal{F}\,f \in S$, since $\widehat{f}(\xi)$ is differentiable in every order, and therefore,

$$D^\alpha\widehat{f}(\xi) = (-1)^{|\alpha|}\int_{\mathbb{R}^n} x^\alpha f(x)e^{-i\xi\cdot x}\,\mathrm{d}x.$$

Therefore $D^\alpha\widehat{f} \in S$ for every $\alpha \in \mathbb{N}^n$ and therefore $\widehat{f} \in C^\infty(\mathbb{R}^n)$. On the other hand,

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

Taking supremum over $\xi \in \mathbb{R}^n$ we get that $\sup_{\xi\in\mathbb{R}^n}|\xi^\alpha\widehat{f}(\xi)| \leq C < \infty$. □

Since $S \subseteq L^1(\mathbb{R}^n)$ we can actually define the Fourier transform for every $f \in L^1(\mathbb{R}^n)$.

Proposition 159. *Let $f \in L^1(\mathbb{R}^n)$. Then $\widehat{f}(\xi)$ is continuous for $\xi \in \mathbb{R}^n$; if $x_jf(x) \in L^1(\mathbb{R}^n)$ for $j \in \{1,\ldots,n\}$ then,*

$$\frac{\partial}{\partial \xi_j}\widehat{f}(\xi) = (2\pi)^{-n/2}\int_{\mathbb{R}^n} -ix_jf(x)e^{-i\xi\cdot x}\,\mathrm{d}x$$

Exists and continuous.

Proof. For $h \in \mathbb{R}^n$ we have that:

$$\begin{aligned}
\widehat{f}(\xi+h) - \widehat{f}(\xi) &= (2\pi)^{-n/2}\int_{\mathbb{R}^n} f(x)\left(e^{-i(\xi+h)\cdot x)} - e^{-i\xi\cdot x}\right)\,\mathrm{d}x \\
&= (2\pi)^{-n/2}\int_{\mathbb{R}^n} f(x)e^{-\xi\cdot x}\left(e^{-ix\cdot h} - 1\right)\,\mathrm{d}x \to 0
\end{aligned}$$

Where the last convergence is due to bounded convergence theorem as $h \to 0$. For the second part, we assume for simplicity that $n = 1$. Now we have that:

$$\begin{aligned}\frac{\widehat{f}(\xi+h) - \widehat{f}(\xi)}{h} &= (2\pi)^{-1/2} \int_{\mathbb{R}} f(x) \frac{e^{i(\xi+h)x} - e^{i\xi x}}{h} \,\mathrm{d}x \\ &= (2\pi)^{-1/2} \int_{\mathbb{R}} f(x) e^{-i\xi x} \frac{e^{-ixh} - 1}{h} \,\mathrm{d}x\end{aligned}$$

Now as $h \to 0$ the integrand approaches to $(-ix)f(x)e^{-i\xi x}$. But for every $\theta \in \mathbb{R}$ we have that $|e^{i\theta} - 1| \le |\theta|$. So $|e^{ixh} - 1| \le |x||h|$ and finally,

$$\left| \frac{e^{-ixh} - 1}{h} \right| \le |x|.$$

So by the bounded convergence theorem $C|x||f(x)|$ is an upper bound for the integrand as wanted. $\square$

10.3 Transform Fourier and Variable Changes

Suppose that $f(x) \in S$ and $f(x+h) \in S$. Then we have that:

$$\widehat{f(x+h)}(\xi) = (2\pi)^{-n/2} \int_{\mathbb{R}^n} f(x+h) e^{-i\xi\cdot x} \,\mathrm{d}x = (2\pi)^{-n/2} \int_{\mathbb{R}^n} f(x) e^{-i\xi\cdot(x-h)} \,\mathrm{d}x = e^{-i\xi\cdot h} \widehat{f}(\xi)$$

Now if $\varepsilon > 0$ then we similarly calculate that:

$$\widehat{f(\varepsilon x)}(\xi) = \varepsilon^{-n} \widehat{f}\left(\frac{1}{\varepsilon}\xi\right).$$

Theorem 160 (Fourier inversion theorem)**.** *Let $f \in S$ such that $\widehat{f} \in S$. then the following equality holds:*

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

Proof. Let $\psi(\xi) \in S$ and consider the following:

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

Now for $\varepsilon > 0$ we consider $\psi(\varepsilon\xi)$ instead of $\psi(\xi)$ and we have that:

$$\int_{\mathbb{R}^n} \widehat{f}(\xi)\psi(\varepsilon\xi)e^{i\xi\cdot x} \,\mathrm{d}\xi = \int_{\mathbb{R}^n} \varepsilon^{-n} \widehat{\psi}\left(\frac{y}{\varepsilon}\right) \,\mathrm{d}y.$$

Suppose now that $\psi(0) = 1$ and $\varepsilon \to 0$, we have in the right hand side,

$$\int_{\mathbb{R}^n} f(x+y)\varepsilon^{-n}\widehat{\psi}\left(\frac{y}{\varepsilon}\right)\,\mathrm{d}y = \int_{\mathbb{R}^n} f(x+\varepsilon z)\widehat{\psi}(z)\,\mathrm{d}z \to f(x)\int_{\mathbb{R}^n}\widehat{\psi}(z)\,\mathrm{d}z.$$

So the equality turns into

$$\int_{\mathbb{R}^n}\widehat{f}(\xi)e^{i\xi\cdot x}\,\mathrm{d}\xi = f(x)\int_{\mathbb{R}^n}\widehat{\psi}(z)\,\mathrm{d}z.$$

Now we take $\psi(\xi) = e^{-\xi^2/2}$ and we calculate $\widehat{\psi}(y)$:

$$\widehat{\psi}(y) = (2\pi)^{-n/2}\int_{\mathbb{R}^n} e^{-\xi^2/2}e^{-i\xi\cdot y}\,\mathrm{d}\xi = \prod_{j=1}^{n}\left((2\pi)^{-1/2}\int_{\mathbb{R}} e^{-\xi_j^2/2}e^{-i\xi_j y_j}\,\mathrm{d}\xi_j\right).$$

So we only need to calculate the single dimensional integral of the Gaussian on $\mathbb{R}$.

$$\begin{aligned}(2\pi)^{-1/2}\int_{\mathbb{R}} e^{-\xi^2/2}e^{-i\xi y}\,\mathrm{d}\xi &= (2\pi)^{-1/2}\int_{\mathbb{R}} e^{-(\xi+iy)^2/2}\,\mathrm{d}\xi e^{-y^2/2}\\ &= (2\pi)^{-1/2}\int_{\mathbb{R}} e^{-\xi^2/2}\,\mathrm{d}\xi e^{-y^2/2} = e^{-y^2/2}.\end{aligned}$$

Therefore $\widehat{e^{-\xi^2/2}}(y) = e^{-y^2/2}$. So going back to the $\mathbb{R}^n$ context, we have that the product becomes 1. So we have the wanted conclusion as wanted. $\square$

Corollary 161. *The operator $\mathcal{F}$ is a bijection.* $\square$

Proposition 162. *Let $\varphi, \psi \in S$. Then:*

1. $\int_{\mathbb{R}^n}\widehat{\varphi}\psi = \int_{\mathbb{R}^n}\varphi\widehat{\psi}$.
2. $\int_{\mathbb{R}^n}\widehat{\psi}\overline{\widehat{\psi}} = \int_{\mathbb{R}^n}\varphi\overline{\psi}$. $\square$

Corollary 163. *If we take $\varphi = \psi$, then we get that $\int_{\mathbb{R}^n}|\widehat{\varphi}|^2 = \int_{\mathbb{R}^n}|\varphi|^2$. It means that $\mathcal{F}$ is an isometry on S as well as induced from $L^2(\mathbb{R}^n)$.* $\square$

Corollary 164. *We can extend $\mathcal{F}$ to an isometric isomorphism to $L^2(\mathbb{R}^n)$.* $\square$

This leads us to the following definition.

Definition 165. If T is an isometric isomorphism on L^2 we say that it is a unitary operator.

10.4 The Extension of the Fourier Transform

We will write $\mathcal{F}$ for the extension of the Fourier transform to L^2, although $\widehat{f}$ is not necessarily defined on functions in L^2.

Proposition 166. *Let $f \in L^1(\mathbb{R}^n) \cap L^2(\mathbb{R}^n)$. Then $\mathcal{F}f(\xi) = (2\pi)^{-n/2}\int_{\mathbb{R}^n}(x)e^{-i\xi\cdot x}\,\mathrm{d}x$ almost everywhere when the integral is well-defined.*

This means that we can find a continuous representative for $\mathcal{F}\, f$ if we can calculate the integral, at least in the case that $f \in L^1(\mathbb{R}^n)$ as well.

Proof. First we deal with the case that $\operatorname{supp} f \Subset \mathbb{R}^n$, and without loss of generality $\operatorname{supp} f \Subset B_r(0)$ for some $r > 0$. Then we can find a sequence $\{\varphi_j\}_{j=1}^{\infty} \subseteq C_0^{\infty}(B_r(0))$ such that $\varphi_j \to f$ in L^2. From Hölder's inequality it follows that $\varphi_j \to f$ in L^1 as well.

Because $\varphi_j \to f$ in L^2 it follows that $\mathcal{F}\,\varphi_j \to \mathcal{F}\, f$, but we know how $\mathcal{F}\,\varphi_j$ looks like,

$$\mathcal{F}\,\varphi = (2\pi)^{-n/2} \int_{\mathbb{R}^n} \varphi_j(x) e^{-i\xi\cdot x}\,\mathrm{d}\,x.$$

But now, by the convergence in L^1 we have that

$$(2\pi)^{-n/2} \int_{\mathbb{R}^n} \varphi_j(x) e^{-i\xi\cdot x}\,\mathrm{d}\,x \to (2\pi)^{-n/2} \int_{\mathbb{R}^n} f(x) e^{-i\xi\cdot x}\,\mathrm{d}\,x.$$

We calculate the approximation to see that:

$$\left| \int_{\mathbb{R}^n} (\varphi_j(x) - f(x)) e^{-i\xi\cdot x}\,\mathrm{d}\,x \right| \le \int_{\mathbb{R}^n} |\varphi_j(x) - f(x)|\,\mathrm{d}\,x \to 0.$$

Therefore we have that $\mathcal{F}\, f = (2\pi)^{-n/2} \int_{\mathbb{R}^n} f(x) e^{-i\xi\cdot x}\,\mathrm{d}\,x$ almost everywhere, thus the proof is complete for the case where the integral is defined and f has a compact support.

Now we deal in the general case. Define $f_k = f \cdot \chi_{\{x||x|\le k\}}$, then $f_k \to f$ in L^2 and therefore $\mathcal{F}\, f_k \to \mathcal{F}\, f$ as well (again in L^2). By the previous case, we already know that:

$$\mathcal{F}\, f_k = (2\pi)^{-n/2} \int_{\mathbb{R}^n} f_k(x) e^{-i\xi\cdot x}\,\mathrm{d}\,x$$

But we also have that $f_k \to f$ in L^1 and for every $\xi \in \mathbb{R}^n$ we have that

$$\mathcal{F}\, f(\xi) \to (2\pi)^{-n/2} \int_{\mathbb{R}^n} f(x) e^{-i\xi\cdot x}\,\mathrm{d}\,x$$

(and this convergence is uniform in ξ) and the result follows. □

Corollary 167. *We can extend the Fourier transform to $L^1(\mathbb{R}^n)$ given by:*

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

Proposition 168. *For every $f \in L^1(\mathbb{R}^n)$:*

1. *$\widehat{f}(\xi)$ is continuous and bounded.*

2. *(Riemann-Lebesgue Lemma)* $\lim_{R\to 0} \sup_{|\xi|>R} |\widehat{f}(\xi)| = 0$.

Proof. Let $\varepsilon > 0$ and $\varphi \in S$ such that $\|\varphi - f\|_1 < \varepsilon$. Then we have that

$$|\widehat{\varphi}(\xi) - \widehat{f}(\xi)| \le (2\pi)^{-n/2} \int_{\mathbb{R}^n} |\varphi(x) - f(x)|\,\mathrm{d}\,x < \varepsilon.$$

On the other hand, $\widehat{\varphi} \in S$, so $\widehat{f}$ is a limit of a uniformly convergent sequence of sequences which are continuous and bounded, so it is also continuous and bounded.

For the second part, take φ like in the previous part, and $R \gg 1$ such that $\sup_{|\xi|>R} |\widehat{\varphi}(\xi)| < \varepsilon$, then we have that

$$|\widehat{f}(\xi)| \le |\widehat{\varphi}(\xi) - \widehat{f}(\xi)| + |\widehat{\varphi}(\xi)| < 2\varepsilon.$$ □

We have already proved the Riemann-Lebesgue Lemma in the context of weak topologies. We have seen that for $f \in L^2(S^1)$ it holds that $\int_{S^1} f(x)e^{-inx} \, \mathrm{d}\, x \to 0$ when $n \to 0$. This is of course true for $f \in L^1$ as well, by approximations.

Proposition 169. *Let $f \in L^2(\mathbb{R}^n)$ and define $\widehat{f}_k(\xi) = (2\pi)^{-n/2} \int_{|x| \le k} f(x)e^{-i\xi \cdot x} \, \mathrm{d}\, x$. Then in L^2 we have that $\widehat{f}_k(\xi) = \mathcal{F} f(\xi)$ in L^2.*

Proof. We know that $f_k \to f$ in L^2, so $\mathcal{F} f_k \to \mathcal{F} f$ in L^2 as well. On the other hand, we know that

$$\mathcal{F} f_k(\xi) = (2\pi)^{-n/2} \int_{\mathbb{R}^n} f_k(x)e^{-i\xi \cdot x} \, \mathrm{d}\, x. \qquad \square$$

Definition 170. In the above situation, where $f \in L^2$ we sometimes write that:

$$\mathcal{F} f(\xi) = \mathrm{l.i.m.}(2\pi)^{-n/2} \int_{\mathbb{R}^n} f(x)e^{-i\xi \cdot x} \, \mathrm{d}\, x$$

where the $\mathrm{l.i.m.}$ is an acronym for "limit in mean".

Recall the inversion theorem. We saw that for $f \in S$ it holds by the inversion theorem that $f(-x) = (2\pi)^{-n/2} \int_{\mathbb{R}^n} \widehat{f}(\xi)e^{-i\xi \cdot x} \, \mathrm{d}\, x$. So if we define $Rg(x) = g(-x)$, then $\mathcal{F}^2 = R$ so $\mathcal{F}^4 = R^2 = I$.

Corollary 171. *By a unitary continuation we get that $\mathcal{F}^2 = R$ on $L^2(\mathbb{R}^n)$.* $\square$

10.5 Extending the Inversion Theorem

Theorem 172. *Let $f \in L^1(\mathbb{R}^n)$ such that $\widehat{f} \in L^1(\mathbb{R}^n)$. Then the inversion theorem hold for f. Namely,*

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

Proof. First note that $\int \widehat{f} g = \int f \widehat{g}$ when $f, g \in L^1$, since for $g \in L^1$ we have $\widehat{g} \in L^\infty$ and so on, now by Fubini's theorem the equality follows.

Now define $f_0(x) = (2\pi)^{-n/2} \int_{\mathbb{R}^n} \widehat{f}(\xi)e^{i\xi \cdot x} \, \mathrm{d}\, \xi$. Take $\varphi \in S$ and consider the following equation:

$$\begin{aligned}
\mathbb{R}^n f_0(x)\widehat{\varphi}(x) \, \mathrm{d}\, x &= (2\pi)^{-n/2} \int_{\mathbb{R}^n} \left(\int_{\mathbb{R}^n} \widehat{f}(\xi)e^{i\xi \cdot x} \, \mathrm{d}\, \xi \right) \widehat{\varphi}(x) \, \mathrm{d}\, x \\
&= \int_{\mathbb{R}^n} \widehat{f}(\xi) \left((2\pi)^{-n/2} \int_{\mathbb{R}^n} \widehat{\varphi}(x)e^{i\xi \cdot x} \, \mathrm{d}\, x \right) \mathrm{d}\, \xi \\
&= \int_{\mathbb{R}^n} \widehat{f}(\xi)\varphi(\xi) \, \mathrm{d}\, \xi = \int_{\mathbb{R}^n} f(x)\widehat{\varphi}(x) \, \mathrm{d}\, x.
\end{aligned}$$

Now we use the fact that $\int_{\mathbb{R}^n} (f_0(x) - f(x))\widehat{\varphi}(x) \, \mathrm{d}\, x = 0$ for every $\varphi \in S$. And therefore $f_0(x) = f(x)$ almost everywhere. $\square$

What we have shown, in conclusion, is that whenever the formulas are well-defined we can use the formulas even when f is not in S.

10.6 Fourier Transform and Convolutions

Proposition 173. *If $f, g \in L^1(\mathbb{R}^n)$, then $\widehat{f * g}(\xi) = (2\pi)^{n/2}\widehat{f}(\xi)\widehat{g}(\xi)$.*

Proof. Recall that $f * g(x) = \int_{\mathbb{R}^n} f(y)g(x-y)\,\mathrm{d}y$. Therefore we have that

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

Corollary 174. *Let $X = \mathcal{F}(L^1(\mathbb{R}^n)))$, the image of the Fourier transform on L^1. Then $X \subseteq C(\mathbb{R}^n)$, and its elements rapidly decreasing, and X is closed under multiplication.*

For example when $n = 1$, take $\varphi_\alpha(\xi) = (1+|\xi|)^{-\alpha}$ for $\alpha > 0$. If $\alpha > \frac{1}{2}$ then $\varphi_\alpha \in L^2(\mathbb{R})$ so for some $g_\alpha \in L^2(\mathbb{R})$ we have that $\widehat{g_\alpha} = \varphi_\alpha$. And if $\alpha > 1$ we can also write it as $\varphi_\alpha(\xi) = \varphi_{\alpha/2}(\xi)\varphi_{\alpha/2}(\xi)$ so it is the pointwise product of two functions in L^2.

10.7 Tempered Distributions Strike Back!

Recall that $S'(\mathbb{R}^n)$ is the dual space of $S(\mathbb{R}^n)$. We observe that $S'(\mathbb{R}^n) \subseteq \mathcal{D}'(\mathbb{R}^n)$, namely every continuous linear functional on $S(\mathbb{R}^n)$ will be continuous when restricted to $\mathcal{D}(\mathbb{R}^n)$. On the other hand, $\mathcal{D}'(\mathbb{R}^n)$ is *much* larger. If $f \in L^1_{loc}(\mathbb{R}^n)$ then $\varphi \mapsto \int_{\mathbb{R}^n} f\varphi$ is continuous on $\mathcal{D}(\mathbb{R}^n)$; on the other hand, this functional need not be extendable to $S(\mathbb{R}^n)$.

Recall that $T_f(\varphi) = \int_{\mathbb{R}^n} f\varphi$ is a linear functional if the integral is always defined for $\varphi \in S$.

Remark.
1. *For T_f to be defined it is sufficient that f has a polynomial growth, namely $|f(x)| \leq C(1+|x|)^k$ for some constants C, k.*
2. *If $\varphi \in S$ and $\psi \in C^\infty(\mathbb{R}^n)$ then $\psi\varphi \in S$ and $D^\alpha\psi$ has polynomial growth.*

Definition 175. For every $T \in S'(\mathbb{R}^n)$ and every multi-index α we define

$$D^\alpha T(\varphi) = T\left((-1)^{|\alpha|}D^\alpha\varphi\right).$$

For every $\psi \in C^\infty(\mathbb{R}^n)$ that has polynomial growth (as do its derivatives), and $\varphi \in S, T \in S'$ we define

$$\psi T(\varphi) = T(\psi\varphi).$$

Note that as $D^\alpha\colon S \to S$ is a continuous operator, $D^\alpha T \in S'(\mathbb{R}^n)$ as well. Similarly as $\varphi \mapsto \psi\varphi$ is continuous, $\psi T \in S'$ as well.

What else can we find in $S'(\mathbb{R}^n)$? Every $u \in \mathcal{D}'(\mathbb{R}^n)$ which has compact support, for example $D^\alpha\delta_{x_0}$ or other Borel measures with compact support.

10.8 Fourier Transform of Tempered Distributions

Suppose that $T \in S'(\mathbb{R}^n)$ and $\varphi \in S$. We define $\widehat{T}(\varphi) = T(\widehat{\varphi})$. So we have that $\widehat{T} \in S'$ and if f is suitable (e.g. $f \in L^1$), then $\widehat{T_f} = T_{\widehat{f}}$.

Example 176. *We calculate $\widehat{\delta_{x_0}}$:*

$$\begin{aligned}\widehat{\delta_{x_0}}(\varphi) &= \delta_{x_0}(\widehat{\varphi}) \\ &= (2\pi)^{-n/2} \int_{\mathbb{R}^n} \varphi(\xi) e^{-i\xi\cdot x_0} \,\mathrm{d}\xi \\ &= (2\pi)^{-n/2} T_{e^{-i\xi\cdot x_0}}(\varphi)\end{aligned}$$

So we can write[2] *that* $\widehat{\delta_{x_0}} = (2\pi)^{-n/2} e^{-i\xi\cdot x_0}$.

It follows from the above example that for $x_0 = 0$, $\widehat{\delta_0} = (2\pi)^{-n/2} T_1$.

Proposition 177. $\widehat{D_j T} = \xi_j \widehat{T}$.

Proof.

$$\begin{aligned}\widehat{D_j T}(\varphi) &= D_j T(\widehat{\varphi}) \\ &= -T(D_j \widehat{\varphi}(x)) \\ &= -T\left(D_j\left((2\pi)^{-n/2} \int_{\mathbb{R}^n} \varphi(\xi) e^{-i\xi\cdot x} \,\mathrm{d}\xi\right)\right) \\ &= -T\left(\frac{1}{i}(2\pi)^{-n/2} \int_{\mathbb{R}^n} \varphi(\xi)(-i\xi_j) e^{-i\xi\cdot x} \,\mathrm{d}\xi\right) \\ &= T\left((2\pi)^{-n/2} \int_{\mathbb{R}^n} \varphi(\xi)\xi_j e^{-i\xi\cdot x} \,\mathrm{d}\xi\right) \\ &= T(\widehat{x_j\varphi}) = \widehat{T}(\xi_j\varphi) = \xi_j \widehat{T}(\varphi).\end{aligned}$$

□

The above generalizes to $\widehat{D^\alpha T} = \xi^\alpha \widehat{T}$.

Proposition 178. *Suppose that $f \in C^1(\mathbb{R}^n)$ and $f, \frac{\partial f}{\partial x_j}$ have polynomial growth, then*

$$\frac{\partial}{\partial x_j} T_f = T_{\frac{\partial}{\partial x_j} f}. \quad \square$$

10.9 Heat Equations

Suppose that $u(x,t)$ satisfies that for $t \geq 0$, $\partial_t u = \partial_x^2 u$ with initial condition $u(x,0) = u_0(x) \in S(\mathbb{R})$. For every $t \geq 0$ we consider the Fourier transform of $u(x,t)$ with respect to x. Namely,

$$\widehat{u}(\xi,t) = (2\pi)^{-1/2} \int_{\mathbb{R}} u(x,t) e^{-i\xi x} \,\mathrm{d}x.$$

[2] This is a sloppy notation which is mathematically incorrect, but scientifically concise.

We differentiate with respect to t to obtain:

$$\begin{aligned}\frac{\partial}{\partial t}\widehat{u}(\xi,t) &= (2\pi)^{-1/2}\int_{\mathbb{R}} \frac{\partial}{\partial t}u(x,t)e^{-i\xi x}\,\mathrm{d}x \\ &= (2\pi)^{-1/2}\int_{\mathbb{R}} \frac{\partial^2}{\partial x^2}u(x,t)e^{-i\xi x}\,\mathrm{d}x \\ &= \widehat{frac\partial^2\partial x^2 u(x,t)}(\xi,t) = -\xi^2\widehat{u}(\xi,t).\end{aligned}$$

Therefore we have a differential equation $\frac{\partial}{\partial t}\widehat{u}(\xi,t) = -\xi^2\widehat{u}(\xi,t)$ with an initial condition $\widehat{u}(\xi,0) = \widehat{u_0}(\xi) \in S(\mathbb{R})$. The unique solution is $\widehat{u}(\xi,t) = e^{-t\xi^2}\widehat{u_0}(\xi)$.

Now suppose that $G(x,t) \in S(\mathbb{R})$ for $t > 0$ such that $\widehat{G}(\xi,t) = e^{-t\xi^2}$. Then we have

$$\widehat{G(\cdot,t) * u_0}(\cdot) = (2\pi)^{-1/2}\widehat{G}(\xi,t)\widehat{u_0}(\xi) = (2\pi)^{-1/2}e^{-t\xi^2}\widehat{u_0}(\xi) = (2\pi)^{-1/2}\widehat{u}(\xi,t).$$

In conclusion $u(x,t) = (2\pi)^{-1/2}G(\cdot,t) * u_0(\cdot)$ or in other words,

$$u(x,t) = (2\pi)^{-1/2}\int_{\mathbb{R}} G(x-y,t)u_0(y)\,\mathrm{d}y.$$

10.10 Final Remarks on Tempered Distributions and the Fourier Transform

The Fourier transform $\mathcal{F}$ can be extended to an isometric isomorphism of $L^2(\mathbb{R}^n)$ such that whenever the following integrals are defined equality holds,

$$\int_{\mathbb{R}^n} |\,\mathcal{F}\,f(\xi)|^2\,\mathrm{d}\xi = \int_{\mathbb{R}^n} |f(x)|^2\,\mathrm{d}x. \qquad \text{(Parseval's identity)}$$

Moreover, we have seen that $\int_{\mathbb{R}^n}\varphi\overline{\psi} = \int_{\mathbb{R}^n}\overline{\varphi}\psi$, at least when $\varphi,\psi \in S(\mathbb{R}^n)$. If we have $f,g \in L^2(\mathbb{R}^n)$ we can find sequences $\{\varphi_k\}_{k=1}^{\infty}$ and $\{\psi_k\}_{k=1}^{\infty}$ such that in $L^2(\mathbb{R}^n)$, $\varphi_k \to f$ and $\psi_k \to \psi$. Then we have that $\mathcal{F}\,\varpi_f \to \mathcal{F}\,f$ and $\mathcal{F}\,\psi_k \to \mathcal{F}\,g$. And by the identity we know that $\int_{\mathbb{R}^n}\varphi_k\overline{\psi_k} = \int_{\mathbb{R}^n}\widehat{\varphi_k}\overline{\widehat{\psi_k}}$. Combining all these we obtain that for every $f,g \in L^2$ it holds that,

$$\int_{\mathbb{R}^n} \mathcal{F}\,f(\xi)\overline{\mathcal{F}\,g(\xi)}\,\mathrm{d}\xi = \int_{\mathbb{R}^n} f(x)\overline{g(x)}\,\mathrm{d}x. \qquad \text{(Plancherel's theorem)}$$

While $\mathcal{F}\colon S(\mathbb{R}^n) \to S(\mathbb{R}^n)$ and $\mathcal{F}\colon L^2(\mathbb{R}^n) \to L^2(\mathbb{R}^n)$ were continuous, we also extended $\mathcal{F}$ to $S'(\mathbb{R}^n)$ by the formula $\widehat{T}(\varphi) = T(\widehat{\varphi})$, and this extension is also continuous when we consider $S'(\mathbb{R}^n)$ with the weak-$*$ topology.

Recall the heat equation $u_t = u_{xx}$ and $u(x,0) = u_0(x) \in S(\mathbb{R})$. We saw that $\partial_t\widehat{u}(\xi,t) = -\xi^2\widehat{u}(\xi,t)$. So we got that $\widehat{u}(\xi,t) = e^{-t\xi^2}\widehat{u_0}(\xi)$. Suppose now that $e^{-t\xi^2} = \widehat{G(x,t)}(\xi)$, then $\widehat{G(x,t) * u_0}(x) = (2\pi)^{-1/2}e^{-t\xi^2}\widehat{u_0}(\xi)$. Or in other words, $\widehat{\frac{1}{\sqrt{2\pi}}G * u_0} = \widehat{u}(\xi,t)$. We also saw that in that case $u(x,t) = (2\pi)^{-1/2}\int_{\mathbb{R}} G(x-y,t)u_0(y)\,\mathrm{d}y$.

Since $e^{-t\xi^2} \in S(\mathbb{R})$, for $t > 0$ we get that $G(x,t) \in S(\mathbb{R})$. So we have that,

$$\begin{aligned} G(x,t) &= (2\pi)^{-1/2} \int_{\mathbb{R}} e^{-t\xi^2} e^{-i\xi x} \,\mathrm{d}\xi \\ &= (2\pi)^{-1/2} \int_{\mathbb{R}} e^{-\frac{1}{2}y^2} e^{-i\frac{y}{\sqrt{2t}}x} \frac{1}{\sqrt{2t}} \,\mathrm{d}y \\ &= \frac{1}{\sqrt{2t}} e^{-\frac{1}{2}\frac{x^2}{2t}}. \end{aligned}$$

Therefore we get that,

$$u(x,t) = \frac{1}{\sqrt{4\pi t}} \int_{\mathbb{R}} G(x-y,t)\sqrt{2t}u_0(y) \,\mathrm{d}y = \frac{1}{\sqrt{4\pi t}} \int_{\mathbb{R}} e^{-\frac{|x-y|^2}{4t}} u_0(y) \,\mathrm{d}y.$$

So if we have that $u_0(y) \in S(\mathbb{R})$, then $\widehat{u_0}(\xi) \in S(\mathbb{R})$ also. In particular, as $t \to 0$, $e^{-t\xi^2}\widehat{u_0}(\xi) \to \widehat{u_0}(x_0)$ in $S(\mathbb{R})$.

In conclusion, if $u_0 \in S$, then $\frac{1}{\sqrt{4\pi t}} \int_{\mathbb{R}} e^{-\frac{|x-y|^2}{4t}} u_0(y) \,\mathrm{d}\, y \to u_0(x)$ in S, as $t \to 0$. Additionally, $e^{-t\xi^2} \to T_1$ in $S'(\mathbb{R})$ (recall that $\int_{\mathbb{R}} e^{-t\xi^2}\varphi(\xi) \,\mathrm{d}\,\xi \to \int_{\mathbb{R}} \varphi(\xi) \,\mathrm{d}\,\xi$). So we get that $\widehat{e^{-t\xi^2}} \to \widehat{T_1}$, or in other words $G(x,t) \to \widehat{T_1}$ in $S'(\mathbb{R})$. Recall that

$$\widehat{T_1}(\varphi) = T_1(\widehat{\varphi}) = \int \widehat{\varphi}(\xi) \,\mathrm{d}\xi = (2\pi)^{1/2}\varphi(0),$$

so $\widehat{T_1} = (2\pi)^{1/2}\delta_0$. In other words, $G(x,t) \to (2\pi)^{1/2}$ in $S'(\mathbb{R})$. In particular $\{G(x,t)\}_{t>0}$ is an approximation for δ_0.

Chapter 11

Compact Operators on Banach Spaces

Suppose that X is a Banach space, recall that $B(X)$ is the space of all bounded linear operators $T\colon X \to X$. We agree that the norms will not bear subscript, since we only work in X and it will be clear when something is an operator norm or a norm in X. We will use T to denote an operator in $B(X)$ without mentioning it often.

Proposition 179. *If $T \in B(X)$ such that $T(X) = X$ and $\ker T = \{0\}$ (so T is a bijection), then $T^{-1} \in B(X)$ as well.* □

11.1 The Spectrum of an Operator

Definition 180. Let $\lambda \in \mathbb{C}$ such that $T - \lambda I\colon X \to X$ is a bijective operator (so there is $(T - \lambda I)^{-1} \in B(X)$). The set $\{\lambda \in \mathbb{C} \mid (T - \lambda I)^{-1} \in B(X)\}$ is called the *resolvent set* of T.

Proposition 181. *Suppose that $\lambda \in \mathbb{C}$ and $|\lambda| > \|T\|$, then λ is in the resolvent set of T.*

Proof. Note that $T - \lambda I) = \lambda(\lambda^{-1}T - I)$, and now $\|\lambda^{-1}T\| < 1$. By Neumann's series there is an inverse operator, $(\lambda^{-1}T - I)^{-1}$. □

Definition 182. The *spectrum* of T is the complement, in $\mathbb{C}$, of the resolvent set. We denote the specturm by $\sigma(T)$.

From the previous proposition we know that $\lambda \in \sigma(T)$ implies that $|\lambda| \leq \|T\|$. So it is a bounded set.

Example 183. *Suppose that $X = C[0,1]$. Consider T defined by $(Tf)(x) = xf(x)$. We know that*

$$\|Tf\| \leq 1 \cdot \|f\| = \max\{|f(x)| \mid x \in [0,1]\} \implies \|T\| \leq 1.$$

Let $\lambda \in \mathbb{C}$ and we try to solve the equation $(T - \lambda I)f = g$. Namely $xf(x) - \lambda f(x) = g(x)$. If $\lambda \notin [0,1]$ we have that $f(x) = \frac{g(x)}{x-\lambda}$ which is a continuous function. So $\sigma(T) \subseteq [0,1]$. What happens if $\lambda_0 \in [0,1]$?

So for $x \in [0,1]$ we have $(x - \lambda_0)f(x) = g(x)$, and therefore $g(\lambda_0) = 0$. So we cannot take any g to be the inverse. It follows, if so, that $\sigma(T) = [0,1]$.

We saw that $T - \lambda_0 I$ is not surjective, but is it injective? Suppose that $f_0 \in \ker(T - \lambda_0 I)$. Therefore for every $x \in [0,1]$ satisfies $xf_0(x) = \lambda_0 f_0(x)$ so $f(x) = 0$ for all $x \neq \lambda_0$ and by continuity $f = 0$ everywhere. So $T - \lambda I$ is indeed injective.

Definition 184. We say that $\lambda \in \sigma(T)$ is an *eigenvalue* of T such that for some $h \neq 0$ we have that $Th = \lambda h$.

The example above shows that we can have T which has a nontrivial spectrum, but only 0 is an eigenvalue of T.

11.2 Compact Operators

Definition 185. We say that $T \in B(X)$ is a *compact operator*[1] if for every bounded sequence $\{x_n\}_{n=1}^{\infty}$ there is a convergent subsequence of $\{Tx_n\}_{n=1}^{\infty}$.

Proposition 186. *T is a compact operator if and only if $\overline{T(B)}$ is compact, where B is the unit ball.*[2]

Proof. Take $\{z_n\}_{n=1}^{\infty} \subseteq \overline{T(B)}$, then for every n there is some $y_n \in B$ such that $\|Ty_n - z_n\| < \frac{1}{n}$, so there is some subsequence $Ty_{n'}$ which is convergent, and it is clear that $z_{n'}$ is convergent. The other direction is trivial. $\square$

Proposition 187. *The requirement that $T \in B(X)$ is unnecessary. Namely, every compact operator is bounded.*

Proof. Suppose that T is not bounded and take $\{x_n\}_{n=1}^{\infty}$ on the unit sphere such that $\|Tx_n\| \to \infty$. So there is no convergent subsequence of $\{Tx_n\}_{n=1}^{\infty}$ and therefore T is not compact. $\square$

We will denote by $K(X)$ the subspace of $B(X)$ of the compact operators on X.

Proposition 188. *$K(X)$ is a closed subspace of $B(X)$.*

Proof. Let $\{T_n\}_{n=1}^{\infty} \subseteq K(X)$ a convergent sequence in $B(X)$ with T its limit there. For every $\varepsilon > 0$ there is some $N \in \mathbb{N}$ such that $\|T_n - T\| < \varepsilon$ for all $n > N$.

Let $\{x_k\}_{k=1}^{\infty} \subseteq X$ a bounded sequence. By diagonalization there is a subsequence $x_{k'}$ such that $\lim_{k' \to \infty} T_j x_{k'}$ exists for all j. We will now show that $Tx_{k'}$ is convergent. Fix $\varepsilon > 0$ and N as in the definition, and let $n > N$.

$$\begin{aligned}
|Tx_{k'} - Tx_{l'}\| &\leq \|(T - T_n)x_{k'}\| + \|T_n(x_{k'} - x_{l'})\| + \|(T - T_n)x_{l'}\| \\
&\leq \frac{2\varepsilon}{3} \sup_{k'} \|x_{k'}\| + \|T_n(x_{k'} - x_{l'})\|
\end{aligned}$$

And this goes to 0 as wanted, so the sequence converges. $\square$

Proposition 189. *$K(X)$ is a two-sided ideal. Namely, if $T \in K(X)$ and $S \in B(X)$ then $ST, TS \in K(X)$.* $\square$

Example 190. *Consider $C[0,1]$ again and look at $Tf(x) = \int_0^x f(s)\,\mathrm{d}s$. We know that for all x, $|Tf(x)| \leq \|f\|$ so $\|T\| \leq 1$. Suppose that $\lambda \neq 0$, is it an eigenvalue? Suppose that $Tf = \lambda f$, then for every x*

$$\lambda f(x) = \int_0^x f(s)\,\mathrm{d}s \implies \lambda f'(x) = f(x),\ f(0) = 0.$$

[1] Also called *completely continuous*

[2] This is the common definition.

And we know there is no such f other than $f(x) = 0$.

Is $\lambda \neq 0$ in the spectrum of T? We need to check if $T - \lambda I$ is surjective. Suppose $h \in C[0,1]$, we ask if there is some f such that $\int_0^x f(s)\,\mathrm{d}\,s - \lambda f(x) = h(x)$? By solving the differential equation, the existence and uniqueness of the solution guarantee that indeed $\lambda \notin \sigma(T)$. Therefore $\sigma(T) = \{0\}$.

Finally, is T compact? Suppose that $\{f_n\}_{n=1}^{\infty}$ is a sequence in the closed unit ball of $C[0,1]$, then $\|Tf_n\| \leq 1$ as well. But we even have that,

$$|Tf_n(x_2) - Tf_n(x_1)| = \left| \int_{x_1}^{x_2} f_n(s)\,\mathrm{d}\,s \right| \leq |x_2 - x_1|.$$

Therefore the sequence $\{Tf_n\}_{n=1}^{\infty}$ is equiconsistent and by the Arzela-Ascoli theorem it has a convergent subsequence.

Proposition 191. *If $T \in K(X)$ and $\lambda \neq 0$, then the image of $T - \lambda I$ is closed in X.*

Proof. Suppose that $\{x_n\}_{n=1}^{\infty}$ is a sequence and $(T - \lambda I)x_n \to y$, then we need to show that there is some x such that $Tx = y$.

We can assume that $y \neq 0$, since $(T - \lambda I)0 = 0$ by linearity. If $\{x_n\}$ is bounded, then there is a convergent subsequence $Tx_{n'}$, therefore $Tx_{n'} - \lambda x_{n'} \to y$ implies that $\lambda x_{n'} \to \lambda x$ is convergent. But this means that $(T - \lambda I)x_{n'} \to (T - \lambda I)x$ which means $Tx = y$ as wanted.

It remains to show the claim when $\|x_n\| \to \infty$ (if there was a bounded subsequence, then by passing to it we reduce to the previous case). let $E = \ker(T - \lambda I)$, then $x_n \notin E$ for all but finitely many $n \in \mathbb{N}$ as in that case $y = 0$ and we assumed this is not the case. Let F_n be the space spanned by $E \cup \{x_n\}$, then there is $z_n \in F_n$ which is almost orthogonal to E, namely $d(z_n, E) \geq \frac{1}{2}$ and $\|z_n\| = 1$. We can write $z_n = a_n x_n + u_n$ where $a_n \in \mathbb{C}$ and $u_n \in E$.

We observe that $\{a_n\}_{n=1}^{\infty}$ has no subsequence converging to 0. Otherwise $a_{n'} \to 0$ and we have that $(T - \lambda I)z_{n'} = a_{n'}(T - \lambda I)x_{n'} \to (\lim a_{n'})y = 0$. But now $\{z_{n'}\}$ is a bounded sequence so there is a subsequence $\{z_{n''}\}$ such that $z_{n''} \to z$ and therefore $(T - \lambda I)z_{n''} \to (T - \lambda I)z$ which implies that $Tz = 0$ and so $z \in E$, which is a contradiction since $d(z_{n''}, E) \to d(z, E) > 0$.

And so we have that,

$$(T - \lambda I)\frac{z_n}{a_n} = (T - \lambda I)x_n \to y.$$

Moreover $|a_n| \geq c > 0$ for some c, and therefore $\frac{1}{a_n}$ is bounded so $\{\frac{z_n}{a_n}\}$ is bounded and we reduced to the first part of the proof. □

Proposition 192. *Let $T \in K(X)$ then the following three statements cannot occur simultaneously:*

1. *There is $\{x_n \mid n \in \mathbb{N}\}$ which is linearly independent.*
2. *There is $\lambda_n \to \lambda \neq 0$.*
3. *$Tx_n = \lambda_n x_n$.*

Namely, there is no sequence of linearly independent eigenvectors whose eigenvalues do not converge to 0.

Proof. Assume towards contradiction that the three statements happen. For every $n \in \mathbb{N}$ let $F_n = \operatorname{span}\{x_1, \dots, x_n\}$, then $F_n \subsetneq F_{n+1}$. Then there is some $y_{n+1} \in F_{n+1}$ such that $\|y_{n+1}\| = 1$ and $d(y_{n+1}, F_n) \geq \frac{1}{2}$.

Note that $T(F_n) \subseteq F_n$, so if we take $m > n$ we get the follow equality:

$$Ty_m - Ty_n = Ty_m - \lambda_m y_m - Ty_n + \lambda_m y_m.$$

And $Ty_m - \lambda_m y_m - Ty_n \in F_n \subseteq F_{m-1}$, so for large enough m we get

$$\|Ty_m - Ty_n\| \geq d(\lambda_m y_m, F_{m-1}) = |\lambda_m| d(y_m, F_{m-1} \geq \frac{|\lambda}{2}\frac{1}{2}.$$

But T is compact, so $\{Ty_n\}$ has a convergent subsequence which is a contradiction as the above shows that $\{Ty_n\}$ is discrete. □

Corollary 193. *Suppose that $\lambda \neq 0$ is an eigenvalue of T, then it has finite dimensional eigenspace.* □

11.3 The Spectral Theorem

Theorem 194. *Let $T \in K(X)$, then $\sigma(T) \setminus \{0\}$ is a sequence $\{\lambda_n\}_{n=1}^{\infty}$ of eigenvalues and* $\lim_{n\to\infty} \lambda_n = 0$.

Proof. Suppose $\lambda \neq 0$ is not an eigenvalue, then $\ker(T - \lambda I) = \{0\}$, so if $T - \lambda I$ is surjective, $\lambda \notin \sigma(T)$.

Let F_n be the range of $(T - \lambda I)^n$ for $n \in \mathbb{N}$. Then $F_{n+1} \subseteq F_n$, and we claim that each F_n is closed. To see that, note that $(T - \lambda I)^n = \sum c_\beta T^\beta + (-\lambda)^n I$, but the sum is compact so by a previous claim F_n is closed. Assume towards contradiction that $F_{n+1} \subsetneq F_n$ for all n, find $x_n \in F_n$ such that $\|x_n\| = 1$ and $d(x_n, F_{n+1}) \geq \frac{1}{2}$. Take $m > n$ and we have that

$$Tx_m - Tx_n = Tx_m - \lambda x_n - (Tx_n - \lambda x_n) = Tx_m - (Tx_n - \lambda x_n) - \lambda x_n.$$

As in the previous proof, $Tx_m - (Tx_n - \lambda x_n) \in F_{n+1}$, so

$$\|Tx_m - Tx_n\| \geq |\lambda| d(x_n, F_{n+1}) \geq \frac{1}{2}|\lambda|.$$

But by the compactness of T we obtain a contradiction, as $\{Tx_n \mid n \in \mathbb{N}\}$ is discrete.

Therefore the sequence of F_n's has to stabilize. There is some n for which $F_n = F_{n+1}$. Now we claim that actually $F_{n-1} = F_n$ as well. Suppose that $x \in F_{n-1}$ then there is $y \in X$ such that $x = (T - \lambda I)^{n-1} y$. Therefore,

$$(T - \lambda I)x = (T - \lambda I)^n y = (T - \lambda I)^{n+1} z.$$

But λ is not an eigenvalue, then $x = (T - \lambda I)^n z \in F_n$. Therefore $F_{n-1} = F_n$, so it has to be the case that $F_n = X$ for all $n \in \mathbb{N}$ and $T - \lambda I$ is surjective as wanted.

We have proved, if so, that every $\lambda \in \sigma(T)$ is an eigenvalue. We will prove that for every $n \in \mathbb{N}$ there are at most finitely many $\lambda \in \sigma(T)$ such that $|\lambda| \geq \frac{1}{n}$. But this follows easily from the last proposition, since between $\frac{1}{n}$ and $\|T\|$ there can be at most finitely many $|\lambda|$'s. □

Corollary 195 (Fredholm alternative)**.** *Let $T \in K(X)$ and $\lambda \neq 0$. If λ is not an eigenvalue there is a unique solution for every y to the equation $(T - \lambda I)x = y$.*