

Functional Analysis

github.com/dan-gc/analysis

These are preparation notes for a course on Functional Analysis at IMPA, summer 2024. They are based on Folland, *Real Analysis: Modern Techniques and Their Applications*; Bachman and Narici, *Functional Analysis*; Brezis, *Functional Analysis, Sobolev Spaces and Partial Differential Equations*.

Contents

1	Measure theory	3
1.1	σ -algebras	3
1.2	Measures	4
1.3	Outer measures	5
1.4	Measurable functions	7
1.5	Riesz and Nagy's Functional Analysis	10
1.6	Integration	11
1.7	Riemann and Lebesgue integrals	14
1.8	Modes of convergence	15
1.9	Product measures	16
2	Point set topology	19
2.1	Metric spaces	19
2.2	Topological spaces	20
2.3	Compact spaces	22
2.4	Locally Compact Hausdorff spaces	24
2.5	Three compactness theorems	25
2.6	The Stone-Weierstrass Theorem	25
3	Inner product spaces	27
3.1	Inner products	27
3.2	Orthogonal projections	28
3.3	Riesz representation theorem	29
3.4	Adjoint operator	29
3.5	Spectral theorem for normal transformations	30
3.6	Unitary and orthogonal transformations	30
3.7	Normed spaces	31
4	Hahn-Banach theorems	33
4.1	Analytic form of the Hahn-Banach theorem	33
4.2	Geometric forms of the Hahn-Banach theorem	36
4.3	The bidual E^{**} . Orthogonality relations.	38
5	The Uniform Boundedness Principle and the Closed Graph Theorem	38
5.1	The Baire Category Theorem	38

5.2	The Uniform Boundedness Principle	39
5.3	The Open Mapping Theorem and the Closed Graph Theorem	42
5.4	An introduction to Unbounded Linear Operators. Definition of the Adjoint	43
6	Weak topologies. Reflexive Spaces. Separable Spaces. Uniformly Convex Spaces.	46
6.1	The weak topology $\sigma(E, E^*)$	46
6.2	The weak* topology $\sigma(E^*, E)$	48
6.3	Reflexive spaces	49
6.4	Separable spaces	52
6.5	Uniformly Convex Spaces	53
6.6	Exercises	54
7	L^p spaces.	55

1 Measure theory

A measure should surely satisfy:

1. If E_1, E_2, \dots is a finite or infinite sequence of disjoint sets,

$$\mu(E_1 \cup E_2 \cup \dots) = \mu(E_1) + \mu(E_2) + \dots$$

2. If E is congruent to F ,

$$\mu(E) = \mu(F)$$

3. If Q is the unit cube,

$$\mu(Q) = 1$$

1.1 σ -algebras

Let X be a nonempty set.

- An *algebra of sets* on X is a nonempty collection \mathcal{A} of subsets of X that is closed under finite unions and complements, that is,
 1. If $E_1, \dots, E_n \in \mathcal{A}$, then $\bigcup_{i=1}^n E_i \in \mathcal{A}$.
 2. If $E \in \mathcal{A}$, then $E^c \in \mathcal{A}$.
- A σ -*algebra* is an algebra of sets closed under countable unions.
- The intersection of all σ -algebras containing any subset $\mathcal{E} \subset \mathcal{P}(X)$ is the σ -*algebra generated by* \mathcal{E} .
- The σ -algebra generated by the open sets of a topological (or metric) space X is the *Borel algebra* \mathcal{B}_X .
- Let $\{X_\alpha\}_{\alpha \in A}$ is a collection of nonempty sets, $X = \prod_{\alpha} X_\alpha$ and $\pi_\alpha : X \rightarrow X_\alpha$ the coordinate functions. If \mathcal{M}_α is a σ -algebra on X_α , the *product σ -algebra* on X is the σ -algebra generated by

$$\{\pi_\alpha(E_\alpha) : E_\alpha \in \mathcal{M}_\alpha, \alpha \in A\}$$

We denote this σ -algebra by $\bigotimes_{\alpha \in A} \mathcal{M}_\alpha$.

Proposition 1.1. $\bigotimes_{\alpha \in A} \mathcal{M}_\alpha$ is generated by $\{\prod_{\alpha \in A} E_\alpha : E_\alpha \in \mathcal{M}_\alpha\}$.

Proposition 1.2. Let X_1, \dots, X_n be metric spaces and let $X = \prod_i X_i$ be equipped with the product metric. Then $\bigotimes_i \mathcal{B}_i \subset \mathcal{B}_X$. If every X_i is separable equality holds.

Corollary 1.3. $\mathcal{B}_{\mathbb{R}^n} = \bigotimes_{i=1}^n \mathcal{B}_{\mathbb{R}}$.

1.2 Measures

Let X be a set equipped with a σ -algebra \mathcal{M} .

- A **measure** on \mathcal{M} (or on (X, \mathcal{M}) , or simply on X) is a function $\mu : \mathcal{M} \rightarrow [0, \infty)$ such that
 1. $\mu(\emptyset) = 0$.
 2. if $\{E_j\}_1^\infty$ is a sequence of disjoint sets in \mathcal{M} , then $\mu(\bigcup_1^\infty E_j) = \sum_1^\infty \mu(E_j)$.
- If X is a set and $\mathcal{M} \subset \mathcal{P}(X)$ is a σ -algebra, (X, \mathcal{M}) is called a **measurable space** and the sets in \mathcal{M} are **measurable sets**. If μ is a measure on (X, \mathcal{M}) , then (X, \mathcal{M}, μ) is called a **measure space**.
- If $\mu(X) < \infty$ (and hence $\mu(E) < \infty$ for all $E \in \mathcal{M}$), μ is called **finite**. If $X = \bigcup_1^\infty E_j$, with $\mu(E_j) < \infty$, μ is called **σ -finite**. If for every $E \in \mathcal{M}$ with $\mu(E) = \infty$ there exists $F \in \mathcal{M}$ such that $F \subset E$ and $0 < \mu(F) < \infty$, μ is called **semifinite**.

Theorem 1.4 (Properties of measure spaces). Let (X, \mathcal{M}, μ) be a measure space.

1. **(Monotonicity.)** If $E, F \in \mathcal{M}$ and $E \subset F$, then $\mu(E) \leq \mu(F)$.
 2. **(Subadditivity.)** If $\{E_j\}_1^\infty \subset \mathcal{M}$, then $\mu(\bigcup_1^\infty E_j) \leq \sum_1^\infty \mu(E_j)$.
 3. **(Continuity from below.)** If $\{E_j\}_1^\infty \subset \mathcal{M}$, $E_1 \subset E_2 \subset \dots$, then $\mu(\bigcup_1^\infty E_j) = \lim_{j \rightarrow \infty} \mu(E_j)$.
 4. **(Continuity from above.)** If $\{E_j\}_1^\infty \subset \mathcal{M}$, $E_1 \supset E_2 \supset \dots$, and $\mu(E_1) < \infty$, then $\mu(\bigcap_1^\infty E_j) = \lim_{j \rightarrow \infty} \mu(E_j)$.
- If $E \in \mathcal{M}$ and $\mu(E) = 0$, E is called a **null set**. If a statement about points in X is true except for points in a null set, we say it is true **almost everywhere**.
 - If $\mu(E) = 0$ and $F \subset E$, then $\mu(F) = 0$ provided $F \in \mathcal{M}$. A measure whose domain contains all subsets of null sets is **complete**. Completeness may help avoid technical difficulties, and it can always be achieved by enlarging the domain of μ :

Theorem 1.5. Let (X, \mathcal{M}, μ) be a measure. Let $\mathcal{N} = \{N \in \mathcal{M} : \mu(N) = 0\}$ and $\overline{\mathcal{M}} = \{E \cup F : E \in \mathcal{M} \text{ and } F \subset N \text{ for some } N \in \mathcal{N}\}$. Then $\overline{\mathcal{M}}$ is a σ -algebra, and there is a unique extension $\bar{\mu}$ of μ to a complete measure on $\overline{\mathcal{M}}$, called the **completion of μ** .

- Exercises
 1. If μ_1, \dots, μ_n are measures on (X, \mathcal{M}) and $a_1, \dots, a_n \in [0, \infty)$, then $\sum_1^n a_j \mu_j$ is also a measure on (X, \mathcal{M}) .
 2. $\mu(E \cup F) = \mu(E) + \mu(F) - \mu(E \cap F)$.
 3. A set $E \subset X$ is called **locally measurable** if $E \cap A \in \mathcal{M}$ whenever $A \in \mathcal{M}$ and $\mu(A) < \infty$. If $\tilde{\mathcal{M}}$ equals the collection of the saturated sets \tilde{M} , it is called **saturated**. The saturated measure $\tilde{\mu}$ on $\tilde{\mathcal{M}}$ defined by $\tilde{\mu}(E) = \mu(E)$ for $E \in \mathcal{M}$ and $\tilde{\mu}(E) = \infty$ otherwise is called the **saturation of μ** .

1.3 Outer measures

This is used to construct measures. The key idea is to approximate the measure of a set by simpler enclosing sets, like with the Riemann integral.

- An **outer measure** on a nonempty set X is a function $\mu^* : \mathcal{P}(X) \rightarrow [0, \infty]$ that satisfies

1. $\mu^*(\emptyset) = 0$,
2. $\mu^*(A) \leq \mu^*(B)$ if $A \subset B$, and
3. $\mu^*(\bigcup_1^\infty A_j) \leq \sum_1^\infty \mu^*(A_j)$.

Proposition 1.6. Let $\mathcal{E} \subset \mathcal{P}(X)$ and $\rho : \mathcal{E} \rightarrow [0, \infty]$ be such that $\emptyset \in \mathcal{E}$, $X \in \mathcal{E}$ and $\rho\emptyset = 0$. For any $A \subset X$, define

$$\mu^*(A) = \inf \left\{ \sum_1^\infty \mu(E_j) : E_j \in \mathcal{E} \text{ and } A \subset \bigcup_1^\infty E_j \right\} \quad (1.1)$$

then μ^* is an outer measure.

- A set $A \subset X$ is called **μ -measurable** if

$$\mu^*(E) = \mu^*(E \cap A) + \mu^*(E \cap A^c) \quad \text{for all } E \subset X$$

Which makes sense if we think E is a *well-behaved* set such that $A \subset E$, so that $\mu^*(A) = \mu^*(E) - \mu^*(E \cap A^c)$.

Theorem 1.7 (Carathéodory). If μ^* is an outer measure on X , the collection \mathcal{M} of μ^* -measurable sets is a σ -algebra, and the restriction of μ^* to \mathcal{M} is a complete measure.

- If $\mathcal{A} \subset \mathcal{P}(X)$ is an algebra, the function $\mu_0 : \mathcal{A} \rightarrow [0, \infty]$ is a **premeasure** if
 - $\mu_0(\emptyset) = 0$,
 - If $\{A_j\}_1^\infty$ is a sequence of disjoint sets in \mathcal{A} such that $\bigcup_1^\infty A_j \in \mathcal{A}$, then $\mu_0(\bigcup_1^\infty A_j) = \sum_1^\infty \mu_0(A_j)$.

Proposition 1.8. If μ_0 is a premeasure on \mathcal{A} and μ^* is defined by eq. (1.1) by taking $\rho = \mu_0$, then

1. $\mu^*|_{\mathcal{A}} = \mu_0$,
2. every set in \mathcal{A} is μ^* measurable.

Theorem 1.9. Let $\mathcal{A} \subset \mathcal{P}(X)$ be an algebra, μ_0 a premeasure on \mathcal{A} and \mathcal{M} the σ -algebra generated by \mathcal{A} . There exists a measure μ on \mathcal{M} whose restriction to \mathcal{A} is μ_0 —namely $\mu = \mu^*|_{\mathcal{M}}$, where μ^* is given by eq. (1.1). (This is a consequence of Carathéodory's theorem and the last proposition.)

If ν is another measure on \mathcal{M} that extends μ_0 , then $\nu(E) \leq \mu(E)$ for all $E \in \mathcal{M}$ with equality when $\mu(E) < \infty$. If μ_0 is σ -finite, then μ is the unique extension of μ_0 to a measure on \mathcal{M} .

- A **Borel measure on \mathbb{R}** is a measure on \mathbb{R} whose domain is the Borel σ -algebra $\mathcal{B}_{\mathbb{R}}$.

Proposition 1.10. Let $F : \mathbb{R} \rightarrow \mathbb{R}$ be increasing and right continuous. If $(a_j, b_j]$, $j = 1, \dots, n$ are disjoint **half-open intervals**, so that $0 \leq a_j, b_j \leq \infty$, or (a_j, ∞) , or they are empty, define

$$\mu_0 \left(\bigcup_{j=1}^n (a_j, b_j] \right) = \sum_{j=1}^n [F(b_j) - F(a_j)]$$

and let $\mu_0(\emptyset) = 0$. Then μ_0 is a premeasure on the algebra \mathcal{A} of finite disjoint unions of half-open intervals.

Theorem 1.11. If $F : \mathbb{R} \rightarrow \mathbb{R}$ is any increasing, right continuous function, there is a unique Borel measure μ_F on \mathbb{R} such that $\mu_F((a, b]) = F(b) - F(a) \forall a, b$.

If G is another such function, $\mu_F = \mu_G$ if and only if $F - G$ is constant. Conversely, given a μ Borel measure on \mathbb{R} that is finite on all bounded Borel sets we may define an increasing right continuous function F such that $\mu = \mu_F$.

- By theorem 1.9, there is a complete measure $\bar{\mu}_F$ whose domain includes $\mathcal{B}_{\mathbb{R}}$. In fact $\bar{\mu}_F$ is the completion of μ_F and its domain is strictly larger than $\mathcal{B}_{\mathbb{R}}$. This complete measure is called the **Lebesgue-Stieltjes measure** and is also denoted by μ_F .

In the following μ is the Lebesgue-Stieltjes measure associated to some increasing, right-continuous function F , and \mathcal{M}_{μ} is the domain of μ .

Theorem 1.12. If $E \in \mathcal{M}_{\mu}$, then

$$\begin{aligned} \mu(E) &= \inf \{ \mu(U) : E \subseteq U \text{ and } U \text{ is open} \} \\ &= \sup \{ \mu(K) : K \subseteq E \text{ and } K \text{ is compact} \} \end{aligned}$$

Theorem 1.13. If $E \subset \mathbb{R}$, the following are equivalent:

1. $E \in \mathcal{M}_{\mu}$.
2. $E = V \setminus N_1$ where V is a G_{δ} (countable intersection of open sets) and $\mu(N_1) = 0$.
3. $E = H \cup N_2$ where H is an F_{σ} (countable union of closed sets) set and $\mu(N_2) = 0$.

Proposition 1.14. If $E \in \mathcal{M}_{\mu}$ and $\mu(E) < \infty$, then for every $\varepsilon > 0$, there is a set A that is a finite union of open intervals such that $\mu(E \triangle A) < \varepsilon$.

- The **Lebesgue measure** is the Lebesgue-Stieltjes measure of $F(x) = x$. We denote it by m and its domain by \mathcal{L} .

Theorem 1.15. Is $E \in \mathcal{L}$, then $E + s \in \mathcal{L}$ and $rE \in \mathcal{L}$ for all $s, r \in \mathbb{R}$. Moreover, $m(E + s) = m(E)$ and $m(rE) = |r|m(E)$.

Proposition 1.16. Let C be the Cantor set. C is compact, nowhere dense and totally disconnected (ie. the only connected subsets of C are single points). C has no isolated points. $m(C) = 0$. $\text{card}(C) = \mathfrak{c}$.

1.4 Measurable functions

- Recall that a mapping $f : X \rightarrow Y$ between two sets induces a mapping $f^{-1} : \mathcal{P}(Y) \rightarrow \mathcal{P}(X)$ defined by the inverse image, which preserves unions, intersection and complements, so that if \mathcal{N} is a σ -algebra on Y , then $\{f^{-1}(E) : E \in \mathcal{N}\}$ is a σ -algebra on X .

If (X, \mathcal{M}) and (Y, \mathcal{N}) are measurable spaces, a mapping $f : X \rightarrow Y$ is called *measurable* if $f^{-1}(E) \in \mathcal{M}$ for all $E \in \mathcal{N}$.

Proposition 1.17. A function whose codomain is a product measure space is measurable if precomposing with every projection is measurable.

Corollary 1.18. A function $f : X \rightarrow \mathbb{C}$ is \mathcal{M} -measurable if and only if $\text{Re } f$ and $\text{Im } f$ are \mathcal{M} -measurable.

Proposition 1.19. If \mathcal{N} is generated by \mathcal{E} , then $f : X \rightarrow Y$ is measurable if and only if $f^{-1}(E) \in \mathcal{M}$ for all $E \in \mathcal{E}$.

Corollary 1.20. If X and Y are metric (or topological spaces), every continuous function is $(\mathcal{B}_X, \mathcal{B}_Y)$ -measurable.

Proposition 1.21. If $f, g : X \rightarrow \mathbb{C}$ are \mathcal{M} -measurable, then so are $f + g$ and fg .

Proposition 1.22. If $\{f_j\}$ is a sequence of $\overline{\mathbb{R}}$ -valued measurable functions on (X, \mathcal{M}) , then the functions

$$\begin{aligned} \sup_j f_j(x) & \quad \limsup_{j \rightarrow \infty} f_j(x) \\ \inf_j f_j(x) & \quad \liminf_{j \rightarrow \infty} f_j(x) \end{aligned}$$

are measurable. If

$$f(x) = \lim_{j \rightarrow \infty} f_j(x)$$

exists for every $x \in X$, then f is measurable.

Corollary 1.23. If $f, g : X \rightarrow \overline{\mathbb{R}}$ is measurable, then so are $\max(f, g)$ and $\min(f, g)$.

If $f : X \rightarrow \overline{\mathbb{R}}$ we define the *positive* and *negative* parts of f as:

$$f^+(x) = \max(f(x), 0) \quad f^-(x) = \max(-f(x), 0)$$

Then $f = f^+ - f^-$, and if f is measurable, so are f^+ and f^- by corollary 1.23.

- Let (X, \mathcal{M}) be a measurable space. If $E \subset X$, the **characteristic or indicator function** χ_E of E is

$$\chi(x) = \begin{cases} 1 & \text{if } x \in E \\ 0 & \text{if } x \notin E \end{cases}$$

χ_E is measurable if and only if $E \in \mathcal{M}$.

A **simple function on X** is a finite linear combination, with complex coefficients, of characteristic functions of sets in \mathcal{M} . (We do not allow simple functions to assume values $\pm\infty$.) Equivalently, $f : X \rightarrow \mathbb{C}$ is simple if and only if f is measurable and the range of f is a finite subset of \mathbb{C} . The **standard representation of f** is

$$f = \sum_{j=1}^n z_j \chi_{E_j} \quad \text{where } E_j = f^{-1}(\{z_j\}) \text{ and } \text{range}(f) = \{z_1, \dots, z_n\}$$

If f and g are simple functions, then so are $f + g$ and fg .

Theorem 1.24. Let (X, \mathcal{M}) be a measurable space.

1. If $f : X \rightarrow [0, \infty]$ is measurable, there is a sequence $\{\varphi_n\}$ of simple functions such that $0 \leq \varphi_1 \leq \dots \leq f$, $\varphi_n \rightarrow f$ pointwise and $\varphi_n \rightarrow f$ uniformly on any set on which f is bounded.
2. If $f : X \rightarrow \mathbb{C}$ is measurable, there is a sequence $\{\varphi_n\}$ of simple functions such that $0 \leq |\varphi_1| \leq \dots \leq |f|$, $\{\varphi_n\} \rightarrow f$ pointwise and $\varphi_n \rightarrow f$ uniformly on any set on which f is bounded.

Proof.

1. Let $n \in \mathbb{N}$ and $0 \leq k \leq 2^{2n} - 1$. Set

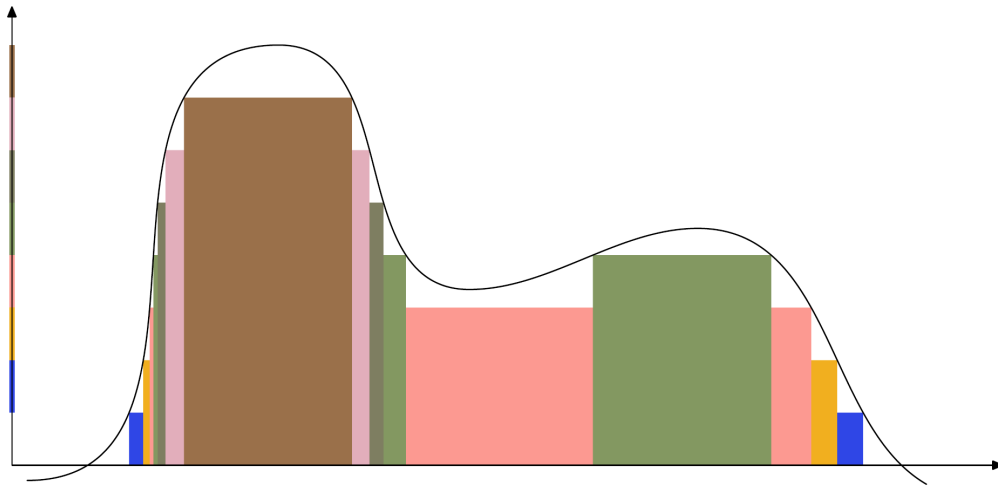
$$E_n^k = f^{-1}((k2^{-n}, (k+1)2^{-n}]) \quad \text{and} \quad F_n = f^{-1}((2^n, \infty]),$$

and finally

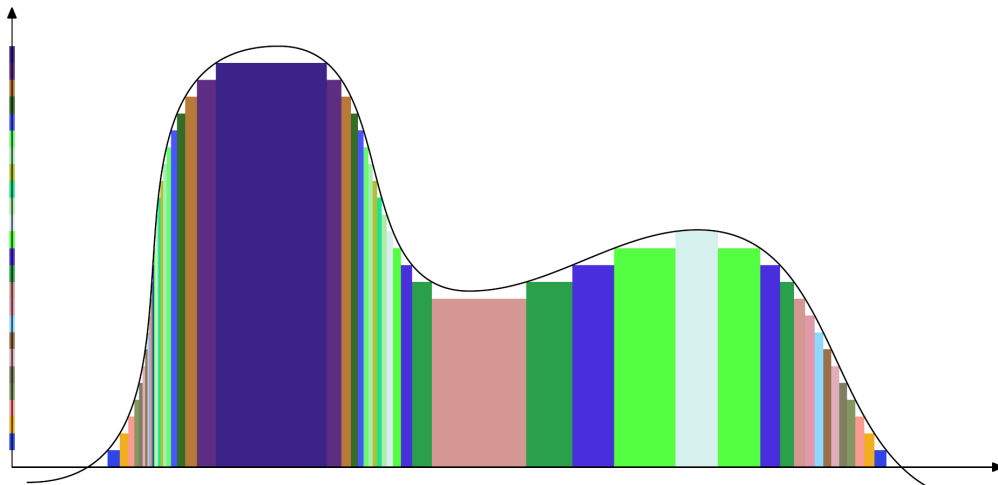
$$\varphi_n = \sum_{k=0}^{2^{2n}-1} k2^{-n} \chi_{E_n^k} + 2^n \chi_{F_n}.$$

Then $\varphi_n \leq \varphi_{n+1} \forall n$ and $0 \leq f - \varphi_n \leq 2^{-n}$ on the set where $f \leq 2^n$.

2. If $f = g + ih$ we can apply part *a* to the positive and negative parts of g and h .



(a) φ_7 .



(b) φ_{24} .

Figure 1: Approximation by simple functions.

□

Proposition 1.25. Let (X, \mathcal{M}, μ) be a measure space and let $(X, \overline{\mathcal{M}}, \overline{\mu})$ be its completion. If f is a $\overline{\mathcal{M}}$ -measurable function on X , there is an \mathcal{M} -measurable function g such that $f = g\overline{\mu}$ -almost everywhere.

1.5 Riesz and Nagy's Functional Analysis

This section is a short digression to show the approach of Riesz and Nagy, *Functional Analysis*.

Theorem 1.26 (Lebesgue). Every monotonic function $f(x)$ possesses a finite derivative at every point with the possible exception of the points x of a set of measure zero, or, as it is often phrased, almost everywhere.

A set of *measure zero* is a set of values x which can be covered by a finite number or by a denumerable sequence of intervals whose total length is arbitrarily small.

Functions $f(x)$ continuous or not, for which the sum

$$\sum_{ab} = \sum_1^n |f(x_k) - f(x_{k-1})|$$

defined in terms of a decomposition of the interval (a, b) into partial intervals (x_{k-1}, x_k) $k = 1, 2, \dots, n$, does not surpass a finite bound, independent of the particular choice of decomposition, are called **functions of bounded variation**. The least upper bound is called the **total variation of $f(x)$ in the interval (a, b)** .

Theorem 1.27 (Lebesgue). Every function of bounded variation possesses a finite derivative almost everywhere.

Theorem 1.28 (Fubini). Let

$$f_1(x) + f_2(x) + \dots = s(x)$$

be a convergent series all of whose terms are monotonic functions of the same type, defined on the interval $a \leq x \leq b$. Then

$$f'_1(x) + f'_2(x) + \dots = s'(x)$$

except perhaps on a set of measure zero; that is, term by term differentiation is possible almost everywhere.

Theorem 1.29 (Lebesgue). Almost all points of an arbitrary linear set are density points of that set.

We include the following quote regarding Riemann integrals:

[p. 23] (...) we arrive at a necessary condition that $f(x)$ be integrable in the Riemann sense, namely, that the function $f(x)$ be continuous almost everywhere.

A **step function** defined on an interval (a, b) is a function having a constant value c_k in each of a finite number of intervals i_k of finite length $|i_k|$ and vanishing outside these intervals. We suppose the integral defined for these functions, as usual, by the sum $\sum c_k |i_k|$.

Lemma 1.30. For every sequence $\{\varphi_n(x)\}$ of step functions which decreases to 0 almost everywhere, the sequence of values of their integrals also tends to zero.

Lemma 1.31. If for an increasing sequence of step functions $\{\varphi_n(x)\}$ the values of their integrals have a common bound, then the sequence $\{\varphi_n(x)\}$ tends almost everywhere to a finite limit.

Denote the class of stepfunctions by C_0 and the class of functions which are limits almost everywhere of the sequences $\{\varphi_n\}$ referred to in lemma 1.31. For $f(x) = \lim \varphi_n(x)$ define

$$\int_a^b f(x)dx = \lim_{n \rightarrow \infty} \int_a^b \varphi_n(x)dx$$

which does not depend on the particular choice of functions φ_n . In fact, if $\{\psi_n\}$ is increases almost everywhere to a limit function $g(x) \geq f(x)$, we also have

$$\lim_{n \rightarrow \infty} \int_a^b \psi_n(x)dx \geq \lim_{n \rightarrow \infty} \int_a^b \varphi_n(x)dx$$

so that if $g(x) = f(x)$ almost everywhere is implied.

1.6 Integration

In this section we continue reading Folland, [Real Analysis: Modern Techniques and Their Applications](#).

Fix a measure space (X, \mathcal{M}, μ) and define

$$L^+ = \text{the space of all measurable functions from } X \text{ to } [0, \infty].$$

If φ is a simple function in L^+ with standard representation $\varphi = \sum_{j=1}^n a_j \chi_{E_j}$, we define the *integral of φ* by

$$\int \varphi d\mu = \sum_{j=1}^n a_j \mu(E_j)$$

with the convention that $0 \cdot \infty = 0$. It may happen that $\int \varphi d\mu = \infty$.

Proposition 1.32. Let φ and ψ be simple functions in L^+ .

1. If $c \geq 0$, then $\int c\varphi = c \int \varphi$.
2. $\int(\varphi + \psi) = \int \varphi + \int \psi$.
3. If $\varphi \leq \psi$ then $\int \varphi \leq \int \psi$.
4. The map $A \mapsto \int_A d\mu$ is a measure on \mathcal{M} .

We now extend the integral to all functions in L^+ by defining

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

Remark 1.1. In Riesz and Nagy, *Functional Analysis* the integral of f is the *limit* of the integrals of simple functions.

Remark 1.2. It is obvious from the definition that if $f \leq g$ then $\int f \leq \int g$ and $\int cf = c \int f$ for all $c \in [0, \infty)$.

Theorem 1.33 (Monotone convergence, Beppo Levi, as in Folland). If $\{f_n\}$ is a sequence in L^+ (the set of measurable functions with values on $[0, \infty]$) such that $f_j \leq f_{j+1}$ for all j and $f = \lim_{n \rightarrow \infty} f_n$, then $\int f_n \rightarrow \int f$.

Theorem 1.34 (Monotone convergence, Beppo Levi, as in Brezis). Let (f_n) be a sequence of functions in L^1 that satisfy

1. $f_1 \leq f_2 \leq \dots \leq f_n \leq f_{n+1} \leq \dots$ a.e. on Ω ,
2. $\sup_n \int f_n < \infty$.

Then $f_n(x)$ converges a.e. on Ω to a finite limit, which we denote by $f(x)$. The function f belongs to L^1 and $\|f_n - f\|_1 \rightarrow 0$.

Proof of the first statement. Since $\int f_n \leq \int f$ for all n , we have $\lim \int f_n \leq \int f$. To establish the reverse inequality fix $\alpha \in (0, 1)$, let φ be a simple function with $0 \leq \varphi \leq f$ and $E_n = \{x : f_n(x) \geq \alpha\varphi(x)\}$. That is E_n is the set of points in X where the approximating function f_n has a value greater than $\alpha\varphi(x)$.

Then $\{E_n\}$ is an increasing sequence of measurable sets ... □

Theorem 1.35 (Additivity of the integral). If $\{f_n\}$ is a finite or infinite sequence in L^+ and $f = \sum_n f_n$ then $\int f = \sum_n \int f_n$.

Proposition 1.36. If $f \in L^+$ then $\int f = 0$ if and only if $f = 0$ a.e.

Corollary 1.37. If $\{f_n\} \subset L^+$ and $f_n(x)$ increases to $f(x)$ for a.e. x , then $\int f = \lim \int f_n$.

Lemma 1.38 (Fatou). If $\{f_n\}$ is any sequence in L^+ , then

$$\int \liminf f_n \leq \liminf \int f_n.$$

Proof. For each $k \geq 1$, we have $\inf_{n \geq k} f_n \leq f_j$ for $j \geq k$, so that $\int \inf_{n \geq k} f_n \leq \int f_j$ for $j \geq k$, and hence $\int \inf_{n \geq k} f_n \leq \inf_{j \geq k} \int f_j$. As $k \rightarrow \infty$,

$$\int \liminf f_n = \int \lim_{k \rightarrow \infty} \inf_{n \geq k} f_n \stackrel{\text{m.c.t.}}{=} \lim_{k \rightarrow \infty} \int \inf_{n \geq k} f_n \leq \lim_{k \rightarrow \infty} \inf_{j \geq k} \int f_j = \liminf \int f_n$$

by the monotone convergence theorem. □

Corollary 1.39. If $\{f_n\} \subset L^+$, $f \in L^+$ and $f_n \rightarrow f$ a.e., then $\int f \leq \liminf \int f_n$.

Proposition 1.40. If $f \in L^+$ and $\int f < \infty$, then $\{x : f(x) = \infty\}$ is a null set and $\{x : f(x) > 0\}$ is σ -finite.

We may extend the former integral to real-valued functions as follows. If f^+ and f^- are the positive and negative parts of f and at least one of $\int f^+$ and $\int f^-$ is finite, we define

$$\int f = \int f^+ - \int f^-.$$

If $\int f^+$ and $\int f^-$ are both finite we say f is **integrable**. Since $|f| = f^+ + f^-$, it is clear that f is integrable if and only if $\int |f| < \infty$.

Proposition 1.41. The set of integrable real-valued functions on X is a real vector space and the integral is a linear functional on it.

If f is a complex-valued measurable function, we say that f is **integrable** if $\int |f| < \infty$. Since $|f| \leq |\operatorname{Re} f| + |\operatorname{Im} f| \leq 2|f|$, f is integrable if and only if $\operatorname{Re} f$ and $\operatorname{Im} f$ are both integrable, and we define

$$\int f = \int \operatorname{Re} f + i \int \operatorname{Im} f.$$

The set of all complex-valued integrable functions is a complex vector space and the integral is a complex-linear functional on it. We denote this space by $L^1, L^1(\mu), L^1(X, \mu)$ or $L^1(X)$.

Proposition 1.42. If $f \in L^1$, then $|\int f| \leq \int |f|$.

Proposition 1.43.

1. If $f \in L^1$, then $\{x : f(x) \neq 0\}$ is σ -finite.
2. If $f, g \in L^1$, then $\int_E f = \int_E g$ for all $E \in \mathcal{M}$ if and only if $\int |f - g| = 0$ if and only if $f = g$ almost everywhere.

This proposition shows that for the purposes of integration it makes no difference if we alter functions on null sets. With this in mind, we redefine L^1 **as the set of equivalence classes of a.e.-defined integrable functions on X** , where f and g are considered equivalent if $f = g$ a.e.

This new definition has two advantages. First, that if $\bar{\mu}$ is the completion of μ , there is one-to-one correspondance between $L^1(\bar{\mu})$ and $L^1(\mu)$ by ??, and we shall identify those spaces. Second, L^1 is a metric space with distance function $\rho(f, g) = \int |f - g|$: the triangle inequality is easily verified, and obviously $\rho(f, g) = \rho(g, f)$, but to prove that $\rho(f, g) = 0$ only when $f = g$ we must identify functions that are equal a.e. We shall refer to convergence with respect to this metric as **convergence in L^1** .

Theorem 1.44 (Dominated Convergence, Lebesgue, as in Folland). Let $\{f_n\}$ be a sequence in L^1 such that $f_n \rightarrow f$ a.e. and there exists a nonnegative $g \in L^1$ such that $|f_n| \leq g$ a.e. for all n . Then $f \in L^1$ and $\int f_n \rightarrow \int f$.

Proof. By Fatou's Lemma. □

Theorem 1.45 (Dominated Convergence, Lebesgue, as in Brezis). Let (f_n) be a sequence of functions in L^1 that satisfy

1. $f_n(x) \rightarrow f(x)$ a.e. on Ω ,
2. there is a function $g \in L^1$ such that for all n , $|f_n(x)| \leq g(x)$ a.e. on Ω .

Then $f \in L^1$ and $\|f_n - f\|_1 \rightarrow 0$.

Remark 1.3. By comparing the statements of the Monotone Convergence and Dominated Convergence theorems in Folland and Brezis, we suspect that $\|f_n - f\|_1 \rightarrow 0 \iff \int f_n \rightarrow \int f$. At least one of the implications is straightforward:

$$\begin{aligned}
\|f_n - f\|_1 \rightarrow 0 &\iff \int |f_n - f| \rightarrow 0 \\
&\implies \left| \int (f_n - f) \right| \rightarrow 0 \\
&\iff \left| \int f_n - \int f \right| \rightarrow 0 \\
&\iff \int f_n \rightarrow \int f
\end{aligned}$$

Theorem 1.46. Suppose that $\{f_j\}$ is a sequence in L^1 such that $\sum_{j=1}^{\infty} \int |f_j| < \infty$. Then $\sum_{j=1}^{\infty} f_j$ converges a.e. to a function in L^1 , and $\int \sum_{j=1}^{\infty} f_j = \sum_{j=1}^{\infty} \int f_j$.

Proof. By the Dominated Convergence Theorem. □

Theorem 1.47. Integrable simple functions are dense in L^1 . That is, for every $f \in L^1$ and $\varepsilon > 0$ there is an integrable simple function $\varphi = \sum a_j \chi_{E_j}$ such that $\int |f - \varphi| d\mu < \varepsilon$.

If μ is a Lebesgue-Stieltjes measure on \mathbb{R} , the sets E_j in the definition of φ can be taken to be finite unions of open intervals, and there is a continuous function g **that vanishes outside a bounded interval** such that $\int |f - g| d\mu < \varepsilon$.

Proof. By the Dominated Convergence Theorem. □

The *Lebesgue integral* is the integral we have developed then the measure is the Lebesgue measure.

1.7 Riemann and Lebesgue integrals

In this section we briefly compare the Riemann and Lebesgue integrals using Darboux's characterization of the Riemann integral in terms of upper and lower sums, which we now recall. Let $[a, b]$ be a compact interval. By a *partition of $[a, b]$* we mean an infinite sequence $P = \{t_j\}_0^n$ such that $a = t_0 < t_1 < \dots < t_n = b$. Let f be an arbitrary real-valued function on $[a, b]$. For each partition P we define

$$S_P f = \sum_{j=0}^n M_j (t_j - t_{j-1}), \quad s_P f = \sum_{j=0}^n m_j (t_j - t_{j-1}),$$

where M_j and m_j are the supremum and infimum of f on $[t_{j-1}, t_k]$. Then we define

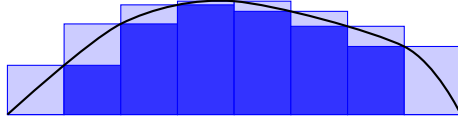
$$\bar{I}_a^b = \inf_P S_P f, \quad \underline{I}_a^b(f) = \sup_P s_P f.$$

If $\bar{I}_a^b(f) = \underline{I}_a^b(f)$, their common value is the **Riemann integral** $\int_a^b f$ and f is called **Riemann integrable**.

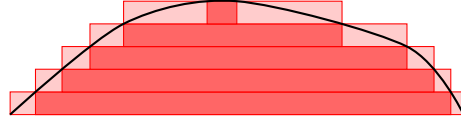
Theorem 1.48. Let f be a bounded real-valued function on $[a, b]$.

1. If f is Riemann integrable, then f is Lebesgue measurable (and hence integrable on $[a, b]$ since it is bounded), and the integrals coincide.
2. f is Riemann integrable if and only if $\{x \in [a, b] : f \text{ is discontinuous at } x\}$ has Lebesgue measure zero.

Proof. Uses the Dominated Convergence Theorem. □



(a) Riemann integral



(b) Lebesgue integral

Remark 1.4. There are two main advantages of Lebesgue integration over Riemann integration. First, the convergence theorems, which do not hold for Riemann integrals. Second, that a wider class of functions may be integrated, which ultimately results in the completeness of certain function spaces.

1.8 Modes of convergence

Of course uniform convergence implies pointwise convergence, which in turn implies almost everywhere convergence, but these modes of convergence do not imply convergence in L^1 or vice versa.

For example, $f_n = \frac{1}{n}\chi_{(0,n)}$ converges uniformly, pointwise and almost everywhere to 0, but $\int |f_n| = 1$ for all n . It is possible to find a function that converges in L^1 but does not converge pointwise.

The dominated convergence theorem states that if $f_n \rightarrow f$ a.e. and $|f_n| \leq g \in L^1 \forall n$, then $f \in L^1$. We also have that $f_n \rightarrow f$ in L^1 , since $|f_n - f| \leq 2g$. We shall see below that if $f_n \rightarrow f$ in L^1 then some subsequence converges to f a.e.

Now we introduce another mode of convergence. We say that a sequence $\{f_n\}$ of measurable complex-valued functions on (X, \mathcal{M}, μ) is **Cauchy in measure** if for every $\varepsilon > 0$,

$$\mu(\{x : |f_n(x) - f_m(x)| \geq \varepsilon\}) \rightarrow 0 \text{ as } m, n \rightarrow \infty,$$

and that $\{f_n\}$ *converges in measure* to f if for every $\varepsilon > 0$,

$$\mu(\{x : |f_n(x) - f_m(x)| \geq \varepsilon\}) \rightarrow 0 \text{ as } n \rightarrow \infty.$$

While our example sequence, $f_n = \frac{1}{n}\chi_{(0,n)}$ converges to zero in measure, the sequence $f_n = \chi_{[n,n+1]}$ is not Cauchy in measure.

Proposition 1.49. If $f_n \rightarrow f$ in L^1 , then $f_n \rightarrow f$ in measure.

Remark 1.5. The converse is false by the sequence $f_n = \frac{1}{n}\chi_{(0,n)}$.

Theorem 1.50. Suppose that $\{f_n\}$ is Cauchy in measure. Then there is a measurable function f such that $f_n \rightarrow f$ in measure, and there is a subsequence $\{f_{n_j}\}$ that converges to f a.e. Moreover, if also $f_n \rightarrow g$ in measure, then $g = f$ a.e.

Remark 1.6. Bear in mind that a metric space is complete if all Cauchy sequences converge, and that we expect the space of real-valued functions to be complete since \mathbb{R} is complete.

Remark 1.7. As an extra reminder, we recall that Cauchy sequences of real numbers are bounded, so that by the Bolzano-Weierstrass theorem they have convergent subsequences, which implies that they are convergent (since they are Cauchy).

Theorem 1.51. If $f_n \rightarrow f$ in L^1 , then there is a subsequence f_{n_j} that converges to f a.e.

Remark 1.8. This result may be thought as a kind of L^1 -to-a.e. *sequential compactness*.

While, as we have said, a.e. convergence does not imply convergence in measure, the conclusion is true in a finite measure space. In fact, something stronger holds:

Theorem 1.52 (Egoroff). Suppose that $\mu(X) < \infty$ and f_1, f_2, \dots and f are measurable complex-valued functions on X such that $f_n \rightarrow f$ a.e. Then f_n converges to f *almost uniformly*, that is, for every $\varepsilon > 0$ there exists $E \subset X$ such that $\mu(E) < \varepsilon$ and $f_n \rightarrow f$ on E^c . (Almost uniform convergence implies a.e. convergence and convergence in measure.)

1.9 Product measures

First we recall the construction of the product σ -algebra seen in section 1.1. Let $\{X_\alpha\}_{\alpha \in A}$ is a collection of nonempty sets, $X = \prod_{\alpha \in A} X_\alpha$ and $\pi_\alpha : X \rightarrow X_\alpha$ the coordinate functions. If \mathcal{M}_α is a σ -algebra on X_α , the **product σ -algebra** on X is the σ -algebra generated by

$$\{\pi_\alpha(E_\alpha) : E_\alpha \in \mathcal{M}_\alpha, \alpha \in A\}$$

We denote this σ -algebra by $\bigotimes_{\alpha \in A} \mathcal{M}_\alpha$.

Proposition 1.53. $\bigotimes_{\alpha \in A} \mathcal{M}_\alpha$ is generated by $\{\prod_{\alpha \in A} E_\alpha : E_\alpha \in \mathcal{M}_\alpha\}$.

In this section we consider two measure spaces (X, \mathcal{M}, μ) and (Y, \mathcal{N}, ν) and construct a measure on $\mathcal{M} \otimes \mathcal{N}$, the *product measure of μ and ν* .

To begin with we define a (*measurable*) *rectangle* to be a set of the form $A \times B$ where $A \in \mathcal{M}$ and $B \in \mathcal{N}$. As we have just recalled, the collection of finite disjoint unions of rectangles \mathcal{A} is an algebra, and it generates $\mathcal{M} \otimes \mathcal{N}$.

Constructing the product measure is done by constructing first a premeasure. We consider a rectangle $A \times B$ that is finite or countable disjoint union of rectangles $A_j \times B_j$. Using additivity of the integral we find that $\mu(A)\nu(B) = \sum \mu(A_j)\nu(B_j)$. Then we define for an element $E \in \mathcal{A}$ that is the disjoint union of rectangles $A_1 \times B_1, \dots, A_n \times B_n$, the function $\pi(E) = \sum_{j=1}^n \mu(A_j)\nu(B_j)$. Then π is well defined on \mathcal{A} and is a premeasure.

By ??, there is an outer measure on $X \times Y$ whose restriction to $\mathcal{M} \times \mathcal{N}$ is a measure that extends π , which we call the *product measure $\mu \times \nu$ of μ and ν* . Moreover, if μ and ν are σ -finite (they can be seen as a countable union of sets of finite measure), so is $\mu \times \nu$. In this case, $\mu \times \nu$ is the unique measure on $\mathcal{M} \otimes \mathcal{N}$ such that $\mu \times \nu(A \times B) = \mu(A)\nu(B)$ for all rectangles $A \times B$.

The same construction works for a finite number of factor spaces $(X_j, \mathcal{M}_j, \mu_j)$, yielding the product measure $\mu_1 \times \dots \times \mu_n$ on $\mathcal{M}_1 \otimes \dots \otimes \mathcal{M}_n$. And again, if the μ_j are σ -finite, the product measure is uniquely determined.

Let (X, \mathcal{M}, μ) and (Y, \mathcal{N}, ν) be two measure spaces. If $E \subset X \times Y$, define for $x \in E$ and $y \in Y$ the *x-section* E_x and the *y-section* E^y by

$$E_x = \{y \in Y : (x, y) \in E\} \quad E^y = \{x \in X : (x, y) \in E\}.$$

Also, if f is a function on $X \times Y$ we define the *x-section* f_x and the *y-section* f^y of f by

$$f_x(y) = f^y(x) = f(x, y).$$

Proposition 1.54.

1. If $E \in \mathcal{M} \times \mathcal{N}$, then $E_x \in \mathcal{N}$ for all $x \in X$ and $E^y \in \mathcal{M}$ for all $y \in Y$.
2. If f is in $\mathcal{M} \otimes \mathcal{N}$ -measurable, then f_x is \mathcal{N} -measurable for all $x \in X$ and f^y is \mathcal{M} -measurable for all $y \in Y$.

Theorem 1.55. Suppose (X, \mathcal{M}, μ) and (Y, \mathcal{N}, ν) are σ -finite measure spaces. If $E \in \mathcal{M} \otimes \mathcal{N}$, then the functions $x \mapsto \nu(E_x)$ and $y \mapsto \mu(E^y)$ are measurable on X and Y , respectively, and

$$\mu \times \nu(E) = \int \nu(E_x) d\mu(x) = \int \mu(E^y) d\nu(y).$$

Proof. Uses the *Monotone Class Lemma*, and the Monotone Convergence and Dominated Convergence theorems. \square

Theorem 1.56 (Fubini-Tonelli, as in Folland). Suppose that (X, \mathcal{M}, μ) and (Y, \mathcal{N}, ν) are σ -finite measure spaces.

1. **(Tonelli.)** If $f \in L^+(X \times Y)$ = the set of measurable functions from $X \times Y$ to $[0, \infty]$, then the functions

$$g(x) = \int f_x d\nu \quad \text{and} \quad h(y) = \int f^y d\mu$$

are in $L^+(X)$ and $L^+(Y)$ respectively, and

$$\int f d(\mu \times \nu) = \int \left(\int f(x, y) d\nu(y) \right) d\mu(x) = \int \left(\int f(x, y) d\mu(x) \right) d\nu(y). \quad (1.2)$$

2. **(Fubini.)** If $f \in L^1(\mu \times \nu)$, then $f_x \in L^1(\nu)$ for a.e. $x \in X$, $f^y \in L^1(\mu)$ for a.e. $y \in Y$, the a.e.-defined functions g and h are $L^1(\mu)$ and $L^1(\nu)$, respectively, and eq. (1.2) holds.

Proof. By theorem 1.55 and the Monotone Convergence Theorem. \square

Theorem 1.57 (Tonelli, as in Brezis). Let $F(x, y) : \Omega_1 \times \Omega_2 \rightarrow \mathbb{R}$ be a measurable function satisfying

$$\int_{\Omega_1} |F(x, y)| d\mu_2 < \infty \quad \text{for a.e. } x \in \Omega_1$$

and

$$\int_{\Omega_1} d\mu_1 \int_{\Omega_2} |F(x, y)| d\mu_2 < \infty.$$

Then $F \in L^1(\Omega_1 \times \Omega_2)$.

Roughly speaking: if a slice of F is integrable, then F is integrable.

Theorem 1.58 (Fubini, as in Brezis). Assume that $F \in L^1(\Omega_1 \times \Omega_2)$. Then for a.e. $x \in \Omega_1$, $F(x, y) \in L^1_y(\Omega_2)$ and $\int_{\Omega_2} F(x, y) d\mu_2 \in L^1_x(\Omega_1)$. Similarly, for a.e. $y \in \Omega_2$, $F(x, y) \in L^1_x(\Omega_1)$ and $\int_{\Omega_1} F(x, y) d\mu_1 \in L^1_y(\Omega_2)$.

Moreover, one has

$$\int_{\Omega_1} d\mu_1 \int_{\Omega_2} F(x, y) d\mu_2 = \int_{\Omega_2} d\mu_2 \int_{\Omega_1} F(x, y) d\mu_1 = \iint_{\Omega_1 \times \Omega_2} F(x, y) d\mu_1 d\mu_2.$$

Remark 1.9. These theorems show that the integral of a function over a cartesian product may be obtained by fixing one variable, integrating the resulting one-variable function, and then varying the first variable.

Remark 1.10. The difference between the theorems appears to concern whether or not we suppose that the function is non-negative.

Theorem 1.59 (2.44).

$$\int f(x) dx = |\det T| \int f \circ T(x) dx$$

Theorem 1.60 (2.47, diffeomorphisms). content...

2 Point set topology

2.1 Metric spaces

A **metric** on a set X is a function $\rho : X \times X \rightarrow [0, \infty)$ such that

1. $\rho(x, x) = 0$ if and only if $x = 0$.
2. $\rho(x, y) = \rho(y, x)$ for all $x, y \in X$.
3. $d(x, z) \leq d(x, y) + d(y, z)$ for all $x, y, z \in X$.

Two metrics ρ_1 and ρ_2 on a set X are **equivalent** if $C\rho_1 \leq \rho_2 \leq C'\rho_1$ for some $C, C' > 0$.

Theorem 2.1. Let (X, d) and (Y, d') be metric spaces. If $f : X \rightarrow Y$, the following are equivalent conditions for f to be **continuous**:

1. $\forall x \in X \forall \varepsilon > 0 \exists \delta > 0 : f(B_\delta(x)) \subseteq B_\varepsilon(f(x))$
2. $\forall x \in X \forall x_n \rightarrow x : f(x_n) \rightarrow f(x)$.
3. $\forall F \subseteq Y$ open, $f^{-1}(F)$ is open.
4. $\forall F \subseteq Y$ closed, $f^{-1}(F)$ is closed.

If $f : (X, \rho) \rightarrow (Y, \rho')$ is such that

$$\forall \varepsilon > 0 \exists \delta > 0 \forall x \in X : f(B_\delta(x)) \subseteq B_\varepsilon(f(x))$$

we say f is **uniformly continuous**.

Exercise. If (Y, ρ') is complete and $f : A \rightarrow Y$ is uniformly continuous on $A \subset X$ and $\overline{A} = X$, then f has a unique continuous extension $g : X \rightarrow Y$ which is uniformly continuous on X . Show that this is not true in general if Y is not complete.

If $f : (X, \rho) \rightarrow (Y, \rho')$ is a bijective function such that for any $x, y \in X$, $\rho(x, y) = \rho'(f(x), f(y))$ we say f is an **isometry** and the two spaces are **isometric**. A function $f : (X, \rho) \rightarrow (X, \rho)$ is a **contraction** if there exists a $0 < a < 1$ such that $\rho(f(x), f(y)) \leq a\rho(x, y)$ for any $x, y \in X$. Every contraction is continuous, and if X is complete then any contraction has a unique fixed point.

A sequence $\{x_n\}$ in X **converges** to x if $\lim_{n \rightarrow \infty} \rho(x_n, x) = 0$. A sequence $\{x_n\}$ in X is called **Cauchy** if $\lim_{n, m \rightarrow \infty} \rho(x_n, x_m) = 0$, that is

$$\forall \varepsilon > 0 \exists N \in \mathbb{N} \forall n, m > N : \rho(x_n, x_m) < \varepsilon.$$

A subset $E \subseteq X$ is **complete** if every Cauchy sequence in E converges to its limit in E . If (X, ρ) and (X^*, ρ^*) are metric spaces and

1. (X, ρ) is isometric to a subspace (X, ρ^*) of (X^*, ρ^*) ,
2. The closure of X_0 is all of X^* (X_0 is **everywhere dense** or simply **dense**),

we say (X^*, ρ^*) is the *completion* of (X, ρ) .

Theorem 2.2. Every metric space (X, ρ) has a completion (X^*, ρ^*) . If (X^{**}, ρ^{**}) is also a completion of (X, ρ) , then (X^*, ρ^*) is isometric to (X^{**}, ρ^{**}) ; that is, the completion of a space is unique up to isometry.

Proof. Consider equivalence classes of Cauchy sequences. □

Proposition 2.3. A closed subset of a metric space is complete, and a complete subset of an arbitrary metric space is closed.

2.2 Topological spaces

Let X be a nonempty set. A *topology* on X is a family \mathcal{T} of subsets of X that contains \emptyset and X and is closed under arbitrary unions and finite intersections. (That is, if $\{U_\alpha\}_{\alpha \in A} \subset \mathcal{T}$ then $\bigcup_{\alpha \in A} U_\alpha \in \mathcal{T}$ and if $U_1, \dots, U_n \in \mathcal{T}$ then $\bigcap_{j=1}^n U_j \in \mathcal{T}$.) The pair (X, \mathcal{T}) is called a *topological space*.

The members of \mathcal{T} are called *open sets* and their complements are called *closed sets*. If $A \subseteq X$, the union of all open sets contained in A is called the *interior* of A , denoted by A° ; so that a point $x \in A^\circ$ when there is an open set O contained in A such that $x \in O$. The intersection of all closed sets containing A is called the *closure* of A . The difference $\bar{A} \setminus A^\circ = \bar{A} \cap \bar{A}^c$ is called the *boundary* of A and is denoted by ∂A .

Following Bachman and Narici, *Functional Analysis*, a point $x \in X$ is an *adherence point* of the subset E if every open set containing x contains a point of E . The set of all adherence points of E is called the *closure* of E and it is denoted by \bar{E} . It is immediate that $E \subset \bar{E}$.

Proposition 2.4. Let X be a topological space and A and B subsets of X . Then

1. $A \subset B \implies \bar{A} \subset \bar{B}$.
2. $\overline{A \cup B} \implies \bar{A} \cup \bar{B}$,
3. $\bar{A} = \bar{\bar{A}}$.
4. A subset F of X is closed if and only if $F = \bar{F}$.

If X is a topological space and A is a subset of X , a point $x \in X$ is called a *limit point* of A if every open set containing x contains a point of A distinct from x . The set of all limit points of A is denoted by A' and is called the *derived set* of A . Clearly, $\bar{A} = A \cup A'$, and, in view of the last item in the last proposition, A is closed if and only if $A' \subset A$.

A sequence of points $\{x_n\}$ in a topological space X *converges* to the point $x \in X$ if for every open set O containing x there exists an index N (depending on O) such that $x_n \in O$ for all $n > N$. That is, every open set containing x must contain almost all (but a finite number) of the x_n .

Proposition 2.5. Let X_1, X_2 and X_3 be topological spaces. Let $f : X_1 \rightarrow X_2$ and $g : X_2 \rightarrow X_3$ be mappings.

1. If f and g are continuous, then the composite mapping gf is continuous.
2. f is continuous if and only if for every subset A of X , $f\bar{A} \subset \overline{f(A)}$.
3. Suppose f is a 1:1, onto mapping. Then f is a homeomorphism of X_1 onto X_2 is and only if, for all subsets A of X_1 , $f(\bar{A}) = \overline{f(A)}$.

If \mathcal{T}_1 and \mathcal{T}_2 are two topologies on a set X , we say \mathcal{T}_1 is *weaker* (or *coarser*) and \mathcal{T}_2 *stronger* (or *finer*) if $\mathcal{T}_1 \subseteq \mathcal{T}_2$. $E \subseteq X$ is called *dense* if $\bar{E} = X$ and *nowhere dense* if \bar{E} has empty interior. X is called *separable* if it has a countable dense subset.

- T_0 If $x \neq y$, there is an open set containing x but not y , or an open set containing y but not x .
- T_1 If $x \neq y$, there is an open set containing y but not x . Equivalently, $\{x\}$ is closed for every $x \in X$.
- T_2 (**Hausdorff.**) If $x \neq y$ there are disjoint open sets U and V such that $x \in U$ and $y \in V$.
- T_3 (**Regular.**) X is T_1 and for any closed set $A \subset X$ and any $x \in A^c$ there are disjoint open sets U, V with $x \in U$ and $A \subseteq V$.
- $T_{3\frac{1}{2}}$ (**Tychonoff, Completely regular.**) X is T_1 and for each closed $A \subseteq X$ and each $x \notin A$ there exists $f \in C(X, [0, 1])$ such that $f(x) = 1$ and $f = 0$ on A .
- T_4 (**Normal.**) X is T_1 and for any disjoint closed sets A, B in X there are disjoint open sets U, V with $A \subseteq U$ and $B \subseteq V$.

If X is any set and $\{f_\alpha : X \rightarrow Y_\alpha\}_{\alpha \in A}$ is a family of maps from X into some topological spaces Y_α , there is a unique weakest topology τ on X that makes all the f_α continuous called the *weak topology generated by* $\{f_\alpha\}_{\alpha \in A}$. An example of this topology is the *product topology* on $X = \prod_{\alpha \in A} X_\alpha$ with the projections.

Proposition 2.6.

- If X_α is Hausdorff for each $\alpha \in A$ then $X = \prod_{\alpha \in A} X_\alpha$ is Hausdorff.
- If X_α and Y are topological spaces, a function $f : Y \rightarrow X = \prod_{\alpha \in A} X_\alpha$ is continuous if and only iff $\pi_\alpha \circ f$ is continuous for each α .
- If X is a topological space, A is a nonempty set and $\{f_n\}$ is a sequence in X^A , then $f_n \rightarrow f$ in the product topology if and only if $f_n \rightarrow f$ pointwise.

If X is any set and $K = \mathbb{R}$ or \mathbb{C} , denote by $B(X, K)$ the *set of bounded K -valued functions on X* , $C(X, K)$ the set of *continuous K -valued functions on X* , and $BC(X, K)$ the *set of bounded continuous functions on X* . If no field is specified we take it to be \mathbb{C} .

For $f \in B(X)$ define the *uniform norm* of f to be

$$\|f\|_u = \sup\{|f(x)| : x \in X\}$$

Then the function $\rho(f, g) = |Vert f - g|_{Vert_u}$ is a metric on $B(X)$. Convergence in this metric is simply uniform convergence:

$$\{f_n\} \xrightarrow{u} f \iff \forall \varepsilon > 0 \exists N \in \mathbb{N} \forall n > N \forall x \in X : |f_n(x) - f(x)| < \varepsilon$$

$B(X)$ is complete with this metric since \mathbb{C} is complete.

Proposition 2.7. If X is a topological space, $BC(X)$ is a closed subspace of $B(X)$ in the uniform metric; in particular $BC(X)$ is complete.

Lemma 2.8 (Urysohn). Let X be a normal space. If A and B are disjoint closed sets in X , there exists $f \in C(X, [0, 1])$ such that $f = 0$ on A and $f = 1$ on B .

Theorem 2.9 (Tietze Extension Theorem). Let X be a normal space. If A is a closed subset of X and $f \in C(A, [a, b])$, there exists $F \in C(X, [a, b])$ such that $F|_A = f$.

Corollary 2.10. If X is normal, $A \subseteq X$ is closed and $f \in C(A)$, there exists $F \in C(X)$ such that $F|_A = f$.

Urysohn's lemma shows that every T_4 space is completely regular ($T_{3\frac{1}{2}}$).

Theorem 2.11 (Dugundji). Sea X un espacio metrizable, $A = \bar{A} \subset X$ y L un espacio vectorial localmente convexo, y $V \subset L$ convexo. Entonces cualquier función $f : A \rightarrow V$ admite una extensión F .

$$\begin{array}{ccc} A & \xrightarrow{f} & V \\ \downarrow & \nearrow F & \\ X & & \end{array}$$

Además $\text{img } F \subset \text{conv img } f$.

2.3 Compact spaces

A topological space X is called **compact** if whenever $\{U_\alpha\}_{\alpha \in A}$ is an open cover of X there is a finite subset B of A such that $X = \bigcup_{\alpha \in B} U_\alpha$. A subset $Y \subseteq X$ is called **compact** if it is compact in the relative topology and **precompact** if its closure is compact.

A family $\{F_\alpha\}_{\alpha \in A}$ of subsets of X has the **finite intersection property** if $\bigcap_{\alpha \in B} F_\alpha \neq \emptyset$ for all finite $B \subseteq A$.

Proposition 2.12.

- A topological space X is compact if and only if for every family $\{F_\alpha\}_{\alpha \in A}$ of closed sets with the finite intersection property, $\bigcap_{\alpha \in A} F_\alpha \neq \emptyset$.
- A closed subset of a compact space is compact.
- If K is a compact subset of a Hausdorff space X and $x \notin K$ then there are disjoint open sets U, V such that $x \in U$ and $K \subseteq V$.
- Every compact subset of a Hausdorff space is closed.

- Every compact Hausdorff space is normal.
- If X is compact and $f : X \rightarrow Y$ is continuous then $f(X)$ is compact.
- If X is compact, then $C(X) = BC(X)$.
- If X is compact and Y is Hausdorff, then any continuous bijection $f : X \rightarrow Y$ is an homeomorphism.

A topological space X is *countably compact* if every countable open cover of X has a finite subcover, and *sequentially compact* if every sequence in X has a convergent subsequence. For metric spaces compactness and sequential compactness are the equivalent. There exists no general relation between compactness and sequential compactness.

Theorem 2.13 (Heine-Borel). A subset of \mathbb{R}^n is compact if and only if it is closed and bounded.

Theorem 2.14 (Bolzano-Weierstrass). Every bounded sequence in \mathbb{R}^n has a convergent subsequence. Equivalently, a subset of \mathbb{R}^n is sequentially compact if and only if it is closed and bounded.

Theorem 2.15. If E is a subset of a metric space (X, ρ) , the following are equivalent:

1. E is complete and *totally bounded* (it can be covered by finitely many balls of radius ε).
2. Every sequence in E has a subsequence that converges to a point in E .
3. If $\{V_\alpha\}_{\alpha \in A}$ is an open cover of E , then there is a finite subset $F \subseteq A$ such that $\{V_\alpha\}_{\alpha \in F}$ covers E .

Theorem 2.16. If (X, ρ) is a metric space and A is compact, then A is closed and bounded.

We conclude with the approach of Narici.

If (X, ρ) is a metric space, $A \subseteq X$ is *relatively compact* if \overline{A} is compact. If $\varepsilon > 0$, a subset $N \subset X$ is an ε -*net with respect to* A if $\forall x \in A \exists n \in N : \rho(x, n) < \varepsilon$. A is *totally boundad* if for any $\varepsilon > 0$ there exists a finite ε -net with respect to A .

Theorem 2.17. Let (X, ρ) be a metric space and $A \subseteq X$. If for every sequence of points from A one can select a convergent subsequence, then A is totally bounded.

A set A is *countably compact* if every infinite subset of A has a limit point in A . All compact sets are countably compact. A is *sequentially compact* if every sequence in A has a subsequence that converges to a point in A . In a metric space, compactness is equivalent to countable and sequential compactness.

Theorem 2.18. Let (X, ρ) be a metric space and $A \subseteq X$.

1. A is relatively compact if and only if a convergent subsequence can be selected from every sequence of points in A . (We do not claim that the limit point is a member of A .)
2. If A is relatively compact, it is also totally bounded.

3. If (X, ρ) is complete and A is totally bounded, then A is relatively compact.
4. If A is compact then A is closed and totally bounded.

2.4 Locally Compact Hausdorff spaces

A topological space is called *locally compact* if every point has a compact neighbourhood (a set $A \subset X$ such that $x \in A^\circ$). We call locally compact Hausdorff spaces *LCH* for short.

Proposition 2.19. Let X be a LCH space.

- If $U \subseteq X$ is open and $x \in U$, there is a compact neighbourhood K of x such that $K \subset U$.
- If $K \subseteq U \subseteq X$, with K compact and U open, there exists a precompact open V such that $K \subseteq V \subset \overline{V} \subset U$.
- **(Urysohn's Lemma, Locally Compact Version.)** If $K \subset U \subseteq X$, there exists $f \in C(X, [0, 1])$ such that $f = 1$ on K and $f = 0$ outside a compact subset of U .
- Every LCH space is completely regular.
- **(Tietze Extension Theorem, Locally Compact Version)** If $K \subseteq X$ is compact and $f \in C(K)$, there exists $F \in C(X)$ such that $F|_K = f$. F may be taken to vanish outside a compact set.

If $f \in C(X)$, the *support* of f is the closure of $\{x \in X : f(x) \neq 0\}$ and denote $C_c(X) := \{f \in C(X) : \text{supp } f \text{ is compact}\}$. We say f *vanishes at infinity* if for every $\varepsilon > 0$ the set $\{x : |f(x)| \geq \varepsilon\}$ is compact and define $C_0(X) := \{f \in C(X) : f \text{ vanishes at infinity}\}$.

Proposition 2.20. If X is an LCH space, $C_0(X)$ is the closure of $C_c(X)$ in the uniform metric.

If X is a topological space, there are many ways of topologizing \mathbb{C}^X . One way is the product topology, that is, the topology of pointwise convergence. Another is the *topology of uniform convergence*, which is generated by the sets

$$\left\{ g \in \mathbb{C}^X : \sup_{x \in X} |g(x) - f(x)| < n^{-1} \right\} \quad n \in \mathbb{N}, f \in \mathbb{C}^X.$$

In view of a previous proposition (cite?), we know $C(X)$ is a closed subset of \mathbb{C}^X with the topology of uniform convergence. Another topology is the *topology of uniform convergence on compact sets*, generated by the sets

$$\left\{ g \in \mathbb{C}^X : \sup_{x \in K} |g(x) - f(x)| < n^{-1} \right\} \quad n \in \mathbb{N}, f \in \mathbb{C}^X, K \subseteq X \text{ compact}.$$

Proposition 2.21. Let X be an LCH space.

- If $E \subseteq X$, then E is closed if and only if $E \cap K$ is closed for every compact $K \subseteq X$.
- $C(X)$ is a closed subspace of \mathbb{C}^X in the topology of uniform convergence on compact sets.
- If $\{U_j\}_{j=1}^n$ is an open cover of a compact subset K of X , then there is a partition of unity on K subordinate to $\{U_j\}_{j=1}^n$ consisting of compactly supported functions.

Theorem 2.22 (Urysohn Metrization Theorem). Every second countable normal space is metrizable.

2.5 Three compactness theorems

Recall that if $X = \prod_{\alpha \in A} X_\alpha$, an element $x \in X$ is just a mapping from A to $\bigcup_{\alpha \in A} X_\alpha$, with $x(\alpha)$ the α th coordinate of x .

Theorem 2.23. If $\{X_\alpha\}_{\alpha \in A}$ is a family of compact topological spaces, then $X = \prod_{\alpha \in A} X_\alpha$ is compact with the product topology.

Let X be a topological space and $\mathcal{F} \subseteq C(X)$ a family of complex-valued continuous functions on X . We say \mathcal{F} is **equicontinuous at** $x \in X$ if for every $\varepsilon > 0$ there is a neighbourhood U of x such that $|f(x) - f(y)| < \varepsilon$ for all $y \in U$ and all $f \in \mathcal{F}$; and **equicontinuous** if it is equicontinuous at every $x \in X$. Also, \mathcal{F} is **pointwise bounded** if $\{|f(x)| : f \in \mathcal{F}\}$ is bounded for all $x \in X$.

Theorem 2.24 (Arzelá-Ascoli I). Let X be a compact Hausdorff space. If \mathcal{F} is an equicontinuous, pointwise bounded subset of $C(X)$, then \mathcal{F} is totally bounded in the uniform metric, and the closure of \mathcal{F} in $C(X)$ is compact.

Theorem 2.25 (Arzelá-Ascoli II). Let X be a locally compact Hausdorff space. If $\{f_n\}$ is an equicontinuous, pointwise bounded sequence in $C(X)$, then there exists $f \in C(X)$ and a subsequence of $\{f_n\}$ that converges to f uniformly on compact sets.

2.6 The Stone-Weierstrass Theorem

Recall that the Weierstrass theorem states that any continuous function on a compact interval $[a, b]$ is the uniform limit of polynomials on $[a, b]$. Throughout this subsection, X will denote a compact Hausdorff space, and $C(X)$ is equipped with the uniform metric.

A subset \mathcal{A} of $C(X, \mathbb{R})$ of $C(X)$ is said to **separate points** if for every $x, y \in X$ with $x \neq y$ there exists $f \in \mathcal{A}$ such that $f(x) \neq f(y)$. \mathcal{A} is called an **algebra** if it is a real (resp. complex) vector subspace of $C(X, \mathbb{R})$ (resp. $C(X)$) such that $fg \in \mathcal{A}$ whenever $f, g \in \mathcal{A}$. \mathcal{A} is called a **lattice** if $\max(f, g)$ and $\min(f, g)$ are in \mathcal{A} whenever $f, g \in \mathcal{A}$. If \mathcal{A} is an algebra or a lattice, so is its closure in the uniform metric.

Theorem 2.26 (Stone-Weierstrass Theorem). Let X be a compact Hausdorff space. If \mathcal{A} is a closed subalgebra of $C(X, \mathbb{R})$ that separates points, then either $\mathcal{A} = C(X, \mathbb{R})$ or $\mathcal{A} = \{f \in C(X, \mathbb{R}) : f(x_0) = 0\}$ for some $x_0 \in X$. The first alternative holds if and only if \mathcal{A} contains the constant functions.

Corollary 2.27. Suppose \mathcal{B} is a subalgebra of $C(X, \mathbb{R})$ that separates points. If there exists $x_0 \in X$ such that $f(x_0) = 0$ for all $f \in \mathcal{B}$, then \mathcal{B} is dense in $\{f \in C(X, \mathbb{R}) : f(x_0) = 0\}$. Otherwise, \mathcal{B} is dense in $C(X, \mathbb{R})$.

The classical Weierstrass approximation theorem is the special case of this corollary where X is the compact subset of \mathbb{R}^n and \mathcal{B} is the algebra of polynomials on \mathbb{R}^n (restricted to X); here \mathcal{B} contains the constant functions, so it is dense in $C(X, \mathbb{R})$.

The Stone-Weierstrass theorem, as stated, is false for complex-valued functions. We may show that $f(z) = \bar{z}$ cannot be approximated uniformly by polynomials on the unit circle.

Theorem 2.28 (Complex Stone-Weierstrass Theorem). Let X be a compact Hausdorff space. If \mathcal{A} is a closed complex subalgebra of $C(X)$ that separates points and is closed under complex conjugation, then either $\mathcal{A} = C(X)$ or $\mathcal{A} = \{f \in C(X) : f(x_0) = 0\}$ for some $x_0 \in X$.

Finally, there is a version of the Stone-Weierstrass theorem for noncompact LCH spaces. We state for real functions; the complex analogue is an immediate consequence.

Theorem 2.29 (LCH Stone-Weierstrass Theorem). Let X be a noncompact LCH space. If \mathcal{A} is a closed complex subalgebra of $C_0(X, \mathbb{R})$ that separates points, then either $\mathcal{A} = C_0(X, \mathbb{R})$ or $\mathcal{A} = \{f \in C_0(X, \mathbb{R}) : f(x_0) = 0\}$ for some $x_0 \in X$.

3 Inner product spaces

3.1 Inner products

Let X be a real or complex vector space. An *inner product* on X is a mapping

$$\langle -, - \rangle : X \times X \rightarrow F$$

with the following properties:

- (I₁) if $x, y \in X$ then $\langle x, y \rangle = \overline{\langle y, x \rangle}$;
- (I₂) if α, β are scalars, $\langle \alpha x + \beta y, z \rangle = \alpha \langle x, z \rangle + \beta \langle y, z \rangle$;
- (I₃) $\langle x, x \rangle \geq 0$ for all $x \in X$ and equal to zero if and only if x is the zero vector. (Since, by I₁, $\langle x, x \rangle$ must be real.)

Examples.

- Let $X = C[a, b]$ be complex-valued continuous functions on the closed interval $[a, b]$ with pointwise addition and scalar product. As the inner product of any two vectors f and g in this space take

$$\langle f, g \rangle = \int_a^b f(x) \overline{g(x)} dx$$

- Let $X = l_2$, the set of all sequences of complex numbers (a_1, a_2, \dots) with the property that $\sum_{i=1}^{\infty} |a_i|^2 < \infty$. As the inner product of any two vectors $x = (a_i)$ and $y = (b_i)$ in this space take

$$\langle f, g \rangle = \sum_{i=1}^{\infty} a_i \bar{b}_i$$

which converges by the Hölder inequality.

- Let Y be the closed interval $[a, b]$, S the Lebesgue measurable sets and μ the Lebesgue measure. Then, for the equivalence classes of square-integrable functions (complex-valued) on $[a, b]$ we can take as the inner product of two classes f and g ,

$$\langle f, g \rangle = \int_a^b f(x) \overline{g(x)} dx$$

where the integral is the Lebesgue integral. This space is denoted by $L_2(a, b)$.

Theorem 3.1 (Cauchy-Schwarz inequality). Let X be an inner product space and let $x, y \in X$. Then

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

with equality holding if and only if x and y are linearly independent.

3.2 Orthogonal projections

Two vectors $x, y \in X$ are **orthogonal** if $\langle x, y \rangle = 0$.

Examples.

1. In $L_2(-\pi, \pi)$, the collection (or any subset thereof)

$$x_n = \frac{1}{\sqrt{2\pi}} e^{int}, \quad n = 0, \pm 1, \dots$$

is an orthonormal set of vectors.

Proof. For any $n \in \mathbb{Z}$,

$$\int_{-\pi}^{\pi} x_n \bar{x}_n dt = \frac{1}{2\pi} \int_{-\pi}^{\pi} e^{int} \overline{e^{int}} dt = \frac{1}{2\pi} \int_{-\pi}^{\pi} \|e^{int}\|^2 dt = 1,$$

and if m is another integer,

$$\int_{-\pi}^{\pi} x_n \bar{x}_m dt = \frac{1}{2\pi} \int_{-\pi}^{\pi} e^{int} \overline{e^{imt}} dt = \frac{1}{2\pi} \int_{-\pi}^{\pi} e^{(n-m)it} dt = \frac{1}{2\pi} \left[\frac{e^{(n-m)it}}{(n-m)i} \right]_{-\pi}^{\pi} = 0.$$

□

2. If we restrict our attention to only real-valued functions that are square-integrable on the interval $[-\pi, \pi]$, then the collection (or any subset thereof)

$$\begin{aligned} & \frac{1}{\sqrt{2\pi}}, \frac{1}{\sqrt{\pi}} \cos t, \frac{1}{\sqrt{\pi}} \cos 2t, \dots \\ & \frac{1}{\sqrt{\pi}} \sin t, \frac{1}{\sqrt{\pi}} \sin 2t, \dots \end{aligned}$$

is an orthonormal set.

Theorem 3.2. If S is an orthonormal subset of an inner product space, then it is linearly independent (where linear independence is defined as finite sums).

Theorem 3.3 (Gram-Schmidt process). Let X be an inner product space. If $\{y_1, y_2, \dots\}$ is a linearly independent set of vectors, then there exists an orthonormal set of vectors $\{x_1, x_2, \dots\}$ such that, for any n ,

$$\langle y_1, y_2, \dots, y_n \rangle = \langle x_1, x_2, \dots, x_n \rangle$$

where the brackets indicate the subspace spanned by the vectors enclosed.

If S is any subset of X , the **orthogonal complement of S in X** is the linear space $S^\perp := \{x \in X : x \perp s \text{ for all } s \in S\}$.

Theorem 3.4. If M is a finite-dimensional subspace of X , then $X = M \oplus M^\perp$.

3.3 Riesz representation theorem

Theorem 3.5 (Riesz). If X is a finite-dimensional inner product space and f is a linear functional on X , then there exists a unique vector $y \in X$ such that $f(x) = \langle x, y \rangle$ for all $x \in X$.

Proof. Given an orthonormal basis e_i of X , consider $y = \sum_i \overline{f(e_i)} e_i$. □

In Riemannian geometry this is called *raising an index* of a 1-form. Indeed, $\omega_p \in \Lambda^1(T_p M)$ is just a linear functional on $T_p M$, and $(\omega)^{\sharp} = g^{ij} \omega_j E_i$ at p is just a vector y such that $\omega_p(x) = \langle x, y \rangle$ for all $x \in T_p M$. So the former theorem may also be stated as “ $y = f^{\sharp}$ exists”. Recall this is given by viewing the inner product as a nonsingular matrix.

3.4 Adjoint operator

Let $A : X \rightarrow X$ be a linear transformation in a finite-dimensional inner product space X . For a given $y \in X$, define the linear functional

$$\begin{aligned} f^y : X &\rightarrow F \\ x &\mapsto \langle Ax, y \rangle \end{aligned}$$

which, by the Riesz representation theorem yields a unique $z \in X$ such that

$$f^y(x) = \langle x, z \rangle$$

Then the *adjoint of A* is the linear map

$$\begin{aligned} A^* : X &\rightarrow X \\ y &\mapsto z \end{aligned}$$

so that $\langle Ax, y \rangle = \langle x, A^*y \rangle$.

Proposition 3.6 (Properties of the adjoint).

1. $(\alpha A)^* = \overline{\alpha} A^*$.
2. $(A + B)^* = A^* + B^*$.
3. $(AB)^* = B^* A^*$.
4. $(A^*)^* = A$.

If $A = A^*$ we say A is *self-adjoint*, and if $AA^* = A^*A$ we say A is *normal*.

Theorem 3.7. If A is self-adjoint, its eigenvalues are real. Eigenvectors associated to distinct eigenvalues of a self-adjoint operator are orthogonal.

Theorem 3.8. If M is an invariant subspace of X under A , then M^{\perp} is invariant under A^* .

Theorem 3.9. If A is a linear transformation on a finite-dimensional inner product space X , then $\text{range}(A)^{\perp} = \text{null}(A^*)$.

3.5 Spectral theorem for normal transformations

Theorem 3.10. Let A be a self-adjoint transformation in a finite-dimensional inner product space X . Then there exists an orthonormal basis of X consisting of eigenvectors of A .

Lemma 3.11. Let A be a normal transformation in a finite-dimensional inner product space X . Then $\|Ax\| = \|A^*x\|$ for all $x \in X$.

Theorem 3.12. Let A be a normal transformation in a complex finite-dimensional inner product space X . Then there exists an orthonormal basis of X consisting of eigenvectors of A .

Theorem 3.13. If A is a normal transformation on a finite-dimensional inner product space. Eigenvectors associated to distinct eigenvalues of a self-adjoint operator are orthogonal.

Recall that the notation $X = M_1 \oplus \dots \oplus M_k$ means that X is the **direct sum** of the M_i , which means that $X = M_1 + \dots + M_k$ and $M_i \cap \{M_1 + \dots + M_i + \dots + M_k\} = \{0\}$, (every element in X is expressed as a unique sum of elements in M_i). If $M_i \perp M_j$ for all $i \neq j$, we say this is an **orthogonal direct sum decomposition of X** , and the **orthogonal projection to M_j** is just taking the corresponding component of a given element in its decomposition.

Theorem 3.14 (Spectral decomposition theorem for normal transformations). To every normal transformation A on a complex finite-dimensional inner product space there correspond scalar $\lambda_1, \dots, \lambda_k$, the distinct eigenvalues of A , and orthogonal projections E_1, \dots, E_k with $k \leq \dim X$, such that

1. E_i is the orthogonal projection on $\text{Null}(A - \lambda_i)$ for $i = 1, \dots, k$.
2. $E_i \neq 0$ and $E_i E_j = 0$ for $i, j = 1, \dots, k$.
3. $\sum_{j=1}^k E_j = 1$.
4. $\sum_{j=1}^k \lambda_j E_j = A$.

If A was self-adjoint, we could weaken the hypotheses to a real inner product space.

3.6 Unitary and orthogonal transformations

Let X be a finite-dimensional inner product space, and $U : X \rightarrow X$ a linear transformation with $U^*U = 1$. We say U is **unitary** if X is complex and **orthogonal** if X is real. The condition $U^*U = 1$ implies that $UU^* = 1$.

Theorem 3.15. Let X be a finite-dimensional inner product space, and $U : X \rightarrow X$ a linear transformation. The following statements are equivalent:

1. $U^*U = 1$.
2. $\langle Ux, Uy \rangle = \langle x, y \rangle$.

3. $\|Ux\| = \|x\|$ for all $x \in X$.

Theorem 3.16. If U is a unitary transformation on the finite-dimensional inner product space X , then each of the eigenvalues of U must have an absolute value equal to 1.

To summarize:

Theorem 3.17. Let A be a normal transformation on a complex finite-dimensional inner product space. Then

1. A is self-adjoint if and only if each eigenvalue of A is real.
2. A is unitary if and only if each eigenvalue of A has absolute value equal to 1.

3.7 Normed spaces

Let X be a real or complex vector space. A **norm** on X is a function $\|\cdot\| : X \rightarrow \mathbb{R}$ such that

1. $\|x\| \geq 0$ and $\|x\| = 0$ if and only if $x = 0$.
2. $\|\lambda x\| = |\lambda| \|x\|$ for all $x \in X$ and $\lambda \in \mathbb{R}$.
3. **(Triangle inequality.)** $\|x + y\| \leq \|x\| + \|y\|$ for all $x, y \in X$.

Every normed space is a metric space with the distance function $\rho(x, y) = \|x - y\|$. Two norms $\|\cdot\|_1$ and $\|\cdot\|_2$ are called **equivalent** if there exist $C_1, C_2 > 0$ such that

$$C_1 \|x\|_1 \leq \|x\|_2 \leq \|x\|_1 \quad \forall x \in X$$

Equivalent norms define the same topology and the same Cauchy sequences.

A normed space that is complete is called a **Banach space**.

Theorem 3.18. For every normed linear space X there is a complete normed linear space X^* such that X is **congruent** (isomorphic and isometric) to a dense subset of X^* and the norm on X^* extends the norm on X .

If $\{x_n\}$ is a sequence in X , the series $\sum_{n=1}^{\infty} x_n$ **converges to** x if $\sum_{n=1}^N x_n \rightarrow x$ as $N \rightarrow \infty$, and it is **absolutely convergent** if $\sum_{n=1}^{\infty} \|x_n\| < \infty$.

Theorem 3.19. A normed vector space X is complete if and only if every absolutely convergent series in X converges.

Examples.

- If X is a topological space, $B(X)$ and $BC(X)$ are Banach spaces with the uniform norm $\|f\|_u = \sup_{x \in X} |f(x)|$.
- If (X, \mathcal{M}, μ) is a measure space, $L^1(\mu)$ is a Banach space with the norm $\|f\|_1 = \int |f| d\mu$. (Observe that $\|\cdot\|_1$ is only a seminorm if we do not identify functions that are equal a.e.)

If X and Y are normed vector spaces, $X \times Y$ is a normed vector space with the **product norm**, $\|(x, y)\| = \max(\|x\|, \|y\|)$. If M is a vector subspace of X , the quotient space X/M consisting of equivalence classes under $x \sim y$ iff $x - y \in M$ is a normed space with the **quotient norm**, $\|x + M\| = \inf_{y \in M} \|x + y\|$.

A linear map $T : X \rightarrow Y$ between two normed vector spaces is **bounded** if there exists $C \geq 0$ such that

$$\|Tx\| \leq C\|x\| \quad \forall x \in X$$

Proposition 3.20. If X and Y are normed vector spaces and $T : X \rightarrow Y$ is a linear map, then T is continuous if and only if it is bounded.

Proof. (\implies) There exists $\delta > 0$ such that $\|x\| \leq \delta$ implies $\|Tx\| \leq 1$. For any nonzero $x \in X$,

$$\|Tx\| = \left\| \frac{\|x\|}{\delta} T\left(\delta \frac{x}{\|x\|}\right) \right\| \leq \frac{1}{\delta} \|x\|.$$

(\impliedby) If $\|x - y\| < \frac{\varepsilon}{C}$,

$$\|T(x) - T(y)\| = \|T(x - y)\| \leq C\|x - y\| < \varepsilon.$$

□

In fact, if T is bounded it is uniformly continuous and even Lipschitz continuous.

We denote by $L(X, Y)$ the space of bounded linear maps from X to Y , which is a normed vector space with the **operator norm**

$$\begin{aligned} \|T\| &= \sup_{\substack{x \in X \\ \|x\| \leq 1}} \|Tx\| \\ &= \sup\{\|Tx\| : \|x\| = 1\} \\ &= \sup\left\{ \frac{\|Tx\|}{\|x\|} : x \neq 0 \right\} \\ &= \inf\{C : \|Tx\| \leq C\|x\| \text{ for all } x \in X\} \end{aligned}$$

Proposition 3.21. If Y is complete, so is $L(X, Y)$.

Proof. Y since $\|T_n x - T_m x\| \leq \|T_n - T_m\| \|x\|$. Define pointwise $Tx = \lim_{n \rightarrow \infty} T_n x$. □

T is **invertible** if it is bijective and T^{-1} is bounded. It is called an isometry if $\|Tx\| = \|x\|$ for all $x \in X$. An isometry is injective but not necessarily surjective.

If X is a vector space over $K = \mathbb{R}, \mathbb{C}$, a **linear functional** is a linear map from X to K .

Proposition 3.22 (Relationship between real and complex linear functionals). Let X be a vector space over \mathbb{C} . If f is a complex linear functional on X , $u := \operatorname{Re} x$ is a real linear functional and $f(x) = u(x) - iu(ix)$. Conversely, if u is a real functional on X , then $f(x) := u(x) - iu(ix)$ is a complex linear functional, and if X is normed, $\|u\| = \|f\|$.

4 Hahn-Banach theorems

4.1 Analytic form of the Hahn-Banach theorem

It is not obvious that there are any nonzero bounded linear functionals on an arbitrary normed vector space. If E is a real vector space, **sublinear** or **Minkoswky functional** on E is a map $p : E \rightarrow \mathbb{R}$ such that

$$p(\lambda x) = \lambda p(x) \quad \forall x \in X \text{ and } \lambda \geq 0 \quad (4.1)$$

$$p(x + y) \leq p(x) + p(y) \quad \forall x, y \in X \quad (4.2)$$

Theorem 4.1 (Helly-Hahn-Banach). Let E be a vector space over \mathbb{R} and $p : E \rightarrow \mathbb{R}$ a sublinear functional. If $G \subseteq E$ is a linear subspace and $g : G \rightarrow \mathbb{R}$ is a linear functional such that

$$g(x) \leq p(x) \quad \forall x \in G,$$

then there exists a linear functional f defined on all of E that extends g , that is, $g(x) = f(x) \forall x \in G$ and such that

$$f(x) \leq p(x) \quad \forall x \in E.$$

For a proof first recall that a partial order P is **inductive** if every totally ordered subset Q in P has an upper bound, and that

Lemma 4.2 (Zorn). Every nonempty order set that is inductive has a maximal element.

Proof. (Of theorem 4.1) Consider the set

$$P = \left\{ h : D(h) \subseteq E \rightarrow \mathbb{R} : \begin{array}{l} D(h) \text{ is a linear subspace of } E, \\ h \text{ is linear, } G \subseteq D(h), \\ h \text{ extends } g, \text{ and } h(x) \leq p(x) \forall x \in D(h) \end{array} \right\}$$

Then P is a partial order with

$$h_1 \leq h_2 \iff D(h_1) \subseteq D(h_2) \text{ and } h_2 \text{ extends } h_1$$

P is nonempty since $g \in P$. To show it is inductive, take $Q \subseteq P$ a partially order subset and write $Q = (h_i)_{i \in I}$. Then define

$$D(h) = \bigcup_{i \in I} D(h_i), \quad h(x) = h_i(x) \quad \text{if } x \in D(h_i) \text{ for some } i \in I$$

which is an upper bound of Q , so that there is a maximal element f in P by Zorn's Lemma. To finish it suffices to show that $D(f) = E$.

For a contradiction suppose that $D(f) \neq E$ and choose $x_0 \notin D(f)$. We shall construct a function $h \in P$ such that $f < h$. Define $D(h) = D(f) + \mathbb{R}x_0$ and, for every $x \in D(f)$, set

$$h(x + \lambda x_0) = f(x) + t\alpha \quad \forall \lambda \in \mathbb{R}$$

where α is a constant that we choose as follows. We must ensure that

$$h(x + \lambda x_0) = f(x) + \lambda\alpha \leq p(x + tx_0) \quad \forall x \in D(f) \quad \text{and} \quad \forall \lambda \in \mathbb{R}$$

For any $x, y \in D(f)$,

$$\begin{aligned} f(x) + f(y) &= f(x + y) \leq p(x + y) \leq p(x + x_0) + p(y - x_0) \\ \implies f(x) - p(y - x_0) &\leq p(x + x_0) - f(y) \end{aligned}$$

So let α satisfy

$$\sup_{y \in D(f)} \{f(y) - p(y - x_0)\} \leq \alpha \leq \inf_{x \in D(f)} \{p(x + x_0) - f(x)\}$$

If $\lambda = 0$, then $h(x) = f(x) \leq p(x)$. If $\lambda \neq 0$ we must be careful since sublinear functionals only satisfy eq. (4.1) for positive scalars.

If $\lambda > 0$, then

$$\begin{aligned} h(x + \lambda x_0) &= \lambda \cdot h(x/\lambda + x_0) \\ &= \lambda \cdot (f(x/\lambda) + \alpha) \\ &\leq \lambda \cdot (f(x/\lambda) + p(x/\lambda + x_0) - f(x/\lambda)) \\ &\leq p(x + \lambda x_0) \end{aligned}$$

and if $\lambda = -\mu < 0$,

$$\begin{aligned} h(x + tx_0) &= (-\lambda) \cdot h(-x/\lambda - x_0) \\ &= \mu \cdot (f(x/\mu) - \alpha) \\ &\leq \mu \cdot (f(x/\mu) - f(x/\mu) + p(x/\mu + x_0)) \\ &\leq p(x - \mu x_0) \\ &= p(x + \lambda x_0). \end{aligned}$$

Then $h \in P$, h extends f and $D(f) \subsetneq D(h)$, which is impossible since f is maximal. \square

An example of a sublinear functional is a **seminorm**, which is a function that satisfies the first two conditions of being norm, but the third. If p is a seminorm,

Theorem 4.3 (Complex Hahn-Banach). Let E be a complex vector space, p a seminorm on E , $G \subseteq E$ a subspace of E and f a complex linear functional on G such that $|f(x)| \leq p(x) \quad \forall x \in G$. Then there exists a complex linear functional f defined on all of E that extends f and $|f(x)| \leq p(x)$.

The space of all continuous linear functionals on E , denoted by $E^* := L(E, K)$, is called the **dual space** of E . Since K is complete, E^* is complete with the operator norm. We denote the **dual norm** (it remains for me to check whether this is the same as the operator norm) by

$$\|f\|_{E^*} = \sup_{\substack{x \in E \\ \|x\| \leq 1}} |f(x)| = \sup_{\substack{x \in E \\ \|x\| \leq 1}} f(x)$$

Given $f \in E^*$ and $x \in E$ we shall often write $\langle f, x \rangle$ instead of $f(x)$ and call $\langle -, - \rangle$ the **scalar product for the duality** E^*, E .

Corollary 4.4. Let $G \subseteq E$ be a linear subspace. If $g : G \rightarrow \mathbb{R}$ is a continuous linear functional, then there exists $f \in E^*$ that extends g and such that

$$\|f\|_{E^*} = \sup_{\substack{x \in G \\ \|x\| \leq 1}} = \|g\|_{G^*}$$

Proof. $g(x) \leq p(x) := \|g\|_{G^*} \|x\|$, which is a sublinear functional. □

Corollary 4.5. For every $x_0 \in E$ there exists $f_0 \in E^*$ such that

$$\|f_0\| = \|x_0\| \quad \text{and} \quad \langle f_0, x_0 \rangle = \|x_0\|^2$$

Proof. By corollary 4.4 with $G = \mathbb{R}x_0$ and $g(tx_0) = t\|x_0\|^2$ so that $\|g\|_{G^*} = \|x_0\|$. □

While the element f_0 in corollary 4.5 is in general not unique (exercise), a sufficient condition is that E^* is **strictly convex**, that is, $\|tx + (1-t)y\| < t\|x\| + (1-t)\|y\|$ for $t \in (0, 1)$, $x, y \in E$ with $\|x\| = \|y\| = 1$ and $x \neq y$. Any Hilbert space and any $L^p(\Omega)$ for $1 < p < \infty$ are.

The (multivalued) **duality map** for $x_0 \in E$ is

$$F(x_0) = \{f_0 \in E^* : \|f_0\| = \|x_0\| \text{ and } \langle f_0, x_0 \rangle = \|x_0\|^2\}$$

Corollary 4.6. For every $x \in E$ we have

$$\|x\| = \sup_{\substack{f \in E^* \\ \|f\| \leq 1}} |\langle f, x \rangle| = \max_{\substack{f \in E^* \\ \|f\| \leq 1}} |\langle f, x \rangle|$$

Proof. Assuming $x \neq 0$, if $\|f\| \leq 1$, $|\langle f, x \rangle| \leq \|f\| \|x\| \leq \|x\|$, so

$$\sup_{\substack{f \in E^* \\ \|f\| \leq 1}} |\langle f, x \rangle| \leq \|x\|.$$

On the other hand, by corollary 4.5 there is some $f_0 \in E^*$ such that $\|f_0\| = \|x\|$ and $\langle f_0, x \rangle = \|x\|^2$. Set $f_1 = f_0/\|x\|$ so that $\|f_1\| = 1$ and $\langle f_1, x \rangle = \|x\|$. □

4.2 Geometric forms of the Hahn-Banach theorem

An *affine hyperplane* is a subset H of E of the form

$$H = \{x \in E : f(x) = \alpha\}$$

where if f is a nonzero linear functional not necessarily continuous and α a real constant. We also write $H = [f = \alpha]$.

Proposition 4.7. The hyperplane $H = [f = \alpha]$ is closed if and only if f is continuous.

Proof. It is clear that if f is continuous then H is closed. Conversely... \square

Let A and B be two subsets of E . We say that the hyperplane $H = [f = \alpha]$ *separates* A and B if

$$f(x) \leq \alpha \quad \forall x \in A \quad \text{and} \quad f(x) \geq \alpha \quad \forall x \in B$$

We say H *strictly separates* A and B if there is some $\varepsilon > 0$ such that

$$f(x) \leq \alpha - \varepsilon \quad \forall x \in A \quad \text{and} \quad f(x) \geq \alpha + \varepsilon \quad \forall x \in B$$

A subset $A \subseteq E$ is *convex* if

$$tx + (1 - t)y \in A \quad \forall x, y \in A, \forall t \in [0, 1].$$

Theorem 4.8 (Hahn-Banach, first geometric form). Let $A \subset E$ and $B \subset E$ be two nonempty convex subsets such that $A \cap B = \emptyset$. Assume that one of them is open. Then there exists a closed hyperplane that separates A and B .

Lemma 4.9. Let $C \subset E$ be an open convex set with $0 \in C$. For every $x \in E$ set

$$p(x) = \inf\{\alpha > 0 : \alpha^{-1}x \in C\}$$

which is called the *Minkowski functional of C* or the *gauge of C* . Then p is a sublinear functional, that is, satisfies eqs. (4.1) and (4.2), and also

1. there is a constant M such that $0 \leq p(x) \leq M\|x\| \quad \forall x \in E$.
2. $C = \{x \in E : p(x) < 1\}$. (In particular, if C is the unit ball centered at 0, then p is just the norm.)

Lemma 4.10. Let $C \subset E$ be a nonempty open convex set and let $x_0 \in E$ with $x_0 \notin C$. Then there exists $f \in E^*$ such that $f(x) < f(x_0) \quad \forall x \in C$. In particular, the hyperplane $[f = f(x_0)]$ separates $\{x_0\}$ and C .

Proof. After a translation we may assume that $0 \in C$ and introduce the gauge p of C . For the linear subspace $G = \mathbb{R}x_0$ and the linear functional $g : G \rightarrow \mathbb{R}$ defined by $g(t_0) = t$ ($t \in \mathbb{R}$) we have that $g(x) \leq p(x) \quad (\forall x \in G)$. We may thus apply the Helly-Hahn-Banach theorem. \square

Proof. (Of theorem 4.8). Set $C = A - B$ so that C is convex (check!). C is also open (since $C = \bigcup_{y \in B} (A - y)$) and $0 \notin C$ (because $A \cap B = \emptyset$). By lemma 4.10 there exists $f \in E^*$ such that

$$f(z) \leq 0 \quad \forall z \in C \quad \Longleftrightarrow \quad f(x) < f(y) \quad \forall x \in A \quad \forall y \in B$$

Fixing a constant α satisfying

$$\sup_{x \in A} f(x) \leq \alpha \leq \inf_{y \in B} f(y)$$

we conclude that the hyperplane $[f = \alpha]$ separates A and B . \square

Theorem 4.11 (Hahn-Banach, second geometric form). Let $A \subset E$ and $B \subset E$ be two nonempty convex subsets such that $A \cap B = \emptyset$. Assume that A is closed and B is compact. Then there exists a closed hyperplane that strictly separates A and B .

Proof. Set $C = A - B$, so that C is convex, closed (check!) and $0 \notin C$. Hence there is some $r > 0$ such that $B(0, r) \cap C = \emptyset$. By theorem 4.8 there is a hyperplane that separates $B(0, r)$ and C . Therefore, there is some $f \in E^*$, $f \neq 0$ such that

$$f(x - y) \leq f(rz) \quad \forall x \in A, \quad \forall y \in B \quad \forall z \in B(0, 1)$$

(Incomplete.) \square

In infinite-dimensional vector spaces it is in general impossible to separate any two nonempty disjoint convex sets. In finite-dimensional vector spaces, however, it is always possible.

Corollary 4.12. Let $F \subset E$ be a linear subspace such that $\overline{F} \neq E$. Then there exists some $f \in E^*$, $f \neq 0$ such that

$$\langle f, x \rangle = 0 \quad \forall x \in F$$

Proof. Let $x_0 \in E$ with $x_0 \notin \overline{F}$. Using theorem 4.11 with $A = \overline{F}$ and $B = \{x_0\}$ we find a closed hyperplane $[f = \alpha]$ that strictly separates \overline{F} and $\{x_0\}$. We thus have

$$\langle f, x \rangle < \alpha \langle f, x_0 \rangle \quad \forall x \in F$$

In particular $\langle f, x \rangle = 0 \quad \forall x \in F$ since $\lambda \langle f, x \rangle < \alpha \quad \forall \lambda \in \mathbb{R}$. \square

By this corollary, we may prove that a linear subspace $F \subset E$ is dense by showing that every continuous linear functional that vanishes on F must vanish in all of E .

4.3 The bidual E^{**} . Orthogonality relations.

Let E be a normed vector space. Recall the norm on its dual space E^* is

$$\|f\|_{E^*} = \sup_{\substack{x \in E \\ \|x\| \leq 1}} |\langle f, x \rangle|.$$

The *bidual* space of E is the dual of E^* with norm

$$\|\xi\|_{E^{**}} = \sup_{\substack{f \in E^* \\ \|f\| \leq 1}} |\langle \xi, f \rangle| \quad \xi \in E^{**}$$

There is a *canonical injection* $J : E \rightarrow E^{**}$ defined as follows: given $x \in E$, the map $f \mapsto \langle f, x \rangle$ is a continuous linear functional on E^* ; thus it is an element of E^{**} which we denote by Jx . We have

$$\langle Jx, f \rangle_{E^{**}, E^*} = \langle f, x \rangle_{E^*, E} \quad \forall x \in E, \forall f \in E^*.$$

It is clear that J is linear and that J is an isometry, that is, $\|Jx\|_{E^{**}} = \|x\|_E$. Indeed, by corollary 4.6 we have

$$\|x\|_{E^{**}} = \sup_{\substack{f \in E^* \\ \|f\| \leq 1}} |\langle Jx, f \rangle| = \sup_{\substack{f \in E^* \\ \|f\| \leq 1}} |\langle f, x \rangle| = \|x\|_E$$

It may happen that J is not surjective from E onto E^{**} . However, it is convenient to identify E with a subspace of E^{**} using J . If J turns out to be surjective we say E is *reflexive* and E^{**} is identified with E . and if $N \subset E^*$ is a linear subspace we set

$$N^\perp = \{x \in E : \langle f, x \rangle = 0 \forall f \in N\}$$

If $M \subset E$ is a linear subspace, we set

$$M^\perp = \{f \in E^* : \langle f, x \rangle = 0 \forall x \in M\}.$$

Notice that N^\perp is a subset of E rather than of E^{**} . Both M^\perp and N^\perp are closed linear subspaces. We say M^\perp (resp. N) is the *space orthogonal to M (resp. N)*.

Proposition 4.13. Let $M \subset E$ be a linear subspace. Then

$$(M^\perp)^\perp = \overline{M}.$$

Let $N \subset E^*$ be a linear subspace. Then

$$(N^\perp)^\perp \supset \overline{N}.$$

5 The Uniform Boundedness Principle and the Closed Graph Theorem

5.1 The Baire Category Theorem

This theorem states that, in a complete metric space, a countable union of closed sets with empty interior has empty interior:

Theorem 5.1 (Baire). Let X be a complete metric space and let $(X_n)_{n \geq 1}$ be a sequence of closed subsets in X . Assume that

$$\text{Int } X_n = \emptyset \text{ for every } n \geq 1$$

Then

$$\text{Int} \left(\bigcup_{n=1}^{\infty} X_n \right) = \emptyset$$

And there's also this other form of the same theorem:

Theorem 5.2 (Baire). Let X be a nonempty complete metric space. Let $(X_n)_{n \geq 1}$ be a sequence of closed subsets such that

$$\bigcup_{n=1}^{\infty} X_n = X.$$

Then there exists some n_0 such that $\text{Int } X_{n_0} \neq \emptyset$.

Remark 5.1. A subset $A \subset X$ has empty interior if and only if its complement A^c is dense in X . Indeed, to see A^c is dense in X notice every point $x \in X$ has a neighbourhood intersecting A^c : since A has empty interior, every open set containing $x \in A$ must intersect A^c . Conversely, if every point $x \in X$ has a neighbourhood intersecting A^c , there cannot be a point in the interior of A .

Proof. (Of the first form.) Set $O_n = X_n^c$ so that O_n is open and dense in X for every $n \geq 1$. To show $\bigcup_{n=1}^{\infty} X_n$ has empty interior, it suffices to show that its complement $G = \bigcap_{n=1}^{\infty} O_n$ is dense in X . Succintly, the procedure is to construct a sequence of a sequence of closed balls inside any given open set ω of X ; one ball for every O_n . This produces a Cauchy sequence, whose limit is a point in $\omega \cap G$. \square

5.2 The Uniform Boundedness Principle

Let E and F be two normed vector spaces. Recall that the space of continuous (=linear) operators from E into F is denoted by $\mathcal{L}(E, F)$ and is equipped with the norm

$$\|T\|_{\mathcal{L}(E, F)} = \sup_{\substack{x \in E \\ \|x\| \leq 1}} \|Tx\| = \inf_{\substack{x \in E \\ \|Tx\| \leq c\|x\|}} c.$$

Theorem 5.3 (Banach-Steinhaus, uniform boundedness principle). Let E and F be two Banach spaces and let $(T_i)_{i \in I}$ be a family (not necessarily countable) of continuous linear operators from E to F . Assume that

$$\sup_{i \in I} \|T_i x\| < \infty \quad \forall x \in X \tag{5.1}$$

Then

$$\sup_{i \in I} \|T_i\|_{\mathcal{L}(E, F)} < \infty.$$

In other words, there exists a constant c such that

$$\|T_i x\| \leq c\|x\| \quad \forall x \in E \quad \forall i \in I$$

Remark 5.2. What's remarkable about this theorem is that from *pointwise* (the norm of all the $T_i x$ is bounded at every x) estimates we obtain a *global (uniform)* estimate: there is a unique bound for all the T_i .

Proof. For every $n \geq 1$, let

$$X_n = \{x \in E : \forall i \in I, \|T_i\| \leq n\}$$

which is closed (why?) and by eq. (5.1) we have

$$\bigcup_{n=1}^{\infty} X_n = E.$$

It follows from the Baire category theorem that there is some $n_0 \geq 1$ such that $\text{Int}(X_0) \neq \emptyset$. Pick a ball $B(x_0, r) \subset X_{n_0}$, so that for every $z \in B(0, 1)$ and $i \in I$,

$$\begin{aligned} \|T_i(z)\| &= r^{-1} \|T(x_0 + rz) - T(x_0)\| \\ &\leq r^{-1}(n_0 + n_0). \end{aligned}$$

So

$$\sup_{i \in I} \|T_i\|_{\mathcal{L}(E, F)} = \sup_{i \in I} \left(\sup_{z \in B(0, 1)} \|T_i z\|_F \right) \leq r^{-1}(n_0 + n_0) < \infty.$$

□

Remark 5.3. Recall that the pointwise limit of continuous functions need not be continuous. The former theorem does not imply that $\|T_n - T\|_{\mathcal{L}(E, F)}$.

Corollary 5.4. Let E and F be two Banach spaces. Let (T_n) be a sequence of continuous linear operators from E to F such that for every $x \in E$, $T_n x$ converges to a limit denoted by Tx . Then

1. $\sup_n \|T_n\|_{\mathcal{L}(E, F)} < \infty$ (uniform boundedness principle),
2. $T \in \mathcal{L}(E, F)$,
3. $\|T\|_{\mathcal{L}(E, F)} \leq \liminf_{n \rightarrow \infty} \|T_n\|_{\mathcal{L}(E, F)}$. (Contrary, I think, to $\lim_{n \rightarrow \infty} \|T_n\| = \|T\|$ if convergence was uniform).

Corollary 5.5. Let G be a Banach space and let B be a subset of G . Assume that

$$\text{for every } f \in G^* \text{ the set } f(B) = \{\langle f, x \rangle : x \in B\} \text{ is bounded} \quad (5.2)$$

then

$$B \text{ is bounded.}$$

Proof. Define a family of linear operators on G^* indexed by B as

$$T_b f = \langle f, b \rangle.$$

By eq. (5.2),

$$\sup_{b \in B} |T_b(f)| < \infty \quad \forall f \in E$$

so that by the uniform boundedness theorem there exists a constant c such that

$$|\langle f, b \rangle| \leq c \|f\| \quad \forall f \in G^* \quad \forall b \in B$$

and by corollary 4.6,

$$\|b\| = \sup_{\substack{f \in G^* \\ \|f\| \leq 1}} |\langle f, b \rangle| \leq c$$

□

Remark 5.4. To prove that a set B is bounded it suffices to *look* at B through the bounded linear functionals.

Corollary 5.6. Let G be a Banach space and let B^* be a subset of G^* . Assume that

$$\text{for every } x \in G \text{ the set } \langle B^*, x \rangle = \{\langle f, x \rangle : f \in B^*\} \text{ is bounded} \quad (5.3)$$

then

$$B^* \text{ is bounded.}$$

Proof. Define for every $b \in B^*$

$$T_b(x) = \langle b, x \rangle \quad \forall x \in G$$

so that there is a constant c such that

$$|\langle b, x \rangle| \leq c \|x\| \quad \forall b \in B^*, \forall x \in G.$$

And this time by definition of operator norm we have

$$\|b\| = \sup_{\substack{x \in G \\ \|x\| \leq 1}} |\langle b, x \rangle| \leq c \quad \forall b \in B^*.$$

□

Remark 5.5. In these last two proofs it has become obvious how the norm of a vector can be described in terms of operators, and, dually, the norm of an operator can be described in terms of points. Recall corollary 4.6 is a consequence of the Hahn-Banach theorem.

5.3 The Open Mapping Theorem and the Closed Graph Theorem

Here are two basic results due to Banach.

Theorem 5.7 (Open mapping theorem). *A surjective continuous linear operator between two Banach spaces is an open map.*

Let E and F be two Banach spaces and T be a linear operator from E into F that is surjective. Then there exists a constant $c > 0$ such that

$$T(B_E(0, 1)) \supset B_F(0, c) \quad (5.4)$$

Remark 5.6. First let us make sure that eq. (5.4) implies that T is open. Let U be open in E and fix any point $y_0 \in T(U)$, so that there is some $x_0 \in U$ such that $Tx_0 = y_0$. Choose a ball $x_0 + B(0, r) = B(x_0, r) \subset U$. Then

$$\begin{aligned} y_0 + T(B(0, r)) &\subset T(U) \\ \implies T(B(0, r)) &\subset -y_0 + T(U) \\ \implies r^{-1}T(B(0, 1)) &\subset -r^{-1}y_0 + r^{-1}T(U). \end{aligned}$$

And by eq. (5.4) we obtain

$$\begin{aligned} r^{-1}B_F(0, c) &\subset -r^{-1}y_0 + r^{-1}T(U) \\ \implies B(0, rc) &\subset -y_0 + T(U) \\ \implies B(y_0, rc) &\subset T(U). \end{aligned}$$

Corollary 5.8. Let E and F be two Banach spaces and let T be a continuous linear operator from E into F that is bijective. Then T^{-1} is also continuous.

Proof of corollary 5.8. Since T is injective, choosing a vector $Tx \in B_F(0, c)$ implies its only preimage x is in $B(0, 1)$, so $\|x\| < 1$. By **homogeneity**,

$$\|x\| \leq \frac{1}{c} \|Tx\|$$

so that T^{-1} is continuous. □

Corollary 5.9. Let E be a vector space provided with two norms $\|\cdot\|_1$ and $\|\cdot\|_2$. Assume that E is a Banach space for *both* norms and that there exists a constant $C \geq 0$ such that

$$\|x\|_2 \leq C\|x\|_1 \quad \forall x \in E.$$

Then the two norms are equivalent, that is, there is a constant $c > 0$ such that

$$\|x\|_1 \leq c\|x\|_2 \quad \forall x \in E.$$

(That is, for both inequalities

$$c^{-1}\|x\|_1 \leq \|x\|_2 \leq C\|x\|_1 \quad \forall x \in E)$$

to hold it suffices that the inequality on the right side to holds.)

Proof of corollary 5.9. By corollary 5.8 with $E = (E, \|\cdot\|_1)$, $F = (E, \|\cdot\|_2)$ and $T = \text{Id}$. (That is, the two norms induce the same topology?) \square

Proof of theorem 5.7.

Claim 5.1. Assume that T is a linear surjective operator from E onto F . Then there exists a constant $c > 0$ such that

$$\overline{T(B(0, 1))} \supset B(0, 2c). \quad (5.5)$$

Proof of claim 5.1. [Uses Baire category theorem.] \square

Claim 5.2. Assume that T is a continuous linear operator from E into F that satisfies eq. (5.5). Then

$$T(B(0, 1)) \supset B(0, c). \quad (5.6)$$

Proof of claim 5.2. [Uses Cauchy sequences.] (Compare eqs. (5.4) and (5.6).) \square

\square

Theorem 5.10 (Closed graph theorem). Let E and F be two Banach spaces. Let T be a linear operator from E into F . Assume that the graph of T , denoted by $G(T)$, is closed in $E \times F$. Then T is continuous.

Remark 5.7. The converse is true: the graph of any continuous map is closed.

Proof. Consider, on E , the two norms

$$\|x\|_1 = \|x\|_E + \|Tx\|_F \quad \text{and} \quad \|x\|_2 = \|x\|_E$$

($\|\cdot\|_1$ is called the **graph norm**).

Then prove these two norms are equivalent. (Check details.)

It is easy to check that E is a Banach space with $\|\cdot\|_1$. Since it is also a Banach space with $\|\cdot\|_2$, it follows from corollary 5.9 that the two norms are equivalent and thus there exists a constant $c > 0$ such that $\|x\|_1 \leq c\|x\|_2$. We conclude that $\|Tx\|_F \leq c\|x\|_E$. \square

5.4 An introduction to Unbounded Linear Operators. Definition of the Adjoint

Let E and F be two Banach spaces. An **unbounded linear operator** from E into F is a linear map $A : D(A) \subset E \rightarrow F$ defined on a linear subspace $D(A) \subset E$ with values in F . The set $D(A)$ is called the **domain** of A . We say that A is **bounded** or **continuous** if $D(A) = E$ and if there is a constant $c \geq 0$ such that

$$\|Au\| \leq c\|u\| \quad \forall u \in E.$$

The norm of a bounded operator is defined as

$$\|A\|_{\mathcal{L}(E,F)} := \sup_{u \neq 0} \frac{\|Au\|}{\|u\|}.$$

The *graph* of A is the set

$$G(A) := \{[u, Au] : u \in D(A)\} \subset E \times F$$

A is *closed* if $G(A)$ is closed in $E \times F$. We denote by $R(A)$ the *range* of A and by $N(A)$ the *kernel* of A .

Remark 5.8. In order to prove that an operator A is closed, one proceeds in general as follows. Take a sequence (u_n) in $D(A)$ such that $u_n \rightarrow u$ in E and $Au_n \rightarrow f$ in F . Then just check that

1. $u \in D(A)$
2. $f = Au$

which amounts to taking a convergent sequence in the graph of A and checking its limit is still in the graph. Notice it does not suffice to take only a sequence convergent to 0.

Remark 5.9. If A is closed, then $N(A)$ is closed but $R(A)$ need not be closed.

Let $A : D(A) \subset E \rightarrow F$ be an unbounded linear operator that is *densely defined*, that is, for which $D(A)$ is dense in E . To introduce an unbounded operator $A^* : D(A^*) \subset F^* \rightarrow E^*$ define its domain as

$$D(A^*) := \{v \in F^* : \exists c \geq 0 \text{ such that } |\langle v, Au \rangle| \leq c\|u\| \ \forall u \in D(A)\},$$

which is clearly a linear subspace of F^* . Just like the pullback a 1-form is defined by pushing forward a vector and then evaluating on the form, define for $v \in D(A^*) \subset F^*$,

$$g(u) = \langle v, Au \rangle \quad \forall u \in D(A).$$

Notice this “pullback” of v is defined only for vectors in $D(A) \subset E$, so that we must extend g to a functional on all of E . By definition, $|g(u)| \leq c\|u\| \ \forall u \in D(A)$, so that we may apply Hahn-Banach (theorem 4.1) to obtain a linear map $f : E \rightarrow \mathbb{R}$ that extends g and $|f(u)| \leq c\|u\| \ \forall u \in E$. It follows that the extension is unique from the fact that $D(A)$ is dense in E (and that f is continuous).

The unbounded linear operator $A^* : D(A^*) \subseteq F^* \rightarrow E^*$ is called the *adjoint* of A . The fundamental relation between A and A^* is

$$\langle v, Au \rangle_{F^*,F} = \langle A^*v, u \rangle_{E^*,E} \quad \forall u \in D(A), \ \forall v \in D(A^*).$$

Which reminds me of

$$\omega(F_*v) = (F^*\omega)v.$$

Remark 5.10. Its not necessary to use Hahn-Banach to extend g .

Remark 5.11. It may happen that $D(A^*)$ is not dense in F^* . This is a rather pathological situation. It is always true that if A is closed then $D(A^*)$ is dense in F^* for the weak* topology $\sigma(F^*, F)$. In particular, if F is reflexive, then $D(A^*)$ is dense in F^* for the usual norm topology.

Remark 5.12. If A is bounded, then A^* is also bounded and $\|A^*\|_{\mathcal{L}(E^*, F^*)} = \|A\|_{\mathcal{L}(E, F)}$.
To see this,.

Proposition 5.11. Let $A : D(A) \subset E \rightarrow F$ be a densely defined unbounded linear operator. Then A^* is closed, that is, $G(A^*)$ is closed in $F^* \times E^*$.

Proof. Suppose $v_n \in D(A^*)$ is such that $v_n \rightarrow v$ in F^* and $A^*v_n \rightarrow f$ in E^* . We must show that $v \in D(A^*)$ and that $A^*v = f$. We know that

$$\langle v_n, Au \rangle = \langle A^*v_n, u \rangle \quad \forall u \in D(A).$$

At the limit we obtain

$$\langle v, Au \rangle = \langle f, u \rangle \quad \forall u \in D(A)$$

so that

$$\langle v, Au \rangle = \langle f, u \rangle \leq \|f\| \|u\| \quad \forall u \in D(A)$$

since $f \in E^*$, which means that $v \in D(A^*)$, and also that $A^*v = f$. \square

The graphs of A and A^* are related by a ver simple orthogonality relation. Consider the isomorphism $I : F^* \times E^* \rightarrow E^* \times F^*$ defined by

$$I([v, f]) = [-f, v].$$

Let $A : D(A) \subset E \rightarrow F$ be a densely defined unbounded linear operator. Then

$$I[G(A^*)] = G(A)^\perp = \{[f, v] \in (E \times F)^* : \langle f, x \rangle = 0 \quad \forall x \in G(A)\}.$$

Indeed, let $[v, f] \in F^* \times E^*$, then

$$\begin{aligned} [v, f] \in G(A^*) &\iff \langle f, u \rangle = \langle v, Au \rangle \quad \forall u \in D(A) \\ &\iff -\langle f, u \rangle + \langle v, Au \rangle = 0 \quad \forall u \in D(A) \\ &\iff [-f, u] \in F(A)^\perp \end{aligned}$$

Corollary 5.12. Orthogonality relations between range and null space of an operator and its adjoint.

6 Weak topologies. Reflexive Spaces. Separable Spaces. Uniformly Convex Spaces.

6.1 The weak topology $\sigma(E, E^*)$

Consider the situation of having a set X and a collection $(Y_i)_{i \in I}$ of topological spaces along with a collection of maps $\varphi_i : X \rightarrow Y_i$ for $i \in I$. Then we embark in finding the topology \mathcal{T} on X with the fewest open sets that makes all the maps $(\varphi_i)_{i \in I}$ continuous. A construction is given by considering finite intersections of sets of the form $\varphi_i^{-1}(\omega_i)$ for ω_i open in Y_i , and then arbitrary unions. Such is the **weakest topology associated to the collection** $(\varphi_i)_{i \in I}$. Then we have

Proposition 6.1. Let (x_n) be a sequence in X . Then $x_n \rightarrow x$ in \mathcal{T} if and only if $\varphi_i(x_n) \rightarrow \varphi_i(x)$ for every $i \in I$.

Proposition 6.2. Let Z be a topological space and let ψ be a map from Z into X . Then ψ is continuous if and only if $\varphi_i \circ \psi$ is continuous from Z into Y_i for every $i \in I$.

Let E be a Banach space. The **weak topology** $\sigma(E, E^*)$ is the weak topology associated to the linear functionals in E^* . Since every functional in E^* is continuous (by definition of E^*), the weak topology is weaker than the usual (norm) topology.

Proposition 6.3. The weak topology $\sigma(E, E^*)$ is Hausdorff.

Proposition 6.4. By Hahn-Banach second geometric form.

Proposition 6.5. Let $x_0 \in E$. Given $\varepsilon > 0$ and a **finite** set $\{f_1, f_2, \dots, f_k\} \subset E^*$, consider

$$V = V(f_1, \dots, f_k; \varepsilon) = \{x \in E : |\langle f_i, x - x_0 \rangle| < \varepsilon \ \forall i = 1, 2, \dots, k\}.$$

Then V is a neighbourhood of x_0 for the topology $\sigma(E, E^*)$. Moreover, we obtain a **basis of neighbourhoods** of x_0 for $\sigma(E, E^*)$ by varying ε, k and the f_i in E .

If a sequence (x_n) in E converges to x in the weak topology we write $x_n \rightharpoonup x$.

Proposition 6.6. Let (x_n) be a sequence in E . Then

1. $x_n \rightharpoonup x \iff \langle f, x_n \rangle \rightarrow \langle f, x \rangle \ \forall f \in E^*$.
2. If $x_n \rightarrow x$, then $x_n \rightharpoonup x$.
3. If $x_n \rightharpoonup x$, then $(\|x_n\|)$ is bounded and $\|x\| \leq \liminf \|x_n\|$.
4. If $x_n \rightharpoonup x$ and $f_n \rightarrow f$, then $\langle f_n, x_n \rangle \rightarrow \langle f, x \rangle$.

Proposition 6.7. When E is finite-dimensional, the weak topology and the usual topology are the same. In particular, a sequence converges weakly if and only if it converges strongly.

Example. The unit sphere $S = \{x \in E : \|x\| = 1\}$ with E infinite-dimensional is never closed in the weak topology $\sigma(E, E^*)$. More precisely, its weak closure is

$$\overline{S}^{\sigma(E, E^*)} = B_E = \{x \in E : \|x\| \leq 1\}.$$

The geometric interpretation of this construction is that when E is infinite-dimensional, every neighbourhood of a point in the weak topology contains a line passing through the point; even a whole affine space passing through the point. The proof uses proposition 6.5.

Example. The unit ball $U = \{x \in E : \|x\| < 1\}$ with E infinite-dimensional is never open in the weak topology. If U was weakly open, its complement U^c would be weakly closed, so that $S = B_E \cap U^c$ would be weakly closed.

Remark 6.1. In infinite-dimensional spaces the weak topology is never metrizable. However, as we shall see, if E^* is separable, one can define a norm on E that induces on bounded sets of E the weak topology.

Remark 6.2. Usually, in infinite-dimensional spaces, there exist sequences that converge weakly and do not converge strongly. For example, if E^* is separable or if E is reflexive, one can construct a sequence (x_n) in E such that $\|x_n\| = 1$ and $x_n \rightharpoonup 0$. However, there are spaces such as ℓ^1 with the unusual property that every weakly convergent sequence is strongly convergent. Recall that two metrizable spaces with the same convergent sequences have identical topologies, but two topological spaces with the same convergent sequences need not have identical topologies.

Theorem 6.8. Let C be a convex subset of E . Then C is closed in the weak topology if and only if it is closed in the strong topology.

Proof. Hahn-Banach. □

Remark 6.3. In the proof we see that every closed convex set C coincides with the intersection of all the closed half-spaces containing C :

Corollary 6.9 (Manzur). Assume (x_n) converges weakly to x . Then there exists a sequence (y_n) made up of convex combinations of the x_n that converges strongly to x .

Corollary 6.10. Assume that $\varphi : E \rightarrow (-\infty, +\infty]$ is convex and l.s.c. in the strong topology. Then φ is l.s.c. in the weak topology.

Theorem 6.11. Let E and F be two Banach spaces and let T be a linear operator from E into F . Assume that T is continuous in the strong topologies. Then T is continuous from E weak into F weak, and conversely.

Remark 6.4. In fact, if a linear operator is continuous from E strong into F weak, then it is continuous from E strong into F strong. The following continuity properties are all the same: $S \rightarrow S$, $W \rightarrow W$, $S \rightarrow W$. On the other hand, very few linear operators are continuous $W \rightarrow S$.

Remark 6.5. Nonlinear maps that are continuous from E strong into F strong are in general not continuous from E weak into F weak.

6.2 The weak* topology $\sigma(E^*, E)$

For every $x \in E$ consider the linear functional in E^{**} defined by $f \mapsto \langle f, x \rangle$. The *weak** $\sigma(E^*, E)$ is the weakest topology associated to the former collection of functionals as x runs through E . It is clear that $\sigma(E^*, E)$ is coarser than the usual weak topology of E^* , denoted by $\sigma(E^*, E^{**})$.

Remark 6.6. The hysteria over weak topologies is explained by the fact that there are more compact sets in a weaker topology. The closed unit ball B_{E^*} is *never compact in the strong topology (unless E is finite-dimensional)*; however, it is always compact in the weak* by the Banach-Algolou-Bourbaki theorem, ??

Proposition 6.12. The weak* topology is Housdorff.

Proof. Hahn-Banach. □

Proposition 6.13. Let $f_0 \in E^*$. Given a **finite** set $\{x_1, x_2, \dots, x_k\}$ in E and $\varepsilon > 0$, consider

$$V = V(x_1, \dots, x_k; \varepsilon) = \{f \in E^* : |\langle f - f_0, x_i \rangle| < \varepsilon \forall i = 1, \dots, k\}.$$

Then V is a neighbourhood of f_0 for the weak* topology. Moreover, we obtain a **basis of neighbourhoods** of f_0 for the weak* topology by varying ε, k and the x_i in E .

If a sequence (f_n) converges to f in the weak* topology, we write $f_n \xrightarrow{*} f$.

Proposition 6.14. Let (f_n) be a sequence in E^* . Then

1. $f_n \xrightarrow{*} f \iff \langle f_n, x \rangle \rightarrow \langle f, x \rangle \forall x \in E$.
2. $f_n \rightarrow f \implies f_n \rightharpoonup f \implies f_n \xrightarrow{*} f$.
3. If $f_n \xrightarrow{*} f$, then $(\|f_n\|)$ is bounded and $\|f\| \leq \liminf \|f_n\|$.
4. $f_n \xrightarrow{*} f$ and $x_n \rightarrow x$, then $\langle f_n, x_n \rangle \rightarrow \langle f, x \rangle$.

Remark 6.7. In finite-dimensional spaces, the three topologies strong, weak and weak* coincide. Indeed, the canonical injection $J : E \rightarrow E^{**}$ is surjective.

Remark 6.8. If the canonical injection J is not surjective, the weak* topology is strictly coarser than the topology $\sigma(E^*, E^{**})$. (This requires two propositions not listed.)

Theorem 6.15 (Banach-Alaoglu-Bourbaki). The closed unit ball $B_{E^*} = \{f \in E^* : \|f\| \leq 1\}$ is compact in the weak* topology.

Proof. Considering the space of all functions from E into \mathbb{R} and equipping it with the product (=Tychonoff, pointwise convergence) topology, in which we may embed E^* homeomorphically (using the weak* topology). Compactness is then checked with the product topology in the space of all functions. □

6.3 Reflexive spaces

Recall that the *canonical injection* $J : E \rightarrow E^{**}$ is defined as follows: given $x \in E$, the map $f \mapsto \langle f, x \rangle$ is a continuous linear functional on E^* , and thus an element of E^{**} which we denote by Jx . We have

$$\langle Jx, f \rangle_{E^{**}, E^*} = \langle f, x \rangle_{E^*, E} \quad \forall x \in E, \forall f \in E^{**}.$$

It is clear that J is linear and that J is an isometry, that is, $\|Jx\|_{E^{**}} = \|x\|_E$. Since it may happen that J is not surjective, we say E is *reflexive* when J is surjective, in which case we usually identify E with E^{**} .

Remark 6.9. Finite-dimensional spaces are reflexive since a space and its dual are of the same dimension ($\dim E = \dim E^* = \dim E^{**}$). L^p and ℓ^p are reflexive for $1 < p < \infty$. Hilbert spaces are reflexive. However, L^1 , L^∞ , ℓ^1 and ℓ^∞ , as well as $C(K)$, the space of continuous functions on an infinite metric compact space K , are not reflexive.

Remark 6.10. There may be other surjective isometries from E to E^{**} , so that our discussion is only true for J .

Theorem 6.16 (Kakutani). Let E be a Banach space. Then E is reflexive if and only if $B_E = \{x \in E : \|x\| \leq 1\}$ is compact in the weak topology.

Proof.

(\implies) Using Banach-Alaoglu.

(\impliedby) Using the following two lemmas.

Lemma 6.17 (Helly). (An equivalence between two conditions that is later used in the context of proposition 6.5.)

Lemma 6.18 (Goldstine). Let E be any Banach space. Then $J(B_E)$ is dense in $B_{E^{**}}$ with respect to the topology $\sigma(E^{**}, E^*)$, and consequently $J(E)$ is dense in E^{**} in the topology $\sigma(E^{**}, E^*)$.

Proof. Helly's lemma, proposition 6.5. □

To conclude notice that since J is always continuous (check!), if B_E is compact in the weak topology, then $J(B_E)$ as well with $\sigma(E^{**}, E^*)$, and in particular it is closed. By Goldstine's lemma, $J(B_E)$ is dense in $B_{E^{**}}$, so that $J(B_E) = B_{E^{**}}$, which in turn implies that $J(E) = E^{**}$, that is, E is reflexive. □

Next we state some properties related to compactness.

Theorem 6.19. Assume that E is a reflexive Banach space and let (x_n) be a bounded sequence in E . Then there exists a subsequence (x_{n_k}) that converges in the weak topology.

Proof. Is given in section 6.4. □

Theorem 6.20 (Eberlein-Šmulian). Assume that E is a Banach space such that every bounded sequence in E admits a weakly convergent subsequence. Then E is reflexive.

Remark 6.11. From Wikipedia: Bolzano-Weirstrass theorem says that every infinite bounded sequence in \mathbb{R}^n has a convergent subsequence. Equivalently, a subset of \mathbb{R}^n is sequentially compact if and only if it is closed and bounded.

Remark 6.12. We recall theorem 2.18. Let (X, ρ) be a metric space and $A \subseteq X$.

1. A convergent subsequence can be selected out of any sequence of points in A if and only if A is relatively compact, that is, its closure \bar{A} is compact. (We do not claim that the limit point is in A .)
2. If A is relatively compact, it is totally bounded (in the sense that there exists a finite ε -net with respect to A for every $\varepsilon > 0$).
3. If (X, ρ) is complete and A is totally bounded, then A is relatively compact.
4. If A is compact then it is closed and totally bounded.

Remark 6.13. Here is Brezis' remark in this respect:

1. If X is a metric space, then X is compact if and only if every sequence in X admits a convergent subsequence.
2. There exist compact topological spaces X and some sequences in X without any convergent subsequence. An example is B_{E^*} , which is compact in the weak* topology; when $E = \ell^\infty$, it is easy to construct a sequence in X without any convergent subsequence.
3. If X is any topological space with the property that every sequence admits a convergent subsequence, X need not be compact.

Proposition 6.21. If X is a reflexive Banach space and $M \subseteq E$ is a closed linear subspace of E , then M is reflexive.

Proof. By Hahn-Banach, every functional on M is the restriction of a functional on E , so that the weak topology on M and the subspace topology of the weak topology on E coincide. By Kakutani's theorem it suffices to show the closed unit ball on M is compact. However, B_E is compact in the weak topology of E and M is closed, too, in the weak topology by theorem 6.8. \square

Corollary 6.22. A Banach space E is reflexive if and only if its dual space E^* is reflexive.

Proof. (\implies) Roughly speaking, $E^{**} = E$ implies $E^{***} = E^*$. Indeed, for $\varphi \in E^{***}$, the map $x \mapsto \langle \varphi, Jx \rangle$ is a continuous linear functional on E , where $J : E \rightarrow E^{**}$ is the canonical isomorphism. Call such a functional f , so that

$$\langle \varphi, Jx \rangle := \langle f, x \rangle \quad \forall x \in E.$$

Now by definition of J , $\langle Jx, f \rangle = \langle f, x \rangle \forall x \in E$, so that

$$\langle \varphi, Jx \rangle = \langle Jx, f \rangle.$$

Since J is surjective, we have

$$\langle \varphi, \xi \rangle = \langle \xi, f \rangle \quad \forall \xi \in E^{**}.$$

If the canonical injection from E^* into E^{***} is the map $f \mapsto \varphi$, we have seen exactly that it is surjective.

(\Leftarrow) Applying the previous step to the space E^* , we have that E^{**} is reflexive. Since $J(E)$ is a closed subspace of E^{**} in the strong topology (why?), we conclude by proposition 6.21 that $J(E)$ is reflexive. This implies that E is reflexive (in general, if E and F are two Banach spaces and T is a linear surjective isometry from E onto F , then E is reflexive iff F is reflexive). \square

Corollary 6.23. Let E be a reflexive Banach space. Let $K \subset E$ be a bounded closed and convex subset of E . Then K is compact in the weak topology.

Proof. Notice K is a weakly closed subset of a weakly compact set: it is weakly closed by being strongly closed and convex, and it is a subset of λB_E . \square

Corollary 6.24. (l.s.c.)

Theorem 6.25. Let E and F be two reflexive Banach spaces. Let $A : D(A) \subset E \rightarrow F$ be an unbounded linear operator that is densely defined and closed. Then $D(A^*)$ is dense in F^* . Thus A^{**} is well defined and

$$A^{**} = A.$$

Proof. ($D(A^*)$ is dense in F^*) By corollary 4.12 it suffices to show that every linear functional that vanishes on $D(A^*)$ is identically zero. Let $\varphi \in F^{**}$ vanish on $D(A^*)$. Since F is reflexive, we may identify φ with an element in F that we denote with the same symbol. Then we have:

$$0 = \langle \varphi, w \rangle = \langle J\varphi, w \rangle = \langle w, \varphi \rangle \quad \forall w \in D(A^*).$$

If $\varphi \neq 0$, then $[0, \varphi] \neq G(A)$ since A is linear and the image of zero must be zero. Since $G(A)$ is closed, we may strictly separate $G(A)$ and $[0, \varphi]$ with a closed hyperplane. That is, there is a functional $[f, v] \in E^* \times F^*$ and a number $\alpha \in \mathbb{R}$ such that

$$\langle f, u \rangle + \langle v, Au \rangle < \alpha < \langle v, \varphi \rangle \quad \forall u \in D(A).$$

I understand this argument shows that the pullback of v is a continuous linear functional on E , that is, it is in the domain of the adjoint. Also it shows that $\langle v, \varphi \rangle \neq 0$, a contradiction when choosing $w = v$ as above.

($A^{**} = A$) **Orthogonality relations.** \square

Remark 6.14. Linear operators of reflexive Banach spaces are self-adjoint.

6.4 Separable spaces

A metric space E is *separable* if there exists a subset $D \subset E$ that is countable and dense.

Finite-dimensional spaces are separable. L^p and ℓ^p are separable for $1 \leq p < \infty$. The space of continuous functions on a compact metric space $C(K)$ is also separable. L^∞ and ℓ^∞ are not separable.

Proposition 6.26. Any subset of a separable space is separable.

Proposition 6.27. Let E be a Banach space. If E^* is separable then E is separable.

Remark 6.15. The converse is not true by the case of L^1 and L^∞ .

Corollary 6.28. Let E be a Banach space. Then

$$E \text{ is reflexive and separable} \iff E^* \text{ is reflexive and separable.}$$

Proof. Using proposition 6.27 and corollary 6.22. □

Recall that a topological space is said to be *metrizable* if there exists a metric on X that induces the topology of X .

Theorem 6.29. Let E be a separable Banach space. Then B_{E^*} is metrizable in the weak* topology. Conversely, if B_{E^*} is metrizable in the weak* topology, then E is separable.

There is a “dual” statement:

Theorem 6.30. Let E be a Banach space such that E^* is separable. Then B_E is metrizable in the weak topology. Conversely, if B_E is metrizable in the weak topology, then E^* is separable.

Proof. The metric is induced by the following norm on E^* :

$$[f] = \sum_{n=1}^{\infty} \frac{1}{2^n} |\langle f, x_n \rangle|$$

for $f \in E^*$. □

Remark 6.16. We emphasize again that in infinite-dimensional spaces the weak topology and weak* is *not* metrizable. In particular, the topology induced by the norm $[\]$ is not the weak* topology.

Corollary 6.31. Let E be a separable Banach space and let (f_n) be a bounded sequence in E^* . Then there exists a subsequence (f_{n_k}) that converges in the weak* topology.

Proof. Follows from Banach-Alaoglu and theorem 6.29 since it suffices to check the property in the closed unit ball. □

Proof of theorem 6.19. Uses the last theorem and other results developed so far. □

6.5 Uniformly Convex Spaces

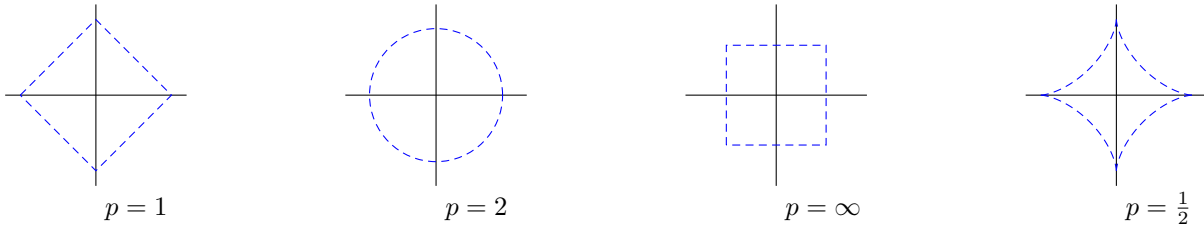
A Banach space E is said to be *uniformly convex* if

$$\forall \varepsilon > 0 \quad \exists \delta > 0 \text{ such that}$$

$$[x, y \in E, \|x\| \leq 1, \|y\| \leq 1 \text{ and } \|x - y\| > \varepsilon] \implies \left[\left\| \frac{x + y}{2} \right\| < 1 - \delta \right]$$

That is, if we slide a rule of length $\varepsilon > 0$ in the unit ball, then its midpoint must stay within a ball of radius $(1 - \delta)$ for some $\delta > 0$. This implies, in particular, that sphere must be “round”, and cannot include any line segment.

Example. In \mathbb{R}^2 , the norm $\|x\|_2 = [|x_1|^2 + |x_2|^2]^{1/2}$ is uniformly convex, while the norms $\|x\|_1 = |x_1| + |x_2|$ and $\|x\|_\infty = \max\{|x_1|, |x_2|\}$ are not.



Example. L^p spaces are uniformly convex for $1 < p < \infty$ as are Hilbert spaces.

Theorem 6.32 (Mulman-Pettis). Every uniformly convex Banach space is reflexive.

Remark 6.17. Uniform convexity is a geometric property, while reflexivity is a topological property: the former is not invariant while interchanging equivalent norms, while the second is. It is striking that a geometric property implies a topological one. While uniform convexity can be a useful tool for proving reflexivity, bear in mind that there are some weird reflexive spaces that admit no uniformly convex equivalent norm!

Proof. It suffices to show that $J(B_E) = B_{E^{**}}$. Let $\xi \in E^{**}$ with $\|\xi\| = 1$ and let us show that $\xi \in J(B_E)$. Since $J(B_E)$ is closed in E^{**} it suffices to prove that there is a sequence in B_E whose images under J converge to ξ , that is,

$$\forall \varepsilon > 0 \quad \exists x \in B_E \text{ such that } \|\xi - J(x)\| \leq \varepsilon.$$

Such an x is constructed using the modulus of uniform convexity δ for any given $\varepsilon > 0$ and Goldstine's lemma (which states that $J(B_E)$ is dense in $B_{E^{**}}$ with respect to $\sigma(E^{**}, E^*)$, and consequently $J(E)$ is dense in E^{**} with respect to $\sigma(E^{**}, E^*)$). \square

Proposition 6.33. Assume that E is a uniformly convex Banach space. Let (x_n) be a sequence in E such that $x_n \rightharpoonup x$ weakly and

$$\limsup \|x_n\| \leq \|x\|.$$

Then $x_n \rightarrow x$ strongly.

6.6 Exercises

Exercise (3.1). If a subset of a Banach space is weakly compact then it is bounded.

Exercise (3.3). The weak and strong closures of a convex subset of a Banach space coincide.

Exercise (3.5). A weakly convergent sequence in a strongly compact subset of a Banach space converges strongly.

Exercise (3.20). Let E be a Banach space.

1. Prove that there exist a compact topological space K and an isometry from E into $C(K)$ equipped with its usual norm. **Hint:** take $K = B_{E^*}$ equipped with $\sigma(E^*, E)$.
2. Assuming E is separable, prove that there exists an isometry from E into ℓ^∞ .

Exercise (3.21). Prove without using metrizability that any bounded sequence in the dual space of a Banach space has a weakly* convergent subsequence. **Hint:** use a diagonal process.

Exercise. Let K be a compact metric space that is not finite. Prove that $C(K)$ is not reflexive. **Hint:** let (a_n) be a sequence in K such that $a_n \rightarrow a$ and $a_n \neq a \forall n$. Consider the linear functional $f(u) = \sum_{n=1}^{\infty} \frac{1}{2^n} u(a_n)$ with $u \in C(K)$ and proceed as in exercises 1.3 and 1.4.

7 L^p spaces.

Let $(\Omega, \mathcal{M}, \mu)$ be a measure space, that is, Ω is a set, \mathcal{M} is a σ -algebra in Ω (a family of subsets containing \emptyset and closed under complements and countable unions), μ is a measure (a function $\mathcal{M} \rightarrow [0, \infty]$ such that $\mu(\emptyset) = 0$ and $\mu(\bigcup_{n=1}^{\infty} A_n) = \sum_{n=1}^{\infty} \mu(A_n)$ whenever A_n are disjoint) and Ω is σ -finite (there is a countable family of measurable sets of finite measure that covers Ω).

We denote by L^1 the space of integrable functions on $(\Omega, \mathcal{M}, \mu)$ (=their integral is finite).

Theorem 7.1. Para todos $p, q \in (1, \infty)$ com $\frac{1}{p} + \frac{1}{q} = 1$,

$$\|\varphi\psi\|_1 \leq \|\varphi\|_p \|\psi\|_q \quad \forall \varphi \in L^p, \psi \in L^q.$$

Proof.

Claim 7.1. Dados $p, q \in [1, \infty]$ com $\frac{1}{p} + \frac{1}{q} = 1$, vale

$$a^p b^q \leq \frac{a}{p} + \frac{b}{q}.$$

Proof. Considere $\alpha > 0$ e $f : (0, \infty) \rightarrow \mathbb{R}$ dada por $f(t) = t^\alpha - \alpha t$. Assim, $f'(t) = \alpha t^{\alpha-1} - \alpha$. Note que o único ponto crítico da função f é $t = 1$, e de fato $f(t) \leq f(1)$, ie.

$$\begin{aligned} t^\alpha &\leq f(1) \quad \forall t \in 1 \\ t^\alpha &\leq \alpha t + (1 - \alpha) \quad \forall t > 0. \end{aligned}$$

Tomando $t = \frac{a}{b}$ e $\alpha = \frac{1}{p}$ obtemos

$$\begin{aligned} \left(\frac{a}{b}\right)^{\frac{1}{p}} &\leq \frac{1}{p} \left(\frac{a}{b}\right) + \frac{1}{q} \\ \iff a^{\frac{1}{p}} b^{\frac{1}{q}-1} &\leq \frac{1}{p} \left(\frac{a}{b}\right) + \frac{1}{q} \\ \iff a^{\frac{1}{p}} b^{\frac{1}{q}} &\leq \frac{a}{p} + \frac{b}{q}. \end{aligned}$$

□

Tomando

$$a = \frac{|\varphi(x)|^p}{\|\varphi\|_p^p} \quad b = \frac{|\psi(x)|^q}{\|\psi\|_q^q}.$$

obtemos

$$\left(\frac{|\varphi(x)|^p}{\|\varphi\|_p^p}\right)^{\frac{1}{p}} \left(\frac{|\psi(x)|^q}{\|\psi\|_q^q}\right)^{\frac{1}{q}} \leq \frac{|\varphi(x)|^p}{p\|\varphi\|_p^p} + \frac{|\psi(x)|^q}{q\|\psi\|_q^q}$$

que uma expressão longa que se simplifica facilmente:

$$\begin{aligned} \frac{|\varphi(x)|}{\|\varphi\|_p} \frac{|\psi(x)|}{\|\psi\|_q} &\leq \frac{|\varphi(x)|^p}{p\|\varphi\|_p^p} + \frac{|\psi(x)|^q}{q\|\psi\|_q^q} \\ \implies |\varphi(x)||\psi(x)| &\leq \left(\frac{|\varphi(x)|^p}{p\|\varphi\|_p^p} + \frac{|\psi(x)|^q}{q\|\psi\|_q^q} \right) \|\varphi\|_p \|\psi\|_q \end{aligned}$$

Integrando,

$$\|\varphi\psi\|_1 \leq \|\varphi\|_p \|\psi\|_q \left(\frac{1}{p} \underbrace{\frac{\int |\varphi|^p d\mu}{\|\varphi\|_p^p}}_1 + \frac{1}{q} \underbrace{\frac{\int |\psi|^q d\mu}{\|\psi\|_q^q}}_1 \right) = \|\varphi\|_p \|\psi\|_q$$

□

Theorem 7.2 (Desigualdade de Minkowski). Dado $p \in [1, \infty)$,

$$\|\varphi + \psi\|_p \leq \|\varphi\|_p + \|\psi\|_p \quad \forall \varphi, \psi \in L^p(\mu).$$

Proof. Observe que

$$\begin{aligned} |\varphi(x) + \psi(x)|^p &\leq (|\varphi(x)| + |\psi(x)|)^p \\ &\leq (2 \max\{|\varphi(x)|, |\psi(x)|\})^p \\ &\leq 2^p \max\{|\varphi(x)|^p, |\psi(x)|^p\} \\ &\leq 2^p (|\varphi(x)|^p + |\psi(x)|^p) \end{aligned}$$

□