NOTES OF FUNCTIONAL ANALYSIS (Fall 2016)

by Martin Brokate (Technische Universität München) and

Augusto Visintin (Università di Trento)

.

0 Program

Questo corso introduce gli elementi essenziali dell'Analisi Funzionale, rinviando le applicazioni alle PDEs al relativo corso della laurea magistrale.

L'Analisi Funzionale rientra tra i requisiti per tutti i percorsi, salvo quelli in crittografia e didattica, per la laurea magistrale in matematica a Trento.

Prerequisiti

Calcolo differenziale ed integrale, con serie di Fourier ed ODE *(Analisi I, II e III).*

Teoria della misura di Lebesgue e dell'integrazione *(Analisi III).*

Algebra lineare *(Geometria I).*

Topologia generale *(Geometria II).*

Programma

— *Spazi di Banach.* Spazi normati e di Banach. Lemma di Riesz e caratterizzazione degli spazi di dimensione finita. Spazio degli operatori lineari e continui. Spazio duale.

Teoremi di Hahn-Banach, di Baire, di Banach-Steinhaus, dell'applicazione aperta, del grafico chiuso. Teoremi di separazione.

Convergenze debole e debole star. Teoremi di Banach-Alaoglu e di Mazur. Seminorme e spazi di Fréchet.

— *Spazi di funzioni o di successioni.* C^k, L^p, ℓ^p, c, ecc..

Definition and main properties of these spaces. Teorema di Ascoli-Arzelà.

— *Spazi di Hilbert.* Prodotto scalare, spazi di Hilbert. Ortogonalità. Proiezione ortogonale su un convesso chiuso.

Teorema di rappresentazione di Riesz-Fréchet. Teorema di Lax-Milgram. Sistemi ortonormali, basi hilbertiane, coefficienti di Fourier.

— *Operatori lineari e continui.* Operatori compatti in spazi di Banach. Teoria di Riesz-Schauder ed alternativa di Fredholm.

Esercitazioni

Queste tratteranno semplici quesiti riguardanti la teoria sopra indicata.

Particolare attenzione verrà dedicata alla discussione di esempi e controesempi. Alcuni controesempi riguarderanno enunciati ottenuti rimuovendo un'ipotesi dai teoremi sopra indicati.

Riferimento essenziale restano le note del corso.

Alcuni testi di consultazione:

R. Bhatia: *Notes on Functional Analysis. (Lez. 1–15.)* Hindustan Book Agency, New Delhi 2009 [un testo introduttivo]

H. Brezis: *Functional Analysis, Sobolev Spaces and Partial Differential Equations.* Springer, New York 2011 [un testo avanzato, con numerosi esercizi svolti]

G. Teschl: *Topics in Real and Functional Analysis. (Part 1)* Note disponibili in rete. [un testo introduttivo]

Modalità di esame

Questo comprende una prova scritta con esercizi e quesiti di teoria, seguito da una prova orale.

0.1 Note

In Heaven there is the golden Book of Mathematics, and of course it includes the proof of every assertion. On the Earth any mathematical text includes unproved statements, as does any other text of finite length.

In these notes, if nothing is mentioned, then it should be understood that the justification is fairly simple and is left to the reader. In some cases the symbol [Ex] is appended, to mean that the justification is not difficult and is explicitly proposed as an exercise. The symbol [] is used to mean that the omitted justification is nontrivial.

A bullet • is used to mark the most important results. The symbol * is used to mark parts that may be skipped.

Chapter I – Banach spaces

Contents: 0. Review of L^p Spaces 1. Basic Notions. 2. Bases, Product and Quotient Spaces. 3. The Hahn-Banach Theorem. 4. Separation. 5. The Baire Theorem and its Consequences.

6. Weak Topologies. 7. Dimension. 8. Compactness. 9. The Ascoli-Arzelà Theorem. 10. Adjoint Operator.

0 Review of L^p Spaces

0.1 Three fundamental inequalities

Lemma 0.1 (*Young Inequality) For any* $p, q > 1$ *such that* $1/p + 1/q = 1$*,*

$$ab \leq \frac{a^p}{p} + \frac{b^q}{q} \qquad \forall a, b \geq 0. \tag{0.1}$$

Proof. Without loss of generality we may assume that $a, b > 0$. By the concavity of the logarithm function we have

$$\log\left(\frac{a^p}{p} + \frac{b^q}{q}\right) \geq \frac{1}{p}\log(a^p) + \frac{1}{q}\log(b^q) = \log a + \log b = \log(ab)\,.$$

As the exponential function is monotone, by passing to the exponentials we get (0.1). □

Let $(A, \mathcal{A}, \mu)$ be a measure space with μ a positive measure, and denote by $\mathcal{M}(A, \mathcal{A}, \mu)$, or just $\mathcal{M}(A)$, the linear space of equivalence classes of μ-a.e. coinciding measurable functions $A \to \mathbb{K}$. The sets

$$\begin{aligned} L^p(A) &:= \{v \in \mathcal{M}(A) : \|v\|_p := \big(\textstyle\int_\Omega |v(x)|^p \, d\mu(x)\big)^{1/p} < \infty\} \ \ (0 < p < \infty), \\ L^\infty(A) &:= \big\{v \in \mathcal{M}(A) : \|v\|_\infty := \operatorname*{ess\,sup}_\Omega |v| < \infty\big\}, \end{aligned} \tag{0.2}$$

where $\operatorname{ess\,sup}_\Omega |v| := \inf_{\mu(N)=0} \sup_{x \in \Omega \setminus N} |v(x)|$, are linear subspaces of $\mathcal{M}(A)$.

• **Theorem 0.2** (*Hölder Inequality) For any* $p, q \in [1, +\infty]$ *with* $1/p + 1/q = 1$*,* [1]

$$uv \in L^1(A), \qquad \int_A |u(x)v(x)| \, d\mu(x) \leq \|u\|_p \|v\|_q \qquad \forall u \in L^p(A), \forall v \in L^q(A). \tag{0.3}$$

[1] Here and in the following, this means that $q = \infty$ if $p = 1$, and $p = \infty$ if $q = 1$.

Proof. We may assume that $u, v \not\equiv 0$ (the null function) and that both p and q are finite and different from 1, since otherwise the result is trivial. After replacing u by $\widetilde{u} := u/\|u\|_p$ and v by $\widetilde{v} := v/\|v\|_q$, we are reduced to proving that

$$\widetilde{u}\widetilde{v} \in L^1(A), \qquad \int_A |\widetilde{u}(x)\widetilde{v}(x)|\, d\mu(x) \leq 1 \qquad \forall u \in L^p(A), \forall v \in L^q(A). \tag{0.4}$$

The Young inequality (0.1) yields

$$|\widetilde{u}(x)\widetilde{v}(x)| \leq \frac{1}{p}|\widetilde{u}(x)|^p + \frac{1}{q}|\widetilde{v}(x)|^q \qquad \text{for a.e. } x \in A.$$

Integrating over A we get $\widetilde{u}\widetilde{v} \in L^1(A)$ and

$$\textstyle\int_A |\widetilde{u}(x)\widetilde{v}(x)|\, d\mu(x) \leq \frac{1}{p}\int_A |\widetilde{u}(x)|^p\, d\mu(x) + \frac{1}{q}\int_A |\widetilde{v}(x)|^q\, d\mu(x) = \frac{1}{p} + \frac{1}{q} = 1\,,$$

that is (0.4). □

Proposition 0.3 (*Minkowski Inequality) For any* $p \in [1, +\infty]$,

$$\|u + v\|_p \leq \|u\|_p + \|v\|_p \qquad \forall u, v \in L^p(A). \tag{0.5}$$

Proof. We may confine ourselves to the case $1 < p < +\infty$, for otherwise the statement is obvious. For a.e. $x \in A$,

$$|u(x) + v(x)|^p \leq 2^p(\max\{|u(x)|, |v(x)|\})^p \leq 2^p\big(|u(x)|^p + |v(x)|^p\big)\,;$$

by integrating over A we may conclude that $u+v \in L^p(A)$. Setting $q := p/(p-1)$, by the Hölder inequality we have

$$\begin{aligned}\|u + v\|_p^p &= \int_A |u(x) + v(x)||u(x) + v(x)|^{p-1}\, d\mu(x)\\ &\leq \int_A |u(x)||u(x) + v(x)|^{p-1}\, d\mu(x) + \int_A |v(x)||u(x) + v(x)|^{p-1}\, d\mu(x)\\ &\leq \big(\|u\|_p + \|v\|_p\big)\|\, |u + v|^{p-1}\|_q = \big(\|u\|_p + \|v\|_p\big)\|u + v\|_p^{p-1}.\end{aligned}$$

By dividing both members by $\|u + v\|_p^{p-1}$, we get the desired inequality. □

The Minkowski inequality is the triangular inequality for L^p spaces, which are thus normed spaces. We shall see that these spaces are also complete, so that they are Banach spaces.

ℓ^p spaces and discrete inequalities. For any $p \in [1, +\infty]$, the L^p space constructed over the set $\mathbb{N}$ equippped with the counting measure (i.e., the measure that associates to any finite subset of $\mathbb{N}$ the number of its elements) is denoted by ℓ^p. These *sequence spaces* play an important role in functional analysis.

Selecting $A = \{1, \ldots, M\}$ and μ equal to the counting measure, the Hölder inequality provides a discrete version for finite sums:

$$\textstyle\sum_{n=1}^M |a_n b_n| \leq \Big(\sum_{n=1}^M |a_n|^p\Big)^{1/p} \Big(\sum_{n=1}^M |b_n|^q\Big)^{1/q} \tag{0.6}$$

for any $a_1, ..., a_M, b_1, ..., b_M \in \mathbb{C}$ and any $M \in \mathbb{N}$. Passing to the limit, one then gets the Hölder inequality for ℓ^p spaces.

Similarly, a discrete version of the Minkowski inequality provides the triangular inequality for ℓ^p spaces.

0.2 Nesting of L^p and ℓ^p spaces

We claim that for any μ-measurable set A of finite measure and any $p, q \in [1, +\infty]$

$$1 \leq p < q \leq +\infty \qquad \Rightarrow \qquad L^q(A) \subset L^p(A). \tag{0.7}$$

This inclusion is easily checked if either p or q equals either 1 or $+\infty$. [Ex] Let us then assume that $1 < p < q < +\infty$. Notice that the exponents $r = q/p$ and $s = q/(q-p)$ are conjugate. For any $v \in L^q(A)$, the Hölder inequality then yields

$$\|v\|_p^p = \int_A |v(x)|^p \cdot 1 \, d\mu(x) \leq \| \, |v(x)|^p \, \|_r \, \|1\|_s = \|v(x)\|_q^p \, \mu(A)^{1/s}.$$

This proves the claim. This inequality actually shows that the injection $L^q(A) \to L^p(A)$ is continuous: if a sequence converges in $L^q(A)$, then it converges to the same limit in $L^p(A)$.

The analogous statement fails if $\mu(A) = +\infty$. For instance, let $A =]-1, +\infty[$ and $1 \leq p < q < +\infty$; there exists $\alpha > 0$ such that $\alpha p < 1 < \alpha q$; then $x^{-\alpha} \in L^q(A) \setminus L^p(A)$.

For ℓ^p spaces the inclusions are reversed:

$$1 \leq p < q \leq +\infty \qquad \Rightarrow \qquad \ell^p \subset \ell^q. \qquad \text{[Ex]} \tag{0.8}$$

Remark. Why are here inclusions reversed? This may be understood considering that functions defined on a finite measure set may have a large L^p-norm only if somewhere they are large. On the other hand sequences may have a large ℓ^p-norm only if they decay at infinity rather slowly. Moreover the behaviors of powers of large real numbers is opposite to that of small values: as p increases, x^p increases for any $x > 1$, and instead decreases for any $0 < x < 1$. (This is clearly seen by drawing the graph of these functions.)

0.3 Some basic properties of L^p and ℓ^p spaces

These spaces play an important role in functional analysis, since they are a large source of examples and counterexamples. Here we state some of their properties without proofs.

From now on, our measure space will be a (possibly unbounded) Euclidean open set (i.e., an open subset of $\mathbb{R}^N$ for some integer N), denoted by Ω, equipped with the standard Lebesgue measure on the Borel σ-algebra. The following is a classical result of measure theory, and is at the basis of the importance of these spaces for analysis.

• **Theorem 0.4** (*Fischer-Riesz*) *For any $p \in [1, +\infty]$, the normed space $L^p(\Omega)$ is complete. []*

L^p-spaces are thus Banach spaces.

We shall denote by $C_c^0(\Omega)$ the linear space of compactly supported continuous functions $\Omega \to \mathbb{R}$; this is obviously a linear subspace of $L^p(\Omega)$ for any $p \in [1, +\infty]$.

• **Theorem 0.5** (*Density*) *For any $p \in [1, +\infty[$, the linear space $C_c^0(\Omega)$ is dense in $L^p(\Omega)$. This fails for $L^\infty(\Omega)$. []*

Proof. Let us fix any $u \in L^p(\Omega)$, which we may assume to be real and nonnegative without loss of generality. The truncated functions

$$\widetilde{u}_n(x) = \min\{u(x), n\} \quad \forall x \in B(0,n), \qquad \widetilde{u}_n(x) = 0 \quad \forall x \in \Omega \setminus B(0,n),$$

are measurable and bounded, satisfy $0 \le u_n \le u$ and converge pointwise a.e. to u. Hence $u_n \in L^p(\Omega)$ and $u_n \to u$ in $L^p(\Omega)$ by dominated convergence.

By a classical theorem of Lusin, [2] for any n there exists a compactly supported continuous function $v_n : \Omega \to \mathbb{R}$ such that, setting $\Omega_n := \{x \in \Omega : u_n(x) \neq v_n(x)\}$ and $|\Omega_n| \le n^{-(p+1)}$. Hence

$$\|u_n - v_n\|_p^p = \int_{\Omega_n} |u_n - v_n|^p \, dx \le \int_{\Omega_n} (2n)^p \, dx \le \frac{2^p}{n} \to 0$$

as $n \to \infty$. Therefore $\|u - v_n\|_p \le \|u - u_n\|_p + \|u_n - v_n\|_p \to 0$ as $n \to \infty$. □

$C^0_c(\Omega)$ is not dense in $L^\infty(\Omega)$. For instance, the sign function ($S(x) = -1$ if $x < 0$, $S(x) = 1$ if $x > 0$) cannot be approximated by continuous functions in the metric of this space. Indeed a uniform limit of continuous functions cannot be continuous.

Theorem 0.6 *For any $p \in [1, +\infty[$, the space $L^p(\Omega)$ is separable (i.e., it has a countable dense subset). This fails for $L^\infty(\Omega)$. []*

Outline of the proof. First notice that Ω can be represented as the union of a countable family $\{K_n\}$ of compact subsets. As on any compact set the L^p-norm is dominated by the uniform norm, by the above density theorem it suffices to approximate uniformly any function of $C^0(K_n)$. By taking coefficients with rational real and imaginary part, and by extending polynomials by null outside the support of the approximated function, by the classical Weierstraß theorem [3] for any n one gets a countable family of functions that uniformly approximate all functions that are supported in K_n. □

In $L^\infty(\Omega)$ counterexamples are easily constructed.

• **Theorem 0.7** (*Fréchet-Riesz*) *Let $p \in]1, +\infty[$ and set $p' = p/(1-p)$ (this is the conjugate exponent). Then*

$$[\Phi_p(f)](v) := \textstyle\int_\Omega f v \, dx \qquad \forall f \in L^{p'}(\Omega), \forall v \in L^p(\Omega) \tag{0.9}$$

defines an isometric isomorphism $\Phi_p : L^{p'}(\Omega) \to L^p(\Omega)'$. []

* *Proof.* (i) By the Hölder inequality (0.3)

$$\Phi_p(f) \in L^p(A)', \qquad \|\Phi_p(f)\|_{L^p(A)'} \le \|f\|_{L^{p'}(A)} \qquad \forall f \in L^{p'}(A); \tag{0.10}$$

thus $\Phi_p(L^{p'}(A)) \subset L^p(A)'$.

(ii) In order to prove the opposite inclusion, let us first show that $\Phi_p(L^{p'}(A))$ is a closed subspace of $L^p(A)'$. As $p = p'/(p'-1)$, by choosing [4] $v = |f|^{p'-2}\overline{f}$ we get $|v|^p = |f|^{p'}$, whence

$$\|f\|_{L^{p'}(A)}^{p'-1} = \Big(\int_A |f|^{p'} \, d\mu \Big)^{(p'-1)/p'} = \Big(\int_A |v|^p \, d\mu \Big)^{1/p} = \|v\|_{L^p(A)}.$$

[2] **Lusin's Theorem:** Let Ω be an open subset of $\mathbb{R}^N$ equipped with the ordinary Lebesgue measure, and $u : \Omega \to \mathbb{C}$ be defined everywhere. Then u is measurable iff for any $\varepsilon > 0$ there exists a compactly supported continuous function $v : \Omega \to \mathbb{C}$ such that $|\{x \in \Omega : u(x) \neq v(x)\}| \le \varepsilon$ and $\sup_\Omega |v| \le \sup_\Omega |u|$.

[3] **Weierstraß Theorem:** If K is a compact subset of $\mathbb{R}^N$, then the linear space of (the restrictions of) polynomials is dense in $C^0(K)$.

[4] By $\overline{f}$ we denote the complex conjugate of f.

Therefore, as $fv = |f|^{p'}$,

$$[\Phi_p(f)](v) = \int_A fv \, d\mu = \|f\|^{p'}_{L^{p'}(A)} = \|f\|_{L^{p'}(A)} \|v\|_{L^p(A)};$$

thus equality holds in (0.10), so that Φ_p is an isometry. As $L^{p'}(A)$ is complete, $\Phi_p(L^{p'}(A))$ is a closed subspace of $L^p(A)'$.

(iii) In order to show that $\Phi_p(L^{p'}(A)) = L^p(A)'$, next we prove that the former space is dense in the second. Because of a Corollary of the Hahn-Banach theorem (see ahead) and of the reflexivity of $L^p(A)$, it suffices to show that

$$v \in L^p(A), \quad [\Phi_p(f)](v) = 0 \quad \forall f \in L^{p'}(A) \quad \Rightarrow \quad v = 0 \quad \text{a.e. in } A.$$

Setting $g = |v|^{p-2}\overline{v} \in L^{p'}(A)$, we get $[\Phi_p(g)](v) = \int_A gv \, d\mu = \int_A |v|^p \, d\mu$. Therefore $[\Phi_p(f)](v) = 0$ entails $v = 0$. □

Next we consider the dual space of $L^1(\Omega)$ and of $L^\infty(\Omega)$.

Theorem 0.8 (*Steinhaus-Nikodým*) *$\Phi_1 : L^\infty(\Omega) \to L^1(\Omega)'$ is an isometric isomorphism. []*

Proposition 0.9 *Define the operator $\Phi_\infty : L^1(\Omega) \to L^\infty(\Omega)'$ as in (0.9), with $\infty' = 1$. This is a nonsurjective isometry. []*

Conclusions. For any $p \in [1, +\infty[$ ($p = \infty$ excluded), denoting the conjugate index by p', we may identify $L^p(\Omega)'$ with $L^{p'}(\Omega)$. This fails for $p = \infty$, since

$$\text{we may just identify } L^1(\Omega) \text{ with a proper closed subspace of } L^\infty(\Omega)'. \tag{0.11}$$

More generally, the same holds for any measure space $(A, \mathcal{A}, \mu)$: [5] we may thus identify $(\ell^p)'$ with $\ell^{p'}$ for any $p \in [1, +\infty[$, but we may just identify ℓ^1 with a proper closed subspace of $(\ell^\infty)'$.
Although in general $L^1(\Omega)$ need not have a predual, we saw that one may identify ℓ^1 with $(c_0)'$. [6] Thus ℓ^1 is the dual of a separable Banach space, at variance with $L^1(\Omega)$.

An Exercise.
(i) For any $p \in [1, 2]$, prove that $L^p(\mathbb{R}) \subset L^1(\mathbb{R}) + L^2(\mathbb{R})$.
(ii) More generally, for any $p, q, r \in [1, +\infty]$, prove that if $p < q < r$ then $L^q(\mathbb{R}) \subset L^p(\mathbb{R}) + L^r(\mathbb{R})$.
(iii) Is $\ell^q \subset \ell^p + \ell^r$ for any $p, q, r \in [1, +\infty]$ with $p < q < r$?

1 Basic Notions

In this section we present the definition of a normed space and some basic constructions.

1.1 Normed and Banach spaces

[5] for $p = 1$ (and just for this index) the measure should be assumed σ-finite; that is, the set A should be representable as a countable union of sets of finite measure.

[6] With standard notation, we shall denote by c the set of converging sequences, by c_0 the set of sequences that tend to 0, and by c_{00} the set the sequences that have just a finite number of nonvanishing terms. All of these are normed subspaces of ℓ^∞; c and c_0 are complete., whereas c is not.

A linear space X over $\mathbb{K}$ ($= \mathbb{C}$ or $\mathbb{R}$) is called a **normed space** iff it is equipped with a **norm**, namely, a function $\|\cdot\| : X \to \mathbb{R}^+$ such that, for any $u, v \in X$ and any $\lambda \in \mathbb{K}$,

$$\begin{aligned} &\|u\| = 0 \quad \Leftrightarrow \quad u = 0, \\ &\|\lambda u\| = |\lambda| \|u\|, \\ &\|u + v\| \leq \|u\| + \|v\|. \end{aligned} \tag{1.12}$$

The function $d(u, v) := \|u - v\|$ is then a distance (or metric) in X. It is easy to check the inverse triangle inequality

$$\|u - v\| \geq \big| \|u\| - \|v\| \big| \qquad \forall u, v \in X. \tag{1.13}$$

Any norm $\|\cdot\|$ on X induces a topology τ, called the **strong topology** or **norm topology**. We say that a sequence $\{u_n\}$ in a normed space X converges iff $d(u_n, u) \to 0$ for a suitable $u \in X$, and then writes $u_n \to u$. If $u_n \to 0$, one calls $\{u_n\}$ a **vanishing** sequence. We say that the series $\sum_{n=1}^{\infty} u_n$ converges iff the sequence of its partial sums $\{S_m := \sum_{n=1}^{m} u_n\}_{m \in \mathbb{N}}$ converges.

A normed space that is complete w.r.t. d is called a **Banach space**. E.g., taking the absolute value (the modulus, resp.) as a norm, $\mathbb{R}$ ($\mathbb{C}$ resp.) becomes a Banach space. More examples are given below.

If X is a linear space we define a **seminorm** as a mapping $p : X \to \mathbb{R}^+$ such that, for any $u, v \in X$ and any $\lambda \in \mathbb{K}$,

$$\begin{aligned} &p(\lambda u) = |\lambda| p(u), \\ &p(u + v) \leq p(u) + p(v), \end{aligned} \tag{1.14}$$

whence $p(0) = 0$. Thus p has all the properties of a norm, except that $p(u) = 0$ does not entail $u = 0$. A linear space equipped with a seminorm is called a seminormed space.

Proposition 1.1 *Any linear subspace M of a normed space X is a normed subspace of X. If X is a Banach space, then so is M iff it is closed. [Ex]*

Henceforth, a **closed subspace** of a normed space X is understood to be a closed linear subspace of X. (Thus, a closed subspace is closed with respect to addition, scalar multiplication, and convergence.) Note that the closure $\overline{M}$ of a linear subspace M is a closed subspace.

1.2 Examples of normed spaces

(i) First we consider the spaces $\mathbb{R}^N$ and $\mathbb{C}^N$ for any integer $N > 0$. [7] For any $p \in [1, \infty[$,

$$\|u\|_p := \Big(\textstyle\sum_{k=1}^{N} |u_k|^p \Big)^{1/p} \qquad \forall u = (u_1, \dots, u_N) \in \mathbb{K}^N, \quad \mathbb{K} = \mathbb{R} \text{ or } \mathbb{C}, \tag{1.15}$$

defines a norm. For $p > 1$ the triangle inequality is just a restatement of the classical **Minkowski inequality,** see Proposition 0.3:

$$\Big(\textstyle\sum_{k=1}^{N} |u_k + v_k|^p \Big)^{1/p} \leq \Big(\textstyle\sum_{k=1}^{N} |u_k|^p \Big)^{1/p} + \Big(\textstyle\sum_{k=1}^{N} |v_k|^p \Big)^{1/p}. \tag{1.16}$$

[7] By convention, $\mathbb{R}^0 := \{0\}$; this is referred to as the trivial space, and will always be excluded.

For $p = \infty$, we define

$$\|u\|_\infty := \max_{1\le k\le N} |u_k| \qquad \forall u = (u_1, \ldots, u_N) \in \mathbb{K}^N,\ \mathbb{K} = \mathbb{R} \text{ or } \mathbb{C}. \tag{1.17}$$

In $\mathbb{R}^N$, the special case $p = 2$ corresponds to the Euclidean distance $d(u, v) := \big(\sum_{k=1}^N |u_k - v_k|^2\big)^{1/2}$, and $\mathbb{R}^N$ is called the Euclidean space of dimension N in this case.

(ii) For any $p \in [1, \infty[$, the **sequence space** $\ell^p_{\mathbb{K}}$ (or ℓ^p, shortly) consists of all sequences $u = \{u_n\}$ such that

$$\|u\|_p := \left(\sum_{k=1}^\infty |u_k|^p \right)^{1/p} < +\infty. \tag{1.18}$$

Passing to the limit as $N \to \infty$ in (1.16) first on the right- and then on the left-hand side, we see that $\ell^p_{\mathbb{K}}$ is closed under addition and that $\|\cdot\|_p$ fulfills the triangle inequality; thus we conclude that $\ell^p_{\mathbb{K}}$ is a normed space. [8]

(iii) For any nonempty set A, let us denote by $B(A;\mathbb{K})$ the set of bounded functions $A \to \mathbb{K}$. This is a linear subspace of $\mathbb{K}^A$, and is a Banach space equipped with the so-called **uniform norm** (or **sup-norm**)

$$\|f\|_\infty = \sup_{x\in A} |f(x)|. \tag{1.19}$$

This topology induces the uniform convergence in A.

(iv) Let A be a compact topological space; $A = [0, 1]$, say. Since the uniform limit of continuous functions is again continuous, the linear space of continuous functions $C^0([0, 1];\mathbb{K})$ is a closed subspace of $B([0, 1];\mathbb{K})$, hence also a Banach space.

(v) For $A = \mathbb{N}$, $B(A;\mathbb{K})$ can be identified with ℓ^∞. Furthermore, the convergent sequences and the vanishing sequences form two closed subspaces of ℓ^∞, which are respectively denoted by c and c_0; these are themselves Banach spaces.

(vi) We define c_{00} to be the linear space of finite sequences, that is, of sequences $u = \{u_n\}$ such that $u_n \neq 0$ for at most finitely many n. Then c_{00} is a linear subspace of all sequence spaces defined above in (ii) and (v); it is not closed (hence, not complete) in the respective norms. The closure of c_{00} equals ℓ^p for $p < \infty$, and c_0 (not ℓ^∞!) for $p = \infty$.

(vii) Let us denote by $\widetilde{\mathcal{M}}(\mathbb{R};\mathbb{K})$ the space of all functions on $\mathbb{R}$ that are measurable with respect to the Lebesgue measure, and by $\mathcal{M}(\mathbb{R};\mathbb{K})$ the quotient obtained by identifying functions that

[8] To show that $\ell^p_{\mathbb{K}}$ is complete, let $\{u^n\}$ with $u^n = \{u^n_k\}_{k\in\mathbb{N}}$ be a Cauchy sequence in $\ell^p_{\mathbb{K}}$. Since $|u^n_k - u^m_k| \le \|u^n - u^m\|_p$, for any $k \in \mathbb{N}$ the sequence of components $\{u^n_k\}_{n\in\mathbb{N}}$ is a Cauchy sequence in $\mathbb{K}$; hence it converges. Let $u = \{u_k\}$ be the sequence defined by $u_k = \lim_{n\to\infty} u^n_k$. For any given $\varepsilon > 0$, let us choose $\widetilde{n}$ such that

$$\|u^n - u^m\|_p = \left(\sum_{k=1}^\infty |u^n_k - u^m_k|^p \right)^{1/p} \le \varepsilon \qquad \forall m, n \ge \widetilde{n}.$$

Moreover, for any N choose $\widetilde{m} \ge \widetilde{n}$ such that $\big(\sum_{k=1}^N |u^{\widetilde{m}}_k - u_k|^p\big)^{1/p} \le \varepsilon$. The Minkowski inequality (1.16) then implies that

$$\Big(\sum_{k=1}^N |u^n_k - u_k|^p\Big)^{1/p} \le \Big(\sum_{k=1}^N |u^n_k - u^{\widetilde{m}}_k|^p\Big)^{1/p} + \Big(\sum_{k=1}^N |u^{\widetilde{m}}_k - u_k|^p\Big)^{1/p} \le 2\varepsilon$$

for all $n \ge \widetilde{n}$ and $N \in \mathbb{N}$. Passing to the limit $N \to \infty$, first we see that $u^n - u \in \ell^p_{\mathbb{K}}$, whence $u \in \ell^p_{\mathbb{K}}$, and then that $\|u^n - u\|_p \to 0$. □

coincide almost everywhere. (We shall systematically assume this identification.) Let us set

$$
\begin{aligned}
&L^p(\mathbb{R};\mathbb{K}) := \Big\{ v \in \mathcal{M}(\mathbb{R};\mathbb{K}) : \|v\|_p := \Big(\int_{\mathbb{R}} |v(x)|^p \, dx \Big)^{1/p} < +\infty \Big\} \quad \forall p \in [1,\infty], \\
&L^\infty(\mathbb{R};\mathbb{K}) := \big\{ v \in \mathcal{M}(\mathbb{R};\mathbb{K}) : \|v\|_\infty := \operatorname*{ess\,sup}_{x\in\mathbb{R}} |v(x)| < +\infty \big\}.
\end{aligned} \tag{1.20}
$$

For any measure space $(A, \mathcal{A}, \mu)$, one similarly defines $L^p(A, \mathcal{A}, \mu; \mathbb{K})$, which is often abbreviated to $L^p(A;\mathbb{K})$ or even $L^p(A)$. In Section XXX we shall see that these are Banach spaces for any $p \in [1, \infty]$.

Other examples of noncomplete normed spaces are: $C^1(K)$ equipped with the norm of $C^0(K)$; $C^0(K)$ equipped with the norm of $L^p(K)$ with $1 \le p < +\infty$; $L^p(A)$ equipped with the norm of $L^q(A)$ with $1 \le q < p$ if $\mu(A) < +\infty$; ℓ^p equipped with the norm of ℓ^q with $p < q$.

1.3 Linear and continuous mappings

If X is a linear space over $\mathbb{K}$, then the set of all linear functionals from X to $\mathbb{K}$ is called the **algebraic dual space** of X; we shall denote it by $X^{\#}$.

If X is a normed space, then the set of all linear and continuous **functionals** [9] we shall denote it by X'. The space X is then called a **pre-dual** of X'. Ahead we shall see that not all spaces have a pre-dual, and that this need not be unique.

The dual $(X')'$ of the dual X' of a normed space X is called the **bidual space** (or simply the **bidual**) of X, and is denoted by X''. A canonical embedding $J : X \to X''$ is defined by setting

$$
(Ju)(f) = f(u) \qquad \forall u \in X, \forall f \in X. \tag{1.21}
$$

If $J(X) = X''$, the space X is called **reflexive**.

More generally, if X_1 and X_2 are normed spaces over the same field, [10] the set of all linear and continuous mappings from X_1 to X_2 is denoted by $\mathcal{L}(X_1; X_2)$, or $\mathcal{L}(X_1)$ if $X_1 = X_2$.

A subset A of a normed space is called **bounded** iff $\sup_{u\in A} \|u\|$ is finite. A mapping $L : X_1 \to X_2$ between normed spaces is called **bounded** iff it maps bounded subsets of X_1 to bounded subsets of X_2.

Proposition 1.2 *Let $L : X_1 \to X_2$ be a linear mapping between normed spaces X_1 and X_2. Let us denote the norm of X_i by $\|\cdot\|_i$, $i = 1, 2$. Then the next three properties are mutually equivalent:*

$$
L \ \ \textit{is continuous,} \tag{1.22}
$$

$$
L \ \ \textit{is bounded,} \tag{1.23}
$$

$$
\exists C > 0 : \forall u \in X_1, \quad \|Lu\|_2 \le C\|u\|_1, \tag{1.24}
$$

$$
\exists C > 0 : \forall u, v \in X_1, \quad \|Lu - Lv\|_2 \le C\|u - v\|_1. \tag{1.25}
$$

Proof. First we show by contradiction that (1.22) implies (1.23). If (1.23) does not hold, then $\|Lv_n\|_2 > n$ for some sequence $\{v_n\}$ with $\|v_n\|_1 \ne 0$ for any n. Hence $u_n = v_n/(n\|v_n\|_1) \to 0$ while $\|Lu_n\|_2 > 1$. This contradicts (1.22).

If (1.23) holds then

$$
\sup\{\|Lv\|_2 : u \in X_1, \|v\|_1 = 1\} \le C < +\infty.
$$

[9] We shall use this terminology: functions map numbers to numbers, functionals map functions to numbers, operators map functions to functions.

[10] Henceforth this will be implied.

As any $u \in X_1$ is of the form $u = \|u\|_1 v$ with $\|v\|_1 = 1$, (1.24) follows.

By replacing u by $u - v$ in (1.24), we get (1.25). Finally, (1.25) obviously yields (1.22). □

Proposition 1.3 *For any normed spaces X_1 and X_2, if X_1 is isomorphic to $\mathbb{K}^N$ for some N, then any linear mapping $L : X_1 \to X_2$ is continuous. [Ex]*

Theorem 1.4 *For any normed spaces X_1 and X_2, $\mathcal{L}(X_1; X_2)$ is a normed space equipped with the norm*

$$\|L\|_{\mathcal{L}(X_1;X_2)} := \sup\{\|Lu\| : u \in X_1, \|u\| \leq 1\} \qquad \forall L \in \mathcal{L}(X_1; X_2). \tag{1.26}$$

If X_2 is a Banach space, then $\mathcal{L}(X_1; X_2)$ is a Banach space.

Proof. We leave the proof of the first statement to the reader, and just prove the second one.

Let $\{L_n\}$ be a Cauchy sequence in $\mathcal{L}(X_1; X_2)$. For any $u \in X_1$, by (1.25) $\{L_n u\}$ is then a Cauchy sequence in X_2. By the completeness of this space, $Lu := \lim_{n\to\infty} L_n u$ then exists. It is easily seen that the linearity of the L_ns implies that of L. By the continuity of the norm, $\|Lu\| = \lim_{n\to\infty} \|L_n u\| \leq \sup_n \|L_n\| \|u\|$. By the boundedness of Cauchy sequences, $C = \sup_n \|L_n\|$ is finite. (1.24) is thus fulfilled, and L is then continuous.

Finally, we show that $L_n \to L$. As $\{L_n\}$ is a Cauchy sequence, for any $\varepsilon > 0$ there exists $\widetilde{n} \in \mathbb{N}$ such that $\|L_m - L_n\| \leq \varepsilon$ for any $m > n \geq \widetilde{n}$. For any $u \in X_1$ we thus have $\|L_m u - L_n u\| \leq \varepsilon \|u\|$. Taking the limit as $m \to \infty$, we get $\|Lu - L_n u\| \leq \varepsilon \|u\|$, whence $\|L - L_n\| \leq \varepsilon$ for any $n \geq \widetilde{n}$. Thus $L_n \to L$ in $\mathcal{L}(X_1; X_2)$. □

In the special case $X_2 = \mathbb{K}$, setting $X_1 = X$ in (1.26) defines the **dual norm**

$$\|f\|_{X'} := \sup\{|f(u)| : u \in X, \|u\| \leq 1\} \qquad \forall f \in X'.$$

Corollary 1.5 *The dual of any normed space is a Banach space.*

Henceforth we shall assume that $X_1 \neq \{0\}$, [11] so that

$$\begin{aligned} \|L\|_{\mathcal{L}(X_1;X_2)} &= \sup\{\|Lu\| : u \in X_1, \|u\| = 1\} \\ &= \sup\{\|Lu\|/\|u\| : u \in X_1, \|u\| \neq 0\}. \end{aligned} \tag{1.27}$$

Notice that

$$\|Lu\| \leq \|L\| \|u\| \qquad \forall u \in X, \tag{1.28}$$

as $\|L\|$ is the smallest constant to be used in the place of C in (1.24). From (1.28) it easily follows that, for any $L_1 \in \mathcal{L}(X_1; X_2)$ and $L_2 \in \mathcal{L}(X_2; X_3)$,

$$\|L_2 \circ L_1\|_{\mathcal{L}(X_1;X_3)} \leq \|L_2\|_{\mathcal{L}(X_2;X_3)} \|L_1\|_{\mathcal{L}(X_1;X_2)}.$$

Proposition 1.6 *Let M be a linear and dense subspace of a normed space X_1, and X_2 be a Banach space. Then every $L \in \mathcal{L}(M; X_2)$ can be uniquely extended to an operator $\widetilde{L} \in \mathcal{L}(X_1; X_2)$.*

Proof. For any $u \in X_1$ and any sequence $\{u_n\}$ in M converging to u, let us set $\widetilde{L}u = \lim_{n\to\infty} Lu_n$. The mapping $\widetilde{L} : X_1 \to X_2$ is a well-defined and continuous, since L is uniformly continuous on the dense subspace M and X_2 is complete. The linearity of $\widetilde{L}$ directly follows from its definition. □

[11] This also is a Banach space!

A linear mapping L between two linear spaces X_1 and X_2 is called an **algebraic isomorphism** iff it is linear and bijective; the same then holds for its inverse L^{-1}. We then write $X_1 \sim X_2$.

A mapping $L \in \mathcal{L}(X_1; X_2)$ between normed spaces X_1 and X_2 is called a **topological isomorphism** (or simply an **isomorphism**) [12] iff it is bijective and its inverse L^{-1} is also continuous, that is, $L^{-1} \in \mathcal{L}(X_2; X_1)$. We then write $X_1 \simeq X_2$. [13]

A linear mapping L between normed spaces X_1 and X_2 is called an **isometry** iff $\|Lu\| = \|u\|$ for every $u \in X_1$. (This makes sense since then $Lu = 0$ only if $u = 0$, so L is injective.) If moreover L is surjective, L is called an **isometric isomorphism**. The spaces X_1 and X_2 are then called isometrically isomorphic, and we write $X_1 \cong X_2$. In this case the spaces X_1 and X_2 have the same metric structure, as $\|Lu - Lv\| = \|L(u-v)\| = \|u - v\|$ for every $u, v \in X_1$. Both L and L^{-1} are then linear and continuous, and $\|L\| = \|L^{-1}\| = 1$.

1.4 Examples of topological dual spaces

(i) For any $N \geq 1$, the dual $(\mathbb{K}^N)'$ is isometrically isomorphic to $\mathbb{K}^N$ via the canonical isomorphism $L : \mathbb{K}^N \to (\mathbb{K}^N)'$, which is defined by

$$(Lu)(v) = \textstyle\sum_{k=1}^N u_k v_k \qquad \forall u, v \in \mathbb{K}^N. \ \ [Ex]$$

(ii) The dual of c_0 is isometrically isomorphic to ℓ^1, because of the next statement.

Theorem 1.7 *The mapping $L : \ell^1 \to c_0'$ defined by*

$$(Lu)(v) = \textstyle\sum_{k=1}^\infty u_k v_k \qquad \forall u \in \ell^1, \forall v \in c_0 \tag{1.29}$$

is a surjective isometry.

Proof. This series is absolutely convergent and bounded by $\|u\|_1 \|v\|_\infty$; thus

$$Lu \in c_0', \qquad \|Lu\|_{c_0'} \leq \|u\|_1. \tag{1.30}$$

For any given $u \in \ell^1$ with $u \neq 0$ and any $n \in \mathbb{N}$, let us define $v^n \in c_0$ by $v_k^n = \text{sign}(u_k)$ if $k \leq n$ and $v_k^n = 0$ otherwise. Then $\|v^n\|_{c_0} = 1$ for any n, and

$$\begin{aligned} \|u\|_1 &= \lim_{n\to\infty} \sum_{k=1}^n |u_k| = \lim_{n\to\infty} \sum_{k=1}^n u_k v_k^n \\ &\leq \limsup_{n\to\infty} \|Lu\|_{c_0'} \|v^n\|_{c_0} = \|Lu\|_{c_0'}. \end{aligned} \tag{1.31}$$

Hence $\|u\|_1 = \|Lu\|_{c_0'}$.

It remains to show that L is surjective. Let $f \in c_0'$ be given, for any $k \in \mathbb{N}$ let us denote by e_k the k-th unit vector (i.e., $e_{k,j} = \delta_{kj}$), and set $u_k = f(e_k)$. For any n, $\| \sum_{k=1}^n \text{sign}(u_k) e_k \|_{c_0} = 1$. We then have

$$\textstyle\sum_{k=1}^n |u_k| = \sum_{k=1}^n \text{sign}(u_k) \cdot f(e_k) = f\left(\sum_{k=1}^n \text{sign}(u_k) e_k\right) \leq \|f\|_{c_0'}.$$

So $u = \{u_k\} \in \ell^1$ and

$$f(v) = \textstyle\sum_{k=1}^\infty v_k f(e_k) = \sum_{k=1}^\infty v_k u_k = (Lu)(v) \qquad \forall v \in c_0. \square$$

[12] The terms **isomorphism of normed spaces** and **linear homeomorphism** are also used in the literature.

[13] Thus, an isomorphism of normed spaces is both an algebraic isomorphism and a homeomorphism in the sense of topology.

(iii) The mapping defined by (1.29) also provides a surjective isometry between $\ell^{p'}$ and the dual of ℓ^p, if $1 \le p < +\infty$ and $p' = p/(p-1)$ which is called the *dual exponent* of p. The structure of this proof is the same as for the Proposition above; for $p > 1$ one uses the Hölder inequality. [Ex] Notice that we are excluding $p = +\infty$; we shall come back to this issue ahead.

(iv) Let A be a compact subset of $\mathbb{R}^N$. It can be proved that the dual space of $C^0(A;\mathbb{K})$ is isometrically isomorphic to (hence, identifiable with) the space of finite regular signed measures on the Borel σ-algebra on A. Thus, any continuous linear functional f on $C^0(A;\mathbb{R})$ can be represented as an integral

$$f(u) = \int_A u(x)\, d\mu(x)$$

for some finite regular signed measure μ.

(v) By the Fréchet-Riesz theorem, the dual of the space $L^p(A)$ is isometrically isomorphic to $L^{p'}(A)$ for $1 \le p < \infty$, where p' is the dual exponent.

* **(vi)** It can be proved that the dual of $L^\infty(A)$ is isometrically isomorphic to the space of finitely additive measures on A and contains $L^1(A)$ as a proper closed subspace.

1.5 Examples of bounded operators

(i) As we already stated, for any matrix $A \in \mathbb{K}^{M,N}$, the associated linear mapping

$$L : \mathbb{K}^N \to \mathbb{K}^M, \qquad (Lu)_j = \sum_{k=1}^N a_{jk} u_k, \quad 1 \le j \le M$$

is bounded.

(ii) The **right** (or **forward**) **shift** S_r and the **left** (or **backward**) **shift** S_l

$$(S_r u)_k = u_{k-1}, \quad (S_\ell u)_k = u_{k+1}, \tag{1.32}$$

are most naturally defined on doubly infinite sequences $\{u_k\}_{k\in\mathbb{Z}}$; obviously they are isometric isomorphisms on $\ell^p_{\mathbb{K}}(\mathbb{Z})$ for any $p \in [1,\infty]$. For unilateral sequences $u = (u_1, u_2, \dots)$ one sets

$$S_r(u_1, u_2, \dots) = (0, u_1, u_2, \dots), \quad S_\ell(u_1, u_2, \dots) = (u_2, u_3, \dots).$$

In this case S_r, S_ℓ still belong to $\mathcal{L}(X)$ for $X = \ell^p_{\mathbb{K}}(\mathbb{N})$, but they are no longer isomorphisms.

(iii) For any $p \in [1,\infty]$, if $a \in L^\infty(A, \mathcal{A}, \mu)$, then the **multiplication operator** defined by

$$(L_a u)(x) = a(x)u(x) \qquad \text{for a.e. } x \in A$$

belongs to $\mathcal{L}(L^p(A))$, and $\|L_a\| = \|a\|_\infty$. Similarly, if A is a compact metric space and $a \in C^0(A)$, then $L_a \in \mathcal{L}(C^0(A))$ and $\|L_a\| = \|a\|_\infty$.

(iv) Let $(A, \mathcal{A}, \mu)$ and $(B, \mathcal{B}, \nu)$ be two σ-finite measure spaces, $k \in L^2(A \times B)$, and set

$$(Lu)(x) = \int_B k(x,y)u(y)\, d\mu(y) \qquad \text{for a.e. } x \in B, \forall u \in L^2(B). \tag{1.33}$$

By the classical theorems of Tonelli and Fubini and the Hölder inequality, Lu is an a.e. well-defined and measurable function, and

$$\int_A \left|\int_B k(x,y)u(y)\, d\nu(y)\right|^2 d\mu(x) \le \int_A\int_B |k(x,y)|^2\, d\nu(y)\, d\mu(x) \cdot \int_B |u(y)|^2\, d\nu(y).$$

L is thus a bounded linear mapping from $L^2(B)$ to $L^2(A)$, and

$$\|L\| \le \iint_{A\times B} |k(x,y)|^2\, d\mu(y)\, d\nu(x).$$

The function k is called the **kernel** of the **integral operator** L.

If $A = B = [a, b]$ and $\mu = \nu$ is the Lebesgue measure, then the operators

$$(L_1 u)(x) = \int_a^b k(x,y)u(y)\,dy, \quad (L_2 u)(x) = \int_a^x k(x,y)u(y)\,dy \quad \forall x \in [a,b],$$

are respectively called **Fredholm** and **Volterra integral operators**.

(v) An **infinite matrix** $A = (a_{jk})$ defines a bounded linear mapping between sequence spaces by the formula

$$(Lu)_j = \sum_{k=1}^{\infty} a_{jk} u_k, \quad 1 \le j < \infty. \tag{1.34}$$

The estimate

$$\sum_{j=1}^{\infty} \left(\sum_{k=1}^{\infty} |a_{jk} u_k|\right)^2 \le \left(\sum_{j,k=1}^{\infty} |a_{jk}|^2\right) \sum_{k=1}^{\infty} |u_k|^2 \tag{1.35}$$

follows from the Hölder inequality. If $\sum_{j,k} |a_{jk}|^2 < +\infty$, this entails that $L \in \mathcal{L}(X)$ for $X = \ell^2$, with $\|L\|^2 \le \sum_{j,k} |a_{jk}|^2$. This condition, however, is not necessary: e.g., the unit matrix does not satisfy it. Indeed, in the diagonal case $(Lu)_k = \alpha_k u_k$, we have $L \in \mathcal{L}(X)$, $X = \ell^p$, iff $\|\alpha\|_\infty < \infty$. One may ask for conditions, in terms of the elements of A, which are necessary as well as sufficient in order that (15.1) defines a bounded linear mapping from ℓ^p to ℓ^q. However, no "useful" ones are known for $p, q \in]1, \infty[$.

2 Bases and Finite Dimensional Spaces

2.1 Bases and dimension

Hamel bases. A subset S of a linear space X is called a Hamel basis (or **algebraic basis)** iff every element of X has a representation as a linear combination of elements of S, [14] and this representation is unique. This holds iff S is linearly independent (i.e., any linear combination of its elements vanishes only if all its coefficients vanish), and S is maximal among all linearly independent subsets of X.

Let us recall the classical *Hausdorff maximality principle:* [15]

Theorem 2.1 *In a partially ordered set, every totally ordered subset is contained in a maximal totally ordered subset.*

Proposition 2.2 *Any linear space X has a Hamel basis.*

Proof. Let us first order by inclusion the family $\mathcal{A}$ $(\subset \mathcal{P}(S))$ of all linearly independent subsets of X. (Thus $\Sigma \subset \mathcal{P}(S)$ for any $\Sigma \in \mathcal{A}$.) By the Hausdorff principle, there exists a maximal totally ordered subset Σ of $\mathcal{A}$. Thus $\bigcup\{A : A \in \Sigma\}$ $(\subset \mathcal{P}(S))$ is linearly independent and maximal among all linearly independent subsets of X. □

If a Hamel basis is either finite, or countably infinite or uncountably infinite, the same holds for any other Hamel basis; in the finite case, moreover, any pair of Hamel bases have the same number of elements. [Ex] We can thus define the **(Hamel) dimension** of the space to be either finite, or countably infinite, or uncountably infinite, depending on any of its Hamel bases. (This is thus a purely algebraic notion.) In the finite case, the number of elements of any Hamel basis is called the dimension of the space.

[14] For us linear combinations are always finite. Notice that, without a notion of convergence, there is no natural way to define infinite linear combinations.

[15] We shall often apply this principle, which is equivalent to the Zorn Lemma (3.1) as well as to the axiom of choice, and might be often replaced by either of them.

In Proposition 5.2 we shall see that no Banach space has countable infinite dimension.

Schauder bases. A sequence $\{u_n\}$ in X is called a **Schauder basis** (or a topological basis) iff every $u \in X$ has a unique representation of the form

$$u = \sum_{n=1}^{\infty} a_n u_n \qquad \text{with } a_n \in \mathbb{K}, \forall n.$$

This may depend on the order in which the elements u_n are enumerated; in other terms, a reordering of a Schauder basis need not be a Schauder basis. [] Obviously, any Schauder basis is a linearly independent subset.

A topological space is called **separable** iff it has a countable dense subset. [16]

Proposition 2.3 *If a normed space has a Schauder basis, then it is separable.*

Proof. (Finite) linear combinations of elements of the basis with rational coefficients form a countable dense subset of the set of all linear combinations of elements of the basis, and by hypothesis this latter set is dense X. ☐

On the other hand, unexpectedly there exist separable Banach spaces without any Schauder basis. (Counterexamples are nontrivial.) []

2.2 Isomorphisms and comparison of norms

Let a linear space X be equipped with two norms $\|\cdot\|_1$ and $\|\cdot\|_2$. We say that $\|\cdot\|_1$ is **weaker** than $\|\cdot\|_2$ on X (and $\|\cdot\|_2$ is **stronger** than $\|\cdot\|_1$) [17] iff there exists $C > 0$ such that

$$\|u\|_1 \le C\|u\|_2 \qquad \forall u \in X, \tag{2.1}$$

or equivalently, iff the identity mapping j from $(X, \|\cdot\|_2)$ to $(X, \|\cdot\|_1)$ is continuous. The two norms are called **equivalent** iff j is a topological isomorphism, or equivalently, iff there exist constants $c, C > 0$ such that

$$c\|u\|_1 \le \|u\|_2 \le C\|u\|_1 \qquad \forall u \in X. \tag{2.2}$$

In terms of the induced topologies τ_1 in $(X, \|\cdot\|_1)$ and τ_2 in $(X, \|\cdot\|_2)$, this means that $\|\cdot\|_2$ is stronger than $\|\cdot\|_1$ iff $\tau_2 \supset \tau_1$ (i.e., τ_2 is finer than τ_1), and the two norms are equivalent iff $\tau_1 = \tau_2$.

Remark. An infinite-dimensional Banach space may be isomorphic to a proper subspace of itself. Indeed, the space c_0 of vanishing sequences is a proper subspace of the space c of convergent sequences, and defining the mapping $L : c \to c_0$ by

$$(Lu)_1 = \lim_{n\to\infty} u_n, \quad (Lu)_k = u_{k-1} - \lim_{n\to\infty} u_n \quad \forall k > 1, \tag{2.3}$$

for any $u = \{u_n\} \in c$, L is an isomorphism. [Ex]

2.3 Product spaces

Let $(X_j, \|\cdot\|_j)$ be normed spaces over the field $\mathbb{K}$, where $1 \le j \le N$ and $N \in \mathbb{N}$. We may equip the product space $X = \prod_{j=1}^{N} X_j$ with a norm as follows, for any $p \in [1, \infty]$:

$$\|u\|_p = \begin{cases} \Big(\sum_{j=1}^{N} \|u_j\|_j^p \Big)^{1/p} & \text{if } 1 \le p < +\infty, \\ \max_{1\le j\le N} \|u_j\|_j & \text{if } p = \infty. \end{cases} \tag{2.4}$$

[16] Despite of the similarity of these terms, there is no relation between separability and the separation properties (e.g., the Hausdorff property) of topological spaces.

[17] By a standard terminology, "weaker" actually means "weaker or equal"; and similarly for "stronger".

Proposition 2.4 *The formula (2.4) defines a norm on the product space $X = \prod_{j=1}^N X_j$. This is a Banach spaces iff so are all the X_js. [Ex]*

Since all norms on $\mathbb{R}^N$ are equivalent, different choices of p in (2.4) lead to equivalent norms on the product space X. So whether another normed space Y is isomorphic to X does not depend on p. On the other hand, whether the two spaces are isometrically isomorphic depends on p. In this context there is no canonical choice of the norm of the product space.

Associated with the product there are the canonical projections

$$p_j : X \to X_j, \quad p_j(u_1, \dots, u_N) = u_j, \tag{2.5}$$

and the canonical injections

$$i_j : X_j \to X, \quad i_j(u_j) = (0, \dots, 0, u_j, 0, \dots, 0). \tag{2.6}$$

It is obvious from the definitions that, for any $j \in \{1, \dots, N\}$, the injection i_j is an isometry, and that $p_j \in \mathcal{L}(X; X_j)$ with $\|p_j\|_{\mathcal{L}(X;X_j)} = 1$ (unless X_j is the trivial space).

The dimension of X is related to the dimension of its factors X_j by the formula

$$\dim\,(X) = \sum_{j=1}^N \dim\,(X_j). \tag{2.7}$$

(This sum is assumed to be infinite if so is at least one of the addenda.) Indeed, if V_j is a Hamel basis of X_j, $1 \le j \le N$, then $V := \bigcup_{j=1}^N i_j(V_j)$ is a Hamel basis of X.

2.4 Quotient spaces

If X is a linear space and M is a linear subspace of X, the **quotient space** X/M consists of all equivalence classes $[u]$ of elements $u \in X$ with $[u] = [v]$ iff $u - v \in M$. Setting

$$[u] + [v] = [u + v], \quad \lambda[u] = [\lambda u] \qquad \forall u, v \in X, \forall \lambda \in \mathbb{K},$$

The linear mapping $q : X \to X/M$ defined by $q(u) = [u]$ is called the **quotient mapping**. The dimension of X/M is called the **codimension** of M in X [18].

Proposition 2.5 *Let M be a closed subspace of a normed space X. Then* [19]

$$\|[u]\| := \inf_{v \in M} \|u - v\| = \inf_{v \in [u]} \|v\| \tag{2.8}$$

defines a norm on X/M with $\|[u]\| \le \|u\|$ for all $u \in X$. This norm is called the ***quotient norm****. The quotient mapping $q : X \to X/M : u \mapsto [u]$ is continuous.*

If X is a Banach space, then so is X/M.

Proof. If $\|[u]\| = 0$, then $\|u - v_n\| \to 0$ for some sequence $\{v_n\}$ in M; so $u \in M$ since M is closed, and thus $[u] = 0$. The seminorm properties (1.14) are obvious from the definitions, as well as the continuity of q.

Let now $\{[u_n]\}$ be a Cauchy sequence in X/M. Possibly passing to a subsequence and relabeling, we may assume that $\|[u_n] - [u_{n-1}]\| \le 2^{-n}$. We may successively choose $\widetilde{u}_n \in [u_n]$ for any n such that

$$\|\widetilde{u}_n - \widetilde{u}_{n-1}\| \overset{(2.8)}{\le} 2\|[u_n] - [u_{n-1}]\| \le 2^{1-n}.$$

[18] We will see later that the dimension of X/M equals the dimension of any complement of M.

[19] $\|[u]\|$ equals the distance $d(u, M)$ between u and M, according to the customary definition of the distance between a point and a set.

Then $\{\widetilde{u}_n\}$ is a Cauchy sequence in X. If X is complete, then there exists $u = \lim \widetilde{u}_n$, and we get $[u_n] - [u] = [\widetilde{u}_n] - [u] = [\widetilde{u}_n - u] \to 0$. $\square$

If M is a linear but nonclosed subspace of X, then (2.8) defines a seminorm; this is not a norm since $\|[u]\| = 0$ for any $u \in \overline{M} \setminus M$.

2.5 Series

A series $\sum_{n=1}^{\infty} u_n$ in a normed space X is called **convergent** iff the sequence formed by its partial sums $\{S_m = \sum_{n=1}^{m} u_n\}$ converges in X. The series is called **totally** (or **absolutely) convergent** iff the numerical series $\sum_{n=1}^{\infty} \|u_n\|$ converges (in $\mathbb{R}$).

Proposition 2.6 *Prove that a normed space X is complete iff any totally convergent series in X is convergent.*

Proof. (i) "If"-part. Let $\{u_j\}$ be a Cauchy sequence in X, and recursively construct a subsequence $\{u_{n_j}\}$ as follows: for any j, n_j is selected so that $n_j > n_{j-1}$ and $\|u_{n_j} - u_m\| \leq 2^{-j}$ for any $m > n_j$. Therefore $\|u_{n_{j+2}} - u_{n_{j+1}}\| \leq 2^{-j}$ for any j, whence $\sum_{j=1}^{\infty} \|u_{n_{j+1}} - u_{n_j}\| < +\infty$. The sequence $\{u_{n_j}\}$ then converges.

(ii) " Only if"-part. It suffices to notice that a sequence $\{u_j\}$ is convergent if so is $\{\|u_j\|\}$, since

$$\Big\| \sum_{j=n}^{m} u_j \Big\| \leq \sum_{j=n}^{m} \|u_j\| \qquad \forall n, m \in \mathbb{N}\ (n < m).\square$$

2.6 Algebraically complements and direct sums

The notion of projection somehow bridges linear, Banach and Hilbert spaces. First we address the algebraic side. Let M_1 and M_2 be two linear subspaces of a linear space V, and set

$$V = M_1 \oplus M_2 \quad \Leftrightarrow \quad M_1 + M_2 = V, \quad M_1 \cap M_2 = \{0\}. \tag{2.9}$$

This holds iff for any $x \in V$ there exists one and only one pair $(x_1, x_2) \in M_1 \times M_2$ such that $x = x_1 + x_2$. We then say that V is the **algebraic direct sum** of M_1 and M_2, that M_1 and M_2 **algebraically complement** (or *supplement*) each other, and that they are algebraically complemented. Notice that then M_1 is isomorphic to the quotient space X/M_2, and symmetrically M_2 is isomorphic to the quotient space X/M_1. The dimension of M_2 is called the **codimension** of M_1. For instance, a hyperplane is a linear subspace of codimension 1. Notice that then M_1 is (linearly) isomorphic to the linear quotient space X/M_2. Thus

$$V = M_1 \oplus M_2 \quad \Rightarrow \quad \text{dim }(M_2) = \text{codim }(M_1) = \text{codim }(X/M_2). \tag{2.10}$$

Any linear subspace A of a linear space V is algebraically complemented. (This may be proved via transfinite induction.) []

For any operator $L : V_1 \to V_2$ between linear spaces, let us set

$$\begin{aligned} &\mathcal{R}(L) := L(V_1)\ (\subset V_2) \quad \text{(range or image of } L), \\ &\mathcal{N}(L) := L^{-1}(0)\ (\subset V_1) \quad \text{(kernel or nullspace of } L). \end{aligned} \tag{2.11}$$

Proposition 2.7 *If V_1 and V_2 are linear spaces and $L : V_1 \to V_2$ is a linear operator, then*

$$\text{codim}(\mathcal{N}(L)) = \text{dim}(V_1/\mathcal{N}(L)) = \text{dim}(\mathcal{R}(L)). \tag{2.12}$$

Proof. It suffices to notice that $V_1 = \mathcal{N}(L) \oplus V_1/\mathcal{N}(L)$ (by an obvious identitification), and the operator L induces a linear isomorphism between $V_1/\mathcal{N}(L)$ and $\mathcal{R}(L)$. □

Remark. On the other hand, in general $\dim(\mathcal{N}(L)) \neq \operatorname{codim}(\mathcal{R}(L))$. Actually, an isomorphism preserves the dimension, but need not preserve the codimension.

2.7 * Projections in linear spaces

A linear operator $P : V \to V$ is called a **projection** on V iff it is **idempotent**, that is, $P^2 = P$ (or equivalently $P(I - P) = 0$).

Proposition 2.8 *Let V be a linear space.*
If P is a projection on V, then $\widetilde{P} := I - P$ is also a projection, and

$$V = \mathcal{R}(P) \oplus \mathcal{N}(P) = \mathcal{N}(\widetilde{P}) \oplus \mathcal{R}(\widetilde{P}), \qquad P\widetilde{P} = \widetilde{P}P = 0. \tag{2.13}$$

Conversely, if M_1, M_2 are two linear subspaces of V and $V = M_1 \oplus M_2$, then there exists a unique pair of projections P and $\widetilde{P}$ on V such that $P + \widetilde{P} = I$ and

$$M_1 = \mathcal{R}(P) = \mathcal{N}(\widetilde{P}), \qquad M_2 = \mathcal{R}(\widetilde{P}) = \mathcal{N}(P), \qquad P\widetilde{P} = \widetilde{P}P = 0. \tag{2.14}$$

Proof. If P is a projection then obviously $I - P$ is a projection, too. For any $x \in V$, $x = P(x) + [x - P(x)]$ with $P(x) \in \mathcal{R}(P) = \mathcal{N}(\widetilde{P})$ and $x - P(x) \in \mathcal{R}(\widetilde{P}) = \mathcal{N}(P)$. Moreover if $x \in \mathcal{R}(P) \cap \mathcal{N}(P)$ then $x = P(x) = 0$; (2.13) thus holds.
Let us now assume that $V = M_1 \oplus M_2$. For any $x \in V$ let (x_1, x_2) be the unique pair of $\in M_1 \times M_2$ such that $x = x_1 + x_2$, and set $P(x) := x_1$. It is straightforward to see that P is a projection on V, $M_1 = \mathcal{R}(P)$ and $M_2 = \mathcal{N}(P)$. By setting $\widetilde{P}(x) := x_2$ we then get (2.14).
Finally, it is clear that P is the unique projection such that $M_1 = \mathcal{R}(P)$ and $M_2 = \mathcal{N}(P)$. □

2.8 * Projections in Banach spaces

We define *continuous* projections in normed spaces.

Proposition 2.9 *Any projection P on a Banach space X is continuous iff both $\mathcal{R}(P)$ and $\mathcal{N}(P)$ are closed.* [20]

Proof. The "only if"-part is obvious, as $\mathcal{R}(P) = \mathcal{N}(I - P)$.
Let us come to the "if"-part. Let $u_n \to u$ and $Pu_n \to w$. As $Pu_n \in \mathcal{R}(P)$ and this set is closed, we infer that $w \in \mathcal{R}(P)$, whence $Pw = w$. Similarly, as $u - Pu_n \in \mathcal{N}(P)$ and this set is closed, we have $u - w \in \mathcal{N}(P)$, whence $Pu = Pw$. Thus $Pu = w$, that is, the graph of P is closed. It then suffice to apply the Closed Graph Theorem 5.12. □

If M_1, M_2 are closed subspaces of a Banach space X and $X = M_1 \oplus M_2$, then X is called a **topological direct sum**; in this case one says that M_1 and M_2 **topologically complement** (or *topologically supplement*) each other in X, and that they are topologically complemented. Notice that then M_1 is (topologically) isomorphic to the quotient space X/M_2.
At variance with what we saw in the purely algebraic set-up,

$$\text{a closed subspace of a Banach space need not be topologically complemented;} \tag{2.15}$$

[20] This may be compared with the following statement: A linear functional $f : X \to \mathbb{R}$ is continuous if (and only if) $f^{-1}(0)$ is closed. [Ex]

i.e., it need not be either the range or the nullspace of any continuous projection. For instance, c_0 has no topological complement in ℓ^∞ (Phillips's theorem), [] although, as we saw, it has an algebraic complement. [21] However a closed subspace of a Banach space is topologically complemented whenever either its dimension or its codimension are finite. [] Moreover, a Banach space X is topologically isomorphic to a Hilbert space iff any closed subspace of X is topologically complemented (Lindenstrauss-Tzafriri's Theorem). []

Remark. Although c_0 is not complemented in ℓ^∞, it can be strongly separated from any $u \in \ell^\infty \setminus c_0$ by a closed affine hyperplane, because of the Tukey and Klee Theorem 4.6.

3 The Hahn-Banach Theorem

In this section we state and prove the real and complex forms of the Hahn-Banach theorem, and draw a number of consequences. This result is one of the pillars of functional analysis; for instance, it entails that the dual X' of any normed space $X \neq \{0\}$ is nontrivial: this is at the basis of the feasibility of *functional* analysis.

3.1 Sublinear functionals

Let X be a linear space over the field $\mathbb{K}$. A functional $p : X \to \mathbb{R}$ is called **sublinear** iff

$$p(\lambda_1 v_1 + \lambda_2 v_2) \leq \lambda_1 p(v_1) + \lambda_2 p(v_2) \qquad \forall v_1, v_2 \in X, \forall \lambda_1, \lambda_2 \geq 0. \tag{3.1}$$

This holds iff p is subadditive and *positively* homogeneous of degree 1, that is,

$$\begin{aligned} &p(v_1 + v_2) \leq p(v_1) + p(v_2) \qquad \forall v_1, v_2 \in X, \\ &p(\lambda v) = \lambda p(v) \qquad \forall v \in X, \forall \lambda \geq 0, \end{aligned} \tag{3.2}$$

or equivalently p is convex and positively homogeneous of degree 1. [Ex]

A more restricted class consists of the functionals $p : X \to \mathbb{R}^+$ that are subadditive and *absolutely* homogeneous of degree 1, that is,

$$\begin{aligned} &p(u + v) \leq p(u) + p(v) \quad \forall v_1, v_2 \in X, \\ &p(\lambda v) = |\lambda| p(v) \qquad \forall v \in X, \forall \lambda \in \mathbb{K}. \end{aligned} \tag{3.2'}$$

These coincide with what we already called **seminorms.** Seminorms are obviously sublinear, but for instance any nontrivial linear functional is sublinear without being a seminorm.

3.2 Hahn-Banach theorem: real version

The Hahn-Banach theorem has several variants. A general formulation applies to any real linear space, and does not assume the presence of any topology. Loosely speaking, it states that any linear functional defined on a linear subspace may be extended to the whole space preserving linearity and without increasing its *size* (suitably defined).

The argument rests upon transfinite induction, that here we apply in the following form:

[21] Let us set $C^0_0([0,1]) := \{v \in C^0([0,1]) : v(0) = v(1) = 0\}$, $C^0_0(\mathbb{R}) := \{v \in C^0(\mathbb{R}) : v(x) \to 0 \text{ as } x \to \pm\infty\}$, and equip these spaces with the uniform norm. Note the analogies:

$$c_{00} \text{ is dense in } c_0; \quad c_0, c \text{ are closed subspaces of } \ell^\infty; \tag{2.16}$$

$$C^0_c(\mathbb{R}) \text{ is dense in } C^0_0(\mathbb{R}); \quad C^0_0(\mathbb{R}), C^0_b(\mathbb{R}) \text{ are closed subspaces of } L^\infty(\mathbb{R}); \tag{2.17}$$

$$C^0_c(]0,1[) \text{ is dense in } C^0_0([0,1]); \quad C^0_0([0,1]), C^0_b([0,1]) \text{ are closed subspaces of } L^\infty(0,1). \tag{2.18}$$

Theorem 3.1 (*Zorn's lemma) If in a partially ordered set S any (nonempty) totally ordered subset has an upper bound,* [22] *then S itself has a maximal element.*

• **Theorem 3.2** (*Hahn-Banach theorem for real vector spaces) Let X be a real linear space, $p : X \to \mathbb{R}$ be a sublinear functional, M be a linear subspace of X, and $f : M \to \mathbb{R}$ be a linear functional such that $f(v) \leq p(v)$ for any $v \in M$. Then there exists a linear functional $\widetilde{f} : X \to \mathbb{R}$ such that $\widetilde{f} = f$ in M and $\widetilde{f}(v) \leq p(v)$ for any $v \in X$.* [23]

(Trivial examples show that the extension need not be unique.)

Proof. [24] The set

$$\Phi := \big\{ g : \mathrm{Dom}(g) \to \mathbb{R} \text{ linear} : \ M \subset \mathrm{Dom}(g) \subset X, g = f \text{ in } M, g \leq p \big\}$$

can be partially ordered by setting

$$g_1 \preceq g_2 \quad \Leftrightarrow \quad \mathrm{Dom}(g_1) \subset \mathrm{Dom}(g_2), \ \ g_1 = g_2 \text{ in } \mathrm{Dom}(g_1).$$

We claim that this ordering is inductive. In fact, if $\{g_i\}_{i \in I}$ is a totally ordered subset of Φ, then, setting $\mathrm{Dom}(g) := \bigcup_i \mathrm{Dom}(g_i)$ and $g(u) = g_i(u)$ for any $u \in \mathrm{Dom}(g_i)$, we get $g \in \Phi$ and $g_i \preceq g$ for any i. By Zorn's Lemma (3.1) then there exists a maximal element $\widetilde{f} \in \Phi$.

At this point it suffices to prove that $\mathrm{Dom}(\widetilde{f}) = X$. By contradiction, let $u_0 \in X \setminus \mathrm{Dom}(\widetilde{f})$, fix an $\alpha \in \mathbb{R}$, and define a linear functional h by setting

$$\begin{aligned} \mathrm{Dom}(h) &:= \{v + \lambda u_0 : v \in \mathrm{Dom}(\widetilde{f}), \lambda \in \mathbb{R}\}, \\ h(v + \lambda u_0) &:= \widetilde{f}(v) + \lambda\alpha \qquad \forall v \in \mathrm{Dom}(\widetilde{f}), \forall \lambda \in \mathbb{R}. \end{aligned} \tag{3.3}$$

We claim that α may be chosen in such a way that $h \leq p$. This will entail that $\widetilde{f} \preceq h$, contradicting the maximality of $\widetilde{f}$, and will thus complete the proof.

By the positive homogeneity of p, it suffices to prove that

$$\widetilde{f}(v) + \lambda\alpha \leq p(v + \lambda u_0) \qquad \forall v, w \in \mathrm{Dom}(\widetilde{f}), \lambda = \pm 1. \tag{3.4}$$

For any $v, w \in \mathrm{Dom}(\widetilde{f})$, the linearity of $\widetilde{f}$, the inequality $\widetilde{f} \leq p$ in $\mathrm{Dom}(\widetilde{f})$, and the subadditivity of p yield

$$\widetilde{f}(v) + \widetilde{f}(w) = \widetilde{f}(v + w) \leq p(v + w) \leq p(v - u_0) + p(w + u_0),$$

whence

$$\sup_{v \in \mathrm{Dom}(\widetilde{f})} \widetilde{f}(v) - p(v - u_0) \leq \inf_{w \in \mathrm{Dom}(\widetilde{f})} p(w + u_0) - \widetilde{f}(w).$$

Thus there exists $\alpha \in \mathbb{R}$ such that

$$\widetilde{f}(v) - p(v - u_0) \leq \alpha \leq p(w + u_0) - \widetilde{f}(w).$$

This coincides with (3.4), as we had to prove. □

[22] Whenever this holds one says that S is *inductive*.

[23] If we drop any reference to the sublinear functional p in the theorem (and its proof), we obtain the standard extension theorem of linear functionals in linear spaces.

[24] This proof is nonconstructive. No constructive argument is actually known.

3.3 Hahn-Banach theorem: complex version

Next we provide a different formulation of the Hahn-Banach theorem, that also applies to complex linear spaces, but assumes that the sublinear functional is a seminorm. (Notice that for $\mathbb{K} = \mathbb{C}$ the homogeneity property $(1.12)_2$ is stronger than the positive homogeneity.)

First, for any complex linear space $V_{\mathbb{C}}$, let us denote by $V_{\mathbb{R}}$ the associated real linear space. Notice that $Im(z) = Re(-iz) = -Re(iz)$ for any complex number z.

Lemma 3.3 *The real part g of any linear functional f on a linear space $V_{\mathbb{C}}$ over $\mathbb{C}$ is a linear functional on the associated linear space $V_{\mathbb{R}}$ over $\mathbb{R}$. Viceversa, any linear functional g on $V_{\mathbb{R}}$ is the real part of the linear functional $f(v) = g(v) - ig(iv)$ on $V_{\mathbb{C}}$. Moreover, if $V_{\mathbb{C}}$ is a normed space, then f is continuous iff so is g, and $\|f\| = \|g\|$ in that case.*

Proof. We just prove the final assertion, as the remainder is easily checked. For any $v \in X$, $|g(v))| \leq |f(v)|$. On the other hand, if $f(v) = re^{i\theta}$ (with $r, \theta \geq 0$), then $g(e^{-i\theta}v) = \text{Re}(f(e^{-i\theta}v)) = \text{Re}(e^{-i\theta}f(v)) = r = |f(v)|$. Hence $\|f\| = \|g\|$. □

Theorem 3.4 (*Hahn-Banach theorem for seminormed spaces*) [25] *Let X be a linear space over $\mathbb{K}$ equipped with a seminorm p, M be a linear subspace of X, and $f : M \to \mathbb{K}$ be a linear functional such that $|f(v)| \leq p(v)$ for any $v \in M$. Then there exists a linear functional $\widetilde{f} : X \to \mathbb{K}$ such that $\widetilde{f} = f$ in M and $|\widetilde{f}(v)| \leq p(v)$ for any $v \in X$.*

Proof. If $\mathbb{K} = \mathbb{R}$ the statement directly follows from Theorem 3.2; let us then assume that $\mathbb{K} = \mathbb{C}$. Let us extend $Re(f)$ to a linear functional $g : X \to \mathbb{R}$ with $g \leq p$, as we did in Theorem 3.2. By the previous lemma the functional $\widetilde{f} : X \to \mathbb{C} : v \mapsto g(v) - ig(iv)$ is then linear and extends f.

For any fixed $v \in X$, we have $\widetilde{f}(v) = re^{i\theta}$ for some $r, \theta \geq 0$. Hence $\widetilde{f}(e^{-i\theta}v) = e^{-i\theta}\widetilde{f}(v) = r \geq 0$, and therefore $|\widetilde{f}(e^{-i\theta}v)| = \widetilde{f}(e^{-i\theta}v) = g(e^{-i\theta}v)$. As $g \leq p$ in X, by the positive homogeneity of p we then have

$$\begin{aligned} |\widetilde{f}(v)| &= |e^{-i\theta}\widetilde{f}(v)| = |\widetilde{f}(e^{-i\theta}v)| = g(e^{-i\theta}v) \\ &\leq p(e^{-i\theta}v) = |e^{-i\theta}|p(v) = p(v) \qquad \forall v \in X. \square \end{aligned} \tag{3.5}$$

3.4 Some consequences of the Hahn-Banach theorem.

Corollary 3.5 *Let M be a linear subspace of a normed space X. Any functional $f \in M'$ can then be extended to a functional $\widetilde{f} \in X'$ such that $\|\widetilde{f}\|_{X'} = \|f\|_{M'}$.*

(The extension $f \mapsto \widetilde{f}$ is said to be norm-preserving.)

Proof. As $|f(v)| \leq \|f\|_{M'}\|v\|_X$ for any $v \in M$, we may apply Theorem 3.4 with $p(v) = \|f\|_{M'}\|v\|_X$. There exists thus an extension $\widetilde{f} \in X'$ such that $|\widetilde{f}(v)| \leq \|f\|_{M'}\|v\|_X$ for all $v \in X$. Hence $\|\widetilde{f}\|_{X'} \leq \|f\|_{M'}$; as $\widetilde{f}$ extends f, this is an equality. □

Remark. The Hahn-Banach theorem concerns the extension of linear and continuous functionals. No analogous extension exists in general for linear and continuous operators between Banach spaces. For instance, the identity operator $c \to c$ cannot be extended to a linear and continuous operator $\ell^\infty \to c$.

[25] Due to Bohnenblust-Sobczyk-Soukhomlinoff.

Corollary 3.6 *Let $X \neq \{0\}$ be a real normed space. For any $u \in X$, then there exists $f \in X'$ such that $\|f\| = 1$ and $f(u) = \|u\|$.*

Thus

$$\|u\| = \max\{f(u) : f \in X', \|f\| \le 1\} \qquad \forall\, u \in X. \tag{3.6}$$

Proof. For $u = 0$ the thesis is trivial. For any fixed $u \in X \setminus \{0\}$, let us define the linear subspace $M = \mathbb{K}u$ and set $g(\lambda u) = \lambda\|u\|$ for any $\lambda \in \mathbb{K}$. As $g \in M'$ and $\|g\|_{M'} = 1$, by Corollary 3.5 there exists a norm-preserving extension $f \in X'$. □

By the foregoing corollary, *functional* analysis in normed spaces has plenty of linear and continuous functionals at its disposal. Because of the next result, the dual of a normed space separates points.

Corollary 3.7 *Let X be a normed space, M be a closed subspace of X, and $u \in X \setminus M$. Then there exists $f \in X'$ such that $f(u) = 1$ and $f(v) = 0$ for any $v \in M$.*

Proof. As M is a closed subspace, we may define the quotient space X/M. For any $u \in X\setminus M$, by Corollary 3.6 there exists $g \in (X/M)'$ such that $g([u]) = \|[u]\| \neq 0$. Then $f : X \to \mathbb{K} : v \mapsto g([v])/g([u])$ has the required properties. □

Corollary 3.8 *A linear subspace M of a normed space X is dense in X if (and only if) $f = 0$ is the only element of X' such that $f(v) = 0$ for any $v \in M$.*

Proof. By Corollary 3.7, if $\overline{M} \neq X$ then for any $u \in X \setminus \overline{M}$ there exists a functional $f \in X'$ such that $f \neq 0$ and $f(v) = 0$ for any $v \in M$. (The converse implication is obvious and is not related to the Hahn-Banach theorem.) □

Corollary 3.9 *Let M be a linear subspace of the dual X' of a normed space X. If M is dense in X', then $u = 0$ is the only element of X such that $f(u) = 0$ for any $f \in M$.*

Proof. Let $u \in X$ be such that $f(u) = 0$ for any $f \in M$. By the density of M, then $f(u) = 0$ for any $f \in X'$. By Corollary 3.6 then $\|u\| = 0$, i.e., $u = 0$. □

Proposition 3.10 *The canonical embedding $J : X \to X''$ (see (1.21)) is an isometric isomorphism between X and $J(X)$. Moreover, $J(X)$ is a closed subspace of X'' iff X is a Banach space.*

Proof. The first statement holds since

$$\|Ju\| = \sup_{\|f\|\le 1} |Ju(f)| \overset{(3.6)}{=} \max_{\|f\|\le 1} |f(u)| = \|u\| \qquad \forall u \in X. \tag{3.7}$$

Therefore $J(X)$ is complete iff so is X, and the final assertion follows. □

A Banach space $\hat{X}$ is called a **completion** of a normed space X, if there exists an isometry $\Phi : X \to \hat{X}$ such that $\Phi(X)$ is dense in $\hat{X}$. The bidual may be used to construct this completion.

Proposition 3.11 *Any normed space X has a completion $\hat{X}$. Any two completions of X are isometrically isomorphic.*

Proof. Let $J : X \to X''$ be the canonical embedding. Then the closure of $J(X)$ in X'' is a completion of X, because of Proposition 3.10. Now let $\hat{X}$ and $\widetilde{X}$ be two completions of X with corresponding isometries $\Phi : X \to \hat{X}$ and $\widetilde{\Phi} : X \to \widetilde{X}$. By Proposition 1.6, the mapping $L_0 = \widetilde{\Phi}\circ\Phi^{-1} : \Phi(X) \to \widetilde{\Phi}(X)$ has an extension $L : \hat{X} \to \widetilde{X}$, and this is an isometric isomorphism, too. □

4 Separation

4.1 Separation in linear spaces

Let us distinguish between linear separation and (linear and) topological separation. The former is a purely algebraic concept, and concerns convex subsets of a linear space X; the latter applies to normed spaces.

We say that $H \subset X$ is a **hyperplane** iff $H = f^{-1}(0)$, for some linear functional $f : X \to \mathbb{R}$ with $f \neq 0$; obviously H determines f up to a factor $\lambda \neq 0$. We define an **affine hyperplane** as the translate of a hyperplane: $H = f^{-1}(0) + v$ for any $f \neq 0$ and $v \in X$; or equivalently, setting $\alpha = f(v)$, $H = f^{-1}(\alpha)$ for any $\alpha \in \mathbb{K}$. A hyperplane is a proper linear subspace that is maximal, in the sense that X is the only linear subspace that strictly contains it. Hyperplanes are linear subspaces of codimension one.

Let A, B be two nonempty subsets of a real linear space X. We say that a nonzero linear functional $f : X \to \mathbb{R}$ **separates** A and B iff $f(u) \le f(v)$ for any $u \in A$ and any $v \in B$, that is,

$$\sup_{u \in A} f(u) \le \inf_{v \in B} f(v). \tag{4.1}$$

Any $\alpha \in [\sup_A f, \inf_B f]$ then determines an affine hyperplane $H = f^{-1}(\alpha)$ such that A and B are respectively contained in the half-spaces $f^{-1}(]-\infty, \alpha])$ and $f^{-1}([\alpha, +\infty[)$. Notice that A and/or B might intersect H, or even be contained in it (even $A = B$ is not excluded!).

We say that f **strongly separates** [26] A and B iff

$$\sup_{u \in A} f(u) < \inf_{v \in B} f(v). \tag{4.2}$$

If X is a normed space, one is mainly interested in the case where the separating functional f is continuous, i.e., $f \in X'$. Then the affine hyperplane $H = f^{-1}(\alpha)$ is closed. Moreover, $f \in X'$ separates two nonempty subsets A and B of X iff it separates their closures; the same holds for the strong separation. [Ex]

Gauges. For any subset M of a linear space X over $\mathbb{K}$, the functional

$$p_M : X \to [0, +\infty] : u \mapsto \inf\{\lambda > 0 : u \in \lambda M\} \tag{4.3}$$

is called the **Minkowski functional** (or the **gauge**) of M. For instance, if $M = B(0, R)$ then $p_M = \|\cdot\|/R$. Note that $p_\emptyset = \inf \emptyset = +\infty$, $p_X = 0$, and

$$M_1 \subset M_2 \quad \Longrightarrow \quad p_{M_1} \ge p_{M_2} \qquad \forall M_1, M_2 \subset X. \tag{4.4}$$

The set M is called **absorbing** iff $p_M(u) < +\infty$ for any $u \in X$ (or equivalently, $X = \bigcup_{\lambda > 0} \lambda M$), and **balanced** iff $\lambda M \subset M$ whenever $\lambda \in \mathbb{K}$ and $|\lambda| = 1$.

Lemma 4.1 *Let X be a linear space. If p is a seminorm, then the p-*unit ball *$M_p := \{u \in X : p(u) \le 1\}$ is absorbing, balanced and convex. Conversely, the gauge p_M of any absorbing, balanced and convex set $M \subset X$ is a seminorm. Moreover $p_{M_p} = p$ and $\{u : p_M(u) < 1\} \subset M \subset \{u : p_M(u) \le 1\}$. [Ex]*

Lemma 4.2 *An absorbing, balanced and convex subset M of a normed space X is a neighbourhood of the origin iff its gauge p_M is continuous on X. Moreover* $\text{int}(M) = \{u : p_M(u) < 1\}$ *in this case.*

[26] Here *strongly* does not refer to the strong topology!

Proof. If the seminorm p_M is continuous on X, then $\{u : p_M(u) < 1\}$ is an open subset of M, and thus $0 \in \{u : p_M(u) < 1\} \subset \text{int}(M)$. In order to prove the converse, let us assume that $0 \in \text{int}(M)$. By Lemma 4.1, then

$$0 \in \text{int}(\varepsilon M) \subset \{u : p_M(u) \leq \varepsilon\} \qquad \forall \varepsilon > 0 \,. \tag{4.5}$$

As $\{\varepsilon M : \varepsilon > 0\}$ is a basis of neighbourhoods of the origin, this implies that $p_M(u_n) \to 0$ whenever $u_n \to 0$. Thus p_M is continuous at the origin, hence on the whole X because of the inverse triangle inequality.

It remains to show that $\text{int}(M) \subset \{u : p_M(u) < 1\}$. Indeed, if $u \in X$ with $p_M(u) = 1$, then for any $\lambda > 1$ we have $p_M(\lambda u) = \lambda p_M(u) > 1$ and thus $\lambda u \notin M$ by Lemma 4.1. Thus, $u \notin \text{int}(M)$. □

4.2 Separation in normed spaces

Let us first deal with real normed spaces, and begin with separation of a point from a convex set.

Theorem 4.3 *Let X be a real normed space, A be a convex subset with nonempty interior, and $u \in X \setminus \text{int}(A)$. Then*

$$\exists f \in X' : \quad f(v) \leq f(u) \quad \forall v \in A, \quad f(v) < f(u) \quad \forall v \in \text{int}(A). \tag{4.6}$$

In particular f thus separates A and $\{u\}$.

* *Proof.* We may assume that $0 \in \text{int}(A)$ without loss of generality. Let us define $g : \mathbb{R}u \to \mathbb{R}$ by $g(tu) = t$, so that $|g| \leq p_A$ on $\mathbb{R}u$ by Lemma 4.1. Because of Theorem 3.4, g can be extended to a linear functional $f : X \to \mathbb{R}$ with $|f(v)| \leq p_A(v)$ for any $v \in X$. By Lemma 4.2, p_A is bounded; hence f is bounded, thus $f \in X'$. By Lemma 4.1 and Lemma 4.2, $p_A(v) < 1$ for any $v \in \text{int}(A)$ and $p_A(v) \leq 1$ for any $v \in A$. As $f(u) = 1$, the assertion follows. □

A counterexample. In the foregoing theorem the hypothesis that A has nonempty interior cannot be dropped. A counterexample is provided by any proper dense linear subspace Y of a real Banach space X; e.g., $Y = \ell^1$ and $X = c_0$. As Y is dense in X, for any nonzero $f \in X'$ there exists $v \in X$ such that $f(v) \neq 0$; by the linearity of f, then $f(Y) = \mathbb{R}$. No point of $X \setminus Y$ can thus be separated from Y.

Lemma 4.4 *Let A and B be two nonempty subsets of a normed space X. Then:*
(i) if A and B are convex, then $A + B$ is convex;
(ii) if A is open, then $A + B$ is open;
(iii) if A is compact and B is closed, then $A + B$ is closed.

(The same properties obviously hold for $A - B = A + (-B)$, too.)

Proof. Parts (i) and (ii) are straightforward. Let us prove part (iii). For any point $w \in \overline{A + B}$, there exist sequences $\{u_n\} \subset A$ and $\{v_n\} \subset B$ such that $u_n + v_n \to w$. As A is compact, there exists a convergent subsequence $\{u_{n'}\}$ whose limit u belongs to A. Hence $v_{n'} \to v := w - u$, and $v \in B$ since B is closed. Thus $w = u + v \in A + B$. □

The set $A + B$ need not be closed if A and B are just closed. For instance, let $A_\pm = \{(x, \pm 1/x) \in \mathbb{R} : x > 0\}$ and set $C = A_+ + A_-$; then $(0, 0) \in \bar{C} \setminus C$. Indeed, e.g., $C \ni (1/n, n) + (1/n, -n) = (2/n, 0) \to (0, 0)$.

• **Theorem 4.5** (*Separation – Eidelheit*) *Let A and B be two disjoint nonempty convex subsets of a real normed space X, with A open. Then A and B can be separated by a closed affine hyperplane.*

Proof. By Lemma 4.4 the set $A - B$ is convex and open. As A and B are disjoint, $0 \notin A - B$. By Theorem 4.3, the closed subspace $\{0\}$ can then be separated from $A - B$; that is, there exists $f \in X'$ such that $f(A - B) \leq f(0) = 0$. As $f(A) - f(B) = f(A - B)$, we conclude that $f(A) \leq f(B)$. (One may also show that $f(A) < f(B)$.) □

• **Theorem 4.6** (*Strong Separation – Tukey and Klee*) *Let X be a real normed space, and A and C be two disjoint nonempty convex subsets of X, with A compact and C closed. Then A and C can be strongly separated by a closed affine hyperplane.*

Proof. As A is compact, $\varepsilon = \text{dist}(A, C) := \inf \big\{ \|u - v\| : u \in A, v \in C \big\} > 0$. [Ex] Thus, denoting by B_ε the open ball with radius ε centered in 0, the sets C and $A_\varepsilon = A + B_\varepsilon$ are disjoint. By Lemma 4.4, A_ε is open and convex. By Theorem 4.5, then there exists $f \in X'$ such that $\sup f(A_\varepsilon) \leq \inf f(C)$. By selecting $u \in A$ with $f(u) = \max f(A)$ and $v \in B_\varepsilon$ with $f(v) > 0$, we have

$$\max f(A) < f(u) + f(v) = f(u + v) \leq \sup f(A_\varepsilon) \leq \inf f(C). \square$$

An example. By Theorem 4.6, any $\bar{u} \in c \setminus c_0$ can be strongly separated from c_0 by a closed affine hyperplane. In this case the hyperplane is actually a closed subspace of the form: $f^{-1}(0)$, for a suitable $f \in c'$.

If $\mathbb{K} = \mathbb{R}$, then we may select $f(v) = \lim_{j\to\infty} v_j$ for any $v = (v_1, ..., v_j, ...) \in c$; indeed $f(\bar{u}) \neq 0$ and $f(v) = 0$ for any $v \in c_0$.

If $\mathbb{K} = \mathbb{C}$, then we select the same functional if $Re\big(\lim_{j\to\infty} v_j\big) \neq 0$. Otherwise we select $f(u) = i \lim_{j\to\infty} v_j$, so that $Re(f(\bar{u})) \neq 0$, as $\bar{u} \notin c_0$.

Corollary 4.7 *Any closed convex subset A of a real normed space X is the intersection of the closed half-spaces that contain it.*

Proof. It suffices to notice that any point of the complementary set of A can be strongly separated from A, since singletons are convex and compact. □

Remarks. (i) In Theorem 4.6 the hypothesis of compactness cannot be replaced by closedness. For instance, in $\mathbb{R}^2$ the set $\{(x, y) : x > 0, xy \geq 1\}$ and the x-axis cannot be strongly separated.

(ii) Theorems 4.5 and 4.6 are based on somehow diverging hypotheses. The first one states that two disjoint nonempty convex subsets of X can be separated, provided that one of the two sets is open. The second result instead states that two disjoint nonempty convex subsets can be strongly separated, provided that one of them is compact and the other one is closed.

Separation in complex normed spaces. Let X be a complex linear space, and $X_\mathbb{R}$ be the corresponding real linear space. Any norm on X is also a norm on X_R. By Lemma 3.3, $f \in X'$ iff $Re(f) \in X'_R$, and in that case $\|f\| = \|Re(f)\|$. This allows us to extend the notion of separation to any complex linear space X:

$$\begin{aligned}&\text{a nonzero linear functional } f : X \to \mathbb{C} \text{ separates two}\\ &\text{nonempty subsets of } X \text{ iff } Re(f) \text{ separates them in } X_\mathbb{R}.\end{aligned} \tag{4.7}$$

We similarly extend strong separation. The previous separation results are then readily extended to complex normed spaces. We leave this reformulation to the reader.

5 The Baire Theorem and its Consequences

5.1 The Baire theorem

The results of this section stem from the following classical result, which concerns the topology of complete metric spaces.

• **Theorem 5.1** (*Baire Theorem) Let X be a complete metric space. If X is a countable union of closed subsets, then at least one of them has nonempty interior. [Dually and equivalently: the intersection of any countable family of open dense subsets of X is dense.]*

Proof. Let $\{X_n\}$ be any sequence of (possibly nondisjoint) closed subsets of X with empty interior. It suffices to show that $X \neq \bigcup_{n\in\mathbb{N}} X_n$.

As X_0 is closed and has no interior point, it cannot coincide with X. Thus $X \setminus X_0$ is open and contains a closed ball, say $B(x_0, \varepsilon_0)$. As X_1 is closed and has no interior point, the set $(X \setminus X_1) \cap \mathrm{int}(B(x_0, \varepsilon_0))$ is open and contains a closed ball $B(x_1, \varepsilon_1)$ with $\varepsilon_1 \leq \varepsilon_0/2$. By iterating this procedure, we construct a nested sequence $\{B(x_n, \varepsilon_n)\}$ of closed balls; each of them does not intersect X_n, and $\varepsilon_n \to 0$. By the completeness of X, the sequence $\{x_n\}$ then converges to some $x \in X$. By construction $B(x_m, \varepsilon_m) \cap \big(\bigcup_{n\leq m} X_n\big) = \emptyset$ for any m. As $x \in \bigcap_{n\in\mathbb{N}} B(x_n, \varepsilon_n)$, we infer that $x \notin \bigcup_{n\in\mathbb{N}} X_n$; thus $X \neq \bigcup_{n\in\mathbb{N}} X_n$. □

This theorem has several consequences, including the next result.

Corollary 5.2 *No Banach space has countable infinite dimension.*

Proof. By contradiction, let $\{u_n\}$ be a countable Hamel basis, and let X_m be the linear span of $\{u_1, ..., u_m\}$ for any $m \in \mathbb{N}$. These finite-dimensional subspaces are closed and have empty interior. As the whole space is the union of these sets, by the Baire theorem it cannot be complete. □

5.2 The principle of uniform boundedness

This result provides the uniform boundedness of pointwise bounded families $\mathcal{F}$ of bounded operators.

• **Theorem 5.3** (*Banach-Steinhaus) Let X_1 be a Banach space, X_2 a normed space, and $\mathcal{F} \subset \mathcal{L}(X_1; X_2)$. Then*

$$\sup_{L\in\mathcal{F}} \|Lu\|_2 < +\infty \quad \forall u \in X_1 \quad \Rightarrow \quad \sup_{L\in\mathcal{F}} \|L\|_{\mathcal{L}(X_1;X_2)} < +\infty. \tag{5.1}$$

In other terms, denoting by B_{X_1} the open unit ball in X_1,

$$\begin{aligned} &\forall u \in B_{X_1}, \exists C > 0 : \forall L \in \mathcal{F}, \|Lu\|_2 \leq C \\ \Rightarrow \quad &\exists \hat{C} > 0 : \forall u \in B_{X_1}, \forall L \in \mathcal{F}, \|Lu\|_2 \leq \hat{C}. \end{aligned} \tag{5.2}$$

Proof. Let us set

$$A_n = \{u \in X_1 : \forall L \in \mathcal{F}, \|Lu\|_2 \leq n\} = \textstyle\bigcap_{L\in\mathcal{F}} \{u \in X_1 : \|Lu\|_2 \leq n\}$$

for any n. Because of the continuity of the operators of $\mathcal{F}$, this set is closed. By the assumption of pointwise boundedness, each $u \in X_1$ belongs to some A_n; thus, $\bigcup_n A_n = X_1$. By the Baire theorem, for some $\widetilde{n} \in \mathbb{N}$ the interior of $A_{\widetilde{n}}$ is then nonempty. So let $w \in A_{\widetilde{n}}$ and $r > 0$ be such that $w + rB(0,1) \subset A_{\widetilde{n}}$. Therefore

$$r\|L(v)\|_2 \leq \widetilde{n} + \|L(w)\|_2 \leq 2\widetilde{n} \qquad \forall L \in \mathcal{F}, \forall v \in B(0,1), \tag{5.3}$$

whence $\sup_{L\in\mathcal{F}} \|L\|_{\mathcal{L}(X_1;X_2)} \leq 2\widetilde{n}/r$. □

Remark. The Banach-Steinhaus theorem fails if X_1 is not complete. A counterexample is provided by the family of functionals $\{f_n\}_{n\in\mathbb{N}}$, with $f_n(x) = nx_n$ for any $x := (x_1, x_2, ...) \in c_{00}$ and any n.

The Banach-Steinhaus theorem concerns linear and continuous operators, in particular functionals.

Corollary 5.4 *Let X be a Banach space. Any set $C \subset X'$ is bounded if (and only if) for any $x \in X$, $\{f(x) : f \in C\}$ is a bounded subset of $\mathbb{K}$.*
Let X be a normed space. Any set $B \subset X$ is bounded if (and only if) for any $f \in X'$, $\{f(x) : x \in B\}$ is a bounded subset of $\mathbb{K}$.

Another consequence is that the pointwise limit of a sequence of bounded operators is again a bounded operator.

Corollary 5.5 *Let X_1 be a Banach space, X_2 be a normed space, and $\{L_n\}$ be a sequence in $\mathcal{L}(X_1; X_2)$. Assume that for any $u \in X_1$ the sequence $\{L_n u\}$ converges in X_2, and denote this limit by Lu. This defines an operator $L \in \mathcal{L}(X_1; X_2)$ satisfying*

$$\|L\| \leq \liminf_{n\to\infty} \|L_n\|. \tag{5.4}$$

Proof. The linearity of L is straightforward. Moreover, since $\{L_n u\}$ is bounded for any $u \in X_1$, by the Banach-Steinhaus theorem $C = \sup_n \|L_n\| < \infty$. For any $u \in X_1$, then

$$\|Lu\| = \big\| \lim_{n\to\infty} L_n u\big\| = \lim_{n\to\infty} \|L_n u\| \leq \limsup_{n\to\infty} \|L_n\|\|u\| \leq C\|u\|. \tag{5.5}$$

This proves the continuity of L as well as the inequality (5.4). □

5.3 The open mapping theorem

A linear mapping between two normed spaces is called **open** iff it maps open sets to open sets.

Lemma 5.6 *Let X_1 and X_2 be normed spaces. A mapping $L \in \mathcal{L}(X_1; X_2)$ is open if (and only if) 0 is an interior point of $L(\overset{\circ}{B}_{X_1})$.* [27]

Proof. Let U be open in X_1, $v \in L(U)$, and $u \in U \cap L^{-1}(v)$. For any $\varepsilon > 0$ such that $u + \varepsilon\overset{\circ}{B}_{X_1} \subset U$, we have $v + \varepsilon L(\overset{\circ}{B}_{X_1}) = L(u + \varepsilon\overset{\circ}{B}_{X_1}) \subset L(U)$. We conclude that $L(U)$ is open. □

As any neighborhood of the origin is absorbing, this lemma entails that any open mapping is surjective. The open mapping theorem establishes the converse, whenever X_1 and X_2 are complete.

[27] $\overset{\circ}{B}_{X_1}$ is the open unit ball in X_1.

• **Theorem 5.7** (*Banach's Open Mapping Theorem*) *Let X_1 and X_2 be Banach spaces. Any surjective mapping $L \in \mathcal{L}(X_1; X_2)$ is open.*

Proof. Set $V_m = L(mB_{X_1})$. We have $X_2 = \bigcup_{m\geq 1} V_m$ as L is surjective. By Baire's theorem, some $\overline{V_k}$ has nonempty interior, so it contains an open ball $v + B_\delta = v + \delta B_{X_1}$. We infer that $-v + B_\delta \subset \overline{V_k}$ by symmetry, and $B_\delta \subset \overline{V_k}$ by convexity; consequently $B_\varepsilon \subset \overline{V_1}$ for $\varepsilon = \delta/k$. Because of Lemma 5.6, it now suffices to prove that $\overline{V_1} \subset V_2$, since 0 will then also be an interior point of $V_1 = L(B_{X_1})$. To this purpose, let $v \in \overline{V_1}$. Set $v_0 = v$ and choose $u_1 \in B_{X_1}$ with $v_1 := v_0 - Lu_1 \in B_{\varepsilon/2} \subset \overline{L(B_{1/2})}$. Proceeding inductively, choose $u_n \in B_{2^{1-n}}$ with $v_n := v_{n-1} - Lu_n \in B_{2^{-n}\varepsilon} \subset \overline{L(B_{2^{-n}})}$. Therefore $\sum_{n=0}^\infty \|u_i\| < +\infty$; hence $\sum_{n=0}^\infty u_i$ converges in X_2. We thus get

$$u := \sum_{n=1}^\infty u_n \in B_2, \quad v - \sum_{j=1}^n Lu_j = v_n \to 0;$$

thus $v = Lu \in L(B_2) = V_2$. □

Remark. One might wonder whether all surjective mappings $L \in \mathcal{L}(X_1; X_2)$ map closed sets to closed sets. This fails even in $\mathbb{R}^2$. For instance the linear, continuous and surjective functional $L : \mathbb{R}^2 \to \mathbb{R} : (x, y) \mapsto x$ maps the closed set $A := \{(x, y) : xy \geq 1\}$ to the nonclosed set $L(A) = \mathbb{R} \setminus \{0\}$. (On the other hand, continuous mappings map compact sets to compact sets, also without any linearity assumption.)

Some consequences of the Open Mapping Theorem.

Corollary 5.8 (*Inverse Mapping Theorem*) *Let X_1 and X_2 be Banach spaces. The inverse of any linear, continuous and bijective mapping from X_1 to X_2 is linear and continuous.*

A linear mapping $L : X_1 \to X_2$ between normed spaces X_1 and X_2 is called **bounded below** iff there exists $c > 0$ such that

$$c\|u\| \leq \|Lu\| \qquad \forall u \in X_1. \tag{5.6}$$

Theorem 5.9 (*Closed Range Theorem*) *Let X_1 and X_2 be Banach spaces. For any $L \in \mathcal{L}(X_1; X_2)$ the following statements are equivalent:*
(i) L is bounded below,
(ii) L is injective and $\mathcal{R}(L)$ is closed,
(iii) $L : X_1 \to \mathcal{R}(L)$ is an isomorphism.

Proof. First we prove that "(i)⇒(ii)". Let us assume that L is bounded below, so that $Lu = 0$ implies $u = 0$. If $\{Lu_n\}$ is a Cauchy sequence in $\mathcal{R}(L)$, then by (5.6) $\{u_n\}$ is a Cauchy sequence in X_1. Hence $u_n \to u$ for some $u \in X_1$, and $Lu_n \to Lu$ by continuity. Thus $\mathcal{R}(L)$ is complete, hence closed.

Next we prove that "(ii)⇒(iii)". By the linearity of L, $\mathcal{R}(L)$ is a linear space; if it is also closed in X_2, then it is a Banach space. By applying the inverse mapping theorem (Corollary 5.8) with $\mathcal{R}(L)$ in place of X_2, we thus get (iii).

The implication "(iii)⇒(i)" is obvious. □

Corollary 5.10 *Let X_1 and X_2 be Banach spaces. For any $L \in \mathcal{L}(X_1; X_2)$ the following statements are equivalent:*
(i) L is bounded below and $\mathcal{R}(L)$ is dense,
(ii) L is an isomorphism.

Let us recall that, whenever $\|\cdot\|_1$ and $\|\cdot\|_2$ are two norms over the same linear space X, we say that the former norm is weaker than the latter one (and that this is stronger than the other one) iff there exists a constant $C > 0$ such that $\|u\|_1 \leq C\|u\|_2$ for all $u \in X$. The topology induced by $\|\cdot\|_2$ is then finer than that induced by $\|\cdot\|_1$.

Corollary 5.11 *Let a linear space X be a Banach space when equipped with either of two norms $\|\cdot\|_1$ and $\|\cdot\|_2$. If one of these norms is stronger than the other one, then they are equivalent.*

Proof. Denoting X_1 and X_2 the two normed spaces, it suffices to apply Corollary 5.8 to the identity mapping. □

Remark. The foregoing assertion fails in noncomplete normed spaces. For instance, the linear space c_{00} equipped with any of the norms of ℓ^p $(1 \leq p \leq \infty)$ is a noncomplete normed space. The norm of ℓ^p is finer than that of ℓ^q whenever $p < q$, but these norms are not equivalent.

• **Theorem 5.12** (*Closed Graph Theorem*) *Let X_1 and X_2 be Banach spaces. A linear mapping $L : X_1 \to X_2$ is continuous iff its graph $G_L := \{(v, Lv) : v \in X_1\}$ is closed in $X_1 \times X_2$.*

Proof. The proof of the "only if"-part is straightforward. Let us then prove the "if"-part. If the linear subspace G_L of $X_1 \times X_2$ is closed, then it is a Banach space by itself. The projections

$$p_i : G_L \to X_i, \qquad p_1(v, Lv) = v, \qquad p_2(v, Lv) = Lv$$

are continuous. As p_1 is bijective, p_1^{-1} also is continuous by the inverse mapping theorem (Corollary 5.8). Therefore $L = p_2 \circ p_1^{-1}$ is continuous. □

Remarks. (i) Let $L : X_1 \to X_2$ be a linear mapping between two Banach spaces X_1 and X_2, and let us consider the following statements:

$$(a)\ u_n \to u \ \text{ in } X_1, \qquad (b)\ Lu_n \to w \ \text{ in } X_2, \qquad (c)\ w = Lu.$$

The mapping L is closed iff "(a) and (b) together imply (c)", whereas it is continuous off "(a) implies (b) and (c)" (for any sequence $\{u_n\}$ and any u in X_1). The fact that the former property entails the latter is a remarkable consequence of the open mapping theorem.

(ii) Let X_1 and X_2 be Banach spaces, and $L : X_1 \to X_2$ be a linear mapping. The graph of L may be closed (that is, L may be continuous by the latter theorem) even if its range is not closed. Whenever $X_1 \subset X_2$ with proper and continuous injection, the identity mapping $X_1 \to X_2$ is a counterexample. See however Theorem 5.9.

6 Weak Topologies

In this section we introduce the weak and the weak star topology.

In any normed space X the topology generated by the norm is called the norm topology or the **strong topology.** The so-called **weak topology** is defined as follows:

$$\begin{array}{l}\text{the weak topology on } X \text{ is the coarsest topology on } X \\ \text{among those that make all functionals in } X' \text{ continuous.}\end{array} \tag{6.1}$$

Any weakly closed (weakly open, resp.) set is closed (open, resp.). [28] Let us denote by $B(0, \varepsilon)$ the ball of $\mathbb{K}$ with center 0 and radius $\varepsilon > 0$. The family

$$\mathcal{S} = \{f^{-1}(B(0,\varepsilon)) : f \in X', \varepsilon > 0\} \tag{6.2}$$

is a subbasis of the system of the neigborhoods of the origin in the weak topology. Any weakly open subset of X is thus the union of a family of elements, and each of these elements is the intersection of a finite subfamily of $\bigcup_{u \in X}(u + \mathcal{S})$.

The same construction applies to the dual space X':

$$\begin{array}{l}\text{the weak topology on } X' \text{ is the coarsest topology on } X' \\ \text{among those that make all functionals in } X'' \text{ continuous.}\end{array} \tag{6.3}$$

In X' the **weak star topology** is also defined. [29] This topology on X' is generated by the family $\{(\mathbb{K}, Ju) : u \in X\}$, where by J we denote the canonical imbedding $X \to X''$. Thus

$$\begin{array}{l}\text{the weak star topology on } X' \text{ is the coarsest topology on } X' \\ \text{among those that make all functionals of } J(X) \text{ continuous.}\end{array} \tag{6.4}$$

The family

$$\mathcal{S} = \{(Ju)^{-1}(B(0,\varepsilon)) : u \in X, \varepsilon > 0\} \tag{6.5}$$

is a subbasis of the system of neigborhoods of the origin in this topology.

If X is reflexive, then on X' the weak star topology coincides with the weak topology. Otherwise the former is strictly coarser than the latter, since any element of $X'' \setminus X$ is weakly continuous on X' but not weakly star continuous.

6.1 Weak and weak star convergence

The weak and weak star topologies induce the following notions of convergence for sequences (and for nets...).

Proposition 6.1 *Let X be a normed space. For any sequence $\{u_n\}$ and u in X, and for any sequence $\{f_n\}$ and f in X', the following holds*

$$u_n \to u \quad \text{weakly in } X \quad \Leftrightarrow \quad f(u_n) \to f(u) \quad \forall f \in X', \tag{6.6}$$

$$f_n \to f \quad \text{weakly in } X' \quad \Leftrightarrow \quad F(f_n) \to F(f) \quad \forall F \in X'', \tag{6.7}$$

$$f_n \to f \quad \text{weakly star in } X' \quad \Leftrightarrow \quad f_n(u) \to f(u) \quad \forall u \in X. \tag{6.8}$$

One says that two disjoint open subsets of a topological space T **separate** two distinct points $u, v \in T$ iff they respectively include u and v. The space T is called a **Hausdorff space**, and its topology is said to be *Hausdorff,* iff any pair of distinct points of T can be separated. It is easy to see that this holds iff the limit of any convergent sequence in T is unique. The next result thus entails the uniqueness of both weak and weak star limits.

Proposition 6.2 *The weak topology of a normed space X and the weak star topology of X' are Hausdorff. The weak and weak star limits are then unique.*

[28] The norm topology is our default topology unless otherwise specified. So "open" stands for "strongly open", "closed" for "strongly closed", and so on.

[29] This is meaningful only in a dual space. The *star* refers to the fact that in the literature the dual space is often denoted by X^*, with a star instead of a prime.

Proof. Let us first consider the weak topology. If $u, v \in X$ with $u \neq v$, then $f(u) \neq f(v)$ for some $f \in X'$ by Theorem 4.6. For any open subsets I and J of $\mathbb{K}$ that separate $f(u)$ and $f(v)$, thus $f^{-1}(I)$ and $f^{-1}(J)$ separate u and v.

Let us next come to the weak star topology. Without using any consequence of the Hahn-Banach theorem, by definition of function, if $f, g \in X'$ with $f \neq g$, then $f(u) \neq g(u)$ for some $u \in X$. The argument then proceeds as above. □

Proposition 6.3 *(i) Let X be a normed space. Then any weakly convergent sequence $\{u_n\}$ in X is bounded. Moreover, denoting its weak limit by u,*

$$\|u\| \leq \liminf_{n\to\infty} \|u_n\|. \tag{6.9}$$

(ii) Let X be a normed space. Then any weakly convergent sequence $\{f_n\}$ in X' is bounded. Moreover, denoting its weak limit by f,

$$\|f\| \leq \liminf_{n\to\infty} \|f_n\|. \tag{6.10}$$

(iii) Let X be a Banach space. Then any weakly star convergent sequence $\{f_n\}$ in the dual space X' is bounded. Moreover, its weak star limit f fulfills (6.10).

Proof. Let us first prove part (iii). Let $f_n \to f$ weakly star in X'. By Proposition 6.1, $\{f_n(u)\}$ is convergent and hence bounded for any $u \in X$. By of the principle of uniform boundedness (Theorem 5.3), the set $\{\|f_n\|\}$ is then bounded. Inequality (6.10) follows from the estimate

$$|f(u)| = \lim_{n\to\infty} |f_n(u)| \leq \liminf_{n\to\infty} \|f_n\|\|u\| \qquad \forall u \in X. \tag{6.11}$$

In order to prove (i), let us consider the canonical imbedding $J : X \to X''$ and note that $u_n \to u$ weakly in X iff $Ju_n \to Ju$ weakly star in X''. As J is an isometry, part (i) follows from part (iii).

Part (ii) is just a reformulation of part (i) in X'. □

Remark. Part (iii) of the preceding theorem fails if X is not complete. Here is a counterexample for $X = c_{00}$. For any $n \in \mathbb{N}$, let us define $f_n \in {c_{00}}'$ by setting $f_n(u) := nu_n$ for any $u = (u_1, u_2, ...) \in c_{00}$. The sequence $\{f_n\}$ weakly star vanishes in ${c_{00}}'$ as $n \to \infty$, but is unbounded.

However, this sequence $\{f_n\}$ does not weakly star vanish in ${c_0}'$, although c_0 is the completion of c_{00}.

Next we show that, whenever existing, the pointwise weak limit of a sequence of bounded operators defines a bounded operator. In particular, this applies to a pointwise weakly convergent sequence in X'.

Proposition 6.4 *Let X_1 be a normed space, X_2 be a Banach space, and $\{L_n\}$ be a sequence in $\mathcal{L}(X_1; X_2)$. If the sequence $\{L_n u\}$ weakly converges in X_2 for any $u \in X_1$, then, denoting by Lu this limit, a mapping $L : X_1 \to X_2$ is defined such that*

$$L \in \mathcal{L}(X_1; X_2)\,, \qquad \|L\| \leq \liminf_{n\to\infty} \|L_n\|. \tag{6.12}$$

Proof. The linearity of L is straightforward, and we just show the continuity. By Proposition 6.3(i), the sequence $\{L_n u\}$ is bounded in the Banach space X_2 for any $v \in X_1$. By the principle of uniform boundedness (Theorem 5.3), then $\sup_n \|L_n\| < \infty$. By (6.9), then

$$\|Lu\| \leq \liminf_{n\to\infty} \|L_n u\| \leq \liminf_{n\to\infty} \|L_n\|\|u\| \qquad \forall v \in X_1.$$

(6.12) is thus established. □

6.2 * Two results concerning weak convergence

*** Weak closedness and weak sequential closedness.** In Proposition 7.4 we shall see that the weak topology is not metrizable (i.e., it is not induced by a metric), whenever X has infinite dimension. The closure of a subset S of X may thus include elements that cannot be represented as weak limit of any sequence in S. Therefore weak continuity and sequential weak continuity need not coincide.

For any $n \in \mathbb{N}$ let us denote by e_n the sequence $(0, ..., 0, 1, 0, ...)$, with 1 at the nth place and 0 elsewhere, and define the set

$$A := \{\sqrt{n} e_n : n \in \mathbb{N}\} \subset X = \ell^2.$$

By Proposition 6.3, no sequence of distinct elements of A weakly vanishes in X, since any such sequence is unbounded. On the other hand, we claim that the null element lies in the weak closure of A or, equivalently, that any weak neighborhood of 0 intersects A. In order to prove this, according to (6.2) it suffices to show that A intersects any set of the form

$$U = \bigcap_{k=1}^{m} \{u \in X : |f_k(u)| < \varepsilon\},$$

where $\varepsilon > 0$, $m \in \mathbb{N}$ and $f_1, \ldots, f_m \in X^*$. Let such a set U be given, let

$$f_k(u) = \sum_{j=1}^{\infty} f_k^j u_j, \quad \sum_{j=1}^{\infty} |f_k^j|^2 < \infty, \quad 1 \le k \le m.$$

Let us define $g^j := \sum_{k=1}^{m} |f_k^j|$ for $j \in \mathbb{N}$. Hence $g = (g^j) \in \ell^2$, and thus there exists an $n \in \mathbb{N}$ such that $\sqrt{n} g^n < \varepsilon$ (otherwise $\sum |g^j|^2 \ge \varepsilon^2 \sum_j (1/j) = \infty$). For this n we have

$$|f_k(\sqrt{n} e_n)| = \sqrt{n} |f_k^n| \le \sqrt{n} g^n < \varepsilon, \qquad 1 \le k \le m.$$

Thus, $\sqrt{n} e_n \in U$ and the claim is proved.

Another example is due to von Neumann. For any $n \in \mathbb{N}$ let us denote by e_n the sequence $(0, ..., 0, 1, 0, ...)$ (with 1 at the nth place and 0 elsewhere), and define the unbounded set

$$A := \{e_m + m e_n : m, n \in \mathbb{N}, 1 \le m < n\}.$$

No sequence of distinct elements of A weakly vanishes in ℓ^2. [Ex] Nevertheless one may show that the null element is in the weak closure of A, or equivalently that any weak neighborhood of 0 intersects A. []

*** The Schur property.** Next we outline a surprising phenomenon. We saw that in any infinite-dimensional normed space the weak topology is strictly weaker than the strong one. Nevertheless the following occurs.

Theorem 6.5 (*Schur*) *In ℓ^1 any weakly convergent sequence is also strongly convergent.*

Proof. Let $\{u_n\} \subset \ell^1$ be such that $u_n \to 0$ weakly in ℓ^1. Hence for any m, denoting by u_n^m the mth component of u_n, $u_n^m \to 0$ as $n \to \infty$. It is then clear that $u_n \to 0$ in measure in any finite subset of $\mathbb{N}$, that is, in any subset of finite measure, if we equip $\mathbb{N}$ with the counting measure. By the next (nontrivial) lemma of measure theory, it then follows that $u_n \to 0$ in ℓ^1. □

Lemma 6.6 *Let* $(A, \mathcal{A}, \mu)$ *be a measure space,* $\{u_n\}$ *be a sequence in* $L^1(A)$*, and* $u_n \to u$ *weakly in* $L^1(A)$*. Then* $u_n \to u$ *strongly in* $L^1(A)$ *iff* $u_n \to u$ *in measure in every set* $\widetilde{A} \subset A$ *of finite measure. []*

Any Banach space that has the property of part (i) is said to have the *Schur property.* So ℓ^1 has it, but $L^1(0,1)$ does not. (In passing notice that $L^1(0,1)$ is not topologically isomorphic to ℓ^1: e.g., ℓ^1 has a predual, whereas $L^1(0,1)$ has not.)

Ahead we shall see that in infinite-dimensional Banach spaces, the weak topology is not determined by the weakly convergent sequences. Therefore the Schur property does not entail that the weak and strong topology coincide. This fact actually only holds in finite-dimensional normed space, as we shall see ahead.

6.3 Two convexity results of Mazur

If X is a linear space, then a function $f : X \to \mathbb{R} \cup \{+\infty\}$ is said to be **quasiconvex** iff for any $a \in \mathbb{R}$ the sublevel set $\{v \in X : f(v) \leq a\}$ is convex. On the other hand, $f : X \to \mathbb{R} \cup \{+\infty\}$ is said to be **convex** iff the *epigraph* $\{(v, a) \in X \times \mathbb{R} : f(v) \leq a\}$ is convex. Any convex function is quasiconvex, but not conversely; e.g., the real function $x \mapsto \sqrt{|x|}$ is quasiconvex but not convex.

If X is a topological space, then a function $f : X \to \mathbb{R} \cup \{+\infty\}$ is said to be **lower semicontinuous** (**sequentially lower semicontinuous,** resp.) iff for any $a \in \mathbb{R}$ the sublevel set $\{v \in X : f(v) \leq a\}$ is closed (sequentially closed, resp.), or equivalently iff the epigraph of f is closed (sequentially closed, resp.) in the product topology of $X \times \mathbb{R}$.

• **Theorem 6.7** (*Mazur) Let* X *be a normed space. Then:*
(i) A convex subset of X *is closed iff it is weakly closed.*
(ii) A quasiconvex function $f : X \to \mathbb{R} \cup \{+\infty\}$ *is lower semicontinuous iff it is weakly lower semicontinuous.*
(iii) A linear functional $X \to \mathbb{C}$ *is continuous iff it is weakly continuous.*

Proof. Any closed convex subset is weakly closed, by Corollary 4.7 and because any closed half-space is weakly closed. The converse is trivial. Part (i) is thus established.

Part (ii) follows from part (i), since a function $f : X \to \mathbb{R} \cup \{+\infty\}$ is lower semicontinuous (weakly lower semicontinuous, resp.) iff its epigraph is closed (weakly closed, resp.) in the product topology of $X \times \mathbb{R}$.

For any real functional f, part (iii) follows from part (ii), as f and $-f$ are both convex. By applying this property to the real and imaginary parts, this also holds for any complex functional. □

The next result is somehow surprising.

Corollary 6.8 ** [Mazur] Let* X *be a normed space, and* $u_n \to u$ *weakly in* X*. Then* u *is the strong limit of a sequence* $\{\widetilde{u}_m\}$ *of suitable (finite) convex combinations of elements of the sequence* $\{u_n\}$*. That is, for any* $m \in \mathbb{N}$ *there exist* $\ell_m \in \mathbb{N}$ *and* $\lambda_{m,k} \geq 0$ *for* $1 \leq k \leq \ell_m$*, with* $\sum_{k=1}^{\ell_m} \lambda_{m,k} = 1$*, such that*

$$\widetilde{u}_m := \sum_{k=1}^{\ell_m} \lambda_{m,k} u_k \;\to\; u \qquad \text{in } X, \text{ as } m \to \infty. \tag{6.13}$$

Proof. Let us set $S := \{u_n\}$, and notice that $\overline{\text{co}}(S)$ (the closure of the convex hull of S) is convex, as the closure of any convex set is convex. [Ex] By Theorem 6.7, any weak limit of a sequence in S belongs to $\overline{\text{co}}(S)$. Thus $u \in \overline{\text{co}}(S)$, that is, u is the strong limit of a sequence of elements of $\text{co}(S)$. □

7 Dimension

In this section we characterize finite-dimensional normed spaces in several ways. We begin with a statement which does not involve weak topologies.

Theorem 7.1 *(i) On $\mathbb{K}^N$ all norms are equivalent.*

(ii) For any normed spaces X_1 and X_2, if X_1 has finite dimension then any linear mapping $L : X_1 \to X_2$ is continuous.

(iii) A normed space X over $\mathbb{K}$ has finite dimension iff it is topologically isomorphic to $\mathbb{K}^N$ for some $N \geq 0$.

(iv) Any two normed spaces X_1 and X_2 of finite dimension are topologically isomorphic iff they have the same dimension.

(v) Let X be a normed space over the field $\mathbb{K}$ has finite dimension if all linear functionals $X \to \mathbb{K}$ are continuous.

Proof. (i) Let us denote by $\{e_i : i = 1, ..., N\}$ the family of unit vectors, and set $u_i = u \cdot e_i$ for any $u \in \mathbb{K}^N$ and any i. It suffices to show that any norm $\| \cdot \|$ on $\mathbb{K}^N$ is equivalent to the maximum norm $\|u\|_\infty = \max_i |u_i|$.

By setting $C = \sum_{i=1}^N \|e_i\|$, we have

$$\|u\| \leq \textstyle\sum_{i=1}^N |u_i| \|e_i\| \leq C\|u\|_\infty \qquad \forall u \in \mathbb{K}^N. \tag{7.1}$$

On the other hand, $f(u) = \|u\|$ defines a real-valued function which is continuous on $(\mathbb{K}^N, \| \cdot \|_\infty)$, since

$$|f(u) - f(v)| = \big|\|u\| - \|v\|\big| \leq \|u - v\| \overset{(7.1)}{\leq} C\|u - v\|_\infty \quad \forall u, v \in \mathbb{K}^N.$$

Hence f attains its minimum c on $S = \{\|u\|_\infty = 1\}$, which is compact for the topology induced by the norm $\| \cdot \|_\infty$. Moreover $c > 0$ since $f > 0$ on S. It follows that

$$c \leq \| \|u\|_\infty^{-1} u\| \quad \text{i.e.,} \quad c\|u\|_\infty \leq \|u\| \qquad \forall u \in \mathbb{K}^N.$$

The norms $\| \cdot \|$ on $\mathbb{K}^N$ and $\| \cdot \|_\infty$ are thus equivalent.

(ii) This is left to the Reader.

(iii) If $X \simeq \mathbb{K}^N$, then any topological isomorphism $L : \mathbb{K}^N \to X$ is also an algebraic isomorphism. Hence $\dim(X) = \dim(\mathbb{K}^N) = N$.

For the converse, let $\dim(X) = N$. By part (ii), any algebraic isomorphism between X and $\mathbb{K}^N$ is also a topological isomorphism.

(iv) If X_1 and X_2 have the same dimension, then they are algebraically isomorphic to $\mathbb{K}^N$. By part (i) then they are mutually topologically isomorphic.

Converse, if they are mutually topologically isomorphic, then by part (iii) they are topologically isomorphic to the same $\mathbb{K}^N$; hence they have the same dimension.

(v) By contradiction, let us assume that X is infinite-dimensional. Let $\{u_n\}$ be a sequence of linearly independent elements of X of unit norm, and set $f(u_n) = n$ for any n. Hence $n^{-1}u_n \to 0$, but $f(n^{-1}u_n) = 1$ for all n. By a simple procedure based on the Hausdorff maximality principle (2.1), f may be extended to a linear functional $\overline{f} : X \to \mathbb{K}$, which is then unbounded, hence discontinuous. □

The next two results cast some light upon the weak topology in infinite-dimensional normed spaces.

Proposition 7.2 *A normed space X has infinite dimension iff all nonempty weakly open subsets of X are unbounded. In this case, any weakly open set that includes 0 contains a closed subspace of infinite dimension.*

If X has a predual then the same holds for the weak star topology.

Proof. If $X = \mathbb{K}^N$, then the set $A = \{u : \ |u_i| < \varepsilon \text{ for } i = 1, ..., N\}$ is bounded, nonempty and weakly open. This proves the "if"-part.

Next let us assume that X has infinite dimension. Let A be a nonempty weakly open subset of X; without loss of generality we may assume that A contains the origin. By construction of the weak topology from the subbasis (6.2), there exists a finite set $\{f_1, \ldots, f_m\} \subset X'$ such that $M := \bigcap_i f_i^{-1}(0) \subset A$. As M is the kernel of the operator $(f_1, \ldots, f_m) \in \mathcal{L}(X; \mathbb{K}^m)$, it is an infinite-dimensional closed subspace of X.

This also proves the second statement of the thesis.

For the weak star topology the proof is similar. □

Corollary 7.3 *A normed space X has finite dimension iff the strong and the weak topology coincide on X.*

Remarks. The following holds in any infinite-dimensional normed space X.

(i) As weakly open sets are unbounded, the open unit ball B^0_X is not weakly open. It even has no interior point for the weak topology.

(ii) By Mazur's Theorem 6.7, the closed unit ball is weakly closed. This set coincides with its weak boundary, as it has no interior point in the weak topology.

(iii) The sphere, $\{v \in X : \|v\| = 1\}$, is closed but not weakly closed. For example, if $X = \ell^2$ the canonical sequence of unit vectors $\{e_n\}$ weakly converges to the origin.

If X has a predual, then analogous properties hold for the weak star topology.

Here is a further property along the same line.

Proposition 7.4 *In any infinite-dimensional normed space X, the unit sphere $\{v \in X : \|v\| = 1\}$ is weakly dense in the unit ball B_X.*

Analogously, the unit sphere $\{f \in X' : \|f\| = 1\}$ is weakly star dense in the unit ball B'_X.

Proof. Let $u \in B_X$, and A be a weak neighborhood of u. By Proposition 7.2, the weakly open set $A - u$ $(:= \{v - u : v \in A\})$ contains a straight line $\{tw : t \in \mathbb{R}\}$ for some $w \in X$, $w \neq 0$. As $\|u\| < 1$, it follows that $\|u + tw\| = 1$ for a suitable $t \in \mathbb{K}$. Any weak neighborhood of any point of B_X thus contains a point of the unit sphere. We conclude that the unit sphere of B_X is weakly dense in B_X.

The proof of the second statement is similar. □

Proposition 7.5 *(i) The weak topology of a normed space X is metrizable (i.e., it is induced by a metric) iff the space is finite-dimensional.*

(ii) If X is a Banach space, then the weak star topology of its dual X' is metrizable iff X is finite-dimensional.

Proof. The "if" part of (i) is already known. Let us assume that $\dim(X) = \infty$ and that the weak topology is induced by a metric d. By Proposition 7.2, for any $n \in \mathbb{N}$ then there exists $u_n \in X$ with $d(u_n, 0) < 1/n$ and $\|u_n\| > n$. The sequence $\{u_n\}$ thus weakly converges to 0, although it is unbounded in norm; this contradicts part (i) of Proposition 6.3.

The proof of part (ii) is based on part (ii) of Proposition 6.3 and is analogous. □

If X is not complete, it may happen that $\dim(X) = \infty$ and nevertheless the weak star topology of X' is metrizable.

7.1 A Riesz theorem

Lemma 7.6 (*Riesz) Let M be a proper closed subspace of a normed space X over $\mathbb{K}$, and let $\theta < 1$. Then*

$$\exists u \in X : \quad \|u\| = 1, \quad d(u, M) = \inf\big\{\|u - v\| : v \in M\big\} \geq \theta. \tag{7.2}$$

Proof. Let us fix any $w \in X \setminus M$. As M is closed, $c := \inf\{\|w - v\| : v \in M\} > 0$. Hence there exists $z \in M$ such that $\|w - z\| \leq c/\theta$. Let us set $u := (w - z)/\|w - z\|$, and select any $v \in M$. As $z + \|w - z\|v \in M$, we have $\big\|w - z - \|w - z\|v\big\| \geq c$. Therefore

$$\|u - v\| = \frac{1}{\|w - z\|}\big\|w - z - \|w - z\|v\big\| \geq \frac{\theta}{c}\, c = \theta. \square \tag{7.3}$$

The Riesz Lemma holds for $\theta = 1$ for reflexive spaces. [] For instance, it fails in $X = \{v \in C^0([0,1]) : v(0) = 0\}$ for $M = \{v \in X : \int_0^1 v(s)\, ds = 0\}$. []

Theorem 7.7 (*Riesz) The closed unit ball of a normed space is compact iff the space is finite-dimensional.*

Therefore any compact subset of an infinite-dimensional normed space has empty interior.

Proof. The "if" part is straightforward, since in $X = \mathbb{K}^N$ ($N \geq 1$), and thus in any finite-dimensional space, a set is compact iff it is closed and bounded.

Let us next assume that X has infinite dimension, and prove the "only if" part. Via the Riesz Lemma 7.6 we inductively construct a sequence $\{u_n\} \subset X$ such that $\|u_n\| = 1$ for any n and $\|u_n - u_m\| > 1/2$ whenever $n \neq m$. This has no convergent subsequence. Thus the closed unit ball of X is not compact. □

In conclusion, we have seen that for any normed space over the field $\mathbb{K}$ the following conditions are mutually equivalent:

— the space has finite dimension;
— the space is topologically isomorphic to $\mathbb{K}^N$ for some $N \geq 0$;
— the weak topology is metrizable;
— the algebraic dual coincides with the topological dual;
— the origin has a bounded weak neighbourhood;
— the weak and the strong topology coincide;
— a closed ball (equivalently, any closed ball) is compact.

Anyway in some infinite-dimensional spaces any weakly convergent sequence is strongly convergent, by the Schur phenomen.

8 Compactness

8.1 Strong compactness

The family of relatively compact subsets of a normed space includes bounded subsets of finite-dimensional subspaces, convergent sequences, finite unions of sets of these two classes, and their convex hull.

The next nontrivial result provides a precise characterization.

Theorem 8.1 (*Grothendieck) A subset K of a Banach space X is relatively compact iff there exists a vanishing sequence $\{u_n\}$ in X such that $K \subset \overline{\text{co}}\,(\{u_n\})$.* [30] *[]*

Proof of the "if" part. Let $\{u_n\}$ be as prescribed by the theorem. As $u_n \to 0$, for any $\varepsilon > 0$ there exist $m \in \mathbb{N}$ such that $\{u_n\} \subset \{u_1, ..., u_m\} + B(0, \varepsilon)$. Let us denote the convex hull of $\{u_1, ..., u_m\}$ by A_ε; this is a bounded subset of a finite-dimensional subspace, hence it is totally bounded. Since $A_\varepsilon + B(0, \varepsilon)$ is convex and includes $\{u_n\}$, we infer that $K \subset A_\varepsilon + B(0, \varepsilon)$. We have thus shown that

$$\begin{aligned} &\forall \varepsilon > 0, \text{ there exists a bounded subset } A_\varepsilon \text{ of a} \\ &\text{finite-dimensional subspace such that } K \subset A_\varepsilon + B(0, \varepsilon). \end{aligned} \tag{8.1}$$

Therefore K is totally bounded, hence relatively compact. □

• **Corollary 8.2** *A subset K of a Banach space X is relatively compact iff* (8.1) *is fulfilled.*

Proof. By the Grothendieck theorem, there exists a vanishing sequence $\{u_n\}$ such that $K \subset \overline{\text{co}}\,(\{u_n\})$. In the foregoing proof we have shown that this entails (8.1), and that in turn (8.1) entails that K is relatively compact. □

Grothendieck's characterization shows that compact subsets in infinite-dimensional Banach spaces are somehow rare. The prominent role of compactness in applications, e.g., in the analysis of partial differential equations, induces one to search for a larger family of compact subsets, corresponding to a coarser topology. As we are going to see, an answer is provided by the use of the weak star topology in dual spaces.

Some topological notions. A topological space A is called **compact** iff every open covering of A has a finite subcovering. On the other hand A is called **sequentially compact** iff every sequence in A has a subsequence that converges to an element of A. In metrizable spaces the two properties are equivalent, but in nonmetrizable topological spaces in general there is no implication between them.

A subset A of a Hausdorff space H is called **relatively compact in** H iff its closure $\overline{A}$ is compact. A is called **relatively sequentially compact** iff every sequence in A has a convergent subsequence (whose limit however may not belong to A).

8.2 Weak and weak star compactness

Bounded subsets of $L^1(0, 1)$ need not be relatively weakly compact: consider e.g. the sequence $\{n\chi_{[0,1/n]}\}$. [31] A further weakening of the topology is then in order in our search for compactness. For dual spaces the weak star topology may be regarded as a sensible point of arrival, because of the next classical result.

[30] By $\overline{\text{co}}\,(A)$ we denote the closed convex hull of any set A.

[31] The Schwartz theory of distributions shows that this sequence converges to δ_0 (the celebrated *delta of Dirac)* in a space which is larger than $L^1(0, 1)$.

• **Theorem 8.3** (*Banach-Alaoglu) For any normed space X, the closed unit ball $B_{X'}$ of the dual space is weakly star compact.*

* *Proof.* The family $\mathcal{F}$ of mappings $f : X \to \mathbb{K}$ such that $|f(x)| \leq \|x\|$ for any x can be identified with the set of their graphs, namely, $P := \prod_{x \in X}[\{x\} \times B(0, \|x\|)]$. $B_{X'}$ equipped with the weak star topology, i.e. the topology of pointwise convergence, is homeomorphic to a subset $\widetilde{P}$ of P equipped with the restriction of the product topology of P. The space P is compact because of the Tychonoff lemma below. If we show that $\widetilde{P}$ is closed in P, we can then conclude that $B_{X'}$ is weakly star compact.

In order to prove this property, let us consider the sets

$$A_{x,y} = \{f \in \mathcal{F} : f(x+y) - f(x) - f(y) = 0\} \qquad \forall x, y \in X,$$
$$S_{\lambda,x} = \{f \in \mathcal{F} : f(\lambda x) - \lambda f(x) = 0\} \qquad \forall \lambda \in \mathbb{K}, \forall x \in X,$$

so that

$$B_{X'} = \Big(\bigcap_{x,y \in X} A_{x,y}\Big) \cap \Big(\bigcap_{\lambda \in \mathbb{K}, x \in X} S_{\lambda,x}\Big).$$

As each $A_{x,y}$ and each $S_{\lambda,x}$ are identified with closed subsests of P, the same holds for $B_{X'}$. □

Lemma 8.4 (*Tychonoff) Any Cartesian product of compact topological spaces is compact, if it is equipped with the product topology. []*

Corollary 8.5 *For any normed space X, any bounded subset of X' is relatively weakly star compact.*

By the next statement, any normed space may be identified with a closed subspace of $C^0(K)$, for a suitable compact K.

* **Proposition 8.6** *Let X be a Banach space, equip $B_{X'}$ with the weak star topology and $C^0(B_{X'})$ with the max norm. Then X is isometrically isomorphic to a closed subspace of $C^0(B_{X'})$.*

Proof. By the Banach-Alaoglu Theorem 8.3, $B_{X'}$ is weakly star compact; hence $C^0(B_{X'})$ is a Banach space. Denoting by J the canonical imbedding $X \to X''$, the operator $\widetilde{J} : X \to C^0(B_{X'}) : u \mapsto Ju\big|_{B_{X'}}$ (the restriction of Ju to $B_{X'}$) is an isometry, because for any $u \in X$ we have

$$\|\widetilde{J}u\|_\infty = \max_{f \in B_{X'}} |(\widetilde{J}u)(f)| = \max_{f \in B_{X'}} | < f, u > | = \|u\|_X.$$

As X is complete, $\widetilde{J}(X)$ is a closed subspace of $C^0(B_{X'})$. □

"Thus, in a sense the entire theory of normed spaces is contained in the theory of subspaces of normed spaces $C^0(K)$ such that K is a compact Hausdorff space. This by no means trivializes the theory of normed spaces, but rather serves to point out the richness of the theory of the spaces $C^0(K)$." [Megginson p. 231] (verbatim).

Remarks. (i) Whenever X is an infinite-dimensional Banach space, X' is not locally compact w.r.t. the weak star topology, in spite of the Banach-Alaoglu theorem. Actually, any weak star neighbourhood of the origin is unbounded, hence noncompact.

(ii) The net theorem is strictly related to the Banach-Alaoglu Theorem 8.3, and is often used in the analysis of partial differential equations. (This is the original formulation of the theorem, which was then extended by Alaoglu.)

• **Corollary 8.7** (*Banach) For any separable normed space X, $B_{X'}$ is weakly star sequentially compact.*

Any bounded subset of X' then fulfills the same property.

9 The Ascoli-Arzelà theorem

In this section we state and prove the Ascoli-Arzelà theorem. This is one of the main results of compactness, jointly with the Banach-Alaoglu theorem.

9.1 The Ascoli-Arzelà theorem

The following classical result conveys an important characterization of (relative) compactness in $C^0(K)$ (K being a compact metric space). This is relevant since the space $C^0(K)$ has no predual, and thus here one cannot use the Banach-Alaoglu theorem.

• **Theorem 9.1** (*Ascoli-Arzelà) Let K be a compact metric space. A subset $\mathcal{F}$ of $C^0(K)$ is relatively compact if (and only if) it is (equi)bounded as well as uniformly equicontinuous in $C^0(K)$, that is,* [32]

$$\sup\big\{|u(x)| : x \in K, u \in \mathcal{F}\big\} < +\infty, \tag{9.1}$$

$$\sup\big\{|u(x) - u(y)| : x, y \in K, d(x, y) \le h, u \in \mathcal{F}\big\} \to 0 \qquad \text{as } h \to 0. \tag{9.2}$$

**Proof.* It suffices to show that a Cauchy subsequence may be extracted from any sequence in $\mathcal{F}$. Let $\{x_j\}_{j\in\mathbb{N}}$ be a (countable) dense subset in K. Because of the boundedness of $\mathcal{F}$, from $\{u_n\}$ one may extract a subsequence $\{u_{n_1}\}$ such that $\{u_{n_1}(x_1)\}$ converges. Similarly, for $j = 2, 3, ...,$ from $\{u_{n_{j-1}}\}$ one may iteratively extract a subsequence $\{u_{n_j}\}$ in such a way that $\{u_{n_j}(x_\ell)\}$ converges for all $\ell \le j$. By a diagonalization procedure, for any $m \in \mathbb{N}$ let us then define $\widetilde{u}_m$ as the mth element of the sequence u_{n_m}. Thus $\{\widetilde{u}_m\}$ is a subsequence extracted not only from the initial sequence $\{u_n\}$ but also from $\{u_{n_j}\}$ for any $j \in \mathbb{N}$; moreover $\{\widetilde{u}_m(x_j)\}$ converges for any $j \in \mathbb{N}$.

Let us now fix any $\varepsilon > 0$. By equicontinuity there exists a $\delta > 0$ such that

$$|\widetilde{u}_m(x_j) - \widetilde{u}_m(y)| \le \varepsilon \qquad \forall y \in K \cap B(x_j, \delta), \forall j, m \in \mathbb{N}. \tag{9.3}$$

By the compactness of K, a finite subcovering $\{B(x_j, \delta)\}_{j\in J}$ may be extracted from the family of the open balls $\{B(x_j, \delta)\}_{j\in\mathbb{N}}$. Therefore, for any m', m'',

$$\begin{aligned} &|\widetilde{u}_{m'}(y) - \widetilde{u}_{m''}(y)| \\ &\le |\widetilde{u}_{m'}(y) - \widetilde{u}_{m'}(x_j)| + |\widetilde{u}_{m'}(x_j) - \widetilde{u}_{m''}(x_j)| + |\widetilde{u}_{m''}(x_j) - \widetilde{u}_{m''}(y)| \\ &\overset{(9.3)}{\le} \varepsilon + |\widetilde{u}_{m'}(x_j) - \widetilde{u}_{m''}(x_j)| + \varepsilon \qquad \forall y \in K \cap B(x_j, \delta), \forall j \in J, \end{aligned}$$

whence

$$\max_{y\in K} |\widetilde{u}_{m'}(y) - \widetilde{u}_{m''}(y)| \le \max_{j\in J} |\widetilde{u}_{m'}(x_j) - \widetilde{u}_{m''}(x_j)| + 2\varepsilon.$$

[32] Because of the compactness of K, (9.2) is equivalent to the following (otherwise weaker) condition of *pointwise* equicontinuity:

$$\sup_{u\in\mathcal{F}} \sup\big\{|u(x) - u(y)| : y \in K, d(x, y) \le h, u \in \mathcal{F}\big\} \to 0 \qquad \text{as } h \to 0, \forall x \in K. \quad [Ex]$$

For any j, $\{\widetilde{u}_m(x_j)\}$ is a Cauchy sequence (in $\mathbb{R}$) as it converges. $\{\widetilde{u}_m\}$ is then a Cauchy sequence in $C^0(K)$. □

• **Corollary 9.2** *Let Ω be a bounded convex subset of $\mathbb{R}^N$. For any nonnegative integers m, n with $m > n$, any bounded subset $C^m(\bar{\Omega})$ is relatively compact in $C^n(\bar{\Omega})$.*

Proof. It is easily checked that it suffices to prove the thesis for $m = 1$ and $n = 0$, since then the general statement follows by applying this result to the derivatives. For any $u \in C^1(\bar{\Omega})$ and any $x, y \in \Omega$, by the convexity of Ω and the mean-value theorem, there exists $\lambda \in [0, 1]$ such that, setting $\xi_\lambda = \lambda x + (1 - \lambda)y$ $(\in \Omega)$,

$$|u(x) - u(y)| = |\nabla u(\xi_\lambda) \cdot (y - x)| \le \sup_{\bar{\Omega}} |\nabla u|\, |y - x| \le \|u\|_{C^1(\bar{\Omega})}\, |y - x|.$$

Any bounded subset $\mathcal{F}$ of $C^1(\bar{\Omega})$ is thus equicontinuous. As $\mathcal{F}$ is also bounded in $C^0(\bar{\Omega})$, by the Ascoli-Arzelà theorem it is then relatively compact in $C^0(\bar{\Omega})$. □

10 Adjoint Operator

Throughout this section X and Y will be Banach spaces over $\mathbb{K}$ $(= \mathbb{R}$ or $\mathbb{C})$.

10.1 Definition and basic properties
Let $L \in \mathcal{L}(X; Y)$. By setting

$$L'f := f \circ L \ (\in X') \qquad \forall f \in Y', \tag{10.1}$$

we define a linear mapping $L' : Y' \to X'$, called the **Banach adjoint** (or just the **adjoint**) of L. Using duality pairings, an equivalent definition of L' is provided by

$$\langle L'f, u\rangle := \langle f, Lu\rangle \qquad \forall f \in Y', \forall u \in X. \tag{10.2}$$

Since

$$\begin{aligned} \sup_{\|f\|\le 1} \|L'f\| &= \sup_{\|f\|\le 1} \sup_{\|u\|\le 1} |\langle L'f, u\rangle| \\ &= \sup_{\|u\|\le 1} \sup_{\|f\|\le 1} |\langle f, Lu\rangle| = \sup_{\|u\|\le 1} \|Lu\| = \|L\|, \end{aligned} \tag{10.3}$$

we have $L' \in \mathcal{L}(Y'; X')$ with $\|L'\| = \|L\|$. By iterating this process, we obtain the adjoint of the adjoint: $L'' = (L')' : X'' \to Y''$.

Next we state some simple properties, and leave the proofs to the reader.

Proposition 10.1 *Let X and Y be Banach spaces and $L \in \mathcal{L}(X; Y)$. Then:*
(i) The mapping $L \mapsto L'$ is linear, that is,

$$(L_1 + L_2)' = L_1' + L_2', \quad (\alpha L)' = \alpha L', \tag{10.4}$$

for any $L, L_1, L_2 \in \mathcal{L}(X; Y)$ and any scalar α.
(ii) Denoting by $J_X : X \to X''$ and $J_Y : Y \to Y''$ the canonical embeddings,

$$L'' \circ J_X = J_Y \circ L \qquad \forall L \in \mathcal{L}(X; Y). \tag{10.5}$$

(iii) For any Banach spaces X_1, X_2, and X_3,

$$(L_2 \circ L_1)' = L_1' \circ L_2' \qquad \forall L_1 \in \mathcal{L}(X_1; X_2), \forall L_2 \in \mathcal{L}(X_2; X_3). \tag{10.6}$$

(iv) For any $L \in \mathcal{L}(X;Y)$, L' is invertible iff so is L. In this case

$$(L')^{-1} = (L^{-1})'. \tag{10.7}$$

Moreover, L' is an isometric isomorphism iff so is L.
(v) If X and Y are reflexive, then the mapping $\mathcal{L}(X;Y) \to \mathcal{L}(Y';X') : L \mapsto L'$ is surjective.

To prove (10.7), notice that $(L^{-1})L = I = L(L^{-1})$, whence by (10.6)

$$L'(L^{-1})' = I' = I = (L^{-1})'L'.$$

Examples. **(i)** Let $X = \mathbb{K}^N$, $Y = \mathbb{K}^M$, let $L \in \mathcal{L}(X;Y)$ be represented by the matrix $A \in \mathbb{K}^{M\times N}$ with respect to the canonical basis. Through the isomorphism between X and X' given by $\langle x', u\rangle = x^T u = \sum_j x_j u_j$, $x \in \mathbb{K}^N$, and analogously for Y, the adjoint $L' \in \mathcal{L}(Y';X')$ is represented by the transposed matrix $A^T \in \mathbb{K}^{N\times M}$ of A.

(ii) Let us compute the adjoints $S_r', S_\ell' \in \mathcal{L}(X')$ of the shift operators S_r, S_ℓ for $X = \ell^p_{\mathbb{K}}$, $1 \le p < +\infty$. Let us make use of the isometry between $(\ell^p_{\mathbb{K}})'$ and $\ell^q_{\mathbb{K}}$, $\langle f, u\rangle = \sum_{k\in\mathbb{N}} f_k u_k$ for any $u \in \ell^p_{\mathbb{K}}$, where $q = p/(p-1)$ if $p \ne 1$ and $q = \infty$ if $p = 1$. Then for any $u \in \ell^p_{\mathbb{K}}$ and $f \in \ell^q_{\mathbb{K}}$ we have

$$\langle S_r' f, u\rangle = \langle f, S_r u\rangle = \sum_{k=2}^{\infty} f_k u_{k-1} = \sum_{h=1}^{\infty} f_{h+1} u_h = \langle S_\ell f, u\rangle, \tag{10.8}$$

and analogously $\langle S_\ell' f, u\rangle = \langle S_r f, u\rangle$. Thus $S_r' = S_\ell$, $S_\ell' = S_r$. Incidentally, notice that $S_\ell S_r = I$, but $S_r S_\ell \ne I$.

10.2 Annihilators

For any subset M of X, we define the **right annihilator** of M in X' as

$$M^0 = \{f \in X' : \langle f, v\rangle = 0, \ \forall v \in M\}.\text{[33]} \tag{10.9}$$

For any subset N of X', we similarly define the right annihilator of N in X'', and also the **left annihilator** of N in X as

$${}^0N = \{v \in X : \langle f, v\rangle = 0, \ \forall f \in N\}. \tag{10.10}$$

Note that M^0 is weakly star closed in X', and ${}^0N$ is weakly closed in X.

*** Proposition 10.2** *Let X be a Banach spaces. Then:*
(i) For any linear subspace M of X,

$${}^0(M^0) = \overline{M}^{weak} = \overline{M}^{strong}. \tag{10.11}$$

(ii) For any linear subspace N of X',

$$({}^0N)^0 = \overline{N}^{weak*} \supset \overline{N}^{strong}, \tag{10.12}$$

possibly with strict inclusion. However if X is reflexive, then the latter inclusion is an equality: $({}^0N)^0 = \overline{N}^{weak} = \overline{N}^{strong}$.

[33] Some authors use of the symbol "$\perp$" for both the annihilator (a subset of X') and, in Hilbert space, the orthogonal complement (a subset of X) is a standard practice, although it might lead to confusion. Some authors actually name annihilators *orthogonal complements* in Banach space, too.

**Proof.* Since ${}^0(M^0)$ and $({}^0N)^0$ are weakly resp. weakly star closed, the inclusions "$\supset$" that are implicit in the equalities in (10.11) and (10.12) directly follow from the definitions. The reverse inclusion "$\subset$" holds in (10.11), since for any $u \notin \overline{M}^{\text{strong}}$ there is an $f \in M^0$ with $f(u) \neq 0$ by a corollary of the Hahn-Banach theorem. It remains to prove "$\subset$" for (10.12). Let $f \in X'$ with $f \notin \overline{N}^{\text{weak}*}$; it suffices to find a $v \in {}^0N$ such that $f(v) \neq 0$.

To this end, let U be a weak star neighbourhood of 0 such that $(f+U) \cap N = \emptyset$, or equivalently $(N-f) \cap U = \emptyset$. According to the construction of the weak star topology, we may assume U to be of the form $\{h \in X' : |h(u_k)| < \varepsilon,\ 1 \leq k \leq n\}$ where $u_1, \ldots, u_n \in X$ and $\varepsilon > 0$. Let us define $T : X' \to \mathbb{R}^n$ by $(Tg)_k = g(u_k)$ for $1 \leq k \leq n$. Since $\mathcal{N}(T) \subset U$ and $(N-f) \cap U = \emptyset$, we have $Tf \notin T(N)$. By the Hahn-Banach theorem there exists $w \in (\mathbb{R}^n)' \simeq \mathbb{R}^n$ such that $\langle w, Tf \rangle \neq 0$ and $\langle w, Tg \rangle = 0$ for every $g \in N$. Setting $v = \sum_k w_k u_k$ we obtain that $f(v) = \sum_k w_k f(u_k) = \langle w, Tf \rangle \neq 0$ and that $g(v) = \langle w, Tg \rangle = 0$ for every $g \in N$. Thus $v \in {}^0N$ and the proof is complete.

The final statement is a simple consequence of (10.12). □

The range $\mathcal{R}(L)$ and the null space $\mathcal{N}(L)$ of an operator L are related to the range and null space of its adjoint L' as follows. [34]

• **Theorem 10.3** *For any $L \in \mathcal{L}(X;Y)$,*

$$\mathcal{R}(L)^0 = \mathcal{N}(L'), \qquad {}^0\mathcal{R}(L') = \mathcal{N}(L), \tag{10.13}$$

$$\mathcal{N}(L)^0 = \overline{\mathcal{R}(L')}^{weak*}, \qquad {}^0\mathcal{N}(L') = \overline{\mathcal{R}(L)}. \tag{10.14}$$

Proof. Let us prove (10.13). If $v \notin \mathcal{N}(L)$, then by the Hahn-Banach theorem $\langle L'g, v \rangle = \langle g, Lv \rangle \neq 0$ for some $g \in Y'$. Thus $v \notin {}^0\mathcal{R}(L')$ and consequently ${}^0\mathcal{R}(L') \subset \mathcal{N}(L)$.

The other three inclusions in (10.13) directly follow from the definitions. [Ex] For instance, for any $v \in \mathcal{N}(L)$ and any $g \in Y'$, $\langle L'g, v \rangle = \langle g, Lv \rangle = g(0) = 0$; thus $v \in {}^0\mathcal{R}(L')$ and consequently $\mathcal{N}(L) \subset {}^0\mathcal{R}(L')$.

The equalities (10.14) are a consequence of those in (10.13) and those in Proposition 10.2. Namely, $\mathcal{N}(L)^0 = ({}^0\mathcal{R}(L'))^0 = \overline{\mathcal{R}(L')}^{\text{weak}*}$, and ${}^0\mathcal{N}(L') = {}^0(\mathcal{R}(L)^0) = \overline{\mathcal{R}(L)}$. □

[34] If $\mathcal{R}(L)$ is closed, the closed range theorem given below will provide additional information.

Chapter II — Hilbert Spaces

Contents: 1. The inner product. 2. Orthogonality and projections. 3. The representation theorem 4. Orthonormal systems and Hilbert bases.

11 The Inner Product

11.1 Inner products and basic properties

Let H be a linear space over the field $\mathbb{K}$. A mapping $(\cdot,\cdot) : H \times H \to \mathbb{K}$ is called an **inner product** (or a **scalar product)** [35] over H iff it fulfills the following properties:

$$\text{the functional } H \to \mathbb{K} : u \mapsto (u,v) \text{ is linear over } \mathbb{K} \qquad \forall v \in H, \tag{11.15}$$

$$(u,v) = \overline{(v,u)} \qquad \forall u,v \in H, \tag{11.16}$$

$$(u,u) > 0 \qquad \forall u \in H \setminus \{0\}. \tag{11.17}$$

[36]

Hence $(u,u) \in \mathbb{R}$ for all $u \in H$, and

$$(u,u) = 0 \quad \Leftrightarrow \quad u = 0. \tag{11.18}$$

The properties (11.15) and (11.16) obviously entail that

$$\text{the functional } H \to \mathbb{K} : v \mapsto \overline{(u,v)} \text{ is linear over } \mathbb{K} \qquad \forall u \in H. \tag{11.19}$$

A linear space over $\mathbb{C}$ ($\mathbb{R}$, resp.) equipped with an inner product is called a complex (real, resp.) **inner-product space**, or a **pre-Hilbert space**. Here is some further terminology:
(11.16) $\Leftrightarrow$ $(\cdot,\cdot)$ is **Hermitian** (or **skew-symmetric**) if $\mathbb{K} = \mathbb{C}$, **symmetric** if $\mathbb{K} = \mathbb{R}$;
(11.19) $\Leftrightarrow$ $(u,\cdot)$ is **antilinear**, or **conjugate-linear**, or **skew-linear**;
(11.15) and (11.19) $\Leftrightarrow$ $(\cdot,\cdot)$ is **sesquilinear** if $\mathbb{K} = \mathbb{C}$, **bilinear** if $\mathbb{K} = \mathbb{R}$.

Henceforth, when dealing with an inner-product space, we set

$$\|u\| := \sqrt{(u,u)} \qquad \forall u \in H. \tag{11.20}$$

By (11.15) – (11.17) above and by (11.23) ahead, we infer that $\|\cdot\|$ is indeed a norm over H, which is then called a **Hilbert norm**. Dealing with an inner-product space we shall always refer to this norm, if not otherwise specified.

Proposition 11.1 *If H is an inner-product space* [37] *over the field $\mathbb{K}$, then*

$$|(u,v)| \le \|u\|\|v\| \qquad \forall u,v \in H \quad \text{(Cauchy-Schwarz inequality)}, \tag{11.21}$$

$$\|u+v\|^2 + \|u-v\|^2 = 2\|u\|^2 + 2\|v\|^2 \quad \forall u,v \in H \text{ (parallelogram identity)}, \tag{11.22}$$

$$\|u+v\| \le \|u\| + \|v\| \qquad \forall u,v \in H \quad \text{(triangle inequality)}, \tag{11.23}$$

$$\textit{the mapping } (\cdot,\cdot) : H \times H \to \mathbb{K} \textit{ is continuous.} \tag{11.24}$$

[35] The notation $(\cdot,\cdot)$ is traditional. Unfortunately it is also used to denote pairs.
[37] Actually, here (11.15) and (11.16) suffice, the definiteness (11.17) is not needed as the proof shows.

Proof. (i) Let us prove (11.21). Without loss of generality, we may assume that $\mathbb{K} = \mathbb{C}$ and $v \neq 0$. By (11.15) and (11.16), [38]

$$\begin{aligned} 0 &\leq (u + \lambda v, u + \lambda v) = (u, u) + |\lambda|^2 (v, v) + (\lambda v, u) + (u, \lambda v) \\ &= \|u\|^2 + |\lambda|^2 \|v\|^2 + 2Re[\lambda(v, u)] \qquad \forall \lambda \in \mathbb{C}. \end{aligned}$$

By taking $\lambda = -(u, v)/\|v\|^2$, we then get

$$0 \leq \|u\|^2 + \frac{|(u, v)|^2}{\|v\|^2} - 2\frac{|(u, v)|^2}{\|v\|^2} = \|u\|^2 - \frac{|(u, v)|^2}{\|v\|^2},$$

and this yields (11.21).

(ii) In order to check (11.22), notice that

$$\begin{aligned} \|u + v\|^2 = (u + v, u + v) &= \|u\|^2 + (u, v) + (v, u) + \|v\|^2 \\ &= \|u\|^2 + 2Re(u, v) + \|v\|^2, \end{aligned} \tag{11.25}$$

and similarly $\|u - v\|^2 = \|u\|^2 - 2Re(u, v) + \|v\|^2$. Summing these equalities we get (11.22).

(iii) By (11.25) and by the Cauchy-Schwarz inequality we have

$$\|u + v\|^2 \leq \|u\|^2 + 2\|u\|\|v\| + \|v\|^2 = \big(\|u\| + \|v\|\big)^2 \qquad \forall u, v \in H,$$

that is (11.23).

(iv) Let two sequences $\{u_n\}$ and $\{v_n\}$ in H be such that $u_n \to u$ and $v_n \to v$ in the norm topology. We have

$$\begin{aligned} |(u, v) - (u_n, v_n)| &\leq |(u, v) - (u_n, v)| + |(u_n, v) - (u_n, v_n)| \\ &\leq \|u - u_n\|\|v\| + \|u_n\|\|v - v_n\|; \end{aligned}$$

as $\|u_n\|$ is uniformly bounded, the latter sum vanishes as $n \to \infty$. □

It is promptly checked that the Cauchy-Schwarz inequality (11.21) is reduced to an equality iff u and v are linearly dependent. The same holds for the triangle inequality (11.23).

Remark. Some properties of inner-product spaces involve just a finite number of elements, have a geometric content, and may be expressed in terms of the linear subspace that is spanned by those elements; this may make their proof especially simple. This is the case e.g. for (11.21)–(11.23).

11.2 The polarization identity

The denomination *parallelogram identity* of (11.22) is easily understood by considering the parallelogram of vertices $0, u, v, u + v$ in the two-dimensional subspace spanned by u and v (assuming $u, v \neq 0$ and $u \neq v$). By (11.22), it is promptly checked that the sum of the squared lengths of the sides equals the sum of the squared lengths of the diagonals. In the plane this is known as the *Apollonius theorem.*

We saw that the inner product determines a norm which fulfills the parallelogram identity (11.22). The polarization identity that is displayed in the next lemma relates the inner product to the corresponding norm. More generally, Theorem 11.2 below shows that any normed space in which the norm fulfills the parallelogram identity (11.22) is an inner-product space, in which the inner product is defined via the polarization identity. The parallelogram identity thus characterizes inner-product spaces in the class of normed spaces. [39]

[38] We shall still denote by $Re(z)$ the real part of any complex number z.
[39] The are several other characterizations. See e.g. [de Figueiredo-Karlovitz].

Theorem 11.2 (*P. Jordan, von Neumann) If H is a complex normed space equipped with a norm $\|\cdot\|$ that fulfills the parallelogram identity* (11.22)*, then*

$$\begin{aligned}(u,v) &= \tfrac{1}{4}\textstyle\sum_{n=1}^{4} i^n\|u+i^n v\|^2 \\ &= \tfrac{1}{4}\big[(\|u+v\|^2-\|u-v\|^2)+i(\|u+iv\|^2-\|u-iv\|^2)\big] \quad \forall u,v\in H\end{aligned} \tag{11.26}$$

is an inner product, and is related to the norm $\|\cdot\|$ by (11.20)*.*

If H is a real space, then (11.26) *is replaced by*

$$\begin{aligned}(u,v) &= \frac{1}{4}\big(\|u+v\|^2-\|u-v\|^2\big) \\ &= \frac{1}{2}\big(\|u+v\|^2-\|u\|^2-\|v\|^2\big) \qquad \forall u,v\in H.\end{aligned} \tag{11.27}$$

For $\mathbb{K}=\mathbb{R}$ the proof is straightforward. For $\mathbb{K}=\mathbb{R}$ it is a slightly technical (and boring). □

11.3 Hilbert spaces

An inner-product space is called a **Hilbert space** whenever it is complete w.r.t. the induced norm. A (closed) Banach subspace of a Hilbert space is itself a Hilbert space, and is called a Hilbert subspace. Henceforth we shall confine ourselves to Hilbert spaces, which is the most relevant class for applications, although for several results the completeness is not really needed.

Two Hilbert space H_1 and H_2 are called **isometrically isomorphic** iff there exists a linear surjective operator $U: H_1 \to H_2$ such that $(Uu,Uv)_{H_2} = (u,v)_{H_1}$ for any $u,v\in H_1$. This is tantamount to $\|Uu\|_{H_2} = \|u\|_{H_1}$ for any $u\in H_1$ Such an operator is called **unitary.**

11.4 Examples

(i) For any $N\geq 1$, $\mathbb{K}^N$ is a Hilbert space over $\mathbb{K}$ equipped with the inner product

$$(u,v)_{\mathbb{K}^N} := \textstyle\sum_{n=1}^{N} u_n\overline{v_n} \qquad \forall u,v\in\mathbb{K}^N. \tag{11.28}$$

$\mathbb{C}^N$ can also be equipped with the structure of Hilbert space over $\mathbb{R}$: this corresponds to identifying $\mathbb{R}^{2N}$ with $\mathbb{C}^N$ via the mapping $(u_1,...,u_{2N})\mapsto(u_1+iu_2,...,u_{2N-1}+iu_{2N})$.

(ii) As a particular case of the example (ii), the sequence space ℓ^2 $(=\ell^2_{\mathbb{K}})$ is a Hilbert space over $\mathbb{K}$ equipped with the inner product

$$(u,v) := \textstyle\sum_{n=1}^{\infty} u_n\overline{v_n} \qquad \forall u=\{u_n\}, v=\{v_n\}\in\ell^2. \tag{11.29}$$

(iii) For any measure space $(A,\mathcal{A},\mu)$, $L^2(A,\mathcal{A},\mu;\mathbb{K})$ is a Hilbert space over $\mathbb{K}$ equipped with the inner product

$$(u,v) := \textstyle\int_A u(x)\overline{v(x)}\,d\mu(x) \qquad \forall u,v\in L^2(A,\mathcal{A},\mu). \tag{11.30}$$

(iv) For any $N\geq 1$, one may also define $L^2(A,\mathcal{A},\mu;\mathbb{K}^N)$, $L^2(A,\mathcal{A},\mu;\mathbb{K}^{N\times N})$, [40] ℓ^N, $\ell^{N\times N}$ and so on in a obvious way; these are also Hilbert spaces.

(v) The construction of example (iii) can be extended as follows. For a (nonempty) index set A, let $\{H_\alpha\}_{\alpha\in A}$ be a family of Hilbert spaces over the field $\mathbb{K}$, each H_α being equipped with the inner product $(\cdot,\cdot)_\alpha$ and the associated norm $\|\cdot\|_\alpha$. Let us define the linear space $H := \prod_\alpha H_\alpha$ of all mappings $\alpha\mapsto u_\alpha$ such that the real family $\{\|u_\alpha\|_\alpha : \alpha\in A\}$ is square summable, i.e., $\sum_{\alpha\in A}\|u_\alpha\|_\alpha^2 < +\infty$. Any element of this space has only a finite number of nonvanishing components u_α.

[40] This is a space of matrix-valued functions.

This is a Hilbert space over $\mathbb{K}$, called a **Hilbert direct sum,** if it is equipped with the natural scalar product

$$(u, v) := \textstyle\sum_{\alpha\in A}(u, v)_\alpha \qquad \forall u, v \in H. \tag{11.31}$$

This topology coincides with the product topology iff the index set A is finite.

(vi) The linear space of sequences of $\mathbb{K}$ that only contain a finite number of nonvanishing elements is a noncomplete inner-product subspace of ℓ^2. Its completion coincides with the Hilbert space ℓ^2. (This is a particular case of the previous example.)

(vii) Let Ω be a bounded domain of $\mathbb{R}^N$ ($N \geq 1$). Let us equip the linear space of continuous functions $\bar{\Omega} \to \mathbb{K}$ with the inner product

$$(u, v) := \textstyle\int_\Omega u(x)\overline{v(x)}\,dx \qquad \forall u, v \in C^0(\bar{\Omega}).$$

This space is not complete. For instance, for $\Omega = \,]-1,1[$, $\{u_n : x \mapsto \arctan(nx)\}_{n\in\mathbb{N}}$ is a Cauchy sequence in this space, but it does not converge to any continuous function (it converges a.e. to the discontinuous function $\frac{\pi}{2}$sign). The completion of this space coincides with the Hilbert space $L^2(-1,1)$. [Ex]

On the other hand, ℓ^p, $L^p(0,1)$ with $p \neq 2$ and $C^k(\Omega)$ ($k \in \mathbb{N}$) are not Hilbert spaces. Actually, their respective norms do not fulfill the parallelogram identity, and these spaces are not Hilbertizable.

11.5 Exercises

1. Let V be a complex linear space and $b_1(\cdot,\cdot)$, $b_2(\cdot,\cdot)$ be two sesquilinear mappings $V\times V \to \mathbb{C}$ such that $b_1(v,v) = b_2(v,v)$ for any $v \in V$. Show that these two mappings then coincide on the whole $V \times V$.

2. Let V be a complex linear space and a mapping $b(\cdot,\cdot) : V \times V \to \mathbb{C}$ be either sesquilinear and symmetric (rather than skew-symmetric), or bilinear and skew-symmetric (rather than symmetric). Show that then in either case $b(u,v) = 0$ for any $u, v \in V$.

3. Let V be a complex linear space. Show that a sesquilinear mapping $(\cdot,\cdot) : V \times V \to \mathbb{C}$ is Hermitian iff the associated quadratic mapping $V \to \mathbb{C} : v \mapsto (v,v)$ is real-valued.

   Notice that this fails in real spaces (with "symmetric" in place of "Hermitian")!

4. Let H be the set of all complex sequences $\{x_n\}$ such that

   $$\|\{x_n\}\| := \Big(\sum_{n=1}^{\infty} \|x_n\|^4\Big)^{1/4} + \Big(\sum_{n=1}^{\infty} \|x_n\|^2\Big)^{1/2} < +\infty.$$

   Check that this is a norm on H.

   (i) Is this norm equivalent to a Hilbert norm?
   (ii) Does H coincide with any of the known sequence spaces?
   (iii) Formulate an analogous exercise in terms of Lebesgue functions $]0,1[\to \mathbb{R}$ (instead of sequences, with integrals instead of series), and answer the analogous questions.

5. Let H be an inner-product space and $x, y \in H$. Show that

   $$x \perp y \quad \Leftrightarrow \quad \|x + \lambda y\| \geq \|x\| \quad \forall \lambda \in \mathbb{K}.$$

6. Let H be an inner-product space. Show that for any $x, y, z \in H$

   $$\|x - z\| = \|x - y\| + \|y - z\| \quad \Leftrightarrow \quad \exists \lambda \in \,]0,1[:\ y = \lambda x + (1-\lambda)z.$$

7. Let H be an inner-product space and $x, y \in H$. Show that if H is a real space then

$$x \perp y \quad \Leftrightarrow \quad \|x+y\|^2 = \|x\|^2 + \|y\|^2,$$

and find a counterexample for a complex space.

8. Let H be a complex inner-product space and $x, y \in H$. Show that

$$x \perp y \quad \Leftrightarrow \quad \|ax+by\|^2 = a^2\|x\|^2 + b^2\|y\|^2 \quad \forall a, b \in \mathbb{C}.$$

9. Let $\{u_n\}$ be a sequence in a Hilbert space H and $u \in H$ be such that $u_n \to u$ weakly and $\|u_n\| \to \|u\|$. Prove that then $u_n \to u$.

   Hint: Develop the square $\|u_n - u\|^2$ ...

10. Let H be a Hilbert space, $L : \mathrm{Dom}(L) \subset H \to H$ be a closed unbounded linear operator, and equip the linear space $X = \{v \in H : Lv \in H\}$ with the norm $\|v\|_X = \|v\|_H + \|Lv\|_H$ [which is named *the norm of the graph* of L in H].

    (i) Check that this is a Banach space.
    (ii) Is this space Hilbertizable?

11. Let $1 \le p \le +\infty$. In the linear space ℓ^p set $|||\cdot||| = \|\cdot\|_{\ell^2} + \|\cdot\|_{\ell^p}$.

    (i) $|||\cdot|||$ is a norm for some p?
    (ii) Is the corresponding space complete? (if not so, indicate the associated completed space.)
    (iii) Is $|||\cdot|||$ a Hilbert norm for some p?
    (iv) Is $|||\cdot|||$ a Hilbertizable norm for some p?

12. * Exhibit a nonseparable Hilbert space.

    Hint: A standard example uses the Cartesian product of a continuous families of copies of the field $\mathbb{K}$...

13. Consider the following classes:

    $\mathcal{B}$: Banach spaces, $\mathcal{H}$: Hilbert spaces,

    $\mathcal{E}$: Euclidean spaces, $\mathcal{P}$: Normed spaces with a predual,

    $\mathcal{F}$: Fréchet spaces, $\mathcal{N}$: Normed spaces, $\mathcal{R}$: Reflexive spaces.

    Which inclusions hold among these spaces?

12 Orthogonality and Projections

The norm provides a distance that is positively homogeneous of degree one and invariant by translation. The inner product allows one to define angles, in particular orthogonality, and then orthogonal projections.

12.1 Orthogonality

Let H be an inner-product space. We shall say that two elements $u, v \in H$ are **orthogonal**, and write $u \perp v$, iff $(u, v) = 0$. More generally, we shall say that two (nonempty) subsets U and V of H are orthogonal, and write $U \perp V$, iff $(u, v) = 0$ for any $u \in U$ and any $v \in V$. We define the **orthogonal complement** of any (nonempty) subset U of H as

$$U^\perp := \{v \in H : (v, u) = 0 \;\; \forall u \in U\};$$

At variance with what occurs in general Banach spaces, in real inner-product spaces one can measure angles: for any unit vectors $u, v \in H$ we define the (nonoriented) angle formed by u and v to be $\arccos(u, v)$.

12.2 Orthogonal projection on a convex set

• **Theorem 12.1** *Let K be a nonempty closed convex subset of a Hilbert space H. For any $u \in H$ there exists one and only one* ***(orthogonal) projection*** *$w \in K$ such that*

$$\|u - w\| = \inf\{\|u - v\| : v \in K\}. \tag{12.1}$$

This condition is equivalent to the ***variational inequality***

$$Re\,(u - w, v - w) \le 0 \qquad \forall v \in K. \tag{12.2}$$

The projection operator $P_K : u \mapsto w$ is ***nonexpansive****, that is,*

$$\|P_K u_1 - P_K u_2\| \le \|u_1 - u_2\| \qquad \forall u_1, u_2 \in H. \tag{12.3}$$

Proof. (i) Let $\{v_n\} \subset K$ be a **minimizing sequence** for the distance from K, that is,

$$d_n := \|u - v_n\| \to \inf\{\|u - v\| : v \in K\} =: d \qquad \text{as } n \to \infty.$$

The parallelogram identity yields

$$\begin{aligned} &2\|u - (v_n + v_m)/2\|^2 + 2\|(v_n - v_m)/2\|^2 \\ &= \|u - v_n\|^2 + \|u - v_m\|^2 = d_n^2 + d_m^2. \end{aligned}$$

As $(v_n + v_m)/2 \in K$ we have $\|u - (v_n + v_m)/2\| \ge d$, so that by the preceding equality

$$2\|(v_n - v_m)/2\|^2 \le d_n^2 + d_m^2 - 2d^2 \to 0.$$

Thus $\{v_n\}$ is a Cauchy sequence in H; by the completeness of H, it converges to some $w \in K$. By the continuity of the norm then $\|u - w\| = \lim_{n\to\infty} \|u - v_n\| = d$, namely (12.1).

(ii) Let w fulfil (12.1). For any $v \in K$ and any $t \in \,]0, 1]$, $w + t(v - w) \in K$ by the convexity of K. Hence

$$\begin{aligned} \|u - w\|^2 &\le \|u - [w + t(v - w)]\|^2 \\ &= \|u - w\|^2 - 2tRe(u - w, v - w) + t^2\|v - w\|^2, \end{aligned}$$

that is, $2t\,Re\,(u - w, v - w) \le t^2\|v - w\|^2$. Dividing by t and passing to the limit as $t \to 0$, we then get (12.2). Conversely,

$$\begin{aligned} \|u - v\|^2 &= \|(u - w) - (v - w)\|^2 \\ &= \|u - w\|^2 + \|v - w\|^2 - 2Re(u - w, v - w) \\ &\overset{(12.2)}{\ge} \|u - w\|^2 + \|v - w\|^2 \ge \|u - w\|^2 \qquad \forall v \in K. \end{aligned}$$

(iii) For any given $u_1, u_2 \in H$, let $w_1, w_2 \in H$ satisfy (12.2) with $u = u_1$ resp. $u = u_2$, and take $v = w_2$ resp. $v = w_1$. Summing the two inequalities we get

$$\|w_1 - w_2\|^2 \le Re\,(u_1 - u_2, w_1 - w_2) \le \|u_1 - u_2\|\,\|w_1 - w_2\|,$$

whence $\|w_1 - w_2\| \leq \|u_1 - u_2\|$, i.e. (12.3). In particular, the choice $u_1 = u_2$ shows that (12.2) defines a unique element $w = P_K u$ for any given $u \in H$. □

Remarks. (i) A geometric interpretation provides a clear understanding of this theorem. For instance, by drawing the intersection of K with the plane that contains u, w and v (assuming that they are distinct and nonaligned), explains why (12.2) characterizes the orthogonal projection, at least if H is a real Hilbert space.

(ii) For any normed space X, we call projection any mapping $P \in \mathcal{L}(X)$ such that $P^2 = P$. Although one cannot define orthogonal projections in normed spaces, it is easily seen that if K is a nonempty closed convex subset of a reflexive Banach space X, then for any $u \in X$ there exists a $w \in K$ such that $\|u - w\| = \inf\{\|u - v\| : v \in K\}$; this is unique if X is uniformly convex.

(iii) Let us denote by $\mathbb{R}^2_\infty$ the space $\mathbb{R}^2$ equipped with the norm $\|(u_1, u_2)\|_\infty = \max\{\|u_1\|, \|u_2\|\}$; this is obviously a non-uniformly-convex reflexive Banach space. In this case the projection exists but need not be unique. Nevertheless $\mathbb{R}^2_\infty$ is a Hilbert space, since the non-Hilbert norm $\|\cdot\|_\infty$ is equivalent to the Euclidean norm $\|\cdot\|_\infty$.

This does not contradict the theorem above, since it is understood that when we deal with a Hilbert space we refer to the Hilbert norm, if not otherwise specified. □

12.3 Orthogonal projection on a subspace

Corollary 12.2 *Let V be a closed subspace of a Hilbert space H. The projection operator P_V is then linear and continuous. Moreover, for any $u \in H$,*

$$w = P_V u \qquad \Leftrightarrow \qquad (w - u, v) = 0 \quad \forall v \in V. \tag{12.4}$$

The latter is called a **variational equation**, and the linear and continuous operator P_V is named an **orthogonal projection,** or just a projection.

Proof. Let us assume that $w = P_V u$. For any $\widetilde{v} \in V$, by selecting $v = w \pm \widetilde{v}$ $(\in V)$ in (12.2), we have $Re(w - u, \widetilde{v}) = 0$ for any $\widetilde{v} \in V$. By taking $\widetilde{v} = v$ and $\widetilde{v} = iv$ for any $v \in V$, we then get $Im(w - u, v) = 0$. Therefore $(w - u, v) = 0$ for any $v \in V$. The converse implication and the linearity of P_V are straightforward. The continuity follows from (12.3). □

Remarks. (i) In Theorem 12.1 the distance from the nonempty closed convex set K is minimized without assuming any compactness property for K.

(ii) Theorem 12.1 rests upon the completeness of the set K, rather than that of H. Therefore this result remains valid in noncomplete inner-product spaces, provided that the convex subset K is complete. In particular the orthogonal projection either on a finite-dimensional linear subspace V, or on a (nonempty) closed convex subset of V, thus exists also in noncomplete inner-product spaces.

(iii) Variational inequalities and variational equations are extensively used in analysis, in particular in convex analysis, in optimization, in the theory of PDEs, and so on. They also have a large number of applications in mathematical physics, in economics, in operational research, and so on.

12.4 Orthogonal decomposition

• **Theorem 12.3** *Let A be a nonempty subset of a Hilbert space H, and M be the closure of its linear span (i.e., $M := \overline{\text{span}}(A)$). Then $A^\perp = M^\perp$ is a subspace of H, and* [41]

$$\bar{M} = \mathcal{R}(P_{\bar{M}}), \qquad M^\perp = \mathcal{N}(P_{\bar{M}}), \tag{12.5}$$

$$u = P_{\bar{M}}(u) + P_{M^\perp}(u) \qquad \forall u \in H, \tag{12.6}$$

$$H = \bar{M} \oplus M^\perp. \tag{12.7}$$

Moreover $(M^\perp)^\perp = M$.

Because of (12.7), $M^\perp$ is called the *orthogonal complement* of M. (Dealing with Hilbert spaces, often we shall refer to it as the complement of M).

Proof. For any sequence $\{u_n\}$ in H, if $u_n \perp A$ for any n and $u_n \to u$, then $u \perp M$, by the continuity of the inner product; thus $A^\perp$ is closed and $A^\perp \subset M^\perp$. The opposite inclusion is trivial.

For any $u \in H$, $u - P_M u \in M^\perp$ by (12.4); thus $u = P_M u + (u - P_M u) \in M + M^\perp$. Moreover, if $u \in M \cap M^\perp$ then $(u, u) = 0$, that is, $u = 0$; (12.6) is thus established, and this yields (12.7).

By applying (12.7) to $M^\perp$ we get $H = M^\perp \oplus (M^\perp)^\perp$. Comparing this equality with (12.7), we conclude that $(M^\perp)^\perp = \bar{M}$. □

12.5 Characterizations of orthogonal projections

Theorem 12.4 *For any closed subspace M of a Hilbert space H, the projection operator P_M is continuous, and*

*(i) P_M is **idempotent**, i.e., $P_M^2 = P_M$;*

*(ii) P_M is **self-adjoint**, i.e., $(P_M u, v) = (u, P_M v)$ for any $u, v \in H$.*

Conversely, any idempotent, self-adjoint, linear operator $P : H \to H$ coincides with the projection on the closed subspace $\mathcal{R}(P)$ (in particular, it is thus continuous).

Proof. The continuity directly follows from the nonexpansiveness (12.3).

For any $u \in H$, $P_M u = P_M u + 0 \in M + M^\perp$, whence $P_M(P_M u) = P_M u$. Thus (i) holds. For any $u, v \in H$, (12.4) yields $(P_M u, v) = (P_M u, P_M v) = (u, P_M v)$, i.e. (ii) is fulfilled.

Let us now assume that $P : H \to H$ is an idempotent, self-adjoint, linear operator. Properties (i) and (ii) and the Cauchy-Schwarz inequality yield

$$\|Pu\|^2 = (Pu, Pu) = (P^2 u, u) = (Pu, u) \le \|Pu\|\,\|u\|,$$

whence $\|Pu\| \le \|u\|$; thus P is continuous. Let us set $M := P(H)$; it suffices to show that $P = P_M$. For any sequence $\{u_n\}$ in the linear subspace M, if $u_n \to u$ then $Pu = \lim Pu_n = \lim u_n \in M$. Thus M is a closed subspace of H. For any $u, v \in H$, as P is self-adjoint we have

$$(u - Pu, Pv) = (Pu - P^2 u, v) = (Pu - Pu, v) = 0.$$

Hence $u - Pu \in M^\perp$; thus $P = P_M$. □

Here is another characterization.

[41] As usual, for any linear mapping L, we denote its range by $\mathcal{R}(L)$ and its null space by $\mathcal{N}(L)$. For any linear subspaces A and B, by "$H = A \oplus B$" we mean that $H = A + B$ and $A \cap B = \{0\}$.

Theorem 12.5 *An operator* $P \in \mathcal{L}(H)$ *is an orthogonal projection iff* $P^2 = P$ *and it is nonexpansive, i.e.,* $\|P\|_{\mathcal{L}(X)} \le 1$. *Equality holds iff* $P \not\equiv 0$. *[]*

Proposition 12.6 *Let* H *be a Hilbert space and* M, N *be two subspaces. Then:*

(i) The composition $P_M P_N$ *is an orthogonal projection iff* P_M *and* P_N *commute, i.e.* $P_M P_N = P_N P_M$. *Either property entails* $P_M P_N = P_{M \cap N}$.

(ii) The sum $P_M + P_N$ *is an orthogonal projection iff* P_M *and* P_N *reciprocally annihilate, i.e.* $P_M P_N = P_N P_M = 0$, *or equivalently* $M \perp N$. *Any of these properties entails* $P_M + P_N = P_{M \oplus N}$.

(iii) $M \subset N$ *iff either* $P_M P_N = P_M$, *or* $P_N P_M = P_M$, *or* $\|P_M x\| \le \|P_N x\|$ *for any* $x \in H$. *[Ex]*

12.6 Overview of projections

So far we have seen three notions of linear projection:

(i) in linear spaces any *idempotent linear* operator from the space to itself is called a projection. The range is a linear subspace. In this setup of course there is no distance to be minimized.

(ii) in Banach spaces one introduces *continuous* projections (or just projections). Their range is a (closed) subspace.

(iii) in Hilbert spaces self-adjoint continuous projections are called *orthogonal* projections (or just projections). This class coincides with that of nonexpansive projections.

In Hilbert spaces one can also deal with orthogonal projections on (nonempty) closed convex subsets: they map any element of the space to the point of the subset that has minimal distance from that element. These projections may be expressed via variational inequalities, whereas projections on closed subspaces are characterized by variational equations.

12.7 Exercises

— Let H be the set of all complex sequences $\{x_n\}$ such that

$$\|\{x_n\}\| := \Big(\sum_{n=1}^{m} \|x_n\|^4\Big)^{1/4} + \Big(\sum_{n=1}^{m} \|x_n\|^2\Big)^{1/2} < +\infty.$$

(i) Show that this is a norm.
(ii) Is it equivalent to Hilbert norm?
(iii) Does H coincide with any of the known sequence spaces?

— Let $\{u_n\}$ be a sequence in a Hilbert space H and $u \in H$ be such that $u_n \to u$ weakly and $\|u_n\| \to \|u\|$. Prove that then $u_n \to u$.
Hint: Develop the square $\|u_n - u\|^2$ ...

— Let K $(\neq \emptyset)$ be a closed convex subset of a reflexive Banach space X.
(i) Show that there exists a (possibly-valued!) projection operator P_K (characterized by the property of minimizing the distance).
(ii) Assuming that P_K is single-valued, need P_K be nonexpansive?
(iii) Do these statements hold in infinite dimension, too?

13 The Representation Theorem

13.1 Riesz-Fréchet Theorem

Via orthogonality next we infer that any Hilbert space is isometric isomorphic to its dual.

• **Theorem 13.1** (*Riesz-Fréchet's Representation Theorem) Let H be a Hilbert space over the field $\mathbb{K}$. The operator*

$$\Gamma : H \to H' \qquad \textit{defined by} \quad \Gamma_v(u) := (u, v) \qquad \forall u, v \in H \tag{13.1}$$

is bijective and isometric. H' is thus a Hilbert space, and is isometrically isomorphic to H.

*If $\mathbb{K} = \mathbb{R}$ the mapping $v \to \Gamma_v$ is linear, whereas if $\mathbb{K} = \mathbb{C}$ it is **antilinear,** i.e.,*

$$\Gamma_{\lambda_1 v_1 + \lambda_2 v_2} = \bar{\lambda}_1 \Gamma_{v_1} + \bar{\lambda}_2 \Gamma_{v_2} \qquad \forall \lambda_1, \lambda_2 \in \mathbb{C}, \forall v_1, v_2 \in H. \tag{13.2}$$

The inverse operator $\Gamma^{-1} : H' \to H$ is often called the **Riesz isomorphism**.

Proof. For any $v \in H$, the Cauchy-Schwarz inequality (11.21) yields

$$|\Gamma_v(u)| = |(u, v)| \le \|u\| \|v\| \qquad \forall u \in H;$$

thus $\Gamma_v \in H'$ and $\|\Gamma_v\|_{H'} \le \|v\|$. As $\|v\|^2 = \Gamma_v(v) \le \|\Gamma_v\|_{H'} \|v\|$, the opposite inequality is also fulfilled. Thus Γ is an isometry.

Let us now fix any $f \in H'$ and show that $f = \Gamma_v$ for some $v \in H$. Obviously, $0 = \Gamma_0$. Let us assume that $f \ne 0$, and choose any $z \in [f^{-1}(0)]^\perp$ such that $f(z) = 1$. For any $u \in H$, $w := u - f(u)z \in f^{-1}(0)$. Hence $(w, z) = 0$, i.e. $(u, z) - f(u)\|z\|^2 = 0$. Setting $v = z/\|z\|^2$ we then get $(u, v) = f(u)$. Thus $R_v = f$ and therefore R_v is onto H'.

The antilinearity of Γ follows from the antilinearity of the inner product w.r.t. to the second argument. □

Remarks. (i) This representation theorem entails that $L^2(A)'$ may be identified with $L^2(A)$, see Theorem 0.7.

(ii) For any $u \in H$ one may also consider the functional $\widetilde{\Gamma}_u : H \to \mathbb{K} : v \mapsto (u, v)$. If $\mathbb{K} = \mathbb{R}$, $\widetilde{\Gamma}_u = \Gamma_u$ and is linear. On the other hand if $\mathbb{K} = \mathbb{C}$, for any $u \in H$, $\widetilde{\Gamma}_u$ is continuous and antilinear:

$$\widetilde{\Gamma}_u(\lambda_1 v_1 + \lambda_2 v_2) = \bar{\lambda}_1 \widetilde{\Gamma}_u(v_1) + \bar{\lambda}_2 \widetilde{\Gamma}_u(v_2) \qquad \forall \lambda_1, \lambda_2 \in \mathbb{C}, \forall u, v_1, v_2 \in H.$$

Thus the map $u \mapsto \widetilde{\Gamma}_u$ is linear, but it maps H to its **antidual** $\widetilde{H}$, namely the linear space of continuous and antilinear functionals $H \to \mathbb{C}$. □

Proposition 13.2 *Any Hilbert space H is reflexive.*

Proof. We shall show that the canonical embedding $J : H \to H''$ is surjective. Let us first define the antilinear operator $\Gamma : H \to H'$ as in (13.1). For any $u'' \in H''$, the functional $v \mapsto f(v) := \overline{\langle u'', \Gamma_v \rangle}$ is an element of H'. By Theorem 13.1, then $f = \Gamma_u$ for some $u \in H$. Denoting by J the canonical isomorphism $H \to H''$, we then have

$$\langle u'', \Gamma_v \rangle = \overline{f(v)} = \overline{\Gamma_u(v)} = \overline{(v, u)} = (u, v) = \Gamma_v(u) = \langle J(u), \Gamma_v \rangle \qquad \forall v \in H.$$

As Γ is surjective, Γ_v is any element of H'. We thus have $\langle u'', g \rangle = \langle J(u), g \rangle$ for any $g \in H'$, that is, $u'' = J(u)$. As u'' was arbitrary, we conclude that J is surjective. □

13.2 The Lax-Milgram Theorem

• **Theorem 13.3** (*Lax-Milgram*) *Let H be a Hilbert space, and $L \in \mathcal{L}(H)$ be such that, for some $\alpha > 0$,*

$$(Lv, v) \geq \alpha\|v\|^2 \qquad \forall v \in H \quad \text{(coerciveness)}. \tag{13.3}$$

Then L is bijective and $L^{-1} \in \mathcal{L}(H)$; more precisely,

$$\|L^{-1}w\| \leq \alpha^{-1}\|w\| \qquad \forall w \in H. \tag{13.4}$$

Proof. By the continuity and the coerciveness of L,

$$\alpha\|v\|^2 \leq (Lv, v) \leq \|Lv\|\,\|v\| \quad \text{whence} \quad \alpha\|v\| \leq \|Lv\| \qquad \forall v \in H.$$

This entails that:

(i) L is injective;

(ii) if L^{-1} exists then (13.4) is fulfilled;

(iii) Any sequence $\{v_n\}$ in H is Cauchy if so is $\{Lv_n\}$.

To conclude the proof it then suffices to show that L is surjective. By (iii) and by the continuity of L, $L(H)$ is closed. For any $v \in L(H)^\perp$ we have $\alpha\|v\|^2 \leq (Lv, v) = 0$, whence $v = 0$; thus $L(H)^\perp = \{0\}$. As $H = \overline{L(H)} \oplus L(H)^\perp = L(H) \oplus \{0\}$, we conclude that $L(H) = H$. □

Remark. If L is self-adjoint, that is, if $(Lu, v) = (u, Lv)$ for any $u, v \in H$, then the thesis of the Lax-Milgram theorem directly follows from the Riesz-Fréchet representation Theorem 13.1.

We check this for the case of a *real* Hilbert space, for the sake of simplicity. In that case, $(u, v) \mapsto ((u, v)) := (u, Lv)$ defines an inner product over H. Since L is continuous and coercive, the corresponding norm is equivalent to the original one, hence the dual H' is the same for either choice of inner product.

To prove that L is surjective, let $b \in H$ define $f \in H'$ by setting $f(v) = (v, b)$ for any v. By the representation theorem, there exists $u \in H$ with $f(v) = ((v, u)) = (v, Lu)$ for any $v \in H$; thus $Lu = b$ and consequently L is surjective. Above we already derived the injectivity L and (13.4) from (13.3). □

The Lax-Milgram theorem is widely used to derive existence and uniqueness results for linear boundary value problems, written as an equation $Au = b$ between suitably chosen function spaces.

13.3 Exercises

1. = Prove the following extension of the classical Pythagoras's theorem to Hilbert spaces. For any finite orthogonal system $\{u_n\}_{n=1,\ldots,m}$ of an inner-product space H,

$$\left\|\sum_{n=1}^{m} u_n\right\|^2 = \sum_{n=1}^{m} \|u_n\|^2.$$

For $m = 2$ (and only in this case), conversely this formula holds only if u_1 and u_2 are orthogonal.

2. Prove directly the Hahn-Banach Theorem I.4.1 in a Hilbert space, assuming that M is a closed subspace, without using the Zorn Lemma (neither any equivalent statement).

3. Let M be a closed subspace of a Hilbert space H, X be a Banach space, and $L : M \to X$ be a linear and continuous operator. Prove that L has a linear and continuous extension to the whole H.

   This may be regarded as a sort of Hahn-Banach-type Theorem for operators. Notice that this property may fail if H is just a Banach space. Consider for instance $M = X = c_0$, $H = \ell^\infty$ and L equal to the identity operator.

4. Let M be a linear subspace of a Hilbert space H. Prove that M is dense in H iff $M^\perp = \{0\}$.

5. Under which assumptions is $M^\perp = [(M^\perp)^\perp]^\perp$ in a Hilbert space?

6. Let H be the linear space of functions $f : \mathbb{R} \to \mathbb{R}$ such that $\{x \in \mathbb{R} : f(x) \neq 0\}$ is at most countable. Is this a Hilbert space w.r.t. the inner product $(f, g) := \sum_{x\in\mathbb{R}} f(x)g(x)$?

7. * Prove that, if H is a Hilbert space and $L \in \mathcal{L}(H)$ is symmetric, that is,

$$(Lu, v) = \overline{(u, Lv)} \; \big(= (Lv, u)\big) \qquad \forall u, v \in H,$$

   then the thesis of the Lax-Milgram theorem follows from the Riesz-Fréchet representation Theorem 13.1.

   Hint: The mapping $(u, v) \mapsto ((u, v)) := (u, Lv)$ defines an inner product over H. Notice that the dual H' is the same for the original and this newly defined inner product.

   Since L is continuous and coercive, the corresponding norm is equivalent to the original one, hence the dual H' is the same for either choice of inner product.

   To prove that L is surjective, let $b \in H$ define $f \in H'$ be setting $f(v) = (v, b)$ for any v. By the representation theorem, there exists $u \in H$ with $f(v) = ((v, u)) = (v, Lu)$ for any $v \in H$; thus $Lu = b$ and consequently L is surjective. That L is injective, and that (13.4) holds, follows as above by virtue of the inequality $\alpha\|v\| \leq \|Lv\|$, valid for any $v \in H$.

14 Orthonormal Systems and Hilbert Bases

14.1 Orthonormal Systems

A subset $A \neq \emptyset$ of an inner-product space H is called an **orthogonal system** iff $(u, v) = 0$ for any two distinct elements $u, v \in A$; A is said to be **orthonormal** iff moreover $\|u\| = 1$ for any $u \in A$. (The origin 0 may thus belong to orthogonal but not to orthonormal systems.)

Proposition 14.1 (*Gram-Schmidt Orthonormalization***)**
Let $\{u_n\}$ be an either finite or countable linearly independent subset of an inner-product space H, and set $v_1 = u_1/\|u_1\|$. For any integer $n > 1$, by induction let us assume that $v_1, ..., v_n$ are known, and set

$$\begin{aligned} w_{n+1} &:= u_{n+1} - \sum_{j=1}^{n} (u_{n+1}, v_j) v_j, \\ v_{n+1} &:= w_{n+1}/\|w_{n+1}\| \qquad \forall n \geq 1. \end{aligned} \tag{14.1}$$

This entails that $\{v_n\}$ is an orthonormal subset of H, and

$$\text{span}\{v_1, ..., v_n\} = \text{span}\{u_1, ..., u_n\} \qquad \forall n.$$

Proof. For any $n \in \mathbb{N}$, let us denote by V_n the span of $\{u_1, ..., u_n\}$, and define w_{n+1} and v_{n+1} as above. By induction hypothesis, let us assume that V_n coincides with the span of $\{v_1, ..., v_n\}$. Notice that $w_{n+1} \neq 0$ as $u_{n+1} \notin V_n$. For any n, by construction v_{n+1} is of unit norm and is orthogonal to $v_1, ..., v_n$. The sequence $\{u_n\}$ is thus orthonormal, and V_{n+1} coincides with the span of $\{v_1, ..., v_{n+1}\}$. □

In passing, notice that $\sum_{j=1}^{n}(u_{n+1}, v_j)v_j$ coincides with the projection of u_{n+1} on the span V_n of $\{v_1, ..., v_n\}$.

For instance, the Gram-Schmidt procedure transforms the set of monomials $\{f_n(x) := x^n : n \in \mathbb{N} \cup \{0\}\}$ to an orthonormal system of $L^2(-1,1)$, more specifically the classical family of Legendre polynomials:

$$\{P_n(x) = (2^n n!)^{-1}(d/dx)^n[(x^2-1)^n],\ \forall x \in [-1,1]\}. \quad []$$

The next result illustrates the relevance of orthogonal sequences in Hilbert spaces. [42]

Theorem 14.2 *For any orthogonal sequence $\{u_n\}_{n\in\mathbb{N}}$ in a Hilbert space H, the following properties are mutually equivalent:*

$$\textstyle\sum_{n=1}^{\infty} u_n \quad \textit{converges unconditionally (in } H\textit{)}, \tag{14.2}$$

$$\textstyle\sum_{n=1}^{\infty} u_n \quad \textit{converges weakly unconditionally (in } H\textit{)}, \tag{14.3}$$

$$\textstyle\sum_{n=1}^{\infty} \|u_n\|^2 \quad \textit{converges (in } \mathbb{R}\textit{)}. \tag{14.4}$$

Proof. (14.2) ⇒ (14.3): this is obvious.

Let us show that (14.3) ⇒ (14.4). By (14.3) the sequence of the partial sums $\{\sum_{n=1}^{m} u_n\}_{m\in\mathbb{N}}$ is bounded. By the orthogonality of the sequence $\{u_n\}$, then

$$\sum_{n=1}^{m} \|u_n\|^2 = \left\| \sum_{n=1}^{m} u_n \right\|^2 \leq \text{Constant (independent of } m\text{)}.$$

This yields (14.4).

Let us next prove that (14.4) ⇒ (14.2). By (14.4) the sequence of partial sums $\{\sum_{n=1}^{m} \|u_n\|^2\}_{m\in\mathbb{N}}$ is Cauchy in $\mathbb{R}$, and by the orthogonality of $\{u_n\}$

$$\left\| \sum_{n=m}^{\ell} u_n \right\|^2 = \sum_{n=m}^{\ell} \|u_n\|^2 \qquad \forall \ell, m \in \mathbb{N}, m < \ell.$$

Hence the sequence of partial sums $\{\sum_{n=1}^{m} u_n\}_{m\in\mathbb{N}}$ is Cauchy in H. By the completeness of H, the series $\sum_{n=1}^{\infty} u_n$ then converges. As the convergence in (14.4) is unconditional, the same then holds for $\sum_{n=1}^{\infty} u_n$. □

14.2 Bessel inequality

For any (nonempty) subset A of a normed space X, we denote by $\overline{\text{span}}\,(A)$ the closure of the set of finite linear combinations of elements of A. For any Hilbert space H,

$$\overline{\text{span}}\,(A) = (A^{\perp})^{\perp} \qquad \forall A \subset H.[Ex] \tag{14.5}$$

The next statement rests upon finite-dimensional geometry, and is essentially a reformulation of the classical Pythagoras theorem in Hilbert spaces.

[42] In chapter XXX we called *unconditionally convergent* a series $\sum_{n=1}^{\infty} u_n$ in a Banach space iff $\sum_{n=1}^{\infty} u_n = \sum_{n=1}^{\infty} u_{\pi(n)}$ for all permutations $\pi : \mathbb{N} \to \mathbb{N}$.

Lemma 14.3 *Let $\{u_n\}_{1\le n\le m}$ be a finite orthonormal subset of an inner-product space H. Then for any $u \in H$*

$$\big\|u - \textstyle\sum_{n=1}^m (u,u_n)u_n\big\|^2 = \|u\|^2 - \sum_{n=1}^m |(u,u_n)|^2 = \|u\|^2 - \big\|\sum_{n=1}^m (u,u_n)u_n\big\|^2. \tag{14.6}$$

Moreover $\sum_{n=1}^m (u,u_n)u_n$ coincides with the orthogonal projection of u on $\mathrm{span}\,(\{u_n\}_{1\le n\le m})$.

Proof. Let us fix any $u \in H$, and set $\alpha_n := (u,u_n)$ for any $n \in \mathbb{N}$. By the orthonormality of $\{u_n\}$, for any $m \in \mathbb{N}$ we have

$$\big\|\textstyle\sum_{n=1}^m \alpha_n u_n\big\|^2 = \sum_{n=1}^m \|\alpha_n u_n\|^2 = \sum_{n=1}^m |\alpha_n|^2 \qquad \forall \ell > m, \tag{14.7}$$

$$\big(u, \textstyle\sum_{n=1}^m \alpha_n u_n\big) = \sum_{n=1}^m \|\alpha_n u_n\|^2 = \sum_{n=1}^m |\alpha_n|^2. \tag{14.8}$$

Hence

$$\begin{aligned}\big\|u - \textstyle\sum_{n=1}^m \alpha_n u_n\big\|^2 &= \|u\|^2 - 2\mathrm{Re}\big(u, \textstyle\sum_{n=1}^m \alpha_n u_n\big) + \big\|\textstyle\sum_{n=1}^m \alpha_n u_n\big\|^2 \\ &\overset{(14.7),(14.8)}{=} \|u\|^2 - 2\textstyle\sum_{n=1}^m |\alpha_n|^2 + \sum_{n=1}^m |\alpha_n|^2 \\ &= \|u\|^2 - \textstyle\sum_{n=1}^m |\alpha_n|^2 \qquad \forall m \in \mathbb{N}.\end{aligned} \tag{14.9}$$

For any m, let us define the partial sum $s_m = \sum_{n=1}^m \alpha_n u_n$, and notice that

$$\|s_\ell - s_m\|^2 = \big\|\textstyle\sum_{n=m+1}^\ell \alpha_n u_n\big\|^2 = \sum_{n=m+1}^\ell |\alpha_n|^2 \qquad \forall \ell, m \in \mathbb{N}, m < \ell; \tag{14.10}$$

by the completeness of H, $\widetilde{u} := \sum_{n=1}^\infty \alpha_n u_n$ then converges.
(14.7) and (14.9) also yield

$$\big\|u - \textstyle\sum_{n=1}^m \alpha_n u_n\big\|^2 = \|u\|^2 - \big\|\textstyle\sum_{n=1}^m \alpha_n u_n\big\|^2 \qquad \forall m \in \mathbb{N}\,. \tag{14.11}$$

A straightforward calculation shows that $(u - \widetilde{u}, \widetilde{u}) = 0$, and this yields the final statement. □

Next we extend Lemma 14.3 to infinite orthonormal subsets in Hilbert spaces.

Proposition 14.4 *Let $\{u_n\}_{n\in\mathbb{N}}$ be an orthonormal sequence in a Hilbert space H. Then for any $u \in H$, the series $\sum_{n=1}^\infty (u,u_n)u_n$ converges and*

$$\left\|u - \sum_{n=1}^\infty (u,u_n)u_n\right\|^2 = \|u\|^2 - \sum_{n=1}^\infty |(u,u_n)|^2 = \|u\|^2 - \left\|\sum_{n=1}^\infty (u,u_n)u_n\right\|^2. \tag{14.12}$$

Hence

$$\|u\|^2 \ge \sum_{n=1}^\infty |(u,u_n)|^2 \qquad \forall u \in H (\textbf{\textit{Bessel inequality}}). \tag{14.13}$$

Moreover $\sum_{n=1}^\infty (u,u_n)u_n$ coincides with the orthogonal projection of u on the subspace $\overline{span}\,(\{u_n\}_{n\in\mathbb{N}})$.

Proof. By the completeness of H, the series $\sum_{n=1}^\infty (u,u_n)u_n$ converges. Passing to the limit as $m \to \infty$ in (14.6), we then get (14.12). □

14.3 Hilbert bases

An orthonormal subset A of a Hilbert space H is called a **Hilbert basis** (or an orthonormal basis) iff $H = \overline{\mathrm{span}}\,(A)$. In this case A cannot be extended to any larger orthonormal subset of H; one then says that the orthonormal subset A is **complete.** [43] Notice that a Hilbert basis may be finite, countable or also uncountable.

[43] This should not be confused with the completeness of the space.

Theorem 14.5 *Any Hilbert space has a Hilbert basis.*

Outline of the Proof. The argument may be based on the Zorn lemma or also the Hausdorff's maximal chain principle. [44] [Ex] □

Proposition 14.6 *Let J be an index set and $A := \{u_j : j \in J\}$ be an orthonormal subset of a Hilbert space H. The following properties are mutually equivalent:*
(i) A is a Hilbert basis;
(ii) A is maximal (w.r.t. the ordering by inclusion) among orthonormal subsets;
(iii) 0 is the only element of H orthogonal to A. [Ex]

Proposition 14.7 *Let J be an index set and $A := \{u_j : j \in J\}$ be a Hilbert basis of a Hilbert space H. For any $u \in H$, $u = \sum_{j\in J}(u, u_j)u_j$ unconditionally, and $J_u := \{j \in J : (u, u_j) \neq 0\}$ is at most countable.* [45] *[]*

14.4 Fourier coefficients

We saw that, whenever $A := \{u_j : j \in J\}$ is an orthonormal subset of a Hilbert space H, $u = \sum_{j\in J}(u, u_j)u_j$ unconditionally. The (u, u_j)s are called the (generalized) **Fourier coefficients** of u w.r.t. the orthonormal subset $\{u_j\}$. This denomination arose from and refers to the Fourier analysis of periodic functions of a real variable, that we shall briefly illustrate ahead.

Next we provide more precise information about these coefficients.

Theorem 14.8 *If $\{u_j\}_{j\in J}$ is a Hilbert basis of a Hilbert space H, then*

$$u = \sum_{j\in J}(u, u_j)u_j \qquad \forall u \in H \ \textbf{\textit{(Fourier expansion)}}; \tag{14.14}$$

$$(u, v) = \sum_{j\in J}(u, u_j)\overline{(v, u_j)} \qquad \forall u, v \in H \ \textbf{\textit{(Parseval identity)}}; \tag{14.15}$$

$$\|u\|^2 = \sum_{j\in J} |(u, u_j)|^2 \qquad \forall u \in H \ \textbf{\textit{(Parseval formula)}}. \tag{14.16}$$

If J is infinite countable, H is thus isometrically isomorphic to ℓ^2.

Proof. For any $u \in H$, by the Bessel inequality (14.12), $\sum_{\ell\in L} |(u, u_\ell)|^2 < +\infty$ for any finite set $L \subset J$. As the set $J_u := \{j \in J : (u, u_j) \neq 0\}$ is at most countable, (14.14) then follows from the final part of Proposition 14.4.

For any finite set $L \subset J$, we have

$$\left(\sum_{j\in J}(u, u_j)u_j, \sum_{\ell\in L}(v, u_\ell)u_\ell\right) = \sum_{j\in J}(u, u_j)\overline{(v, u_j)} \qquad \forall u, v \in H;$$

(14.14) thus yields (14.15). Selecting $v = u$ in (14.15) we get (14.16). □

Remark. We have just see that if a Hilbert space has a countable Hilbert basis (equivalently, if it is separable) then it is isometrically isomorphic to ℓ^2. We can then rephrase Megginson's remark that we stated after Proposition 8.6:

[44] This states that in any partially ordered set, every totally ordered subset is contained in a maximal totally ordered subset. This is equivalent to the Zorn lemma as well as to the axiom of choice.

[45] ... Refer to infinite sums in the Banach chapter.

In a sense the entire theory of normed spaces is contained in the theory of ℓ^2. This by no means trivializes the theory of normed spaces, but rather serves to point out the richness of the theory of the spaces ℓ^2.

14.5 Hilbert dimension

*** Proposition 14.9** *Any two Hilbert bases of the same Hilbert space have the same cardinality (this is called the* ***Hilbert dimension****).*

Proof. If one of the two bases is finite, the other one is also finite, and the result is straightforward. Let us then assume that $\{u_j\}_{j\in J}$ and $\{v_\ell\}_{\ell\in L}$ are two infinite bases of a Hilbert space. For any $\ell \in L$, as we saw the set $J_\ell := \{j \in J : (v_\ell, u_j) \neq 0\}$ is at most countable.

Let us denote by card(A) the cardinality of any set A. Any $j \in J$ is element of some J_ℓ, as $\{v_\ell\}_{\ell\in L}$ is a Hilbert basis; hence $\mathrm{card}\,(J) \le \mathrm{card}\,(\bigcup_{\ell\in L} J_\ell)$. Moreover $\mathrm{card}\,(\bigcup_{\ell\in L} J_\ell) \le \mathrm{card}\,(L)$, as each J_ℓ is at most countable and card$\,(L)$ is infinite. Thus $\mathrm{card}\,(J) \le \mathrm{card}\,(L)$. By the symmetry of the argument, we infer that this is an equality. □

A countable Hilbert basis is a Schauder basis as well, and the space is then separable. Conversely, any separable Hilbert space has a countable Hilbert basis. [Ex] (On the other hand, as we already remarked, a separable Banach spaces may not have any Schauder basis.)

14.6 Fourier series in L^2

For any $n \in \mathbb{N}$, let us set $(e_n)_n := 1$ and $(e_n)_m := 0$ if $m \neq n$. The sequence of unit vectors $\{e_n\}$ is the canonical Hilbert basis of ℓ^2.

Another example, which is at the basis of the terminology we just introduced, is provided by reformulating the theory of Fourier series in the setup of Hilbert spaces.

Proposition 14.10 *Let us set*

$$u_k(x) = e^{ikx}/\sqrt{2\pi} \qquad \text{for a.e. } x \in\,]-\pi,\pi[, \forall k \in \mathbb{Z}. \tag{14.17}$$

The family $\{u_k\}_{k\in\mathbb{Z}}$ is a Hilbert basis of $L^2(-\pi,\pi;\mathbb{C})$.

Proof. It is straightforward to check that $\{u_k\}_{k\in\mathbb{Z}}$ is an orthonormal system in H. By an obvious bijection between $\mathbb{Z}$ and $\mathbb{N}$, this orthonormal system indexed by $k \in \mathbb{Z}$ may be transformed to one indexed by $n \in \mathbb{N}$. By the classical Stone-Weierstrass theorem, $\{u_k\}_{k\in\mathbb{Z}}$ is dense in $C^0([-\pi,\pi];\mathbb{C})$. [] By the density of the canonic injection $C^0([-\pi,\pi]) \to L^2(-\pi,\pi)$, this family is also dense in the latter space, too. □

The following formulas define the transform $L^2(-\pi,\pi) \to \ell^2 : f \mapsto \{\hat{f}_k\}_{k\in\mathbb{N}}$ and its inverse:

$$\hat{f}_k = (f, u_k) \quad \forall k \in \mathbb{Z}, \qquad f(x) = \sum_{k\in\mathbb{Z}} \hat{f}_k u_k(x) \quad \text{for a.e. } x \in\,]-\pi,\pi[, \tag{14.18}$$

the convergence of the latter series being understood in the sense of $L^2(-\pi,\pi)$. More explicitly, these two formulas read [46]

$$\begin{aligned} &\hat{f}_k = \frac{1}{\sqrt{2\pi}} \int_{-\pi}^{\pi} f(x) e^{-ikx}\, dx \qquad \forall k \in \mathbb{Z}, \\ &\lim_{m\to\infty} \int_{-\pi}^{\pi} \Big| f(x) - \sum_{k=-m}^{k=m} \hat{f}_k u_k(x) \Big|^2 dx = 0. \end{aligned} \tag{14.19}$$

The operator $L^2(-\pi,\pi) \to \ell^2 : f \mapsto \{\hat{f}_k\}$ is also called the *Fourier series transform in* L^2.

14.7 Overview of bases

So far we have introduced three types of bases:

(i) Hamel bases for linear spaces: they exist for any space;

(ii) Schauder bases for separable Banach spaces: for some spaces they do not exist;

(iii) Hilbert bases for Hilbert spaces: they exist for any space.

A Schauder basis of a separable Hilbert space is a Hilbert basis iff it is orthonormal. In this case the Schauder basis is unconditional. [Ex]

14.8 Overview of Hilbert spaces

We introduced the axioms of the inner product, and derived some basic properties, in particular the Cauchy-Schwarz inequality and the parallelogram identity. By means of the Cauchy-Schwarz inequality, we showed that a norm can be associated with any inner product. Thus any Hilbert space is also a Banach space. Conversely, an inner product is associated with any norm which fulfills the parallelogram identity.

By means of the inner product, we defined the concepts of orthogonality and of orthogonal projection. The completeness entails the existence of the orthogonal projection on any (nonempty) closed convex subset, in particular on closed subspaces. Orthogonal projections are characterized as idempotent, self-adjoint, linear operators of the space to itself. Orthogonal projections also provide a surjective isometric isomorphism between the antidual of any Hilbert space and the space itself (Riesz-Fréchet representation theorem).

We then dealt with orthonormal systems of a Hilbert space and derived the Bessel inequality. We defined Hilbert bases, and saw that any Hilbert space is endowed with such a basis, and thus has a Hilbert dimension. Finally, we derived the Fourier expansion of any element of a Hilbert space w.r.t. to a Hilbert basis.

14.9 Exercises

1. = Prove the following extension of the classical Pythagoras's theorem to Hilbert spaces. For any finite orthogonal system $\{u_n\}_{n=1,\dots,m}$ of an inner-product space H,

$$\Big\| \sum_{n=1}^{m} u_n \Big\|^2 = \sum_{n=1}^{m} \|u_n\|^2.$$

For $m = 2$ (and only in this case), conversely this formula holds only if u_1 and u_2 are orthogonal.

[46] A priori the L^2-convergence entails the a.e. convergence just for a suitable subsequence. In 1966 Carleson was able to prove the convergence of the whole sequence: a highly nontrivial result! On this basis in 2006 he was awarded the prestigious Abel prize.

2. Prove directly the Hahn-Banach Theorem I.4.1 in a Hilbert space, assuming that M is a closed subspace, without using the Zorn Lemma (neither any equivalent statement).

3. Let M be a closed subspace of a Hilbert space H, X be a Banach space, and $L : M \to X$ be a linear and continuous operator. Prove that L has a linear and continuous extension to the whole H.

   This may be regarded as a sort of Hahn-Banach-type Theorem for operators. Notice that this property may fail if H is just a Banach space. Consider for instance $M = X = c_0$, $H = \ell^\infty$ and L equal to the identity operator.

4. Let M be a linear subspace of a Hilbert space H. Prove that M is dense in H iff $M^\perp = \{0\}$.

5. Under which assumptions is $M^\perp = [(M^\perp)^\perp]^\perp$ in a Hilbert space?

6. Let H be the linear space of functions $f : \mathbb{R} \to \mathbb{R}$ such that $\{x \in \mathbb{R} : f(x) \neq 0\}$ is at most countable. Is this a Hilbert space w.r.t. the inner product $(f, g) := \sum_{x\in\mathbb{R}} f(x)g(x)$?

7. * Prove that, if H is a Hilbert space and $L \in \mathcal{L}(H)$ is symmetric, that is,

$$(Lu, v) = \overline{(u, Lv)} \ \big(= (Lv, u)\big) \qquad \forall u, v \in H,$$

   then the thesis of the Lax-Milgram theorem follows from the Riesz-Fréchet representation Theorem 13.1.

   Hint: The mapping $(u, v) \mapsto ((u, v)) := (u, Lv)$ defines an inner product over H. Notice that the dual H' is the same for the original and this newly defined inner product.

   Since L is continuous and coercive, the corresponding norm is equivalent to the original one, hence the dual H' is the same for either choice of inner product.

   To prove that L is surjective, let $b \in H$ define $f \in H'$ be setting $f(v) = (v, b)$ for any v. By the representation theorem, there exists $u \in H$ with $f(v) = ((v, u)) = (v, Lu)$ for any $v \in H$; thus $Lu = b$ and consequently L is surjective. That L is injective, and that (13.4) holds, follows as above by virtue of the inequality $\alpha\|v\| \leq \|Lv\|$, valid for any $v \in H$.

Chapter III — Operators

Contents: 1. Bounded linear operators. 2. Compact operators. 3. The Riesz and Fredholm theory. 4. Introduction to spectral theory.

15 Bounded linear operators

15.1 Examples

(i) For any matrix $A \in \mathbb{K}^{M,N}$, the associated linear mapping $L : \mathbb{K}^N \to \mathbb{K}^M$

$$(Lu)_j = \sum_{k=1}^{N} a_{jk} u_k, \quad 1 \leq j \leq M$$

defines a bounded linear operator.

(ii) In some cases an **infinite matrix** $A = (a_{jk})$ defines a bounded linear mapping between sequence spaces by the formula

$$(Lu)_j = \sum_{k=1}^{\infty} a_{jk} u_k, \quad 1 \leq j < \infty. \tag{15.1}$$

For example, the estimate

$$\sum_{j=1}^{\infty} \left(\sum_{k=1}^{\infty} |a_{jk} u_k| \right)^2 \leq \left(\sum_{j,k=1}^{\infty} |a_{jk}|^2 \right) \sum_{k=1}^{\infty} |u_k|^2 \quad [Ex] \tag{15.2}$$

entails that $L \in \mathcal{L}(\ell^2)$, with $\|L\|^2 \leq \sum_{j,k} |a_{jk}|^2$, if the latter sum is finite. This condition, however, is not necessary: e.g., the unit matrix does not satisfy it. Indeed, in the diagonal case $(Lu)_k = \alpha_k u_k$, we have $L \in \mathcal{L}(\ell^2)$ iff $\|\alpha\|_\infty < \infty$. One may ask for conditions, in terms of the elements of A, which are necessary as well as sufficient in order that (15.1) defines a bounded linear mapping from ℓ^p to ℓ^q. However, no "useful" conditions are known for $1 < p, q < \infty$.

(iii) The **right** (or **forward**) **shift** S_r and the **left** (or **backward**) **shift** S_l

$$(S_r u)_k = u_{k-1}, \quad (S_\ell u)_k = u_{k+1}, \tag{15.3}$$

are most naturally defined on bilateral sequences $\{u_k\}_{k\in\mathbb{Z}}$; obviously they are isometries on $\ell^p_{\mathbb{K}}(\mathbb{Z})$ for any $p \in [1, \infty]$. For unilateral sequences $u = (u_1, u_2, \dots)$ one sets

$$S_r(u_1, u_2, \dots) = (0, u_1, u_2, \dots), \quad S_\ell(u_1, u_2, \dots) = (u_2, u_3, \dots).$$

In this case S_r and S_ℓ still belong to $\mathcal{L}(X)$ for $X = \ell^p$ $(:= \ell^p_{\mathbb{K}}(\mathbb{N})$, but they are no longer isomorphisms.

(iv) For any $p \in [1, \infty]$, if a is a bounded measurable function on a measure space $(A, \mathcal{A}, \mu)$, then the **multiplication operator** define by

$$(Lu)(x) = a(x)u(x) \qquad \text{for a.e. } x \in A$$

is an operator $L \in \mathcal{L}(L^p(A))$ and $\|L\| = \|a\|_\infty$. Similarly, if A is a compact metric space and $a \in C^0(A)$, then $L \in \mathcal{L}(L^p(A))$ and $\|L\| = \max_A |a|$.

(v) Let $(A, \mathcal{A}, \mu)$ and $(B, \mathcal{B}, \nu)$ be two σ-finite measure spaces, $k \in L^2(A \times B)$, and set

$$(Lu)(x) = \int_B k(x,y)u(y)\, d\mu(y) \qquad \text{for a.e. } x \in A, \forall u \in L^2(B). \tag{15.4}$$

By the theorems of Tonelli and Fubini and the Hölder inequality, Lu is an a.e. well-defined and measurable function, and, by the Cauchy-Schwarz inequality,

$$\int_A \left| \int_B k(x,y)u(y)\, d\mu(y) \right|^2 d\nu(x) \leq \int_A \int_B |k(x,y)|^2\, d\mu(y)\, d\nu(x) \cdot \int_B |u(y)|^2\, d\mu(y).$$

L is thus a bounded linear mapping $L^2(B) \to L^2(A)$, and

$$\|L\| \leq \iint_{A \times B} |k(x,y)|^2\, d\mu(y)\, d\nu(x).$$

The function k is called the **kernel** of the **integral operator** L.

If $A = B = [a,b]$ and $\mu = \nu$ is the Lebesgue measure, then the operators

$$(L_1 u)(x) = \int_a^b k(x,y)u(y)\, dy, \quad (L_2 u)(x) = \int_a^x k(x,y)u(y)\, dy \quad \forall x \in [a,b]$$

are respectively called **Fredholm** and **Volterra integral operators**.

15.2 Adjoints in Hilbert spaces

Let H_1, H_2 be Hilbert spaces over $\mathbb{K}$ ($= \mathbb{R}$ or $\mathbb{C}$). For any $L \in \mathcal{L}(H_1; H_2)$, the **(Hilbert) adjoint** $L^* \in \mathcal{L}(H_2; H_1)$ is defined by

$$(L^* u, v)_{H_1} = (u, Lv)_{H_2}, \qquad \forall\, u \in H_2 \quad \forall\, v \in H_1\,. \tag{15.5}$$

We state several properties that are easily derived via the procedure that we used for Banach spaces. We leave it to the reader to verify these statements.

L^* is well defined and

$$\begin{aligned} &(L_1 + L_2)^* = L_1^* + L_2^*, \qquad (\alpha L)^* = \overline{\alpha} L^*, \\ &(L_2 \circ L_1)^* = L_1^* \circ L_2^*, \qquad L^{**} = L, \qquad \|L^*\| = \|L\|, \end{aligned} \tag{15.6}$$

for all operators L, L_1, L_2 with appropriate domain and range, and all scalars α.

Let us define the canonical isomorphism $\Gamma_i : H_i \to H_i'$ $(i = 1, 2)$ as in (13.1). The Hilbert adjoint $L^* : H_2 \to H_1$ is related to the adjoint $L' : H_2' \to H_1'$ by

$$L^* = \Gamma_1^{-1} L' \Gamma_2, \qquad L' = \Gamma_1 L^* \Gamma_2^{-1}. \tag{15.7}$$

In the complex case some formulae for L^* may slightly differ from those for L', since $(\alpha L)' = \alpha L'$ but $(\alpha L)^* = \overline{\alpha} L^*$ for scalars $\alpha \in \mathbb{C}$.

An operator $L \in \mathcal{L}(H)$ is invertible iff L^* is invertible, and $(L^*)^{-1} = (L^{-1})^*$ in that case. A computation analogous to (10.3) shows that $\|L^*\| = \|L\|$. Moreover,

$$\|L^* L\| = \|L\|^2 \qquad \forall L \in \mathcal{L}(H). \tag{15.8}$$

Indeed, the inequality $\|L^* L\| \leq \|L\|^2$ is obvious. The converse inequality follows from the estimate

$$\|Lu\|^2 = (Lu, Lu) = (u, L^* Lu) \leq \|u\|.$$

15.3 Exercises

1. Let X, Y be Banach spaces and $T \in \mathcal{L}(X, Y)$.

   (i) Is it true that the range of T is dense iff T^* is injective?
   * (ii) Is it true that T is surjective iff T^* is injective?

2. * Let X, Y be Banach spaces and $T \in \mathcal{L}(X, Y)$. Show that T is injective iff there exists $C > 0$ such that $\|Tu\| \geq C\|u\|$ for any $u \in X$.

16 Compact Operators

Throughout this section X and Y will denote Banach spaces over $\mathbb{K}$.

An operator $K : X \to Y$ is called **compact** (or **completely continuous**) iff it maps any bounded subset of X to a relatively compact subset of Y, or equivalently iff $K(B_X)$ (the image of the unit ball) is relatively compact in Y. We shall denote the set of all compact operators by $\mathcal{K}(X;Y)$, or $\mathcal{K}(X)$ if $X = Y$. It is easy to see that $K : X \to Y$ is compact iff for every bounded sequence $\{u_n\}$ in X, the sequence $\{Ku_n\}$ has a convergent subsequence (in Y). This entails that any compact operator maps weakly (or weakly star) convergent sequences to convergent sequences.

Proposition 16.1 *The composition $M \circ L$ of two linear continuous operators is compact if either M or L is compact. (The converse fails.)*

Proof. This holds since continuous operators map bounded sets to bounded sets, and relatively compact sets to relatively compact sets. □

Proposition 16.2 *$\mathcal{K}(X;Y)$ is a closed subspace of the Banach space $\mathcal{L}(X;Y)$.*

Proof. One readily checks that $\mathcal{K}(X;Y)$ is a linear subspace. Thus it suffices to prove that the uniform limit of compact operators is a compact operator.

Let $L \in \mathcal{L}(X;Y)$ and $\{K_m\}$ be a sequence in $\mathcal{K}(X;Y)$ such that $K_m \to L$. For any m, the image $K_m(B_X)$ of the unit ball is relatively compact, hence totally bounded. [47] For any $\varepsilon > 0$ and any m, $K_m(B_X)$ may thus be covered by the union of finitely many balls of radius $\varepsilon/2$. As $\|K_m - L\|_{\mathcal{L}(X;Y)} \leq \varepsilon/2$ for m large enough, $L(B_X)$ may then be covered by the union of finitely many balls of radius ε. This means that $L(B_X)$ is totally bounded. Since Y is complete, this set is thus relatively compact. □

Theorem 16.3 (*Schauder) An operator $L \in \mathcal{L}(X;Y)$ is compact iff its adjoint L' is compact. (If X, Y are Hilbert spaces, then the same holds for the Hilbert adjoint L^*.)*

* *Proof.* If L is compact, then $A = \overline{L(B_X)}$ is compact. Let $\{g_n\}$ be a sequence in Y' with $\|g_n\| \leq 1$. Since $|g_n(v) - g_n(w)| \leq \|v - w\|$ for all $v, w \in A$, the sequence $\{g_n|_A\}$ is bounded and equicontinuous in $C^0(A)$. By the Ascoli-Arzelà theorem, some subsequence $\{g_{n_k}\}$ then converges uniformly on A, and thus is Cauchy w.r.t. the uniform norm. As

$$\|L'g_{n_k} - L'g_{n_j}\|_{X'} = \sup_{x \in B_X} \langle g_{n_k} - g_{n_j}, Lx \rangle \leq \|g_{n_k} - g_{n_j}\|_{C^0(A)} \|L\|_{\mathcal{L}(X,Y)} \quad \forall k, j,$$

the subsequence $\{L'g_{n_k}\}$ is Cauchy and hence convergent in X'. Thus L' is compact.

[47] By definition, a subset of a metric space is totally bounded iff for any $\varepsilon > 0$ it can be covered by a union of finitely many balls of radius ε.

If conversely L' is compact, then so is $L^{**} : X'' \to Y''$. Since $J_Y(L(B_X)) = L^{**}(J_X(B_X))$, the set $J_Y(L(B_X))$ is relatively compact in Y''. As J_Y is an isometry we conclude that $L(B_X)$ is relatively compact in Y. □

16.1 Examples

(i) If either X or Y has finite dimension, then every $L \in \mathcal{L}(X;Y)$ is compact.

(ii) The shift operators S_r and S_ℓ are not compact on ℓ^p, since $S_\ell \circ S_r = I$ is not compact.

(iii) Let A be a Euclidean set with nonempty interior and $X = C^0_b(A)$. The multiplication operator

$$(L_a u)(x) = a(x)u(x) \qquad \text{for a given } a \in C^0_b(A),$$

is compact only if $a \equiv 0$ identically in A. [Ex] An analogous result holds for $X = L^p(A)$.

(iv) Depending on the properties of the kernel k, the integral operator

$$(Lu)(x) = \int_A k(x,y)u(y)\,d\nu(y) \tag{16.1}$$

not only belongs to $\mathcal{L}(X;Y)$ for suitable function spaces X and Y, but actually is compact. E.g., let $X = L^2(A_1;\mu)$, $Y = L^2(A_2;\nu)$ for two σ-finite measures μ and ν, and let $k \in L^2(A_1 \times A_2; \mu \otimes \nu)$.

If k has the product form $k(x,y) = g(x)h(y)$ with $g \in L^2(A_2)$ and $h \in L^2(A_1)$, then

$$(Lu)(x) = g(x) \int_{A_1} h(y)u(y)\,d\nu(y) \qquad \text{for } \mu\text{-a.e. } x \in A_2.$$

Thus $\mathcal{R}(L)$ equals the one-dimensional subspace spanned by g; in this case L is thus compact.

For the general case, it turns out that for any $\varepsilon > 0$ there is a $k_\varepsilon \in L^2(A_1 \times A_2)$, which is a finite linear combination of kernels in product form and satisfies $\|k - k_\varepsilon\|_{L^2(A_1 \times A_2)} \le \varepsilon$. [Ex] Let us denote by L_ε the integral operator associated to k_ε. The range $\mathcal{R}(L_\varepsilon)$ has finite dimension, and satisfies $\|L - L_\varepsilon\| \le \|k - k_\varepsilon\|_{L^2(A_1 \times A_2)} \le \varepsilon$. As $\mathcal{K}(X;Y)$ is a closed subspace of $\mathcal{L}(X;Y)$, L is thus compact.

(v) As another example of operator of the form (16.1), let us assume that $X = C^0(A_1)$ and $Y = C^0(A_2)$, with A_1 and A_2 compact subsets of $\mathbb{R}^N$ and $\mathbb{R}^M$ (resp.) equipped with the Lebesgue measure, and let $k \in C^0(A_1 \times A_2)$. Let us denote by B_1 the unit ball of $C^0(A_1)$. Since k is uniformly continuous, $L(B_1)$ is not only bounded but also equicontinuous, hence relatively compact in $C^0(A_2)$ by the Ascoli-Arzelà theorem; so L is compact in this case, too.

16.2 Exercises

1. Let X, Y be Banach spaces, and let a linear operator $L : X \to Y$ map the unit ball B_X to a relatively compact subset of Y. Is then L continuous?

2. Let us fix any $f \in C^0_b(\mathbb{R}^2)$ and set $[T(v)](x) = \int_0^x f(x,y)v(y)\,dy$ for any x.

   (i) Is T a bounded operator $C^0([0,1]) \to C^1([0,1])$?
   (ii) Is T a compact operator $C^0([0,1]) \to C^1([0,1])$?
   (iii) Is T a compact operator $C^0([0,1]) \to C^0([0,1])$?
   (iv) Is T a bounded operator $C^0_b(\mathbb{R}) \to C^0_b(\mathbb{R})$?

3. Let us fix any $k \in C^0_b(\mathbb{R}^2)$ and set
   $[T(v)](x) := x^{-1} \int_0^x k(x,y)v(y)\,dy$ for any $x \neq 0$, $\quad [T(v)](0) := k(0,0)v(0)$.

(This is the most natural way of defining $[T(v)](0)$: why?)

(i) Is T a bounded operator $C^0([0,1]) \to C^0([0,1])$?
(ii) Is T a compact operator $C^0([0,1]) \to C^1([0,1])$?
(iii) Is T a compact operator $C^0_b(\mathbb{R}) \to C^0_b(\mathbb{R})$?

4. Let us fix any $k \in L^\infty(\mathbb{R}^2)$ and set
$[T(v)](x) := x^{-1} \int_0^x k(x,y)v(y)\,dy$ for any $x \neq 0$, $\quad [T(v)](0) := k(0,0)v(0)$.

   (i) Is T a bounded operator $L^1(0,1) \to L^1(0,1)$?
   * (ii) Is T a compact operator $L^1(0,1) \to L^1(0,1)$?

5. Let X, Y, Z be Banach spaces and $L \in \mathcal{L}(X,Y), M \in \mathcal{L}(Y,Z)$. Show that if either of these operators is compact, then their composition ML is also compact.

6. * Let X, Y be Banach spaces and $L \in \mathcal{L}(X,Y)$ be compact and bijective.

   (i) Give an example.
   (ii) In which cases is L^{-1} compact?

7. Are the left and right shift operators compact in the spaces ℓ^p $(1 \leq p \leq +\infty)$?

8. Are the inclusions among the spaces ℓ^p $(1 \leq p \leq +\infty)$ compact?

9. Are the inclusions among the spaces $L^p(0,1)$ $(1 \leq p \leq +\infty)$ compact?

10. Are the canonical injections $C^{k+1}([0,1]) \to C^k([0,1])$ $(k \in \mathbb{N})$ compact?

11. Is the canonical injection $C^1_b(\mathbb{R}) \to C^0_b(\mathbb{R})$ compact?

12. Is the canonical injection $C^0([0,1]) \to L^p(0,1)$ compact for some $1 \leq p < +\infty$?

13. Is the canonical injection $C^0([0,1]) \to L^\infty(0,1)$ compact?

14. Let $1 \leq p < +\infty$. Are the canonical injections $\ell^p \to c_0$, $\ell^p \to c$, $\ell^p \to \ell^\infty$ compact?

15. * Let X, Y be two Banach spaces, $L \in \mathcal{L}(X,Y)$, and denote by B_X the unit ball of X.

    (i) Show that if X is reflexive, then $L(B_X)$ is closed.
    (ii) Show that if X is reflexive and L is compact, then $L(B_X)$ is compact.
    (iii) Check that if $X = Y = C^0([0,1])$ and $(Lu)(x) = \int_0^x u(t)\,dt$ for any x, then $L(B_X)$ is not closed.

    (This exercise has been taken from [Brezis] p. 171.)

16. Let X, Y be two Banach spaces, with X of infinite dimension, and let $L \in \mathcal{L}(X,Y)$ be compact. Show that there exists a sequence $\{u_n\}$ in X such that $\|u_n\| = 1$ for any n and $Lu_n \to 0$ in Y.

    (This exercise has been taken from [Brezis] p. 171.)

17 The Riesz and Fredholm theory

We say that a linear operator $L : X \to Y$ has **finite rank** iff its range $\mathcal{R}(L)$ has finite dimension. [48] It is promptly seen that any finite-rank operator acting between Banach spaces is compact.

[48] Some vocabulary:
English *range* (of a function) $\leftrightarrow$ Italian *insieme immagine* (di una funzione).
English *rank* (of a function) = dimension of its range $\leftrightarrow$ Italian *rango* (di una funzione) = dimensione del suo insieme immagine.

A celebrated theorem of F. Riesz extends some known properties of linear operators in Euclidean spaces (thus of matrices) to compact perturbations of linear operators in Banach spaces. For finite-rank operators this result is quite simple; first we illustrate it in that case, as a preliminary step towards the Riesz theorem.

Proposition 17.1 *If X is a linear space and $L : X \to X$ is a linear mapping of finite rank, then*

$$(i)\ \mathcal{N}(I-L) \text{ has finite dimension,} \tag{17.1}$$

$$(ii)\ \mathcal{R}(I-L) \text{ has finite codimension.} \tag{17.2}$$

Proof. Notice that

$$\mathcal{N}(I-L) \subset \mathcal{R}(L), \qquad \mathcal{N}(L) \subset \mathcal{R}(I-L). \tag{17.3}$$

Thus (17.1) holds. By the second inclusion of (17.3), $\mathrm{codim}(\mathcal{R}(I-L)) \le \mathrm{codim}(\mathcal{N}(L))$. By Proposition 2.7, $\mathrm{codim}(\mathcal{N}(L)) = \dim(\mathcal{R}(L))$. (17.2) then follows. □

$\mathcal{R}(I-L))$ is obviously closed, since it is a finite-dimensional linear spaces.

For any linear operator A from a finite dimensional space to itself, the codimension of the kernel $\mathcal{N}(A)$ coincides with the rank of the associated matrix, and this is invariant by transposition. This entails that $\mathcal{N}(I-L)$ and $\mathcal{N}(I-L')$ have the same dimension.

The following classical result extends the above properties to compact perturbations of the identity (more generally, of any linear isomorphism) in a Banach space X. [49]

Theorem 17.2 (*Riesz) Let X be a Banach space and $K : X \to X$ be a compact operator. Then*

$$(i)\ \mathcal{N}(I-K) \text{ has finite dimension,} \tag{17.4}$$

$$(ii)\ \mathcal{R}(I-K) \text{ is closed and has finite codimension,} \tag{17.5}$$

$$(iii)\ \mathrm{codim}\,(\mathcal{R}(I-K)) = \dim(\mathcal{N}(I-K)) = \dim(\mathcal{N}(I-K')), \tag{17.6}$$

$$(iv)\ \mathcal{N}(I-K) = \{0\} \ \Leftrightarrow\ \mathcal{R}(I-K) = X. \tag{17.7}$$

We shall not prove this deep result, and just interpret it. Anyway, part of the thesis is easily checked. The restriction of K to $\mathcal{N}(I-K)$ coincides with the identity and is compact; hence $\mathcal{N}(I-K)$ has finite dimension. Part (iv) directly follows from part (iii), see below.

Remark. The Riesz theorem 17.2 applies to all operators of the form $A - K$, for any linear isomorphism $A : X \to X$ and any compact operator $K : X \to X$. Actually, $A^{-1}K$ is also compact, so the theorem holds for $I - A^{-1}K$. Therefore it also holds for $A(I - A^{-1}K) = A - K$. □

Corollary 17.3 *Let X be a Banach space and $K : X \to X$ be a compact operator. The thesis of the Riesz Theorem 17.2 then holds also for K'. Moreover,*

$$\mathcal{R}(I-K) = {}^0\mathcal{N}(I-K'), \tag{17.8}$$

and $\mathcal{R}(I-K') = [\mathcal{N}(I-K)]^0$ if X is reflexive.

[49] This is one of the milestones of functional analysis.

Proof. The first statement follows from the Riesz theorem, since by the Schauder theorem the operator $K' : X' \to X'$ is also compact.

(17.8) stems from the closedness of $\mathcal{R}(I-K)$ and from (10.14). The final statement follows from (17.8). □

17.1 The Fredholm Alternative

The statement (17.7) expresses the Fredholm alternative:

$$I - K \text{ is injective iff it is surjective.}$$

This is a basic result of linear algebra, namely of the theory of linear operators in finite-dimensional spaces. By the Riesz theorem, here we extend to compact perturbations of the identity in infinite-dimensional spaces.

Let us consider an equation of the form

$$\text{for a prescribed } b \in Y, \text{ find } u \in X \text{ such that} \quad u - Ku = b. \tag{17.9}$$

Corollary 17.4 (*Fredholm alternative) Let X be a Banach space and $K : X \to X$ be a compact operator. Then*

either (i) *$u - Ku = b$ has a unique solution $u \in X$ for each $b \in X$,*

or (ii) *the homogeneous equation $u - Ku = 0$ has a nontrivial solution $u \in X$, and there exists $b \in X$ such that $u - Ku = b$ has no solution.*

In the second case, the inhomogeneous equation $u - Ku = b$ is solvable iff $f(b) = 0$ for all solutions f of the homogeneous adjoint equation $(I - K')f = 0$.

Proof. The dichotomy between the cases (i) and (ii) directly follows from (17.7).
The final statement of this corollary stems from (17.8). □

Remarks. (i) The two cases above respectively correspond to

$$(i)\ \mathcal{N}(I-K) = \{0\} \text{ and } \mathcal{R}(I-K) = X, \tag{17.10}$$

$$(ii)\ 1 \le \dim(\mathcal{N}(I-K)) = \operatorname{codim}(\mathcal{R}(I-K)) = \dim(\mathcal{N}(I-K')) < +\infty. \tag{17.11}$$

(ii) $\mathcal{R}(I-K)$ may be characterized as the space of the elements of X that fulfill a finite number of linearly independent linear equations; these are often regarded as *constraints.* The equality $\dim(\mathcal{N}(I-K)) = \operatorname{codim}(\mathcal{R}(I-K))$ thus means that the (finite) number of linearly independent solutions $u \in X$ of $u - Ku = 0$ equals the number of linearly independent constraints that define $\mathcal{R}(I-K)$.

(iii) The equality $\dim(\mathcal{N}(I-K)) = \dim(\mathcal{N}(I-K'))$ obviously means that the (finite) number of linearly independent solutions $u \in X$ of $u - Ku = 0$ equals the number of linearly independent solutions $f \in X'$ of $f - K'f = 0$.

(iv) We already pointed out that the Riesz Theorem 17.2 applies to all operators of the form $A - K$, for any linear isomorphism $A : X \to X$ and any compact operator $K : X \to X$. The same then applies to the Fredholm alternative.

(v) If K is a compact operator is a Hilbert space H, the Riesz Theorem 17.2 may be reformulated as follows, in terms of the Hilbert adjoint K^*:

$$\dim(\mathcal{N}(I-K)) = \dim(\mathcal{N}(I-K^*)) < +\infty, \tag{17.12}$$

$$H = \mathcal{N}(I-K) \oplus \mathcal{R}(I-K^*).\square \tag{17.13}$$