3 Basic Functional Analysis

Functional Analysis emerged as a coherent field of mathematics in the first four decades of the 20th century. It provided a unified framework to treat different objects using abstraction and axiomatization. The main idea is to view functions as points, respectively elements, of an abstract space endowed with certain structures that are axiomatically defined. This way mathematicians were able to "escape" from the usual finite dimensional Euclidean spaces and consider infinite dimensional function spaces. The starting point was the thesis of Fréchet in 1906 who introduced the abstract notion of "metric space" – a concept that was influential in the development of both functional analysis and point set topology. The work of Fréchet was the culmination of the efforts and contributions of many prominent mathematicians from France, Germany, and Italy. Combined with the revolution of measure theory this provided a fertile ground for the development of functional analysis. The prominent figure in the story is that of the Polish mathematician Stefan Banach (1892–1945).

In this chapter, we review the basic notions and results of "Linear Functional Analysis." Moreover, we touch on "Operator Theory" and in particular, we discuss the spectral properties of compact self-adjoint operators on a Hilbert space.

3.1 Topological Vector Spaces, Hahn–Banach Theorem

We start with the basic notion of a topological vector space. Recall that a **vector space** or **linear space** is a set X equipped with two operations $+ : X \times X \to X$ defined by $(x, u) \to x + u$ called the vector addition and $\cdot : \mathbb{K} \times X \to X$ defined by $(\lambda, x) \to \lambda \cdot x$ called the scalar multiplication where $\mathbb{K} = \mathbb{R}$ or $\mathbb{K} = \mathbb{C}$.

Definition 3.1.1. A **topological vector space** is a vector space endowed with a Hausdorff topology τ, which makes the two vector space operations above continuous. Then we say that τ is a vector topology on X.

Remark 3.1.2. Continuity of vector addition means that if $x, u \in X$ and $V \in \tau$ is a neighborhood of $x + u$, that is, $V \in \mathcal{N}(x + u)$, then there exist $U_x \in \mathcal{N}(x)$ and $U_u \in \mathcal{N}(u)$ such that $U_x + U_u \subseteq V$. Similarly the continuity of the scalar multiplication implies that if $(\lambda, x) \in \mathbb{K} \times X$ with $\mathbb{K} = \mathbb{R}$ or $\mathbb{K} = \mathbb{C}$ and $V \in \mathcal{N}(\lambda x)$, then there exist $\varepsilon > 0$ and $U_x \in \mathcal{N}(x)$ such that $\mu U_x \subseteq V$ for all $|\mu - \lambda| < \varepsilon$. Moreover, for a given $x \in X$ and a given $\lambda \in \mathbb{K}$ we introduce

$$\hat{T}_x(u) = x + u \quad \text{for all } u \in X \qquad \text{(the translation operator)}\,,$$
$$\hat{M}_\lambda(u) = \lambda u \quad \text{for all } u \in X \qquad \text{(the scalar multiplication operator)}\,.$$

Clearly, these operators are homeomorphisms of X onto X. It follows that the vector topology τ is **translation invariant**, that is, $U \in \tau$ if and only if $x + U \in \tau$ for all $x \in X$.

https://doi.org/10.1515/9783110532982-003

Hence, τ is completely determined by any local basis, in particular by the local basis at the origin. If the vector topology is induced by a metric d, then the metric is invariant, that is, $d(x+v, u+v) = d(x,u)$ for all $x, u, v \in X$.

An immediate consequence of these observations is the following simple lemma.

Lemma 3.1.3. *Let (X, τ) be a topological vector space.*
(a) *For all $U, V \in \tau$ and for all $\lambda \in \mathbb{K}$ it follows $U + V \in \tau$ and $\lambda U \in \tau$.*
(b) *If $A \subseteq X$ and $U \in \tau$, then $\overline{A} + U = A + U$ and it is open.*
(c) *If $K \subseteq X$ is compact and $C \subseteq X$ is closed, then $K + C \subseteq X$ is closed.*
(d) *If $K_1, K_2 \subseteq X$ are compact sets, then $K_1 + K_2 \subseteq X$ is compact.*
(e) *If $\varphi\colon X \to \mathbb{R}$ is linear, then φ is continuous if and only if φ is continuous at $x = 0$.*

Proof. (a), (b), and (e) are clear. (c) Let $\{v_\alpha\}_{\alpha\in I} \subseteq K + C$ be a net such that $v_\alpha \to v$. We have that $v_\alpha = x_\alpha + u_\alpha$ with $x_\alpha \in K$ and $v_\alpha \in C$ for all $\alpha \in I$. The compactness of K implies that there exists a subnet $\{x_\beta\}_{\beta\in J}$ of $\{x_\alpha\}_{\alpha\in I}$ such that $x_\beta \to x \in K$; see Proposition 1.4.45(c). Then $u_\beta = v_\beta - x_\beta \to v - x = u \in C$ since C is closed; see Proposition 1.2.36. Therefore $v = x + u$ with $x \in K$ and $u \in C$. Hence, we conclude that $K + C$ is closed.

(d) Since "+" is continuous on $X \times X$ and $K_1 \times K_2 \subseteq X \times X$ is compact (see Theorem 1.4.56) we conclude that $+(K_1 \times K_2) = K_1 + K_2 \subseteq X$ is compact; see Theorem 1.4.51. □

Remark 3.1.4. The algebraic sum of two closed sets need not be closed. In $\mathbb{R}^2$ equipped with the usual Euclidean metric, we consider the sets

$$C_1 = \left\{\left(x, \frac{1}{x}\right) : x \in \mathbb{R} \setminus \{0\}\right\} \quad \text{and} \quad C_2 = \{(u, 0) : u \in \mathbb{R}\}\,.$$

Then both are closed in $\mathbb{R}^2$ but $C_1 + C_2 = \{(x+u, 1/x) : x \in \mathbb{R} \setminus \{0\}, u \in \mathbb{R}\}$ is not closed in $\mathbb{R} \times \mathbb{R}$.

Remark 3.1.5. Let X, Y be two vector spaces. Recall that a map $A\colon X \to Y$ is called a linear function if it is additive and homogeneous, that is,

$$\begin{aligned} &A(x+y) = A(x) + A(y) \quad \text{for all } x, y \in X\,,\\ &A(\lambda x) = \lambda A(x) \quad \text{for all } \lambda \in \mathbb{K} \text{ and for all } x \in X\,. \end{aligned}$$

By $N(A)$ we denote the kernel of A, that is, $N(A) = \{x \in X : A(x) = 0\}$ and by $R(A)$ the range of A, that is, $R(A) = \{A(x) : x \in X\}$.

Now we introduce certain classes of sets that are important in the study of topological vector spaces.

Definition 3.1.6. Let X be a vector space and $A \subseteq X$.
(a) We say that A is **convex** if for all $x, u \in A$ and $\lambda \in [0,1]$, it holds $(1-\lambda)x + \lambda u \in A$.
(b) We say that A is **absorbing** if for any $x \in X$ there is $t = t(x) > 0$ such that $x \in tA$. So every absorbing set contains the origin.

(c) We say that A is **balanced** if $\lambda A \subseteq A$ for all $\lambda \in \mathbb{K}$ with $|\lambda| \leq 1$.
(d) We say that A is **symmetric** if $A = -A$.

Lemma 3.1.7. *If (X, τ) is a topological vector space and $V \in \mathcal{N}(0)$, then there exists a symmetric set $U \in \mathcal{N}(0)$ such that $U + U \subseteq V$.*

Proof. The continuity of the vector addition operation implies that there exist $U_1, U_2 \in \mathcal{N}(0)$ such that $U_1 + U_2 \subseteq V$. Let $U = U_1 \cap (-U_1) \cap U_2 \cap (-U_2)$. Then $U \in \mathcal{N}(0)$ is symmetric and $U + U \subseteq V$. □

Proposition 3.1.8. *If (X, τ) is a topological vector space, $K \subseteq X$ is compact, $C \subseteq X$ is closed, and $K \cap C = \emptyset$, then there exists $U \in \mathcal{N}(0)$ such that $(K + U) \cap (C + U) = \emptyset$.*

Proof. We assume that $K \neq \emptyset$ or otherwise the result is obvious. Let $x \in K$. Applying Lemma 3.1.7 there is a symmetric $U_x \in \mathcal{N}(0)$ such that $(x + U_x + U_x + U_x) \cap C = \emptyset$. Exploiting the symmetry of U_x it follows that $(x + U_x + U_x) \cap (C + U_x) = \emptyset$. The compactness of K implies that there exist $\{x_n\}_{n=1}^m \subseteq K$ such that $K \subseteq \bigcup_{n=1}^m (x_n + U_{x_n})$. Let $U = \bigcap_{n=1}^m U_{x_n} \in \mathcal{N}(0)$. Then

$$K + U \subseteq \bigcup_{n=1}^{m} (x_n + U_{x_n} + U) \subseteq \bigcup_{n=1}^{m} (x_n + U_{x_n} + U_{x_n}) .$$

We conclude that $(K + U) \cap (C + U) = \emptyset$. □

Note that $K + U$ is an open set containing K and $C + U$ is an open set containing C; see Lemma 3.1.3(b). Taking K to be a singleton we obtain the following result.

Corollary 3.1.9. *Every topological vector space is regular; see Definition* 1.2.7.

Proposition 3.1.10. *Let (X, τ) be a topological vector space.*
(a) *If $A \subseteq X$, then $\overline{A} = \bigcap_{U \in \mathcal{N}(0)} (A + U)$.*
(b) *If $A, C \subseteq X$, then $\overline{A} + \overline{C} \subseteq \overline{A + C}$.*
(c) *If $A \subseteq X$ is convex, then $\operatorname{int} A$ and $\overline{A}$ are convex.*
(d) *If $A \subseteq X$ is balanced, then $\overline{A}$ is balanced and when $0 \in \operatorname{int} A$, then $\operatorname{int} A$ is balanced.*

Proof. (a) We know that $x \in \overline{A}$ if and only if $(x + U) \cap A \neq \emptyset$ for all $U \in \mathcal{N}(0)$. Hence, $x \in \overline{A}$ if and only if $x \in A - U$ for every $U \in \mathcal{N}(0)$. But $U \in \mathcal{N}(0)$ if and only if $-U \in \mathcal{N}(0)$.

(b) Let $x \in \overline{A}$, $u \in \overline{C}$ and let $V \in \mathcal{N}(x + u)$. Then there exist $V_x \in \mathcal{N}(x)$, $V_u \in \mathcal{N}(u)$ such that $V_x + V_u \subseteq V$. Then choose $x' \in A \cap V_x$ and $u' \in C \cap V_u$. The existence follows since $x \in \overline{A}$ and $u \in \overline{C}$. Then $x' + u' \in (A + C) \cap V$. Since $V \in \mathcal{N}(x + u)$ we conclude that $x + u \in \overline{A + C}$, thus $\overline{A} + \overline{C} \subseteq \overline{A + C}$.

(c) Since $\operatorname{int} A \subseteq A$ and A is convex, it follows that

$$(1 - \lambda) \operatorname{int} A + \lambda \operatorname{int} A \subseteq A \quad \text{for all } \lambda \in (0, 1) . \tag{3.1.1}$$

Note that the left-hand side in (3.1.1) is an open set and so

$$(1 - \lambda) \operatorname{int} A + \lambda \operatorname{int} A \subseteq \operatorname{int} A \quad \text{for all } \lambda \in (0, 1) .$$

Hence, $\operatorname{int} A$ is convex. For $\lambda \in (0,1)$, due to part (b) and since A is convex, one gets

$$(1-\lambda)\overline{A} + \lambda\overline{A} = \overline{(1-\lambda)A} + \overline{\lambda A} \subseteq \overline{(1-\lambda)A + \lambda A} \subseteq \overline{A}\,.$$

Therefore $\overline{A}$ is convex.

(d) The proof that $\overline{A}$ is balanced is similar to the proof of part (c).

Let $\lambda \in \mathbb{K}$ be such that $0 < |\lambda| \le 1$. Since A is balanced, we derive $\lambda \operatorname{int} A = \operatorname{int} \lambda A \subseteq \lambda A \subseteq A$, which shows that $\lambda \operatorname{int} A \subseteq A$. Moreover, since $0 \in \operatorname{int} A$, for $\lambda = 0$, it follows that $\lambda \operatorname{int} A \subseteq \operatorname{int} A$ and so $\operatorname{int} A$ is balanced. □

This leads to the following structural result for the topology of X.

Proposition 3.1.11. *Let (X,τ) be a topological vector space.*
(a) *Every $V \in \mathcal{N}(0)$ contains a balanced $U \in \mathcal{N}(0)$.*
(b) *Every convex $V \in \mathcal{N}(0)$ contains a balanced convex $U \in \mathcal{N}(0)$.*

Proof. (a) Let $V \in \mathcal{N}(0)$. Exploiting the continuity of the scalar multiplication operation, there exist $\delta > 0$ and $\tilde{U} \in \mathcal{N}(0)$ such that $\lambda\tilde{U} \subseteq V$ for all $\lambda \in \mathbb{K}$ with $|\lambda| < \delta$. Let U be the union of all these sets $\lambda\tilde{U}$. Evidently, $U \in \mathcal{N}(0)$, U is balanced and $U \subseteq V$.

(b) Let $V \in \mathcal{N}(0)$ be convex. Let $A = \bigcap_{|\lambda|=1} \lambda V$. Applying part (a), let $\hat{U} \in \mathcal{N}(0)$ be balanced such that $\hat{U} \subseteq V$. We have $\lambda^{-1}\hat{U} = \hat{U}$ for all $\lambda \in \mathbb{K}$ with $|\lambda| = 1$. Hence $\hat{U} \subseteq \lambda V$ and thus $\hat{U} \subseteq A$. This means that $\hat{U} \subseteq \operatorname{int} A \in \mathcal{N}(0)$. Moreover, $\operatorname{int} A \subseteq V$. The set A is convex, being the intersection of convex sets. Hence, $\operatorname{int} A$ is convex; see Proposition 3.1.10(c). We claim that $\operatorname{int} A$ is balanced. According to Proposition 3.1.10(d) it suffices to show that A is balanced. To this end, let $t \in [0,1]$ and $\mu \in \mathbb{K}$ with $|\mu| = 1$. Then, since $\lambda V \in \mathcal{N}(0)$ is convex,

$$t\mu A = \bigcap_{|\lambda|=1} t\mu\lambda V = \bigcap_{|\lambda|=1} t\lambda V \subseteq \bigcap_{|\lambda|=1} \lambda V\,.$$

Therefore, $t\mu A \subseteq A$ and so A is balanced. We conclude that $U = \operatorname{int} A \in \mathcal{N}(0)$ is the desired balanced and convex neighborhood of the origin. □

Corollary 3.1.12. *Every topological vector space has a local basis consisting of balanced sets.*

We introduce some particular types of topological vector spaces depending on the structure of the local basis.

Definition 3.1.13. Let (X,τ) be a topological vector space.
(a) A set $A \subseteq X$ is said to be **bounded** if for every $U \in \mathcal{N}(0)$ there is a $t_U > 0$ such that $A \subseteq tU$ for all $t > t_U$.
(b) We say that X is **locally convex** if it has a local basis $\mathcal{B}$ consisting of convex sets.
(c) We say that X is **locally bounded** if it has a bounded set in $\mathcal{N}(0)$.
(d) We say that X is **Fréchet** if it is locally convex and the topology τ is induced by a complete translation invariant metric d.
(e) A **norm** on X is a real function $\|\cdot\|$ such that

(e)$_1$ $\|x\| \geq 0$ for all $x \in X$ and $\|x\| = 0$ if and only if $x = 0$;
(e)$_2$ $\|\lambda x\| = |\lambda|\|x\|$ for all $(\lambda, x) \in \mathbb{K} \times X$;
(e)$_3$ $\|x + u\| \leq \|x\| + \|u\|$ for all $x, u \in X$, which is called triangle inequality.
X equipped with a norm is called a **normed space**. The norm defines a translation invariant metric $d(x, u) = \|x - u\|$. If (X, d) is complete, then X is a **Banach space**.

(f) We say that X is **normable** if τ is generated by the metric induced by a norm.

Remark 3.1.14. If X is locally bounded, then it is first countable. Indeed, if $U \in \mathcal{N}(0)$ is bounded and $r_n \to 0^+$, then $\{r_n U\}_{n\in\mathbb{N}}$ is a local basis for the origin.

Finite dimensional vector spaces exhibit some distinguishing properties. The **Euclidean norm** on X being finite dimensional with $\dim X = n$ is defined by

$$\|x\|_2 = \left(\sum_{k=1}^{n} |x_k|^2\right)^{\frac{1}{2}} \quad \text{for all } x = (x_k)_{k=1}^n \in X .$$

The topology on X induced by $\|\cdot\|_2$ is known as the **Euclidean topology**. It turns out that the Euclidean space is the prototype of a n-dimensional vector space.

Definition 3.1.15. Let X be a vector space and let $\|\cdot\|, |\cdot|$ be two norms on X. We say that these norms are **equivalent** if there exist constants $\eta > m > 0$ such that

$$m\|x\| \leq |x| \leq \eta\|x\| \quad \text{for all } x \in X .$$

Remark 3.1.16. Equivalence of norms is an equivalence relation and equivalent norms generate the same topology on X.

Proposition 3.1.17. *In a finite dimensional vector space any two norms are equivalent.*

Proof. Let X be the n-dimensional vector space with norm $\|\cdot\|$ and consider $\mathbb{R}^n$ equipped with the norm $\|\cdot\|_2$. Let $\{e_k\}_{k=1}^n \subseteq X$ be a basis for X and consider the linear map $A \colon \mathbb{R}^n \to X$ defined by

$$A(\lambda) = \sum_{k=1}^{n} \lambda_k e_k \quad \text{for all } \lambda = (\lambda_k)_{k=1}^n \in \mathbb{R}^n .$$

It is easy to see that A is an isomorphism. Moreover, we obtain the estimate

$$\|A(\lambda)\| \leq \sum_{k=1}^{n} |\lambda_k|\|e_k\| \leq \left(\sum_{k=1}^{n} |\lambda_k|^2\right)^{\frac{1}{2}} \left(\sum_{k=1}^{n} \|e_k\|^2\right)^{\frac{1}{2}} \leq \eta\|\lambda\|_2 \tag{3.1.2}$$

with $\eta = \left(\sum_{k=1}^n \|e_k\|^2\right)^{1/2}$. Therefore, A is continuous.

In addition, let $\xi = \|\cdot\| \circ A \colon \mathbb{R}^N \to \mathbb{R}$, that is,

$$\xi(\lambda) = \left\|\sum_{k=1}^{n} \lambda_k e_k\right\| \quad \text{for all } \lambda = (\lambda_k)_{k=1}^n \in \mathbb{R}^N . \tag{3.1.3}$$

Of course, ξ is continuous. Moreover, $\partial B_1 = \{\lambda \in \mathbb{R}^N : \|\lambda\|_2 = 1\}$ is closed and bounded, and thus compact; see Theorem 1.5.38. Hence, there exists $\lambda^* \in \partial B_1$ such that

$$\xi(\lambda^*) = \inf_{\lambda \in \partial B_1} \xi(\lambda) = m \geq 0\,;$$

see Theorem 1.4.52. If $m = 0$, then $\left\|\sum_{k=1}^n \lambda_k^* e_k\right\| = 0$ (see (3.1.3)), a contradiction since $\lambda^* \in \partial B_1$. Hence, $m > 0$ and we get

$$m\|\lambda\|_2 \leq \|A(\lambda)\| \quad \text{for all } \lambda \in \mathbb{R}^n\,. \tag{3.1.4}$$

From (3.1.2) and (3.1.4) we infer that X and $\mathbb{R}^n$ are linearly homeomorphic and so we conclude that any two norms on X are equivalent. □

Corollary 3.1.18. *Every finite dimensional normed space is complete, thus a Banach space.*

Corollary 3.1.19. *Every finite dimensional subspace of a normed space is closed.*

Next we will give a characterization of finite dimensional normed spaces in terms of the topological properties of the closed unit ball $\overline{B}_1 = \{x \in X : \|x\| \leq 1\}$. First we need an auxiliary result known as the "Riesz Lemma."

Lemma 3.1.20 (Riesz Lemma). *If X is a normed space, $Y \subseteq X$ is a proper, closed vector subspace, and $0 < \vartheta < 1$, then there exists $x_\vartheta \in (X \setminus Y) \cap \partial B_1$ such that $d(x_\vartheta, Y) \geq \vartheta$.*

Proof. Let $u \in X \setminus Y$. Since Y is closed it holds that $d(u, Y) = m > 0$. We choose $y \in Y$ such that $\|u - y\| \leq m/\vartheta$ and set $x_\vartheta = (u - y)/(\|u - y\|) \in \partial B_1$. Then for every $v \in Y$ it follows that

$$\|x_\vartheta - v\| = \frac{1}{\|u - y\|}\|u - (y + v\|u - y\|)\|\,. \tag{3.1.5}$$

Note that $y + v\|u - y\| \in Y$. Therefore, from (3.1.5) and the choice of $y \in Y$, it results in $\|x_\vartheta - v\| \geq m/(m/\vartheta) = \vartheta$. □

Applying this lemma, we have the following characterization of finite dimensional normed spaces.

Theorem 3.1.21. *A normed space X is finite dimensional if and only if $\overline{B}_1$ is compact.*

Proof. $\Longrightarrow$: This direction follows from Theorem 1.5.38.

$\Longleftarrow$: The set $\overline{B}_1$ is totally bounded; see Remark 1.5.32. Hence, there is $\{x_k\}_{k=1}^n \subseteq \overline{B}_1$ such that

$$\overline{B}_1 \subseteq \bigcup_{k=1}^n \left(x_k + B_{\frac{1}{2}}\right) \tag{3.1.6}$$

with $B_{1/2} = \{x \in X : \|x\| < 1/2\}$. Let $Y = \operatorname{span}\{x_k\}_{k=1}^n$. The Corollary 3.1.19 implies that $Y \subseteq X$ is closed. Suppose that $Y \neq X$. Then by Lemma 3.1.20 we find $\hat{x} \in (X \setminus Y) \cap \partial B_1$

such that

$$d(\hat{x}, Y) \geq \vartheta > \frac{1}{2} . \tag{3.1.7}$$

Comparing (3.1.6) and (3.1.7) we reach a contradiction. Hence $X = Y$ and so X is finite dimensional. □

Proposition 3.1.22. *If X is a finite dimensional normed space, Y is a normed space and $L: X \to Y$ is a linear map, then L is continuous.*

Proof. Suppose $\dim X = n$ and let $\{e_k\}_{k=1}^n$ be a basis for X. Since L is linear, we derive for $x = \sum_{k=1}^n \lambda_k e_k \in X$ with $\lambda_k \in \mathbb{K}$ that $L(x) = \sum_{k=1}^n \lambda_k L(e_k)$. Hence $\|L(x)\|_Y \leq \sum_{k=1}^n |\lambda_k| \|e_k\|_X \leq M \left(\sum_{k=1}^n |\lambda_k|^2\right)^{1/2}$ by the Bunyakowsky–Cauchy–Schwarz inequality for finite sums with $M = \left(\sum_{k=1}^n \|L(e_k)\|_Y^2\right)^{1/2}$. On the other hand we know from Proposition 3.1.17 the existence of $m > 0$ such that $\|\lambda\|_2 \leq 1/m \|x\|_X$ where $\lambda = (\lambda_k)_{k=1}^n \in \mathbb{R}^n$. Therefore, it follows $\|L(x)\|_Y \leq M/m \|x\|_X$. Hence L is continuous. □

Remark 3.1.23. In particular, if X is a finite dimensional normed space, then every linear functional $f: X \to \mathbb{R}$ is continuous. In fact the converse is true as well.

We conclude our discussion of finite dimensional topological vector spaces with a result closely related to Theorem 3.1.21. It says that there are no infinite dimensional locally compact topological vector spaces.

Proposition 3.1.24. *A topological vector space (X, τ) is locally compact if and only if X is finite dimensional.*

Proof. $\Longrightarrow$: Let $U \in \mathcal{N}(0)$ be relatively compact. So there is $\{x_k\}_{k=1}^n \subseteq U$ such that

$$\overline{U} \subseteq \bigcup_{k=1}^{n} \left(x_k + \frac{1}{2} U\right) = \{x_1, \ldots, x_n\} + \frac{1}{2} U . \tag{3.1.8}$$

Let $\operatorname{span}\{x_k\}_{k=1}^n$. Then from (3.1.8) it follows

$$\frac{1}{2}\overline{U} \subseteq \frac{1}{2}\left[Y + \frac{1}{2}U\right] = Y + \frac{1}{2^2} U .$$

By induction we have

$$\overline{U} \subseteq Y + \frac{1}{2^n} U \quad \text{for all } n \in \mathbb{N} . \tag{3.1.9}$$

We fix $x \in \overline{U}$. Then from (3.1.9) we see that $x = y_n + 1/2^n u_n$ with $y_n \in Y$, $u_n \in U$ and $n \in \mathbb{N}$. Since U is relatively compact we find a subnet $\{u_\beta\}_{\beta \in J}$ of $\{u_n\}_{n \in \mathbb{N}}$ such that $u_\beta \to u$. Moreover, $1/2^\beta \to 0$. Hence, $y_\beta = x - (1/2^\beta) u_\beta \to x \in Y$. Therefore $\overline{U} \subseteq Y$ and since U is absorbing, we conclude that $X = Y$. Hence, X is finite dimensional.

$\Longleftarrow$: Since X is finite dimensional, we see that X is linearly homeomorphic to $(\mathbb{R}^n, \|\cdot\|_2)$. As X is a normed space, invoking Theorem 3.1.21, we get that $\overline{B}_1$ is compact. Thus, X is locally compact. □

Proposition 3.1.25. *If (X, τ) is a topological vector space and $A \subseteq X$, then the following statements are equivalent:*
(a) *A is bounded; see Definition* 3.1.13*(a).*
(b) *If $\{x_n\}_{n\geq 1} \subseteq A$ and $\{\lambda_n\}_{n\geq 1} \subseteq \mathbb{K}$ with $\lambda_n \to 0$, then $\lambda_n x_n \to 0$ in X.*

Proof. (a) $\Longrightarrow$ (b): Let $U \in \mathcal{N}(0)$ be balanced; see Corollary 3.1.12. Then $A \subseteq tU$ for some $t > 0$. Suppose $\{x_n\}_{n\geq 1} \subseteq A$ and $\{\lambda_n\}_{n\geq 1} \subseteq \mathbb{K}$ such that $\lambda_n \to 0$. Then there exists $n_0 \in \mathbb{N}$ such that $|\lambda_n| t < 1$ for all $n > n_0$. Since U is balanced, it follows $\lambda_n x_n = \lambda_n t 1/t x_n \in U$ for all $n > n_0$. We conclude that $\lambda_n x_n \to 0$ in X as $n \to \infty$.

(b) $\Longrightarrow$ (a): Arguing by contradiction suppose that A is not bounded. Then there exist $t_n \to +\infty$ and $U \in \mathcal{N}(0)$ such that $(X \setminus t_n U) \cap A \neq \emptyset$ for all $n \in \mathbb{N}$. Let $x_n \in A$ with $x_n \notin t_n U$ for all $n \in \mathbb{N}$. We have $1/t_n x_n \notin U$ for all $n \in \mathbb{N}$. Hence $1/t_n u_n$ does not converge to 0, a contradiction to our hypothesis. □

Next we take a closer look at convex sets. In Proposition 3.1.10(c) we saw that the interior and the closure of a convex set remain convex. In fact we can say more.

Proposition 3.1.26. *If X is a topological vector space, $C \subseteq X$ is a convex set and* $0 \leq t < 1$, *then* $(1-t)\operatorname{int} C + t\overline{C} \subseteq \operatorname{int} C$.

Proof. For $t = 0$, the result is trivially true. So, suppose that $0 < t < 1$ and let $x \in \operatorname{int} C$ and $u \in \overline{C}$. Then there exists $U \in \mathcal{N}(0)$ such that $x + U \subseteq C$. Note that $u - (1-t)/tU \in \mathcal{N}(u)$ and so there exists $y \in C \cap (u - (1-t)/tU)$. Therefore $t(u - y) \in (1-t)U$. Let $V = (1-t)(x + U) + ty = (1-t)x + (1-t)U + ty$. This is a nonempty open set and $V \subseteq C$ due to the convexity of C. One gets

$$(1-t)x + tu = (1-t)x + t(u-y) + ty \in (1-t)x + (1-t)U + ty = V \subseteq C\,,$$

which gives $(1-t)x + tu \in \operatorname{int} C$. □

Proposition 3.1.27. *If X is a topological vector space and $C \subseteq X$ is convex, then* $\overline{\operatorname{int} C} = \overline{C}$ *and* $\operatorname{int} \overline{C} = \operatorname{int} C$.

Proof. From Proposition 3.1.26 it follows $(1-t)\operatorname{int} C + t\overline{C} \subseteq \operatorname{int} C$ for all $0 \leq t \leq 1$. Letting $t \to 1^-$ gives $\overline{C} = \overline{\operatorname{int} C}$.

Let $u \in \operatorname{int} C$ and $x \in \operatorname{int} \overline{C}$. Then there exists $U \in \mathcal{N}(0)$ such that $x + U \subseteq \overline{C}$. Since U is absorbing there exists $\vartheta \in (0, 1)$ such that $\vartheta(x-u) \in U$. Then $x+\vartheta(x-u) \in \overline{C}$. Applying Proposition 3.1.26 gives $x - \vartheta(x-u) = (1-\vartheta)x + \vartheta u \in \operatorname{int} C$. Applying Proposition 3.1.26 again yields

$$x = \frac{1}{2}\,[x - \vartheta(x-u)] + \frac{1}{2}\,[x + \vartheta(x-u)] \in \operatorname{int} C\,.$$

This shows $\operatorname{int} \overline{C} \subseteq \operatorname{int} C \subseteq \operatorname{int} \overline{C}$ and so $\operatorname{int} \overline{C} = \operatorname{int} C$. □

Remark 3.1.28. Usually, sets C satisfying $\overline{\operatorname{int} C} = \overline{C}$ and $\operatorname{int} \overline{C} = \operatorname{int} C$ are called **regular**.

Clearly the intersection of any family of convex sets is again convex. So we can state the following definition.

Definition 3.1.29. Let X be a vector space and $A \subseteq X$ a nonempty set. The **convex hull** of A, denoted by conv A, is the intersection of all convex sets that contain A. Therefore, conv A is the smallest convex set containing A. An alternative description is given by

$$\operatorname{conv} A = \left\{ x \in X \colon \exists x_k \in A, t_k \geq 0, k = 1, \ldots, n \text{ with } \sum_{k=1}^{n} t_k = 1, x = \sum_{k=1}^{n} t_k x_k \right\} .$$

That is, conv A is the set of all convex combinations of elements in A. If X is a topological vector space, the **closed convex hull** of A, denoted by $\overline{\operatorname{conv}}\, A$, is the set $\overline{\operatorname{conv} A}$.

For finite dimensional vector spaces, the convex hull of a set is described more precisely by the so-called "Carathéodory Convexity Theorem."

Theorem 3.1.30 (Carathéodory Convexity Theorem). *If X is an m-dimensional vector space, $A \subseteq X$, and $x \in \operatorname{conv} A$, then x is the convex combination of at most $(m+1)$-elements of A.*

Proof. From Definition 3.1.29 we know that $x = \sum_{k=1}^{n} t_k x_k$ with $t_k \geq 0, x_k \in A, k = 1, \ldots, n$ and $\sum_{k=1}^{n} t_k = 1$. Without any loss of generality we may assume that $t_k > 0$ for all $k = 1, \ldots, n$.

Suppose that $n > m+1$, then $\{x_k - x_1\}_{k=2}^{n}$ must be linearly dependent. Hence, there exist $\beta_2, \ldots, \beta_m \in \mathbb{R}$ not all of them equal to zero such that

$$\sum_{k=2}^{n} \beta_k x_k - \left(\sum_{k=2}^{n} \beta_k \right) x_1 = 0 .$$

Thus, there are $\eta_1, \ldots, \eta_n \in \mathbb{R}$ not all of them equal to zero such that $\sum_{k=1}^{n} \eta_k x_k = 0$ and $\sum_{k=1}^{n} \eta_k = 0$.

We set

$$I_+ = \{k \in \{1, \ldots, n\} \colon \eta_k > 0\} , \qquad I_- = \{k \in \{1, \ldots, n\} \colon \eta_k < 0\} ,$$
$$\mu = \min_{k \in I_+} \frac{t_k}{\eta_k} , \qquad J = \{k \in I_+ \colon t_k - \mu \eta_k = 0\} .$$

The sets I_+, I_- and J are nonempty and $\mu > 0$. One obtains

$$x = \sum_{k=1}^{n} t_k x_k = \sum_{k=1}^{n} (t_k - \mu \eta_k) x_k = \sum_{k \notin J} (t_k - \mu \eta_k) x_k . \tag{3.1.10}$$

If $k \in I_+$, then $t_k - \mu\eta_k \geq 0$. If $k \in I_-$, then $t_k - \mu\eta_k > 0$. If $k \in I_+ \setminus J$, then $t_k - \mu\eta_k > 0$. Moreover, we get

$$\sum_{k=1}^{n} (t_k - \mu \eta_k) = \sum_{k=1}^{n} t_k - \mu \sum_{k=1}^{n} \eta_k = 1 . \tag{3.1.11}$$

From (3.1.10) and (3.1.11) we see that x is written as a convex combination with positive weights of n' elements with $n' < n$. We repeat this process until $n' \leq m+1$. $\square$

Corollary 3.1.31. *If X is an m-dimensional topological vector space and $K \subseteq X$ is compact, then* conv $K \subseteq X$ *is compact as well.*

Proof. Let $D = \{(t_1, \dots, t_{m+1})\colon t_k \geq 0, k = 1, \dots, m+1, \sum_{k=1}^{m+1} t_k = 1\} \subseteq \mathbb{R}^{m+1}$ and consider the map $\xi\colon \mathbb{R}^{m+1} \times \left(\prod_{k=1}^{m+1} X_k = X\right) \to X$ defined by

$$\xi\left((t_k)_{k=1}^{m+1}, x_1, \dots, x_{m+1}\right) = \sum_{k=1}^{m+1} t_k x_k \,.$$

It is easy to see that ξ is continuous. Since $D \subseteq \mathbb{R}^{m+1}$ and $\prod_{k=1}^{m+1}(C_k = K) \subseteq \prod_{k=1}^{m+1}(X_k = X)$ are both compact, we get that $D x \left(\prod_{k=1}^{m+1} C_k = K\right)$ is compact as well and so $\xi(D, K, \dots, K) \subseteq X$ is also compact. But according to Theorem 3.1.30, $\xi(D, K, \dots, K) = \operatorname{conv} K$. Hence, conv $K \subseteq X$ is compact. ☐

The corollary fails in infinite dimensional topological vector spaces.

Example 3.1.32. Let $c_0 = \{(x_n)_{n\geq 1}\colon x_n \in \mathbb{R}$ for all $n \in \mathbb{N}$ with $x_n \to 0\}$ furnished with the norm $\|(x_n)_{n\geq 1}\| = \sup\{|x_n|\colon n \in \mathbb{N}\}$. Then c_0 is a Banach space. Let $\hat{u}_n = (\delta_{k,n} 1/n)$ with $\delta_{k,n}$ being the Kronecker delta. Evidently $\hat{u}_n \in c_0$ for all $n \in \mathbb{N}$. Let $K = \{\hat{u}_n\} \cup \{0\}$. Then $K \subseteq c_0$ is compact, but

$$\hat{u} = \sum_{n\geq 1} \frac{1}{2^k} \hat{u}_n \in \overline{\operatorname{conv}} K \,, \quad \hat{u} \notin \operatorname{conv} K \,.$$

Thus, conv K is not closed, hence it is not compact.

In the next definition we extend the notion of total boundedness (see Definition 1.5.31), to general topological vector spaces that are not necessarily metrizable.

Definition 3.1.33. Let X be a topological vector space with a local basis $\mathcal{B}$. A set $A \subseteq X$ is said to be **totally bounded** if for every $U \in \mathcal{B}$ there exists a finite subset $F \subseteq X$ such that $A \subseteq F + U$.

Remark 3.1.34. The following assertions are easy to see:
(a) A totally bounded set is bounded; see Definition 3.1.13(a).
(b) The closure of a totally bounded set is totally bounded.
(c) Compact sets are totally bounded.

Proposition 3.1.35. *If X is a locally convex space and $A \subseteq X$ is totally bounded, then* conv A *is totally bounded.*

Proof. Let $U \in \mathcal{N}(0)$ be convex. Since A is totally bounded, there exists a finite $F \subseteq X$ such that $A \subseteq F + 1/2U$. Corollary 3.1.31 implies that conv F is compact. Let $x \in$ conv A. Then

$$x = \sum_{k=1}^{n} t_k x_k \quad \text{with} \quad t_k \geq 0, x_k \in A, k = 1, \dots, n, \sum_{k=1}^{n} t_k = 1 \,.$$

For every $k \in \{1, \dots, n\}$ there is $u_k \in F$ such that $x_k \in u_k + 1/2U$. Then one gets

$$x = \sum_{k=1}^{n} t_k(x_k - u_k) + \sum_{k=1}^{n} t_k u_k \in \frac{1}{2}U + \operatorname{conv} F$$

since U is convex. Hence

$$\operatorname{conv} A \subseteq \operatorname{conv} F + \frac{1}{2}U\,. \tag{3.1.12}$$

As we already remarked, $\operatorname{conv} F \subseteq X$ is compact. Thus, we find a finite set $E \subseteq \operatorname{conv} F$ such that $\operatorname{conv} F \subseteq E + 1/2U$, which gives, due to (3.1.12) and the fact that U is convex, that $\operatorname{conv} A \subseteq E + U$ and so we see that $\operatorname{conv} A$ is totally bounded in the sense of Definition 3.1.33. □

From this proposition we deduce the following useful result.

Theorem 3.1.36. *If X is a Fréchet space and $A \subseteq X$ is compact, then $\overline{\operatorname{conv}}\,A \subseteq X$ is compact as well.*

Proof. Since $A \subseteq X$ is compact, it is totally bounded; see Theorem 1.5.36. Then Proposition 3.1.35 implies that $\operatorname{conv} A$ is totally bounded, hence $\overline{\operatorname{conv}}\,A$ is totally bounded; see Remark 3.1.34. Then Theorem 1.5.36 implies that $\overline{\operatorname{conv}}\,A$ is compact. □

Next we introduce an important class of convex functionals that describes locally convex spaces.

Definition 3.1.37. Let X be a vector space. A function $\rho\colon X \to \mathbb{R}$ is a **seminorm** if the following hold:
(a) ρ is subadditive, that is, $\rho(x + u) \le \rho(x) + \rho(u)$ for all $x, u \in X$.
(b) ρ is absolutely homogeneous, that is, $\rho(\lambda x) = |\lambda|\rho(x)$ for all $\lambda \in \mathbb{K}$ and for all $x \in X$.
If $\rho(x) \neq 0$ for $x \neq 0$, then the seminorm is a norm; see Definition 3.1.13(e). A family $\mathcal{P}$ of seminorms on X is said to be **separating** if for each $x \neq 0$ there exists $\rho \in \mathcal{P}$ such that $\rho(x) \neq 0$. Given an absorbing set $A \subseteq X$, the real functional $\rho_A\colon X \to \mathbb{R}$ defined by $\rho_A(x) = \inf[t > 0\colon x \in tA]$ is the **Minkowski functional of** A (or **gauge of** A).

Proposition 3.1.38. *If X is a vector space and $\rho\colon X \to \mathbb{R}$ is a seminorm, then the following hold:*
(a) *$\rho(0) = 0$, $|\rho(x) - \rho(u)| \le \rho(x - u)$ for all $x, u \in X$, $\rho(x) \ge 0$ for all $x \in X$;*
(b) *$N(\rho) = \{x \in X\colon \rho(x) = 0\}$ is a vector subspace of X;*
(c) *$B_1 = \{x \in X\colon \rho(x) < 1\}$ is convex, absorbing, balanced, and $\rho = \rho_{B_1}$.*

Proof. (a) From Definition 3.1.37 we have $\rho(0) = \rho(\lambda 0) = |\lambda|\rho(0)$ for all $\lambda \in \mathbb{K}$, hence $\rho(0) = 0$. Moreover,

$$\rho(x) = \rho(x - u + u) \le \rho(x - u) + \rho(u) \quad \text{for all } x, u \in X\,,$$

hence, $|\rho(x) - \rho(u)| \le \rho(x - u)$ by interchanging the roles of x and u. If $u = 0$, then we see that $\rho(x) \ge 0$ for all $x \in X$.

(b) Let $\lambda \in \mathbb{K}$ and $x, u \in N(p)$. Then

$$0 \le \rho(\lambda x + u) \le |\lambda|\rho(x) + \rho(u) = 0 .$$

Hence $\lambda x + u \in N(p)$ and so $N(p)$ is a vector subspace of X.

(c) Let $x, u \in B_1$ and $t \in (0, 1)$. Then $\rho((1-t)x + tu) \le (1-t)\rho(x) + t\rho(u) < 1$, which implies that B_1 is convex.

If $x \in X$ and $\vartheta > \rho(x)$, then $\rho(1/\vartheta x) = 1/\vartheta \rho(x) < 1$ and so B_1 is absorbing. Moreover, it is clear that B_1 is balanced. From the previous argument we see that $\rho_{B_1} \le \rho$. Next let $0 < \eta \le \rho(x)$. Then $1 \le \rho(1/\eta x)$ and so $1/\eta x \notin B_1$. Therefore, $\rho \le \rho_{B_1}$ and we conclude that $\rho = \rho_{B_1}$. □

For the Minkowski functional we obtain the following result.

Proposition 3.1.39. *If X is a vector space and $A \subseteq X$ is convex and absorbing, then the following hold:*
(a) *ρ_A is subadditive and positively homogeneous, that is, ρ_A is sublinear;*
(b) *ρ_A is a seminorm if A is in addition balanced;*
(c) *if $B = \{x \in X : \rho_A(x) < 1\}$ and $C = \{x \in X : \rho_A(x) \le 1\}$, then $B \subseteq A \subseteq C$ and $\rho_B = \rho_A = \rho_C$.*

Proof. (a) For every $x \in X$, let $A(x) = \{t > 0 : x \in tA\}$. Pick $t \in A(x)$ and $\vartheta > t$. Since $0 \in A$ and A is convex, it holds $\vartheta \in A(x)$. Therefore, $A(x)$ is a half-line starting at $\rho_A(x)$. Suppose that $\rho_A(x) < \vartheta$ and $\rho_A(u) < \mu$. Let $\tau = \vartheta + \mu$. Then it follows that $1/\vartheta x \in A$, $1/\mu u \in A$ and since A is convex

$$\frac{1}{\tau}(x + u) = \left(\frac{\vartheta}{\tau}\right)\frac{1}{\vartheta}x + \left(\frac{\mu}{\tau}\right)\frac{1}{\mu}u \in A .$$

This gives $\rho_A(x + u) \le \tau$ and so ρ_A is subadditive. Of course, ρ_A is also positively homogeneous.

(b) This is immediate from Definition 3.1.6(c) and Definition 3.1.37.

(c) Suppose $\rho_A(x) < 1$. Then $1 \in A(x)$ and so $x \in X$. On the other hand if $x \in A$, then $\rho_A(x) \le 1$ and so we conclude that $B \subseteq A \subseteq C$. It follows that $B(x) \subseteq A(x) \subseteq C(x)$ for every $x \in X$ and so $\rho_B(x) \le \rho_A(x) \le \rho_C(x)$. Suppose $\rho_C(x) < \vartheta < \mu$. Then $1/\vartheta x \in C$ and so $\rho_A(1/\vartheta x) \le 1$, hence

$$\rho_A\left(\frac{1}{\mu}x\right) = \rho_A\left(\frac{\vartheta}{\mu}\frac{1}{\vartheta}x\right) = \frac{\vartheta}{\mu}\rho_A\left(\frac{1}{\vartheta}x\right) \le \frac{\vartheta}{\mu} < 1 .$$

Therefore, $1/\mu x \in B$, $\rho_B(1/\mu x) \le 1$, hence $\rho_B(x) \le \mu$. We conclude that $\rho_B = \rho_A = \rho_C$. □

Seminorms characterize locally convex topologies. The following theorem can be found in Yosida [311, p. 26].

Theorem 3.1.40. *If X is a vector space and $\{\rho_\alpha\}_{\alpha \in I}$ is a separating family of seminorms on X, then there is a weakest locally convex topology on X making all the seminorms continuous. Conversely, any locally convex space is topologized by the seminorms defined* *by the Minkowski functionals of the convex, absorbing, balanced sets. Such sets are often called* ***barrels****. Moreover, if $\mathcal{F} \subseteq \mathbb{R}^X$ is a set of $\mathbb{R}$-valued linear functionals on X, then the weakest topology on X making all elements of $\mathcal{F}$ continuous is locally convex.*

What about normable spaces, the next more restrictive class of vector spaces after locally convex spaces? We have the following theorem known as "Kolmogorov's Normability Criterion."

Theorem 3.1.41 (Kolmogorov's Normability Criterion). *A topological vector space X is normable if and only if it is locally convex and locally bounded, that is, it possesses a bounded convex neighborhood of the origin.*

Proof. $\Longrightarrow$: The open unit ball $B_1 = \{x \in X : \|x\| < 1\}$ is a bounded convex neighborhood of the origin.

$\Longleftarrow$: Let $U \in \mathcal{N}(0)$ be bounded convex. We may also assume that U is balanced; see Corollary 3.1.12. Let $\|x\| = \rho_U(x)$ for all $x \in X$ with ρ_U being the Minkowski functional of U. Note that $\{tU\}_{t>0}$ is a local basis for the topology of X. If $x \neq 0$, then there exists $t > 0$ such that $x \notin tU$ and so $\|x\| = \rho_U(x) \geq t$. Then, from Proposition 3.1.39(b) we infer that $\|\cdot\|$ is a norm on X. Moreover, from Proposition 3.1.39(c) we conclude that $\{x \in X : \|x\| < t\} = tU$ for all $t > 0$. Therefore the norm topology coincides with the initial locally convex topology on X. ☐

Now we are ready for one of the most important results in analysis with far-reaching consequences. This is the celebrated "Hahn–Banach Extension Theorem."

Theorem 3.1.42 (Hahn–Banach Extension Theorem). *If X is a vector space, $\rho: X \to \mathbb{R}$ is subadditive and positively homogeneous, that is, sublinear, $V \subseteq X$ is a vector subspace, $f: V \to \mathbb{R}$ is linear and $f(x) \leq \rho(x)$ for all $x \in V$, then there exists $\hat{f}: X \to \mathbb{R}$ being linear such that $\hat{f}|_V = f$ and $\hat{f}(x) \leq \rho(x)$ for all $x \in X$.*

Proof. We assume that $V \neq X$ and let $u \in X \setminus V$. Let $Y = \operatorname{span}\{V \cup \{u\}\}$. Then each $y \in Y$ can be written in a unique way as $y = x + \lambda u$ with $x \in V$ and $\lambda \in \mathbb{R}$. Then any extension $\hat{f}$ of f on Y must be of the form $\hat{f}(x + \lambda u) = f(x) + \lambda \hat{f}(u)$. So, the main problem is to define $\hat{f}(u)$. Recall that the extension $\hat{f}$ must satisfy $\hat{f} \leq \rho$ on Y. Therefore

$$f(x) + \lambda \hat{f}(u) \leq \rho(x + \lambda u)\,. \tag{3.1.13}$$

Taking $\lambda = 1$ in (3.1.13) yields $\hat{f}(u) \leq \rho(x+u) - f(x)$. Similarly, if we take $\lambda = -1$ and replace x by $-x$ in (3.1.13) we infer $-f(x) - \hat{f}(u) \leq \rho(-x-u)$ because the subadditivity of ρ implies $-f(x) \leq f(-x)$. It follows that

$$-f(v) - \rho(-v-u) \leq \hat{f}(u) \leq -f(x) + \rho(x+u) \quad \text{for all } v, x \in V\,. \tag{3.1.14}$$

Therefore the value $\hat{f}(u)$ cannot be chosen arbitrarily but it must satisfy (3.1.14). However, in order to make (3.1.14) possible, we need to have

$$-f(v) - \rho(-v-u) \leq -f(x) + \rho(x+u) \quad \text{for all } v, x \in V\,. \tag{3.1.15}$$

But note that $f(x) - f(v) = f(x - v) \leq \rho(x - v) = \rho(x + u + (-v - u)) \leq \rho(x + u) + \rho(-v - u)$ and so (3.1.15) holds. Now we can define the extension $\hat{f}$ of f on Y. We can take for example $\hat{f}(u) = \inf[-f(x) + \rho(x + u) : x \in V]$ and obtain $\hat{f}(x + \lambda u) = f(x) + \lambda \hat{f}(u)$. Clearly $\hat{f}$ is linear on Y and $\hat{f}|_V = f$. We need to show that $\hat{f} \leq \rho$. Since $\hat{f}|_V = f$ we get $\hat{f} \leq \rho$ when $\lambda = 0$. So, let $\lambda \neq 0$. Then we replace $v, x \in V$ by $1/\lambda x \in V$ in (3.1.14). This gives

$$-f\left(\frac{1}{\lambda}x\right) - \rho\left(-\frac{1}{\lambda}x - u\right) \leq \hat{f}(u) \leq -f\left(\frac{1}{\lambda}x\right) + \rho\left(\frac{1}{\lambda}x + u\right),$$

which implies

$$f\left(\frac{1}{\lambda}x\right) + \hat{f}(u) \leq \rho\left(\frac{1}{\lambda}x + u\right) \quad -f\left(\frac{1}{\lambda}x\right) - \hat{f}(u) \leq \rho\left(-\frac{1}{\lambda}x - u\right). \tag{3.1.16}$$

If $\lambda > 0$, then if we multiply the first inequality in (3.1.16) with λ, we obtain $\hat{f}(x + \lambda u) \leq \rho(x + \lambda u)$. If $\lambda < 0$, then multiplying the second inequality in (3.1.16) with $-\lambda$ gives $\hat{f}(x + \lambda u) \leq \rho(x + \lambda u)$. In summary we have showed that $\hat{f} \leq \rho$.

Now let

$$\mathcal{L} = \{(Y, \hat{f}) : Y \text{ is a subspace of } X \text{ containing } V \text{ and } \\ \hat{f} \text{ is a linear extension of } f \text{ on } Y \text{ with } \hat{f} \leq \rho\}.$$

We order $\mathcal{L}$ as follows: $(Y, \hat{f}) \leq (Y', \hat{f}')$ if $Y \subseteq Y'$ and $\hat{f}'|_Y = \hat{f}$. Then every chain $\mathcal{D}$ of $\mathcal{L}$ has an upper bound in $\mathcal{L}$ namely if $\mathcal{D} = \{(Y_\alpha, \hat{f}_\alpha)_{\alpha \in I}\}$, then $Y = \bigcup_{\alpha \in I} Y_\alpha$ is a linear subspace of X and $\hat{f}(x) = \hat{f}_\alpha(x)$ for $x \in Y_\alpha$ is a well-defined linear functional on Y. Evidently, $(Y, \hat{f}) \in \mathcal{L}$ and $(Y_\alpha, \hat{f}_\alpha) \leq (Y, \hat{f})$ for all $\alpha \in I$. By Zorn's Lemma (see Section 1.4), $\mathcal{L}$ admits a maximal element $(Y, \hat{f})$. We must have $Y = X$ or otherwise we repeat the construction in the first part of the proof and contradict the maximality of $(Y, \hat{f})$. □

Remark 3.1.43. It should be noted that the extension $\hat{f}$ is in general not unique.

A careful reading of the proof of Theorem 3.1.42 reveals that the complex variant of the result requires a modification of the condition on ρ since positive homogeneity of ρ makes in that case no sense.

Theorem 3.1.44 (Hahn–Banach Extension Theorem (Complex Variant)). *If X is a complex vector space, $\rho : X \to \mathbb{R}$ is a seminorm, $V \subseteq X$ is a vector subspace, $f : V \to \mathbb{C}$ is linear and $|f(x)| \leq \rho(x)$ for all $x \in X$, then there exists $\hat{f} : X \to \mathbb{C}$ being linear such that $\hat{f}|_V = f$ and $|\hat{f}(x)| \leq \rho(x)$ for all $x \in X$.*

From now on, unless otherwise stated, all vector spaces will be over the reals.

Definition 3.1.45. Let X, Y be normed spaces. A linear operator $A : X \to Y$ is **bounded** if $\|A(x)\|_Y \leq M\|x\|_X$ for some $M > 0$ and for all $x \in X$. The smallest $M \geq 0$ for which the inequality above holds, is called the **operator norm** of A and it is denoted by

$$\|A\|_L = \sup\left[\frac{\|A(x)\|_Y}{\|x\|_X} : x \in X, x \neq 0\right].$$

By $L(X, Y)$ we denote the vector space of all bounded, linear operators from X into Y. Evidently $(L(X, Y), \|\cdot\|_L)$ is a normed space and the resulting norm topology is called the **uniform operator topology**. If $Y = \mathbb{R}$, then $L(X, \mathbb{R}) = X^*$ is the **topological dual** and its elements are called **bounded linear functionals**. If $x \in X$ and $x^* \in X^*$ we usually write $\langle x^*, x\rangle$ instead of $x^*(x)$ and call $\langle\cdot,\cdot\rangle$ the **duality brackets** for the pair (X^*, X).

The proof of the next proposition is straightforward and so its proof is omitted.

Proposition 3.1.46. *If X, Y are normed spaces and $A\colon X \to Y$ is a linear operator, then the following properties are equivalent:*
(a) *A is bounded.*
(b) *A is continuous.*
(c) *A is continuous at $x = 0$.*

Proposition 3.1.47. *If X is a normed space and Y is a Banach space, then $(L(X, Y), \|\cdot\|_L)$ is a Banach space.*

Proof. Suppose that $\{A_n\}_{n\ge 1} \subseteq L(X, Y)$ is a $\|\cdot\|_L$-Cauchy sequence. Then it follows

$$\|(A_n - A_m)(x)\|_Y \le \|A_n - A_m\|_L\|x\| \quad \text{for all } n, m \in \mathbb{N} \text{ and for all } x \in X\,.$$

Since Y is complete, one gets that $A(x) = \lim_{n\to\infty} A_n(x)$ exists for all $x \in X$. Of course, $A\colon X \to Y$ is linear and

$$\|A(x) - A_n(x)\|_Y = \lim_{m\to\infty} \|A_m(x) - A_n(x)\|_Y \le \limsup_{m\to\infty} \|A_m - A_n\|_L\|x\|\,.$$

So, for given $\varepsilon > 0$, there exists $n_0 = n_0(\varepsilon) \in \mathbb{N}$ such that

$$\|A(x) - A_n(x)\|_Y \le \varepsilon\|x\| \quad \text{for all } x \in X \text{ and for all } n \ge n_0\,. \tag{3.1.17}$$

Hence

$$\|A(x)\|_Y = \|A(x) - A_{n_0}(x)\|_Y + \|A_{n_0}(x)\|_Y \le (\varepsilon + \|A_{n_0}\|_L)\|x\|_X\,.$$

This implies that $A \in L(X, Y)$ and $\|A_n - A\|_L \to 0$ as $n \to \infty$; see (3.1.17). ☐

Corollary 3.1.48. *If X is a normed space, then X^* is a Banach space and $\|x^*\|_* = \sup\{|\langle x^*, x\rangle|\colon \|x\| \le 1\} = \sup\{\langle x^*, x\rangle\colon \|x\| \le 1\}$.*

Proposition 3.1.49. *If X is a normed space, $V \subseteq X$ is a vector subspace, and $u^* \in V^*$, then there exists $x^* \in X^*$ such that $x^*\big|_V = u^*$ and $\|x^*\|_* = \|u^*\|_{V^*}$.*

Proof. Applying Theorem 3.1.42 with $\rho(x) = \|u^*\|_{V^*}\|x\|$ for all $x \in X$ yields the assertion. ☐

Proposition 3.1.50. *If X is a normed space and $x_0 \in X$, then there exists $x_0^* \in X^*$ such that $\|x_0^*\|_* = \|x_0\|$ and $\langle x_0^*, x_0\rangle = \|x_0\|^2$.*

Proof. Applying Proposition 3.1.49 with $V = \mathbb{R}x_0$ and $x_0^*(tx_0) = \langle x_0^*, tx_0\rangle = t\|x_0\|^2$ gives the desired result $\|x_0^*\|_* = \|x_0\|$. ☐

Remark 3.1.51. The element $x_0^* \in X^*$ is not unique in general. In order to have uniqueness we need additional structure on X^*, for example, strict convexity; see Section 3.4. The multivalued map $\mathcal{F}\colon X \to 2^{X^*} \setminus \{\emptyset\}$ defined by $\mathcal{F}(x) = \{x^* \in X^*\colon \|x^*\|_* = \|x\|$ and $\langle x^*, x\rangle = \|x\|^2\}$ is called the **duality map** from X into X^*. It is important in Nonlinear Analysis and we will encounter it again in Section 6.1.

Proposition 3.1.52. *If X is a normed space and $x \in X$, then*

$$\|x\| = \sup\left[|\langle x^*, x\rangle|\colon x^* \in X^*, \|x^*\|_* \le 1\right] = \sup\left[\langle x^*, x\rangle\colon x^* \in X^*, \|x^*\|_* \le 1\right] .$$

Proof. We assume that $x \neq 0$. Note that

$$\sup\left[|\langle x^*, x\rangle|\colon x^* \in X^*, \|x^*\|_* \le 1\right] \le \|x\| . \tag{3.1.18}$$

On the other hand from Proposition 3.1.50 we know that there is $x_0^* \in X^*$ such that $\|x_0^*\|_* = \|x\|$ and $\langle x_0^*, x\rangle = \|x\|^2$. Let $\hat{x}_0^* = x_0^*/\|x\|$. Then $\|\hat{x}_0^*\|_* = 1$ and $\langle \hat{x}_0^*, x\rangle = \|x\|$. This combined with (3.1.18) implies the assertion of the proposition. □

Now we will produce some important geometric interpretations of the Hahn–Banach Extension Theorem; see Theorem 3.1.42. These are the well-known "Separation Theorems" for convex sets.

Definition 3.1.53. Let X be a vector space. A **hyperplane** is a set of the form $\{f = \vartheta\} = \{x \in X\colon f(x) = \vartheta\}$ with $f\colon X \to \mathbb{R}$ being linear and $\vartheta \in \mathbb{R}$. A hyperplane determines two **half-spaces**, namely $\{f \ge \vartheta\} = \{x \in X\colon f(x) \ge \vartheta\}$ and $\{f \le \vartheta\} = \{x \in X\colon f(x) \le \vartheta\}$. Given two sets $A, C \subseteq X$, we say that the hyperplane $H = \{f = \vartheta\}$ **separates** A **and** C, if $A \subseteq H_- = \{f \le \vartheta\}$ and $C \subseteq H_+ = \{f \ge \vartheta\}$. We say that H **strongly separates** A and C if there exists $\varepsilon > 0$ such that

$$A \subseteq H_-^\varepsilon = \{f \le \vartheta - \varepsilon\} \quad \text{and} \quad C \subseteq H_+^\varepsilon = \{f \ge \vartheta + \varepsilon\} .$$

Proposition 3.1.54. *If X is a topological vector space, then a hyperplane $H = \{f = \vartheta\}$ is either closed or dense. H is closed if and only if f is continuous while H is dense if and only if f is discontinuous.*

Proof. Due to the linearity of f, we may assume that $\vartheta = 0$. If f is continuous, then H is closed while if H is dense, then clearly f is not continuous. Now assume that H is closed. Suppose that $\{x_\alpha\}_{\alpha\in I} \subseteq X$ and $x_\alpha \to 0$. In addition, let $u \in X$ with $f(u) = 1$. Arguing by contradiction, suppose that $f(x_\alpha) \not\to 0$; see Proposition 3.1.46. Then, for at least a subnet, we have $|f(x_\alpha)| \ge \varepsilon$ for all $\alpha \in I$. Let $v_\alpha = u - f(u)/(f(x_\alpha))x_\alpha$. Then $v_\alpha \in H$ since $\vartheta = 0$ and $v_\alpha \to u$. So, $u \in H$, a contradiction. Therefore $f(x_\alpha) \to 0$ and so f is continuous; see Proposition 3.1.46.

Now suppose that f is discontinuous. Then there exist a net $\{x_\alpha\}_{\alpha\in I} \subseteq X$ and $\varepsilon > 0$ such that

$$x_\alpha \to 0 \quad \text{and} \quad |f(x_\alpha)| \ge \varepsilon \quad \text{for all } \alpha \in I .$$

Given any $u \in X$, let $v_\alpha = u - f(u)/(f(x_\alpha))x_\alpha \in H$ for all $\alpha \in I$. We have $v_\alpha \to u$ and so we conclude that H is dense. □

Definition 3.1.55. Let X be a vector space and $A \subseteq X$. A point $x \in A$ is said to be an **absorbing point** of A, if $A - x \subseteq X$ is absorbing; see Definition 3.1.6(b).

Remark 3.1.56. If X is a topological vector space and $\operatorname{int} A \neq \emptyset$, then every $x \in \operatorname{int} A$ is an absorbing point. However, the set A can have absorbing points even if $\operatorname{int} A = \emptyset$. Suppose that X is a normed space and $A = \partial B_1 \cup \{0\}$ where $\partial B_1 = \{x \in X : \|x\| = 1\}$. Then $x = 0$ is absorbing but $\operatorname{int} A = \emptyset$.

Next we present the "First Separation Theorem."

Theorem 3.1.57 (First Separation Theorem). *If X is a vector space, $A, C \subseteq X$ are two nonempty convex sets, $A \cap C = \emptyset$ and one of them has an absorbing point, then they can be separated by a hyperplane $H = \{f = \vartheta\}$ with $f \neq 0$ and $A \cup C$ is not included in H.*

Proof. Suppose A has an absorbing point. Then $A - C$ has an absorbing point x. Since $A \cap C = \emptyset$, we see that $x \neq 0$. Moreover, the set $E = A - C - x$ is nonempty, convex, and absorbing, and $-x \notin E$ since $A \cap C = \emptyset$. Then Proposition 3.1.39 implies that ρ_E is sublinear.

Suppose that $\rho_E(-x) < 1$. Then there exist $0 \leq t < 1$ and $e \in E$ such that $x = te$. Note that $0 \in E$ being absorbing. So we have $-x = te + (1 - t)0 \in E$, a contradiction. Therefore

$$\rho_E(-x) \geq 1 . \tag{3.1.19}$$

Let $V = \mathbb{R}(-x)$ and let $f : V \to \mathbb{R}$ be defined by $f(t(-x)) = t$. Clearly, f is linear and $f \leq \rho_E$ on V. Indeed, if $t \geq 0$, then $\rho_E(t(-x)) = t\rho_E(-x) \geq t$; see (3.1.19). If $t < 0$, then $f(t(-x)) < 0 \leq \rho_E(t(-x))$. Invoking Theorem 3.1.42 implies the existence of $\hat{f} : X \to \mathbb{R}$ being linear such that $\hat{f}|_V = f$ and $\hat{f} \leq \rho_E$. Note that $\hat{f}(x) = -1$ and so $\hat{f} \neq 0$.

We claim that $\hat{f}$ separates A and C. To see this, let $a \in A$ and $c \in C$. It holds

$$\begin{aligned}\hat{f}(a) &= \hat{f}(a - c - x) + \hat{f}(x) + \hat{f}(c) \leq \rho_E(a - c - x) + \hat{f}(x) + \hat{f}(c)\\ &= \rho_E(a - c - x) - 1 + \hat{f}(c) \leq 1 - 1 + \hat{f}(c) = \hat{f}(c) .\end{aligned}$$

Since $a \in A$ and $c \in C$ are arbitrary, we see that $\hat{f}$ separates A and C. Finally, since $0 \in E$, we have $x = a - c$ with $a \in A$ and $c \in C$. Recall that $\hat{f}(x) = -1$. Then $\hat{f}(a) \neq \hat{f}(c)$ and so we cannot have A and C to be subsets of the same hyperplane. □

Lemma 3.1.58. *If X is a topological vector space, $f : X \to \mathbb{R}$ is linear, and f is bounded above or bounded below on a neighborhood of the origin, then f is continuous.*

Proof. Let $U \in \mathcal{N}(0)$ be symmetric and assume that $f \leq M$ on U. Then, for given $\varepsilon > 0$, one gets, since U is symmetric, that $x - u \in \varepsilon/MU$ implies $|f(x) - f(u)| = |f(x - u)| \leq \varepsilon/MM = \varepsilon$. Hence, f is continuous. □

Using this lemma, we can state a topological version of Theorem 3.1.57.

Theorem 3.1.59. *If X is a topological vector space, $A, C \subseteq X$ are nonempty convex sets, $A \cap C = \emptyset$ and one of them has nonempty interior, then they can be separated by a closed hyperplane H and $A \cup C$ is not included in H.*

Proof. Applying Theorem 3.1.57, we obtain a separating hyperplane $H = \{f = \vartheta\}$ with $f \neq 0$. We only need to show that f is continuous. Suppose that $\operatorname{int} A \neq \emptyset$. Then $f(a) \leq \vartheta \leq f(c)$ for all $a \in A$ and for all $c \in C$. Note that if $x \in \operatorname{int} A$, then $U = \operatorname{int} A - x \in \mathcal{N}(0)$ and so $f|_U$ is bounded above, hence f is continuous; see Lemma 3.1.58. □

Next we present the "Second Separation Theorem" called "Strong Separation Theorem."

Theorem 3.1.60 (Strong Separation Theorem). *If X is a locally convex space and $A, C \subseteq X$ are nonempty, disjoint, convex sets, then A and B can be strongly separated by a closed hyperplane if and only if there exists $U \in \mathcal{N}(0)$ being convex such that $(A + U) \cap C = \emptyset$.*

Proof. $\Longrightarrow$: Let f be the linear functional associated with the closed separating hyperplane. Then f is continuous; see Proposition 3.1.54. Moreover, taking $\varepsilon > 0$ from the strong separation (see Definition 3.1.53), $U = \{x \in X : |f(x)| < \varepsilon\}$ is a convex neighborhood of the origin and $(A + U) \cap C = \emptyset$.

$\Longleftarrow$: The set $A + U$ is convex and open. So, we can apply Theorem 3.1.59 and find a linear, continuous functional $f : X \to \mathbb{R}$ and $\vartheta \in \mathbb{R}$ as well as $\varepsilon > 0$ such that $f(a) \leq \vartheta - \varepsilon$ for all $a \in A$ and $f(c) \geq \vartheta + \varepsilon$ for all $c \in C$. Hence A and C are strongly separated by $H = \{f = \vartheta\}$. □

Corollary 3.1.61. *If X is locally convex, $A, C \subseteq X$ are nonempty, disjoint, convex sets and A is compact as well as C is closed, then A and C can be strongly separated by a closed hyperplane.*

Proof. The set $X \setminus C$ is open and $A \subseteq X \setminus C$. The compactness of A implies that there exists a convex neighborhood $U \in \mathcal{N}(0)$ such that $A + U \subseteq X \setminus C$. Hence $(A + U) \cap C = \emptyset$. Applying Theorem 3.1.60 gives the assertion. □

Proposition 3.1.62. *If X is a normed space, $V \subseteq X$ is a vector subspace, and $\overline{V} \neq X$, then there exists $x^* \in X^*$ with $x^* \neq 0$ such that $\langle x^*, v \rangle = 0$ for all $v \in V$.*

Proof. Let $u \in X \setminus \overline{V}$. Then apply Corollary 3.1.61 with $A = \{x_0\}$ and $C = \overline{V}$. Thus, we find $x^* \in X^*$ with $x^* \neq 0$ and $\vartheta \in \mathbb{R}$ such that $\langle x^*, x_0 \rangle < \vartheta < \langle x^*, v \rangle$ for all $v \in \overline{V}$. But since $\overline{V}$ is a vector space, we see that $\langle x^*, v \rangle = 0$ for all $v \in \overline{V}$ since $\lambda \langle x^*, v \rangle > \vartheta$ for all $\lambda \in \mathbb{R}$, hence $\vartheta < 0$. □

Remark 3.1.63. This proposition is useful for determining whether a linear subspace V is dense in X. We must have that the only element of X^* vanishing on V is $x^* = 0$.

3.2 Three Fundamental Theorems

In this section we present three basic theorems that are the core results of linear functional analysis. These are the “Uniform Boundedness Principle,” the “Open Mapping Theorem,” and the “Closed Graph Theorem.” All three depend on the Baire Category Theorem; see Theorem 1.5.68. We recall that the Baire Category Theorem, roughly speaking, provides conditions for a set to be large in the sense that it has a nonempty interior.

We start with the “Uniform Boundedness Principle.” This theorem asserts that for any family of bounded linear operators, pointwise boundedness implies uniform boundedness, that is, boundedness in the operator norm. As before, we consider real vector spaces.

Theorem 3.2.1 (Uniform Boundedness Principle). *If X is a Banach space, Y is a normed space, and $\mathcal{L} \subseteq L(X, Y)$ satisfies*

$$\sup\,[\|A(x)\|_Y\colon A \in \mathcal{L}] = M(x) < \infty\,,$$

then there exists $M_0 > 0$ such that $\sup\,[\|A\|_L\colon A \in \mathcal{L}] \le M_0$.

Proof. For every $n \in \mathbb{N}$ let $E_n = \{x \in X\colon \|A(x)\|_Y \le n \text{ for all } A \in \mathcal{L}\}$. The hypothesis implies that

$$X = \bigcup_{n\ge 1} E_n\,. \tag{3.2.1}$$

Moreover, we claim that for every $n \in \mathbb{N}$, $E_n \subseteq X$ is closed. To see this, let $\{x_m\}_{m\ge 1} \subseteq E_n$ and assume that $x_m \to x$ in X. We obtain $\|A(x_m)\|_Y \le n$ for all $A \in \mathcal{L}$ and for all $m \in \mathbb{N}$. The continuity of A (see Proposition 3.1.46) implies that $\|A(x_m)\|_Y \to \|A(x)\|_Y$ as $m \to \infty$ for every $A \in \mathcal{L}$. Therefore, $\|A(x)\|_Y \le n$ for all $A \in \mathcal{L}$ and so $x \in E_n$, which implies that $E_n \subseteq X$ is closed for every $n \in \mathbb{N}$.

From (3.2.1) and the Baire Category Theorem (see Theorem 1.5.68 and Corollary 1.5.67), we infer that there exists $n_0 \in \mathbb{N}$ such that $\operatorname{int} E_{n_0} \neq \emptyset$. Hence, there exists $\varepsilon > 0$ such that

$$\overline{B}_\varepsilon(x_0) \subseteq E_{n_0} \quad \text{with} \quad \overline{B}_\varepsilon(x_0) = \{x \in X\colon \|x - x_0\|_X \le \varepsilon\}\,. \tag{3.2.2}$$

Let $x \in X$ with $\|x\|_X \le \varepsilon$ and $A \in \mathcal{L}$. Then, due to (3.2.2),

$$\begin{aligned}\|A(x)\|_Y &= \|A(x + x_0) - A(x_0)\|_Y \le \|A(x + x_0)\|_Y + \|A(x_0)\|_Y\\ &\le n_0 + n_0 = 2n_0\,.\end{aligned} \tag{3.2.3}$$

Thus, for all $u \in X$ with $\|u\|_X = 1$, it follows, because of (3.2.3), that

$$\|A(u)\|_Y = \frac{1}{\varepsilon}\|A(\varepsilon u)\|_Y \le \frac{2n_0}{\varepsilon} \quad \text{for all } A \in \mathcal{L}\,.$$

Hence,

$$\sup\big[\|A(u)\|_Y\colon \|u\|_X \le 1\big] = \|A\|_L \le \frac{2n_0}{\varepsilon} \quad \text{for all } A \in \mathcal{L}\,. \qquad \square$$

Theorem 3.2.1 leads to the so-called "Banach–Steinhaus Theorem," which says that the pointwise limit of a sequence of bounded linear operators is a bounded linear operator.

Theorem 3.2.2 (Banach–Steinhaus Theorem). *If X, Y are Banach spaces and $\{A_n\}_{n\geq 1} \subseteq L(X, Y)$ is a sequence such that*

$$A_n(x) \to A(x) \quad \textit{in } Y \textit{ as } n \to \infty \textit{ for all } x \in X\,,$$

then the following hold:
(a) $A \in L(X, Y)$ *and* $\sup_{n\geq 1} \|A_n\|_L < \infty$;
(b) $\|A\|_L \leq \liminf_{n\to\infty} \|A_n\|_L$.

Proof. (a) Clearly, $A\colon X \to Y$ is linear. Since $\{A_n(x)\}_{n\geq 1} \subseteq Y$ is convergent, it holds that

$$\sup_{n\in\mathbb{N}} \|A_n(x)\|_Y = M(x) < \infty\,.$$

Applying Theorem 3.2.1, there exists $M_0 > 0$ such that $\sup_{n\in\mathbb{N}} \|A_n\|_L \leq M_0 < \infty$, which implies $\|A_n(x)\|_Y \leq M_0\|x\|_X$ for all $x \in X$ and for all $n \in \mathbb{N}$. Therefore, we derive $\|A(x)\|_Y = \lim_{n\to\infty} \|A_n(x)\|_Y \leq M_0\|x\|_X$ for all $x \in X$, which, due to Proposition 3.1.46, results in $A \in L(X, Y)$.

(b) It holds that $\|A_n(x)\|_Y \leq \|A_n\|_L\|x\|_X$ for all $x \in X$ and for all $n \in \mathbb{N}$. This gives $\|A(x)\|_Y \leq \liminf_{n\to\infty} \|A_n\|_L\|x\|_X$ for all $x \in X$ and so, $\|A\|_L \leq \liminf_{n\to\infty} \|A_n\|_L$. □

Example 3.2.3. (a) Theorems 3.2.1 and 3.2.2 fail if X is only a normed space. To see this, let us define the following subspaces:

$$\begin{aligned} l^\infty &= \left\{\hat{x} = (x_n)_{n\geq 1} \in \mathbb{R}^{\mathbb{N}} : \sup_{n\geq 1} |x_n| < \infty\right\}\,, \\ c_0 &= \left\{\hat{x} = (x_n)_{n\geq 1} \in \mathbb{R}^{\mathbb{N}} : x_n \to 0 \text{ as } n \to \infty\right\}\,, \\ X &= \left\{\hat{x} = (x_n)_{n\geq 1} \in \mathbb{R}^{\mathbb{N}} : \text{ there exists } n_0 \in \mathbb{N} \text{ such that } x_n = 0 \text{ for } n \geq n_0\right\}\,. \end{aligned}$$

Evidently, $X \subseteq c_0 \subseteq l^\infty$ and we furnish l^∞ with the supremum norm $\|\hat{x}\| = \sup_{n\in\mathbb{N}} |x_n|$. With this norm, l^∞ is a Banach space, c_0 is a closed subspace hence a Banach space itself, but $\overline{X}^{\|\cdot\|} = c_0$. Let $A_n\colon X \to X$ with $n \geq 1$ and $A\colon X \to X$ be defined by

$$A_n(\hat{x}) = (x, 2x_2, \ldots, nx_n, 0, 0, \ldots)\,, \quad A(\hat{x}) = (kx_k)_{k\geq 1}\,.$$

Then $A_n(\hat{x}) \to A(\hat{x})$ as $n \to \infty$ for all $\hat{x} \in X$ and $\|A_n\|_L = n$ for all $n \in \mathbb{N}$. Precisely, $\{A_n\}_{n\geq 1}$ is pointwise convergent, hence pointwise bounded as well, but $\sup_{n\geq 1} \|A_n\|_L = \infty$ and thus, A is not bounded.

(b) In Theorem 3.2.1(b) the inequality can be strict. Let

$$l^2 = \left\{\hat{x} = (x_n)_{n\geq 1} \subseteq \mathbb{R}^{\mathbb{N}} : \sum_{n\geq 1} x_n^2 < \infty\right\}$$

furnished with the norm,

$$\|\hat{x}\| = \left(\sum_{n\geq 1} x_n^2\right)^{\frac{1}{2}} .$$

With this norm, l^2 becomes a Banach space. In fact it becomes a Hilbert space; see Section 3.5. Let $X = l^2$, $Y = \mathbb{R}$, and consider the bounded linear operators $A_k \colon l^2 \to \mathbb{R}$ with $k \geq 1$ defined by $A_k(\hat{x}) = x_k$ for every $\hat{x} = (x_n)_{n\geq 1} \in l^2$ and for every $k \in \mathbb{N}$. Evidently, $A_k(\hat{x}) \to 0$ as $k \to \infty$ for every $\hat{x} \in l^2$ but $\|A_k\|_L = 1$ for all $n \in \mathbb{N}$.

Theorem 3.2.1 leads to interesting characterizations of bounded sets in a Banach space X and in its dual X^*; see Definition 3.1.45. In the next section we will interpret these results in terms of weak and weak* topologies, respectively.

Proposition 3.2.4. *If X is a normed space and $B \subseteq X$ is nonempty, then B is bounded if and only if $x^*(B) = \{\langle x^*, u\rangle \colon u \in B\} \subseteq \mathbb{R}$ is bounded for every $x^* \in X^*$.*

Proof. $\Longrightarrow$: This follows from the fact that $|\langle x^*, u\rangle| \leq \|x^*\|_* \|u\|$ for every $x^* \in X^*$ and for all $u \in B$. So, if B is bounded, then $\|u\| \leq M$ for some $M > 0$ and for all $u \in B$. Therefore, $x^*(B) \subseteq [-\varrho, \varrho]$ with $\varrho = \|x^*\|_* M$.

$\Longleftarrow$: For every $u \in B$, let $A_u(x^*) = \langle x^*, u\rangle$ for all $x^* \in X^*$ where $\langle\cdot,\cdot\rangle$ denotes the duality brackets for the pair (X^*, X). Then $A_u \in L(X^*, \mathbb{R})$ for all $u \in B$ and by hypothesis,

$$\sup_{u\in B} |A_u(x^*)| = \sup_{u\in B} |\langle x^*, u\rangle| < +\infty .$$

Since X^* is a Banach space (see Corollary 3.1.48), we can apply Theorem 3.2.1 and find $M > 0$ such that

$$|A_u(x^*)| = |\langle x^*, u\rangle| \leq M\|x^*\|_* \quad \text{for all } x^* \in X^* \text{ and for all } u \in B .$$

Because of Proposition 3.1.52 we infer that $\|u\| \leq M$, which shows that B is bounded. $\square$

There is also a "dual" version of this result.

Proposition 3.2.5. *If X is a Banach space and $B^* \subseteq X^*$ is nonempty, then B^* is bounded if and only if $x(B^*) = \{\langle u^*, x\rangle \colon u^* \in B^*\} \subseteq \mathbb{R}$ is bounded for every $x \in X$.*

Proof. $\Longrightarrow$: This is as in the previous proof.

$\Longleftarrow$: For every $u^* \in B^*$, let $A_{u^*}(x) = \langle u^*, x\rangle$ for all $x \in X$. Then $A_{u^*} \in L(X, \mathbb{R})$ for all $u^* \in B^*$ and by hypothesis,

$$\sup_{u^*\in B^*} |A_{u^*}(x)| = \sup_{u^*\in B^*} |\langle u^*, x\rangle| < \infty .$$

Since X is a Banach space, we can apply Theorem 3.2.1 and find $M > 0$ such that

$$|A_{u^*}(x)| = |\langle u^*, x\rangle| \leq M\|x\| \quad \text{for all } x \in X \text{ and for all } u^* \in B^* .$$

Then, Corollary 3.1.48 implies that $\|u^*\|_* \leq M$ for all $u^* \in B^*$. $\square$

Next we will prove the "Open Mapping Theorem," which asserts that a surjective bounded linear operator between Banach spaces is an open map.

In order to prove this theorem, we will need two auxiliary results.

Lemma 3.2.6. *If X, Y are Banach spaces and $A \in L(X, Y)$ surjective, then there exists $\vartheta > 0$ such that for any $\varepsilon > 0$ and $y \in Y$ we find $x \in X$ such that*

$$\|A(x) - y\|_Y \le \varepsilon \quad \textit{and} \quad \|x\|_X \le \frac{1}{\vartheta}\|y\|_Y \,.$$

Proof. Let $B_1^X = \{x \in X\colon \|x\|_X < 1\}$. The surjectivity of A implies that

$$Y = \bigcup_{n\ge 1} A(nB_1^X)\,.$$

Then by the Baire Category Theorem there is $n \in \mathbb{N}$ such that $\operatorname{int}\overline{A(nB_1^X)} \neq \emptyset$. This implies $B_\eta(y_0) \subseteq \overline{A(nB_1^X)}$ for some $\eta > 0$ and $y_0 \in Y$. Here $B_\eta(y_0) = \{y \in Y\colon \|y - y_0\|_Y < \eta\}$. Given $y \in Y$ with $\|y\|_Y < \eta$, let $\{x_k\}_{k\ge1}, \{u_k\}_{k\ge1} \subseteq nB_1^X$ such that

$$A(x_k) \to y_0 \quad \text{and} \quad A(u_k) \to y_0 + y \text{ in } Y \text{ as } k \to \infty\,.$$

Let $v_k = u_k - x_k$ for $k \in \mathbb{N}$. Then

$$A(v_k) \to y \quad \text{in } Y \text{ as } k \to \infty \quad \text{and} \quad \|v_k\|_X < 2n \quad \text{for all } k \in \mathbb{N}\,. \tag{3.2.4}$$

Let $\mathrm{w} \in Y \setminus \{0\}$ and let $z = (\eta/2)\cdot(\mathrm{w}/\|w\|)$. Then $z \in Y$ and $\|z\|_Y < \eta$. From (3.2.4) we know that there exist $\{\tilde{v}_k\}_{k\ge1} \subseteq X$ such that

$$A\left(\tilde{v}_k\right) \to z = \frac{\eta}{2}\frac{\mathrm{w}}{\|w\|_X} \quad \text{in } Y \text{ as } k \to \infty \quad \text{and} \quad \|\tilde{v}_k\|_X < 2n \quad \text{for all } k \in \mathbb{N}\,.$$

Hence,

$$A\left(\frac{2}{\eta}\|w\|_X\tilde{v}_k\right) \to w \quad \text{in } Y \text{ as } k \to \infty\,. \tag{3.2.5}$$

Note that

$$\frac{2}{\eta}\|w\|_X\,\|\tilde{v}_k\|_X < \frac{4n}{\eta}\|w\|_X \quad \text{for all } k \in \mathbb{N}\,.$$

Finally let $\vartheta = \eta/(4n)$ and apply (3.2.5) to obtain the result of the lemma. □

Using this lemma, we can prove the following proposition.

Proposition 3.2.7. *If X, Y are Banach spaces, $B_1^X = \{x \in X\colon \|x\|_X < 1\}$, $B_1^Y = \{y \in Y\colon \|y\|_Y < 1\}$, and $A \in L(X, Y)$ is surjective, then there exists $\delta > 0$ such that $\delta B_1^Y \subseteq A(B_1^X)$.*

Proof. Let $\vartheta > 0$ be as postulated by Lemma 3.2.6. Let $y \in \vartheta B_1^Y$ and $\varepsilon = 1/2\vartheta > 0$. Using Lemma 3.2.6, there exists $x_1 \in X$ such that

$$\|A(x_1) - y\|_Y \le \frac{\vartheta}{2} \quad \text{and} \quad \|x_1\|_X \le \frac{1}{\vartheta}\|y\|_Y < 1\,. \tag{3.2.6}$$

Now consider $y - A(x_1) \in Y$ and $\varepsilon = \vartheta/4$. A new application of Lemma 3.2.6 gives $x_2 \in X$ such that

$$\|A(x_2) - (y - A(x_1))\|_Y \le \frac{\vartheta}{4} \quad \text{and} \quad \|x_2\|_X \le \frac{1}{\vartheta}\|y - A(x_1)\|_Y < \frac{1}{2}\,,$$

see (3.2.6). Suppose that we have produced $\{x_k\}_{k\ge1}^n \subseteq X$ such that

$$\left\| A\left(\sum_{k=1}^n x_k\right) - y \right\|_Y \le \frac{\vartheta}{2^n} \quad \text{and} \quad \|x_k\|_X \le \frac{1}{2^{k-1}} \quad \text{for all } k = 1, \ldots, n\,. \tag{3.2.7}$$

Using Lemma 3.2.6, we obtain $x_{n+1} \in X$ such that

$$\left\| A\left(\sum_{k=1}^{n+1} x_k\right) - y \right\|_Y \le \frac{\vartheta}{2^{n+1}} \text{ and } \|x_{n+1}\|_X \le \frac{1}{\vartheta}\left\| A\left(\sum_{k=1}^{n} x_k\right) - y \right\|_Y \le \frac{1}{2^n}\,; \tag{3.2.8}$$

see (3.2.7). By induction we have a sequence $\{x_n\}_{n\ge1} \subseteq X$ such that (3.2.8) holds.

Let $u_n = \sum_{k=1}^n x_k \in X$ with $n \in \mathbb{N}$. For $m > n$ one gets

$$\|u_m - u_n\|_X = \left\| \sum_{k=n+1}^m x_k \right\|_X \le \sum_{k=n+1}^m \frac{1}{2^k}\,;$$

see (3.2.8). This implies that $\{u_n\}_{n\ge1} \subseteq X$ is a Cauchy sequence. Since X is a Banach space, we obtain $u_n \to u$ in X. Then

$$\|u\|_X \le \sum_{k\ge1} \|x_k\|_X \le \sum_{k\ge1} \frac{1}{2^{k-1}} = 2\,,$$

which shows that $u \in 2B_1^X$. From (3.2.8) it follows $\|A(u_n)-y\|_Y \le \vartheta/2^n$, hence $A(u_n) \to y$ in Y. But we also have $A(u_n) \to A(u)$ in Y. Therefore, $y = A(u)$. Recall that $y \in \vartheta B_1^Y$ is arbitrary and $x \in 2B_1^X$. That means $\vartheta/2B_1^Y \subseteq A(B_1^X)$. Choosing $\delta = \vartheta/2 > 0$, we obtain the assertion of the proposition. □

Remark 3.2.8. This proposition provides estimates for the solutions $x \in X$ of $A(x) = y \in Y$ in terms of y. That the equation $A(x) = y$ always has a solution for all $y \in Y$ is a consequence of the surjectivity of A.

Once we have this proposition, we can easily prove the “Open Mapping Theorem.”

Theorem 3.2.9 (Open Mapping Theorem). *If X, Y are Banach spaces and $A \in L(X, Y)$ is surjective, then A is an open map, that is, it maps open sets in X to open sets in Y.*

Proof. Let $U \subseteq X$ be nonempty and open, and let $x_0 \in U$. Let $V = U - x_0 \in \mathcal{N}(0)$. Then there exists $\xi > 0$ such that $\xi B_1^X \subseteq V$. Using Proposition 3.2.7 we find $\delta > 0$ such that

$$A(V) \supseteq A(\xi B_1^X) = \xi A(B_1^X) \supseteq \xi\delta B_1^Y\,,$$

which implies

$$A(U) = A(V + x_0) = A(x_0) + A(V) \supseteq A(x_0) + \xi\delta B_1^Y\,.$$

The last set is open in Y centered at $A(x_0)$ with a radius of $\xi\delta > 0$. This means that $A(U) \subseteq Y$ is open. $\square$

As an easy consequence of the Open Mapping Theorem we obtain the so-called "Banach Theorem."

Theorem 3.2.10 (Banach Theorem). *If X, Y are Banach spaces and $A \in L(X, Y)$ is a bijection, that is, A is surjective and injective, then $A^{-1} \in L(Y, X)$.*

Proof. First note that $A^{-1}\colon Y \to X$ is a well-defined linear map. Let $U \subseteq X$ be open. Due to Theorem 3.2.9 it follows that $(A^{-1})^{-1}(U) = A(U) \subseteq Y$ is open. Then Proposition 3.1.46 implies that $A^{-1} \in L(Y, X)$. $\square$

Definition 3.2.11. Let X be a vector space and let $\|\cdot\|, |\cdot|$ be two norms on X. We say that the two norms are **equivalent** if there exists a constant $\vartheta \geq 1$ such that

$$\frac{1}{\vartheta}\|x\| \leq |x| \leq \vartheta\|x\| \quad \text{for all } x \in X\,.$$

Remark 3.2.12. This notion defines an equivalence relation on the set of all possible norms on X. The norms $\|\cdot\|, |\cdot|$ on X are equivalent if and only if $\mathrm{id}\colon (X, \|\cdot\|) \to (X, |\cdot|)$ and $\mathrm{id}\colon (X, |\cdot|) \to (X, \|\cdot\|)$ are both bounded linear operators. Two norms are equivalent if and only if they generate the same metric topology on X. Finally, if $\|\cdot\|, |\cdot|$ are equivalent norms, then $(X, \|\cdot\|)$ is a Banach space if and only if $(X, |\cdot|)$ is a Banach space.

Proposition 3.2.13. *If V is a vector space, $\|\cdot\|$ and $|\cdot|$ are two norms on V with V being a Banach space for both norms and there exists $\eta > 0$ such that*

$$|x| \leq \eta\|x\| \quad \textit{for all } x \in V\,,$$

then $\|\cdot\|$ and $|\cdot|$ are equivalent norms on V.

Proof. Let $X = (V, \|\cdot\|)$, $Y = (V, |\cdot|)$, and $A = \mathrm{id}\colon X \to Y$ with $\mathrm{id}(x) = x$ for all $x \in X$. Then $A \in L(X, Y)$ is bijective and we can apply Theorem 3.2.10 and infer that $A^{-1} = \mathrm{id}\colon Y = (V, |\cdot|) \to X = (V, \|\cdot\|)$ is continuous. So, it follows that the norms $\|\cdot\|$ and $|\cdot|$ are equivalent; see Remark 3.2.12. $\square$

Recall that a continuous map $f\colon X \to Y$ has a closed graph $\mathrm{Gr}\, f = \{(x, y) \in X \times Y : y = f(x)\}$. The converse is not true in general. To see this, let $X = Y = \mathbb{R}_+$ and consider the function $f\colon \mathbb{R}_+ \to \mathbb{R}_+$ defined by

$$f(x) = \begin{cases} 0 & \text{if } x = 0\,, \\ \frac{1}{x} & \text{if } x > 0\,. \end{cases}$$

Then $\mathrm{Gr}\, f$ is closed but f is not continuous at $x = 0$. For linear operators between Banach spaces, the situation changes and we have the third basic theorem of linear functional analysis, which is called the "Closed Graph Theorem."

Theorem 3.2.14 (Closed Graph Theorem). *If X, Y are Banach spaces and $A\colon X \to Y$ is a linear operator, then $A \in L(X, Y)$ if and only if* $\operatorname{Gr} A = \{(x, y) \in X \times Y \colon y = A(x)\} \subseteq X \times Y$ *is closed.*

Proof. $\Longleftarrow$: The graph of any continuous map (linear or not) is closed.

$\Longrightarrow$: On X we consider the following norms

$$\|x\| = \|x\|_X \quad \text{and} \quad |x| = \|x\|_X + \|A(x)\|_Y \quad \text{for all } x \in X .$$

Note that $|\cdot|$ is called the **graph norm**. Since $\operatorname{Gr} A \subseteq X \times Y$ is closed, $(X, |\cdot|)$ is a Banach space. Moreover, the inequality $\|x\| \le |x|$ for all $x \in X$ is clearly satisfied. Invoking Proposition 3.2.13, we conclude that $\|\cdot\|$ and $|\cdot|$ are equivalent norms. Thus, there exists $M > 0$ such that $|x| \le M\|x\|$ for all $x \in X$, which implies $\|A(x)\|_Y \le M\|x\|_X$ for all $x \in X$. Then Proposition 3.1.46 finally gives $A \in L(X, Y)$. □

We can apply these results to quotient spaces (see Section 1.3), which in turn will lead us to complemented spaces.

So, let X be a normed vector space and $V \subseteq X$ a closed subspace. We define the equivalence relation $\sim$ on X by

$$x \sim u \quad \text{if and only if} \quad x - u \in V . \tag{3.2.9}$$

Let $[x]$ denote the equivalence class corresponding to $x \in X$. Then $[x] = x + V = \{x + v \colon v \in V\}$ and let X/V be the quotient space, that is, the set of all equivalence classes under $\sim$ defined by (3.2.9). So, the whole subspace V is collapsed in the quotient space X/V and identified with the zero vector. The quotient space X/V becomes a vector space under the following operations:

$$\begin{aligned} &\text{vector addition:} && [x_1] + [x_2] = x_1 + V + x_2 + V = x_1 + x_2 + V , \\ &\text{scalar multiplication:} && \lambda(x + V) = \lambda x + V , \end{aligned}$$

for all $x_1, x_2, x \in X$ and for all $\lambda \in \mathbb{R}$. As we already mentioned, the zero vector in X/V is $0 + V = V$. We can define a norm on X/V by setting

$$\|[x]\| = \inf[\|x + v\| \colon v \in V] .$$

It is easy to check that this is a norm on X/V. Note that

$$\|[x]\| = \inf[\|x + v\| \colon v \in V] = \inf[\|x - v\| \colon v \in V] \quad \text{for all } x \in X . \tag{3.2.10}$$

Proposition 3.2.15. *If X is a normed space and $V \subseteq X$ is a closed subspace, then the following hold:*

(a) $\|x\| \ge \|[x]\|$ *for all $x \in X$;*

(b) *if $x \in X$ and $\varepsilon > 0$, then there exists $u \in X$ with $u \sim x$, that is $[x] = [u]$, such that $\|u\| \le \|[x]\| + \varepsilon$.*

Proof. (a) This is an immediate consequence from (3.2.10).

(b) Let $v \in V$ be such that $\|x - v\| \le d(x, M) + \varepsilon = \|[x]\| + \varepsilon$; see (3.2.10). Set $u = x - v \in [x]$. Then $\|u\| \le \|[x]\| + \varepsilon$. □

Remark 3.2.16. Suppose that $x, y \in X$ be such that $\|[x-y]\| < \vartheta$ for some $\vartheta > 0$. Then according to Proposition 3.2.15(b), there exists $y' \in X$ such that $[x-y] = [x-y']$ and $\|x-y'\| < \vartheta$.

Proposition 3.2.17. *If X is a Banach space and $V \subseteq X$ is a closed subspace, then X/V is a Banach space as well.*

Proof. Suppose that $\{\|[x_n]\|\}_{n\geq 1} \subseteq X/V$ is a Cauchy sequence. By passing to a subsequence if necessary we may assume that

$$\|[x_n - x_{n+1}]\| < \frac{1}{2^n} \quad \text{for all } n \in \mathbb{N}\,.$$

According to Remark 3.2.16 we can find $x_2' \in X$ such that $[x_1 - x_2] = [x_1 - x_2']$ and $\|x_1 - x_2'\| < 1/2$. Then $[x_2] = [x_2']$ and so we may assume that $x_2' = x_2$. Now again by Remark 3.2.16, there exists $x_3' \in X$ such that $[x_2 - x_3] = [x_2 - x_3']$ and $\|x_2 - x_3'\| < 1/2^2$. As for x_2', we may assume that $x_3' = x_3$. Inductively we obtain that $\|x_n - x_{n+1}\| < 1/2^n$ for all $n \in \mathbb{N}$. So, $\{x_n\}_{n\geq 1} \subseteq X$ is a Cauchy sequence and we may say that $x_n \to x \in X$. Then, Proposition 3.2.15(a) gives

$$\|[x_n] - [x]\| = \|[x_n - x]\| \leq \|x_n - x\|\,.$$

Hence, $[x_n] \to [x]$ and so X/V is a Banach space. □

Remark 3.2.18. In fact there is a kind of converse to the result above. Namely, if X is a normed space, $V \subseteq X$ is a closed subspace, and both V and X/V are complete, then X is a Banach space; see Problem 3.10.

Definition 3.2.19. Let X be a normed space and let $V \subseteq X$ be a closed subspace. The map $p\colon X \to X/V$ defined by $p(x) = [x]$ is called the **quotient map**.

Proposition 3.2.20. *If X is a normed space and $V \subseteq X$ is a closed subspace, then the quotient map $p \in L(X, X/V)$ is surjective and open, and $N(p) = V$, and if $V \neq X$, then $\|p\|_L = 1$.*

Proof. We only need to show that p is open. Let $U \subseteq X$ be open, $x \in U$, and let $B_1^X = \{u \in X\colon \|u\| < 1\}$. Then we find $\vartheta > 0$ such that $x + \vartheta B_1^X \subseteq U$, hence $p(x) + \vartheta p(B_1^X) \subseteq p(U)$. We claim that $p(B_1^X) = B_1^{X/V} = \{[x] \in X/V\colon \|[x]\| < 1\}$. To see this, let $x \in B_1^X$. Then $\|p(x)\| = \|[x]\| \leq \|x\| < 1$; see Proposition 3.2.15(a). Therefore, $p(B_1^X) \subseteq B_1^{X/V}$. On the other hand if $[u] \in B_1^{X/V}$, then there is $u' \in B_1^X$ such that $p(u') = [u'] = [u]$ (see Proposition 3.2.15(b)), and so $B_1^{X/V} \subseteq p(B_1^{X/V})$. Thus finally $p(B_1^X) = B_1^{X/V}$ and so $p(x) + \vartheta B_1^{X/V} \subseteq p(U)$. Hence, p is open. □

Proposition 3.2.21. *If X, Z are normed spaces, $V \subseteq X$ is a closed subspace and $A \in L(X, Z)$ satisfies $N(A) = \{x \in X\colon A(x) = 0\} \supseteq V$, then there exists a unique $\hat{A} \in L(X/V, Z)$ such that $A = \hat{A} \circ p$.*

Proof. The operator $\hat{A}: X/V \to Z$ defined by $\hat{A}([x]) = A(x)$ is well-defined since $V \subseteq N(A)$. Clearly $\hat{A}$ is linear and

$$\left\|\hat{A}([x])\right\|_Z = \|A(x+v)\|_Z \le \|A\|_L \|x+v\|_X \quad \text{for all } v \in V ,$$

since $V \subseteq N(A)$. Hence,

$$\left\|\hat{A}([x])\right\|_Z \le \|A\|_L \inf\left[\|x+v\|_X : v \in V\right] = \|A\|_L \|[x]\| .$$

This shows that $\hat{A} \in L(X/V, Z)$ and $A = \hat{A} \circ p$. Clearly $\hat{A}$ is unique. □

Remark 3.2.22. This is a factorization theorem and it can be better remembered if we use the following figure:

$$\begin{tikzcd} X \arrow[r, "A"] \arrow[d, "p"'] & Z \\ X/V \arrow[ur, "\hat{A}"'] & \end{tikzcd}$$

Proposition 3.2.23. *If X, Z are Banach spaces, $A \in L(X, Z)$ is surjective and $V = N(A) = \{x \in X : A(x) = 0\}$, then X/V and Z are isomorphic, that is, there exists $\mathfrak{L}: X/V \to Z$ being a linear, continuous bijection with a continuous inverse.*

Proof. From Proposition 3.2.21 we know that there exists a unique $\hat{A} \in L(X/V, Z)$ such that $A = \hat{A} \circ p$. If $\hat{A}([x]) = \hat{A}([u])$, then $A(x) = A(u)$ and so $x - u \in N(A)$, which means that $\hat{A}$ is one-to-one. Let $z \in Z$ and recall that A is surjective. Then we can find $x \in X$ such that $A(x) = z$. Thus, $\hat{A}([x]) = z$, which implies that $\hat{A}$ is surjective, that is, a bijection. Invoking Theorem 3.2.10, we conclude that $\hat{A}$ is an isomorphism. □

Definition 3.2.24. Let X be a normed space and let $D \subseteq X$. The **annihilator** of D is defined by

$$D^\perp = \left\{x^* \in X^* : \langle x^*, d\rangle = 0 \text{ for all } d \in D\right\} .$$

Evidently, $D^\perp$ is a closed vector subspace of X^*.

Using this notion we can characterize the dual of a quotient space.

Proposition 3.2.25. *If X is a normed space and $V \subseteq X$ is a closed subspace, then $(X/V)^*$ and $V^\perp$ are isometrically isomorphic.*

Proof. Let $l \in (X/V)^*$ and let $x^* = l \circ p: X \to \mathbb{R}$. Then $x^* \in X^*$ and $x^*|_V = 0$. So, $x^* \in V^\perp$. Conversely, let $x^* \in V^\perp$. Then according to Proposition 3.2.21, there exists a unique $l \in (X/V)^*$ such that $x^* = l \circ p$. So, the linear map $\xi: (X/V)^* \to V^\perp$ defined by $\xi(l) = l \circ p$ is a bijection and

$$l([x]) = \langle \xi(l), x\rangle = \langle \xi(l), x+v\rangle \le \|\xi(l)\|_* \|x+v\| \quad \text{for all } v \in V .$$

Thus

$$\|l\|_{(X/V)^*} \le \|\xi(l)\|_* . \tag{3.2.11}$$

On the other hand, thanks to Proposition 3.2.15(a), one gets

$$\langle \xi(l), x\rangle = l([x]) \le \|l\|_{(X/V)^*}\|[x]\| \le \|l\|_{(X/V)^*}\|x\| .$$

This gives

$$\|\xi(l)\|_* \le \|l\|_{(X/V)^*} . \tag{3.2.12}$$

From (3.2.11) and (3.2.12) we infer that $\|\xi(l)\|_* = \|l\|_{(X/V)^*}$ and so ξ is an isometric isomorphism. □

We present some additional properties of closed subspaces in Banach spaces.

Proposition 3.2.26. *If X is a Banach space and $V, W \subseteq X$ are closed subspaces of X such that $V + W$ is closed, then there exists $\hat{c} > 0$ such that every $u \in V + W$ admits a decomposition $u = v + w$ with $v \in V$ and $\mathrm{w} \in W$ as well as*

$$\|v\| \le \hat{c}\|u\| \quad \textit{and} \quad \|w\| \le \hat{c}\|u\| .$$

Proof. We consider the Cartesian product $V \times W$ furnished with the norm $\|(v, w)\| = \|v\| + \|w\|$. Moreover, we consider on $V + W$ the norm inherited from X. Let $A\colon V \times W \to V + W$ be defined by $A((v, w)) = v + w$. Evidently, $A \in L(V \times W, V + W)$ and is surjective. Since $V \times W$ and $V + W$ are Banach spaces, invoking the Open Mapping Theorem (see Theorem 3.2.9), there exists $c > 0$ such that $u \in V + W$ with $\|u\| < c$ implies $u = v + w$ with $v \in V$, $w \in W$ and $\|v\| + \|w\| < 1$. By the homogeneity, there holds for every $u \in V + W$ that $u = v + w$ with $v \in V$, $w \in W$ and $\|v\| + \|w\| \le 1/c\|u\|$. Then for $\hat{c} = c^{-1}$ we have the result. □

Definition 3.2.27. Let X be a normed space. A closed subspace $V \subseteq X$ is called **complemented** (or we say that it admits a **topological complement**), if there exists a closed subspace $W \subseteq X$ such that $V \cap W = \{0\}$ and $X = V + W$ (we write $X = V \oplus W$). Then we say that V and W are **complementary** subspaces of X.

The next results shows that finite dimensional subspaces or subspaces with finite codimension, are complemented.

Proposition 3.2.28. *If X is a normed space and $V \subseteq X$ is a closed subspace such that* $\dim V < \infty$ *or* $\dim (X/V) < \infty$, *then V is complemented.*

Proof. Let $n = \dim V < \infty$ and let $\{e_k\}_{k=1}^n$ be a basis of V. According to Proposition 3.1.49, there exists $\{e_m^*\}_{m=1}^n \subseteq X^*$ such that

$$\langle e_m^*, e_k\rangle = \delta_{mk} = \begin{cases} 1 & \text{if } m = k ,\\ 0 & \text{if } m \neq k .\end{cases}$$

Let $W = \{x \in X\colon \langle e_m^*, x\rangle = 0 \text{ for all } m \in \{1, \dots, n\}\}$. Clearly $W \subseteq X$ is a closed subspace and $X = V \oplus W$ since $x - \sum_{m=1}^n \langle e_m^*, x\rangle e_m \in W$ for all $x \in X$.

Next let $n = \dim(X/V) < \infty$. We choose $\{x_k\}_{k=1}^n \subseteq X$ such that $\{[x_k]\}_{k=1}^n$ is a basis of X/V. Then $W = \mathrm{span}\{x_k\}_{k=1}^n \subseteq X$ is closed (see Corollary 3.1.19) and satisfies $X = V \oplus W$. □

Remark 3.2.29. It is not true that every closed subspace of an infinite dimensional Banach space is complemented. For example, $c_0 \subseteq l^\infty$ is a closed subspace, but it is not complemented; see Phillips [237]. In fact a result due to Lindenstrauss-Tzafriri [201] says that every Banach space that is not a Hilbert space admits a closed subspace that is not complemented.

3.3 Weak and Weak* Topologies

In this section we study the weak topology on a normed space X and the weak* topology on X^*, which is always a Banach space; see Corollary 3.1.48. These are locally convex topologies and are special cases of the weak topologies introduced in Definition 1.3.1 when $Y_i = \mathbb{R}$ for all $i \in I$ and $\{f_i\}_{i\in I} = X^*$ (for the weak topology) as well as $\{f_i\}_{i\in I} = X$ (for the weak* topology).

The strong (norm) topology on an infinite dimensional normed space is too strong for many purposes. In particular, note that a strongly compact set in an infinite dimensional normed space has an empty interior. Indeed, if this is not the case, then the space is locally compact, hence by Proposition 3.1.24, it is finite dimensional, a contradiction. The main result of this section is "Alaoglu's Theorem" (see Theorem 3.3.38), which says that the unit ball in the dual space X^* is compact for the relative weak* topology. This result is reminiscent of the classical Heine–Borel Theorem; see Theorem 1.5.38.

Definition 3.3.1. Let X be a normed space. The **weak topology** on X is the weakest topology on X with respect to which every element $x^* \in X^*$ ($x^* : X \to \mathbb{R}$ being norm continuous and linear) is continuous. We denote the weak topology by $\mathrm{w}(X, X^*)$ or simply by w.

Remark 3.3.2. As we already mentioned, the w-topology is a particular case of the weak (initial) topology introduced in Definition 1.3.1 when the initial space is X (the normed space), $Y_i = \mathbb{R}$ for all $i \in I$, $I = X^*$ and $f_{x^*} : X \to \mathbb{R}$ with $x^* \in X^* = I$ is the linear functional $f_{x^*}(x) = \langle x^*, x\rangle$. Recall that $\langle\cdot,\cdot\rangle$ denotes the duality brackets for the pair (X^*, X). Evidently the weak topology w is weaker than the norm (metric) topology on X.

Proposition 3.3.3. *The weak topology* $\mathrm{w}(X, X^*)$ *is Hausdorff.*

Proof. From Corollary 3.1.61, we know that $\{f_{x^*}\}_{x^*\in X^*=I}$ is separating and so Proposition 1.3.7 implies that $\mathrm{w}(X, X^*)$ is Hausdorff. □

The weak topology on X is clearly linear, that is, both operations, vector addition and scalar multiplication, are continuous. Moreover, it is locally convex; see Theorem 3.1.40. Note that $\mathbb{R}$ is regular and recall that regularity is hereditary and topological (see

Proposition 1.2.10), and it is preserved in Cartesian products; see Proposition 1.3.13. Therefore, we can improve Proposition 3.3.3 in the following way.

Proposition 3.3.4. *The weak topology* $\mathrm{w}(X, X^*)$ *is regular; see Definition* 1.2.7.

Remark 3.3.5. In fact for the same reasons, $\mathrm{w}(X, X^*)$ is completely regular; see Definition 1.2.19.

The linearity of the weak topology implies that in order to describe it we only need to specify a local basis at the origin. Then by translation we obtain a local basis at any other point. Remark 1.3.2 allows us to give a precise description of the local basis at the origin.

Proposition 3.3.6. *A typical basic weak neighborhood of the origin is given by*

$$U(0; x_1^*, \dots, x_n^*, \varepsilon) = \{x \in X \colon |\langle x_k^*, x\rangle| < \varepsilon \text{ for all } k = 1, \dots, n\}$$

with $\{x_k^*\}_{k=1}^n \subseteq X^*$, $n \in \mathbb{N}$ *and* $\varepsilon > 0$. *As* $\varepsilon > 0$, $n \in \mathbb{N}$ *and* $\{x_k^*\}_{k=1}^n$ *vary, we cover a local basis for the weak topology at the origin. At any other point* $x_0 \in X$ *the local basis consists of sets of the form*

$$x_0 + U(0; x_1^*, \dots, x_n^*, \varepsilon) = \{x \in X \colon |\langle x_k^*, x - x_0\rangle| < \varepsilon \text{ for all } k = 1, \dots, n\}\,.$$

In infinite dimensional normed spaces the weak topology and the strong (norm) topology never coincide. To see this we will need to recall some simple facts from linear algebra. The first is an algebraic variant of the factorization result stated in Proposition 3.2.21.

Lemma 3.3.7. *If* X, Y, Z *are vector spaces,* $f \colon X \to Z$ *and* $g \colon X \to Y$ *are linear maps and* $N(g) \subseteq N(f)$, *where* $N(g) = \{x \in X \colon g(x) = 0\}$, $N(f) = \{x \in X \colon f(x) = 0\}$, *then there exists a linear map* $\xi \colon Y \to Z$ *such that* $f = \xi \circ g$.

Proof. Let $\xi \colon g(X) \to Z$ be defined by $\xi(g(x)) = f(x)$ for all $x \in X$. This linear map is well-defined since if $g(x_1) = g(x_2)$, then $x_1 - x_2 \in N(g) \subseteq N(f)$ and so $f(x_1) = f(x_2)$. Extending ξ to a linear map on all of Y gives $f = \xi \circ g$. □

Using this lemma, we can prove the second auxiliary result from linear algebra.

Lemma 3.3.8. *If* X *is a vector space,* $f, f_1, \dots, f_n \colon X \to \mathbb{R}$ *are linear maps and* $\bigcap_{k=1}^n N(f_k) \subseteq N(f)$, *then* f *is a linear combination of the* f_k'*s.*

Proof. Let $X = X$, $Y = \mathbb{R}^n$, $Z = \mathbb{R}$, $f = f$ and $g = (f_k)_{k=1}^n$ and apply Lemma 3.3.7 to produce a linear functional $\xi \colon \mathbb{R}^n \to \mathbb{R}$ such that $f = \xi \circ g$. Then $\xi(\hat{y}) = \sum_{k=1}^n \lambda_k y_k$ with $\lambda_1, \dots, \lambda_n \in \mathbb{R}$, $\hat{y} = (y_k)_{k=1}^n \in \mathbb{R}^n$. It follows that $f(x) = \sum_{k=1}^n \lambda_k f_k(x)$ for all $x \in X$. □

These auxiliary results lead to the following important observations about the weak topology.

Proposition 3.3.9. *If* X *is an infinite dimensional normed space and* $U \subseteq X$ *is nonempty and* w*-open, then* U *is not bounded.*

Proof. Translating U if necessary, we may assume that $0 \in U$. By Proposition 3.3.6 there exist $x_1^*, \dots, x_n^* \in X^*$ and $\varepsilon > 0$ such that $U(0; x_1^*, \dots, x_n^*, \varepsilon) \subseteq U$. Note that $V = \bigcap_{k=1}^n N(x_k^*) \subseteq U$. Of course, V is a vector subspace of X and we claim that $V \neq \{0\}$. Indeed, if $V = \{0\}$, then it holds that $V \subseteq N(x^*)$ for all $x^* \in X^*$ and so Lemma 3.3.8 implies that x^* is a linear combination of the x_k^*'s. This means that $X^* = \operatorname{span}\{x_k^*\}_{k=1}^n$ and so X^* is finite dimensional, and hence X is finite dimensional, a contradiction. Therefore U is not bounded since it contains V. $\square$

Remark 3.3.10. This proposition implies that weakly open sets are large. In particular, if $x \in V$ (see the previous proof), $x \neq 0$, then $\mathbb{R}x \subseteq U$. Therefore the open unit ball $B_1 = \{x \in X : \|x\| < 1\}$ is never w-open in an infinite dimensional normed space X.

Corollary 3.3.11. *If X is an infinite dimensional normed space, then the weak and strong (norm) topology do not coincide.*

In finite dimensional normed spaces, which are then of course Banach spaces, the two topologies coincide.

Proposition 3.3.12. *If X is a finite dimensional normed space, then the weak topology and the strong (norm) topology coincide.*

Proof. By definition, the weak topology is smaller than the strong topology. So, in order to prove the proposition, it suffices to show that every strongly open set is weakly open. Let $x_0 \in X$ and let U be a strongly open set containing x_0. Then there exists $\varrho > 0$ such that

$$B_\varrho(x_0) = \{x \in X : \|x - x_0\| < \varrho\} \subseteq U\,. \tag{3.3.1}$$

Let $\{e_k\}_{k=1}^n$ be a basis for X with $\|e_k\| = 1$ for all $k = 1, \dots, n$. Then every $x \in X$ admits an expression $x = \sum_{k=1}^n \lambda_k e_k$ with $\lambda_k \in \mathbb{R}$. For every $k = 1, \dots, n$ the coordinate map $x \to \lambda_k$, denoted by x_k^*, is linear and continuous for every $k = 1, \dots, n$. We consider $U(x_0; x_1^*, \dots, x_n^*, \varrho/n)$ being the basic weak neighborhood of x_0 determined by these coordinate maps. Then it follows

$$\|x - x_0\| \le \sum_{k=1}^n |\langle x_k^*, x - x_0\rangle| \le n\frac{\varrho}{n} = \varrho \quad \text{for all } x \in U\left(x_0; x_1^*, \dots, x_n^*, \frac{\varrho}{n}\right)\,,$$

which implies

$$U\left(x_0; x_1^*, \dots, x_n^*, \frac{\varrho}{n}\right) \subseteq B_\varrho(x_0) \subseteq U\,,$$

see (3.3.1). That means that U is w-open and so the two topologies coincide. $\square$

In what follows, we denote the convergence in the weak topology by $\xrightarrow{w}$ and the convergence in the strong (norm) topology by $\to$.

Proposition 3.3.13. *If X is a normed space and $\{x_\alpha\}_{\alpha\in I} \subseteq X$ is a net, then the following hold:*

(a) $x_\alpha \xrightarrow{w} x$ *if and only if* $\langle x^*, x_\alpha\rangle \to \langle x^*, x\rangle$ *for all* $x^* \in X^*$*;*
(b) $x_\alpha \to x$ *implies* $x_\alpha \xrightarrow{w} x$*;*

(c) $x_\alpha \xrightarrow{w} x$ *implies* $\|x\| \le \liminf_{\alpha\in I} \|x_\alpha\|$ *and a weakly convergent sequence is norm bounded;*
(d) $x_\alpha \xrightarrow{w} X$ *in* X *and* $x_\alpha^* \to x^*$ *in* X^* *imply* $\langle x_\alpha^*, x_\alpha\rangle \to \langle x^*, x\rangle$.

Proof. (a) This is a consequence of Proposition 1.3.3.

(b) For every $x^* \in X^*$, we have

$$|\langle x^*, x_\alpha\rangle - \langle x^*, x\rangle| = |\langle x^*, x_\alpha - x\rangle| \le \|x^*\|_* \|x_\alpha - x\| \to 0 .$$

(c) Suppose that there is a sequence $\{x_n\}_{n\in\mathbb{N}} \subseteq X$ such that $x_n \xrightarrow{w} x$. Then, it follows $\langle x^*, x_n - x\rangle \to 0$ for all $x^* \in X^*$, which implies $\sup_{n\in\mathbb{N}} |\langle x^*, x_n - x\rangle| < \infty$. Taking Theorem 3.2.1 into account there exists $M > 0$ such that $\|x_n\| \le M$ for all $n \in \mathbb{N}$.

Evidently we may assume that $x \ne 0$. According to Proposition 3.1.50, there exists $\hat{x}^* \in X^*$ with $\|\hat{x}^*\|_* = 1$ such that $\langle \hat{x}^*, x\rangle = \|x\|$. So, $\|x\| = \lim_{\alpha\in I} |\langle \hat{x}^*, x_\alpha\rangle|$. Then, for given $\varepsilon > 0$ we can find $\alpha_0 = \alpha_0(\varepsilon) \in I$ such that

$$\|x\| - \varepsilon \le |\langle x^*, x_\alpha\rangle| \le \|x_\alpha\| \quad \text{for all } \alpha \ge \alpha_0 .$$

Hence, $\|x\| \le \liminf_{\alpha\in I} \|x_\alpha\|$.

(d) Applying part (c), we derive, for some $M > 0$ and for every $\alpha \in I$, that

$$\begin{aligned} |\langle x_\alpha^*, x_\alpha\rangle - \langle x^*, x\rangle| &\le |\langle x_\alpha^* - x^*, x_\alpha\rangle| + |\langle x^*, x_\alpha - x\rangle| \\ &\le \|x_\alpha^* - x^*\|_* M + |\langle x^*, x_\alpha - x\rangle| \to 0 . \end{aligned}$$

Thus, $\langle x_\alpha^*, x_\alpha\rangle \to \langle x^*, x\rangle$. ☐

Remark 3.3.14. We emphasize that the boundedness in Proposition 3.3.13(c) holds only for weakly convergent sequences and it fails for nets. Indeed, every infinite dimensional normed space admits a net $\{x_\alpha\}_{\alpha\in I} \subseteq X$ such that $x_\alpha \xrightarrow{w} 0$ in X and $\sup[\|x_\eta\| : \eta \ge a, \eta \in I] = +\infty$. To see this let E denote the collection of all nonempty finite subsets of X^*. This set is directed by the set inclusion, that is, if $\alpha, \eta \in E$, $\alpha \ge \eta$ if and only if $\alpha \supseteq \eta$. For each $\alpha = (x_k^*)_{k=1}^n \in E$ there exist some $x_\alpha \in \bigcap_{k=1}^n N(x_k^*)$ such that $\|x_\alpha\| = \operatorname{card} \alpha$. The net $\{x_\alpha\}_{\alpha\in E}$ has the desired properties.

The weak topology is not metrizable in general and so sequences are not adequate to describe it. In fact we have the following result.

Proposition 3.3.15. *If X is a normed space and the weak topology on X is metrizable, then X is finite dimensional.*

Proof. Since the weak topology is metrizable, it is first countable. Hence, we can find a sequence $\{x_n^*\}_{n\ge1} \subseteq X^*$ such that for any given $U \in \mathcal{N}_w(0)$ being the filter of weak neighborhoods of the origin, there exist $\varepsilon \in (0,1) \cap \mathbb{Q}$ and $n_U \in \mathbb{N}$ such that

$$U\left(0; x_1^*, \dots, x_{n_U}^*, \varepsilon\right) \subseteq U . \tag{3.3.2}$$

For each $x^* \in X^*$, we have $U(0; x^*, 1) \in \mathcal{N}_w(0)$ and so by (3.3.2) it follows that

$$U\left(0; x_1^*, \dots, x_{n(U(0;x^*,1))}^*, \varepsilon\right) \subseteq U(0; x^*, 1) .$$

Then

$$\bigcap_{k=1}^{n(U(0;x^*,1))} N(x_k^*) \subseteq N(x^*) \,,$$

which, due to Lemma 3.3.8, results in

$$x^* \in \operatorname{span}\{x_k^*\}_{k=1}^{n(U(0;x^*,1))} \,.$$

Since $x^* \in X^*$ is arbitrary, it follows that $X^* = \bigcup_{k\geq 1} V_k$ with each V_k being finite dimensional. Recall that X^* is a Banach space. So, invoking Corollary 1.5.67 we see that $\operatorname{int} V_{k_0} \neq \emptyset$ for some $k_0 \in \mathbb{N}$. This means that $V_{k_0} = X^*$ and so X^* is finite dimensional. Hence X is finite dimensional. □

In what follows, we define for a normed space

$$\overline{B}_1 = \{x \in X \colon \|x\| \leq 1\} \quad \text{and} \quad \partial B_1 = \{x \in X \colon \|x\| = 1\} \,.$$

Both sets are strongly closed. However, the situation changes for a weak topology. This is another illustration of the character of weak topology compared with strong (norm) topology, in the case of infinite dimensional normed spaces, of course; see Proposition 3.3.12.

Proposition 3.3.16. *If X is an infinite dimensional normed space, then $\overline{\partial B_1}^w = \overline{B}_1$.*

Proof. First we point out that the set $\overline{B}_1$ is w-closed. Indeed, if $\{x_\alpha\}_{\alpha\in I} \subseteq \overline{B}_1$ is a net such that $x_\alpha \xrightarrow{w} x$, then from Proposition 3.3.13(c) one gets $\|x\| \leq \liminf_{\alpha\in I} \|x_\alpha\| \leq 1$. Hence $x \in \overline{B}_1$ and so $\overline{B}_1$ is w-closed. It follows that

$$\overline{\partial B_1}^w \subseteq \overline{B}_1 \,. \tag{3.3.3}$$

Next let $x_0 \in B_1 = \{x \in X \colon \|x\| < 1\}$ and take $U \in \mathcal{N}_w(x_0)$ being the filter of weak neighborhoods of x_0. We may always assume that U is basic, that is,

$$U = U\left(x_0; x_1^*, \ldots, x_n^*, \varepsilon\right) \quad \text{with} \quad \{x_k^*\}_{k=1}^n \subseteq X^* \quad \text{and} \quad \varepsilon > 0 \,.$$

We fix $u \in \bigcap_{k=1}^n N(x_k^*)$, $u \neq 0$ (see the proof of Proposition 3.3.9) and consider the function $\xi \colon \mathbb{R}_+ \to \mathbb{R}_+$ defined by $\xi(\lambda) = \|x_0 + \lambda u\|$ for all $\lambda \geq 0$. We see that ξ is continuous, $\xi(0) < 1$ and $\lim_{\lambda\to+\infty} \xi(\lambda) = +\infty$. So, by Bolzano's Theorem there exists $\lambda_0 > 0$ such that $\xi(\lambda_0) = \|x_0 + \lambda_0 u\| = 1$, hence $x_0 + \lambda_0 u \in \partial B_1$.

Moreover, for every $k = 1, \ldots, n$ we obtain $|\langle x_k^*, x_0 + \lambda_0 u - x_0\rangle| = 0$, which shows that $x_0 + \lambda_0 u \in \partial B_1 \cap U$. Therefore it follows that $B_1 \subseteq \overline{\partial B_1}^w$ and since the weak topology is smaller we infer that $\overline{B_1} \subseteq \overline{B_1}^w \subseteq \overline{\partial B_1}^w$. Finally, because of (3.3.3), we conclude that $\overline{B}_1 = \overline{\partial B_1}^w$. □

Remark 3.3.17. Consider the infinite dimensional Banach space $l^1 = \{\hat{x} = (x_n)_{n\geq 1} \in \mathbb{R}^{\mathbb{N}} \colon \sum_{n\geq 1} |x_n| < \infty\}$ which is called the space of all absolutely summable sequences in $\mathbb{R}$. One can show that weak and norm convergent sequences coincide in l^1. This is known as "Schur's Theorem" and its proof can be found in the book of Diestel [79, p. 85].

Our previous discussion of the weak topology has established that in an infinite dimensional normed space there are many more strongly closed sets than there are weakly closed sets. In the next theorem we show that for convex sets both notions agree. This is a remarkable result since a purely algebraic property, namely convexity, leads to a purely topological conclusion, namely that weak and strong closures coincide. The result is known as “Mazur’s Theorem.”

Theorem 3.3.18 (Mazur’s Theorem). *If X is a normed space and $C \subseteq X$ is convex, then $\overline{C} = \overline{C}^w$.*

Proof. Since the strong (norm) topology is larger than the weak topology we directly obtain

$$\overline{C} \subseteq \overline{C}^w . \tag{3.3.4}$$

Arguing by contradiction suppose that the inclusion in (3.3.4) is strict. That means there exists $x_0 \in \overline{C}^w \setminus \overline{C}$. Invoking the Strong Separation Theorem (see Theorem 3.1.60), we find $x^* \in X^* \setminus \{0\}$ and $\varepsilon > 0$ such that

$$\langle x^*, x_0 \rangle + \varepsilon \leq \langle x^*, u \rangle \quad \text{for all } u \in \overline{C} .$$

We set $\vartheta = \inf[\langle x^*, u\rangle : u \in \overline{C}]$ and $U = \{x \in X : \langle x^*, x \rangle < \vartheta\}$. Evidently $U \in \mathcal{N}_w(x_0)$ with $\mathcal{N}_w(x_0)$ being the filter of weak neighborhoods of x_0. Then $U \cap C = \emptyset$ and so $x_0 \notin \overline{C}^w$, a contradiction. Therefore from (3.3.4) we conclude that $\overline{C} = \overline{C}^w$. □

Corollary 3.3.19. *If X is a normed space and $V \subseteq X$ is a vector subspace, then $\overline{V} = \overline{V}^w$.*

Corollary 3.3.20. *If X is a normed space and $x_n \xrightarrow{w} x$, then there exists a sequence $\{u_n\}_{n\geq 1} \subseteq X$ consisting of convex combinations of the x_n’s such that $u_n \to x$ in X.*

Proof. Let $C = \overline{\operatorname{conv}}\{x_n\}_{n\geq 1}$. Theorem 3.3.18 gives $x \in \overline{C}^w = \overline{C}$ and so $x \in \overline{\operatorname{conv}}\{x_n\}_{n\geq 1}$. The result follows. □

Remark 3.3.21. This corollary known as “Mazur’s Lemma” says that if $x_n \xrightarrow{w} x$, then for a given $\varepsilon > 0$ there exist $t_1, \ldots, t_m \geq 0$ such that $\sum_{k=1}^m t_k = 1$ and $\|x - \sum_{k=1}^m t_k x_k\| < \varepsilon$.

Corollary 3.3.22. *If X is a normed space and $C \subseteq X$ is convex, then C is closed if and only if C is* w-*closed.*

The next result is a consequence of the projective character of the weak topology.

Proposition 3.3.23. *If X, Y are normed spaces, then $A \in L(X, Y)$ if and only if A is weak-to-weak continuous.*

Proof. Note that $A \in L(X, Y)$ if and only if $A(\overline{B}_1^X) \subseteq Y$ is bounded with $\overline{B}_1^X = \{x \in X : \|x\|_X \leq 1\}$; see Proposition 3.1.46. From Proposition 3.2.4 we know that $A(\overline{B}_1^X) \subseteq Y$ is bounded if and only if $y^*(A(\overline{B}_1^X)) \subseteq \mathbb{R}$ is bounded for every $y^* \in Y^*$. But a linear functional on a normed space is continuous if and only if it is weakly continuous. Invoking Proposition 1.3.4 we conclude that A is continuous if and only if it is weak-to-weak continuous. □

From Proposition 3.2.4 we have the following result about bounded sets.

Proposition 3.3.24. *If X is a normed space and $A \subseteq X$, then A is bounded if and only if A is* w*-bounded.*

Remark 3.3.25. We can formulate this result in a more general form. We say that a locally convex topology τ on X is compatible with the pair (X^*, X) if and only if $(X_\tau)^* = X^*$. Then $A \subseteq X$ is bounded if and only if A is τ-bounded. In short, we can say that boundedness is duality invariant.

On the dual space X^* we can define two topologies. The first is the usual strong (metric) topology induced by the norm and the second is the weak topology $\mathrm{w} = \mathrm{w}(X^*, X^{**})$. Recall that the weak topology w is the weakest topology on X^* such that $(X^*_{\mathrm{w}})^* = X^{**}$. There is a third topology that we can define known as the w^***-topology**. This topology makes sense only on dual spaces.

Definition 3.3.26. Let X be a normed space and X^* is the topological dual, that is, $X^* = L(X, \mathbb{R})$. The **weak* topology** on X^* is the weakest topology w^* on X^* such that $(X^*_{\mathrm{w}^*})^* = X$. Consider now the linear functional $f_x \colon X^* \to \mathbb{R}$ defined by $f_x(x^*) = \langle x^*, x\rangle$. Then the weak* topology is the weakest topology on X^* making the collection $\{f_x\}_{x\in X}$ of maps from X^* into $\mathbb{R}$ continuous. The weak* topology on X^* is denoted by w^* or by $\mathrm{w}(X^*, X)$.

Remark 3.3.27. Since $X \subseteq X^{**}$ it is clear that $\mathrm{w}^* \subseteq w$, that is, the weak* topology has fewer open (resp. closed) sets than the weak topology.

Similarly to the weak topology (see Proposition 3.3.4 and Remark 3.3.5), we have the following result.

Proposition 3.3.28. *If X is a normed space, then X^*, equipped with the weak* topology, is a completely regular locally convex space.*

Moreover, we obtain the next two propositions as a consequence from Proposition 3.3.12.

Proposition 3.3.29. *If X is a normed space, then the* w^**, the* w*, and the strong topologies on X^* coincide if and only if X is finite dimensional.*

Proposition 3.3.30. *If X is a normed space, then the basic weak* neighborhood of the origin has the form*

$$U(0; x_1, \ldots, x_n, \varepsilon) = \{x^* \in X^* : |\langle x^*, x_k\rangle| < \varepsilon \text{ for all } k = 1, \ldots, n\}$$

with $\{x_k\}_{k=1}^n \subseteq X$, $n \in \mathbb{N}$ and $\varepsilon > 0$. Since the weak topology is linear, we obtain the local basis at any other point by translation.*

The proof of Proposition 3.3.13 gives the following result. In what follows we denote the convergence in weak* topology by $\xrightarrow{\mathrm{w}^*}$.

Proposition 3.3.31. *If X is a normed space and $\{x_\alpha^*\}_{\alpha\in I} \subseteq X^*$ is a net, then the following hold:*
(a) $x_\alpha^* \xrightarrow{w^*} x^*$ *if and only if* $\langle x_\alpha^*, x\rangle \to \langle x^*, x\rangle$ *for all* $x \in X$;
(b) $x_\alpha^* \to x^*$ *or* $x_\alpha^* \xrightarrow{w} x^*$ *implies* $x_\alpha^* \xrightarrow{w^*} x^*$;
(c) $x_\alpha^* \xrightarrow{w^*} x^*$ *implies* $\|x^*\|_* \le \liminf_{\alpha\in I} \|x_\alpha^*\|_*$ *and every weakly* convergent sequence is norm bounded;*
(d) $x_\alpha^* \xrightarrow{w^*} x^*$ *and* $x_\alpha \to x$ *in* X *imply* $\langle x_\alpha^*, x_\alpha\rangle \to \langle x^*, x\rangle$.

Remark 3.3.32. From the definition of the weak* topology, we see that any linear functional $f\colon X^* \to \mathbb{R}$, which is continuous for the w^*-topology, has the form $f(x^*) = \langle x^*, \hat{x}\rangle$ for some $\hat{x} \in X$.

Proposition 3.3.33. *If X is a normed space and $H \subseteq X^*$ is a* w^**-closed hyperplane, then there exist $\hat{x} \in X$, $\hat{x} \neq 0$, and $\vartheta \in \mathbb{R}$ such that*

$$H = \{x^* \in X^* : \langle x^*, \hat{x}\rangle = \vartheta\} \ .$$

Proof. We know that $H = \{x^* \in X^* : f(x^*) = \vartheta\}$ with $f\colon X^* \to \mathbb{R}$ being linear and $\vartheta \in \mathbb{R}$; see Definition 3.1.53. Since by hypothesis H is w^*-closed, Proposition 3.1.54 implies that f is w^*-continuous. Finally, using Remark 3.3.32, we conclude that there exists $\hat{x} \in X$ such that $H = \{x^* \in X^* : \langle x^*, \hat{x}\rangle = \vartheta\}$. □

Recall that every $x \in X$ defines in a natural way a linear functional $f_x\colon X^* \to \mathbb{R}$ according to the formula $f_x(x^*) = \langle x^*, x\rangle$. Indeed, we see that $|f_x(x^*)| = |\langle x^*, x\rangle| \le \|x^*\|_*\|x\|$, which shows that f_x is bounded, that is, $f_x \in X^*$, and $\|f_x\|_* \le \|x\|$. Thus we can define the map $j\colon X \to X^{**}$ by $j(x) = f_x$. Clearly j is linear, injective, and $\|j(x)\|_* \le \|x\|$ for all $x \in X$. Additional information about this map is supplied by the next proposition.

Proposition 3.3.34. *If X is a normed space and $j\colon X \to X^{**}$ is the linear map defined above, then j is an isometric isomorphism onto $j(X)$.*

Proof. We already proved that j is an isomorphism onto $j(X)$ and $\|j(x)\|_* \le \|x\|$ for all $x \in X$. On the other hand, from Proposition 3.1.50, we know that there exists $x^* \in X^*$ such that $\|x^*\|_* = 1$ and $j(x)(x^*) = \langle x^*, x\rangle = \|x\|$. This shows that $\|j(x)\|_* \ge \|x\|$ for all $x \in X$. Hence, j is an isometry. □

Definition 3.3.35. The isometry $j\colon X \to X^{**}$ of Proposition 3.3.34 is called the **canonical embedding** of the normed space X into X^{**}.

Remark 3.3.36. Using the canonical embedding we can identify X with a subspace of X^{**}. Moreover, $\overline{j(X)}$ is a closed subspace of the Banach space X^{**}. Hence, $V = \overline{j(X)}$ is a Banach space as well. Therefore j is an isometric isomorphism onto a dense subset of the Banach space V. Hence, the canonical embedding provides a shortcut to the completion of a normed space. Every normed space can be viewed as a dense subspace of a Banach space. When the canonical embedding j is not surjective, then the weak topology $\mathrm{w}(X^*, X^{**})$ is strictly larger than the weak* topology. Indeed let $\hat{u} \in X^{**} \setminus j(X)$

and consider the subspace $H = \{x^* \in X^* : \langle \hat{u}, x^* \rangle = 0\}$. Then H is w-closed, but it is not w^*-closed; see Proposition 3.3.33. In fact this example shows that Mazur's Theorem (see Theorem 3.3.18) fails for the w^*-topology. A strongly closed convex set need not be w^*-closed. Moreover, a normed space and its completion have the same dual space; however, their weak* topologies differ. So, one should be careful when dealing with the weak* topology of the dual of a normed space and that of the dual of the Banach space resulting from its completion.

Since X can be viewed as a subspace of X^{**}, it is natural to ask what kind of subspace it is. The answer is given by the so-called "Goldstine's Theorem." In what follows we set

$$B_1^X = \{x \in X : \|x\| < 1\}, \qquad \overline{B}_1^X = \{x \in X : \|x\| \le 1\}\,,$$
$$B_1^{X^{**}} = \{x^{**} \in X^{**} : \|x^{**}\|_{**} < 1\}, \quad \overline{B}_1^{X^{**}} = \{x^{**} \in X^{**} : \|x^{**}\|_{**} \le 1\}\,.$$

Theorem 3.3.37 (Goldstine's Theorem). *If X is a normed space, then $\overline{j\left(B_1^X\right)}^{w^*} = \overline{B}_1^{X^{**}}$ and $\overline{j(X)}^{w^*} = X^{**}$.*

Proof. Clearly, the second equality is a consequence of the first. So, let us prove the first one.

Let $x^{**} \in X^{**} \setminus \overline{j\left(B_1^X\right)}^{w^*}$. Since $\overline{j\left(B_1^X\right)}^{w^*} \subseteq X^{**}$ is convex and w^*-closed, by the Strong Separation Theorem (see Corollary 3.1.61), there exists $x^* \in (X^{**}_{w^*})^* = X^*$ with $x^* \neq 0$ such that

$$\sup\left[\langle x^*, u^{**}\rangle : u^{**} \in \overline{j\left(B_1^X\right)}^{w^*}\right] < \langle x^*, x^{**}\rangle\,. \tag{3.3.5}$$

We may always assume that $\|x^*\|_* = 1$. Then, from (3.3.5), we have

$$1 = \|x^*\|_* < \langle x^*, x^{**}\rangle \le \|x^*\|_* \|x^{**}\|_{**}\,.$$

Hence, $1 < \|x^{**}\|_{**}$ and so $\overline{j\left(B_1^X\right)}^{w^*} = \overline{B}_1^{X^{**}}$. ☐

The weaker a topology is, the more compact sets it has. The next theorem is the most important feature of the weak* topology. It is reminiscent of the Heine–Borel-Theorem and it is the reason why the weak* topology is important in the theory of Banach spaces. The result is known as "Alaoglu's Theorem."

Theorem 3.3.38 (Alaoglu's Theorem). *If X is a normed space, then $\overline{B}_1^{X^*} = \{x^* \in X^* : \|x^*\|_* \le 1\}$ is w^*-compact. More generally every w^*-closed and bounded subset of X^* is w^*-compact.*

Proof. Suppose that $x^* \in \overline{B}_1^{X^*}$. Then for each $x \in \overline{B}_1^X$ it follows $|\langle x^*, x\rangle| \le 1$. Therefore

$$x^*\left(\overline{B}_1^X\right) \subseteq I = \{\lambda \in \mathbb{R} : |\lambda| \le 1\}\,.$$

We can identify each element of $\overline{B}_1^{X^*}$ with a point in $I^{\overline{B}_1^X}$. From Tychonoff's Theorem, see Theorem 1.4.56, $I^{\overline{B}_1^X}$ equipped with the product topology is compact. Since the weak* topology is by definition the topology of pointwise convergence on $\overline{B}_1^X$, the identification of $\overline{B}_1^{X^*}$ with a subset of $I^{\overline{B}_1^X}$ leaves the weak* topology unchanged. So, it remains to show that $\overline{B}_1^{X^*}$ is closed in $I^{\overline{B}_1^X}$. To this end, let $\{x_\alpha^*\}_{\alpha\in I} \subseteq \overline{B}_1^{X^*}$ be a net and assume that it converges pointwise to $g \in I^{\overline{B}_1^X}$. Evidently g is linear and so g is the restriction on $\overline{B}_1^X$ of a linear functional x^* on X. Moreover, since $|g(x)| \le 1$ for all $x \in \overline{B}_1^X$, it follows that $x^* \in \overline{B}_1^{X^*}$ and this proves that $\overline{B}_1^{X^*}$ is closed in $I^{\overline{B}_1^X}$, and hence w*-compact.

Every bounded set $C \subseteq X^*$ satisfies $C \subseteq r\overline{B}_1^{X^*}$ for some $r > 0$. Since $\overline{B}_1^{X^*}$ is w*-compact and C is w*-closed, we conclude that it is w*-compact. □

Remark 3.3.39. From the theorem above, we derive that if X is a normed space and $C \subseteq X^*$, then C is w*-closed and bounded implies that C is w*-compact.

For the converse to hold, we need to assume that X is a Banach space. To see this, let $X = \{\hat{x} = (a_n)_{n\in\mathbb{N}} : a_n = 0 \text{ for all } n \ge n_0\}$ equipped with the norm $\|\hat{x}\| = \sum_{n\in\mathbb{N}} |a_n|$. Clearly this is a normed space but not a Banach space. Consider a sequence $\{\xi_n\}_{n\in\mathbb{N}} \subseteq \mathbb{R}$ with $\xi_n > 0$ for all $n \in \mathbb{N}$ such that $\xi_n \to +\infty$ as $n \to \infty$. Let $\{\hat{x}_n\}_{n\in\mathbb{N}} \subseteq X^*$ be defined by $\hat{x}_n^*(\hat{x}) = a_n$ for all $n \in \mathbb{N}$.

Let $D = \{0, \xi_1\hat{x}_1, \xi_2\hat{x}_2, \dots, \xi_n\hat{x}_n, \dots\} \subseteq X^*$. This set is unbounded since $\|\xi_n\hat{x}_n\|_* = \xi_n \to +\infty$. However, it holds $\xi_n\hat{x}_n(\hat{x}) = \xi_n a_n \to 0$ for all $\hat{x} \in X$. So, $\xi_n\hat{x}_n \xrightarrow{w^*} 0$ and it follows that $D \subseteq X^*$ is w*-bounded.

From the previous remark, we have the following corollary.

Corollary 3.3.40. *If X is a Banach space and $C \subseteq X^*$, then C is bounded if and only if C is* w*-*bounded, that is, $x(C) \subseteq \mathbb{R}$ is bounded for every $x \in X$.*

We conclude this section with a remarkable result of R. C. James, which provides a necessary and sufficient condition for a set C in a Banach space X to be weakly compact. The result is known as "James' Theorem" and its proof is lengthy and can be found in Holmes [155, p. 157].

Theorem 3.3.41 (James' Theorem). *If X is a Banach space and $C \subseteq X$ is bounded and* w-*closed, then C is* w-*compact if and only if every $x^* \in X^*$ attains its supremum over C.*

3.4 Separable and Reflexive Banach Spaces

In this section we examine two special classes of Banach spaces, namely separable and reflexive Banach spaces. They exhibit special properties, which are important in applications.

Definition 3.4.1. (a) A normed space X is **separable** if it contains a countable dense subset.

(b) A normed space X is **reflexive** if the canonical embedding $j\colon X \to X^{**}$ (see Definition 3.3.35) is surjective. A reflexive normed space is necessarily complete, that is, a Banach space.

Remark 3.4.2. Any subset of a separable normed space is a separable metric space. Many important spaces in analysis are separable and/or reflexive. Every finite dimensional Banach space is separable and reflexive. In the definition of reflexivity it is essential to use the canonical embedding j stated in Definition 3.3.35. R. C. James produced in 1951 a remarkable example of a nonreflexive Banach space X that is isometrically isomorphic to X^{**}. In this example, the image of X under the canonical embedding $j\colon X \to X^{**}$ is a closed subspace of codimension one. A detailed construction of this space can be found in Megginson [212]; see Section 4.5. In what follows, for the sake of notational simplicity, we drop the use of the map j. It is understood that X is embedded into X^{**} via the canonical embedding.

Proposition 3.4.3. *If X is a Banach space and X^* is separable, then X is separable.*

Proof. Let $\{x_n^*\}_{n\ge 1} \subseteq X^*$ be dense. Thanks to Corollary 3.1.48 we know that

$$\|x_n^*\|_* = \sup\left[\langle x_n^*, x\rangle : x \in X, \|x\| \le 1\right]$$

for all $n \in \mathbb{N}$. Hence, there exists $x_n \in X$ such that

$$\|x_n\| = 1 \quad \text{and} \quad \frac{1}{2}\|x_n^*\|_* \le \langle x_n^*, x_n\rangle\,, \qquad n \in \mathbb{N}\,. \tag{3.4.1}$$

Let $V_0 = \operatorname{span}_{\mathbb{Q}}\{x_n\}_{n\in\mathbb{N}}$, that is, V_0 is the set of all finite linear combinations with coefficients in $\mathbb{Q}$ of the vectors $\{x_n\}_{n\in\mathbb{N}}$. This set is countable since $V_0 = \bigcup_{m\ge 1} V_m$ with V_m being the set of linear combinations with coefficients in $\mathbb{Q}$ of $\{x_n\}_{n=1}^m$. Each V_m is countable, and so $V_0 = \bigcup_{m\ge 1} V_m$ is countable as well.

Let $V = \operatorname{span}\{x_n\}_{n\in\mathbb{N}}$. We claim that V is dense in X. To this end, let $x^* \in V^\perp$. Then there exists $\{x_{n_k}^*\}_{k\in\mathbb{N}} \subseteq \{x_n^*\}_{n\in\mathbb{N}}$ such that

$$x_{n_k}^* \to x^* \quad \text{in } X^* \text{ as } k \to \infty\,. \tag{3.4.2}$$

Then, because of (3.4.1) and since $x^* \in V^\perp$, it follows that

$$\left\|x_{n_k}^*\right\|_* \le 2\langle x_{n_k}^*, x_{n_k}\rangle = 2\langle x_{n_k}^* - x^*, x_{n_k}\rangle \le 2\left\|x_{n_k}^* - x^*\right\|_* \left\|x_{n_k}\right\| = 2\left\|x_{n_k}^* - x^*\right\|_*\,.$$

Hence, thanks to (3.4.2), one gets $x_{n_k}^* \to x^* = 0$ in X^*. This shows that $V^\perp = \{0\}$ and so V is dense in X; see Remark 3.1.63. Since V_0 is countable and dense in V, we conclude that X is separable. □

Remark 3.4.4. The converse of this result is not true. Namely, separability of X does not imply separability of X^*. For example, $X = L^1([0,1])$ is separable (see Proposition 2.3.24), but $X^* = L^\infty([0,1])$ is not separable; see Proposition 2.3.29. In Section 4.1 we will show that $L^\infty([0,1]) = L^1([0,1])^*$.

Theorem 3.4.5. *A Banach space X is reflexive if and only if $\overline{B}_1^X = \{x \in X : \|x\| \le 1\}$ is* w*-compact.*

Proof. $\Longrightarrow$: The reflexivity of X implies that $X = X^{**}$. Hence $\overline{B}_1^X = \overline{B}_1^{X^{**}}$. By Alaoglu's Theorem (see Theorem 3.3.38), $\overline{B}_1^{X^{**}}$ is w*-compact and from Proposition 1.3.5, we know that

$$\mathrm{w}(X^{**}, X^*)\big|_X = \mathrm{w}(X, X^*)\,. \tag{3.4.3}$$

Therefore, $\overline{B}_1^X$ is w-compact.

$\Longleftarrow$: Since by hypothesis, $\overline{B}_1^X$ is w-compact, it is w*-closed in X^{**}; see (3.4.3). The Goldstine's Theorem (see Theorem 3.3.37), gives $\overline{\overline{B}_1^X}^{\mathrm{w}^*} = \overline{B}_1^{X^{**}}$ and since $\overline{B}_1^X$ is w*-closed in X^{**}, we obtain $\overline{B}_1^X = \overline{B}_1^{X^{**}}$. Therefore $X = X^{**}$ and so we conclude that X is reflexive. ☐

Proposition 3.4.6. *A Banach space X is reflexive if and only if X^* is reflexive.*

Proof. $\Longrightarrow$: Since X is reflexive, we know that $X = X^{**}$ and so the weak and weak* topologies on X^* coincide. Alaoglu's Theorem (see Theorem 3.3.38) implies that $\overline{B}_1^{X^*} = \{x^* \in X^* : \|x^*\|_* \le 1\}$ is w-compact and so Theorem 3.4.5 implies that X^* is reflexive.

$\Longleftarrow$: Since X^* is reflexive, by the previous part of the proof we have that X^{**} is reflexive as well. Then, Theorem 3.4.5 implies that $\overline{B}_1^{X^{**}} = \{x^{**} \in X^{**} : \|x^{**}\|_{**} \le 1\}$ is w-compact. The set $\overline{B}_1^X$ is closed, convex, hence a w-closed subset of $\overline{B}_1^{X^{**}}$; see Mazur's Theorem (Theorem 3.3.18). Therefore $\overline{B}_1^X$ is w-compact in X^{**}. Since the w*-topology on X^{**} is weaker than the w-topology, it follows that $\overline{B}_1^X$ is w*-compact in X^{**}. Hence it is w-compact in X; see (3.4.3). We conclude by using Theorem 3.4.5. ☐

Proposition 3.4.7. *If X is a reflexive Banach space and V is a closed subspace of X, then V is a reflexive Banach space.*

Proof. We know that

$$\mathrm{w}(V, V^*) = \mathrm{w}(X, X^*)\big|_V\,; \tag{3.4.4}$$

see Proposition 1.3.5. The set $\overline{B}_1^V = \{x \in V : \|x\| \le 1\}$ is a weakly closed subset of the weakly compact set $\overline{B}_1^X$; see Theorem 3.4.5. Combining this with (3.4.4), we infer that $\overline{B}_1^V$ is w-compact in V. Then invoking Theorem 3.4.5 we conclude that V is reflexive. ☐

Combining Propositions 3.4.3 and 3.4.6, we obtain the following.

Proposition 3.4.8. *If X is a Banach space, then X is separable and reflexive if and only if X^* is separable and reflexive.*

Proposition 3.4.9. *If X is a reflexive Banach space and $V \subseteq X$ is a closed subspace, then X/V is reflexive.*

Proof. From Proposition 3.2.25, we know that $(X/V)^*$ and $V^\perp$ are isometrically isomorphic. Let $\xi : (X/V)^* \to V^\perp$ be this isometric isomorphism. If $p : X \to X/V$ is the quotient map (see Definition 3.2.19), then from the proof of Proposition 3.2.25 we know that

$$\xi(l) = l \circ p \quad \text{for all } l \in (X/V)^* .$$

Let $l^* \in (X/V)^{**}$. The map $l^* \circ \xi^{-1} \colon V^\perp \to \mathbb{R}$ is a bounded linear functional on a subspace of X^*. Hence, by Proposition 3.1.49 there exists $x^{**} \in X^{**}$ such that $\langle x^{**}, x^* \rangle = \langle l^*, \xi^{-1}(x^*) \rangle$ for all $x^* \in V^\perp$. This implies that

$$\langle x^{**}, l \circ p \rangle = \langle l^*, l \rangle \quad \text{for all } l \in (X/V)^* . \tag{3.4.5}$$

The reflexivity of X implies that there exists $x \in X$ such that $j(x) = x^{**}$ with j being the canonical embedding. Let $u = [x] = p(x) \in X/V$. Combining Definition 3.3.35 and (3.4.5), it follows that

$$\langle l^*, l \rangle = \langle x^{**}, l \circ p \rangle = \langle j(x), l \circ p \rangle = \langle l \circ p, x \rangle = \langle l, p(x) \rangle = \langle l, u \rangle .$$

Hence, $j(u) = l^*$ with j being the canonical embedding for X/V. Since $l^* \in (X/V)^{**}$ is arbitrary, it follows that j is surjective and so X/V is reflexive; see Definition 3.4.1(b). ☐

We know that on an infinite dimensional normed space and on its dual, the weak and weak* topologies are never metrizable. Nevertheless, the traces of these topologies on certain subspaces can be metrizable. The results that follow investigate this issue. We start with a general topological result.

Lemma 3.4.10. *If (X, τ) is a compact topological space and $\{f_n\}_{n\geq 1}$ is a separating sequence of continuous functions on X (see Definition* 1.3.6*), then the topology τ is metrizable.*

Proof. We may assume that $|f_n(x)| \leq 1$ for all $x \in X$ and for all $n \in \mathbb{N}$. On X we consider the metric d defined by

$$d(x, u) = \sum_{n \in \mathbb{N}} \frac{1}{2n} |f_n(x) - f_n(u)| \quad \text{for all } x, u \in X .$$

Let τ_d be the metric topology induced by this metric on X. For every fixed $u \in X$, $x \to d(x, u)$ is τ-continuous as the uniform limit of τ-continuous functions. So, for every $\varepsilon > 0$, it follows that $B_\varepsilon(u) = \{x \in X \colon d(x, u) < \varepsilon\} \in \tau$, which means that $\tau_d \subseteq \tau$. Using Theorem 1.4.54, we see that the identity map $i_X \colon (X, \tau) \to (X, \tau_d)$ is a homeomorphism. Hence $\tau = \tau_d$. ☐

Using this lemma, we can state the first metrizability result for the weak* topology.

Theorem 3.4.11. *If X is a separable normed space and $C \subseteq X^*$ is* w*-*compact, then* C *equipped with the* w*-*topology is metrizable.*

Proof. Let $\{x_n\}_{n\geq 1} \subseteq X$ be dense in X. If $j \colon X \to X^{**}$ is the canonical embedding, then

$$\langle j(x_n), x^* \rangle = \langle x^*, x_n \rangle \quad \text{for all } n \in \mathbb{N} \text{ and for all } x^* \in X^* ;$$

see Definition 3.3.35. So, if $\langle j(x_n), x^* \rangle = 0$ for all $n \in \mathbb{N}$, we derive that $\langle x^*, x_n \rangle = 0$ for all $n \in \mathbb{N}$ and the density of $\{x_n\}_{n\geq 1}$ in X implies that $x^* = 0$. Therefore $\{j(x_n)\}_{n\geq 1} \subseteq X^{**}$

is separating and each $j(x_n)$ is w*-continuous. Applying Lemma 3.4.10, we conclude that (C, w^*) is metrizable. $\square$

We can improve this result in the following way.

Theorem 3.4.12. *If X is a normed space, then the following hold:*
(a) $\left(\overline{B}_1^{X^*}, w^*\right)$ *is metrizable if and only if X is separable;*
(b) $\left(\overline{B}_1^{X}, w\right)$ *is metrizable if and only if X^* is separable.*

Proof. (a) $\Longrightarrow$: Since $\left(\overline{B}_1^{X^*}, w^*\right)$ is metrizable we can find a countable basis $\{U_n\}_{n\ge1}$ at the origin. We obtain

$$U_n = \left\{x^* \in \overline{B}_1^{X^*} : |\langle x^*, x\rangle| < \varepsilon_n \text{ for all } x \in F_n\right\}, \quad n \in \mathbb{N}$$

with $F_n \subseteq X$ finite and $\varepsilon_1, \ldots, \varepsilon_n > 0$. Let $E = \bigcup_{n\ge1} F_n$. Then $E \subseteq X$ is countable and $x^*(E) = 0$ implies $x^* \in U_n$ for all $n \in \mathbb{N}$ and so $x^* = 0$. Moreover, if $x^*(\overline{\operatorname{span}} E) = 0$, then $x^* = 0$. Therefore $\overline{\operatorname{span}} E = X$ and so we conclude that X is separable.

$\Longleftarrow$: This follows from Theorem 3.4.11.

(b) $\Longrightarrow$: As before, let $\{U_n\}_{n\ge1}$ be a countable local basis at the origin of X. We obtain

$$U_n = \left\{x \in \overline{B}_1^{X} : |\langle x^*, x\rangle| < \varepsilon_n \text{ for all } x^* \in F_n^*\right\}, \quad n \in \mathbb{N} \tag{3.4.6}$$

with $F_n^* \subseteq X^*$ finite and $\varepsilon_1, \ldots, \varepsilon_n > 0$. Let $E^* = \bigcup_{n\ge1} F_n^*$. Then $E^* \subseteq X^*$ is countable and so $\overline{\operatorname{span}} E^*$ is separable. We will show that $X^* = \overline{\operatorname{span}} E^*$. Arguing by contradiction, suppose that there exists $\hat{x}^* \in X^* \setminus \overline{\operatorname{span}} E^*$. Let $d = d(x^*, \overline{\operatorname{span}} E^*)$. Then we can find $\hat{x}^{**} \in X^{**}$ such that

$$\|\hat{x}^{**}\|_{**} = \frac{1}{d}, \quad \hat{x}^{**}(\overline{\operatorname{span}} E^*) = 0 \quad \text{and} \quad \langle \hat{x}^{**}, \hat{x}^*\rangle = 1\,; \tag{3.4.7}$$

see Proposition 3.1.50. We introduce

$$V = \left\{x \in \overline{B}_1^{X} : |\langle \hat{x}^*, x\rangle| < \frac{d}{2}\right\}. \tag{3.4.8}$$

Then V is a weak neighborhood of the origin in X and so $U_{n_0} \subseteq V$ for some $n_0 \in \mathbb{N}$. Note that $d\hat{x}^{**} \in \overline{B}_1^{X^{**}}$ and so by Goldstine's Theorem (see Theorem 3.3.37), there is $\hat{x} \in \overline{B}_1^{X}$ such that

$$|\langle d\hat{x}^{**} - \hat{x}, x^*\rangle| < \varepsilon_{n_0} \text{ for all } x^* \in F_{n_0}^* \quad \text{and} \quad |\langle d\hat{x}^{**} - \hat{x}, \hat{x}^*\rangle < \frac{d}{2}\,.$$

Then, due to (3.4.7),

$$|\langle x^*, \hat{x}\rangle| < \varepsilon_{n_0} \quad \text{for all } x^* \in F_{n_0}^* \quad \text{and} \quad |\langle \hat{x}^*, \hat{x}\rangle| > \frac{d}{2}\,.$$

This gives, with view to (3.4.6) and (3.4.8), that $\hat{x} \in U_{n_0}$ and $\hat{x} \notin V$, a contradiction to the fact that $U_{n_0} \subseteq V$. Therefore $X^* = \overline{\operatorname{span}} E^*$, and so X^* is separable.

$\Longleftarrow$: According to Alaoglu's Theorem (see Theorem 3.3.38), we know that $\overline{B}_1^{X^{**}}$ is w^*-compact. Since X^* is separable, from part (a) we derive that $(\overline{B}_1^{X^{**}}, w^*)$ is metrizable. Since $\overline{B}_1^X \subseteq \overline{B}_1^{X^{**}}$ via the canonical embedding and $w(X^{**}, X^*)|_X = w(X, X^*)$, we conclude that $(\overline{B}_1^X, w)$ is metrizable. $\square$

Remark 3.4.13. In particular, this theorem says that if X (resp. X^*) is separable and $C \subseteq X^*$ (resp. $C \subseteq X$) is bounded, then (C, w^*) (resp. (C, w)) is metrizable.

A subset C of a normed space X is said to be **weakly sequentially compact** (resp. **weakly countably compact**, **weakly limit point compact**) if it is sequentially compact (resp. countably compact, limit point compact) in the weak topology; see Definition 1.4.57.

A remarkable result known as the "Eberlein–Smulian Theorem" says that all these notions are equivalent to weak compactness. The proof of this result is lengthy and can be found in Dunford–Schwartz [94, p. 430] and Megginson [212, p. 248].

Theorem 3.4.14 (Eberlein–Smulian Theorem). *If X is a normed space and $C \subseteq X$, then the following properties are equivalent:*
(a) *C is (relatively) weakly compact.*
(b) *C is (relatively) weakly sequentially compact.*
(c) *C is (relatively) weakly countably compact.*
(d) *C is (relatively) weakly limit point compact.*

Remark 3.4.15. The theorem above is not true for the weak* topology.

Combining Theorems 3.4.5 and 3.4.14, we infer the following sequential characterization of reflexivity.

Theorem 3.4.16. *A Banach space X is reflexive if and only if every bounded sequence in X admits a weakly convergent subsequence.*

Two other consequences of Theorem 3.4.14 are the following two results.

Theorem 3.4.17. *If X is a separable normed space and $C \subseteq X$ is weakly compact, then (C, w) is metrizable.*

Theorem 3.4.18. *If X is a reflexive Banach space, $C \subseteq X$ is bounded, and $x \in \overline{C}^{w}$, then there exists a sequence $\{x_n\}_{n\geq 1} \subseteq C$ such that $x_n \xrightarrow{w^*} x$ in X.*

The next proposition provides a way to identify weakly compact sets.

Proposition 3.4.19. *If X is a Banach space, $C \subseteq X$ is* w*-closed, and for every $\varepsilon > 0$ there is a weakly compact set $K_\varepsilon \subseteq X$ such that $C \subseteq K_\varepsilon + \varepsilon \overline{B}_1^X$, then C is weakly compact.*

Proof. Viewing C as a subset of X^{**} via the canonical embedding, we directly obtain

$$\overline{C}^{w^*} \subseteq \overline{K_\varepsilon + \varepsilon \overline{B}_1^X}^{w^*} = \overline{K}_\varepsilon^{w^*} + \varepsilon \overline{\overline{B}_1^X}^{w^*} = K_\varepsilon + \varepsilon \overline{B}_1^{X^{**}},$$

since K_ε is w-compact and due to Theorem 3.3.37. Therefore

$$\overline{C}^{\mathrm{w}^*} \subseteq \bigcap_{\varepsilon>0} \left(K_\varepsilon + \varepsilon \overline{B}_1^{X^{**}}\right) \subseteq X\,,$$

which shows that C is w-compact since C is w-closed. □

Continuing with weakly compact sets, we show that this property is preserved if we take the closed convex hull of the set.

Proposition 3.4.20. *If X is a Banach space and $C \subseteq X$ is* w*-compact, then $\overline{\operatorname{conv}}\, C \subseteq X$ is* w*-compact as well.*

Proof. Let $x^* \in X^*$. Then

$$\sup\left[\langle x^*, x\rangle : x \in C\right] = \sup\left[\langle x^*, u\rangle : u \in \overline{\operatorname{conv}}\, C\right]. \tag{3.4.9}$$

Because $C \subseteq X$ is w-compact, there exists $\hat{x} \in C$ such that

$$\langle x^*, \hat{x}\rangle = \sup\left[\langle x^*, x\rangle : x \in C\right].$$

This implies, due to (3.4.9), that

$$\langle x^*, \hat{x}\rangle = \sup\left[\langle x^*, u\rangle : u \in \overline{\operatorname{conv}}\, C\right].$$

Since $x^* \in X^*$ is arbitrary, invoking James's Theorem (see Theorem 3.3.41), we conclude that $\overline{\operatorname{conv}}\, C$ is w-compact. Note that $\overline{\operatorname{conv}}\, C$ is w-closed by Theorem 3.3.18. □

Next we introduce some new classes of Banach spaces based on some geometric properties of the unit ball.

Definition 3.4.21. Let X be a Banach space.

(a) We say that X is **strictly convex** if for all $x, u \in X$ with $x \neq u$ and $\|x\| = \|u\| = 1$ it holds $\|(1-t)x + tu\| < 1$ for all $t \in (0,1)$.

(b) We say that X is **uniformly convex** if for every $\varepsilon > 0$ there exists $\delta = \delta(\varepsilon) > 0$ such that

$$x, u \in X,\ \|x\| \le 1,\ \|u\| \le 1,\ \|x-u\| \ge \varepsilon \quad \text{imply} \quad \frac{1}{2}\|x+u\| \le 1-\delta\,.$$

(c) We say that X is **locally uniformly convex** if for every $\varepsilon > 0$ and $x \in X$ with $\|x\| = 1$ there exists $\delta = \delta(\varepsilon, x) > 0$ such that

$$u \in X,\ \|u\| = 1,\ \|x-u\| \ge \varepsilon \quad \text{imply} \quad \frac{1}{2}\|x+u\| \le 1-\delta\,.$$

Remark 3.4.22. Evidently it holds

$$\text{Uniformly convex} \Longrightarrow \text{Locally uniformly convex} \Longrightarrow \text{Strictly convex}\,.$$

Note that these implications are not reversible in general. For finite dimensional spaces, the three notions are equivalent.

Proposition 3.4.23. *Let X be a Banach space. The following properties are equivalent:*
(a) *X is strictly convex.*
(b) *The boundary of the unit ball called the unit sphere contains no line segments.*
(c) *$x \neq u$ and $\|x\| = \|u\| = 1$ implies $\|x + u\| < 2$.*
(d) *If $\|x - y\| = \|x - u\| + \|u - y\|$ for $x, u, y \in X$, then there exists $t \in [0, 1]$ such that $u = (1 - t)x + ty$.*
(e) *Every $x^* \in X^* \setminus \{0\}$ attains its supremum on $\overline{B}_1^X$ on at most one point.*

Proof. (a) $\Longrightarrow$ (b): This is obvious from Definition 3.4.21(a).

(b) $\Longrightarrow$ (a): Arguing by contradiction suppose that we can find $x, u \in X$, $x \neq u$, $\|x\| = \|u\| = 1$ and $t_0 \in (0, 1)$ such that $\|(1 - t_0)x + t_0 u\| = 1$. Let $t \in (0, t_0)$. Then we obtain

$$(1 - t_0)x + t_0 u = \frac{1 - t_0}{1 - t}((1 - t)x + tu) + \frac{t_0 - t}{1 - t}u\,,$$

which gives

$$1 \leq \frac{1 - t_0}{1 - t}\|(1 - t)x + tu\| + \frac{t_0 - t}{1 - t}\,.$$

Hence $\|(1 - t)x + tu\| \geq 1$ and so $\|(1 - t)x + tu\| = 1$.

Similarly we treat the case $t \in (t_0, 1)$. Therefore the line segment $[x, u]$ is on the unit sphere of X, a contradiction to the hypothesis.

(a) $\Longrightarrow$ (c) and (c) $\Longrightarrow$ (b): These implications are obvious.

(a) $\Longrightarrow$ (d): Let $x, u, y \in X$ be such that $\|x - y\| = \|x - u\| + \|u - y\|$. We may assume that $\|x - u\| \neq 0$, $\|u - y\| \neq 0$ and $\|x - u\| \leq \|u - y\|$. Then we derive

$$\begin{aligned}
&\left\|\frac{1}{2}\frac{x - u}{\|x - u\|} + \frac{1}{2}\frac{u - y}{\|u - y\|}\right\| \\
&\geq \left\|\frac{1}{2}\frac{x - u}{\|x - u\|} + \frac{1}{2}\frac{u - y}{\|x - u\|}\right\| - \left\|\frac{1}{2}\frac{u - y}{\|x - u\|} - \frac{1}{2}\frac{u - y}{\|u - y\|}\right\| \\
&= \frac{1}{2}\frac{\|x - y\|}{\|x - u\|} - \frac{1}{2}\frac{\|u - y\| - \|x - u\|}{\|x - u\|} \\
&= \frac{1}{2}\frac{1}{\|x - u\|}2\|x - u\| = 1\,.
\end{aligned}$$

Hence we obtain

$$\left\|\frac{x - u}{\|x - u\|} + \frac{u - y}{\|u - y\|}\right\| = 2,$$

which finally gives

$$\frac{x - u}{\|x - u\|} = \frac{u - y}{\|u - y\|}\,.$$

Therefore $u = (1 - t)x + ty$ with $t = (\|x - u\|)/(\|x - y\|) \in (0, 1)$.

(d) $\Longrightarrow$ (c): Let $x, y \in X$, $x \neq y$ with $\|x\| = \|y\| = 1/2\|x + y\| = 1$. Then $\|x + y\| = \|x\| + \|y\|$, which gives $u = 0 = (1 - t)x - ty$ for some $t \in (0, 1)$. Hence $x = t/(1 - t)y$ and so $t = 1/2$, that is, $x = y$, a contradiction. Therefore we conclude that $\|x + y\| < 2$.

(a) $\Longrightarrow$ (e): Let $x^* \in X^* \setminus \{0\}$, and suppose that there exist $x, u \in X$ with $\|x\| = \|u\| = 1$ such that $\langle x^*, x\rangle = \langle x^*, u\rangle = \|x^*\|_*$. For $t \in (0, 1)$ it follows that

$$\|x^*\|_* = (1-t)\langle x^*, x\rangle + t\langle x^*, u\rangle = \langle x^*, (1-t)x + tu\rangle \le \|x^*\|_*\|(1-t)x + tu\| ,$$

which implies $1 \le \|(1-t)x + tu\| < 1$, a contradiction. Thus, $x^* \in X^* \setminus \{0\}$ has at most one maximizer on the closed unit ball of X.

(e) $\Longrightarrow$ (c): Suppose that there are $x, u \in X$, $x \ne u$ with $\|x\| = \|u\| = 1$, $\|x + u\| = 2$. Invoking Proposition 3.1.50, there exists $x^* \in X^*$ such that $\|x^*\|_* = 1$ and $\langle x^*, 1/2(x+u)\rangle = 1/2\|x + u\| = 1$. Hence

$$\langle x^*, x\rangle + \langle x^*, u\rangle = 2 . \tag{3.4.10}$$

It holds $\langle x^*, x\rangle \le 1$ and $\langle x^*, u\rangle \le 1$. So, from (3.4.10) it follows that $\langle x^*, x\rangle = \langle x^*, u\rangle = 1$, which contradicts the hypothesis. Therefore $\|x + u\| < 2$. □

From the proposition above and its proof we directly obtain the following corollary.

Corollary 3.4.24. *Let X be a Banach space. The following properties are equivalent:*
(a) *X is strictly convex.*
(b) *If $x, u \in X$, $\|x\| = \|u\| = 1$ and $\|x + u\| = 2$, then $x = u$.*
(c) *If $x, u \in X$ satisfy $2\|x\|^2 + 2\|u\|^2 = \|x + u\|^2$, then $x = u$.*
(d) *If $x, u \in X \setminus \{0\}$ satisfy $\|x + u\| = \|x\| + \|u\|$, then $x = tu$ for some $t > 0$.*

A sequential reformulation of Definition 3.4.21(b),(c) gives the following characterization of uniform convexity and local uniform convexity.

Proposition 3.4.25. *Let X be a Banach space.*
(a) *X is uniformly convex if and only if for every $\{x_n\}_{n\ge1}, \{u_n\}_{n\ge1} \subseteq \overline{B}_1^X$ such that $\|x_n + u_n\| \to 2$, we have $\|x_n - u_n\| \to 0$ as $n \to \infty$.*
(b) *X is locally uniformly convex if and only if for any $x \in X$, $\|x\| = 1$ and for every sequence $\{x_n\}_{n\ge1} \subseteq X$ with $\|x_n\| = 1$ for all $n \in \mathbb{N}$ such that $\|x_n + x\| \to 2$, we have $\|x_n - x\| \to 0$.*

Remark 3.4.26. In the characterizations above, the sequence can be replaced by nets.

Another characterization of uniform convexity is given by the next proposition.

Proposition 3.4.27. *If X is a Banach space, then X is uniformly convex if and only if for every sequences $\{x_n\}_{n\ge1}, \{u_n\}_{n\ge1} \subseteq X$ with $\{x_n\}_{n\ge1}$ bounded such that*

$$2\|x_n\|^2 + 2\|u_n\|^2 - \|x_n + u_n\|^2 \to 0 \quad \text{as } n \to \infty ,$$

we have $\|x_n - u_n\| \to 0$ as $n \to \infty$.

Proof. $\Longrightarrow$: Note that

$$\begin{aligned}(\|x_n\| - \|u_n\|)^2 &= 2\|x_n\|^2 + 2\|u_n\|^2 - (\|x_n\| + \|u_n\|)^2 \\ &\le 2\|x_n\|^2 + 2\|u_n\|^2 - \|x_n + u_n\|^2\end{aligned}$$

for all $n \in \mathbb{N}$. Hence $\|x_n\| - \|u_n\| \to 0$ as $n \to \infty$. Therefore if $\|x_n\| \to 0$ or $\|u_n\| \to 0$, then $\|x_n - u_n\| \to 0$. So we may assume that there exists $\varepsilon > 0$ such that

$$\|x_n\| \geq \varepsilon \quad \text{and} \quad \|u_n\| \geq \varepsilon \quad \text{for all } n \in \mathbb{N}\,.$$

Let $y_n = x_n/\|x_n\|$, $v_n = u_n/\|u_n\|$ with $n \in \mathbb{N}$. Then $\|y_n\| = \|v_n\| = 1$ for all $n \in \mathbb{N}$ and $\|y_n + v_n\| \to 2$. It follows that $\|y_n - v_n\| \to 0$ and so $\|x_n - u_n\| \to 0$.

$\Longleftarrow$: This implication is obvious; see Proposition 3.4.25. □

Uniformly convex Banach spaces are reflexive. The result is known as the "Milman–Pettis Theorem."

Theorem 3.4.28 (Milman–Pettis Theorem). *If X is a uniformly convex Banach space, then X is reflexive.*

Proof. Let $x^{**} \in \overline{B}_1^{X^{**}}$. Invoking the Goldstine's Theorem (see Theorem 3.3.37), we can find a net $\{x_\alpha\}_{\alpha\in I} \subseteq \overline{B}_1^X$ such that $x_\alpha \xrightarrow{w^*} x^{**}$ in X^{**}. Exploiting the w*-lower semicontinuity of the norm $\|\cdot\|_{**}$ on X^{**} (see Proposition 3.3.31(c)), we see that $\|x_\alpha + x_\beta\| \to 2$. Applying Proposition 3.4.25(a) gives $\|x_\alpha - x_\beta\| \to 0$, which implies that $\{x_\alpha\}_{\alpha\in I} \subseteq X$ is a Cauchy net. The completeness of X implies that $x_\alpha \to x^{**} \in X$ and so $X = X^{**}$, that is, X is reflexive. □

In Remark 3.3.17 we mentioned that in the Banach space l^1 for sequences, weak and norm convergences are equivalent. More generally, any Banach space having this property is said to have the **Schur property**.

Example 3.4.29. The Banach space (in fact Hilbert space; see Section 3.5) $l^2 = \{\hat{x} = (x_n)_{n\geq 1} \in \mathbb{R}^{\mathbb{N}} :\ \sum_{n\geq 1} x_n^2 < \infty\}$ does not have the Schur property. Since l^2 is a Hilbert space, we have $(l^2)^* = l^2$, see Theorem 3.5.21. Let $e_n = (0, \ldots, 1, 0, \ldots)$ with 1 at the nth-spot. Then for every $\hat{x}^* \in (l^2)^* = l^2$ we have $\langle x^*, e_n\rangle \to 0$, that is, $e_n \xrightarrow{w^*} 0$. On the other hand $\|e_n\| = 1$ for all $n \in \mathbb{N}$ and so $e_n \not\to 0$ in the norm topology.

However l^2 as well as every Hilbert space has the following weakened version of the Schur property.

Definition 3.4.30. A normed space X is said to have the **Kadec–Klee property** if it satisfies the following condition:

For every sequence $\{x_n\}_{n\geq 1} \subseteq X$ such that $x_n \xrightarrow{w} x$ in X
and $\|x_n\| \to \|x\|$, we have $x_n \to x$ in X.

Remark 3.4.31. The names **Radon–Riesz property** or **property (H)** are also used in the literature.

Proposition 3.4.32. *If X is a locally uniformly convex Banach space, then X has the Kadec–Klee property.*

Proof. Consider $x_n \xrightarrow{w} x$ in X. Evidently we may assume that $x \neq 0$. Let $u \in X$, $u \neq 0$. Let $y_n = x_n/\|x_n\|$, $y = x/\|x\|$ with $n \in \mathbb{N}$. Then $\|y_n\| = \|y\| = 1$ for all $n \in \mathbb{N}$ and $y_n \xrightarrow{w} y$ in X. Hence

$$2 = 2\|y\| \le \liminf_{n\to\infty} \|y_n + y\| \le \limsup_{n\to\infty} \|y_n + y\| \le \lim_{n\to\infty} \|y_n\| + \|y\| = 2 ;$$

see Proposition 3.3.13(c). Then $\lim_{n\to\infty} \|y_n + y\| = 2$. Proposition 3.4.25(b) implies that $\|y_n - y\| \to 0$ since X is locally uniformly convex. ◻

3.5 Hilbert Spaces

In this section we turn our attention to Hilbert spaces, which are Banach spaces with some additional structure, resulting from the presence of an inner product. The inner product supplies a very rich structure, which leads to important simplifications and makes Hilbert spaces the infinite dimensional analog of Euclidean spaces.

Definition 3.5.1. Let H be a vector space over the field $\mathbb{F}$ with $\mathbb{F} = \mathbb{R}$ or $\mathbb{F} = \mathbb{C}$. An **inner product** on X is a map $(\cdot,\cdot)\colon H \times H \to \mathbb{F}$ such that
(a) $(\lambda x + u, y) = \lambda(x, y) + (u, y)$ for all $x, u, y \in H$ and for all $\lambda \in \mathbb{F}$ (linearity);
(b) $(x, u) = \overline{(u, x)}$ for all $x, u \in H$ (conjugate symmetry);
(c) $(x, x) \ge 0$ and $(x, x) = 0$ if and only if $x = 0$ (positive definiteness).

Remark 3.5.2. Linearity in (a) in fact means linearity in the first argument. In the second argument the map is conjugate linear. Property (b) is sometimes called **Hermitian symmetry**.

The next result is of fundamental importance and is known as the "Cauchy–Bunyakowsky-Schwarz inequality."

Proposition 3.5.3 (Cauchy–Bunyakowsky-Schwarz inequality). *If H is a vector space with inner product $(\cdot,\cdot)$, then $|(x, u)|^2 \le (x, x)(u, u)$ for all $x, u \in H$.*

Proof. Let $x, u \in H$ and let $\lambda \in \mathbb{F}$. Then it follows that

$$0 \le (x - \lambda u, x - \lambda u) = (x, x) - \overline{\lambda}(x, u) - \lambda\overline{(x, u)} + |\lambda|^2(u, u) .$$

Choosing $\lambda = (x, u)/\vartheta$ with $\vartheta > 0$ results in

$$0 \le (x, x) - \frac{1}{\vartheta}\left(2 - \frac{(u, u)}{\vartheta}\right)|(x, u)|^2 .$$

If $u \neq 0$, then choose $\vartheta = (u, u)$ to get the desired inequality. If $u = 0$, then $(x, u) = 0$ and so the inequality holds trivially. ◻

Proposition 3.5.4. *If H is a vector space with inner product $(\cdot,\cdot)$, then $\|x\| = (x, x)^{1/2}$ for all $x \in H$ defines a norm on H.*

Proof. We only need to verify the triangle inequality. So, let $x, u \in H$. Then, using Proposition 3.5.3, it follows that

$$\begin{aligned}\|x+u\|^2 &= (x+u, x+u) = (x,x) + (x,u) + (u,x) + (u,u) \\ &= \|x\|^2 + 2\,\mathrm{Re}(x,u) + \|u\|^2 \le \|x\|^2 + 2|(x,u)| + \|u\|^2 \\ &\le \|x\|^2 + 2\|x\|\|u\| + \|u\|^2 = (\|x\| + \|u\|)^2 .\end{aligned}$$

This shows the assertion. □

Remark 3.5.5. A vector space with an inner product will be referred as an **inner product space**. Usually we will not explicitly mention the inner product unless we want to distinguish between different inner products defined on H. The norm $\|\cdot\|$ defined in Proposition 3.5.4 is the norm defined (induced or generated) by the inner product $(\cdot,\cdot)$.

At this point it is natural to ask when a norm is defined by an inner product. The next proposition will lead to a necessary and sufficient condition for this to happen.

Proposition 3.5.6. *If H is an inner product space, then the following hold:*
(a) *Parallelogram law: For all $x, u \in H$ we have*

$$\|x+u\|^2 + \|x-u\|^2 = 2\left(\|x\|^2 + \|u\|^2\right) .$$

(b) *Polarization identities: For all $x, u \in H$ we have*

$$(x,u) = \frac{1}{4}\left[\|x+u\|^2 - \|x-u\|^2 + i\|x+iu\|^2 - i\|x-iu\|^2\right] \quad \textit{if } \mathbb{F} = \mathbb{C} ,$$

$$(x,u) = \frac{1}{4}\left[\|x+u\|^2 - \|x-u\|^2\right] \quad \textit{if } \mathbb{F} = \mathbb{R} .$$

Proof. (a) For all $x, u \in H$ and for all $\lambda \in \mathbb{F}$ one gets

$$\begin{aligned}\|x+\lambda u\|^2 &= \|x\|^2 + 2\,\mathrm{Re}(\overline{\lambda}(x,u)) + |\lambda|^2\|u\|^2 \\ &= \|x\|^2 + 2\big[\,\mathrm{Re}\,\lambda\,\mathrm{Re}(x,u) - \mathrm{im}\,\lambda\,\mathrm{im}(x,u)\big] + |\lambda|^2\|u\|^2 .\end{aligned} \tag{3.5.1}$$

Choosing $\lambda = 1$ and $\lambda = -1$ in (3.5.1) and adding these equalities, we obtain the desired parallelogram law.

(b) Choosing $\lambda = 1$ and $\lambda = -1$ in (3.5.1) and subtracting, we get the real polarization identity, that is, the case $\mathbb{F} = \mathbb{R}$. Choosing $\lambda = i$ and $\lambda = -i$ in (3.5.1) and subtracting, we obtain the complex polarization identity, that is, the case $\mathbb{F} = \mathbb{C}$. □

The next theorem provides a necessary and sufficient condition for a norm to be generated by an inner product. For a proof of this result, we refer to Weidmann [307, p. 9].

Theorem 3.5.7. *A norm on a vector space H is defined by an inner product if and only if it satisfies the parallelogram law. Moreover, if the norm on H satisfies the parallelogram law, then the unique inner product defining the norm is given by the polarization identities; see Proposition* 3.5.6*(b).*

Definition 3.5.8. A **Hilbert space** is a complete inner product space.

Remark 3.5.9. So, according to Theorem 3.5.7, a Hilbert space is a Banach space whose norm satisfies the parallelogram law.

Theorem 3.5.10. *Every Hilbert space H is uniformly convex, hence reflexive; see Theorem* 3.4.28.

Proof. Let $\varepsilon > 0$ and let $x, u \in H$ with $\|x\| \leq 1$, $\|u\| \leq 1$ and $\|x - u\| \geq \varepsilon$. Using the parallelogram law (see Proposition 3.5.6(a)), we derive $\|(x+u)/2\|^2 \leq 1 - \varepsilon^2/4$, which implies that

$$\frac{1}{2}\|x + u\| \leq 1 - \delta \quad \text{with} \quad \delta = 1 - \left(1 - \frac{\varepsilon^2}{4}\right)^{\frac{1}{2}} > 0 \,.$$

Therefore H is uniformly convex; see Definition 3.4.21(b). □

The next notion is particular to inner product spaces and gives them the extra structure with respect to general Banach spaces.

Definition 3.5.11. Let H be an inner product space and $x, u \in H$. We say that x, u are **orthogonal** denoted by $x \perp u$ if $(x, u) = 0$. If $x \in H$ and $C \subseteq H$, then we say that x is orthogonal to C denoted by $x \perp C$ if $x \perp u$ for all $u \in C$. Finally if $C, D \subseteq X$, we say that the two sets are orthogonal, denoted by $C \perp D$ if $x \perp u$ for all $x \in C$ and for all $u \in D$. We say that $C \subseteq X$ is an **orthogonal set** if $x \perp u$ for all $x, u \in C$ with $x \neq u$.

Remark 3.5.12. Clearly, $x \perp u$ if and only if $u \perp x$. Hence $C \perp D$ if and only if $D \perp C$. Moreover, $C \perp D$ implies $C \cap D = \{0\}$.

The next result is an extension of the classical "Pythagorean Theorem."

Theorem 3.5.13 (Generalized Pythagorean Theorem). *If H is an inner product space and $\{x_k\}_{k=0}^n \subseteq H$ is a finite orthogonal set, then*

$$\left\|\sum_{k=0}^n x_k\right\|^2 = \sum_{k=0}^n \|x_k\|^2 \,.$$

Proof. First suppose that $n = 1$, that is, we have a pair $x_0, x_1 \in X$ of orthogonal vectors. Since $x_0 \perp x_1$ we derive

$$\|x_0 + x_1\|^2 = (x_0 + x_1, x_0 + x_1) = \|x_0\|^2 + 2\,\mathrm{Re}(x_0, x_1) + \|x_1\|^2 = \|x_0\|^2 + \|x_1\|^2 \,.$$

So, the result holds for $n = 1$. Proceeding by induction, suppose that it holds for some $n \in \mathbb{N}$, that is

$$\left\|\sum_{k=0}^n x_k\right\|^2 = \sum_{k=0}^n \|x_k\|^2 \quad \text{for every orthogonal set } \{x_k\}_{k=0}^n \subseteq H \,. \tag{3.5.2}$$

Let $\{x_k\}_{k=0}^{n+1} \subseteq H$ be an arbitrary orthogonal set. Since $x_{n+1} \perp \{x_k\}_{k=0}^{n}$ it follows that $x_{n+1} \perp \sum_{k=0}^{n} x_k$, and hence

$$\left\| \sum_{k=0}^{n+1} x_k \right\|^2 = \left\| \sum_{k=0}^{n} x_k + x_{n+1} \right\|^2 = \left\| \sum_{k=0}^{n} x_k \right\|^2 + \|x_{n+1}\|^2 = \sum_{k=0}^{n+1} \|x_k\|^2 ;$$

see (3.5.2). So, the induction is complete and the Generalized Pythagorean Theorem holds. □

We can state an infinite version of the Pythagorean Theorem.

Theorem 3.5.14. *If H is an inner product space and $\{x_k\}_{k\geq 1} \subseteq H$ is an orthogonal sequence, then the following hold:*

(a) *$\sum_{k\geq 1} x_k$ exists in X implies that $\sum_{k\geq 1} \|x_k\|^2 < \infty$ and $\left\|\sum_{k\geq 1} x_k\right\|^2 = \sum_{k\geq 1} \|x_k\|^2$.*

(b) *If H is a Hilbert space and $\sum_{k\geq 1} \|x_k\|^2 < \infty$, then $\sum_{k\geq 1} x_k$ exists in H.*

Proof. (a) By hypothesis we have

$$\sum_{k=1}^{n} x_k \to \sum_{k\geq 1} x_k \quad \text{in } H \text{ as } n \to \infty ,$$

which implies that

$$\left\| \sum_{k=1}^{n} x_k \right\|^2 \to \left\| \sum_{k\geq 1} x_k \right\|^2 \quad \text{as } n \to \infty . \tag{3.5.3}$$

From Theorem 3.5.13 we obtain

$$\left\| \sum_{k=1}^{n} x_k \right\|^2 = \sum_{k=1}^{n} \|x_k\|^2 \quad \text{for every } n \in \mathbb{N} .$$

Hence, due to (3.5.3)

$$\sum_{k=1}^{n} \|x_k\|^2 \to \left\| \sum_{k\geq 1} x_k \right\|^2 \quad \text{as } n \to \infty .$$

Therefore

$$\left\| \sum_{k\geq 1} x_k \right\|^2 = \sum_{k\geq 1} \|x_k\|^2 < \infty .$$

(b) For $m > n$, it holds

$$\left\| \sum_{k=1}^{m} x_k - \sum_{k=1}^{n} x_k \right\|^2 = \left\| \sum_{k=n+1}^{m} x_k \right\|^2 = \sum_{k=n+1}^{m} \|x_k\|^2 ;$$

see Theorem 3.5.13. Hence

$$\left\| \sum_{k=1}^{m} x_k - \sum_{k=1}^{n} x_k \right\|^2 \to 0 \quad \text{as } n \to \infty .$$

Therefore, $\{\sum_{k=1}^{n} x_k\}_{n\in\mathbb{N}} \subseteq H$ is a Cauchy sequence. Since H is a Hilbert space it follows that $\sum_{k=1}^{n} x_k \to \sum_{k\geq 1} x_k$ in H as $n \to \infty$. □

Corollary 3.5.15. *If H is a Hilbert space and $\{x_k\}_{k\geq 1} \subseteq H$ is an orthogonal sequence, then $\sum_{k\geq 1} \|x_k\|^2 < +\infty$ if and only if $\sum_{k\geq 1} x_k$ exists in H and $\|\sum_{k\geq 1} x_k\|^2 = \sum_{k\geq 1} \|x_k\|^2$.*

Example 3.5.16. Two classical examples of Hilbert spaces are the following ones:
(a) $\mathbb{R}^N$ equipped with the Euclidean inner product

$$(\hat{x}, \hat{u}) = \sum_{k=1}^{N} x_k u_k \quad \text{with } \hat{x} = (x_k)_{k=1}^{N},\ \hat{u} = (u_k)_{k=1}^{N} \in \mathbb{R}^N .$$

(b) The Banach space $l^2 = \{\hat{x} = (x_k)_{k\geq 1} \in \mathbb{R}^{\mathbb{N}} : \sum_{k\geq 1} x_k^2 < \infty\}$ equipped with the inner product

$$(\hat{x}, \hat{u}) = \sum_{k\geq 1} x_k u_k \quad \text{for all } \hat{x}, \hat{u} \in l^2 .$$

Remark 3.5.17. The other sequence Banach spaces $l^p = \{\hat{x} = (x_k)_{k\geq 1} \in \mathbb{R}^{\mathbb{N}} : \sum_{k\geq 1} |x_k|^p < \infty\}$ with $1 < p < \infty$ and $p \neq 2$ are not Hilbert spaces. We can easily see that the parallelogram law fails; see Theorem 3.5.7.

Now we present a basic property of closed convex sets in a Hilbert space. From now on all Hilbert spaces considered are real, that is, $\mathbb{F} = \mathbb{R}$.

Theorem 3.5.18. *If H is a Hilbert space and $C \subseteq H$ is nonempty, closed, and convex, then for any given $x \in H$ there exists a unique element $p_C(x) \in C$ such that $\|x - p_C(x)\| \leq \|x - u\|$ for all $u \in C$.*

Proof. By translating things if necessary, we assume that $x = 0$. Let $\eta = \inf[\|u\| : u \in C]$ and consider the minimizing sequence $\{u_n\}_{n\geq 1} \subseteq C$, that is, $\|u_n\| \searrow \eta$ as $n \to \infty$. From the parallelogram law (see Proposition 3.5.6(a)), one gets for $m > n$, that

$$\|u_m - u_n\|^2 = 2\|u_m\|^2 + 2\|u_n\|^2 - 4\left\|\frac{u_m + u_n}{2}\right\|^2 \leq 2\|u_m\|^2 + 2\|u_n\|^2 - 4\eta^2 ,$$

since C is convex. Hence $\|u_m - u_n\|^2 \to 0$ as $m, n \to \infty$ and so, $\{u_n\}_{n\geq 1} \subseteq C$ is a Cauchy sequence. Thus $u_n \to u$ in H and $\|u\| = \eta$.

Now we show the uniqueness of this best approximation (minimum norm) point u. Suppose that some $v \in C$ satisfies $\|v\| = \eta$. A new application of the parallelogram law gives

$$0 \leq \|u - v\|^2 = 2\|u\|^2 + \|v\|^2 - 4\left\|\frac{u + v}{2}\right\|^2 \leq 4\eta^2 - 4\eta^2 = 0 ,$$

recall again the convexity of C. Then $u = v$. So, $u = p_C(x)$ is the unique best approximation of x in C. □

Definition 3.5.19. The map $p_C : H \to C$ assigning to each $x \in H$ its unique best approximation from C is called the **metric projection of H onto C.**

The next proposition establishes the main properties of the metric projection map.

Proposition 3.5.20. *If H is a Hilbert space, $C \subseteq H$ is nonempty, closed, and convex, and $p_C\colon X \to C$ is the metric projection map. Then the following hold:*
(a) $p_C|_C = \mathrm{id}_C$;
(b) *if* $x \in X \setminus C$, *then* $p_C(x) \in \operatorname{bd} C$;
(c) $(x - p_C(x), u - p_C(x)) \leq 0$ *for all* $u \in C$;
(d) $\|p_C(x) - p_C(y)\| \leq \|x - y\|$ *for all* $x, y \in H$;
(e) *if C is a closed vector subspace of H, then* $x - p_C(x) \perp C$ *and* $p_C \in \mathcal{L}(H)$.

Proof. (a) This is obvious.

(b) Let $t \in (0, 1)$ and let $x_t = (1 - t)x + tp_C(x)$. We get

$$\|x - x_t\| = t\|x - p_C(x)\| < \|x - p_C(x)\| .$$

So, if $p_C(x) \in \operatorname{int} C$, then for $t \in (0, 1)$ close to 1 it follows $x_t \in C$, a contradiction. Hence $p_C(x) \in \operatorname{bd} C$.

(c) Let $x \in X$, $u \in C$ and $t \in (0, 1)$. The convexity of C implies

$$\begin{aligned}\|x - p_C(x)\|^2 \leq \|x - ((1 - t)p_C(x) + tu)\|^2 &= \|x - p_C(x) - t(u - p_C(x))\|^2 \\ &= \|x - p_C(x)\|^2 - 2t(x - p_C(x), u - p_C(x)) + t^2\|u - p_C(x)\|^2 ,\end{aligned}$$

which implies $2(x - p_C(x), u - p_C(x)) \leq t\|u - p_C(x)\|^2$. We let $t \searrow 0$ and obtain $(x - p_C(x), u - p_C(x)) \leq 0$ for all $u \in C$.

(d) Let $x, y \in H$. Using part (c) with $u = p_C(y) \in C$ it follows that

$$(x - p_C(x), p_C(y) - p_C(x)) \leq 0 . \tag{3.5.4}$$

Reversing the roles of $x, y \in H$ we also obtain

$$(y - p_C(y), p_C(x) - p_C(y)) \leq 0 . \tag{3.5.5}$$

Adding (3.5.4) and (3.5.5) yields

$$(x - y, p_C(y) - p_C(x)) + (p_C(y) - p_C(x), p_C(y) - p_C(x)) \leq 0 ,$$

which leads to

$$\|p_C(x) - p_C(y)\|^2 \leq \|x - y\|\|p_C(x) - p_C(y)\| ;$$

see Proposition 3.5.3. This finally gives $\|p_C(x) - p_C(y)\| \leq \|x - y\|$ for all $x, y \in H$.

(e) For every $u \in C$ and $\vartheta \in \mathbb{R}$ we get

$$\begin{aligned}\|x - p_C(x)\|^2 &\leq \|x - [p_C(x) + \vartheta(\pm u)]\|^2 \\ &= \|x - p_C(x)\|^2 \mp 2\vartheta(x - p_C(x), u) + \vartheta^2\|u\|^2 ,\end{aligned}$$

which turns into

$$\pm 2(x - p_C(x), u) \leq \vartheta\|u\|^2 .$$

Letting $\vartheta \searrow 0$, it results in $\pm(x - p_C(x), u) \leq 0$ for all $u \in C$ and so $(x - p_C(x), u) = 0$ for all $u \in H$ since $C \subseteq H$ is a subspace. This gives

$$x - p_C(x) \perp C\,. \tag{3.5.6}$$

Finally note that using (3.5.6), for all $u, x \in H$, leads to

$$(p_C(x+y) - (p_C(x) + p_C(y)), u) = 0 \quad \text{for all } u \in C\,.$$

Hence, $p_C(x+y) = p_C(x) + p_C(y)$, that is, p_C is additive. Clearly $p_C(0) = 0$ and for all $\lambda \in \mathbb{R} \setminus \{0\}$ it follows that $(p_C(\lambda x) - \lambda p_C(x), u) = 0$ for all $u \in C$, which shows that $p_C(\lambda x) = \lambda p_C(x)$, that is, p_C is homogeneous. Therefore $p_C \in L(H)$. □

A remarkable application of this result is a characterization of the topological dual of a Hilbert space. The result is known as the "Riesz-Fréchet Representation Theorem for Hilbert Spaces."

Theorem 3.5.21 (Riesz-Fréchet Representation Theorem for Hilbert Spaces). *If H is a Hilbert space and $x^* \in H^*$, then there exists a unique $x_0 \in H$ such that $\langle x^*, y\rangle = (x_0, y)$ for all $y \in H$ and $\|x^*\|_* = \|x_0\|$.*

Proof. Let $V = (x^*)^{-1}(0)$. This is a closed subspace of H. We may assume that $V \neq H$ otherwise $x^* = 0$ and the result is trivially true with $x_0 = 0$. Let $u_0 \in H \setminus V$, $u_1 = p_V(u_0)$ and $u = (u_0 - u_1)/(\|u_0 - u_1\|)$. Then $\|u\| = 1$ and $(u, x) = 0$ for all $x \in V$; see Proposition 3.5.20(e). Therefore $u \notin V$. For any $y \in H$, we set

$$z = y - tu \quad \text{with} \quad t = \frac{\langle x^*, y\rangle}{\langle x^*, u\rangle}\,.$$

Note that $\langle x^*, u\rangle \neq 0$ since $u \notin V$. Then $\langle x^*, z\rangle = 0$ and so $z \in V$. Therefore $(u, z) = 0$, which implies that

$$\langle x^*, y\rangle = \langle x^*, u\rangle(u, y) \quad \text{for all } y \in H\,. \tag{3.5.7}$$

So if we set $x_0 = \langle x^*, u\rangle u$, then it follows that $\langle x^*, y\rangle = (x_0, y)$ for all $y \in H$. Clearly this x_0 is unique. Evidently, thanks to Proposition 3.5.3 one gets $\|x_0\| \leq |\langle x^*, u\rangle|\|u\| = |\langle x^*, u\rangle| \leq \|x^*\|_*$. Moreover, from (3.5.7) and Proposition 3.5.3 we conclude that $\|x^*\|_* \leq |\langle x^*, u\rangle|\|u\| = \|x_0\|$, which implies $\|x^*\|_* = \|x_0\|$. □

Remark 3.5.22. According to this theorem there is a surjective linear isometry from H^* into H. This means that we can identify H^* with H, that is, a Hilbert space is self-dual. However, it is not always possible to do this identification. This is the case of evolution triples, which we will discuss in Section 4.2.

Definition 3.5.23. Let H be a Hilbert space and $C \subseteq H$. The **orthogonal complement** $C^\perp$ of C is the set

$$C^\perp = \{x \in H : (x, u) = 0 \quad \text{for all } u \in C\}\,.$$

On account of Theorem 3.5.21 the orthogonal complement of C is simply the annihilator of C introduced in Definition 3.2.24. Evidently $C^\perp$ is a closed vector subspace of H. Moreover, $C^{\perp\perp} = (C^\perp)^\perp$.

Remark 3.5.24. Clearly $\{0\}^\perp = H$, $H^\perp = \{0\}$. Moreover $C \perp C^\perp$, $C \cap C^\perp \subseteq \{0\}$ and if $0 \in C$, then $C \cap C^\perp = \{0\}$. Also, if $C, D \subseteq H$ are nonempty sets, then $C \perp D$ if and only if $C \subseteq D^\perp$. Since $\perp$ is a symmetric relation, that is, $C \perp D$ if and only if $D \perp C$, we also obtain that $D \subseteq C^\perp$. Moreover, $C \perp D$ implies that $C \cap D \subseteq \{0\}$. We can easily see that

$$\begin{aligned} &C \subseteq D \text{ implies that } D^\perp \subseteq C^\perp \text{ and } C^{\perp\perp} \subseteq D^{\perp\perp}\,, \\ &C^\perp = (\operatorname{span} C)^\perp = (\overline{\operatorname{span}} C)^\perp\,. \end{aligned} \tag{3.5.8}$$

In addition, since $C \perp C^\perp$ and $C^\perp \perp C^{\perp\perp}$, we derive that $C \subseteq C^{\perp\perp}$ and $C^\perp \subseteq C^{\perp\perp\perp}$, here $C^{\perp\perp\perp} = (C^{\perp\perp})^\perp$. Therefore we have $C \subseteq C^{\perp\perp}$ and $C^\perp = C^{\perp\perp\perp}$. Finally, if $C \subseteq H$ is a vector subspace, then $C^{\perp\perp} = \overline{C}$ and $C^\perp = \{0\}$ if and only if C is dense in H.

Proposition 3.5.25. *If H is a Hilbert space and V is a closed vector subspace of H, then $H = V \oplus V^\perp$; see Definition* 3.2.27.

Proof. It is easy to see that $V \oplus V^\perp$ is a closed vector subspace of H. Suppose that $H \neq V \oplus V^\perp$. Then there exists $u \in H$, $u \neq 0$ such that $u \perp V \oplus V^\perp$. We have $u \in V^\perp \cap V^{\perp\perp} = \{0\}$, a contradiction. Therefore, $H = V \oplus V^\perp$. ☐

From Propositions 3.5.20 and 3.5.25 we infer at once the so-called “Projection Theorem.”

Theorem 3.5.26 (Projection Theorem). *If H is a Hilbert space and V is a closed vector subspace of H, then there exists a unique pair of continuous linear operators $P\colon H \to V$ and $Q\colon H \to V^\perp$ such that*
(a) $x \in V$ *implies that* $P(x) = x$, $Q(x) = 0$ *and* $y \in V^\perp$ *implies that* $P(y) = 0$, $Q(y) = y$*;*
(b) $P(x) = p_V(x)$ *and* $Q(x) = p_{V^\perp}(x)$*;*
(c) *for all* $x \in H$ *one has* $\|x\|^2 = \|P(x)\|^2 + \|Q(x)\|^2$.

Now we turn our attention to orthogonal sets that lead to bases for Hilbert spaces. First we recall the following basic notion from linear algebra.

Definition 3.5.27. Let X be a vector space and $C \subseteq X$. We say that C is **linearly independent** if every $x \in C$ is not a linear combination of vectors in $C \setminus \{x\}$, that is, $x \notin \operatorname{span}[C \setminus \{x\}]$. A set $C \subseteq X$ that is not linearly independent is said to be **linearly dependent**.

Remark 3.5.28. The empty set $\emptyset$ is linearly independent. Also, the singleton $C = \{x\}$, $x \neq 0$ is linearly independent. Any set $C \subseteq X$, which contains the origin, is linearly independent. Finally, $C \subseteq X$ is linearly independent if and only if every finite subset of C is linearly independent.

Proposition 3.5.29. *If H is an inner product space and $C \subseteq H$ is an orthogonal set consisting of nonzero vectors, then C is linearly independent.*

Proof. Arguing by contradiction, suppose that there is a sequence $\{x_k\}_{k=0}^n$ with $n \geq 1$ such that $x_0 = \sum_{k=1}^n \lambda_k x_k$ with $\lambda_k \in \mathbb{R}$, $k = 1, \ldots, n$; see Remark 3.5.28. Exploiting the orthogonality of the set C yields $\|x_0\|^2 = \sum_{k=1}^n \lambda_k (x_k, x_0) = 0$, a contradiction. □

Definition 3.5.30. Let H be an inner product space and $C \subseteq X$. We say that C is an **orthonormal set** if it is an orthogonal set consisting of vectors with unit norm, that is, unit vectors.

Remark 3.5.31. Every orthogonal set consisting of nonzero vectors can be normalized. Indeed, if C is an orthogonal set such that $x \neq 0$ for all $x \in C$, then $\{x/\|x\| : x \in C\}$ is an orthonormal set.

From Proposition 3.5.29 we directly obtain the following result.

Proposition 3.5.32. *If H is an inner product space and $C \subseteq H$ is an orthonormal set, then C is linearly independent.*

The next proposition is an immediate consequence of Definition 3.5.30.

Proposition 3.5.33. *If H is an inner product space, $C \subseteq X$ is an orthonormal set, and $x \in H$ with $\|x\| = 1$ and $x \perp C$, then $C \cup \{x\}$ is an orthonormal set as well.*

Definition 3.5.34. Let H be an inner product space and let $\mathcal{L}$ be the family of all orthonormal subsets of H. Evidently $\mathcal{L} \neq \emptyset$ since $C = \{x\}$ with $\|x\| = 1$ is orthonormal. A set $C \in \mathcal{L}$ is **maximal orthonormal** if there is no set $C' \in \mathcal{L}$ such that $C' \neq C$ and $C \subseteq C'$.

Proposition 3.5.35. *If H is an inner product space and $C \subseteq H$ is an orthonormal set, then the following statements are equivalent:*

(a) *C is a maximal orthonormal set.*
(b) *There is no unit vector $x \in X$, that is, $\|x\| = 1$, such that $C \cup \{x\}$ is an orthonormal set.*
(c) $C^\perp = \{0\}$.

Proof. (a) $\Longrightarrow$ (b): Otherwise $C \cup \{x\}$ contradicts the maximality of C.

(b) $\Longrightarrow$ (c): Let $y \in H$ be such that $y \perp C$. Let $x = y/\|y\|$. Then $\|x\| = 1$ and $C \cup \{x\}$ is an orthonormal set, a contradiction.

(c) $\Longrightarrow$ (a): Arguing by contradiction, suppose that there exists an orthonormal set $C' \subseteq X$ such that $C' \setminus C \neq \emptyset$. Let $x \in C' \setminus C$. Then $\|x\| = 1$ with $x \perp C$, a contradiction. □

Proposition 3.5.36. *If H is an inner product space and $C \subseteq H$ is an orthonormal set, then there exists a maximal orthonormal set $C_0 \subseteq H$ such that $C \subseteq C_0$.*

Proof. Let $\mathcal{L}_C = \{D \in 2^H : D \text{ is an orthonormal set}, C \subseteq D\}$ and let $\mathcal{D}$ be a chain in $\mathcal{L}_C$. Let $\bigcup \mathcal{D} = \bigcup_{D \in \mathcal{D}} D$ and consider $x, u \in \bigcup \mathcal{D}$ with $x \neq u$. Then $x \in D_x \in \mathcal{D}$ and $u \in D_u \in \mathcal{D}$. Since $\mathcal{D}$ is a chain, we may assume that $D_x \subseteq D_u$. Hence $x, u \in D_u$ and so $\bigcup \mathcal{D} \in \mathcal{L}_C$. Invoking Zorn's Lemma (see Section 1.4), $\mathcal{L}_C$ has a maximal element C_0 such that $C \subseteq C_0$ and C_0 is orthogonal. If we can find a unit vector $x \in H$ such that $C_0 \cup \{x\}$ is orthonormal, then $C_0 \cup \{x\} \in \mathcal{L}_C$ and this contradicts the maximality of C_0. This proves that C_0 is a maximal orthonormal set. □

Now that we have established that maximal orthonormal sets exist, we can show that they span H.

Proposition 3.5.37. *If H is an inner product space and $C \subseteq H$ is an orthonormal set, then the following hold:*
(a) $\overline{\operatorname{span}}C = H$ *implies that C is maximal orthonormal.*
(b) *If H is a Hilbert space and $C \subseteq X$ is maximal orthonormal, then* $\overline{\operatorname{span}}C = H$.

Proof. (a) From (3.5.8) we know that $C^\perp = (\overline{\operatorname{span}}C)^\perp = H^\perp = \{0\}$. Then Proposition 3.5.35 implies that C is maximal orthonormal.

(b) From Proposition 3.5.35 and (3.5.8), we deduce that $0 = C^\perp = (\overline{\operatorname{span}}C)^\perp$. Hence $\overline{\operatorname{span}}C = H$; see Remark 3.5.24. □

Definition 3.5.38. Let H be an inner product space. A set $B \subseteq H$ is an **orthonormal basis** of H if the following hold:
(a) B is an orthonormal set.
(b) $\overline{\operatorname{span}}B = H$.

Remark 3.5.39. According to Proposition 3.5.37, every Hilbert space admits an orthonormal basis. In fact, for Hilbert spaces, the notions of maximal orthonormal set and of orthonormal basis coincide. That is, if H is a Hilbert space, then $B \subseteq H$ is a maximal orthonormal set if and only if $B \subseteq H$ is an orthonormal set. In finite dimensional Hilbert space all orthonormal bases are finite and have cardinality equal to the dimension of the space.

The next proposition establishes a fundamental inequality for inner product spaces known as "Bessel's inequality." First let us see how we interpret summation over an arbitrary index set.

Definition 3.5.40. Let (X, τ) be a Hausdorff topological vector space, I be an arbitrary index set, and $I \ni \alpha \to x_\alpha \in X$ be a map. Then the sum $\sum_{\alpha \in I} x_\alpha$ is defined as follows: Let $\mathcal{F}$ be the family of all finite subsets of I ordered by inclusion. Then $\sum_{\alpha\in I} x_\alpha = x$ if and only if the net $\{\sum_{\alpha \in F} x_\alpha\}_{F \in \mathcal{F}}$ τ-converges to x. This is called **unconditional convergence** since it does not depend on any ordering on the index set I. If $I = \mathbb{N}$, then $\sum_{n\geq 1} x_n = x$ means that $\sum_{n=1}^m x_n \xrightarrow{\tau} x$ as $m \to \infty$. Then the series $\sum_{n \geq 1}(-1)^n 1/n$ is convergent but not unconditionally convergent.

Remark 3.5.41. If $X = \mathbb{R}$, then $\sum_{\alpha\in I} x_\alpha = x \in \mathbb{R}$ means that for a given $\varepsilon > 0$ there exists a finite set $F \subseteq I$ such that $\left|x - \sum_{\alpha \in G} x_\alpha\right| \leq \varepsilon$ for all finite $F \subseteq G \subseteq I$. On the other hand $\sum_{\alpha \in I} x_\alpha = +\infty$ means that for any given $M > 0$ we can find a finite set $F \subseteq I$ such that $\sum_{\alpha\in G} x_\alpha \geq M$ for all $F \subseteq G \subseteq I$. Also recall that absolutely convergent series can be rearranged (see Amann–Escher [8, p. 201]) and we mention a remarkable result known as the "Orlicz–Pettis Theorem," which says that a series $\sum_{n\geq 1} x_n$ in a Banach space X is weakly unconditionally convergent if and only if it is strongly unconditionally convergent. Finally we mention another important result, the "Dvoretzky–Rogers Theorem," which says that if X is infinite dimensional, then there exists a sequence $\{x_n\}_{n\geq 1} \subseteq X$ such that $\sum_{n\geq 1} x_n$ is unconditional convergent and $\sum_{n\geq 1} \|x_n\| = +\infty$.

Lemma 3.5.42. *If $X = \mathbb{R}$ and $\{x_\alpha\}_{\alpha\in I} \subseteq [0, +\infty)$, then*

$$\sum_{\alpha\in I} x_\alpha = \sup\left[\sum_{\alpha\in F} x_\alpha : F \subseteq I \text{ is finite}\right] .$$

Proof. First suppose that $\sum_{\alpha\in I} x_\alpha < +\infty$. Then for a given $\varepsilon > 0$ there exists a finite set $F \subseteq I$ such that

$$\sum_{\alpha\in F} x_\alpha \geq \sum_{\alpha\in I} x_\alpha - \varepsilon .$$

Hence,

$$\sum_{\alpha\in I} x_\alpha \geq \sum_{\alpha\in G} x_\alpha \geq \sum_{\alpha\in F} x_\alpha \geq \sum_{\alpha\in I} x_\alpha - \varepsilon \quad \text{for all finite } F \subseteq G \subseteq I .$$

Therefore

$$\sum_{\alpha\in I} x_\alpha = \sup\left[\sum_{\alpha\in F} x_\alpha : F \subseteq I \text{ finite}\right] .$$

Now assume that $\sum_{\alpha\in I} x_\alpha = +\infty$. Then for any given $M > 0$ there exists a finite $F \subseteq I$ such that $\sum_{\alpha\in F} x_\alpha \geq M$, which implies that $\sum_{\alpha\in G} x_\alpha \geq M$ for all finite $F \subseteq G \subseteq I$. Hence,

$$\sup\left[\sum_{\alpha\in F} x_\alpha : F \subseteq I \text{ finite}\right] = +\infty . \qquad \square$$

Remark 3.5.43. If I is uncountable and uncountably many x_α are different from zero, then $\sum_{\alpha\in I} x_\alpha$ cannot converge to a finite limit.

The next result is a fundamental inequality in the theory of Hilbert spaces and is known as "Bessel's inequality."

Proposition 3.5.44 (Bessel's inequality). *If H is an inner product space and $\{x_\alpha\}_{\alpha\in I} \subseteq H$ is an orthonormal set, then $\sum_{\alpha\in I} |(x, x_\alpha)|^2 \leq \|x\|^2$ for all $x \in H$.*

Proof. On account of Lemma 3.5.42, we may assume that I is finite. Let $u = \sum_{\alpha\in I}(x, x_\alpha)x_\alpha$. Then $(x, u) = \sum_{\alpha\in I}(x, x_\alpha)^2 = (u, u)$; see Theorem 3.5.13. Therefore $x - u \perp u$ and so $\|x\|^2 = \|x - u\|^2 + \|u\|^2$ due to the Generalized Pythagorean Theorem; see Theorem 3.5.13. Hence, $\|x\|^2 \geq \|u\|^2 = (u, u) = \sum_{\alpha\in I}(x, x_\alpha)^2$. $\square$

Corollary 3.5.45. *If H is an inner product space and $\{x_\alpha\}_{\alpha\in I} \subseteq H$ is an orthonormal set, then for every $x \in H$ the set $\{\alpha \in I : (x, x_\alpha) \neq 0\}$ is countable.*

Remark 3.5.46. We have already mentioned that every Hilbert space has an orthonormal basis. In fact all orthonormal bases of a Hilbert space have the same cardinality, that is, every maximal orthonormal set in an inner product space has the same cardinality.

Proposition 3.5.47. *If H is a separable inner product space, then every orthonormal set in H is countable.*

Proof. Let $B = \{x_\alpha\}_{\alpha\in I} \subseteq H$ be an orthonormal set and let $D = \{u_n\}_{n\geq 1} \subseteq H$ be dense, which is possible since H is separable. Then for any $\alpha \in I$, $\overline{B}_{1/2}(x_\alpha) \cap D \neq \emptyset$. So, there exists $n_\alpha \in \mathbb{N}$ such that $\|x_\alpha - u_{n_\alpha}\| \leq 1/2$.

Let $\varphi\colon I \to \mathbb{N}$ be defined by $\varphi(\alpha) = n_\alpha$. We claim that φ is injective. Using the parallelogram law and the Generalized Pythagorean Theorem, it follows that

$$\begin{aligned}\sqrt{2} &= \|x_\alpha - x_\beta\| = \|x_\alpha - u_{n_\alpha} - x_\beta + u_{n_\beta} + u_{n_\alpha} - u_{n_\beta}\| \\ &\leq \|x_\alpha - u_{n_\alpha}\| + \|x_\beta - u_{n_\beta}\| + \|u_{n_\alpha} - u_{n_\beta}\| \leq 1 + \|u_{n_\alpha} - u_{n_\beta}\| \,.\end{aligned}$$

Hence, $\sqrt{2} - 1 \leq \|u_{n_\alpha} - u_{n_\beta}\|$ for all $\alpha, \beta \in I$ with $\alpha \neq \beta$. This proves the injectivity of φ, which means that $\operatorname{card} I \leq \operatorname{card} \mathbb{N}$ and so I is countable. □

This leads to the following useful characterization of separable Hilbert spaces.

Theorem 3.5.48. *A Hilbert space H is separable if and only if it has a countable orthonormal basis.*

Given a linearly independent sequence one can produce an orthonormal set with the same linear span. The process to achieve this is known as the “Gram–Schmidt Orthonormalization Process.”

Proposition 3.5.49 (Gram–Schmidt Orthonormalization Process). *If H is an inner product space and $\{u_n\}_{n\geq 1} \subseteq H$ are linearly independent, then there exists an orthonormal sequence $\{x_n\}_{n\geq 1} \subseteq H$ such that* $\operatorname{span}\{u_n\}_{n\geq 1} = \operatorname{span}\{x_n\}_{n\geq 1}$.

Proof. Let $x_1 = u_1/\|u_1\|$. So, the result holds for $n = 1$. Proceeding by induction suppose that we have produced $x_1, \ldots, x_{n-1}$. Then we set

$$h_n = u_n - \sum_{k=1}^{n-1} (u_k, x_k) x_k \,.$$

Evidently $h_n \perp x_k$ for all $k = 1, \ldots, n-1$ and $h_n \neq 0$ since $u_n \notin \operatorname{span}\{u_k\}_{k=1}^{n-1}$, due to the linear independence of the sequence $\{u_n\}_{n\geq 1} \subseteq H$. According to the induction hypothesis, we have $\operatorname{span}\{u_k\}_{k=1}^{n-1} = \operatorname{span}\{x_k\}_{k=1}^{n-1}$. Let $x_n = h_n/\|h_n\|$. Then by induction we have produced the desired orthonormal set $\{x_n\}_{n\geq 1} \subseteq H$. □

We conclude this section with a brief look at the notion of the basis for a vector space X. If X is finite dimensional, then it is well-known that a basis is a set $\{e_k\}_{k=1}^n$ such that every $x \in X$ can be written in a unique way as $x = \sum_{k=1}^n \lambda_k e_k$ with $\lambda_k \in \mathbb{R}$ known as the coordinates of x for the given basis. How do we extend this notion to infinite dimensional vector spaces?

Definition 3.5.50. (a) Given a vector space X, a **Hamel basis** is a set $\{e_\alpha\}_{\alpha\in I} \subseteq X$ such that every $x \in X$ can be written in a unique way as $x = \sum_{\alpha\in I} \lambda_\alpha x_\alpha$ with only finite numbers of the real λ_α different from zero. If X is finite dimensional, then a Hamel basis is the usual basis. But in infinite dimensional spaces there are no obvious Hamel bases although they can be shown to exist via Zorn's Lemma.

(b) Let X be a Banach space. A sequence $\{x_n\}_{n\geq 1} \subseteq X$ is a **Schauder basis** for X if for each $x \in X$ there exists a unique sequence $\{\lambda_n\}_{n\geq 1} \subseteq \mathbb{R}$ such that $x = \sum_{n\geq 1} \lambda_n x_n$.

Remark 3.5.51. The Hamel basis is an algebraic notion that does not relate to any topology. A Banach space with a Schauder basis is necessarily separable. Banach [25, p. 111] asked if every infinite dimensional separable Banach space has a Schauder basis. This question was settled in the negative by Enflo [104] who produced a separable reflexive Banach space with no Schauder basis.

3.6 Bounded and Unbounded Linear Operators

Let X, Y be Banach spaces. Recall that by $L(X, Y)$ we denote the Banach space of all bounded linear operators from X into Y. The norm of $L(X, Y)$ is defined by

$$\|A\|_L = \sup \left[\frac{\|A(x)\|_Y}{\|x\|_X} : x \in X \setminus \{0\} \right] ; \tag{3.6.1}$$

see Definition 3.1.45. If $X = Y$, we write $L(X, X) = L(X)$.

Definition 3.6.1. (a) The norm (metric) topology induced on $L(X, Y)$ by the norm $\|\cdot\|_L$ (see (3.6.1)) is called the **uniform operator topology** or simply the **norm topology**.

(b) The **strong operator topology** on $L(X, Y)$ is the weakest topology on $L(X, Y)$ for which the maps $e_x \colon L(X, Y) \to Y$ with $x \in X$ defined by $e_x(A) = A(x)$ for all $A \in L(X, Y)$ are continuous. Then a local basis at the origin consists of the sets

$$\{A \in L(X, Y) \colon \|A(x_k)\|_Y < \varepsilon \text{ for } k = 1, \dots, n\}$$

with $n \in \mathbb{N}$ and $\varepsilon > 0$. A net $\{A_\alpha\}_{\alpha\in I} \subseteq L(X, Y)$ converges to $A \in L(X, Y)$ in this topology if and only if $\|A_\alpha(x) - A(x)\|_Y \to 0$ for all $x \in X$. We write $A_\alpha \xrightarrow{s} A$ in $L(X, Y)$.

(c) The **weak operator topology** on $L(X, Y)$ is the weakest topology on $L(X, Y)$ for which the maps $e_{x,y^*} : L(X, Y) \to \mathbb{R}$ with $x \in X$ and $y^* \in Y^*$ defined by $e_{x,y^*}(A) = \langle y^*, A(x)\rangle$ are continuous. Then a local basis at the origin consists of the sets

$$\{A \in L(X, Y) \colon |\langle y_i^*, A(x_k)| < \varepsilon \text{ for } k = 1, \dots, n,\ i = 1, \dots, m\}$$

with $n, m \in \mathbb{N}$ and $\varepsilon > 0$. A net $\{A_\alpha\}_{\alpha\in I} \subseteq L(X, Y)$ converges to $A \in L(X, Y)$ in this topology if and only if $|\langle y^*, A_\alpha(x)\rangle - \langle y^*, A(x)\rangle| \to 0$ for all $x \in X, y^* \in Y^*$. We write $A_\alpha \xrightarrow{w} A$ in $L(X, Y)$.

Remark 3.6.2. Evidently it holds that

$$\text{weak topology} \subseteq \text{strong topology} \subseteq \text{norm topology.}$$

We should not confuse the weak operator topology with the weak topology that we can define on the Banach space $L(X, Y)$. Let V be a third Banach space and consider the map $\vartheta\colon L(X, Y) \times L(Y, V) \to L(X, V)$ defined by $\vartheta(A, B) = B \circ A$. Then ϑ is jointly continuous for the uniform operator topology but only separately continuous for the strong and weak operator topologies. In general, the strong and weak operator topologies are not first countable and this complicates their study.

Proposition 3.6.3. *If H is a Hilbert space and $\{A_n\}_{n\geq 1} \subseteq L(H)$ is a sequence such that $\{(y, A_n(x))\}_{n\geq 1}$ is convergent for all $x, y \in H$, then there exists $A \in L(H)$ such that $A_n \xrightarrow{\mathrm{w}} A$.*

Proof. For given $x, y \in H$ we derive $\sup_{n\geq 1} |(y, A_n(x))| < \infty$. Invoking Theorem 3.2.1 we obtain that $\sup_{n\geq 1} \|A_n(x)\| < \infty$. A second application of Theorem 3.2.1 gives $\sup_{n\geq 1} \|A_n\|_L < \infty$.

Let $\xi(x, y) = \lim_{n\to\infty}(y, A_n(x))$. Evidently ξ is bilinear and

$$|\xi(x, y)| \leq \limsup_{n\to\infty} |(y, A_n(x))| \leq \|y\|\|x\|(\sup_{n\geq 1} \|A_n\|_L) .$$

Hence ξ is bounded. Then there exists $A \in L(H)$ such that $(y, A(x)) = \xi(x, y)$; see Theorem 3.5.21. Therefore we get $A_n \xrightarrow{\mathrm{w}} A$ in $L(X, Y)$. □

In a similar way we obtain the corresponding result for the strong operator topology.

Proposition 3.6.4. *If X, Y are Banach spaces, $\{A_n\}_{n\geq 1} \subseteq L(X, Y)$ and $\{A_n(x)\}_{n\geq 1} \subseteq Y$ is a Cauchy sequence for each $x \in X$, then there exists $A \in L(X, Y)$ such that $A_n \xrightarrow{s} A$.*

Remark 3.6.5. Both results fail for nets of operators.

Definition 3.6.6. Let X, Y be normed spaces and $A \in L(X, Y)$. The **adjoint** (or **dual**) operator of A is the unique operator $A^*\colon Y^* \to X^*$ defined by

$$A^*(y^*) = y^* \circ A \quad \text{for all } y^* \in Y^* .$$

Continuing, the **second adjoint** (or **second dual** or **bidual**) $(A^*)^*$ of A is the unique linear map $A^{**}\colon X^{**} \to Y^{**}$ such that

$$A^{**}(x^{**}) = x^{**} \circ A^* \quad \text{for all } x^{**} \in X^{**} .$$

The next proposition summarizes the main properties of A^* and A^{**}.

Proposition 3.6.7. *If X, Y are normed spaces and $A, S, T \in L(X, Y)$, then the following hold:*

(a) *$A^* \in L(Y^*, X^*)$ and $\|A^*\|_L = \|A\|_L$.*

(b) *If $\lambda_1, \lambda_2 \in \mathbb{R}$, then $(\lambda_1 S + \lambda_2 T)^* = \lambda_1 S^* + \lambda_2 T^*$.*

(c) $A^{**}|_X = A$.
(d) *If V is a third normed space and $B \in L(Y, V)$, then* $(B \circ A)^* = A^* \circ B^*$.
(e) *If A is invertible, that is, A^{-1} exists and $A^{-1} \in L(Y, X)$, then A^* is invertible as well and* $(A^*)^{-1} = (A^{-1})^*$.

Proof. (a) For all $x \in X$ and for all $y^* \in Y^*$ one gets

$$\langle A^*(y^*), x\rangle = \langle y^*, A(x)\rangle \le \|y^*\|_* \|A(x)\| \le \|y^*\|_* \|A\|_L \|x\| ,$$

hence $\|A^*(y^*)\| \le \|y^*\|_* \|A\|_L$, and so $\|A^*\|_L \le \|A\|_L$. Given $\varepsilon > 0$ there exists $x_0 \in X$ with $\|x_0\| = 1$ such that $\|A\|_L - \varepsilon \le \|A(x_0)\|$. Let $y^* \in Y^*$ with $\|y^*\|_* = 1$ such that $\langle y^*, A(x_0)\rangle = \|A(x_0)\|$; see Proposition 3.1.50. Then it follows that

$$\langle A^*(y^*), x_0\rangle = \langle y^*, A(x_0)\rangle = \|A(x_0)\| \ge \|A\|_L - \varepsilon ,$$

which gives $\|A^*\|_L \ge \|A\|_L - \varepsilon$. Letting $\varepsilon \searrow 0$, we obtain $\|A^*\|_L \ge \|A\|_L$. Therefore, $A^* \in L(Y^*, X^*)$ and $\|A^*\|_L = \|A\|_L$.

(b) This follows immediately from Definition 3.6.6.

(c) This is also clear from Definition 3.6.6.

(d) For all $x \in X$ and for all $v^* \in V^*$ we derive

$$\langle A^*(B^*(v^*)), x\rangle = \langle B^*(v^*), A(x)\rangle = \langle v^*, B(A(x))\rangle$$

and so we conclude that $A^* \circ B^* = (B \circ A)^*$.

(e) Since A is invertible we have $A^{-1} \circ A = i_X = A \circ A^{-1}$. Then using part (c) we obtain

$$A^* \circ \left(A^{-1}\right)^* = i_X^* = i_{X^*} = \left(A^{-1}\right)^* \circ A^* .$$

Hence, A^* is invertible and $(A^*)^{-1} = \left(A^{-1}\right)^*$. □

Remark 3.6.8. According to this proposition the map $A \to A^*$ from $L(X, Y)$ into $L(Y^*, X^*)$ is an isometric isomorphism. It is also continuous for the weak operator topologies but not for the strong operator topology. When $X = Y = H$ is a complex Hilbert space, that is, over $\mathbb{F} = \mathbb{C}$, then, since H is self-dual, that is, $H = H^*$, we want to define A^* on the space H. From the Riesz-Fréchet Representation Theorem (see Theorem 3.5.21), we know that H is isometric with its dual H^* but the isometry is a conjugate isomorphism $j\colon H \to H^*$. We set $A' = j^{-1} \circ A^* \circ j$ and get that

$$\begin{aligned}(x, A(y)) &= \langle j(x), A(y)\rangle = \langle A^*(j(x)), y\rangle = (j^{-1}(A^*(j(x))), y)\\ &= (A'(x), y)\end{aligned} \tag{3.6.2}$$

for all $x, y \in H$. Then $A' \in L(H)$ is the Hilbert space adjoint and now the map $A \to A'$ is conjugate linear, that is, $\lambda A \to \overline{\lambda} A'$ for all $\lambda \in \mathbb{C}$ because A' is defined on H rather than on H^* and H is identified with H^* by a conjugate isometric isomorphism. However, in what follows for notational uniformity we denote A' by A^* with the understanding that A^* is defined on H. When H is a real Hilbert space, we define again $A' = A^*$ on H as above.

Proposition 3.6.9. *If H is a Hilbert space over $\mathbb{R}$ or $\mathbb{C}$ and if $A \in L(H)$, then $\|A\|_L^2 = \|A^* \circ A\|_L$.*

Proof. Taking Proposition 3.6.7(a) and (3.6.2) into account yields

$$\begin{aligned}\|A\|_L^2 &= \sup\left[\|A(x)\|: \|x\| \le 1\right] = \sup\left[(A(x), A(x)): \|x\| \le 1\right] \\ &= \sup\left[(A^*(A(x)), x): \|x\| \le 1\right] \le \|A^* \circ A\|_L \le \|A^*\|_L \|A\|_L = \|A\|_L^2 .\end{aligned}$$

□

Example 3.6.10. Section 4.1 shows that $(l^1)^* = l^\infty$. Consider the right shift operator $A \in L(l^1)$ defined by $A(\hat{x}) = (0, x_1, x_2, \dots)$ for all $\hat{x} = (x_n)_{n\ge1} \in l^1$. Then $A^*: l^\infty \to l^\infty$ is defined by $A^*(\hat{u}) = (u_2, u_3, \dots)$ for all $\hat{u} = (u_n)_{n\ge1} \in l^\infty$. In this case we have $\|A\|_L = \|A^*\|_L = 1$.

Proposition 3.6.11. *If X, Y are normed spaces and $A \in L(X, Y)$, then $A^* \in L(Y^*, X^*)$ is weak*-to-weak* continuous.*

Conversely, if $T: Y^* \to X^*$ is a weak*-to-weak* continuous linear operator, then there exists $A \in L(X, Y)$ such that $A^* = T$.

Proof. Let $\{y_\alpha^*\}_{\alpha\in I} \subseteq Y^*$ be a net such that $y_\alpha^* \xrightarrow{w^*} y^*$ in Y^*. Then for every $x \in X$, it follows that

$$\langle A^*(y_\alpha^*), x\rangle = \langle y_\alpha^*, A(x)\rangle \to \langle y^*, A(x)\rangle = \langle A^*(y^*), x\rangle ,$$

hence, $A^*(y_\alpha^*) \xrightarrow{w^*} A^*(y^*)$ and so A^* is weak*-to-weak* continuous.

Let $j_X: X \to X^{**}$ and $j_Y: Y \to Y^{**}$ be the canonical embeddings; see Definition 3.3.35. For every $x \in X$, $j_X(x)T$ is a w*-continuous linear functional on Y^*, hence $j_X(x)T \in j_Y(Y)$. Then $j_Y^{-1}(j_X(x)T) \in Y$. So, we can define an operator $A: X \to Y$ by setting $A(x) = j_Y^{-1}(j_X(x)T)$ for all $x \in X$. Clearly A is linear. Moreover, let $\{x_\alpha\}_{\alpha\in I} \subseteq X$ be a net such that $x_\alpha \xrightarrow{w} x$. Then $j_X(x_\alpha) \xrightarrow{w^*} j_X(x)$; see Proposition 3.3.23. Hence, for all $y^* \in Y^*$, we have

$$(j_X(x_\alpha)T)(y^*) \to (j_X(x)T)(y^*) \quad \text{in } \mathbb{R} ,$$

thus $j_X(x_\alpha)T \xrightarrow{w^*} j_X(x)T$ in Y^{**}. Therefore,

$$A(x_\alpha) = j_Y^{-1}(j_X(x_\alpha)T) \xrightarrow{w} j_Y^{-1}(j_X(x)T) = A(x) \quad \text{in } Y .$$

This means that $A: X \to Y$ is weak-to-weak continuous, hence $A \in L(X, Y)$; see Proposition 3.3.23. Moreover, with view to Definition 3.3.35, we get

$$\langle A^*(y^*), x\rangle = \langle y^*, A(x)\rangle = \langle y^*, j_Y^{-1}(j_X(x)T)\rangle = \langle j_X(x)T, y^*\rangle = \langle T(y^*), x\rangle .$$

Thus, $A^* = T$. □

Corollary 3.6.12. *If X, Y are normed spaces and $S: X^* \to Y^*$ is weak*-to-weak* continuous, then $S \in L(X^*, Y^*)$.*

Next we introduce some important special classes of linear operators.

Definition 3.6.13. (a) Let X be a vector space and let $P\colon X \to X$ be a linear operator. We say that P is a **projection** if $P^2 = P$, that is, $P(P(x)) = P(x)$ for all $x \in X$.
(b) Let H be a Hilbert space and $A \in L(H)$. We say that A is **self-adjoint** (or **hermitian**) if $A = A^*$, that is, $(A(x), y) = (x, A(y))$ for all $x, y \in H$.
(c) Let H be a Hilbert space and $P \in L(H)$. We say that P is an **orthogonal projection** if P is a projection and P is self-adjoint.

Proposition 3.6.14. *If H is a Hilbert space and $T, S \in L(H)$ are self-adjoint and commuting, that is, $T \circ S = S \circ T$, then $T \circ S \in L(H)$ is self-adjoint as well.*

Proof. For every $x, y \in H$ we see that

$$(T(S(x)), y) = (S(x), T(y)) = (x, S(T(y))) = (x, T(S(y))) .$$

This shows that $T \circ S$ is self-adjoint. □

Proposition 3.6.15. *If H is a Hilbert space and $A \in L(H)$ is self-adjoint, then for every $m \in \mathbb{N}$, A^m is self-adjoint and $\|A^m\|_L = \|A\|^m$.*

Proof. That A^m is self-adjoint for every $m \in \mathbb{N}$ follows from Proposition 3.6.14. From Proposition 3.6.9 we see that

$$\|A\|_L^2 = \left\|A^* \circ A\right\|_L = \left\|A^2\right\|_L , \quad \left\|A^4\right\|_L = \left\|A^2\right\|_L^2 = \|A\|_L^4$$

and so on. Therefore we obtain

$$\left\|A^{2^n}\right\|_L = \|A\|_L^{2^n} . \tag{3.6.3}$$

If $1 \le m \le 2^n$, then

$$\left\|A^{2^n}\right\|_L = \left\|A^m \circ A^{2^n - m}\right\|_L \le \left\|A^m\right\|_L \|A\|_L^{2^n-m} \le \|A\|_L^m \|A\|_L^{2^n-m} = \|A\|_L^{2^n} ,$$

which, due to (3.6.3), results in

$$\left\|A^m\right\|_L \|A\|_L^{2^n-m} = \|A\|_L^{2^n} .$$

Thus, $\|A^m\|_L = \|A\|_L^m$. □

Proposition 3.6.16. *If H is a Hilbert space and $A \in L(H)$ is self-adjoint, then $\|A\|_L = \sup\left[|(A(x), x)|\colon \|x\| \le 1\right]$.*

Proof. For $x \in H$ with $\|x\| \le 1$ we infer

$$|(A(x), x)| \le \|A(x)\|\|x\| \le \|A\|_L\|x\|^2 \le \|A\|_L ,$$

which gives

$$\sup\left[|(A(x), x)|\colon \|x\| \le 1\right] \le \|A\|_L . \tag{3.6.4}$$

Let $\eta = \sup\left[|(A(x), x)|\colon \|x\| \le 1\right]$. Then $|(A(u), u)| \le \eta\|u\|^2$ for all $u \in H$. For $u \in H$ with $u \ne 0$ let $\lambda = (\|A(u)\|/\|u\|)^{1/2}$ and $y = 1/\lambda A(u)$. Since A is self-adjoint, $(A(\lambda u), y) \in \mathbb{R}$

and, due to the parallelogram law, we obtain

$$\begin{aligned}\|A(u)\|^2 &= (A(u), A(u)) = \left(A(\lambda u), \frac{1}{\lambda}A(u)\right) = (A(\lambda u), y)\\ &= \frac{1}{4}[(A(\lambda u + y), \lambda u + y) - (A(\lambda u - y), \lambda u - y)]\\ &\leq \frac{1}{4}\eta\left(\|\lambda u + y\|^2 + \|\lambda u - y\|^2\right) = \frac{1}{2}\eta\left(\|\lambda u\|^2 + \|y\|^2\right)\\ &= \frac{1}{2}\eta\left(\lambda^2\|u\|^2 + \frac{1}{\lambda^2}\|(A(u)\|^2\right) = \eta\|u\|\|A(u)\|\,,\end{aligned}$$

where we used the fact that $\lambda\|u\| = 1/\lambda\|A(u)\|$. Hence $\|A(u)\| \leq \eta\|u\|$ for every $u \in H$, which gives $\|A\|_L \leq \eta$ and so, because of (3.6.4), the result follows. □

Next we present a useful factorization result.

Proposition 3.6.17. *If X, Y, V are Banach spaces, $A \in L(X, Y)$, $T \in L(V, Y)$, and A is injective, then the following statements are equivalent:*
(a) $R(T) \subseteq R(A)$.
(b) *There exists $S \in L(V, X)$ such that $A \circ S = T$.*

Proof. (a) $\Longrightarrow$ (b): Let $S = A^{-1} \circ T\colon V \to X$, where we recall that A is injective. Then S is linear and $A \circ S = T$. We claim that $\operatorname{Gr} S \subseteq V \times X$ is closed. To this end, let $\{v_n\}_{n\geq 1} \subseteq V$ such that $v_n \to v$ in V and $S(v_n) \to x$ in X. Then $A(x) = \lim_{n\to\infty} A(S(v_n)) = \lim_{n\to\infty} T(v_n) = T(v) = A(S(v))$. Since A is injective it follows that $x = S(v)$ and so $\operatorname{Gr} S \subseteq V \times X$ is closed. Hence, by the Closed Graph Theorem (see Theorem 3.2.14), we conclude that $S \in L(V, X)$.

(b) $\Longrightarrow$ (a): It holds that $R(T) = R(A \circ S) \subseteq R(A)$. □

Next we present two theorems relating operators with the same range space and their adjoints. We start with an auxiliary result.

Lemma 3.6.18. *If X, Y are normed spaces, $A \in L(X, Y)$ and $x^* \in X^*$, then the following statements are equivalent:*
(a) $x^* \in R(A^*)$.
(b) $|\langle x^*, x\rangle| \leq c\|A(x)\|_Y$ *for all $x \in X$ and for some $c > 0$.*

Proof. (a) $\Longrightarrow$ (b): Of course, $x^* = A^*(y^*)$ for some $y^* \in Y^*$. Then

$$|\langle x^*, x\rangle| = |\langle A^*(y^*), x\rangle| = |\langle y^*, A(x)\rangle| \leq \|y^*\|_*\|A(x)\|_Y \quad \text{for all } x \in X\,,$$

which gives $|\langle x^*, x\rangle| \leq c\|A(x)\|_Y$ with $c = \|y^*\|_*$.

(b) $\Longrightarrow$ (a): There exists a continuous, linear functional $g\colon R(A) \to \mathbb{R}$ such that $x^* = g \circ A$. According to Proposition 3.1.49, there exists $y^* \in Y^*$ such that $y^*|_{R(A)} = g$. Then $x^* = y^* \circ A = A^*(y^*)$; see Definition 3.6.6. □

Theorem 3.6.19. *If X, Y, V are Banach spaces and $A \in L(X, Y)$, $T \in L(V, Y)$ with $R(T) \subseteq R(A)$, then $\|T^*(y^*)\|_* \leq c\|A^*(y^*)\|_*$ for all $y^* \in Y^*$ and for some $c > 0$.*

Proof. Let $\hat{X} = X/N(A)$ with $N(A)$ being the kernel of A and $p\colon X \to \hat{X}$ being the quotient map. Then $p^*\colon \hat{X}^* \to X^*$ is an isometric embedding onto $N(A)^\perp \subseteq X^*$; see Proposition 3.2.25. Let $\hat{A}\colon \hat{X} \to Y$ be defined by $\hat{A} \circ p = A$. Then $A^* = p^* \circ \hat{A}^*$, and so

$$\left\|A^*(y^*)\right\|_* = \left\|\hat{A}^*(y^*)\right\|_* \quad \text{for all } y^* \in Y^* .$$

By hypothesis, $R(T) \subseteq R(A) = R(\hat{A})$ and $\hat{A}$ is injective. So, we can use Proposition 3.6.17 and produce $S \in L(V, \hat{X})$ such that $\hat{A} \circ S = T$. Then, since $\hat{A} \circ S = T$,

$$\begin{aligned}
\|T^*(y^*)\|_* &= \sup\left[\frac{\langle T^*(y^*), v\rangle}{\|v\|_V} : v \in V, v \neq 0\right] \\
&= \sup\left[\frac{\langle y^*, T(v)\rangle}{\|v\|_V} : v \in V, v \neq 0\right] \\
&= \sup\left[\frac{\langle \hat{A}^*(y^*), S(v)\rangle}{\|v\|_V} : v \in V, v \neq 0\right] \\
&\leq \sup\left[\frac{\|\hat{A}^*(y^*)\|_* \|S(v)\|_{\hat{X}}}{\|v\|_V} : v \in V, v \neq 0\right] \\
&= \|S\|_L \|\hat{A}^*(y^*)\|_* \quad \text{for all } y^* \in Y^* .
\end{aligned}$$

So, the conclusion of the theorem holds with $c = \|S\|_L$. ☐

Theorem 3.6.20. *If X, Y, V are normed spaces and $A \in L(X, Y)$, $T \in L(X, V)$, then the following statements are equivalent:*
(a) $R(T^*) \subseteq R(A^*)$.
(b) $\|T(x)\|_V \leq c\|A(x)\|_Y$ *for all $x \in X$ and for some $c > 0$.*

Proof. (a) $\Longrightarrow$ (b): Using Theorem 3.6.19, we infer

$$\left\|T^{**}(x^{**})\right\|_{**} \leq c\left\|A^{**}(x^{**})\right\|_{**} \quad \text{for all } x^{**} \in X^{**} \text{ and for some } c > 0 .$$

Applying Proposition 3.6.7 gives

$$\|T(x)\|_V = \left\|T^{**}(x)\right\|_{**} \leq c\left\|A^{**}(x)\right\|_{**} = c\|A(x)\|_Y \quad \text{for all } x \in X .$$

(b) $\Longrightarrow$ (a): Let $x^* \in R(T^*) \subseteq X^*$. Using Lemma 3.6.18 yields

$$|\langle x^*, x\rangle| \leq c_0\|T(x)\|_V \quad \text{for all } x \in X \text{ and for some } c_0 > 0 ,$$

which implies

$$|\langle x^*, x\rangle| \leq c_0 c\|A(x)\|_V \quad \text{for all } x \in X .$$

Hence, with view to Lemma 3.6.18 we see that $x^* \in R(A^*)$. Thus, $R(T^*) \subseteq R(A^*)$. ☐

Theorem 3.6.21. *If X, Y, V are Banach spaces, X is reflexive, $A \in L(X, Y)$, $T \in L(V, Y)$ and*

$$\left\|T^*(y^*)\right\|_* \leq c\left\|A^*(y^*)\right\|_* \quad \textit{for all } y^* \in Y^* \textit{ and for some } c > 0 ,$$

then $R(T) \subseteq R(A)$.

Proof. Applying Theorem 3.6.20 we obtain $R(T^{**}) \subseteq R(A^{**})$. Let $v \in V$ and let $x \in X^{**} = X$ such that $A(x) = A^{**}(x) = T^{**}(v) = T(v)$; see Proposition 3.6.7(c). Hence $R(T) \subseteq R(A)$. □

Motivated from Definition 3.2.24, we introduce a similar notion for sets in X^*.

Definition 3.6.22. Let X be a normed space and $E \subseteq X^*$. The **preannihilator** of E is defined by

$$^{\perp}E = \{x \in X : \langle x^*, x\rangle = 0 \text{ for all } x^* \in E\} .$$

Evidently $^{\perp}E$ is a closed linear subspace of X.

Remark 3.6.23. It is easy to see that if $E \subseteq X^*$ is a vector subspace, then $\overline{E}^{w^*} = (^{\perp}E)^{\perp}$, E is w*-closed if and only if $E = (^{\perp}E)^{\perp}$, and $\overline{E}^{w^*} = X^*$ if and only if $^{\perp}E = \{0\}$.

Moreover, if Y, V are closed vector subspaces of X, then

$$V \cap Y = {}^{\perp}(V^{\perp} + Y^{\perp}) , \quad (V \cap Y)^{\perp} \supseteq \overline{V^{\perp} + Y^{\perp}} ,$$
$$V^{\perp} \cap Y^{\perp} = (V + Y)^{\perp} , \quad {}^{\perp}(V^{\perp} \cap Y^{\perp}) = \overline{V + Y} .$$

Proposition 3.6.24. *If X, Y are normed spaces and $A \in L(X, V)$, then the following hold:*
(a) $R(A)^{\perp} = N(A^*)$ *and* $^{\perp}R(A^*) = N(A)$.
(b) $\overline{R(A)} = Y$ *if and only if A^* is injective.*
(c) *A is injective if and only if* $\overline{R(A^*)}^{w^*} = X^*$.

Proof. (a) Note that

$$\begin{aligned} y^* \in R(A)^{\perp} \quad &\text{if and only if} \quad \langle y^*, A(x)\rangle = 0 \quad \text{for all } x \in X \\ &\text{if and only if} \quad \langle A^*(y^*), x\rangle = 0 \quad \text{for all } x \in X \\ &\text{if and only if} \quad A^*(y^*) = 0 . \end{aligned}$$

Hence $R(A)^{\perp} = N(A^*)$. Similarly, we have

$$\begin{aligned} x \in {}^{\perp}R(A^*) \quad &\text{if and only if} \quad \langle A^*(y^*), x\rangle = 0 \quad \text{for all } y^* \in Y^* \\ &\text{if and only if} \quad \langle y^*, A(x)\rangle = 0 \quad \text{for all } y^* \in Y^* \\ &\text{if and only if} \quad A(x) = 0 . \end{aligned}$$

Thus, $^{\perp}R(A^*) = N(A)$.

(b) $\Longrightarrow$: It holds that $R(A)^{\perp} = \{0\}$ and so with part (a), $N(A^*) = \{0\}$. Hence A^* is injective.

$\Longleftarrow$: It holds that $N(A^*) = \{0\}$ and so with part (a), $R(A)^{\perp} = \{0\}$. Hence $\overline{R(A)} = Y$.

(c) $\Longrightarrow$: It holds that $N(A) = \{0\}$ and so with part (a), $^{\perp}R(A^*) = \{0\}$. Hence, $\overline{R(A^*)}^{w^*} = X^*$; see Remark 3.6.23.

$\Longleftarrow$: It holds that $^{\perp}R(A^*) = \{0\}$ (see Remark 3.6.23), and so with part (a), $N(A) = \{0\}$. Hence, A is injective. □

Remark 3.6.25. If X, Y are Banach spaces with X or Y finite dimensional and $A \in L(X, Y)$, we know from linear algebra that

$$A \text{ is surjective if and only if } A^* \text{ is injective ,}$$
$$A^* \text{ is surjective if and only if } A \text{ is injective .}$$

Indeed in this case $R(A)$ is closed if $\dim Y < \infty$ and $R(A^*)$ is closed if $\dim X < \infty$ and so the equivalences above follow from Proposition 3.6.24. In the general infinite dimensional case we only have the following implications (see Proposition 3.6.24(a))

$$A \text{ is surjective} \quad \Longrightarrow \quad A^* \text{ is injective ,}$$
$$A^* \text{ is surjective} \quad \Longrightarrow \quad A \text{ is injective .}$$

The reverse implications fail. To see this, let $X = Y = H = l^2$, which is a Hilbert space and let $A \in L(H)$ be defined by $A(\hat{x}) = (1/nx_n)_{n\geq 1}$ for all $\hat{x} = (x_n)_{n\geq 1} \in l^2$. Then $A^* = A$ and A is injective but not surjective since $R(A) = R(A^*)$ is only dense in H.

Next we present some results dealing with the basic properties of projections.

Proposition 3.6.26. *If X is a normed space and $P \in L(X)$, then P is a projection if and only if $P^* \in L(X^*)$ is a projection.*

Proof. $\Longrightarrow$: For all $x \in X$ and for all $x^* \in X^*$ we directly obtain

$$\langle P^*(x^*), x\rangle = \langle x^*, P(x)\rangle = \langle x^*, P(P(x))\rangle = \langle P^*(P^*(x^*)), x\rangle \ .$$

This shows that $P^*(x^*) = P^*(P^*(x^*))$ for all $x^* \in X^*$. Hence P^* is a projection as well.

$\Longleftarrow$: This is proven in a similar fashion. ☐

Proposition 3.6.27. *If X is a normed space and $P \in L(X)$, then P is a projection if and only if $I - P$ is a projection.*

Proof. $\Longrightarrow$: For every $x \in X$ one gets

$$(I - P)(I - P)(x) = x - 2P(x) + P(P(x)) = x - P(x) = (I - P)(x) \ .$$

Hence $I - P$ is a projection.

$\Longleftarrow$: Note that $P = I - (I - P)$ and so the implication follows from the previous part. ☐

Proposition 3.6.28. *If X is a normed space and $P \in L(X)$ is a projection, then $N(P) = R(I - P)$ and $R(P) = N(I - P)$.*

Proof. Let $x \in N(p)$. Then $(I - P)(x) = x$ and so $N(P) \subseteq R(I - P)$. Let $u \in R(I - P)$. Then $u = (I - P)(x)$ with $x \in X$. Then $P(u) = P(x - P(x)) = P(x) - P(P(x)) = P(x) - P(x) = 0$ and so $u \in N(p)$. Therefore we conclude that $N(P) = R(I - P)$. Applying this result to the projection $I - P$ we get $R(P) = N(I - P)$. ☐

Corollary 3.6.29. *If X is a normed space and $P \in L(X)$ is a projection, then $R(P) = \{x \in X\colon P(x) = x\}$ and $R(P)$ is closed.*

Corollary 3.6.30. *If X is a Banach space and $P \in L(X)$ is a projection, then $X = N(P) \oplus R(P)$.*

If V and W are complementary subspaces of a Banach space X (see Definition 3.2.27), then we obtain in a unique way, for every $x \in X$, that $x = v + w$ with $v \in V$ and $w \in W$. Let $P_V\colon X \to V$ be the linear operator such that $P_V(x) = v$. Evidently $P_V^2 = P_V$.

Proposition 3.6.31. *$P_V \in L(X)$, that is, P_V is a projection.*

Proof. Suppose that $x_n \to x$ in X and $P_V(x_n) \to v$ in X. Then $(I - P_V)(x_n) \to x - y$ in X. Note that $v \in V$ and $x - v \in W$. So, $v = P_V(x)$ and by the Closed Graph Theorem (see Theorem 3.2.14), it follows that $P_V \in L(X)$. □

Corollary 3.6.32. *If X is a Banach space and $V \subseteq X$ is a subspace, then V is complemented if and only if $V = R(P)$ with $P \in L(X)$ being a projection.*

Corollary 3.6.33. *If X is a Banach space and $V, W \subseteq X$ are complementary subspaces, then V and X/W are isomorphic.*

Next we use complemented subspaces to obtain a kind of Hahn–Banach Extension Theorem for vector valued maps.

Proposition 3.6.34. *If X is a Banach space and $V \subseteq X$ is a subspace, then the following statements are equivalent:*
(a) *For every Banach space Y and every $A \in L(V, Y)$, there is $\hat{A} \in L(X, Y)$ such that $\hat{A}\big|_V = A$.*
(b) *$\overline{V}$ is complemented in X.*

Proof. (a) $\Longrightarrow$ (b): Let $i_0\colon V \to \overline{V}$ be the bounded linear operator defined by $i_0(v) = v$ for all $v \in V$, that is, the identity map on V. Then by hypothesis there exists $\hat{i}_0 \in L(X, \overline{V})$ such that $\hat{i}_0\big|_V = i_0$. Due to the continuity of $\hat{i}_0$ we directly obtain that $\hat{i}_0\big|_{\overline{V}}$ coincides with the identity operator of $\overline{V}$. Therefore, $\hat{i}_0 \in L(X)$ is a projection with $R(\hat{i}_0) = \overline{V}$. Then Corollary 3.6.32 implies that $\overline{V}$ is complemented.

(b) $\Longrightarrow$ (a): Corollary 3.6.32 implies that $\overline{V} = R(P)$ with $P \in L(X)$ being a projection. Let Y be a Banach space and $A \in L(V, Y)$. Then there exists $A_0 \in L(\overline{V}, Y)$ such that $A_0\big|_V = A$; see Theorem 1.5.27. One gets $A_0 \circ P \in L(X, Y)$ and $A_0 \circ P\big|_V = A$. So, $\hat{A} = A_0 \circ P \in L(X, Y)$. □

Proposition 3.6.35. *If X is a Banach space, Y is a normed space, and $A \in L(X, Y)$, then $A^{-1} \in L(Y, X)$ if and only if $R(A)$ is dense in Y and there exists $c > 0$ such that $\|A(x)\|_Y \geq c\|x\|_X$ for all $x \in X$.*

Proof. $\Longrightarrow$: This is obvious.

$\Longleftarrow$: Evidently A is injective and $A^{-1} \in L(V, X)$ with $V = R(A)$. Hence $A^{-1} \in L(Y, X)$ since by hypothesis $\overline{V} = Y$. Moreover, note that $\|A^{-1}\|_L \leq 1/c$. □

Using this proposition we can improve Proposition 3.6.7(e).

Proposition 3.6.36. *If X is a Banach space, Y is a normed space, and $A \in L(X, Y)$, then A is invertible if and only if A^* is invertible.*

Proof. $\Longrightarrow$: This follows from Proposition 3.6.7(e).

$\Longleftarrow$: From Proposition 3.6.24(a) one has that $R(A)^\perp = N(A^*) = \{0\}$ and so $R(A) \subseteq Y$ is dense. Let $x \in X$ and let $x^* \in X^*$ be such that

$$\langle x^*, x\rangle = \|x\|_X \quad \text{and} \quad \|x^*\| = 1\,;$$

see Proposition 3.1.50. Then

$$\begin{aligned}\|x\|_X = \langle x^*, x\rangle &= \langle A^*((A^*)^{-1}(x^*)), x\rangle = \langle (A^*)^{-1}(x^*), A(x)\rangle \\ &\le \|(A^*)^{-1}(x^*)\|_* \|A(x)\|_Y \le \|(A^*)^{-1}\|_L \|A(x)\|_Y\,.\end{aligned}$$

This implies $\|A(x)\|_Y \ge c\|x\|_X$ with $c = (\|(A^*)^{-1}\|_L)^{-1}$. Now we may apply Proposition 3.6.35 and conclude that $A^{-1} \in L(Y, X)$. □

Corollary 3.6.37. *If X is a Banach space, Y is a normed space, and $A \in L(X, Y)$, then the following statements are equivalent:*
(a) *A is invertible.*
(b) *A^* is invertible.*
(c) *There exist $c, \hat{c} > 0$ such that*

$$\begin{aligned}\|A(x)\|_Y &\ge c\|x\|_X && \text{for all } x \in X\,,\\ \|A^*(x^*)\|_* &\ge \hat{c}\|x^*\|_* && \text{for all } x^* \in X^*\,.\end{aligned}$$

In the last part of this section we deal with unbounded linear operators.

Definition 3.6.38. Let X, Y be Banach spaces. An **unbounded linear operator** is a linear map $A\colon D(A) \subseteq X \to Y$ from a linear subspace $D(A)$ into Y. The subspace $D(A)$ is called the **domain** of A. We say that A is **closed** if $\operatorname{Gr} A \subseteq X \times Y$ is closed. By $N(A)$ we denote the kernel of A, that is, $N(A) = \{x \in D(A)\colon A(x) = 0\}$ and by $R(A)$ the range of A, that is, $R(A) = \{A(x)\colon x \in D(A)\}$.

Remark 3.6.39. In this context, A is closed if and only if for every $\{x_n\}_{n\ge1} \subseteq D(A)$ such that $x_n \to x$ in X and $A(x_n) \to y$ in Y, it follows that $x \in D(A)$ and $A(x) = y$. Note that now it is not enough to check that if $x_n \to 0$ in X and $A(x_n) \to y$ in Y, then $y = 0$. Moreover, if A is closed, then $N(A)$ is closed but $R(A)$ need not be closed. In applications most unbounded linear operators are densely defined, that is, $\overline{D(A)} = X$, and closed.

We can extend the notion of adjoint to unbounded linear operators. So, let $A\colon D(A) \subseteq X \to Y$ be an unbounded linear operator that is densely defined, that is, $\overline{D(A)} = X$. Let

$$D(A^*) = \left\{y^* \in Y^*\colon |\langle y^*, A(x)\rangle| \le c\|x\| \text{ for all } x \in D(A) \text{ and for some } c > 0\right\}. \tag{3.6.5}$$

Evidently $D(A^*) \subseteq Y^*$ is a vector subspace. Let $y^* \in D(A^*)$ and consider the functional $f\colon D(A) \to \mathbb{R}$ defined by $f(x) = \langle y^*, A(x)\rangle$ for all $x \in D(A)$. Because of (3.6.5) it follows that $|f(x)| \le c\|x\|$ for all $x \in D(A)$. Since $D(A)$ is dense in X, extending by continuity, there exists a unique functional $\hat{f}: X \to \mathbb{R}$ such that $\hat{f}|_{D(A)} = f$ and $|\hat{f}(x)| \le c\|x\|$ for all $x \in X$. Thus, $\hat{f} \in X^*$. Then we set

$$A^*(y^*) = f^* \, . \tag{3.6.6}$$

Definition 3.6.40. The unbounded linear operator $A^*: D(A^*) \subseteq Y^* \to X^*$ defined by (3.6.6) is called the **adjoint** of A. So, according to the previous construction, we obtain

$$\langle y^*, A(x)\rangle = \langle A^*(y^*), x\rangle \quad \text{for all } x \in D(A) \text{ and for all } y^* \in D(A^*) \, . \tag{3.6.7}$$

Remark 3.6.41. In general, we cannot say that A^* is densely defined. However, if A is also closed, then $D(A^*)$ is w*-dense in Y^*. Therefore, if Y is reflexive and $A: D(A) \subseteq X \to Y$ is closed and densely defined, then $A^*: D(A^*) \subseteq Y^* \to X^*$ is densely defined as well.

Next we show that A^* is always closed.

Proposition 3.6.42. *If X, Y are Banach spaces and $A: D(A) \subseteq X \to Y$ is a densely defined unbounded linear operator, then A^* is closed.*

Proof. Suppose that $y_n^* \to y^*$ in Y^* with $y_n^* \in D(A^*)$ for all $n \in \mathbb{N}$ and $A^*(y_n^*) \to x^*$ in X^*. Thanks to (3.6.7) we have

$$\langle y_n^*, A(x)\rangle = \langle A^*(y_n^*), x\rangle \quad \text{for all } x \in D(A) \text{ and for all } n \in \mathbb{N} \, ,$$

which implies

$$\langle y^*, A(x)\rangle = \langle x^*, x\rangle \quad \text{for all } x \in D(A) \, .$$

This gives

$$|\langle y^*, A(x)\rangle| \le \|x^*\|_* \|x\|_X \quad \text{for all } x \in D(A) \, ,$$

which yields, because of (3.6.5), that $y^* \in D(A^*)$, which in combination with (3.6.7) results in

$$\langle A^*(y^*), x\rangle = \langle x^*, x\rangle \quad \text{for all } x \in D(A) \, .$$

This implies $x^* = A^*(y^*)$. Hence, A^* is closed; see Remark 3.6.39. □

Let $i_0: Y^* \times X^* \to X^* \times Y^*$ be the isomorphism defined by $i_0(y^*, x^*) = (-x^*, y^*)$ for all $y^* \in Y^*$ and for all $x^* \in X^*$.

Proposition 3.6.43. *If X, Y are Banach spaces and $A: D(A) \subseteq X \to Y$ is a densely defined unbounded linear operator, then $i_0(\operatorname{Gr} A^*) = (\operatorname{Gr} A)^\perp$.*

Proof. Let $(y^*, x^*) \in Y^* \times X^*$. Then, thanks to (3.6.7), one has

$$\begin{aligned}
(y^*, x^*) \in \operatorname{Gr} A^* \quad &\text{if and only if} \quad \langle y^*, A(x)\rangle = \langle x^*, x\rangle && \text{for all } x \in D(A)\\
&\text{if and only if} \quad \langle y^*, A(x)\rangle - \langle x^*, x\rangle = 0 && \text{for all } x \in D(A)\\
&\text{if and only if} \quad (-x^*, y^*) \in (\operatorname{Gr} A)^\perp \, .
\end{aligned}$$

□

The next result is an extension of Proposition 3.6.24 to unbounded linear operators. Its proof can be found in Brézis [48, Theorem 2.19, p. 46].

Proposition 3.6.44. *If X, Y are Banach spaces and $A: D(A) \subseteq X \to Y$ is a closed, densely defined, unbounded linear operator, then the following statements are equivalent:*
(a) *$R(A) \subseteq Y$ is closed.*
(b) *$R(A^*) \subseteq X^*$ is closed.*
(c) *$R(A) = {}^{\perp}N(A^*)$.*
(d) *$R(A^*) = N(A)^{\perp}$.*

The next two theorems provide useful characterizations of surjective operators.

Theorem 3.6.45. *If X, Y are Banach spaces and $A: D(A) \subseteq X \to Y$ is a closed, densely defined, unbounded linear operator, then the following statements are equivalent:*
(a) *A is surjective, that is, $R(A) = Y$.*
(b) *$\|y^*\|_* \leq c\|A^*(y^*)\|$ for all $y^* \in D(A^*)$ and for some $c > 0$.*
(c) *$R(A^*) \subseteq X^*$ is closed and $N(A^*) = \{0\}$.*

Proof. (a) $\Longrightarrow$ (b): It suffices to show that

$$D^* = \{y^* \in D(A^*) : \|A^*(y^*)\|_* \leq 1\}$$

is bounded. Then according to Proposition 3.2.5 we need to show that for all $y \in Y$, $\langle D^*, y\rangle \subseteq \mathbb{R}$ is bounded. Exploiting the surjectivity of A, there exists $x \in D(A)$ such that $y = A(x)$. Then

$$\langle y^*, y\rangle = \langle y^*, A(x)\rangle = \langle A^*(y^*), x\rangle ,$$

which implies $|\langle y^*, y\rangle| \leq \|x\|$ for every $y^* \in D^*$. Thus, D^* is bounded.

(b) $\Longrightarrow$ (c): Let $x_n^* \in R(A^*)$ for all $n \in \mathbb{N}$ and assume that $x_n^* \to x^*$ in X^*. We can find $y_n^* \in D(A^*)$ such that $x_n^* = A^*(y_n^*)$ for all $n \in \mathbb{N}$. From (b) we see that

$$\left\|y_m^* - y_n^*\right\|_* \leq c\left\|A^*(y_m^* - y_n^*)\right\|_* = c\left\|A^*(y_m^*) - A^*(y_n^*)\right\|_* .$$

This shows that $\{y_n^*\}_{n\geq 1} \subseteq Y^*$ is a Cauchy sequence and so, we conclude that $y_n^* \to y^*$ in Y^*. But from Proposition 3.6.42, we know that A^* is closed. Hence, $x^* = A^*(y^*)$, and so $R(A^*) \subseteq X^*$ is closed. From (b) it is clear that $N(A^*) = \{0\}$.

(c) $\Longrightarrow$ (a): From Proposition 3.6.44 one has $R(A) = {}^{\perp}N(A^*) = Y$. ☐

In a similar way we can prove a dual version of this theorem.

Theorem 3.6.46. *If X, Y are Banach spaces and $A: D(A) \subseteq X \to Y$ is a closed, densely defined, unbounded linear operator, then the following statements are equivalent:*
(a) *A^* is surjective, that is $R(A^*) = X^*$.*
(b) *$\|x\|_X \leq c\|A(x)\|_Y$ for all $x \in D(A)$ and for some $c > 0$.*
(c) *$R(A) \subseteq Y$ is closed and $N(A) = \{0\}$.*

Definition 3.6.47. Let X, Y be Banach spaces and let $A: D(A) \subseteq X \to Y$ be an unbounded linear operator. We say that A is **closable** if there is a closed unbounded linear operator $\hat{A}: D(\hat{A}) \subseteq X \to Y$ such that

$$D(A) \subseteq D(\hat{A}) \quad \text{and} \quad \hat{A}\big|_{D(A)} = A\,.$$

Every closable operator A has a smallest closed extension called the **closure** of A denoted by $\overline{A}$.

The next proposition characterizes closable operators.

Proposition 3.6.48. *If X, Y are Banach spaces and $A: D(A) \subseteq X \to Y$ is an unbounded linear operator, then the following statements are equivalent:*
(a) *A is closable.*
(b) *If $\{x_n\}_{n\geq 1} \subseteq D(A)$ are such that $x_n \to 0$ in X and $A(x_n) \to y$ in Y, then $y = 0$.*
(c) *The projection map $p_X: \overline{\mathrm{Gr}\,A} \to X$ is injective.*

Proof. (a) $\Longrightarrow$ (b): For every closed extension $\hat{A}$ of A one has $y = \hat{A}(0) = 0$.

(b) $\Longrightarrow$ (c): $\overline{\mathrm{Gr}\,A} \subseteq X \times Y$ is a vector subspace and so $p_X: \overline{\mathrm{Gr}\,A} \to X$ is linear. By hypothesis, $N(p_X) = \{0\}$ and so p_X is injective.

(c) $\Longrightarrow$ (a): Let $D(\hat{A}) = p_X(\overline{\mathrm{Gr}\,A}) \subseteq X$. This is a vector subspace. Let $p_Y: \overline{\mathrm{Gr}\,A} \to Y$ be the projection on the second factor. Then $\hat{A} = p_Y \circ p_X^{-1}: D(\hat{A}) \to Y$ is an unbounded linear operator with $\mathrm{Gr}\,\hat{A} = \overline{\mathrm{Gr}\,A}$ and so $\hat{A}$ is a closed extension of A. □

Proposition 3.6.49. *If X, Y are Banach spaces and $A: D(A) \subseteq X \to Y$ is a closable unbounded linear operator, then* $\mathrm{Gr}\,\overline{A} = \overline{\mathrm{Gr}\,A}$.

Proof. Let $\hat{A}$ be a closed extension of A. Then $\overline{\mathrm{Gr}\,A} \subseteq \mathrm{Gr}\,\hat{A}$ and so if $(0, y) \in \overline{\mathrm{Gr}\,A}$, then $y = 0$. Let $A_0: D(A_0) \to Y$ be defined by $D(A_0) = \{x \in X: (x, y) \in \overline{\mathrm{Gr}\,A}$ for some $y \in Y\}$ and $A_0(x) = y$ with $y \in Y$ being the unique element such that $(x, y) \in \overline{\mathrm{Gr}\,A}$. One has $\mathrm{Gr}\,A_0 = \overline{\mathrm{Gr}\,A}$ and so A_0 is a closed extension of A. But $A_0 \subseteq \hat{A}$, which is an arbitrary closed extension of A. Therefore, $A_0 = \overline{A}$. □

Remark 3.6.50. Note that the domain $D(A)$ of an unbounded linear operator $A: D(A) \subseteq X \to Y$ is a normed space with the graph norm defined by $|x| = \|x\|_X + \|A(x)\|_Y$ for all $x \in D(A)$; see the proof of Theorem 3.2.14. Therefore an unbounded linear operator can be viewed also as a bounded linear operator from its domain equipped with the graph norm. It is easy to see that $A: D(A) \subseteq X \to Y$ is closed if and only if $D(A) \subseteq X$ is a Banach space when furnished with the graph norm.

Example 3.6.51. (a) Let $X = C[0, 1]$ be equipped with the supremum norm. This is a Banach space. Let $A: D(A) \subseteq X \to X$ be the unbounded linear operator defined by $A(u) = u'$ for all $u \in D(A) = C^1[0, 1]$. Evidently A is closed and densely defined. Moreover, the graph norm on $D(A)$ is the usual C^1-norm.

(b) Let H be a separable Hilbert space. From Theorem 3.5.48, we know that H has a countable orthonormal basis $\{e_n\}_{n\geq 1}$. Let $\hat{\lambda} = (\lambda_i)_{i\geq 1} \in \mathbb{R}^{\mathbb{N}}$ and consider the linear operator $A_{\hat{\lambda}}\colon D(A_{\hat{\lambda}}) \subseteq H \to H$ defined by

$$A_{\hat{\lambda}}(x) = \sum_{n\geq 1} \lambda_n (x, e_n) e_n \quad \text{for all } x \in D(A_{\hat{\lambda}}) ,$$

where

$$D(A_{\hat{\lambda}}) = \left\{ x \in H\colon \sum_{n\geq 1} |\lambda_n (x, e_n)|^2 < \infty \right\} .$$

This is a closed, densely defined unbounded linear operator. Note that $A_{\hat{\lambda}} \in L(H)$ if and only if $\hat{\lambda} = (\lambda_n)_{n\geq 1}$ is bounded.

We extend the notion of self-adjoint operator to unbounded linear operators.

Definition 3.6.52. Let H be a Hilbert space and $A\colon D(A) \subseteq H \to H$ is a densely defined unbounded linear operator. Then the **adjoint** of A is the unbounded linear operator $A^*\colon D(A^*) \subseteq H \to H$ defined by

$$D(A^*) = \{u \in H\colon |(u, A(x)| \leq c\|x\| \text{ for all } x \in D(A) \text{ and for some } c > 0\}$$

and

$$(A^*(u), x) = (u, A(x)) \quad \text{ for all } x \in D(A) \text{ and for all } u \in D(A^*) .$$

We say that A is **symmetric**, if $A \subseteq A^*$, that is, $D(A) \subseteq D(A^*)$ and $A^*\big|_{D(A)} = A$, so A^* is an extension of A. We say that A is **self-adjoint** if $A = A^*$.

Remark 3.6.53. Evidently A is symmetric if and only if $(A(u), x) = (u, A(x))$ for all $x, u \in D(A)$. A symmetric operator is always closable (see Proposition 3.6.42). Recall that $D(A^*) \supseteq D(A)$ is dense in H. If A is symmetric, then A^* is a closed extension of A. So, we consider the smallest closed extension A^{**} of A. We have $A^{**} \subseteq A^*$. Therefore for symmetric operators we obtain $A \subseteq A^{**} \subseteq A^*$. If A is closed and symmetric, then $A = A^{**} \subseteq A^*$. Finally, if A is self-adjoint, then $A = A^{**} = A^*$. Therefore, a closed symmetric operator A is self-adjoint if and only if A^* is symmetric.

3.7 Compact Operators – Fredholm Operators

In this section we study a class of operators that closely resemble the operators on finite dimensional spaces. These operators are similar to $N \times N$ matrices and so are small in the sense that they map the closed unit ball to a small set.

Definition 3.7.1. Let X, Y be Banach spaces and let $D \subseteq X$ be nonempty subset. A map $f\colon D \to Y$, not necessarily linear, is said to be **compact** if it is continuous and for every bounded set $B \subseteq D$, the set $\overline{f(B)} \subseteq Y$ is compact. By $K(D, Y)$ we denote the family of all compact maps. If $D = X$, then we define $L_c(X, Y) = K(X, Y) \cap L(X, Y)$.

Remark 3.7.2. If Y is finite dimensional, then every continuous bounded map $f\colon D \to Y$ is compact. If $A \in L_c(X, Y)$, then $R(A)$ is separable.

Another notion closely related to compactness is the following one.

Definition 3.7.3. Let X, Y be Banach spaces and $D \subseteq X$ is nonempty. A map $f\colon D \to Y$ is said to be **completely continuous** if for every sequence $\{x_n\}_{n\geq 1} \subseteq D$ such that $x_n \xrightarrow{w} x$ with $x \in D$, it follows $f(x_n) \to f(x)$ in Y.

Remark 3.7.4. Completely continuous operators $A \in L(X, Y)$ are also known as **Dunford–Pettis Operators**. It is easy to see that a linear operator $A\colon X \to Y$ is completely continuous if and only if $A(C) \subseteq Y$ is compact for every weakly compact $C \subseteq X$.

In general the classes of compact maps and of completely continuous maps are distinct. However, for linear operators we can relate the two classes.

Proposition 3.7.5. *If X, Y are Banach spaces and $A \in L_c(X, Y)$, then A is completely continuous.*

Proof. Let $x_n \xrightarrow{w} x$ in X. Then $\{x_n\}_{n\geq 1} \subseteq X$ is bounded and so $\overline{\{A(x_n)\}}_{n\geq 1} \subseteq Y$ is compact. Thus there exists a subsequence $\{x_{n_k}\}_{k\geq 1}$ of $\{x_n\}_{n\geq 1}$ such that $A(x_{n_k}) \to y$ in Y. From Proposition 3.3.23, one has $A(x_n) \xrightarrow{w} A(x)$ in Y. Therefore $y = A(x)$, and so we conclude that $A(x_n) \to A(x)$ in Y. This proves that A is completely continuous. □

Example 3.7.6. The converse is not true in general. Recall that in l^1, weak and norm convergent sequences coincide; see Remark 3.3.17. Then the identity map $i\colon l^1 \to l^1$ is a completely continuous linear operator, but clearly it is not compact.

However, if we strengthen the structure of X, then the converse of Proposition 3.7.5 holds. In fact we obtain the following result.

Proposition 3.7.7. *If X is a reflexive Banach space, Y is a Banach space, $D \subseteq X$ is nonempty,* w*-closed, and $f\colon D \to Y$ is completely continuous, then $f \in K(D, Y)$.*

Proof. Evidently, f is continuous. Let $B \subseteq D$ be a bounded set. We need to show that $\overline{f(B)} \subseteq Y$ is compact. So, let $\{y_n\}_{n\geq 1} \subseteq f(B) \subseteq Y$. Then $y_n = f(x_n)$ with $\{x_n\}_{n\geq 1} \subseteq B$. The reflexivity of X implies that B is relatively weakly compact. So, the Eberlein–Smulian Theorem, Theorem 3.4.14, says that there exists a subsequence $\{x_{n_k}\}_{k\geq 1}$ of $\{x_n\}_{n\geq 1}$ such that $x_{n_k} \xrightarrow{w} x \in D$. We get $y_{n_k} = f(x_{n_k}) \to f(x) \in \overline{f(B)}$, which means that $\overline{f(B)} \subseteq Y$ is compact. □

Corollary 3.7.8. *If X is a reflexive Banach space, Y is a Banach space, and $A \in L(X, Y)$, then $A \in L_c(X, Y)$ if and only if A is completely continuous.*

The next theorem explains why compact maps resemble maps between finite dimensional spaces. First a simple lemma about relatively compact sets in a Banach space Y.

Lemma 3.7.9. *If Y is a Banach space, $K \subseteq Y$ is nonempty and for every $\varepsilon > 0$, there exists a relatively compact set $K_\varepsilon \subseteq Y$ such that for every $y \in K$ we can find $y_\varepsilon \in K_\varepsilon$ such that $\|y - y_\varepsilon\|_Y < \varepsilon$, then $K \subseteq Y$ is relatively compact.*

Proof. Let $\varepsilon > 0$ be given. There exists a relatively compact set $K_{\varepsilon/2} \subseteq Y$ as postulated by the hypothesis of the lemma. The total boundedness of $K_{\varepsilon/2}$ implies that there exist $\{y_\varepsilon^k\}_{k=1}^m \subseteq K_{\varepsilon/2}$ such that

$$K_{\frac{\varepsilon}{2}} \subseteq \bigcup_{k=1}^m B_{\frac{\varepsilon}{2}}\left(y_\varepsilon^k\right) .$$

By hypothesis, given $y \in K$, there exists $y_{\varepsilon/2} \in K_{\varepsilon/2}$ such that $\|y - y_{\varepsilon/2}\|_Y < \varepsilon/2$. Since $y_{\varepsilon/2} \in B_{\varepsilon/2}(y_\varepsilon^{k_0})$ for some $k_0 \in \{1, \ldots, m\}$ one has $\|y_{\varepsilon/2} - y_\varepsilon^{k_0}\|_Y < \varepsilon/2$. Therefore $\|y - y_\varepsilon^{k_0}\|_Y < \varepsilon$, which implies $K \subseteq \bigcup_{k=1}^m B_\varepsilon(y_\varepsilon^k)$. Hence, K is totally bounded and so relatively compact. □

Theorem 3.7.10. *If X, Y are Banach spaces, $D \subseteq X$ is nonempty, bounded, and $f\colon D \to Y$, then the following two statements are equivalent:*
(a) $f \in K(D, Y)$.
(b) *For every $\varepsilon > 0$ there exists a continuous, bounded map $f_\varepsilon\colon D \to Y$ such that $\|f(x) - f_\varepsilon(x)\|_Y < \varepsilon$ for all $x \in D$ and $f_\varepsilon(D) \subseteq \overline{\operatorname{conv}} f(D)$ as well as* $\dim(\operatorname{span} f_\varepsilon(D)) < \infty$.

Proof. (a) $\Longrightarrow$ (b): Since f is compact, $f(D) \subseteq Y$ is relatively compact. So, for every $\varepsilon > 0$ there exists a sequence $\{y_k\}_{k=1}^m \subseteq f(D)$ such that

$$\min_{k\in\{1,\ldots,m\}} \|f(x) - y_k\|_Y < \varepsilon \quad \text{for all } x \in D . \tag{3.7.1}$$

Recall that $f(D)$ is totally bounded. Let $\lambda_k(x) = \max\{\varepsilon - \|f(x) - y_k\|_Y, 0\}$. Clearly $\lambda_k\colon D \to \mathbb{R}_+$ with $k = 1, \ldots, m$ are continuous functions and do not all vanish simultaneously for $x \in D$, see (3.7.1). We introduce the map $f_\varepsilon\colon D \to Y$ defined by

$$f_\varepsilon(x) = \frac{\sum_{k=1}^m \lambda_k(x) y_k}{\sum_{k=1}^m \lambda_k(x)} . \tag{3.7.2}$$

Evidently f_ε is continuous, bounded, and

$$\|f_\varepsilon(x) - f(x)\|_Y = \left\| \frac{\sum_{k=1}^m \lambda_k(x)(y_k - f(x))}{\sum_{k=1}^m \lambda_k(x)} \right\|_Y < \frac{\sum_{k=1}^m \lambda_k(x)\varepsilon}{\sum_{k=1}^m \lambda_k(x)} = \varepsilon$$

for all $x \in D$; see (3.7.1). Then the boundedness of $f(D)$ implies the boundedness of $f_\varepsilon(D)$ while from (3.7.2) we see that $\dim(\operatorname{span} f_\varepsilon(D)) < \infty$. Therefore, f_ε is compact, and it is clear from (3.7.2) that $f_\varepsilon(D) \subseteq \overline{\operatorname{conv}} f(D)$.

(b) $\Longrightarrow$ (a): Let $\varepsilon = 1/n$ with $n \in \mathbb{N}$. Then there exist continuous, bounded maps $f_{1/n}\colon D \to Y$ such that

$$\left\|f(x) - f_{\frac{1}{n}}(x)\right\|_Y < \frac{1}{n} \quad \text{for all } x \in D ,$$

which shows that f is continuous since it is the uniform limit of continuous maps.

Let $y = f(x)$ with $x \in D$. Then it follows that

$$\|y - y_n\| < \frac{1}{n} \quad \text{with} \quad y_n = f_{\frac{1}{n}}(x) \in f_{\frac{1}{n}}(D)\,.$$

The set $f_{1/n}(D)$ is relatively compact. So, invoking Lemma 3.7.9 we conclude that $f(D) \subseteq Y$ is relatively compact, that is, $f \in K(D, Y)$. □

Definition 3.7.11. Let X, Y be Banach spaces and $A \in L(X, Y)$. We say that A is a **finite rank operator** (or **finite dimensional operator** or **degenerate operator**) if $\dim R(A) < \infty$. We denote the space of all finite rank operators by $L_f(X, Y)$. Clearly $L_f(X, Y) \subseteq L_c(X, Y)$. This inclusion is strict in general.

According to Theorem 3.7.10, every $A \in L_c(X, Y)$ can be approximated uniformly on bounded sets by compact maps with finite dimensional range. However, we cannot say that these approximating maps are in $L_f(X, Y)$. So, we are led to the following definition.

Definition 3.7.12. We say that the Banach space Y has the **approximation property** if for every Banach space X, $\overline{L_f(X, Y)}^{\|\cdot\|_L} = L_c(X, Y)$.

Remark 3.7.13. The first example of a Banach space without the approximation property was produced by Enflo [104] who considered a separable reflexive space. A Banach space with a Schauder basis has the approximation property. So, Enflo's example also showed that not every separable reflexive Banach space has a Schauder basis, another long-standing open problem; see Remark 3.5.51.

Proposition 3.7.14. *If X, Y are Banach spaces, then $L_c(X, Y)$ is a Banach space.*

Proof. We only need to show that $L_c(X, Y)$ is a closed subspace of $L(X, Y)$. So, let $A_n \to A$ in $L(X, Y)$. Then $\sup[\|A_n(x) - A(x)\|_Y : \|x\|_X \leq 1] \to 0$ as $n \to \infty$. Given $\varepsilon > 0$, there exists $n_0 \in \mathbb{N}$ such that $\|A_n(x) - A(x)\|_Y < \varepsilon/2$ for all $x \in \overline{B}_1^X$ and for all $n \geq n_0$. The set $A_{n_0}(\overline{B}_1^X)$ is totally bounded; recall that $A_{n_0} \in L_c(X, Y)$. Hence, there exists a finite $\varepsilon/2$-net $F \subseteq A_{n_0}(\overline{B}_1^X)$; see Definition 1.5.31. Given $x \in \overline{B}_1^X$ there exists $y \in F$ such that $\|A_{n_0}(x) - y\|_Y < \varepsilon/2$. Then

$$\|A(x) - y\|_Y \leq \|A(x) - A_{n_0}(x)\|_Y + \|A_{n_0}(x) - y\|_Y < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon\,.$$

Hence, F is an ε-net for $A(\overline{B}_1^X)$, thus, $A(\overline{B}_1^X)$ is relatively compact. Therefore, $A \in L_c(X, Y)$. □

Proposition 3.7.15. *If X, Y, V are Banach spaces, $A \in L(X, Y)$, $T \in L(Y, V)$, and A or T is compact, then $T \circ A \in L_c(X, Y)$.*

Proof. First suppose that A is compact. Then $\overline{A(\overline{B}_1^X)} \subseteq Y$ is compact, hence $T(\overline{A(\overline{B}_1^X)}) \subseteq V$ is compact. This means that $T \circ A \in L_c(X, V)$.

Now suppose that T is compact. The set $A(\overline{B}_1^X) \subseteq Y$ is bounded. Since $T \in L_c(Y, V)$ we have that $T(A(\overline{B}_1^X)) \subseteq V$ is relatively compact. This means that $T \circ A \in L_c(X, V)$. □

Corollary 3.7.16. *If X is a Banach space, then $L_c(X)$ is a closed ideal of $L(X)$.*

The next characterization of operator compactness is very useful in many occasions and is known as "Schauder's Theorem."

Theorem 3.7.17 (Schauder's Theorem). *If X, Y are Banach spaces and $A \in L(X, Y)$, then $A \in L_c(X, Y)$ if and only if $A^* \in L_c(Y^*, X^*)$.*

Proof. $\Longrightarrow$: Let $K = \overline{A(\overline{B}_1^X)}$. Then $K \subseteq Y$ is compact. Moreover, let $B \subseteq Y^*$ be bounded. Then

$$|\langle y^*, y_1 - y_2 \rangle| \leq c\|y_1 - y_2\| \quad \text{for all } y^* \in B_1\,, \text{ for all } y_1, y_2 \in K\,, \text{ for some } c > 0\,.$$

This shows that $B \subseteq C(K)$ is bounded and equicontinuous. So, invoking the Arzela–Ascoli Theorem (see Theorem 1.6.16), we infer that B is relatively compact. Then, if $\{y_n^*\}_{n\geq 1} \subseteq B$, there exists a subsequence $\{y_{n_k}^*\}_{k\geq 1}$ of $\{y_n^*\}_{n\geq 1}$, which is a uniformly Cauchy sequence on K. This implies that $\{y_{n_k}^* A\}_{k\geq 1}$ is a uniformly Cauchy sequence on $\overline{B}_1^X$. Therefore, $\{y_{n_k}^* A\}_{k\geq 1} \subseteq X^*$ is convergent. But by Definition 3.6.6, $y_{n_k}^* A = A^*(y_{n_k}^*)$. Thus, we conclude that $A^* \in L_c(Y^*, X^*)$.

$\Longleftarrow$: From the previous implication we obtain that $A^{**} \in L_c(X^{**}, Y^{**})$. Let $j_X : X \to X^{**}$ and $j_Y : Y \to Y^{**}$ be the corresponding canonical embeddings. Then $A = j_Y^{-1} \circ A^{**} \circ j_X$ and so Proposition 3.7.15 implies that $A \in L_c(X, Y)$. □

Definition 3.7.18. (a) If X is a vector space and V is a vector subspace of X, then the **codimension** of V in X is the dimension of the quotient vector space X/V.
(b) Let X, Y be Banach spaces and $A \in L(X, Y)$. We say that A is a **Fredholm operator** if $N(A)$ is finite dimensional and $R(A)$ has finite codimension. The number $i(A) = \dim N(A) - \operatorname{codim} R(A) = \dim N(A) - \dim(Y/R(A))$ is called the **index of** A.

Remark 3.7.19. If $A \in L(X, Y)$ is a Fredholm operator, then $X = N(A) \oplus V$ and $A\big|_V$ is an isomorphism of V onto $R(A)$. Moreover, $R(A) \subseteq Y$ is closed.

Lemma 3.7.20. *If X is a Banach space, $A \in L(X)$, $T = i_X - A$, and $V = R(T)$ is a proper closed subspace of X, then for every $\varepsilon > 0$ there exists $x_0 \in \overline{B}_1^X$ such that $d(A(x_0), A(V)) \geq 1 - \varepsilon$.*

Proof. According to the Riesz Lemma (see Lemma 3.1.20), there exists $x_0 \in X$ with $\|x_0\| = 1$ such that $d(x_0, V) \geq 1 - \varepsilon$. One has $T(x_0) \in V$ and $A(V) = (i_X - T)(V) \subseteq Y$. Therefore,

$$d(A(x_0), A(V)) \geq d(A(x_0) + T(x_0), V) = d(x_0, V) \geq 1 - \varepsilon\,.$$ □

Using this lemma we can prove the following theorem, which gives an important class of Fredholm operators.

Theorem 3.7.21. *If X is a Banach space, $A \in L_c(X)$, and $\lambda \neq 0$, then $\lambda i_X - A$ is a Fredholm operator.*

Proof. Clearly, we may assume that $\lambda = 1$. Let $N = N(i_X - A)$. For every $x \in N$ one has $A(x) = x$. Therefore $A\big|_N$ is an isomorphism with a subspace of X and $A\big|_N$ is compact as well. It follows that N is finite dimensional. Proposition 3.2.28 implies that there is a closed subspace V of X such that $X = N \oplus V$. Let $T = i_X - A$ and $\hat{T} = T\big|_V$. We obtain that $R(T) = T(V) = R(\hat{T})$ and $N(\hat{T}) = N \cap V = \{0\}$, hence $\hat{T}$ is injective. We claim that

$$\inf\left[\left\|\hat{T}(x)\right\| : x \in V, \|x\| = 1\right] > 0 \,. \tag{3.7.3}$$

Arguing by contradiction, suppose that (3.7.3) does not hold. Then there exists $x_n \in V$ with $\|x_n\| = 1$ for all $n \in \mathbb{N}$ such that $\left\|\hat{T}(x_n)\right\| \to 0$. Since $A \in L_c(X)$ we may assume that $A(x_n) \to u$ in X. Note that $A(x_n) = x_n$ for all $n \geq 1$, so $\|u\| = 1$. Moreover, $\hat{T}(u) = 0$ and this contradicts the injectivity of $\hat{T}$.

From (3.7.3) we infer that $\|\hat{T}(x)\| \geq c\|x\|$ for all $x \in V$ and for some $c > 0$. Then, Theorem 3.6.45 implies that $R(\hat{T}) = R(T) \subseteq X$ is closed.

We will show that $\operatorname{codim} R(T) < \infty$. Inductively we define

$$T^0 = i_X\,, \quad T^1 = T \quad \text{and} \quad T^{k+1} = TT^k \quad \text{for all } k \in \mathbb{N}_0\,.$$

Moreover we set $N_k = N(T^k)$. Since $T^k = (i_X - T)^k$ and powers of compact operators are again compact operators (see Proposition 3.7.15), we get $T^k = i_X - S_k$ with $S_k \in L_c(X)$. From the first part of the proof we see that $\dim N_k < \infty$ for all $k \in \mathbb{N}_0$.

Let $Z_k = R(T^k) = T^k(V_1)$ with $k \in \mathbb{N}_0$. We have that

$$\{N_k\}_{k\in\mathbb{N}_0} \text{ is increasing and } \{Z_k\}_{k\in\mathbb{N}_0} \text{ is decreasing}\,. \tag{3.7.4}$$

For some $n \in \mathbb{N}_0$ we obtain $Z_n = Z_{n+1}$. Indeed if all the inclusions $Z_k \supseteq Z_{k+1}$ are strict, then with Lemma 3.7.20 there exists $u_n \in \overline{B}_1^{Z_n}$ such that $d(A(u_n), A(Z_{n+1}) \geq 1/2$. Then $\|A(u_n) - A(u_m)\| \geq 1/2$ for $n \neq m$, a contradiction to the fact that $A \in L_c(X)$.

Similarly, for some $m \in \mathbb{N}_0$, it holds $N_m = N_{m+1}$. Indeed if $x \in N_k$, that is, $T^k(x) = 0$, then $T^{k-1}(T(x)) = 0$ and so $T(x) \subseteq N_{k-1} \subseteq N_k$; see (3.7.4). Therefore, again via Lemma 3.7.20, we conclude that $N_m = N_{m+1}$ for some $m \in \mathbb{N}_0$. Thus, we obtain

$$Z_n = Z_{n'} \quad \text{for all } n' \geq n \quad \text{and} \quad N_m = N_{m'} \quad \text{for all } m' \geq m\,.$$

Let $i = \max\{n, m\}$. We claim that $X = N_i \oplus Z_i$. Let $x \in X$. Then $T^i(x) \in Z_i$ and $T^i(Z_i) = T^i(T^i(X)) = T^{2i}(X) = T^i(X) = Z_i$. Therefore there exists $u \in Z_i$ such that $T^i(u) = T^i(x)$, hence $T^i(u - x) = 0$. Therefore, $u - x \in N_i$ and $x = x - u + u$. Since $X = N_i \oplus Z_i$, the codimension of Z_i and also of $Z_1 \supseteq Z_i$ is finite. □

Example 3.7.22. (a) If X, Y are finite dimensional Banach spaces, then every linear operator $A\colon X \to Y$ is a Fredholm operator and $i(A) = \dim X - \dim Y$.

(b) If X, Y are Banach spaces and $A \in L(X, Y)$ is a bijection, then A is a Fredholm operator and $i(A) = 0$.

(c) Let $X = l^p$ with $1 \le p \le \infty$ and let $A \in L(l^p)$ be defined by

$$A(\hat{x}) = (x_{n+k})_{n\ge1} \quad \text{for all } \hat{x} = (x_n)_{n\ge1} \in l^p \text{ and for some } k \in \mathbb{N}\,.$$

Recall that for every $n \in \mathbb{N}$, $e_n = (0, \dots, 0, 1, 0 \dots)$ where 1 is located at the $n \overset{\text{th}}{=}$entry. We see that $N(A) = \operatorname{span}\{e_n\}_{n=1}^k$, $R(A) = \{e_n\}_{n\ge k+1}$, and $R(A) = l^p$. Therefore A is a Fredholm operator and $i(A) = k$.

Let us consider the case where $X = Y$ are Banach spaces and $A \in L_c(X)$. Then according to Theorem 3.7.21, $i_X - A$ is a Fredholm operator. The next theorem, known as the "Fredholm Alternative Theorem," asserts that either the nonhomogenous linear equation $x - A(x) = u$ has a solution $x \in X$ for every $u \in X$ or the corresponding homogeneous equation $x - A(x) = 0$ has a nontrivial solution. The result has interesting applications in boundary values problems.

Theorem 3.7.23 (Fredholm Alternative Theorem). *If X is a Banach space, $A \in L_c(X, Y)$ and $\lambda \ne 0$, then the equation $\lambda x - A(x) = u$ has a solution for every $u \in X$ if and only if the equation $x - A(x) = 0$ only has the trivial solution.*

Proof. Again we may assume that $\lambda = 1$. Let $T = i_X - A$. If $A(x) - x = 0$ only has the trivial solution, then $N = N(T) = \{0\}$ and so T is an isomorphism into. We will show that it is surjective.

Let $V_k = R(T^k)$ for all $k \in \mathbb{N}_0$. From the proof of Theorem 3.7.21 we know that there exists $n \in \mathbb{N}_0$ such that $V_k = V_n$ for all $k \ge n$. We claim that $V_1 = V_0 = X$. If this is not the case, let $m \in \mathbb{N}$ be the smallest integer such that $V_{m-1} \ne V_m = V_{m+1}$. We pick $u \in V_{m-1} \setminus V_m$. Then $T(u) \in V_m = V_{m+1}$. Hence, there exists $v \in V_m$ such that $T(u) = T(v)$ and $u \ne v$ since $u \notin V_m$. But this contradicts the injectivity of T.

Next, assume that T is surjective. Let $N_k = N(T^k)$ for $k \in \mathbb{N}$. We need to show that $N_1 = N(T) = \{0\}$. Recall that $\{N_k\}_{k\ge1}$ is increasing. Arguing by contradiction, suppose that there is $x_1 \ne 0$ such that $x_1 \in N_1$. Inductively we will generate a sequence $\{x_k\}_{k\ge1} \subseteq X$ such that $T(x_{k+1}) = x_k$ and $x_k \in N_k \setminus N_{k-1}$ for all $k \in \mathbb{N}$. Suppose that $x_1, \dots, x_k$ have been constructed. Since $R(T) = X$, there exists $x_{k+1} \in X$ such that $T(x_{k+1}) = x_k$. Then $T^k(x_{k+1}) = T^{k-1}(x_k) = \cdots = x_1 \ne 0$ and $T^k(x_{k+1}) = T(x_1) = 0$. This completes the induction. Since $N_m = N_{m+1}$ for some $m \in \mathbb{N}_0$ (see the proof of Theorem 3.7.21), we have proven the assertion of the theorem. ☐

Next we prove a duality property of Fredholm operators, that is, we show that $A \in L(X, Y)$ is Fredholm if and only if $A^* \in L(Y^*, X^*)$ is Fredholm. We start with a simple lemma.

Lemma 3.7.24. *If X, Y are Banach spaces, $A \in L(X, Y)$ and $\dim(Y/R(A)) < \infty$, then $R(A) \subseteq Y$ is a closed subspace.*

Proof. Let $m = \dim(Y/R(A)) < \infty$. Then there exist vectors $\{y_k\}_{k=1}^m \subseteq Y$ such that

$$[y_k] = y_k + R(A) \in Y/R(A) \quad \text{for all } k \in \{1, \dots, m\}$$

form a basis of $Y/R(A)$. We introduce the space

$$\hat{X} = X \times \mathbb{R}^m \quad \text{with norm } \|(x, \hat{\lambda})\|_{\hat{X}} = \|x\|_X + |\hat{\lambda}|$$

for all $x \in X$ and for all $\hat{\lambda} = (\lambda_k)_{k=1}^m \in \mathbb{R}^m$. Of course $\hat{X}$ with the norm above is a Banach space. Let $\hat{A} \in L(\hat{X}, Y)$ be defined by

$$\hat{A}(x, \hat{\lambda}) = A(x) + \sum_{k=1}^{m} \lambda_k y_k \,.$$

Then $\hat{A}$ is surjective and

$$N(\hat{A}) = \{(x, \lambda) \in X \times \mathbb{R}^m : A(x) = 0, \hat{\lambda} = 0\} = N(A) \times \{0\} \,.$$

Invoking Theorem 3.8.19, there exists $c > 0$ such that

$$\inf[\|x + u\|_X : u \in N(A)] + |\hat{\lambda}| \le c \left\| A(x) + \sum_{k=1}^{m} \lambda_k y_k \right\|_Y \quad \text{for all } x \in X, \hat{\lambda} \in \mathbb{R}^m \,.$$

Let $\hat{\lambda} = 0$. Then

$$\inf[\|x + u\|_X : u \in N(A)] \le c \,\|A(x)\|_Y \quad \text{for all } x \in X \,,$$

which shows that $R(A) \subseteq Y$ is closed; see Theorem 3.8.19. □

Using this lemma, we can prove the duality property for Fredholm operators.

Theorem 3.7.25. *If X, Y are Banach spaces and $A \in L(X, Y)$, then the following hold:*
(a) *A is a Fredholm operator if and only if A^* is a Fredholm operator.*
(b) *If A is a Fredholm operator, then* $\dim N(A^*) = \dim(Y/R(A))$ *and* $\dim N(A) = \dim(X^*/R(A^*))$.

Proof. According to Theorem 3.8.19, $R(A) \subseteq Y$ is closed if and only if $R(A^*) \subseteq X^*$ is closed. So, we may assume that both $R(A) \subseteq Y$ and $R(A^*) \subseteq X^*$ are closed subspaces. Then

$$R(A^*) = N(A)^{\perp} \quad \text{and} \quad R(A)^{\perp} = N(A^*) \,; \tag{3.7.5}$$

see Proposition 3.6.44. Applying Proposition 3.2.25, one has

$$N(A)^* = X^*/N(A)^{\perp} = X^*/R(A^*) \quad \text{and} \quad (Y/R(A))^* = R(A)^{\perp} = N(A^*) \,;$$

see (3.7.5). This completes the proof of both statements (a) and (b). □

The last part of this section is devoted to the spectral theory of bounded linear operators. First, let us recall some standard results about invertible operators. Recall that $A \in L(X, Y)$ is invertible if and only if it is an isomorphism of X onto Y with X, Y being Banach spaces. Moreover, from Proposition 3.6.7(e) we know that $A \in L(X, Y)$ with X, Y being Banach spaces is invertible if and only if A^* is invertible and $(A^{-1})^* = (A^*)^{-1}$. In addition, if X, Y, V are Banach spaces and $A \in L(X, Y)$, $T \in L(Y, V)$ are invertible operators, then $T \circ A \in L(X, V)$ is invertible as well and $(T \circ A)^{-1} = A^{-1}T^{-1}$.

Lemma 3.7.26. *If X is a Banach space, $A \in L(X)$ and $\|A\|_L < 1$, then $i_X - A \in L(X)$ is invertible and $(i_X - A)^{-1} = \sum_{n\geq 0} A^n$ with the series being absolutely convergent.*

Proof. Note that

$$\sum_{n\geq 0} \|A^n\|_L \leq \sum_{n\geq 0} \|A\|_L^n < \infty$$

since by hypothesis $\|A\|_L < 1$. Hence $\sum_{n\geq 0} A^n$ is absolutely convergent in $L(X)$. Then we obtain

$$(i_X - A) \sum_{n\geq 0} A^n = (i_X - A) + (A - A^2) + \ldots = i_X\,,$$

which is called the **telescoping sum**. Similarly we get $(\sum_{n\geq 0} A^n)(i_X - A) = i_X$. Therefore we conclude that $i_X - A \in L(X)$ is invertible and $(i_X - A)^{-1} = \sum_{n\geq 0} A^n$. □

Lemma 3.7.27. *If X is a Banach space, $A, T \in L(X)$, A is invertible and $\|A - T\|_L < 1/\|A^{-1}\|_L$, then T is invertible as well and $\|T^{-1} - A^{-1}\|_L \leq (\|A^{-1}\|_L^2 \|T - A\|_L)/(1 - \|A^{-1}\|_L \|T - A\|_L)$.*

Proof. Note that

$$\|A^{-1}(A - T)\|_L \leq \|A^{-1}\|_L \|T - A\|_L < 1\,.$$

Using Lemma 3.7.26 it follows that $i_X - A^{-1}(A - T) = A^{-1}T \in L(X)$ is invertible. Hence $T \in L(X)$ is invertible since $T = A(A^{-1}T)$. Moreover, we get

$$(i_X - A^{-1}(A - T))^{-1} = \sum_{n\geq 0} (A^{-1}(A - T))^n\,;$$

see Lemma 3.7.26. Therefore

$$T^{-1} = (A - (A - T))^{-1} = (A(i_X - A^{-1}(A - T)))^{-1} = \sum_{n\geq 0} (A^{-1}(A - T))^n A^{-1}\,.$$

Thus,

$$\begin{aligned} \|T^{-1} - A^{-1}\|_L &\leq \sum_{n\geq 1} \left\|(A^{-1}(A - T))^n A^{-1}\right\|_L \leq \|A^{-1}\|_L \sum_{n\geq 1} \left(\|A^{-1}\|_L \|A - T\|_L\right)^n \\ &= \frac{\|A^{-1}\|_L^2 \|A - T\|_L}{1 - \|A^{-1}\|_L \|T - A\|_L}\,. \end{aligned}$$

□

Corollary 3.7.28. *If X is a Banach space and $\mathcal{L} \subseteq L(X)$ is the set of all invertible operators, then $\mathcal{L}$ is an open set in $L(X)$ and the map $A \to A^{-1}$ is a homeomorphism of $\mathcal{L}$ onto $\mathcal{L}$.*

Now we introduce the spectrum of a bounded linear operator. In order to have a complete spectral theory we need to assume that X is a complex Banach space.

Definition 3.7.29. Let X be a complex Banach space and let $A \in L(X)$. The **spectrum** $\sigma(A)$ of A is the set

$$\sigma(A) = \{\lambda \in \mathbb{C}\colon \lambda i_X - A \text{ is not invertible}\}\,.$$

The **resolvent set** $\rho(A)$ of A is the complement of $\sigma(A)$, that is, $\rho(A) = \mathbb{C} \setminus \sigma(A)$. The elements of $\rho(A)$ are called **regular values of** A. Moreover, if $\lambda \in \rho(A)$, then $R(\lambda) = (\lambda i_X - A)^{-1} \in L(X)$ is called the **resolvent of** A **at** λ. The spectrum of A is decomposed in the following way:

$$P\sigma(A) = \{\lambda \in \mathbb{C} \colon \lambda i_X - A \text{ is not injective}\}\,,$$
$$R\sigma(A) = \{\lambda \in \mathbb{C} \colon \lambda i_X - A \text{ is injective but } R(\lambda i_X - A) \subseteq X \text{ is not dense}\}\,,$$
$$C\sigma(A) = \{\lambda \in \mathbb{C} \colon \lambda i_X - A \text{ is injective, } R(\lambda i_X - A) \subseteq X \text{ is dense but } \lambda i_X - A \text{ is not surjective}\}\,.$$

We call $P\sigma(A)$ the **point spectrum of** A, $R\sigma(A)$ is the **residual spectrum of** A, and $C\sigma(A)$ is the **continuous spectrum of** A. Given $\lambda \in \mathbb{C}$ we see that $\lambda \in P\sigma(A)$ if and only if there exists $x \in X \setminus \{0\}$ such that $A(x) = \lambda x$. The elements of $P\sigma(A)$ are called **eigenvectors** for λ and $N(\lambda i_X - A)$ is the **eigenspace** for λ.

Remark 3.7.30. If X is finite dimensional and $n = \dim X$, then $\sigma(A) = P\sigma(A)$ and card $\sigma(A) \le n$. If X is infinite dimensional and $A \in L_c(X)$, then $0 \in \sigma(A)$ or otherwise A would be a compact isomorphism, a contradiction.

Proposition 3.7.31. *If X is a Banach space and $A \in L(X)$, then $\sigma(A) = \sigma(A^*)$.*

Proof. From Proposition 3.6.7(e), we know that $(\lambda i_X - A)$ is invertible if and only if $(\lambda i_X - A)^*$ is invertible. To conclude the proof just note that $(\lambda i_X - A)^* = \lambda i_{X^*} - A^*$. □

On account of Remark 3.6.8, we can state the following corollary concerning operators defined on a Hilbert into itself.

Corollary 3.7.32. *If H is a complex Hilbert space and $A \in L(H)$, then $\sigma(A^*) = \{\overline{\lambda} \colon \lambda \in \sigma(A)\}$.*

Proposition 3.7.33. *If X is a Banach space and $A \in L(X)$, then $\sigma(A) \subseteq \mathbb{C}$ is compact and if $\lambda \in \sigma(A)$, then $|\lambda| \le \|A\|_L$.*

Proof. Corollary 3.7.28 implies that $\rho(A) \subseteq \mathbb{C}$ is open. Hence, $\sigma(A) = \mathbb{C} \setminus \rho(A)$ is closed. Let $\lambda \in \mathbb{C}$ such that $|\lambda| > \|A\|_L$. Then $\lambda i_X - A = \lambda(i_X - 1/\lambda A)$ and so with Lemma 3.7.26, $\lambda i_X - A$ is invertible. Therefore, if $\lambda \in \sigma(A)$, then $|\lambda| \le \|A\|_L$ and $\sigma(A) \subseteq \mathbb{C}$ is compact. □

The next result is valid only for complex Banach spaces. That is why we said that in order to have a complete theory, we need to consider Banach spaces over $\mathbb{C}$.

Proposition 3.7.34. *If X is a complex Banach space and $A \in L(X)$, then $\sigma(A) \neq \emptyset$.*

Proof. We fix $\lambda_0 \in \sigma(A)$ and consider $\lambda \in \mathbb{C}$ such that $|\lambda - \lambda_0| < \|(\lambda_0 i_X - A)^{-1}\|_L^{-1}$. Using Lemma 3.7.27 for the operators $\lambda_0 i_X - A$ and $\lambda i_X - A$, we get

$$\begin{aligned} R(\lambda) = (\lambda i_X - A)^{-1} &= \sum_{n \ge 0} \left[(\lambda_0 i_X - A)^{-1}(\lambda_0 - \lambda) i_X\right]^n (\lambda_0 i_X - A)^{-1} \\ &= \sum_{n \ge 0} (\lambda_0 - \lambda)^n (\lambda_0 i_X - A)^{-(n+1)} = \sum_{n \ge 0} (\lambda_0 - \lambda)^n R(\lambda_0)^{n+1}\,; \end{aligned}$$

see the proof of Lemma 3.7.27. Note that the series is absolutely convergent. So $\lambda \to R(\lambda)$ is an analytic function from $\rho(A)$ into $L(X)$. From the proof of Proposition 3.7.33 we know that if $|\lambda| > \|A\|_L$, then $R(\lambda) = \sum_{n\geq 0} 1/\lambda^{n+1} A^n$, hence $\|R(\lambda)\|_L \leq 1/(|\lambda| - \|A\|_L)$. Arguing by contradiction, suppose that $\rho(A) = \mathbb{C}$. Then $R(\lambda) \to 0$ as $|\lambda| \to +\infty$. So with Liouville's Theorem, we obtain that $R \equiv 0$, a contradiction since the values of R are invertible operators. Therefore $\rho(A) \neq \mathbb{C}$ and so $\sigma(A) \neq \emptyset$. □

As we already pointed out (see Remark 3.7.30), if $\dim X < \infty$ and $A \in L(X)$, then $\sigma(A) = P\sigma(A)$, just recall that in this case A is injective if and only if A is surjective. However, it is not true in general that every point of $\sigma(A)$ is an eigenvalue. For compact operators every nonzero element of $\sigma(A)$ is an eigenvalue.

Proposition 3.7.35. *If X is a Banach space, $A \in L_c(X)$ and $\lambda \in \sigma(A)\setminus\{0\}$, then $\lambda \in P\sigma(A)$.*

Proof. Suppose that $\lambda \neq 0$ is not an eigenvalue of A. Then according to Definition 3.7.29 we obtain $N(\lambda i_X - A) = \{0\}$. Then with the Fredholm Alternative Theorem (see Theorem 3.7.23), we have $R(\lambda i_X - A) = X$. Hence, according to Theorem 3.2.10, $\lambda i_X - A$ is invertible, which means that $\lambda \notin \sigma(A)$. □

Lemma 3.7.36. *If X is a Banach space, $A \in L(X)$, $\{\lambda_k\}_{k=1}^n$ are distinct eigenvalues of A and e_k is an eigenvector corresponding to λ_k for each $k = 1, \dots, n$ with $n \in \mathbb{N}$, then $\{e_k\}_{k=1}^n \subseteq X$ are linearly independent.*

Proof. The proof goes by induction. So, suppose that $\{e_k\}_{k=1}^{n-1}$ are linearly independent. Let $e_n = \sum_{k=1}^{n-1} \vartheta_k e_k$ with $\vartheta_k \in \mathbb{C}$. Then $\sum_{k=1}^{n-1} \lambda_n \vartheta_k e_k = \lambda_n e_n = A(e_n) = \sum_{k=1}^{n-1} \lambda_k \vartheta_k e_k$. Hence, $\sum_{k=1}^{n-1} (\lambda_n - \lambda_k)\vartheta_k e_k = 0$. Since by the induction hypothesis $\{e_k\}_{k=1}^{n-1} \subseteq X$ are linearly independent and $\lambda_n - \lambda_k \neq 0$, we must have $\vartheta_k = 0$ for all $k = 1, \dots, n$. Therefore $\{e_k\}_{k=1}^n \subseteq X$ are linearly independent. □

Proposition 3.7.37. *If X is a Banach space, $A \in L_c(X)$, and $\varepsilon > 0$, then A has only finitely many eigenvalues $\lambda \in \mathbb{C}$ such that $|\lambda| > \varepsilon$.*

Proof. Arguing by contradiction, suppose that there exist distinct eigenvalues $\{\lambda_k\}_{k\geq 1}$ such that $|\lambda_k| > \varepsilon$ for all $k \in \mathbb{N}$. For every eigenvalue λ_k, we choose an eigenvector e_k. For $n \in \mathbb{N}$ let $X_n = \operatorname{span}\{e_k\}_{k=1}^n$. With Lemma 3.7.36 it follows that $A(X_n) = X_n$ and $X_{n-1} \neq X_n$. Invoking the Riesz Lemma (see Lemma 3.1.20), there is a $u_n \in X_n$ such that

$$d(u_n, u_{n+1}) \geq \frac{1}{2} \quad \text{and} \quad \|u_n\| = 1 \quad \text{for all } n \geq 2\,. \tag{3.7.6}$$

Let $y_n = 1/\lambda_n u_n$ and note that $\|y_n\| \leq 1/\varepsilon$. Then $A(y_n) \in X_n$ and $u_n - A(y_n) \in X_{n-1}$. To see this second inclusion, note that $y_n = \sum_{k=1}^n \vartheta_k e_k$ with $\vartheta_k \in \mathbb{C}$. Then

$$u_n - A(y_n) = \sum_{k=1}^{n} \left(1 - \frac{\lambda_k}{\lambda_n}\right) \vartheta_k e_k = \sum_{k=1}^{n-1} \left(1 - \frac{\lambda_k}{\lambda_n}\right) \vartheta_k e_k \in X_{n-1}\,.$$

Let $n > m$. Then $A(y_m) \in X_m \subseteq X_{n-1}$ and $u_n - A(y_n) \in X_{n-1}$. Therefore one has

$$\begin{aligned} \|A(y_n) - A(y_m)\| &\geq d(A(y_n), X_{n-1}) = d(A(y_n) + u_n - A(y_n), X_{n-1}) \\ &= d(u_n, X_{n-1}) \geq \frac{1}{2} \; ; \end{aligned} \tag{3.7.7}$$

see (3.7.6). But $\{y_n\}_{n\geq 1} \subseteq A(\overline{B}_\varepsilon)$ and the latter is relatively compact, a contradiction to (3.7.7). This proves that only finitely many eigenvalues $\lambda \in \mathbb{C}$ satisfy $|\lambda| > \varepsilon$. □

Combining this proposition with Theorem 3.7.21 we obtain the following corollary.

Corollary 3.7.38. *If X is a Banach space and $A \in L_c(X)$, then $\sigma(A) = \{0\} \cup P\sigma(A)$ with $P\sigma(A)$ either a finite set possibly empty or a sequence $\{\lambda_k\}_{k\geq 1} \subseteq \mathbb{C}$ exists such that $\lambda_k \to 0$ as $k \to \infty$ and each λ_k has a corresponding eigenspace that is finite dimensional.*

Now we focus on self-adjoint operators defined on a Hilbert space.

Proposition 3.7.39. *If H is a Hilbert space and $A \in L(H)$ is self-adjoint, then $P\sigma(A) \subseteq \mathbb{R}$ and eigenvectors corresponding to different eigenvalues are orthogonal.*

Proof. Since $A \in L(H)$ is self-adjoint, from Definition 3.6.13(b) it follows that

$$(A(x), y) = (x, A(y)) \quad \text{for all } x, y \in H \, .$$

Suppose $x = y \in H$. Then

$$(A(x), x) = (x, A(x)) = \overline{(A(x), x)} \quad \text{for all } x \in H \, .$$

Hence

$$(A(x), x) \in \mathbb{R} \quad \text{for all } x \in H \, . \tag{3.7.8}$$

Suppose that $\lambda \in P\sigma(A)$. then $(A(x), x) = (\lambda x, x) = \lambda\|x\|^2$, which implies, because of (3.7.8), that $\lambda = (A(x), x)/\|x\|^2 \in \mathbb{R}$.

Next let $\lambda, \mu \in P\sigma(A)$ with $\lambda \neq \mu$ and suppose that $x, u \in H$ are eigenvectors corresponding to λ, μ, respectively. Then one gets

$$\begin{aligned} (A(x), u) &= (\lambda x, u) = \lambda(x, u) \, , \\ (A(x), u) &= (x, A(u)) = (x, \mu u) = \mu(x, u) \end{aligned}$$

since the eigenvalues are real; see above. It follows that $(\lambda - \mu)(x, u) = 0$. As $\lambda \neq \mu$ we conclude that $(x, u) = 0$. □

Proposition 3.7.40. *If H is a Hilbert space and $A \in L(H)$ is self-adjoint, then $\lambda \in \sigma(A)$ if and only if* $\inf [\|\lambda x - A(x)\| : \|x\| = 1] = 0$.

Proof. $\Longrightarrow$: Suppose that $\inf[\|(\lambda i_H - A)(x)\| : \|x\| = 1] > 0$. Then there exists $c > 0$ such that

$$\|(\lambda i_H - A)(x)\| \geq c\|x\| \quad \text{for all } x \in H\,. \tag{3.7.9}$$

We will show that $(\lambda i_H - A)^{-1} \in L(H)$ and so $\lambda \in \rho(A)$. According to Proposition 3.6.35, it suffices to show that $R(\lambda i_H - A)$ is dense in H. If this is not the case, then there exists $\hat{u} \in H \setminus \{0\}$ such that $((\lambda i_H - A)(x), \hat{u}) = 0$ for all $x \in H$. This gives $(x, (\overline{\lambda} i_H - A)\hat{u}) = 0$ for all $x \in H$. Therefore $\overline{\lambda}\hat{u} = A(\hat{u})$, that is, $\overline{\lambda} \in P\sigma(A)$.

But from Proposition 3.7.39 we know that $P\sigma(A) \subseteq \mathbb{R}$. Hence, $\lambda = \overline{\lambda}$ and so $(\lambda i_H - A)(\hat{u}) = 0$, a contradiction to (3.7.9). It follows that $R(\lambda i_H - A)$ is dense in H and so Proposition 3.6.35 implies that $(\lambda i_H - A)^{-1} \in L(H)$, and thus $\lambda \in \rho(A)$.

$\Longleftarrow$: Let $\lambda \in \rho(A)$. Then $(\lambda i_H - A)^{-1} \in L(H)$. So, for $x \in H$ with $\|x\| = 1$ we get

$$\begin{aligned} 1 = \|x\| &= \|(\lambda i_H - A)^{-1}(\lambda i_H - A)(x)\| \leq \|(\lambda i_H - A)^{-1}\|_L \|(\lambda i_H - A)(x)\| \\ &\leq \|\lambda i_H - A\|_L^{-1} \|(\lambda i_H - A)(x)\|\,. \end{aligned}$$

Hence, $\|\lambda i_H - A\|_L \leq \|(\lambda i_H - A)(x)\|$, which gives

$$\|(\lambda i_H - A)^{-1}\|_L^{-1} \leq \inf[\|(\lambda i_H - A)(x)\| : \|x\| = 1]\,.$$

So, if $\inf[\|(\lambda i_H - A)(x)\| : \|x\| = 1] = 0$, then we must have $\lambda \in \sigma(A)$. ☐

Using this proposition we can conclude that the spectrum of a self-adjoint operator is real; compare with Proposition 3.7.39.

Proposition 3.7.41. *If H is a Hilbert space and $A \in L(H)$ is self-adjoint, then $\sigma(A) \subseteq \mathbb{R}$.*

Proof. Let $\lambda = \eta + i\vartheta$ with $\vartheta \neq 0$. For every $x \in H$ with $\|x\| = 1$ we obtain

$$(\lambda x - A(x), x) - (x, \lambda x - A(x)) = (\lambda - \overline{\lambda})\|x\|^2 = 2i\vartheta\,.$$

Hence,

$$\begin{aligned} 2|\vartheta| &= |(\lambda x - A(x), x) - (x, \lambda x - A(x))| \leq |(\lambda x - A(x), x)| + |(x, \lambda x - A(x)| \\ &\leq 2\|(\lambda i_H - A)(x)\|\,. \end{aligned}$$

Therefore,

$$|\vartheta| \leq \inf[\|(\lambda i_H - A)(x)\| : \|x\| = 1]\,. \tag{3.7.10}$$

So, from (3.7.10) and Proposition 3.7.40, we see that $\lambda \in \sigma(A)$ implies that $\vartheta = 0$. Thus, $\sigma(A) \subseteq \mathbb{R}$. ☐

Using this fact we can locate more precisely the spectrum of a self-adjoint operator.

Proposition 3.7.42. *If H is a Hilbert space, $A \in L(H)$ is self-adjoint, and*

$$m_A = \inf[(A(x), x) : \|x\| = 1]\,, \quad M_A = \sup[(A(x), x) : \|x\| = 1]\,,$$

then $\sigma(A) \subseteq [m_A, M_A]$.

Proof. Note that if $T = A + \mu i_H$, then $T \in L(H)$ is self-adjoint and $m_T = m_A + \mu$ as well as $M_T = M_A + \mu$. So, without any loss of generality we may assume that $0 \leq m_A \leq M_A$.

From Proposition 3.6.16, we know that $M_A = \|A\|_L$, while from Proposition 3.7.41, we know that $\sigma(A) \subseteq \mathbb{R}$. We will show that, for every $\vartheta > 0$, $\lambda = M_A + \vartheta \notin \sigma(A)$. According to Proposition 3.7.40, it suffices to show that

$$\inf[\|(\lambda i_H - A)(x)\| : \|x\| = 1] > 0 .$$

For every $x \in H$ with $\|x\| = 1$ one has

$$((\lambda i_H - A)(x), x) = (\lambda x, x) - (A(x), x) \geq (\lambda - M_A)\|x\|^2 = \vartheta\|x\|^2 = \vartheta .$$

This gives $0 < \vartheta \leq \|(\lambda i_H - A)(x)\|$ for all $x \in H$ with $\|x\| = 1$. Therefore, $0 < \vartheta \leq \inf[\|(\lambda i_H - A)(x)\| : \|x\| = 1]$. Then, due to Proposition 3.7.40, this finally proves that $\lambda \notin \sigma(A)$.

Similarly we show that for every $\vartheta > 0$, $\lambda = m_A - \vartheta \notin \sigma(A)$. Hence, we conclude that $\sigma(A) \subseteq [m_A, M_A]$. □

Proposition 3.7.43. *If H is a Hilbert space and $A \in L(H)$ is self-adjoint, then $m_A, M_A \in \sigma(A)$; see Proposition* 3.7.42.

Proof. As before (see the proof of Proposition 3.7.42), we may assume that $0 \leq m_A \leq M_A$. Recall that $M_A = \|A\|_L$; see Proposition 3.6.16. Let $\{x_n\}_{n\geq 1} \subseteq H$ with $\|x_n\| = 1$ for all $n \in \mathbb{N}$ such that

$$(A(x_n), x_n) \to M_A = \|A\|_L \quad \text{as } n \to \infty . \tag{3.7.11}$$

Then, using the fact that $(A(x), x) \geq 0$ for all $x \in H$ and the validity of (3.7.11), it follows that

$$\begin{aligned} 0 &\leq \|(M_A i_H - A)(x_n)\|^2 = (M_A x_n - A(x_n), M_A x_n - A(x_n)) \\ &= M_A^2 + \|A(x_n)\|^2 - 2M_A(A(x_n), x_n) \leq M_A^2 + M_A^2 - 2M_A(A(x_n), x_n)) \to 0 \end{aligned}$$

as $n \to \infty$. Hence $\inf[\|(M_A i_H - A)(x)\| : \|x\| = 1] = 0$, which gives $M_A \in \sigma(A)$; see Proposition 3.7.40.

Similarly we show that $m_A \in \sigma(A)$. □

Next we restrict further ourselves to compact self-adjoint operators.

Proposition 3.7.44. *If H is a Hilbert space and $A \in L_c(H)$ is self-adjoint, then $P\sigma(A) \neq \emptyset$.*

Proof. If $A = 0$, then $\lambda = 0 \in P\sigma(A)$. So, suppose that $A \neq 0$. Then Proposition 3.7.43 gives $\|A\|_L \in \sigma(A)$. Since $\|A\|_L \neq 0$, Corollary 3.7.38 implies that $\|A\|_L \in P\sigma(A)$. □

Proposition 3.7.45. *If H is an infinite dimensional Hilbert space and $A \in L_c(H) \setminus \{0\}$ is self-adjoint, then $\sigma(A) = \{0\} \cup \{\lambda_k\}_{k\geq 1}$ with λ_k being distinct nonzero eigenvalues of A, one of these eigenvalues equals $\|A\|_L$ and $\{\lambda_k\}_{k\geq 1}$ is either finite or a countable sequence such that $\lambda_k \to 0$. Moreover, the Hilbert space H admits an orthonormal basis consisting of eigenvectors corresponding to the eigenvalues of A.*

Proof. This is basically Corollary 3.7.38. From Proposition 3.7.43 we also know that one of the eigenvalues equals $\|A\|_L$. It remains to prove the last part of the proposition concerning the basis of H. Let $\lambda \in P\sigma(A)$ and let $N_\lambda = N(\lambda i_H - A)$. From Theorem 3.7.21 we know that $\dim N_\lambda < +\infty$. Let B_λ be an orthonormal basis for N_λ and let $B = \bigcup_{\lambda \in P\sigma(A)} B_\lambda$. From Proposition 3.7.39 we know that $B \subseteq H$ is an orthonormal set and $\overline{\operatorname{span}} B$ contains all the eigenvectors of A. Suppose that $H \neq \overline{\operatorname{span}} B$ and let $V = (\overline{\operatorname{span}} B)^\perp$. Note that $\overline{\operatorname{span}} B$ is A-invariant. Hence so is V. One has $\sigma(A) = \sigma(A|_{\overline{\operatorname{span}} B}) + \sigma(A|_V)$. But $\sigma(A|_V)$ contains an eigenvalue (see Proposition 3.7.44), and so a corresponding eigenvector u as well. Then u is also an eigenvector of A and so $u \in V \cap \overline{\operatorname{span}} B$, $u \neq 0$, a contradiction. This means that $H = \overline{\operatorname{span}} B$, and so B is an orthonormal basis of H. ☐

Corollary 3.7.46. *If H is a Hilbert space and $A \in L_c(H)$ is self-adjoint, then $\sigma(A) = \overline{P\sigma(A)}$.*

Proof. If H is finite dimensional, then $\sigma(A) = P\sigma(A)$ and it is compact; see Proposition 3.7.33. If H is infinite dimensional, then $P\sigma(A)$ is a countable sequence or a finite sequence. If it is a countable sequence, then the conclusion follows from Proposition 3.7.44. If it is a finite sequence, then since the eigenspaces for the nonzero eigenvalues are finite dimensional (see Corollary 3.7.38), and H is infinite dimensional, then on account of Proposition 3.7.44 we must have that $\lambda = 0 \in P\sigma(A)$. ☐

We have reached the main result on the spectral analysis of compact self-adjoint operators defined on a Hilbert space. The result is known as the "Spectral Decomposition Theorem."

Theorem 3.7.47 (Spectral Decomposition Theorem). *If H is an infinite dimensional separable Hilbert space and $A \in L_c(H)$ is self-adjoint, then there exists an orthonormal basis $\{e_k\}_{k\geq 1} \subseteq H$ consisting of eigenvectors corresponding to the distinct eigenvalues $\{\lambda_k\}_{k\geq 1} \subseteq \mathbb{R}$ and*

$$A(x) = \sum_{k\geq 1} \lambda_k (x, e_k) e_k \quad \textit{for all } x \in H .$$

Moreover, for every $\lambda \in \rho(A)$ and $x \in H$, it holds that

$$R(\lambda)(x) = \sum_{k\geq 1} \frac{(x, e_k)}{\lambda - \lambda_k} e_k .$$

Proof. Let $\{e_k\}_{k\geq 1} \subseteq H$ be an orthonormal basis of H consisting of eigenvectors; see Propositions 3.7.43 and 3.5.47. Then, for $1 \leq n < m$, one has

$$\left\| \sum_{k=n}^{m} \lambda_k (x, e_k) e_k \right\|^2 = \sum_{k=n}^{m} |\lambda_k (x, e_k)|^2 \leq \|A\|_L \sum_{k=n}^{m} |(x, e_k)|^2 \to 0$$

as $n \to \infty$; see Proposition 3.7.33. Hence, $\sum_{k\geq 1} \lambda_k (x, e_k) e_k$ converges in H.

If $\|x\| \le 1$, then, for every $n \in \mathbb{N}$, we derive

$$\begin{aligned}\left\|\sum_{k=1}^{n} \lambda_k (x, e_k) e_k\right\|^2 &= \sum_{k=1}^{n} \lambda_k^2 |(x, e_k)|^2 \le \|A\|_L^2 \sum_{k=1}^{n} |(x, e_k)|^2 \\ &\le \|A\|_L^2 \sum_{k\ge 1} |(x, e_k)|^2 = \|A\|_L^2 \|x\|^2 .\end{aligned}$$

Consider the operator T defined by $T(x) = \sum_{k\ge1} \lambda_k (x, e_k) e_k$. Of course, $T \in L(H)$. Hence, $A(e_k) = T(e_k)$ for all $k \in \mathbb{N}$ and so $A = T$.

Now suppose that $\lambda \in \rho(A)$. Recalling that $\sigma(A) = \mathbb{C} \setminus \rho(A)$ is compact, it follows that $d(\lambda, \sigma(A)) > \vartheta > 0$. Hence, $|\lambda - \lambda_k| > \vartheta$ for all $k \in \mathbb{N}$. Therefore,

$$\left\|\sum_{k=n}^{m} \frac{(x, e_k)}{\lambda - \lambda_k} e_k\right\|^2 = \sum_{k=n}^{m} \frac{|(x, e_k)|^2}{|\lambda - \lambda_k|^2} < \frac{1}{\vartheta^2} \sum_{k=n}^{m} |(x, e_k)|^2 .$$

This shows that $\sum_{k\ge1} (x, e_k)/(\lambda - \lambda_k) e_k$ is convergent in H for all $x \in H$.

Let $T(x) = \sum_{k\ge1} (x, e_k)/(\lambda - \lambda_k) e_k$. Then, for $\|x\| \le 1$, we obtain

$$\left\|\sum_{k=1}^{n} \frac{(x, e_k)}{\lambda - \lambda_k} e_k\right\|^2 \le \frac{1}{\vartheta^2} \sum_{k=1}^{n} |(x, e_k)|^2 = \frac{1}{\vartheta^2} \|x\|^2 \le \frac{1}{\vartheta^2} .$$

Thus, $T \in L(H)$.

Since $x = \sum_{k\ge1} (x, e_k) e_k$, we have $A(x) = \sum_{k\ge1} \lambda_k (x, e_k) e_k$ and

$$(\lambda i_H - A)(x) = \sum_{k\ge1} (\lambda - \lambda_k)(x, e_k) e_k .$$

As $(e_k, e_i) = \delta_{k,i}$ we then get

$$(\lambda i_H - A)(T(x)) = \sum_{k,i\ge1} (\lambda - \lambda_k) \frac{(x, e_i)}{\lambda - \lambda_i} (e_k, e_i) e_k = \sum_{k\ge1} (x, e_k) e_k = x .$$

Similarly we show that $T((\lambda i_H - A)(x)) = x$ for all $x \in H$. Therefore, $T = R(\lambda)$. □

We conclude this section by introducing two more classes of bounded linear operators of Hilbert space into itself.

Definition 3.7.48. Let H be a Hilbert space and $A \in L(H)$.

(a) We say that A is **normal** if $A \circ A^* = A^* \circ A$.

(b) We say that A is **unitary** if A is invertible and $A^{-1} = A^*$.

Remark 3.7.49. Clearly every unitary operator is normal and every self-adjoint operator is normal.

Proposition 3.7.50. *If H is a Hilbert space and $A \in L(H)$, then A is normal if and only if $\|A(x)\| = \|A^*(x)\|$ for all $x \in H$.*

Proof. For every $x \in H$, we derive

$$\begin{aligned} \|A(x)\|^2 - \|A^*(x)\|^2 &= (A(x), A(x)) - (A^*(x), A^*(x)) \\ &= (A^*(A(x)), x) - (A(A^*(x)), x) \\ &= ((A^* \circ A - A \circ A^*)(x), x) . \end{aligned} \tag{3.7.12}$$

From (3.7.12) it follows that A is normal if and only if $\|A(x)\| = \|A^*(x)\|$ for all $x \in H$. □

Proposition 3.7.51. *If H is a Hilbert space and $A \in L(H)$ is surjective, then the following statements are equivalent:*
(a) *A is unitary.*
(b) *$(A(x), A(u)) = (x, u)$ for all $x, u \in H$.*
(c) *A is an isometry.*

Proof. (a) $\Longrightarrow$ (b): For every $x, u \in H$ it holds that

$$(A(x), A(y)) = (A^*(A(x)), u) = (x, u) .$$

(b) $\Longleftrightarrow$ (c): This follows from the polarization identities; see Proposition 3.5.6(b).

(c) $\Longrightarrow$ (a): The operator A is an isometry and surjective, and hence, $A^{-1} \in L(H)$; see Theorem 3.2.10. Moreover, for all $x, u \in H$, one has

$$(A^*(A(x)), u) = (A(x), A(u)) = (x, u) ,$$

since (b) is equivalent to (c). Hence, $A^* \circ A = i_H$ and similarly $A \circ A^* = i_H$, and so $A^{-1} = A^*$. □

Remark 3.7.52. So according to the proposition above, $A \in L(H)$ is unitary if and only if it preserves inner products.

3.8 Remarks

(3.1) The major development of mathematics in the twentieth century was the emphasis on the axiomatic method. This abstract tendency with emphasis on the structural properties led to the development of whole new areas such as “Functional Analysis” with the seminal contributions of Banach, von Neumann, and Riesz to mention only a few major figures and to “Modern Algebra” where prominent figures were Noether and van der Waerden. In this approach, the emphasis is not on the objects but on the rules used to handle them, which are the same for many different classes of objects. The power of the axiomatic method can be traced back in the work of Euclid who provided a model for space locally. The first breakthrough in the abstract axiomatic approach was achieved by Fréchet who introduced abstract metric spaces in this thesis [117]. He was the first to go beyond the familiar concrete Euclidean space setting. The normed space axioms (see Definition 3.1.13(e)) were first introduced by Banach [23] in this thesis. Normed spaces are a subset of metric spaces. The thing that makes normed spaces such a prolific concept is the linkage between the algebraic and the topological structures of the space. This is expressed by the requirement that the two algebraic operations, namely vector addition and scalar multiplication, are continuous. This leads at a higher level of generality to the notion of a topological space. Moreover, the convexity of the balls in a normed space lead to topological vector spaces with a local neighborhood basis consisting of convex sets. These are the locally convex spaces (see Definition 3.1.13(b)) first introduced by von Neumann [303]. Until the mid-forties, the study of functional analysis focused on normed spaces. The first major paper on the theory of locally convex spaces was that of Dieudonné-Schwartz [84] motivated by Schwartz's construction of the theory of distributions. Lemma 3.1.20, called Riesz Lemma, was proved by Riesz [246] and turned out to be a fruitful result for many occasions. Theorem 3.1.30 is due to Carathéodory [62] and Theorem 3.1.41, due to Kolmogorov [179], seems to be the first theorem about locally convex spaces. The Hahn–Banach Theorem (see Theorem 3.1.42) is crucial in the development of the theory of normed spaces. The first version of it was due to Minkowski, who proved that every boundary point in the closed unit ball of a finite dimensional normed space admits at least one supporting hyperplane through it. Later Helly [143] generalized the ideas of Minkowski to certain separable spaces. Fifteen years later, in 1927, Hahn [138] starting from the work of Helly, proved an extension theorem in a more general form without any separability hypothesis. Soon thereafter, we have the result of Banach [24] (see also Banach [25]), who proved the theorem in general vector spaces apart from any topology. The Hahn–Banach Theorem turned out to be a major tool in the development of the theory of locally convex spaces. Although the original proof uses transfinite induction, this part of the argument was later replaced by use of the Zorn's Lemma. The complex version of the result (see Theorem 3.1.44) is due to Bohnenblust–Sobczyk [38] and Suchomlinov [280]. Theorem 3.1.59, the First Separation Theorem, is due to Edelheit [96]. Theorem 3.1.60, the Strong Separation Theorem, is due to Tukey [287] and Klee [176].

(3.2) Theorem 3.2.1, the Uniform Boundedness Principle, was first proved by Hahn [137] for sequences of linear functionals. A more general form was produced by Hildebrandt [148]. The general version of the result and a proof based on the Baire Category Theorem were provided by Banach–Steinhaus [26]; see also Theorem 3.2.2. Theorem 3.2.9, the Open Mapping Theorem, was proved by Schauder [263] for Banach spaces. Banach [25] extended the result to Fréchet spaces; see Definition 3.1.13(d). Theorem 3.2.10 and Theorem 3.2.14, the Closed Graph Theorem, are due to Banach [25]. Banach [25] extended both the Open Mapping Theorem and the Closed Graph Theorem to topological groups. The book of Banach [25] turned out to be one of the most influential books in analysis and remains a reference even today.

Definition 3.8.1. Let P be a property of normed spaces. Suppose that if X is a normed space and $V \subseteq X$ is a closed subspace such that if two of the spaces X, V, X/V have property P, then so does the third. Then we say that P is a **three space property**.

Using this notion we can improve Proposition 3.2.17 in the following way.

Proposition 3.8.2. *Completeness is a three space property.*

(3.3) We point out that Banach [25] worked only with weakly convergent sequences and did not use the notion of "weak topology." In certain occasions this led to unnecessary separability assumptions. The first explicity description of weak neighborhoods in a Hilbert space was given by von Neumann [301] who was the first to recognize that the weak topology is indeed a topology. He also realized the nonmetrizability of the weak topology in an infinite dimensional normed space; see Proposition 3.3.15. Further discussion on this issue can be found in Wehausen [306]. Proposition 3.3.16 was first proven for $X = l^2$ by von Neumann [301]. Theorem 3.3.18 is due to Mazur [210]. Earlier particular versions of this result for the Banach space $C[0,1]$ can be found in Gillespie–Hurwitz [128] and Zalcwasser [312]. That bounded linear operators are weakly continuous was first observed by Banach [24]. The converse (see Proposition 3.3.23) is due to Bade [19]. Theorem 3.3.37, Goldstine's Theorem, is naturally due to Goldstine [132] and Theorem 3.3.38, Alaoglu's Theorem, was proved by Alaoglu [2]. For separable Banach spaces the theorem can be found in Banach [25]. For this reason some people call it the "Banach–Alaoglu Theorem"; see, for example, Megginson [212, p. 229]. Theorem 3.3.41 is due to James [164] and is one of the most influential results in Banach space theory.

Another locally convex topology on X^* being the dual of the normed space X is the bounded weak* topology introduced by Dieudonné [83].

Definition 3.8.3. Let X be a normed space. The **bounded weak* topology** (or the bw***-topology**) is the strongest topology on X^* which coincides with the relative w*-topology on each set $t\overline{B}_1^{X^*} = \{x^* \in X^* : \|x^*\|_* \leq t\}$. Therefore a set $U \subseteq X^*$ is bw*-open if and only if $U \cap t\overline{B}_1^{X^*}$ is relatively w*-open in $t\overline{B}_1^{X^*}$ for every $t > 0$ and $C \subseteq X^*$ is bw*-closed if and only if $C \cap t\overline{B}_1^{X^*}$ is relatively w*-closed in $\overline{B}_1^{X^*}$ for all $t > 0$.

Remark 3.8.4. It can be shown (see, for example Dunford–Schwartz [94, Lemma V.5.4, p. 427]) that a local basis at the origin for the bw*-topology is given by the sets

$$B(S) = \{x^* \in X^* : |\langle x^*, x\rangle| < 1 \text{ for all } x \in S\}\,,$$

where $S = \{x_k\}_{k\geq 1} \subseteq X$ is a sequence converging to zero. We have $w^* \subseteq \text{bw}^* \subseteq \text{norm}$.

These inclusions are strict if X is an infinite dimensional normed space. Directly from Definition 3.8.3 we see that if $\{x^*_\alpha\}_{\alpha\in I} \subseteq X^*$ is a bounded net and $x^* \in X^*$, then $x^*_\alpha \xrightarrow{\text{w}^*} x^*$ if and only if $x^*_\alpha \xrightarrow{\text{bw}^*} x^*$. Of course $X^*_{\text{bw}^*}$ is a locally convex space.

Proposition 3.8.5. *If X is a Banach space, then $X = (X^*_{\text{w}^*})^* = (X^*_{\text{bw}^*})^*$.*

Using this proposition one can show the following theorem known as the "Krein–Smulian Theorem."

Theorem 3.8.6 (Krein–Smulian Theorem). *If X is a Banach space and $C \subseteq X^*$ is a nonempty convex set, then C is* w**-closed if and only if $C \cap t\overline{B}_1^{X^*}$ is* w**-closed for every $t > 0$, that is, C is* w**-closed if and only if C is* bw**-closed.*

Remark 3.8.7. As in Mazur's Theorem (see Theorem 3.3.18), in the theorem above we see that an algebraic property, namely the convexity of C, has topological consequences.

Corollary 3.8.8. *If X is a separable Banach space and $C \subseteq X^*$ is a nonempty convex set, then C is* w**-closed if and only if it is weakly* sequentially closed.*

We can introduce one more locally convex topology on X^*. Recall that the weak* topology is the weakest topology τ on X^* such that $(X^*_\tau)^* = X$. Suppose we ask for the strongest (finest) topology m on X^* for which $(X^*_{\mathrm{m}})^* = X$ is satisfied.

Theorem 3.8.9. *There exists a strongest topology* m *on X^* such that $(X^*_{\mathrm{m}})^* = X$. This is the topology of uniform convergence on all* w*-compact sets, that is, $x^*_\alpha \xrightarrow{\mathrm{m}} x^*$ in X^* if and only if* $\sup[\langle x^*_\alpha - x^*, u\rangle : u \in K] \to 0$ *for all* w*-compact $K \subseteq X$. The space X^*_{m} is locally convex and* m *is called the **Mackey topology** on X^* and is denoted by* $\mathrm{m}(X^*, X)$.

We have already seen how important the notion of convexity is. Next we will see that in some convex sets we can isolate special points of them that in fact generate the set.

Definition 3.8.10. Let X be a topological vector space and $C \subseteq X$ be a nonempty, closed, convex set. A set $E \subseteq C$ is **extremal** in C if E is nonempty, closed, convex and if $x, u \in C$ and $(1-\lambda)x + \lambda u \in E$ for some $\lambda \in (0, 1)$, then $x, u \in E$. An **extreme point** of C is an $x \in C$ such that $\{x\}$ is an extremal subset of C, that is, x is an extreme point of C if it does not lie in the interior of any nontrivial closed line segment of C. By ext C we denote the set of extreme points of C.

The following is the basic theorem about extreme points and it is known as the "Krein–Milman Theorem."

Theorem 3.8.11 (Krein–Milman Theorem). *If X is a locally convex space and $C \subseteq X$ is nonempty, compact, and convex, then* $\operatorname{ext} C \neq \emptyset$ *and* $C = \overline{\operatorname{conv}} \operatorname{ext} C$.

For more on the structure of convex sets, we refer to Giles [127]. The books of Aliprantis-Border [6], Beauzamy [29], Brézis [48], Denkowski-Migórski-Papageorgiou [77], Diestel [79], Fabian et al. [106], Giles [127], Holmes [155], Megginson [212], Rudin [260], and Yosida [311] discuss in detail the weak and weak* topologies.

(3.4) Reflexive Banach spaces were introduced by Hahn [138]. He called them **regular**. The term "reflexive" is due to Lorch [204] and Theorem 3.4.5 is due to James [163]. There are other useful characterizations of reflexivity. We mention three of them. The first is due to Smulian [273].

Theorem 3.8.12. *If X is a Banach space, then X is reflexive if and only if for every decreasing sequence $\{C_n\}_{n\geq 1}$ of nonempty, bounded, closed, convex subsets of X, it holds that $\bigcap_{n\geq 1} C_n \neq \emptyset$.*

The second is due to James [165].

Theorem 3.8.13. *If X is a Banach space, then the following statements are equivalent:*
(a) *X is not reflexive.*
(c) *For every $\lambda \in (0,1)$ there exists a sequence $\{x_n\}_{n\geq 1} \subseteq X$ with $\|x_n\| = 1$ for all $n \in \mathbb{N}$ such that $d(\operatorname{conv}\{x_k\}_{k=1}^n, \overline{\operatorname{conv}}\{x_k\}_{k\geq n+1}) \geq \lambda$ for every $n \in \mathbb{N}$.*
(c) *For some $\lambda \in (0,1)$ there exists a sequence $\{x_n\}_{n\geq 1} \subseteq X$ with $\|x_n\| = 1$ for all $n \in \mathbb{N}$ such that $d(\operatorname{conv}\{x_k\}_{k=1}^n, \overline{\operatorname{conv}}\{x_k\}_{k\geq n+1}) \geq \lambda$ for every $n \in \mathbb{N}$.*

Remark 3.8.14. The interesting feature of the theorem above for reflexivity is that it is intrinsic. Namely, it does not require any knowledge of X^* or X^{**}.

The third is also due to James [165].

Theorem 3.8.15. *If X is a Banach space, then X is reflexive if and only if every $x^* \in X^*$ is norm attaining, that is, there exists $x_0 \in X$, $\|x_0\| \leq 1$ such that $\|x^*\|_* = \langle x^*, x_0\rangle$.*

It is easy to check the next proposition.

Proposition 3.8.16. *Separability and reflexivity are three space properties; see Definition* 3.8.1.

The direct assertions in Theorem 3.4.12 are due to Banach [25] and concerning Theorem 3.4.14, the Eberlein–Smulian Theorem, Smulian [274] showed that weakly compact sets are weakly sequentially compact. Later Eberlein [95] proved the converse. Whitley [308] provided an elementary proof of the theorem. Theorem 3.4.18 reveals the distinctive character of weakly compact sets. They are sequentially compact and each subset of a weakly compact set has a sequentially determined closure. These properties are a particular instance of a more general class of spaces known as **angelic space**; see Floret [113].

Strict convexity and uniform convexity (see Definition 3.4.21(a),(b)) were introduced by Clarkson [68]. Local uniform convexity was introduced by Lovaglia [205]. In the paper of Smith [272], we find examples of reflexive Banach spaces that are locally uniformly convex but not uniformly convex, and of reflexive and nonreflexive Banach spaces that are strictly convex but not locally uniformly convex. The Kadec–Klee property is also called the **Radon–Riesz property** or the H-property; see Day [73].

Proposition 3.8.17. *If X is a uniformly convex Banach space and $V \subseteq X$ is a closed subspace, then X/V is uniformly convex as well.*

(3.5) The notion of abstract Hilbert spaces was introduced by von Neumann [300]. His definition is for a separable space and his aim was to develop the spectral theory for classes of operators on this abstract space. Earlier special realizations of Hilbert spaces were examined by many authors. In particular, Hilbert [147] published between 1904 and 1910 a series of six papers collected in book form developing Hilbert space methods to study integral equations. The name Hilbert space was first used by Riesz [241] for what we know today as l^2. Theorem 3.5.21 was stated by Riesz [242] and Fréchet [118] as separate notes in the same issue of the "Comptes Rendus." In addition to Bessel's inequality (see Proposition 3.5.44), we should also mention the so-called **Parseval's identity**.

Proposition 3.8.18 (Parseval's identity). *If H is a Hilbert space and $\{e_n\}_{n\geq 1} \subseteq H$ is an orthonormal set, then $\{e_n\}_{n\geq 1}$ is an orthonormal basis for H if and only if $\|x\|^2 = \sum_{n\geq 1}(x, e_n)^2$ for all $x \in H$.*

The Gram–Schmidt Orthonormalization Process was first discovered by the Danish statistician Gram. It was elaborated further by Schmidt [265] who demonstrated its usefulness in the study of Hilbert spaces.

(3.6) The operator topologies in Definition 3.6.1 were introduced, in the context of Hilbert spaces, by von Neumann [301]. The notion of adjoint operators (see Definition 3.6.6) was first introduced by Banach [25]. Of course the notion was used earlier in the context of matrix theory. The notion of projection operator (see Definition 3.6.13(a), (c)) is due to Schmidt [265]. The theory of unbounded linear operators was stimulated by attempts in the late 1920s to give quantum mechanics a rigorous mathematical foundation. The first fundamental works on this subject are those of von Neumann [300, 301], [302], and Stone [279]. A more detailed treatment of unbounded linear operators can be found in the books of Goldberg [131], Hille-Phillips [149], Kato [170], Reed-Simon [239], and Weidmann [307].

We state a theorem related to the material of this section.

Theorem 3.8.19. *If X, Y are Banach spaces and $A \in L(X, Y)$, then the following statements are equivalent:*
(a) *$R(A) \subseteq Y$ is closed;*
(b) *$\inf[\|x + v\|_X : A(v) = 0] \leq c\|Ax\|_Y$ for all $x \in X$ and for some $c > 0$;*
(c) *$R(A^*) \subseteq X^*$ is closed;*
(d) *$\inf[\|y^* + x^*\|_{Y^*} : A^*(v^*) = 0] \leq c\|A^*y^*\|_{X^*}$ for all $y^* \in Y^*$ and for some $c > 0$.*

(3.7) The notion of compact operators (see Definition 3.7.1) is essentially due to Hilbert [147]. However, the general definition was given by Riesz [246]. Theorem 3.7.10 is due to Schauder [263]. It is the starting point of the Leray–Schauder degree theory; see Section 6.2. Theorem 3.7.17 is due to Schauder [263]. The terminology "Fredholm Operator" was introduced in recognition of the pioneering work of E. Fredholm on integral equations. The work of Fredholm influenced Hilbert. Fredholm operators exhibit nice composition and stability properties.

Proposition 3.8.20. *If X, Y, V are Banach spaces and $A \in L(X, Y)$, $T \in L(Y, V)$ are Fredholm operators, then $T \circ A \in L(X, V)$ is a Fredholm operator and $i(T \circ A) = i(A) + i(T)$.*

Proposition 3.8.21. *If X, Y are Banach spaces and $A \in L(X, Y)$ is a Fredholm operator, then the following hold:*
(a) *$A + L$ is a Fredholm operator for every $L \in L_c(X, Y)$ and $i(A + L) = i(A)$;*
(b) *there exists $\varepsilon > 0$ such that if $T \in L(X, Y)$ with $\|T\|_L < \varepsilon$, then $A + T$ is a Fredholm operator and $i(A + T) = i(A)$.*

The terminology “spectrum” of $A \in L(X)$ comes from Hilbert who published some papers in book form [147] initiating modern spectral theory. The mathematical setting of self-adjoint operators on a Hilbert space was an important mathematical tool for the development by physicists of the theory of quantum mechanics.

Definition 3.8.22. Let H be a Hilbert space and $A \in L(H)$. We say that A is **positive** (or **monotone**) if $(A(x), x) \geq 0$ for all $x \in H$. Then we write $A \geq 0$. Moreover, if $A, T \in L(H)$, then we write $A \geq T$ if and only if $A - T \geq 0$.

Remark 3.8.23. Every positive $A \in L(H)$ with H being a complex Hilbert space is automatically self-adjoint. This is false for real Hilbert spaces. Moreover, $A^* \circ A \geq 0$ for any $A \in L(H)$.

Proposition 3.8.24. *If H is a Hilbert space, $A \in L(H)$ and $A \geq 0$, then there exists a unique $T \in L(H)$, $T \geq 0$ such that $T^2 = A$. Moreover, T commutes with every bounded linear operator, which commutes with A. We denote T by $A^{1/2}$, the square root of A.*

Definition 3.8.25. Let H be a Hilbert space and $A \in L(H)$. Then $|A| = (A^* \circ A)^{1/2}$; see Proposition 3.8.24.

Finally let us state a result on the usage of unitary operators (see Definition 3.7.48), to identify compact self-adjoint operators.

Proposition 3.8.26. *If H is a separable Hilbert space and $A, T \in L_c(H)$ is self-adjoint, then there exists a unitary operator $U \in L(H)$ such that $U^* \circ T \circ U = A$ if and only if $\dim N(\lambda U - A) = \dim N(\lambda I - T)$ for all $\lambda \in \mathbb{C}$. We say that the operators A and T are **unitarily equivalent**.*

Problems

Problem 3.1. Let X be a vector space and let $\rho\colon X \to \mathbb{R}_+$ be a function such that
(a) $\rho(x) = 0$ if and only if $x = 0$;
(b) $\rho(\lambda x) = |\lambda|\rho(x)$ for all $x \in X$ and for all $\lambda \in \mathbb{F}$.
Show that ρ is a norm if and only if $\overline{B}_1 = \{x \in X\colon \rho(x) \leq 1\}$ is convex.

Problem 3.2. Let X be a vector space and let $\|\cdot\|, |\cdot|$ be two equivalent norms on X, that is, they generate the same topology. Show that $(X, \|\cdot\|), (X, |\cdot|)$ are either both Banach spaces or both are noncomplete.

Problem 3.3. Let X be a topological vector space and let $\{C_k\}_{k=1}^n$ be a finite family of compact, convex subsets of X. Show that $\operatorname{conv}(\bigcup_{k=1}^n C_k)$ is compact.

Problem 3.4. Let X be a normed space, $Y \subseteq X$ be a closed subspace, and let $V \subseteq X$ be a finite dimensional subspace. Show that $Y + V = \{y + v\colon y \in Y, v \in V\} \subseteq X$ is closed.

Problem 3.5. Let X be a normed space and $V \subseteq X$ is a finite dimensional subspace. Show that there exists $x \in X$ with $\|x\| = 1$ such that $1 = d(x, V)$.

Problem 3.6. Let X be a normed space that is a Polish space for the norm topology. Show that X is a Banach space.

Problem 3.7. Show that a normed space X is complete, that is, X is a Banach space, if and only if every absolutely convergent series in X is convergent.

Problem 3.8. Let K be a compact topological space and let $D \subseteq K$ be a closed set. Show that $C(D)$ is isomorphic to a quotient of $C(K)$.

Problem 3.9. Let K, D be compact topological spaces and let $A\colon C(K) \to C(D)$ be a linear operator such that $f \geq 0$ implies $A(f) \geq 0$, that is, A is positive. Show that A is continuous and $\|A\|_L = \|A(1)\|_{C(D)}$ with $1 \in C(K)$ is the constant function equal to 1.

Problem 3.10. Let $X = C[0, 1]$, $u \in X$, and $f\colon X \to \mathbb{R}$ be a linear function defined by $f(y) = \int_0^1 y(t)u(t)dt$ for all $y \in X$. Show that $f \in X^*$ and $\|f\|_* = \int_0^1 |u(t)|dt$.

Problem 3.11. Let X be a normed space and $C \subseteq X$ be a nonempty set. Show that $\overline{\operatorname{conv}} C = \{x \in X\colon \langle x^*, x\rangle \leq \sigma_C(x^*) = \sup\{\langle x^*, c\rangle\colon c \in C\}$, whereby $\sigma_C\colon X^* \to \overline{\mathbb{R}} = \mathbb{R} \cup \{+\infty\}$ is called the **support function of** C.

Problem 3.12. Show that every normed space is isometrically isomorphic to a subspace of $C(K)$ for some compact topological space K.

Problem 3.13. Let X, Y be Banach spaces and let $A \in L(X, Y)$ be surjective. Show that there exists $M > 0$ such that for every $y \in Y$ there is $x \in A^{-1}(y)$ satisfying $\|x\|_X \leq M\|y\|_Y$.

Problem 3.14. Let X, Y be Banach spaces and let $A \in L(X, Y)$ be surjective. Show that Y is isomorphic to $X/N(A)$.

Problem 3.15. Let X be a Banach space and let $C \subseteq X$ be a weakly compact set. Show that C is bounded.

Problem 3.16. Let X be a normed space and $\{x_n^*\}_{n\geq 1} \subseteq X^*$. Suppose that there exists a sequence $\{\varepsilon_n\}_{n\geq 1} \subseteq (0, +\infty)$ with $\varepsilon_n \to 0$ such that for every $x \in X$ there exists $\eta_x > 0$ with $|\langle x_n^*, x\rangle| \leq \eta_x \varepsilon_n$ for all $n \in \mathbb{N}$. Show that $x_n^* \to 0$.

Problem 3.17. Show that separability and reflexivity are three space properties; see Definition 3.8.1.

Problem 3.18. Show that a normed space X is reflexive if and only if each separable, closed subspace $V \subseteq X$ is reflexive.

Problem 3.19. Show that if Y is an infinite dimensional subspace of l^1, then Y is not reflexive.

Problem 3.20. Let X be a separable Banach space. Show that there exists $x_n^* \in X^*$ with $\|x_n^*\|_* = 1$ for all $n \in \mathbb{N}$ such that $\{x_n^*\}_{n\geq 1}$ is separating on X.

Problem 3.21. Let X, Y be Banach spaces with X being reflexive and let $A \in L(X, Y)$ be surjective. Show that Y is reflexive as well.

Problem 3.22. Let X be a Banach space with a separable dual X^*. Show that $\mathcal{B}(X^*) = \mathcal{B}(X^*_{w^*})$. Recall that if Z is a Hausdorff topological space, then $\mathcal{B}(Z)$ denotes the Borel σ-algebra of Z.

Problem 3.23. Let X be a normed space and let $C \subseteq X^*$ be a nonempty, w*-closed set. Show that for any given $x^* \in X^*$ there exists $u_0^* \in C$ such that $\|x^* - u_0^*\|_* = d(x^*, C)$. A set that has this best approximation property for every element in the space is called **proximinal**.

Problem 3.24. Show that a Banach space X is reflexive if and only if every closed convex set is proximinal; see Problem 3.23.

Problem 3.25. Let X, Y be two nontrivial normed spaces and assume that $L(X, Y)$ equipped with the operator norm is a Banach space. Show that Y is a Banach space.

Problem 3.26. Let X be a reflexive Banach space and let Y be another Banach space that is isomorphic to X. Show that Y is reflexive as well.

Problem 3.27. Let X, Y be Banach spaces with X being nonreflexive and Y being reflexive. Suppose that $A \in L(X, Y)$ is injective. Show that $R(A) \subseteq Y$ cannot be closed.

Problem 3.28. Let X be a Banach space and let $C \subseteq X^*$ be a w*-compact set. Show that $\overline{\mathrm{conv}}^{w^*} C$ is w*-compact.

Problem 3.29. Let X be a separable Banach space. Show that X^* is w*-separable.

Problem 3.30. Let H and V be real Hilbert spaces and let $k\colon H \times V \to \mathbb{R}$ be a bilinear form that is bounded, that is, there exists $c > 0$ such that $|k(u, v)| \leq c\|u\|_H\|v\|_V$ for all $u \in H$ and for all $v \in V$. Show that there exists a unique $A \in L(H, V)$ such that $k(u, v) = (A(u), v)_V$ for all $u \in H$ and for all $v \in V$.

Problem 3.31. Let H be a Hilbert space and let $\{e_n\}_{n\geq 1} \subseteq H$ be an orthonormal set. Suppose that $u = \sum_{n\geq 1} a_n e_n$. Show that $a_n = (u, e_n)$ for all $n \in \mathbb{N}$.

Problem 3.32. Let H, V be infinite dimensional separable Hilbert spaces, let $\{e_n\}_{n\geq 1} \subseteq H$ be an orthonormal basis for H, and let $\{\xi_n\}_{n\geq 1} \subseteq V$ be an orthonormal basis for V. Suppose that $A \in L(H, V)$ and $A = (e_n) = \sum_{m\geq 1} \lambda_{nm}\xi_m$ for all $n \in \mathbb{N}$. Show that $\sum_{m\geq 1} |\lambda_{nm}|^2 \leq \|A\|_L^2$ for all $n \in \mathbb{N}$ and $\sum_{n\geq 1} |\lambda_{nm}|^2 \leq \|A\|_L$ for all $m \in \mathbb{N}$.

Problem 3.33. Let H be a Hilbert space and let $A \in L(H)$ be a self-adjoint positive operator. Show that the following statements are equivalent:

(a) $R(A) \subseteq H$ is dense.
(b) $N(A) = \{0\}$.
(c) $(A(x), x) > 0$ for all $x \neq 0$.

Problem 3.34. Let H be a Hilbert space and let $A, T\colon H \to H$ be two linear operators such that $(A(x), u) = (x, T(u))$ for all $x, u \in H$. Show that $A \in L(H)$ and $T = A^*$.

Problem 3.35. Let H be a Hilbert space and let $\{A_n\}_{n\geq 1} \subseteq L(H)$ be such that $\sup_{n\geq 1} |(A_n(x), u)| < \infty$ for all $x, u \in H$. Show that $\sup_{n\geq 1} \|A_n\|_L < \infty$.

Problem 3.36. Let H be a Hilbert space and let $\{A_n\}_{n\geq 1} \subseteq L(H)$ be such that $\lim_{n\to\infty} |(A_n(x), u)| = 0$ for all $x, u \in H$. Can we say that $\|A_n\|_L \to 0$? Justify your answer.

Problem 3.37. Let K, D be compact spaces, let $g \in C(K, D)$, and let $A\colon C(K) \to C(D)$ be the operator defined by $A(f)(t) = f(g(s))$ for all $s \in K$ and for all $t \in D$. Show that
(a) $A \in L(C(K), C(D))$ and find $\|A\|_L$.
(b) $R(A) = C(D)$ if and only if g is injective.
(c) A is an isometry if and only if g is surjective.

Problem 3.38. Let X be a Banach space, let V be a normed space, and let $A \in L(X, V)$. Show that: $A^{-1} \in L(V, X)$ if and only if $R(A) \subseteq V$ is dense and $\|A(x)\|_V \geq c\|x\|_X$ for all $x \in X$ and for some $c > 0$.

Problem 3.39. Let H be a Hilbert space and let $A \in L(H)$ be normal. Show that $\lim_{n\to\infty} \|A^n\|_L^{1/n} = \|A\|_L$.

Problem 3.40. Let H be a Hilbert space and let $P \in L(H)$ be a projection, that is, $P^2 = P$. Show that the following properties are equivalent:
(a) P is an orthogonal projection.
(b) P is normal.
(c) $(P(x), x) = \|P(x)\|^2$ for all $x \in H$.

Problem 3.41. Let X, Y be Banach spaces with X being reflexive, $A \in L_c(X, Y)$, $\|\cdot\|_X$ being the norm of X, and $|\cdot|_X$ being another norm on X, which generates a weaker topology on X. Show that for every $\varepsilon > 0$ there exists $c_\varepsilon > 0$ such that

$$\|A(x)\|_Y \leq \varepsilon\|x\|_X + c_\varepsilon|x|_X \quad \text{for all } x \in X\,.$$

Problem 3.42. Let X be a normed space and let $P \in L_c(X)$ be a projection, that is, $P^2 = P$. Show that $P \in L_f(X)$.

Problem 3.43. Let H be a Hilbert space and let $A \in L(H)$ be self-adjoint. Assume that $A \geq \vartheta i_H$ for some $\vartheta > 0$; see Definition 3.8.22. Show that A is invertible.

Problem 3.44. Let H be a Hilbert space and let $A \in L(H)$ be self-adjoint. Show that the residual spectrum $R\sigma(A)$ of A (see Definition 3.7.29) is empty.

Problem 3.45. Let X be a Banach space and let $A \in L(X)$ and $\lambda \in \mathbb{C}$. Suppose that there exists a sequence $\{x_n\}_{n\geq 1} \subseteq X$ with $\|x_n\| = 1$ for all $n \in \mathbb{N}$ such that $A(x_n) - \lambda x_n \to 0$ in X. Show that $\lambda \in \sigma(A)$.

Problem 3.46. Let H be a Hilbert space and let $P \in L(H)$ be an orthogonal projection. Show that $0 \leq P \leq i_H$; see Definition 3.8.22.

Problem 3.47. Let H be a Hilbert space and let $A \in L(H)$ be such that $(A(x), x) \geq c\|x\|^2$ for all $x \in H$ and for some $c > 0$. Show that A is an isomorphism.

Problem 3.48. Let X be an infinite dimensional Banach space and let $A \in L_c(X)$. Show that there exists $h \in X$ such that there is no $x \in X$ for which we have $A(x) = h$.

Problem 3.49. Let X be a Banach space and let $A\colon D(A) \subseteq X \to X$ be an unbounded linear operator. Suppose there exists $\lambda \in \mathbb{C}$ such that $(A - \lambda I)^{-1} \in L(X)$. Show that A is closed.

Problem 3.50. Let X, Y be Banach spaces and let $A\colon D(A) \subseteq X \to Y$ be an unbounded linear operator such that $\|A(x)\|_Y \geq c\|x\|_X$ for all $x \in D(A)$ and for some $c > 0$. Show that A is closed.

Problem 3.51. Let H be a Hilbert space, let $\{u_n\}_{n\geq 1} \subseteq H$ be an orthonormal set and let $A \in L_c(H)$. Show that $A(u_n) \to 0$ in H.

Problem 3.52. Let X, Y be Banach spaces, let $A\colon X \to Y$ be a linear operator, and suppose that for every $y^* \in Y^*$ one has $y^* \circ A \in X^*$. Show that $A \in L(X, Y)$.

Problem 3.53. Let X be a Banach space and let $P \in L(X)$ be a projection, that is, $P^2 = P$. Show that P^* is a projection in X^*.

Problem 3.54. Let H be a Hilbert space and let $A \in L_c(H)$. Show that there exists $x \in H$ with $\|x\| \leq 1$ such that $\|A(x)\| = \|A\|_L$.

Problem 3.55. Let X, Y be Banach spaces with $Y \neq 0$. Show that X is reflexive if and only if for every $A \in L_c(X, Y)$ there exists $x \in X$ with $\|x\|_X \leq 1$ such that $\|A(x)\|_Y = \|A\|_L$.

Problem 3.56. Let X, Y be Banach spaces and let $A \in L_c(X, Y)$. Show that $R(A) \subseteq Y$ is separable.

Problem 3.57. Let X, Y be Banach spaces and let $A \in L(X, Y)$, which satisfies $\|A(x)\|_Y \geq c\|x\|_X$ for all $x \in X$ and for some $c > 0$. Is it possible for A to be compact? Justify your answer.

Problem 3.58. Let X be an infinite dimensional Banach space and let $A \in L_c(X)$. Show that $0 \in \overline{A(\partial B_1)}$.

Problem 3.59. Let X be a Banach space that is w-separable. Show that X is separable.

Problem 3.60. Let X be an infinite dimensional Banach space and let $K \subseteq X$ be a nonempty, compact set. Show that $\operatorname{int} K = \emptyset$.

Problem 3.61. Let X be a Banach space and assume that there exists an uncountable family $\{U_i\}_{i \in I}$ such that
(a) for each $i \in I$, $U_i \subseteq X$ is nonempty and open;
(b) $U_i \cap U_j = \emptyset$ if $i \neq j$.
Show that X is nonseparable.