Chapter 18
The Krein–Milman Theorem and Its Applications

18.1 Extreme Points of Convex Sets

As we remarked earlier, one of the main merits of the functional analysis-based approach to problems of classical analysis is that it reduces problems formulated analytically to problems of a geometric character. The geometric objects that arise in this way lie in infinite-dimensional spaces, but they can be manipulated by using analogies with figures in the plane or in three-dimensional space. To exploit this analogy more freely, to understand when it helps, rather than mislead us, we have studied above many properties of spaces, subspaces, convex sets, compact and weakly compact sets, linear operators, emphasizing each time the coincidences and differences with the finite-dimensional versions of those objects and properties. In the present chapter we add to the already built arsenal of geometric tools yet another one: the study of convex sets by means of their extreme points. Although extreme points are a direct generalization of the vertices of a polygon or polyhedron, in the framework of classical geometry this purely geometric concept was not used for general figures. The study and application of extreme points to problems of geometry (including finite-dimensional ones), functional analysis, mathematical economics, is one of the achievements of the bygone 20th century.

18.1.1 *Definitions and Examples*

Let A be a convex subset of a linear space X. A point $x \in A$ is called an *extreme point* of the set A if it is not the midpoint of any non-degenerate segment whose endpoints lie in A. The set of extreme points of the set A is denoted by ext A. That is, $x \in$ ext A if and only if for any $x_1, x_2 \in A$, if $(x_1 + x_2)/2 = x$, then $x_1 = x_2$ (and hence both vectors x_1 and x_2 coincide with x).

Theorem 1. *Let A be a convex subset of a topological vector space X. Then none of the interior points of the set A is an extreme point of A.*

Proof. Let $x \in A$ be an interior point. Then there exists a balanced neighborhood U of zero such that $x + U \subset A$. Let $y \in U \setminus \{0\}$. Set $x_1 = x + y$, $x_2 = x - y$. Then $x_1, x_2 \in A$, $(x_1 + x_2)/2 = x$, but $x_1 \neq x_2$.   □

Thus, all extreme points of a set lie on its boundary. Needless to say, this provides rather incomplete information on the positioning of extreme points. Note that ext A depends only on the convex geometry of the set A, but does not depend on the ambient linear space in which A is considered, or on which topology is given on A.

Theorem 2 (examples).

(a) *If A is a convex polygon in the plane, then* ext A *is the set of vertices of A.*

(b) *If A is a disc, then* ext A *is its bounding circle.*

(c) *The set of extreme points of the closed unit ball $\overline{B}_H$ of the Hilbert space H is the unit sphere S_H.*

(d) *The closed unit ball $\overline{B}_{c_0}$ of the space c_0 has no extreme points.*

Proof. Assertions (a) and (b) are obvious. Let us address (c). By Theorem 1, ext $\overline{B}_H \subset S_H$. Let us establish the opposite inclusion. Let $x, x_1, x_2 \in S_H$ and put $(x_1 + x_2)/2 = x$. This means that $\|x_1\| = \|x_2\| = 1$ and $\|x_1 + x_2\| = 2$. But then, by the parallelogram equality,

$$\|x_1 - x_2\|^2 = \|x_1 + x_2\|^2 - 2\|x_1\|^2 - 2\|x_2\|^2 = 4 - 2 - 2 = 0,$$

and so $x_1 = x_2$.

(d) Let us show that no point of $\overline{B}_{c_0}$ is an extreme point. Let $a = (a_1, a_2, \ldots) \in \overline{B}_{c_0}$. This means that all the coordinates a_j are not larger in modulus than 1 and $a_j \to 0$ as $j \to \infty$. From the last fact it follows that there exists an n such that $|a_n| < 1/2$. Consider the following vectors x_1 and x_2, all coordinates of which coincide with those of a, except for the nth one, where they differ from a by $\pm 1/2$:

$$x_1 = (a_1, a_2, \ldots, a_{n-1}, a_n + 1/2, a_{n+1}, \ldots)$$

and

$$x_2 = (a_1, a_2, \ldots, a_{n-1}, a_n - 1/2, a_{n+1}, \ldots).$$

Then x_1 and x_2 lie in $\overline{B}_{c_0}$, $(x_1 + x_2)/2 = a$, but $x_1 \neq x_2$.   □

The extreme points of a Cartesian product of convex sets admit a simple description.

Theorem 3. *Let Γ be an index set, X_γ, $\gamma \in \Gamma$, be linear spaces, and $A_\gamma \subset X_\gamma$ be convex sets. Then* ext $\left(\prod_{\gamma \in \Gamma} A_\gamma \right) = \prod_{\gamma \in \Gamma}$ ext A_γ.

Proof. Let $x = (x_\gamma)_{\gamma \in \Gamma} \in \prod_{\gamma \in \Gamma}$ ext A_γ, i.e., $x_\gamma \in$ ext A_γ for all $\gamma \in \Gamma$. Let us show that $x \in$ ext $\left(\prod_{\gamma \in \Gamma} A_\gamma \right)$. Consider elements $y = (y_\gamma)_{\gamma \in \Gamma}$ and $z = (z_\gamma)_{\gamma \in \Gamma}$ in

$\prod_{\gamma \in \Gamma} A_\gamma$ such that $(y + z)/2 = x$. Then $(y_\gamma + z_\gamma)/2 = x_\gamma$ and $y_\gamma, z_\gamma \in A_\gamma$. Since $x_\gamma \in \text{ext } A_\gamma$, it follows that $y_\gamma = z_\gamma$ for all $\gamma \in \Gamma$, and so $y = z$. This establishes the inclusion $\text{ext}\left(\prod_{\gamma \in \Gamma} A_\gamma\right) \supset \prod_{\gamma \in \Gamma} \text{ext } A_\gamma$.

Now let us establish the opposite inclusion $\text{ext}\left(\prod_{\gamma \in \Gamma} A_\gamma\right) \subset \prod_{\gamma \in \Gamma} \text{ext } A_\gamma$. Let $x = \left(x_\gamma\right)_{\gamma \in \Gamma} \in \left(\prod_{\gamma \in \Gamma} A_\gamma\right) \setminus \left(\prod_{\gamma \in \Gamma} \text{ext } A_\gamma\right)$. Then there is an index $\gamma_0 \in \Gamma$ for which $x_{\gamma_0} \in A_{\gamma_0} \setminus \text{ext } A_{\gamma_0}$. By definition, this means that there exists elements $y_{\gamma_0}, z_{\gamma_0} \in A_{\gamma_0}$, $y_{\gamma_0} \neq z_{\gamma_0}$, such that $(y_{\gamma_0} + z_{\gamma_0})/2 = x_{\gamma_0}$. Now define the elements $y, z \in \prod_{\gamma \in \Gamma} A_\gamma$ as follows: for $\gamma \neq \gamma_0$ put $y_\gamma = z_\gamma = x_\gamma$, while for the index γ_0 take as coordinates precisely the elements y_{γ_0} and z_{γ_0}, respectively. Then $y \neq z$ (they differ in the γ_0 coordinate), but $(y + z)/2 = x$. Therefore, $x \notin \text{ext}\left(\prod_{\gamma \in \Gamma} A_\gamma\right)$. $\square$

An obvious consequence of this theorem is the following descriptions of the extreme points of two important n-dimensional bodies.

Corollary 1. *The extreme points of the n-dimensional cube $[-1, 1]^n$ are precisely the vectors with all coordinates equal to ± 1.* $\square$

Let us recall the notations $\mathbb{C}_1 = \{\lambda \in \mathbb{C} : |\lambda| \leqslant 1\}$ and $\mathbb{T} = \{\lambda \in \mathbb{C} : |\lambda| = 1\}$ for the unit disc and the unit circle.

Corollary 2. *The set of extreme points of the n-dimensional polydisc $(\mathbb{C}_1)^n$ is the skeleton of the polydisc, i.e., the set $\mathbb{T}^n$.* $\square$

Exercises

1. In the real space $C[0, 1]$ the closed unit ball has only two extreme points, the functions $f = 1$ and $g = -1$.

2. In the space $L_1[0, 1]$ the closed unit ball has no extreme points.

3. For $1 < p < \infty$ every element of the unit sphere in the space $L_p[0, 1]$ is an extreme point of the closed unit ball. In other words, $L_p[0, 1]$ is a strictly convex space (see Exercise 6 in Subsection 12.2.1).

4. Using the preceding exercise and the reflexivity, prove the following result: Let $1 < p < \infty$, and let $A \subset L_p[0, 1]$ be a convex closed subset. Then for any $x \in X$ in A there is a unique closest point to x.

5. Let X, Y be linear spaces and $T : X \to Y$ be an injective linear operator. Then for any convex subset $A \subset X$ it holds that $T(\text{ext } A) = \text{ext } T(A)$.

6. Give an example showing that in the preceding exercise the injectivity assumption cannot be discarded.

7. For any convex compact subset $A \subset \mathbb{R}^2$, the set ext A is closed.

8. Give an example of a convex compact subset $A \subset \mathbb{R}^3$ whose set of extreme points is not closed.

18.1.2   The Krein–Milman Theorem

In this subsection we shall prove the main result of this chapter, namely, the existence of extreme points for any convex compact set.

Definition 1. Let A be a convex subset of a linear space X. The set $B \subset A$ is said to be an *extreme subset of the set A* if it meets the following requirements:

— B is not empty;

— B is convex;

— for any two points $x_1, x_2 \in A$, if $(x_1 + x_2)/2 \in B$, then $x_1, x_2 \in B$.

Obviously, a subset consisting of a single point x will be extreme if and only if x is an extreme point. If A is a triangle in the plane, then an example of an extreme subset B is provided by any side of the triangle.

Lemma 1. *Let X, Y be linear spaces, $T: X \to Y$ be a linear operator, and $A \subset X$ be a convex subset. Then for any extreme subset B of the set $T(A)$, the set $T^{-1}(B) \cap A$ (the complete preimage in A of the set B) is an extreme subset of the original set A. In particular, the complete preimage in A of any extreme point of the set $T(A)$ is an extreme subset of A.*

Proof. Suppose $x_1, x_2 \in A$ and $(x_1 + x_2)/2 \in T^{-1}(B)$. Then $Tx_1, Tx_2 \in T(A)$ and $(Tx_1 + Tx_2)/2 \in B$. Since B is an extreme subset of $T(A)$, this means that $Tx_1, Tx_2 \in B$, and so $x_1, x_2 \in T^{-1}(B)$.   □

Lemma 2. *Suppose A is a convex set, B is an extreme subset of A, and C is an extreme subset of B. Then C is an extreme subset of A. In particular, an extreme point of an extreme subset of a set A is an extreme point of A.*

Proof. Suppose $x_1, x_2 \in A$ and $(x_1 + x_2)/2 \in C$. Then, in particular, $(x_1 + x_2)/2 \in B$. Since B is an extreme subset of A, this implies that $x_1, x_2 \in B$. But now recalling that $(x_1 + x_2)/2 \in C$ and C is an extreme subset of B, we conclude that $x_1, x_2 \in C$, as needed.   □

Now let us change the setting from arbitrary linear spaces to locally convex topological vector spaces, and from arbitrary convex sets to convex compact sets.

Lemma 3. *Let A be a convex compact set in a topological vector space X, f a continuous real linear functional on X, and $b = \max_{x \in A} f(x)$. Then the set $M(f, A) = \{x \in A : f(x) = b\}$ of points x in which f attains its maximum on A is an extreme subset of A.*

Proof. The set $f(A)$ is an interval $[a, b]$ joining the minimal and maximal values on A of the functional f. Hence, b is an extreme point of the set $f(A)$. By Lemma 1, $M(f, A) = f^{-1}(b) \cap A$ is an extreme subset. □

Lemma 4. *Let A be a convex compact subset of a topological vector space X and $\mathfrak{M}$ be a centered family of closed extreme subsets of A. Then the intersection $D = \bigcap_{B \in \mathfrak{M}} B$ of all the elements of the family $\mathfrak{M}$ is also a closed extreme subset of A.*

Proof. The compactness of A ensures that the set D is not empty. Convexity and closedness are inherited by intersections of sets, so D is convex and closed. Now let $x_1, x_2 \in A$ and $(x_1 + x_2)/2 \in D$. Then, in particular, $(x_1 + x_2)/2 \in B$ for any $B \in \mathfrak{M}$. Therefore, $x_1, x_2 \in B$ for all $B \in \mathfrak{M}$, whence $x_1, x_2 \in \bigcap_{B \in \mathfrak{M}} B = D$. □

Lemma 5. *Let A be a convex compact subset of a separated locally convex topological vector space, and suppose A consists of more than one point. Then A contains a closed extreme subset B such that $B \neq A$.*

Proof. Suppose $x_1, x_2 \in A$ and $x_1 \neq x_2$. Since the dual of a separated locally convex space separates points, there exists a real continuous linear functional f such that $f(x_1) \neq f(x_2)$. Hence, f is not identically constant on A, and for the required set B we can take the set $M(f, A)$ from Lemma 3. □

Theorem 1 (weak formulation of the Krein–Milman theorem).[1] *Every convex compact set K in a separated locally convex space has extreme points.*

Proof. Consider the family $\mathfrak{Ext}(K)$ of all closed extreme subsets of the compact set K. We equip $\mathfrak{Ext}(K)$ with the decreasing order of sets. By Lemma 4, $\mathfrak{Ext}(K)$ is an inductively ordered set. By Zorn's lemma, there exists a minimal with respect to inclusion closed extreme subset A of the compact set K. By Lemma 5, A consists of exactly one point, which is the sought-for extreme point of K. □

Remark 1. For a convex compact set in a non-locally convex separated topological vector space the assertion of Theorem 1 may fail. A corresponding counterexample was constructed by Roberts [73].

The next result has numerous applications in linear optimization problems and, in particular, in problems of mathematical economics.

Theorem 2. *Let K be a convex compact set in a separated locally convex space X, f a continuous real linear functional on X, and $b = \max_{x \in K} f(x)$. Then there exists a point $x \in \text{ext } K$ in which $f(x) = b$. In other words, when searching for the maximum of a linear functional on a convex compact set, it suffices to consider the values in the extreme points of the compact set under consideration.*

---

[1]Mark Krein and David Milman were Odessa mathematicians. For this reason, in contrast to the theorems of the Lviv school led by Stefan Banach, which became "Ukrainian" only as a result of post-war geopolitical changes, a Ukrainian patriot like me can be proud that the Krein–Milman theorem is "genuinely Ukrainian".

Proof. By Lemma 3, $M(f, K) = \{x \in K : f(x) = b\}$ is an extreme subset of the compact set K, and is closed thanks to the continuity of the functional f. Since the set $M(f, K)$ is convex and compact, it has an extreme point x_0, and in view of the definition of $M(f, K)$, $f(x_0) = b$. It remains to apply Lemma 2: an extreme point of an extreme subset is an extreme point of the original set. □

The application of Theorem 2 becomes especially effective in the case when the set K is a finite-dimensional polyhedron. In this case ext K is a finite set, and the task of calculating the maximum of a linear functional reduces to a finite (admittedly possibly large) item-by-item examination. This examination can be carried out, in particular, by means of the famous *simplex method* of Kantorovich, which these days is presented in every linear programming textbook.

Lemma 6. *Let A and B be convex closed subsets of a locally convex space X. Then the following conditions are equivalent:*

(i) $A = B$;

(ii) $\sup_{x \in A} f(x) = \sup_{x \in B} f(x)$ *for any real linear functional $f \in X^*$.*

Proof. The implication (i) $\Longrightarrow$ (ii) is obvious. Let us prove the converse implication (ii) $\Longrightarrow$ (i). Since the sets A and B play symmetric roles, it suffices to prove that $A \subset B$. Suppose this inclusions does not hold. Then there exists a point $x_0 \in A \setminus B$. Since B is closed, x_0 has a neighborhood U such that $U \cap B = \emptyset$. By the geometric form of the Hahn–Banach theorem, applied to the sets U and B, there exist a continuous real linear functional on X and a constant $a \in \mathbb{R}$ such that $f(x) \leqslant a$ for $x \in B$ and $f(x_0) > a$. Then $\sup_{x \in A} f(x) \geqslant f(x_0) > a \geqslant \sup_{x \in B} f(x)$, which contradicts condition (ii). □

Theorem 3 (Krein–Milman theorem: complete formulation). *Any convex compact set K in a separated locally convex space coincides with the closure of the convex hull of its extreme points.*

Proof. Let $\widetilde{K} = \overline{\text{conv}}\,(\text{ext}\,K)$ and consider an arbitrary continuous real linear functional f on X. Then $\widetilde{K} \subset K$, and so $\sup_{x \in K} f(x) \geqslant \sup_{x \in \widetilde{K}} f(x)$. By Theorem 2, $\sup_{x \in K} f(x) \leqslant \sup_{x \in \text{ext}\,K} f(x) \leqslant \sup_{x \in \widetilde{K}} f(x)$. It remains to apply Lemma 6. □

Thus, we can state that a convex compact set not only has extreme points, but there are "many" such points. For example, if the compact set K is infinite-dimensional, then the set ext K is infinite. Let us give several corollaries.

Corollary 1. *Every convex closed bounded subset of a reflexive space and, in particular, the closed unit ball, has extreme points. If the space is infinite-dimensional, then its closed unit ball has infinitely many extreme points.*

Proof. It suffices to recall that any convex closed bounded subset of a reflexive space is weakly compact. □

This provides another proof of the non-reflexivity of the spaces c_0, $L_1[0, 1]$, and $C[0, 1]$: as we already remarked, in the first two of these spaces the unit ball even has no extreme points, while the unit ball in $C[0, 1]$ has only two extreme points.

If instead of the weak topology we consider the w^*-topology, we obtain another corollary.

Corollary 2. *Let X be a Banach space. Then any convex w^*-closed bounded subset of the space X^* and, in particular, the closed unit ball in X^*, has extreme points. If the space X is infinite-dimensional, then the closed unit ball in the space X^* has infinitely many extreme points.* $\square$

For this reason, the spaces c_0, $L_1[0, 1]$ and $C[0, 1]$ are not just non-reflexive, they actually are not dual to any Banach space (i.e., they are not isometric to any space of the form X^* with X a Banach space).

Exercises

1. Let $A = \text{conv } B$. Then ext $A \subset B$.

2. Let $A \subset B$ and $x \in (\text{ext } B) \cap A$. Then $x \in \text{ext } A$.

3. Let K be a convex compact subset of a strictly convex Banach space. Then a farthest from zero point of the compact set K is an extreme point of K.

4. Let X, Y be Banach spaces, $T \in L(X, Y)$, and K a convex compact set in X. Suppose that $\|Tx\| \leqslant C$ for all $x \in \text{ext } K$. Then $\|Tx\| \leqslant C$ for all $x \in K$.

Using the preceding exercise and the description of the extreme points of the n-dimensional cube, prove the following result:

5. Suppose $x_1, \ldots, x_n$ are elements of the Banach space X and the estimate $\left\|\sum_{k=1}^{n} a_k x_k\right\| \leqslant C$ holds for all $a_k = \pm 1$. Then the same estimate holds for all $a_k \in [-1, 1]$.

6. **Lindenstrauss–Phelps theorem.** *In an infinite-dimensional reflexive Banach space the set of extreme points of the closed unit ball is uncountable.*

The closed unit ball of the space c_0, regarded as a subset of the space $\ell_\infty = \ell_1^*$, is an example of a closed convex and bounded set in a dual space which has no extreme points. Therefore, in Corollary 2 the w^*-closedness assumption cannot be replaced by ordinary closedness. This makes the next result all the more interesting:

7. Let X be a Banach space whose dual X^* is separable. Then any convex closed (in norm) bounded subset of the space X^* has extreme points.

8. None of the spaces c_0, $L_1[0, 1]$ and $C[0, 1]$ can be isomorphically embedded in a separable dual space. In particular, none of these spaces is isomorphic to a dual space.

9. The set of extreme points of a convex metrizable compact subset of a locally convex space is a G_δ-set.

10. For every Banach space X, the identity operator $I \in L(X)$ is an extreme point of the ball $\overline{B}_{L(X)}$.

11. The unit element of any Banach algebra **A** is an extreme point of the ball $\overline{B}_\mathbf{A}$.

18.1.3   Weak Integrals and the Krein–Milman Theorem in Integral Form

Let (Ω, Σ, μ) be a finite measure space and X be a locally convex space. A function $f : \Omega \to X$ is said to be *weakly integrable* if for any $x^* \in X^*$ the composition $x^* \circ f$ is an integrable scalar function and there exists an element $x \in X$ such that

$$\int_\Omega x^* \circ f \, d\mu = x^*(x) \tag{1}$$

for all $x^* \in X^*$. In this case the element x is called the *weak integral* of the function f, and is denoted by the symbol $\int_\Omega f \, d\mu$. With this notation formula (1) takes on the form

$$x^* \left(\int_\Omega f \, d\mu \right) = \int_\Omega x^* \circ f \, d\mu$$

and can be interpreted as saying that a continuous linear functional can be brought under the integral sign.

The weak integral inherits the simplest properties of the ordinary integral:

— $\int_\Omega af_1 + bf_2 \, d\mu = a \int_\Omega f_1 \, d\mu + b \int_\Omega f_2 \, d\mu$ (linearity with respect to the function);

— $\int_\Omega f \, d(a\mu_1 + b\mu_2) = a \int_\Omega f \, d\mu_1 + b \int_\Omega f \, d\mu_2$ (linearity with respect to the measure);

— $\int_{\Omega_1 \sqcup \Omega_2} f d\mu = \int_{\Omega_1} f \, d\mu + \int_{\Omega_2} f \, d\mu$ for any disjoint sets $\Omega_1, \Omega_2 \in \Sigma$ (additivity with respect to the integration domain);

Here, in all the three properties, if the integrals on the right-hand side exist, then so does the integral on the left-hand side.

Let us mention that one of the important properties of the Lebesgue integral, namely, that integrability on a set implies integrability on all its measurable subsets, does not hold for the weak integral (see Exercise 1 below). The root of this unpleasant feature is that the weak topology of a space is not necessarily complete. Let us

examine on some examples how one calculates the weak integral of a vector-valued function.

Example 1. Let X be the sequence space ℓ_p or c_0, and $e_n^* \in X^*$ be the coordinate (evaluation) functionals. Let $f: \Omega \to X$ be a weakly integrable function. For each $t \in \Omega$ denote by $f_n(t)$ the n-th component of the vector $f(t)$: $f(t) = (f_1(t), f_2(t), \ldots)$. Then, by the definition of the weak integral,

$$e_n^*\left(\int_\Omega f \, d\mu\right) = \int_\Omega e_n^* \circ f \, d\mu = \int_\Omega f_n \, d\mu,$$

i.e., $\int_\Omega f \, d\mu$ is the vector with the components $\left(\int_\Omega f_n \, d\mu\right)_{n=1}^\infty$.

Example 2. Let $F: \Omega \to C[0, 1]$ be a weakly integrable function. For each $t \in [0, 1]$ and each $\tau \in \Omega$, we define $f(t, \tau) = (F(\tau))(t)$. Using instead of the coordinate functionals the point evaluation functionals, we obtain the following rule for the calculation of the function $\int_\Omega F \, d\mu \in C[0, 1]$:

$$\left(\int_\Omega F \, d\mu\right)(t) = \int_\Omega f(t, \tau) \, d\mu(\tau).$$

As in the scalar case, a function $f: \Omega \to X$ is said to be *measurable* if $f^{-1}(A) \in \Sigma$ for any Borel subset A of the space X. Let us mention one useful sufficient condition for weak integrability.

Theorem 1. *Let (Ω, Σ, μ) be a probability space, K a convex compact subset of a separated locally convex space X, and $f: \Omega \to K$ a measurable function. Then the function f is weakly integrable and $\int_\Omega f \, d\mu \in K$.*

Proof. Consider the dual pair $((X^*)', X^*)$. Since $K \subset X \subset (X^*)'$, the compact set K can be regarded as a subset of the space $(X^*)'$. Then K will also be compact in the weaker topology $\sigma((X^*)', X^*)$, and so K is a convex $\sigma((X^*)', X^*)$-compact set in $(X^*)'$.

Next, we remark that every functional $x^* \in X^*$ is bounded on K. Consequently, the composition $x^* \circ f$ is a bounded measurable function on Ω, and hence $x^* \circ f$ is an integrable scalar function. Define a linear functional $F: X^* \to \mathbb{C}$ by the formula $F(x^*) = \int_\Omega x^* \circ f \, d\mu$. We claim that $F \in K$. Assuming the contrary, there exists an element $x^* \in X^*$ such that $\operatorname{Re} x^*(s) \leqslant 1$ for all $s \in K$ and $\operatorname{Re} x^*(F) > 1$. Then $\operatorname{Re} x^* \circ f \leqslant 1$ everywhere on Ω, and so

$$\operatorname{Re} x^*(F) = \operatorname{Re} F(x^*) = \int_\Omega \operatorname{Re} x^* \circ f \, d\mu \leqslant 1.$$

The contradiction we reached means that $F \in K$. By construction, $x^*(F) = \int_\Omega x^* \circ f \, d\mu$ for all $x^* \in X^*$, i.e., F is the weak integral of the function f. $\square$

Theorem 2 (Krein–Milman theorem: integral form). *Let K be a convex compact subset of a separated locally convex space X and $x \in K$. Then there exists a regular Borel probability measure μ on $\overline{\mathrm{ext}}\, K$, the closure of the set of extreme points of K, such that*

$$\int_{\overline{\mathrm{ext}}\, K} I \, d\mu = x;$$

here I denotes, as usual, the identity mapping and the integral is understood in the weak sense.

Proof. Recalling the theorem on the general form of linear functionals on the space of continuous functions, we see that the set $\mathrm{M}(\overline{\mathrm{ext}}\, K)$ of all regular Borel probability measures on the compact set $\overline{\mathrm{ext}}\, K$ can be regarded as a subset of the space $C(\overline{\mathrm{ext}}\, K)^*$. Moreover, $\mathrm{M}(\overline{\mathrm{ext}}\, K)$ is the intersection of the closed unit ball of the space $C(\overline{\mathrm{ext}}\, K)^*$ (i.e., of a convex w^*-compact set) with the w^*-closed set $\{F \in C(\overline{\mathrm{ext}}\, K)^* : F(\mathbb{1}) = 1\}$. As such, $\mathrm{M}(\overline{\mathrm{ext}}\, K)$ is a convex w^*-compact subset of $C(\overline{\mathrm{ext}}\, K)^*$.

By the preceding theorem, for each measure $\mu \in \mathrm{M}(\overline{\mathrm{ext}}\, K)$ there exists the weak integral $\int_{\overline{\mathrm{ext}}\, K} I \, d\mu$. Consider the operator $T \colon X^* \to C(\overline{\mathrm{ext}}\, K)^*$ which sends each functional $x^* \in X^*$ into its restriction to $\overline{\mathrm{ext}}\, K$. Let us calculate the action of the adjoint operator $T^* \colon C(\overline{\mathrm{ext}}\, K)^* \to X^{**}$ on the elements of the set $\mathrm{M}(\overline{\mathrm{ext}}\, K)$. For any measure $\mu \in \mathrm{M}(\overline{\mathrm{ext}}\, K)$ and any $x^* \in X^*$ we have

$$\langle T^*\mu, x^* \rangle = \langle \mu, Tx^* \rangle = \int_{\overline{\mathrm{ext}}\, K} x^* \, d\mu = x^*\left(\int_{\overline{\mathrm{ext}}\, K} I \, d\mu \right),$$

i.e.,

$$T^*\mu = \int_{\overline{\mathrm{ext}}\, K} I \, d\mu.$$

Therefore, now our task is reduced to proving the equality

$$T^*(\mathrm{M}(\overline{\mathrm{ext}}\, K)) = K.$$

The inclusion $T^*(\mathrm{M}(\overline{\mathrm{ext}}\, K)) \subset K$ was proved in the preceding theorem. Let us prove the opposite inclusion $T^*\big(\mathrm{M}(\overline{\mathrm{ext}}\, K)\big) \supset K$. Denote by δ_x the probability measure supported at the point x. Then for any $x \in \mathrm{ext}\, K$ we have

$$T^*\delta_x = \int_{\overline{\mathrm{ext}}\, K} I \, d\delta_x = x,$$

i.e., $T^*(\mathrm{M}(\overline{\mathrm{ext}}\, K)) \supset \mathrm{ext}\, K$. Further, $T^*\big(\mathrm{M}(\overline{\mathrm{ext}}\, K)\big)$ is a convex closed set, being the image of the convex w^*-compact set $\mathrm{M}(\overline{\mathrm{ext}}\, K)$ under the w^*-continuous operator T^*.

Therefore, $T^*(\mathrm{M}(\overline{\mathrm{ext}\,K})) \supset \overline{\mathrm{conv}}(\mathrm{ext}\,K)$; but, by Theorem 3 of Subsection 18.1.2, $\overline{\mathrm{conv}}(\mathrm{ext}\,K) = K$.   □

Let us remark that in the metrizable case the measure μ that represents the element x can be selected so that it is supported on the set of extreme points itself, rather than on its closure. Then the integral representation of the element x takes the form $x = \int_{\mathrm{ext}\,K} I\,d\mu$. The proof of this theorem due to G. Choquet and various generalizations thereof can be found in Phelps' monograph [34].

Another, quite fruitful research direction is connected with the consideration of a narrower set than ext K, namely, the set of strongly exposed points. In terms of such points it was possible to characterize the spaces in which the Radon–Nikodým theorem is valid. Results on this theme, as well as many other interesting branches of the geometry of Banach spaces, can be found in the monographs [3, 12].

Exercises

1. On the interval $[0, 1]$ pick a sequence (Δ_n), $n \in \mathbb{N}$, of pairwise disjoint subintervals. Let e_n, $n \in \mathbb{N}$, denote the standard basis of the space c_0. Define the function $f: [0, 1] \to c_0$ as follows: if the point t lies in none of the intervals Δ_n, put $f(t) = 0$; if t lies in an odd-indexed interval Δ_{2n-1}, put $f(t) = (1/|\Delta_{2n-1}|)e_n$; finally, if t lies in an even-indexed interval Δ_{2n}, put $f(t) = -(1/|\Delta_{2n}|)e_n$. Verify that the function f is weakly integrable on $[0, 1]$ with respect to the Lebesgue measure λ and $\int_{[0,1]} f\,d\lambda = 0$. At the same time, on the subset $\Delta = \bigcup_{n=1}^{\infty} \Delta_{2n-1}$ the function f is not weakly integrable: otherwise, we would have the equality $\int_{\Delta} f\,d\lambda = (1, 1, 1, \ldots)$, but there is no such element in c_0.

2. Prove the following theorem of Carathéodory: *If $K \subset \mathbb{R}^n$ is a convex compact set, then every element $x \in K$ admits a representation $x = \sum_{j=1}^{n+1} a_j x_j$, where $x_j \in$* ext K, $a_j \geqslant 0$, *and* $\sum_{j=1}^{n+1} a_j = 1$.

3. Based on Choquet's theorem formulated above, prove that if the convex metrizable compact set K in a locally convex space has a countable number of extreme points, then every element $x \in K$ admits a series expansion $x = \sum_{n=1}^{\infty} a_n x_n$, where $x_n \in$ ext K, $a_n \geqslant 0$, and $\sum_{n=1}^{\infty} a_n = 1$.

4. If one removes the requirement that the set of extreme points in countable, the assertion of the preceding exercise may fail. Provide a counterexample.

18.2 Applications

18.2.1 The Connection Between the Properties
of the Compact Space K and Those of the Space $C(K)$

The space $C(K)$ is more convenient to study than the compact space K, because the elements of a function space can be manipulated more freely than the points of a

topological space. Indeed, in contrast to the points of the compact space K, functions on K can be added and multiplied by scalars; the topology on $C(K)$ is given by a norm, and so one can speak about Cauchy sequences, completeness, convergence of series, and so on. However, all these advantages would depreciate if in the passage from K to $C(K)$ part of the information about the original compact space is lost. Below we will show that actually no such loss occurs and all the properties of the compact space K can be recovered from the properties of the space $C(K)$.

As usual, we will identify the continuous functionals on $C(K)$ with the regular Borel charges that generate them. In particular, δ_x (the probability measure supported at x) can be regarded as the functional on $C(K)$ which acts as $\langle \delta_x, f \rangle = \int_K f \, d\delta_x = f(x)$, i.e., as the evaluation functional at the point x.

A bit more terminology. The *support* of the regular Borel charge σ is defined to be the support of the measure $|\sigma|$ (see Subsection 8.1.2). As in the case of measures, the support of a charge σ is denoted by $\operatorname{supp} \sigma$. Clearly, $\operatorname{supp} \delta_x = \{x\}$, and if $\operatorname{supp} \sigma = \{x\}$, then $\sigma = \lambda \delta_x$, where λ is a non-zero scalar.

For any Borel-measurable bounded function g on K and any Borel charge σ, we denote by $g \times \sigma$ the Borel charge given by

$$(g \times \sigma)(A) = \int_A g \, d\sigma.$$

The functional defined by the charge $g \times \sigma$ acts by the rule

$$\langle g \times \sigma, f \rangle = \int_K fg \, d\sigma.$$

The operation thus introduced enjoys the natural properties of a product:

— $\mathbb{1} \times \sigma = \sigma$;
— $(g + h) \times \sigma = g \times \sigma + h \times \sigma$;
— $(gh) \times \sigma = (hg) \times \sigma = h \times (g \times \sigma)$;
— $g \times (v + \sigma) = g \times v + g \times \sigma$.
— Finally, the norm of the charge $g \times \sigma$ is calculated by the formula

$$\|g \times \sigma\| = \int_K |g| \, d|\sigma|.$$

Theorem 1. *The set of extreme points of the closed unit ball of the space $C(K)^*$ coincides with the set of charges of the form $\lambda \delta_x$, where $x \in K$ and $|\lambda| = 1$.*

Proof. First let us show that charges of the form δ_x are extreme points of the set $\overline{B}_{C(K)^*}$. Since the ball is balanced, this will imply that also $\lambda \delta_x \in \operatorname{ext} \overline{B}_{C(K)^*}$ whenever $|\lambda| = 1$. So, assume $v_1, v_2 \in \overline{B}_{C(K)^*}$ and $(v_1 + v_2)/2 = \delta_x$. Then $(v_1(\{x\}) + v_2(\{x\}))/2 = \delta_x(\{x\}) = 1$. Since both numbers $|v_1(\{x\})|$, $|v_2(\{x\})|$ are not larger than 1, this means that $v_1(\{x\}) = v_2(\{x\}) = 1$. This in turn means that beyond the

point x the charges ν_1 and ν_2 vanish, since otherwise their norms would be strictly larger than 1. That is, $\nu_1 = \nu_2 = \delta_x$.

Now let us show that if the charge $\sigma \in \overline{B}_{C(K)^*}$ is not concentrated at a single point of the compact space K, then it cannot be an extreme point of the unit ball. Indeed, suppose supp σ contains two distinct points $x \neq y$. Surround these points by disjoint neighborhoods U and V. By the definition of the support, the numbers $|\sigma|(U)$ and $|\sigma|(V)$ are different from zero. Let $\varepsilon = \min\{|\sigma|(U), |\sigma|(V)\}$. Consider the function

$$g = \frac{\varepsilon}{|\sigma|(U)} \mathbb{1}_U - \frac{\varepsilon}{|\sigma|(V)} \mathbb{1}_V$$

and the charges

$$\sigma_1 = (1 - g) \times \sigma \quad \text{and} \quad \sigma_2 = (1 + g) \times \sigma.$$

Since $|g| \leqslant 1$, we have $|1 \pm g| = 1 \pm g$. Further, by construction, $\int_K g\,d|\sigma| = 0$. Consequently,

$$\int_K |1 \pm g| d|\sigma| = \int_K d|\sigma| \pm \int_K g\,d|\sigma| = \int_K d|\sigma| = \|\sigma\| \leqslant 1.$$

Hence, $\sigma_1, \sigma_2 \in \overline{B}_{C(K)^*}$. At the same time,

$$\frac{\sigma_1 + \sigma_2}{2} = \sigma$$

and

$$\|\sigma_1 - \sigma_2\| = 2 \int_K |g| d|\sigma| = 4\varepsilon \neq 0;$$

therefore, the charge σ cannot be an extreme point of the unit ball. □

Suppose we are given a Banach space X and we are told that $X = C(K)$ for some compact space K, but not what this compact space is. Can we recover K from the space X? By the preceding theorem, to this end we need to look at the extreme points of the ball $\overline{B}_{X^*}$.

Let us introduce several definitions and notations. The set ext $\overline{B}_{X^*}$ will be regarded as a subspace of the topological space $(X^*, \sigma(X^*, X))$, i.e., we equip ext $\overline{B}_{X^*}$ with the w^*-topology. Further, we introduce on ext $\overline{B}_{X^*}$ the following equivalence relation: $x^* \sim y^*$ if $x^* = \lambda y^*$ for some scalar λ with $|\lambda| = 1$. The equivalence class of the element $x^* \in$ ext $\overline{B}_{X^*}$ is the pair of points $\pm x^*$ in the real case, and the circle passing through x^*, i.e., $[x^*] = \{\lambda x^* : |\lambda| = 1\}$, in the complex case. We denote the set of equivalence classes into which ext $\overline{B}_{X^*}$ decomposes by $\widetilde{K}(X)$, and denote by q the quotient mapping $q \colon$ ext $\overline{B}_{C(K)^*} \to \widetilde{K}(X)$. We equip $\widetilde{K}(X)$ with the strongest topology with respect to which q is w^*-continuous. That is to say, a set $A \subset \widetilde{K}(X)$ is declared to be open in $\widetilde{K}(X)$ if $q^{-1}(A)$ is w^*-open in ext $\overline{B}_{X^*}$. Note that $\widetilde{K}(X)$ is a Hausdorff topological space. Indeed, if $x^*, y^* \in$ ext $\overline{B}_{X^*}$ and $[x^*] \neq [y^*]$, then the functionals x^* and y^* are linearly independent. Hence, the kernel of any of them is not included in the kernel of the other, and so there exists an element $x \in$ Ker $y^* \setminus$ Ker x^*. Multiplying x by a scalar, one can ensure that $x^*(x) = 1$. Then the points $[x^*], [y^*] \in$

$\widetilde{K}(X)$ are separated by the neighborhoods $U = \{[s^*] \in \widetilde{K} : |s^*(x)| > 1/2\}$ and $V = \{[s^*] \in \widetilde{K} : |s^*(x)| < 1/2\}$.

Theorem 2. *Let $X = C(K)$ for some compact space K. Then K is homeomorphic with the topological space $\widetilde{K}(X)$ constructed above.*

Proof. Define the mapping $\delta \colon K \to \text{ext}\,\overline{B}_{X^*}$ by the formula $\delta(t) = \delta_t$. For any function $f \in C(K)$ we have $\langle \delta(t), f \rangle = f(t)$, which depends continuously on t. Since $\text{ext}\,\overline{B}_{X^*}$ is equipped with the w^*-topology, this means that the mapping δ is continuous. Then, as a composition of continuous mappings, the mapping $j \colon K \to \widetilde{K}(X)$, $j = q \circ \delta$, is also continuous. Since $j(t) = [\delta_t]$, Theorem 1 guarantees that the mapping j is bijective. But any bijective continuous mapping of a compact space onto a Hausdorff space is a homeomorphism. $\square$

Corollary 1. *If for two compact spaces K_1 and K_2 the spaces $C(K_1)$ and $C(K_2)$ are isometric, then the spaces K_1 and K_2 are homeomorphic.* $\square$

Theorem 3. *The space $C(K)$ is separable if and only if the compact space K is metrizable.*

Proof. Suppose $C(K)$ is separable. Then (Corollary 4 in Subsection 17.2.4) the w^*-topology is metrizable on the ball $\overline{B}_{C(K)^*}$. The compact space K is homeomorphic to the subset $\{\delta_t : t \in K\}$ of the ball $\overline{B}_{C(K)^*}$, equipped with the w^*-topology (the homeomorphism is provided by the mapping $t \mapsto \delta_t$). Hence, K is metrizable.

Conversely, suppose K is a compact metric space. Then for each $n \in \mathbb{N}$ there exists a cover of the compact space K by balls $U_{n,1}, U_{n,2}, \dots, U_{n,m(n)}$ of radius $1/n$. Denote by $\varphi_{n,1}, \varphi_{n,2}, \dots, \varphi_{n,m(n)}$ a partition of unity subordinate to the cover $U_{n,1}, U_{n,2}, \dots U_{n,m(n)}$ (see Subsection 15.1.3). Let us prove that the system of elements $\{\varphi_{n,j} : n = 1, \dots, \infty, \ j = 1, \dots, m(n)\}$ is complete in $C(K)$. This in turn will establish the desired separability of the space $C(K)$.

Thus, let $f \in C(K)$. For each $\varepsilon > 0$, pick an $n \in \mathbb{N}$ such that for any $t_1, t_2 \in K$, if $\rho(t_1, t_2) < 1/n$, then $|f(t_1) - f(t_2)| < \varepsilon$. Next, in each set $U_{n,j}$ pick one point $t_{n,j}$ and consider the following linear combination f_ε of the functions $\varphi_{n,j}$:

$$f_\varepsilon = f(t_{n,1})\varphi_{n,1} + f(t_{n,2})\varphi_{n,2} + \cdots + f(t_{n,m(n)})\varphi_{n,m(n)}.$$

We claim that $\|f - f_\varepsilon\| < \varepsilon$. Indeed, for any $t \in K$ we have $f(t) = \sum_{j=1}^{m(n)} f(t)\varphi_{n,j}(t)$. Consequently,

$$|f(t) - f_\varepsilon(t)| \leqslant \sum_{j=1}^{m(n)} |f(t) - f(t_{n,j})|\varphi_{n,j}(t).$$

In the last sum, if $\varphi_{n,j}(t) \neq 0$, then $t \in U_{n,j}$, and so $|f(t) - f(t_{n,j})| < \varepsilon$. Continuing the estimate, we conclude that

$$|f(t) - f_\varepsilon(t)| < \sum_{j=1}^{m(n)} \varepsilon \varphi_{n,j}(t) = \varepsilon. \qquad \square$$

Remark 1. Jumping ahead, we observe that in the last part of the proof of Theorem 3, the separability of the space $C(K)$ is an easy consequence of the Stone–Weierstrass theorem. But in our opinion the explicit procedure of approximating a function by a partition of unity is instructive in itself.

Exercises

1. Verify that for any Borel-measurable bounded function g on K and any regular Borel charge σ, the charge $g \times \sigma$ is regular (hint: consider first the case $g = \mathbb{1}_A$, then the case of finitely-valued functions, and finally use the approximation of a bounded function by finitely-valued functions).

2. Verify all the properties, listed at the beginning of the subsection, of the operation $g \times \sigma$ of multiplying a regular Borel charge by a bounded Borel function.

3. The fact that the spaces $C(K_1)$ and $C(K_2)$ are isomorphic does not necessarily imply that the compact spaces K_1 and K_2 are homeomorphic. Example: $K_1 = [0, 1]$ and $K_2 = [0, 1] \cup \{2\}$.

18.2.2 The Stone–Weierstrass Theorem

In this subsection we make acquaintance with an exceptionally beautiful, and at the same time, very useful generalization of Weierstrass' theorem on the approximation of functions by polynomials. This generalization, devised by M.H. Stone, is applicable to functions defined not only on an interval, but also on an arbitrary compact space. The proof given below is due to de Branges (L. de Branges, 1959). The application of the same idea of proof to an even more general result, namely Bishop's theorem, can be found in the book [38].

Theorem 1. *Suppose the linear subspace X of the space $C(K)$ has the following properties:*

(a) $\mathbb{1} \in X$;

(b) *if $f, g \in X$, then $fg \in X$ (in other words, X is a subalgebra of the algebra $C(K)$);*

(c) *for any function $f \in X$, its complex conjugate $\overline{f}$ also belongs to X;*

(d) *for any $t_1, t_2 \in K$, $t_1 \neq t_2$, there exists a function $f \in X$ such that $f(t_1) \neq f(t_2)$ (i.e., X separates the points of the compact space K).*

  Then the subspace X is dense in $C(K)$.

Proof. Suppose the assertion of the theorem is false, i.e., the subspace X is not dense in $C(K)$. Then the annihilator $X^{\perp} \subset C(K)^*$ does not reduce to 0. Recall that $X^{\perp}$ is a w^*-closed subspace in $C(K)^*$ and hence, by Alaoglu's theorem, $\overline{B}_{X^{\perp}} =$

$\overline{B}_{C(K)^*} \cap X^\perp$ is a w^*-compact set. By the Krein–Milman theorem, the ball $\overline{B}_{X^\perp}$ has an extreme point ν. Obviously, $\nu \in S_{X^\perp}$, that is, $\|\nu\| = 1$. We next study the properties of this regular Borel charge ν and show that they are intrinsically contradictory.

We begin with several useful remarks on the properties of the sets X and $X^\perp$:

(i) If $f \in X$, then $\mathrm{Re}\, f \in X$ and $\mathrm{Im}\, f \in X$ (this follows from condition (c) and the formulas $\mathrm{Re}\, f = (f + \bar{f})/2$ and $\mathrm{Im}\, f = (f - \bar{f})/(2i)$).

(ii) If $f \in X$ and $\eta \in X^\perp$, then $f \times \eta \in X^\perp$, where $\times$ is the operation introduced in the preceding subsection. Indeed, for any $g \in X$ the product fg also lies in X, and hence is annihilated by the charge η. We have $\langle f \times \eta, g \rangle = \langle \eta, fg \rangle = 0$, and so $f \times \eta \in X^\perp$.

(iii) If $\eta \in X^\perp$, then supp η contains at least two distinct points. Indeed, if supp $\eta = \{t\}$, then $\eta = a\delta_t$ with $a \in \mathbb{C} \setminus \{0\}$. But then $\langle \eta, \mathbb{1} \rangle = a \neq 0$, i.e., $\eta \notin X^\perp$.

Now let us return to the charge $\nu \in S_{X^\perp}$, a candidate for the role of an extreme point of the ball $\overline{B}_{X^\perp}$. We use property (iii) above. Let $t_1, t_2 \in$ supp ν and $t_1 \neq t_2$. By condition (d), there exists an $f \in X$ such that $f(t_1) \neq f(t_2)$. Then either $\mathrm{Re}\, f(t_1) \neq \mathrm{Re}\, f(t_2)$, or $\mathrm{Im}\, f(t_1) \neq \mathrm{Im}\, f(t_2)$. By property (i), $\mathrm{Re}\, f, \mathrm{Im}\, f \in X$. Therefore, f can be assumed to be a real-valued function: otherwise one can replace it by $\mathrm{Re}\, f$ or by $\mathrm{Im}\, f$. Further, adding to f a large positive constant one can ensure that that f is positive, and then multiplying by a small positive factor we obtain a function whose values lie in the interval $(0, 1)$. Thus, we proved that there exists a function $f \in X$ such that $f(t_1) \neq f(t_2)$ and $0 < f(t) < 1$ for all $t \in K$.

Let us introduce the auxiliary charges $\nu_1 = f \times \nu$ and $\nu_2 = (1 - f) \times \nu$. Then

$$\|\nu_1\| = \int_K f\, d|\nu|, \quad \|\nu_2\| = \int_K (1 - f)\, d|\nu|,$$

and both numbers are different from zero because, by construction, the functions f and $1 - f$ do not take the value 0. Further,

$$\|\nu_1\| + \|\nu_2\| = \int_K d|\nu| = 1.$$

We have the obvious equality

$$\|\nu_1\| \frac{\nu_1}{\|\nu_1\|} + \|\nu_2\| \frac{\nu_2}{\|\nu_2\|} = \nu,$$

the geometric meaning of which is as follows: the vector $\nu \in \overline{B}_{X^\perp}$ is an interior point of the segment that joins the vectors $\frac{\nu_1}{\|\nu_1\|} \in \overline{B}_{X^\perp}$ and $\frac{\nu_2}{\|\nu_2\|} \in \overline{B}_{X^\perp}$ (the fact that the charges ν_1 and ν_2 lie in the subspace $X^\perp$ follows from property (ii)). Since by our assumption ν is an extreme point of the ball $\overline{B}_{X^\perp}$, the endpoints of the segment must coincide with ν: $\frac{\nu_1}{\|\nu_1\|} = \frac{\nu_2}{\|\nu_2\|} = \nu$. In particular, $\nu_1 = \|\nu_1\| \nu$, i.e., $(f - \|\nu_1\|) \times \nu = 0$. Recalling the formula for the norm of a charge, we get that

$$\int_K \big|f - \|v_1\|\big|\, d|v| = 0.$$

In view of the continuity of the function f, the last equality means that $f(t) = \|v_1\|$ for all $t \in \mathrm{supp}\, v$ (Theorem 2 of Subsection 8.1.2). We arrived at a contradiction with the condition $f(t_1) \neq f(t_2)$. $\square$

Exercises

Deduce from the Stone–Weierstrass theorem that:

1. The set of polynomials is dense in $C(K)$, for any compact subset K of $\mathbb{R}$ (in particular, when $K = [a, b]$). Recall that this fact was used in Subsection 13.1.3 to construct functions of a self-adjoint operator.

2. The set of polynomials in n variables is dense in $C(K)$, where K is any compact subset of $\mathbb{R}^n$.

3. The set of "two-sided" polynomials of the form $\sum_{k=-n}^{n} a_k z^k$, $n \in \mathbb{N}$, is dense in the space $C(\mathbb{T})$ of continuous functions on the unit circle $\mathbb{T} = \{z \in \mathbb{C} : |z| = 1\}$ (we used this fact earlier to construct functions of an unitary operator).

Consider the half-line $[0, +\infty]$, i.e., the one-point compactification of the half-line $[0, +\infty)$. The neighborhoods of finite points of $[0, +\infty]$ are defined as usual; the neighborhoods of $+\infty$ are the complements of the bounded sets. Show that:

4. In the described topology $[0, +\infty]$ is compact.

5. The space $C[0, +\infty]$ coincides with the space of continuous functions $f(t)$ on $[0, +\infty)$ that have a limit as $t \to +\infty$.

6. The set of exponential functions e^{-at} with $a \in [0, +\infty)$ is a complete system of elements in $C[0, +\infty]$.

18.2.3 Completely Monotone Functions

An infinitely differentiable function f on $[0, +\infty)$ is said to be *completely monotone* if $(-1)^n f^{(n)}(t) \geqslant 0$ for all $n = 0, 1, 2, \ldots$ and all $t \in [0, +\infty)$. In particular, to be completely monotone the function f must be non-negative ($f(t) \geqslant 0$), non-increasing $((-1)f'(t) \geqslant 0)$, and convex ($f''(t) \geqslant 0$). A typical example of a completely monotone function is $f(t) = e^{-t}$. A well-known theorem of S.N. Bernstein[2]

---

[2]Kharkiv is a city that hosted many famous mathematicians. Sergei Natanovich Bernstein not only worked for a period of time in Kharkiv, he spent a major part of his life there, exerting an invaluable influence on the formation of the Kharkiv mathematical school.

asserts that any completely monotone function can be uniquely represented in the form

$$f(x) = \int_0^\infty e^{-tx} d\mu(t), \tag{1}$$

where μ is a finite regular Borel measure on the half-line. In other words, every completely monotone function is in a sense a combination of exponentials. Differentiating under the integral sign one can readily verify that any function of the form (1) is completely monotone, and thus Bernstein's theorem provides a complete description of the class of completely monotone functions.

   The representation (1) calls forth a natural association with the Krein–Milman theorem in integral form. The first proof of Bernstein's theorem based on this analogy was proposed by Choquet. Below we provide a sufficiently detailed sketch of this proof, leaving its implementation to the reader. A detailed exposition can be found in the short book [34, Chapter 2].

Theorem 1. *If the function $f : [0, +\infty) \to \mathbb{R}$ admits a representation (1), where μ is a finite regular Borel measure, then this representation is unique.*

Proof. Consider μ as a functional on $C[0, +\infty]$. Formula (1) says that we are given the values of this functional on the exponentials e^{-at}: $\langle \mu, e^{-at} \rangle = f(a)$. By Exercise 6 of Subsection 18.2.2, the set of exponentials e^{-at} with $a \in [0, +\infty)$ is complete in $C[0, +\infty]$. Hence, a continuous functional is uniquely determined by its values on this set.   □

   In the space $C^\infty(0, +\infty)$ of infinitely differentiable functions on the open half-line, equipped with the standard topology generated by the seminorms $p_n(f) = \max_{t \in (n^{-1}, n)} |f^{(n-1)}(t)|$, $n \in \mathbb{N}$, consider the set K of all completely monotone functions bounded above by 1. Note that the functions $f \in K$ are defined on the open half-line, but thanks to their monotonicity and boundedness they have limits at 0 and ∞, and consequently can be considered to be defined also at these two points.

Theorem 2. *The set K is convex and compact in $C^\infty(0, +\infty)$.*

Proof. The convexity and closedness are verified directly. Since $C^\infty(0, +\infty)$ is a Montel space (Subsection 16.3.4), to establish the compactness of K it suffices to verify that K is bounded. Now boundedness follows from the following estimate of the n-th derivative of $f \in K$, the proof of which by induction on n is left to the reader:

$$\sup_{a \leqslant t < \infty} |f^{(n)}(t)| \leqslant a^{-n} 2^{n(n+1)/2}$$

for any $a \in (0, 1)$ and any $n = 0, 1, 2, \dots$.   □

Theorem 3. *Suppose the continuous function $f : (0, +\infty) \to \mathbb{R}$ satisfies for all $x, y \in (0, +\infty)$ the functional equation*

$$f(x + y) = f(x)f(y). \tag{2}$$

Then f is an exponential function of the form $f(x) = a^x$.

Proof. Take for a the value $f(1)$. Substituting into (2) $x = 1$ and $y = 1$, we obtain $f(2) = a^2$. Further, if we fix $x = 1$ and take successively $y = 2, 3, \ldots$, we obtain the equality $f(n) = a^n$. Taking in (2) $x = y = n/2$, we conclude that $f(n/2) = a^{n/2}$. Now taking successively $x = y = n2^{-k}$, we obtain the formula $f(x) = a^x$ for all dyadic-rational numbers. To all the remaining positive real values x the equality $f(x) = a^x$ is extended by continuity. $\qquad\square$

Theorem 4. *The set of extreme points of the compact set K introduced above consists of the functions e^{-at}, $a \in [0, +\infty)$, and the null function.*

Proof. Let $f \in \text{ext } K$. Then fix $y > 0$ and consider the auxiliary function $u(x) = f(x + y) - f(x)f(y)$. The reader will be able to verify that the two functions $f_1 = f + u$ and $f_2 = f - u$ lie in K. Since the extreme point f can be written as $f = (f_1 + f_2)/2$, it follows that $u = 0$. This establishes that f satisfies the functional equation (2), and hence that f is an exponential function. But any exponential function that lies in the set K is either 0, or a function of the form e^{-at}.

Now let us show that indeed all the functions indicated in the statement of the theorem lie in ext K. The fact that the functions 0 and $\mathbb{1}$ lie in ext K follows from the condition $0 \leqslant f(t) \leqslant 1$ that we imposed on all $f \in K$. Further, at least one of the functions e^{-a_0t} with $0 < a_0 < \infty$ is an extreme point. Otherwise, the set ext K would consist only of the functions 0 and $\mathbb{1}$ and, by the Krein–Milman theorem, the compact set $K = \overline{\text{conv}}\,\text{ext } K$ would consist only of constants. Now for any $b \in (0, 1)$ the linear operator T which sends each function $f(x)$ into the function $f(bx)$ maps K bijectively onto K. Therefore, the operator T maps extreme points into extreme ones; in particular, the function e^{-a_0bt} is an extreme point. Since b is arbitrary, this shows that $e^{-at} \in \text{ext } K$ for all $0 < a < \infty$. $\qquad\square$

To complete the proof of Bernstein's theorem, consider the bijective mapping $F \colon [0, +\infty] \to \text{ext } K$ defined by the rule $F(0) = \mathbb{1}$, $F(+\infty) = 0$, and $F(a) = e^{-at}$ for $0 < a < \infty$. The reader will readily verify that the mapping F is continuous. Hence, ext K, being the image of a compact set under a continuous mapping, is a closed set. To conclude, F is a continuous bijective mapping of a compact set into a compact set, and hence a homeomorphism.

Let $f \colon [0, +\infty) \to \mathbb{R}$ be a completely monotone function. With no loss of generality, one can assume that $f \in K$: this is readily achieved via multiplication by a factor. By the Krein–Milman theorem in integral form, there exists a regular Borel probability measure ν on ext K such that

$$f = \int_{\text{ext } K} I \, d\nu. \tag{3}$$

Define the measure μ on $[0, +\infty]$ to be the preimage of ν under the mapping F: $\mu(A) = \nu(F(A))$. Changing the variables in (3) yields

$$f = \int_{[0,+\infty]} F(t)\,d\mu(t).$$

Since $F(+\infty) = 0$, the point $+\infty$ can be removed from the integration domain:

$$f = \int_{[0,+\infty)} F(t)\,d\mu(t).$$

Finally, applying to both sides of this equality the evaluation functional δ_x at the point x, we obtain the requisite representation (1):

$$f(x) = \langle \delta_x, f \rangle = \int_{[0,+\infty)} \langle \delta_x, F(t) \rangle\,d\mu(t) = \int_{[0,+\infty)} e^{-tx}\,d\mu(t). \qquad \square$$

18.2.4 *Lyapunov's Theorem on Vector Measures*

We begin this subsection with the "children's" cake-cutting problem. Bart and Todd want to divide a cake in a fair way. The problem is that different parts of the cake have different gastronomical and aesthetical values: some part has marzipan, another candied peel, one carries a chocolate figurine, and so on. An ever bigger issue is the individuality of the children: they may estimate differently the desirability of one and the same piece of the cake. The standard approach to solving this cutting problem goes as follows: Bart cuts the cake into two pieces that from his point of view are equal, and Todd chooses for himself the part that appeals more to him. In this way Bart is convinced that he received exactly one half of the cake, and Todd thinks he received no less than a half. This approach is completely satisfactory as long as Todd does not start bragging that he got a much better part, and Bart is not envious and starts a fight. To avoid such troubles and keep the peace between friends, it is desirable to cut the cakes into two parts such that the parts will be exactly equal from the point of view of Bart, as well as that of Todd. Is this possible? To answer this question, we need a "mature" formulation.[3]

Thus, let Ω be a set (our cake), Σ a σ-algebra of subsets of Ω (the pieces into which one can cut the cake), and μ_1, μ_2 finite countably-additive measures (for each $A \in \Sigma$ the quantity $\mu_1(A)$ (respectively, $\mu_2(A)$) is the "value" that Bart (respectively, Todd) assigns to the cake piece A).[4] Now the problem reads: *is there a set $A \in \Sigma$ such that $\mu_1(A) = \frac{1}{2}\mu_1(\Omega)$ and also $\mu_2(A) = \frac{1}{2}\mu_2(\Omega)$?* These measures μ_1 and μ_2 must also be required to be non-atomic: if some part of the cake cannot be cut into

---

[3] Here simply replacing Bart by Dr. Bartholomew Simpson and Todd by Prof. Todd Flanders is not sufficient to make the formulation "mature".

[4] In principle, μ_1 and μ_2 could also be charges, if some pieces of the cake do not seem appealing to one of the two friends, i.e., have negative value for him.

smaller pieces and both children like very much precisely that part, then the problem is not solvable.

The following theorem of A.A. Lyapunov (1940) shows that the problem has a solution, and in fact not only for two, but also for any finite number of cake lovers. The importance of the theorem is of course not restricted to the fact that it allows a fair cake cutting, regardless of the great importance and applied character of the cake problem. The proof provided below, which uses extreme points, was proposed by Lindenstrauss in 1966.

Theorem 1. *Let $\mu_1, \ldots, \mu_n$ be countably-additive non-atomic real charges on the σ-algebra Σ. Define the vector measure $\mu \colon \Sigma \to \mathbb{R}^n$ by the formula $\mu(A) = (\mu_1(A), \ldots, \mu_n(A))$. Then the set $\mu(\Sigma)$ of all values of the measure μ is convex and compact in $\mathbb{R}^n$.*

Proof. Consider the scalar-valued measure $\nu = |\mu_1| + \cdots + |\mu_n|$, with respect to which all charges μ_k are absolutely continuous. We use the Radon–Nikodým theorem and denote the derivative $d\mu_k/d\nu$ by g_k. Then $g_k \in L_1(\Omega, \Sigma, \nu)$ and $\mu_k(A) = \int_A g_k \, d\nu$ for all $A \in \Sigma$. Consider the operator $T \colon L_\infty(\Omega, \Sigma, \nu) \to \mathbb{R}^n$, acting by the rule

$$Tf = \left(\int_\Omega fg_1 \, d\nu, \ldots, \int_\Omega fg_n \, d\nu \right).$$

The set $\mu(\Sigma)$ of all values of the vector measure μ we are interested in coincides with the image under the operator T of the set of functions $\mathbb{1}_A$ with $A \in \Sigma$.

The space $L_\infty(\Omega, \Sigma, \nu)$ will be regarded as the dual to $L_1(\Omega, \Sigma, \nu)$. Then each of the expressions $\int_\Omega fg_k \, d\nu$ is a w^*-continuous with respect to f functional on $L_\infty(\Omega, \Sigma, \nu)$, and therefore the operator T is w^*-continuous. Now in $L_\infty(\Omega, \Sigma, \nu)$ consider the set W of functions f that satisfy the condition $0 \leqslant f \leqslant 1$ ν-a.e. Then W coincides with the closed ball centered at $f = 1/2$ and of radius $1/2$. By Alaoglu's theorem, the set W is w^*-compact. Moreover, W is convex. Thus, $T(W)$ is a convex compact subset of $\mathbb{R}^n$. Let us show that $T(W) = \mu(\Sigma)$. This will complete the proof of the entire theorem.

Since the functions $\mathbb{1}_A$ with $A \in \Sigma$ lie in W, and the values of the measure μ are vectors of the form $T(\mathbb{1}_A)$, we have $\mu(\Sigma) \subset T(W)$. Let us establish the opposite inclusion. Let $x \in T(W)$ be an arbitrary element; $T^{-1}(x)$ is a w^*-closed subset, hence $T^{-1}(x) \cap W$ is a w^*-compact set. Let $f \in \operatorname{ext}(T^{-1}(x) \cap W)$. We claim that f takes a.e. the value 0 or 1, i.e., $f = \mathbb{1}_A$ for some set $A \in \Sigma$. Because of the equality $x = T(f) = T(\mathbb{1}_A)$, this will establish the requisite inclusion $T(W) \subset \mu(\Sigma)$.

Consider the set $A = \{t \in \Omega : 0 < f(t) < 1\}$. We need to show that $\nu(A) = 0$. Suppose this is not the case. Define

$$A_n = \left\{ t \in \Omega : \frac{1}{n} < f(t) < 1 - \frac{1}{n} \right\}.$$

By our assumption, the union of the sets A_n is not negligible, so $\nu(A_n) \neq 0$ for some $n \in \mathbb{N}$. Then the subspace $L_\infty(A_n) \subset L_\infty(\Omega, \Sigma, \nu)$ of functions with support in A_n is infinite-dimensional (here in all arguments we use the fact that the measure ν is non-atomic). Since T is a finite-dimensional operator, it cannot be injective on an infinite-dimensional space. Hence, there exists a non-zero element $g \in S_{L_\infty(A_n)}$ such that $Tg = 0$. Then both elements $f \pm \frac{1}{n}g$ lie in $T^{-1}(x) \cap W$, which is impossible because f is an extreme point. $\square$

As the next example will show, the direct extension of Lyapunov's theorem to measures with values in an infinite-dimensional space fails.

Example 1. On the interval $[0, 1]$ define the measure μ with values in $L_2[0, 1]$ by the formula $\mu(A) = \mathbb{1}_A$. This measure is non-atomic and countably additive. At the same time, the set $\mu(\mathfrak{B})$ of all values (the range) of the vector measure μ is not convex: $0,\ \mathbb{1} \in \mu(\mathfrak{B})$, but the function identically equal to $1/2$ does not belong to $\mu(\mathfrak{B})$.

Using the fact that for any infinite-dimensional Banach space X there exists an injective operator $T \colon L_2[0, 1] \to X$, one can readily establish the existence of an X-valued non-atomic Borel measure on $[0, 1]$ with non-convex range. Such a measure can be given by the formula $\mu(A) = T(\mathbb{1}_A)$. Nonetheless, infinite-dimensional analogues of the Laypunov theorem do exist, albeit in a weakened form: such generalizations state that the closure $\overline{\mu(\Sigma)}$ of the set of values, rather than the range $\mu(\Sigma)$ itself, is convex.

Definition 1. The Banach space X is said to have the *Lyapunov property* if for any set Ω, any σ-algebra Σ on Ω, and any non-atomic countably-additive measure $\mu \colon \Sigma \to X$, the set $\overline{\mu(\Sigma)}$ is convex.

The same Example 1 above shows that Hilbert spaces do not have the Lyapunov property. At the same time (see [61]), the spaces c_0 and ℓ_p with $p \in [1, 2) \cup (2, +\infty)$ enjoy the Lyapunov property. Thus, with the Lyapunov property we run into a paradoxical situation: with respect to this property, Hilbert spaces are worse than the (rather badly behaved for other problems and, in particular, non-reflexive) space c_0.

Under additional restrictions on the measure, the community of spaces to which the weaker analogue of the Lyapunov theorem extends widens. For instance, if one considers only measures of bounded variation, then according to a theorem of Uhl (see the last chapter of the book [13], and also the paper [60]), the convexity of the set $\overline{\mu(\Sigma)}$ will hold for non-atomic measures taking values in any space with the Radon–Nikodým property (a class of Banach spaces that includes, in particular, all the reflexive spaces).

Comments on the Exercises

Section 18.1.2

Exercise 6. See [66]. As shown in [48], the set of extreme points of the closed unit ball of a reflexive space is not just uncountable, it cannot even have the "small balls property" (concerning this property, see the exercises of Subsection 11.2.1).

Exercise 7. See [3, Corollary 5.12 and Proposition 5.13].

Exercise 9. See [34, P. 15].

Exercise 10. The solution given here was communicated to us by Dirk Werner. Suppose that for some operator $T \in L(X)$ it holds that $\|I \pm T\| \leqslant 1$. Then $\|I^* \pm T^*\| \leqslant 1$, and for any $x^* \in \text{ext } \overline{B}_{X^*}$ we have $\|x^* \pm T^*x^*\| \leqslant 1$. By the definition of an extreme point, this means that $T^*x^* = 0$. Thus, we have proved that T^* maps into 0 all extreme points of the ball $\overline{B}_{X^*}$, and so $T^* = 0$. Consequently, $T = 0$, too.

Exercise 11. Use the previous exercise and Exercise 6 of Subsection 11.1.1.

Section 18.2.1

Exercise 3. By Milyutin's theorem (see the monograph [33]), if K_1 and K_2 are uncountable metrizable compact spaces, then the spaces $C(K_1)$ and $C(K_2)$ are isomorphic. For the concrete case $K_1 = [0, 1]$ and $K_2 = [0, 1] \cup \{2\}$ the corresponding isomorphism of $C[0, 1]$ and $C(K_2)$ can be given without appealing to the highly nontrivial Milutin's construction. Namely, one finds in $C[0, 1]$ a subspace X isomorphic to c_0 (for any sequence of functions $f_n \in S_{C[0,1]}$ with disjoint supports, $X = \overline{\text{Lin}}\{f_n\}$ is a subspace of $C[0, 1]$ isometric to c_0), then one represents $C[0, 1]$ as a direct sum of the form $X \oplus Y$, writing $C(K_2) = \text{Lin}\{\mathbb{1}_{\{2\}}\} \oplus C[0, 1] = \text{Lin}\{\mathbb{1}_{\{2\}}\} \oplus X \oplus Y$, and finally one proves (using the shift operator) that $\text{Lin}\{\mathbb{1}_{\{2\}}\} \oplus X$ is isomorphic to c_0, and thus is isomorphic to X.