CHAPTER 3

CONVEXITY

This chapter deals primarily (though not exclusively) with the most important class of topological vector spaces, namely, the locally convex ones. The highlights, from the theoretical as well as the applied standpoints, are (*a*) the Hahn-Banach theorems (assuring a supply of continuous linear functionals that is adequate for a highly developed duality theory), (*b*) the Banach-Alaoglu compactness theorem in dual spaces, and (*c*) the Krein-Milman theorem about extreme points. Applications to various problems in analysis are postponed to Chapter 5.

The Hahn-Banach Theorems

The plural is used here because the term "Hahn-Banach theorem" is customarily applied to several closely related results. Among these are the *dominated extension theorems* 3.2 and 3.3 (in which no topology is involved), the *separation theorem* 3.4, and the *continuous extension theorem* 3.6. Another separation theorem (which implies 3.4) is stated as Exercise 3.

3.1 Definitions The *dual space* of a topological vector space X is the vector space X^* whose elements are the *continuous* linear functionals on X.

Note that addition and scalar multiplication are defined in X^* by

$$(\Lambda_1 + \Lambda_2)x = \Lambda_1 x + \Lambda_2 x, \qquad (\alpha\Lambda)x = \alpha \cdot \Lambda x.$$

It is clear that these operations do indeed make X^* into a vector space.

It will be necessary to use the obvious fact that every complex vector space is also a real vector space, and it will be convenient to use the following (temporary) terminology: An additive functional Λ on a complex vector space X is called *real-linear* (*complex-linear*) if $\Lambda(\alpha x) = \alpha\Lambda x$ for every $x \in X$ and for every real (complex) scalar α. Our standing rule that any statement about vector spaces in which no scalar field is mentioned applies to both cases is unaffected by this temporary terminology and is still in force.

If u is the real part of a complex-linear functional f on X, then u is real-linear and

$$f(x) = u(x) - iu(ix) \qquad (x \in X) \tag{1}$$

because $z = \text{Re } z - i \text{ Re }(iz)$ for every $z \in \mathcal{C}$.

Conversely, if $u: X \to R$ is real-linear on a complex vector space X and if f is defined by (1), a straightforward computation shows that f is complex-linear.

Suppose now that X is a complex topological vector space. The above facts imply that a complex-linear functional on X is in X^* if and only if its real part is continuous, and that every continuous real-linear $u: X \to R$ is the real part of a unique $f \in X^*$.

3.2 Theorem *Suppose*

(*a*) *M is a subspace of a real vector space X,*

(*b*) *$p: X \to R$ satisfies*

$$p(x + y) \leq p(x) + p(y) \qquad \textit{and} \qquad p(tx) = p(x)$$

if $x \in X$, $y \in X$, $t \geq 0$,

(*c*) *$f: M \to R$ is linear and $f(x) \leq p(x)$ on M.*

Then there exists a linear $\Lambda: X \to R$ such that

$$\Lambda x = f(x) \qquad (x \in M)$$

and

$$-p(-x) \leq \Lambda x \leq p(x) \qquad (x \in X).$$

PROOF. If $M \neq X$, choose $x_1 \in X$, $x_1 \notin M$, and define

$$M_1 = \{x + tx_1: x \in M, t \in R\}.$$

It is clear that M_1 is a vector space. Since

$$f(x) + f(y) = f(x + y) \leq p(x + y) \leq p(x - x_1) + p(x_1 + y),$$

we have

$$(1)\qquad f(x) - p(x - x_1) \le p(y + x_1) - f(y) \qquad (x, y \in M).$$

Let α be the least upper bound of the left side of (1), as x ranges over M. Then

$$(2)\qquad f(x) - \alpha \le p(x - x_1) \qquad (x \in M)$$

and

$$(3)\qquad f(y) + \alpha \le p(y + x_1) \qquad (y \in M).$$

Define f_1 on M_1 by

$$(4)\qquad f_1(x + tx_1) = f(x) + t\alpha \qquad (x \in M,\ t \in R).$$

Then $f_1 = f$ on M, and f_1 is linear on M_1.

Take $t > 0$, replace x by $t^{-1}x$ in (2), replace y by $t^{-1}y$ in (3), and multiply the resulting inequalities by t. In combination with (4), this proves that $f_1 \le p$ on M_1.

The second part of the proof can be done by whatever one's favorite method of transfinite induction is; one can use well-ordering, or Zorn's lemma, or Hausdorff's maximality theorem.

Let $\mathscr{P}$ be the collection of all ordered pairs (M', f'), where M' is a subspace of X that contains M and f' is a linear functional on M' that extends f and satisfies $f' \le p$ on M'. Partially order $\mathscr{P}$ by declaring $(M', f') \le (M'', f'')$ to mean that $M' \subset M''$ and $f'' = f'$ on M'. By Hausdorff's maximality theorem there exists a maximal totally ordered subcollection Ω of $\mathscr{P}$.

Let Φ be the collection of all M' such that $(M', f') \in \Omega$. Then Φ is totally ordered by set inclusion, and the union $\tilde{M}$ of all members of Φ is therefore a subspace of X. If $x \in \tilde{M}$ then $x \in M'$ for some $M' \in \Phi$; define $\Lambda x = f'(x)$, where f' is the function which occurs in the pair $(M', f') \in \Omega$.

It is now easy to check that Λ is well defined on $\tilde{M}$, that Λ is linear, and that $\Lambda \le p$. If $\tilde{M}$ were a proper subspace of X, the first part of the proof would give a further extension of Λ, and this would contradict the maximality of Ω. Thus $\tilde{M} = X$.

Finally, the inequality $\Lambda \le p$ implies that

$$-p(-x) \le -\Lambda(-x) = \Lambda x$$

for all $x \in X$. This completes the proof. ////

3.3 Theorem *Suppose M is a subspace of a vector space X, p is a seminorm on X, and f is a linear functional on M such that*

$$|f(x)| \le p(x) \qquad (x \in M).$$

Then f extends to a linear functional Λ on X that satisfies

$$|\Lambda x| \leq p(x) \qquad (x \in X).$$

PROOF. If the scalar field is R, this is contained in Theorem 3.2, since p now satisfies $p(-x) = p(x)$.

Assume that the scalar field is C. Put $u = \text{Re} f$. By Theorem 3.2 there is a real-linear U on X such that $U = u$ on M and $U \leq p$ on X. Let Λ be the complex-linear functional on X whose real part is U. The discussion in Section 3.1 implies that $\Lambda = f$ on M.

Finally, to every $x \in X$ corresponds an $\alpha \in C$, $|\alpha| = 1$, such that $\alpha\Lambda x = |\Lambda x|$. Hence

$$|\Lambda x| = \Lambda(\alpha x) = U(\alpha x) \leq p(\alpha x) = p(x). \qquad ////$$

Corollary. *If X is a normed space and $x_0 \in X$, there exists $\Lambda \in X^*$ such that*

$$\Lambda x_0 = \|x_0\| \qquad \textit{and} \qquad |\Lambda x| \leq \|x\| \qquad \textit{for all } x \in X.$$

PROOF. If $x_0 = 0$, take $\Lambda = 0$. If $x_0 \neq 0$, apply Theorem 3.3, with $p(x) = \|x\|$, M the one-dimensional space generated by x_0, and $f(\alpha x_0) = \alpha\|x_0\|$ on M. ////

3.4 Theorem *Suppose A and B are disjoint, nonempty, convex sets in a topological vector space X.*

(*a*) *If A is open there exist $\Lambda \in X^*$ and $\gamma \in R$ such that*

$$\text{Re } \Lambda x < \gamma \leq \text{Re } \Lambda y$$

for every $x \in A$ and for every $y \in B$.

(*b*) *If A is compact, B is closed, and X is locally convex, then there exist $\Lambda \in X^*$, $\gamma_1 \in R$, $\gamma_2 \in R$, such that*

$$\text{Re } \Lambda x < \gamma_1 < \gamma_2 < \text{Re } \Lambda y$$

for every $x \in A$ and for every $y \in B$.

Note that this is stated without specifying the scalar field; if it is R, then $\text{Re } \Lambda = \Lambda$, of course.

PROOF. It is enough to prove this for real scalars. For if the scalar field is C and the real case has been proved, then there is a continuous real-linear Λ_1 on X that gives the required separation; if Λ is the unique complex-linear functional on X whose real part is Λ_1, then $\Lambda \in X^*$. (See Section 3.1.) Assume real scalars.

(*a*) Fix $A_0 \in A$, $b_0 \in B$. Put $x_0 = b_0 - a_0$; put $C = A - B + x_0$. Then C is a convex neighborhood of 0 in X. Let p be the Minkowski functional of C. By Theorem 1.35, p satisfies hypothesis (*b*) of Theorem 3.2. Since $A \cap B = \varnothing$, $x_0 \notin C$, and so $p(x_0) \geq 1$.

Define $f(tx_0) = t$ on the subspace M of X generated by x_0. If $t \geq 0$ then

$$f(tx_0) = t \leq tp(x_0) = p(tx_0);$$

if $t < 0$ then $f(tx_0) < 0 \leq p(tx_0)$. Thus $f \leq p$ on M. By Theorem 3.2, f extends to a linear functional Λ on X that also satisfies $\Lambda \leq p$. In particular, $\Lambda \leq 1$ on C, hence $\Lambda \geq -1$ on $-C$, so that $|\Lambda| \leq 1$ on the neighborhood $C \cap (-C)$ of 0. By Theorem 1.18, $\Lambda \in X^*$.

If now $\alpha \in A$ and $b \in B$, we have

$$\Lambda a - \Lambda b + 1 = \Lambda(a - b + x_0) \leq p(a - b + x_0) < 1$$

since $\Lambda x_0 = 1$, $a - b + x_0 \in C$, and C is open. Thus $\Lambda a < \Lambda b$.

It follows that $\Lambda(A)$ and $\Lambda(B)$ are disjoint convex subsets of R, with $\Lambda(A)$ to the left of $\Lambda(B)$. Also, $\Lambda(A)$ is an open set since A is open and since every nonconstant linear functional on X is an open mapping. Let γ be the right end point of $\Lambda(A)$ to get the conclusion of part (*a*).

(*b*) By Theorem 1.10 there is a convex neighborhood V of 0 in X such that $(A + V) \cap B = \varnothing$. Part (*a*), with $A + V$ in place of A, shows that there exists $\Lambda \in X^*$ such that $\Lambda(A + V)$ and $\Lambda(B)$ are disjoint convex subsets of R, with $\Lambda(A + V)$ open and to the left of $\Lambda(B)$. Since $\Lambda(A)$ is a compact subset of $\Lambda(A + V)$, we obtain the conclusion of (*b*). ////

Corollary. *If X is a locally convex space then X^* separates points on X.*

PROOF. If $x_1 \in X$, $x_2 \in X$, and $x_1 \neq x_2$, apply (*b*) of Theorem 3.4 with $A = \{x_1\}$, $B = \{x_2\}$. ////

3.5 Theorem *Suppose M is a subspace of a locally convex space X, and $x_0 \in X$. If x_0 is not in the closure of M, then there exists $\Lambda \in X^*$ such that $\Lambda x_0 = 1$ but $\Lambda x = 0$ for every $x \in M$.*

PROOF. By (*b*) of Theorem 3.4, with $A = \{x_0\}$ and $B = \bar{M}$, there exists $\Lambda \in X^*$ such that Λx_0 and $\Lambda(M)$ are disjoint. Thus $\Lambda(M)$ is a *proper* subspace of the scalar field. This forces $\Lambda(M) = \{0\}$ and $\Lambda x_0 \neq 0$. The desired functional is obtained by dividing Λ by Λx_0. ////

Remark. This theorem is the basis of a standard method of treating certain approximation problems: In order to prove that an $x_0 \in X$ lies in the closure of some subspace M of X it suffices (if X is locally convex) to show that $\Lambda x_0 = 0$ for every continuous linear functional Λ on X that vanishes on M.

3.6 Theorem *If f is a continuous linear functional on a subspace M of a locally convex space X, then there exists $\Lambda \in X^*$ such that $\Lambda = f$ on M.*

Remark. For normed spaces this is an immediate corollary of Theorem 3.3. The general case could also be obtained from 3.3, by relating the continuity of linear functionals to seminorms (see Exercise 8, Chapter 1). The proof given below shows that Theorem 3.6 depends only on the separation property of Theorem 3.5.

PROOF. Assume, without loss of generality, that f is not identically 0 on M. Put

$$M_0 = \{x \in M: f(x) = 0\}$$

and pick $x_0 \in M$ such that $f(x_0) = 1$. Since f is continuous, x_0 is not in the M-closure of M_0, and since M inherits its topology from X, it follows that x_0 is not in the X-closure of M_0.

Theorem 3.5 therefore assures the existence of a $\Lambda \in X^*$ such that $\Lambda x_0 = 1$ and $\Lambda = 0$ on M_0.

If $x \in M$, then $x - f(x)x_0 \in M_0$, since $f(x_0) = 1$. Hence

$$\Lambda x - f(x) = \Lambda x - f(x)\Lambda x_0 = \Lambda(x - f(x)x_0) = 0.$$

Thus $\Lambda = f$ on M. ////

We conclude this discussion with another useful corollary of the separation theorem.

3.7 Theorem *Suppose B is a convex, balanced, closed set in a locally convex space X, $x_0 \in X$, but $x_0 \notin B$. Then there exists $\Lambda \in X^*$ such that $|\Lambda x| \leq 1$ for all $x \in B$, but $\Lambda x_0 > 1$.*

PROOF. Since B is closed and convex, we can apply (*b*) of Theorem 3.4, with $A = \{x_0\}$, to obtain $\Lambda_1 \in X^*$ such that $\Lambda_1 x_0 = re^{i\theta}$ lies outside the closure K of $\Lambda_1(B)$. Since B is balanced, so is K. Hence there exists s, $0 < s < r$, so that $|z| \leq s$ for all $z \in K$. The functional $\Lambda = s^{-1}e^{-i\theta}\Lambda_1$ has the desired properties. ////

Weak Topologies

3.8 Topological preliminaries The purpose of this section is to explain and illustrate some of the phenomena that occur when a set is topologized in several ways.

Let τ_1 and τ_2 be two topologies on a set X, and assume $\tau_1 \subset \tau_2$; that is, every τ_1-open set is also τ_2-open. Then we say that τ_1 is *weaker* than τ_2, or that τ_2 is *stronger* than τ_1. [Note that (in accordance with the meaning of the inclusion symbol $\subset$) the terms "weaker" and "stronger" do not exclude equality.] In this situation, the identity mapping on X is *continuous* from (X, τ_2) to (X, τ_1) and is an *open mapping* from (X, τ_1) to (X, τ_2).

As a first illustration, let us prove that the topology of a compact Hausdorff space has a certain rigidity, in the sense that it cannot be weakened without losing the Hausdorff separation axiom and cannot be strengthened without losing compactness:

(*a*) *If $\tau_1 \subset \tau_2$ are topologies on a set X, if τ_1 is a Hausdorff topology, and if τ_2 is compact, then $\tau_1 = \tau_2$.*

To see this, let $F \subset X$ be τ_2-closed. Since X is τ_2-compact, so is F. Since $\tau_1 \subset \tau_2$, it follows that F is τ_1-compact. (Every τ_1-open cover of F is also a τ_2-open cover.) Since τ_1 is a Hausdorff topology, it follows that F is τ_1-closed.

As another illustration, consider the quotient topology τ_N of X/N, as defined in Section 1.40, and the quotient map $\pi: X \to X/N$. By its very definition, τ_N is the strongest topology on X/N that makes π continuous, and it is the weakest one that makes π an open mapping. Explicitly, if τ' and τ'' are topologies on X/N, and if π is continuous relative to τ' and open relative to τ'', then $\tau' \subset \tau_N \subset \tau''$.

Suppose next that X is a set and $\mathscr{F}$ is a nonempty family of mappings $f: X \to Y_f$, where each Y_f is a topological space. (In many important cases, Y_f is the same for all $f \in \mathscr{F}$.) Let τ be the collection of all unions of finite intersections of sets $f^{-1}(V)$, with $f \in \mathscr{F}$ and V open in Y_f. Then τ is a topology on X, and it is in fact the *weakest* topology on X that makes every $f \in \mathscr{F}$ continuous: If τ' is any other topology with that property, then $\tau \subset \tau'$. This τ is called *the weak topology on X induced by $\mathscr{F}$*, or, more succinctly, the *$\mathscr{F}$-topology of X*.

The best-known example of this situation is undoubtedly the usual way in which one topologizes the cartesian product X of a collection of topological spaces X_α. If $\pi_\alpha(x)$ denotes the αth coordinate of a point $x \in X$, then π_α maps X onto X_α, and the product topology τ of X is, by definition, its $\{\pi_\alpha\}$-topology, the weakest one that makes every π_α continuous. Assume now that every X_α is a *compact Hausdorff space.* Then τ is a compact topology on X (by Tychonoff's theorem), and proposition (*a*) implies that τ cannot be strengthened without spoiling Tychonoff's theorem.

In the last sentence a special case of the following proposition was tacitly used:

(*b*) *If $\mathscr{F}$ is a family of mappings $f: X \to Y_f$, where X is a set and each Y_f is a Hausdorff space, and if $\mathscr{F}$ separates points on X, then the $\mathscr{F}$-topology of X is a Hausdorff topology.*

For if $p \neq q$ are points of X, then $f(p) \neq f(q)$ for some $f \in \mathscr{F}$; the points $f(p)$ and $f(q)$ have disjoint neighborhoods in Y_f whose inverse images under f are open (by definition) and disjoint.

Here is an application of these ideas to a metrization theorem.

(*c*) *If X is a compact topological space and if some sequence $\{f_n\}$ of continuous real-valued functions separates points on X, then X is metrizable.*

Let τ be the given topology of X. Suppose, without loss of generality, that $|f_n| \leq 1$ for all n, and let τ_d be the topology induced on X by the metric

$$d(p, q) = \sum_{n=1}^{\infty} 2^{-n} |f_n(p) - f_n(q)|.$$

This is indeed a metric, since $\{f_n\}$ separates points. Since each f_n is τ-continuous and the series converges uniformly on $X \times X$, d is a τ-continuous function on $X \times X$. The balls

$$B_r(p) = \{q \in X: d(p, q) < r\}$$

are therefore τ-open. Thus $\tau_d \subset \tau$. Since τ_d is induced by a metric, τ_d is a Hausdorff topology, and now (*a*) implies that $\tau = \tau_d$.

The following lemma has applications in the study of vector topologies. In fact, the case $n = 1$ was needed (and proved) at the end of Theorem 3.6.

3.9 Lemma *Suppose $\Lambda_1, \ldots, \Lambda_n$ and Λ are linear functionals on a vector space X. Let*

$$N = \{x: \Lambda_1 x = \cdots = \Lambda_n x = 0\}.$$

The following three properties are then equivalent:

(*a*) *There are scalars $\alpha_1, \ldots, \alpha_n$ such that*

$$\Lambda = \alpha_1 \Lambda_1 + \cdots + \alpha_n \Lambda_n.$$

(*b*) *There exists* $\gamma < \infty$ *such that*

$$|\Lambda x| \le \gamma \max_{1 \le i \le n} |\Lambda_i x| \qquad (x \in X).$$

(*c*) $\Lambda x = 0$ *for every* $x \in N$.

PROOF. It is clear that (*a*) implies (*b*) and that (*b*) implies (*c*). Assume (*c*) holds. Let Φ be the scalar field. Define $\pi\colon X \to \Phi^n$ by

$$\pi(x) = (\Lambda_1 x, \ldots, \Lambda_n x).$$

If $\pi(x) = \pi(x')$, then (*c*) implies $\Lambda x = \Lambda x'$. Hence $f(\pi(x)) = \Lambda x$ defines a linear functional f on $\pi(X)$. Extend f to a linear functional F on Φ^n. This means that there exist $\alpha_i \in \Phi$ such that

$$F(u_1, \ldots, u_n) = \alpha_1 u_1 + \cdots + \alpha_n u_n.$$

Thus

$$\Lambda x = F(\pi(x)) = F(\Lambda_1 x, \ldots, \Lambda_n x) = \sum_{i=1}^{n} \alpha_i \Lambda_i x,$$

which is (*a*). ////

3.10 Theorem *Suppose X is a vector space and X' is a separating vector space of linear functionals on X. Then the X'-topology τ' makes X into a locally convex space whose dual space is X'.*

The assumptions on X' are, more explicitly, that X' is closed under addition and scalar multiplication and that $\Lambda x_1 \ne \Lambda x_2$ for some $\Lambda \in X'$ whenever x_1 and x_2 are distinct points of X.

PROOF. Since R and C are Hausdorff spaces, (*b*) of Section 3.8 shows that τ' is a Hausdorff topology. The linearity of the members of X' shows that τ' is translation-invariant. If $\Lambda_1, \ldots, \Lambda_n \in X'$, if $r_i > 0$, and if

(1) $$V = \{x\colon |\Lambda_i x| < r_i \quad \text{for} \quad 1 \le i \le n\},$$

then V is convex, balanced, and $V \in \tau'$. In fact, the collection of all V of the form (1) is a local base for τ'. Thus τ' is a locally convex topology on X.

If (1) holds, then $\frac{1}{2}V + \frac{1}{2}V = V$. This proves that addition is continuous. Suppose $x \in X$ and α is a scalar. Then $x \in sV$ for some $s > 0$. If $|\beta - \alpha| < r$ and $y - x \in rV$ then

$$\beta y - \alpha x = (\beta - \alpha)y + \alpha(y - x)$$

lies in V, provided that r is so small that

$$r(s + r) + |\alpha| r < 1.$$

Hence scalar multiplication is continuous.

We have now proved that τ' is a locally convex vector topology. Every $\Lambda \in X'$ is τ'-continuous. Conversely, suppose Λ is a τ'-continuous linear functional on X. Then $|\Lambda x| < 1$ for all x in some set V of the form (1). Condition (*b*) of Lemma 3.9 therefore holds; hence so does (*a*): $\Lambda = \sum \alpha_i \Lambda_i$. Since $\Lambda_i \in X'$ and X' is a vector space, $\Lambda \in X'$. This completes the proof. ////

Note: The first part of this proof could have been based on Theorem 1.37 and the separating family of seminorms $p_\Lambda(\Lambda \in X')$ given by $p_\Lambda(x) = |\Lambda x|$.

3.11 The weak topology of a topological vector space Suppose X is a topological vector space (with topology τ) whose dual X^* separates points on X. (We know that this happens in every locally convex X. It also happens in some others; see Exercise 5.) The X^*-topology of X is called *the weak topology of X*.

We shall let X_w denote X topologized by this weak topology τ_w. Theorem 3.10 implies that X_w is a locally convex space whose dual is also X^*.

Since every $\Lambda \in X^*$ is τ-continuous and since τ_w is the weakest topology on X with that property, we have $\tau_w \subset \tau$. In this context, the given topology τ will often be called the *original* topology of X.

Self-explanatory expressions such as original neighborhood, weak neighborhood, original closure, weak closure, originally bounded, weakly bounded, etc., will be used to make it clear with respect to which topology these terms are to be understood.[1]

For instance, let $\{x_n\}$ be a sequence in X. To say that $x_n \to 0$ originally means that every original neighborhood of 0 contains all x_n with sufficiently large n. To say that $x_n \to 0$ weakly means that every weak neighborhood of 0 contains all x_n with sufficiently large n. Since every weak neighborhood of

[1] When X is a Fréchet space (hence, in particular, when X is a Banach space) the original topology of X is usually called its *strong topology*. In that context, the terms "strong" and "strongly" will be used in place of "original" and "originally." For locally convex spaces in general, the term "strong topology" has been given a specific technical meaning. See [15], pp. 256–258; also [14], p. 169. It seems therefore advisable to use "original" in the present general discussion.

0 contains a neighborhood of the form

$$(1) \qquad V = \{x : |\Lambda_i x| < r_i \quad \text{for} \quad 1 \le i \le n\},$$

where $\Lambda_i \in X^*$ and $r_i > 0$, it is easy to see that $x_n \to 0$ *weakly if and only if* $\Lambda x_n \to 0$ *for every* $\Lambda \in X^*$.

Hence every originally convergent sequence converges weakly. (The converse is usually false; see Exercises 5 and 6.)

Similarly, a set $E \subset X$ is *weakly bounded* (that is, E is a bounded subset of X_w) if and only if every V as in (1) contains tE for some $t = t(V) > 0$. This happens if and only if there corresponds to each $\Lambda \in X^*$ a number $\gamma(\Lambda) < \infty$ such that $|\Lambda x| \le \gamma(\Lambda)$ for every $x \in E$. In other words, *a set $E \subset X$ is weakly bounded if and only if every $\Lambda \in X^*$ is a bounded function on E.*

Let V again be as in (1), and put

$$N = \{x : \Lambda_1 x = \cdots = \Lambda_n x = 0\}.$$

Since $x \to (\Lambda_1 x, \ldots, \Lambda_n x)$ maps X into $\mathcal{C}^n$ with null space N, we see that $\dim X \le n + \dim N$. Since $N \subset V$, this leads to the following conclusion.

If X is infinite-dimensional then every weak neighborhood of 0 *contains an infinite-dimensional subspace; hence X_w is not locally bounded.*

This implies in many cases that the weak topology is strictly weaker than the original one. Of course, the two may coincide: Theorem 3.10 implies that $(X_w)_w = X_w$.

We now come to a more interesting result.

3.12 Theorem *Suppose E is a convex subset of a locally convex space X. Then the weak closure $\bar{E}_w$ of E is equal to its original closure $\bar{E}$.*

PROOF. $\bar{E}_w$ is weakly closed, hence originally closed, so that $\bar{E} \subset \bar{E}_w$. To obtain the opposite inclusion, choose $x_0 \in X$, $x_0 \notin \bar{E}$. Part (*b*) of the separation theorem 3.4 shows that there exist $\Lambda \in X^*$ and $\gamma \in R$ such that, for every $x \in \bar{E}$,

$$\operatorname{Re} \Lambda x_0 < \gamma < \operatorname{Re} \Lambda x.$$

The set $\{x : \operatorname{Re} \Lambda x < \gamma\}$ is therefore a weak neighborhood of x_0 that does not intersect E. Thus x_0 is not in $\bar{E}_w$. This proves $\bar{E}_w \subset \bar{E}$. ////

Corollaries. *For convex subsets of a locally convex space,*

(*a*) *originally closed equals weakly closed, and*

(*b*) *originally dense equals weakly dense.*

The proofs are obvious. Here is another noteworthy consequence of Theorem 3.12.

3.13 Theorem *Suppose X is a metrizable locally convex space. If $\{x_n\}$ is a sequence in X that converges weakly to some $x \in X$, then there is a sequence $\{y_i\}$ in X such that*

(*a*) *each y_i is a convex combination of finitely many x_n, and*
(*b*) *$y_i \to x$ originally.*

Conclusion (*a*) says, more explicitly, that there exist numbers $\alpha_{in} \geq 0$, such that

$$\sum_{n=1}^{\infty} \alpha_{in} = 1, \qquad y_i = \sum_{n=1}^{\infty} \alpha_{in} x_n,$$

and, for each i, only finitely many α_{in} are $\neq 0$.

PROOF. Let H be the convex hull of the set of all x_n; let K be the weak closure of H. Then $x \in K$. By Theorem 3.12, x is also in the original closure of H. Since the original topology of X is assumed to be metrizable, it follows that there is a sequence $\{y_i\}$ in H that converges originally to x. ////

To get a feeling for what is involved here, consider the following example.

Let K be a compact Hausdorff space (the unit interval on the real line is a sufficiently interesting one), and assume that f and f_n ($n = 1, 2, 3, \ldots$) are continuous complex functions on K such that $f_n(x) \to f(x)$ for every $x \in K$, as $n \to \infty$, and such that $|f_n(x)| \leq 1$ for all n and all $x \in K$. Theorem 3.13 asserts that there are convex combinations of the f_n that converge *uniformly* to f.

To see this, let $C(K)$ be the Banach space of all complex continuous functions on K, normed by the supremum. Then strong convergence is the same as uniform convergence on K. If μ is any complex Borel measure on K, Lebesgue's dominated convergence theorem implies that $\int f_n \, d\mu \to \int f \, d\mu$. Hence $f_n \to f$ weakly, by the Riesz representation theorem which identifies the dual of $C(K)$ with the space of all regular complex Borel measures on K. Now Theorem 3.13 can be applied.

After this short detour we now return to our main line of development.

3.14 The weak*-topology of a dual space Let X again be a topological vector space whose dual is X^*. For the definitions that follow, it is irrelevant whether X^* separates points on X or not. The important observation to make is that *every* $x \in X$ *induces a linear functional* f_x *on* X^*, *defined by*

$$f_x \Lambda = \Lambda x,$$

and that $\{f_x : x \in X\}$ *separates points on* X^*.

The linearity of each f_x is obvious; if $f_x \Lambda = f_x \Lambda'$ for all $x \in X$, then $\Lambda x = \Lambda' x$ for all x, and so $\Lambda = \Lambda'$ by the very definition of what it means for two functions to be equal.

We are now in the situation described by Theorem 3.10, with X^* in place of X and with X in place of X'.

The X-topology of X^* is called *the weak*-topology of* X^* (pronunciation: weak star topology).

Theorem 3.10 implies that this is a locally convex vector topology on X^* and that *every linear functional on* X^* *that is weak*-continuous has the form* $\Lambda \to \Lambda x$ *for some* $x \in X$.

The weak*-topologies have a very important compactness property to which we now turn our attention. Various pathological features of the weak- and weak*-topologies are described in Exercises 9 and 10.

Compact Convex Sets

3.15 The Banach-Alaoglu theorem *If* V *is a neighborhood of* 0 *in a topological vector space* X *and if*

$$K = \{\Lambda \in X^* : |\Lambda x| \leq 1 \quad \text{for every} \quad x \in V\}$$

then K *is weak*-compact.*

Note: K is sometimes called the *polar* of V. It is clear that K is convex and balanced, because this is true of the unit disc in $\mathbb{C}$ (and of the interval $[-1, 1]$ in R). There is some redundancy in the definition of K, since every linear functional on X that is bounded on V is continuous, hence is in X^*.

PROOF. Since neighborhoods of 0 are absorbing, there corresponds to each $x \in X$ a number $\gamma(x) < \infty$ such that $x \in \gamma(x)V$. Hence

$$(1) \qquad |\Lambda x| \leq \gamma(x) \qquad (x \in X, \Lambda \in K).$$

Let D_x be the set of all scalars α such that $|\alpha| \leq \gamma(x)$. Let τ be the product topology on P, the cartesian product of all D_x, one for each $x \in X$. Since each D_x is compact, so is P, by Tychonoff's theorem. The elements of P are the functions f on X (linear or not) that satisfy

$$(2) \qquad |f(x)| \leq \gamma(x) \qquad (x \in X).$$

Thus $K \subset X^* \cap P$. It follows that K inherits two topologies: one from X^* (its weak*-topology, to which the conclusion of the theorem refers) and the other, τ, from P. We will see that

(*a*) these two topologies coincide on K, and
(*b*) K is a closed subset of P.

Since P is compact, (*b*) implies that K is τ-compact, and then (*a*) implies that K is weak*-compact.

Fix some $\Lambda_0 \in K$. Choose $x_i \in X$, for $1 \leq i \leq n$; choose $\delta > 0$. Put

(3) $$W_1 = \{\Lambda \in X^* : |\Lambda x_i - \Lambda_0 x_i| < \delta \quad \text{for} \quad 1 \leq i \leq n\}$$

and

(4) $$W_2 = \{f \in P : |f(x_i) - \Lambda_0 x_i| < \delta \quad \text{for} \quad 1 \leq i \leq n\}.$$

Let n, x_i, and δ range over all admissible values. The resulting sets W_1 then form a local base for the weak*-topology of X^* at Λ_0 and the sets W_2 form a local base for the product topology τ of P at Λ_0. Since $K \subset P \cap X^*$, we have

$$W_1 \cap K = W_2 \cap K.$$

This proves (*a*).

Next, suppose f_0 is in the τ-closure of K. Choose $x \in X$, $y \in X$, scalars α and β, and $\varepsilon > 0$. The set of all $f \in P$ such that $|f - f_0| < \varepsilon$ at x, at y, and at $\alpha x + \beta y$ is a τ-neighborhood of f_0. Therefore K contains such an f. Since this f is linear, we have

$$f_0(\alpha x + \beta y) - \alpha f_0(x) - \beta f_0(y)$$
$$= (f_0 - f)(\alpha x + \beta y) + \alpha(f - f_0)(x) + \beta(f - f_0)(y),$$

so that

$$|f_0(\alpha x + \beta y) - \alpha f_0(x) - \beta f_0(y)| < (1 + |\alpha| + |\beta|)\varepsilon.$$

Since ε was arbitrary, we see that f_0 is linear. Finally, if $x \in V$ and $\varepsilon > 0$, the same argument shows that there is an $f \in K$ such that $|f(x) - f_0(x)| < \varepsilon$. Since $|f(x)| \leq 1$, by the definition of K, it follows that $|f_0(x)| \leq 1$. We conclude that $f_0 \in K$. This proves (*b*) and hence the theorem. ////

When X is *separable* (i.e., when there is a countable dense set in X), then the conclusion of the Banach-Alaoglu theorem can be strengthened by combining it with the following fact:

3.16 Theorem *If X is a separable topological vector space, if $K \subset X^*$, and if K is weak*-compact, then K is metrizable, in the weak*-topology.*

Warning: It does not follow that X^* itself is metrizable in its weak*-topology. In fact, this is false whenever X is an infinite-dimensional Banach space. See Exercise 15.

PROOF. Let $\{x_n\}$ be a countable dense set in X. Put $f_n(\Lambda) = \Lambda x_n$, for $\Lambda \in X^*$. Each f_n is weak*-continuous, by the definition of the weak*-topology. If $f_n(\Lambda) = f_n(\Lambda')$ for all n, then $\Lambda x_n = \Lambda' x_n$ for all n, which implies that $\Lambda = \Lambda'$, since both are continuous on X and coincide on a dense set.

Thus $\{f_n\}$ is a countable family of continuous functions that separates points on X^*. The metrizability of K now follows from (c) of Section 3.8. ////

3.17 Theorem *If V is a neighborhood of 0 in a separable topological vector space X, and if $\{\Lambda_n\}$ is a sequence in X^* such that*

$$|\Lambda_n x| \le 1 \qquad (x \in V, n = 1, 2, 3, \ldots),$$

then there is a subsequence $\{\Lambda_{n_i}\}$ and there is a $\Lambda \in X^*$ such that

$$\Lambda x = \lim_{i \to \infty} \Lambda_{n_i} x \qquad (x \in X).$$

In other words, the polar of V is sequentially compact in the weak*-topology.

PROOF. Combine Theorems 3.15 and 3.16. ////

The next application of the Banach-Alaoglu theorem involves the Hahn-Banach theorem and a category argument.

3.18 Theorem *In a locally convex space X, every weakly bounded set is originally bounded, and vice versa.*

Part (d) of Exercise 5 shows that the local convexity of X cannot be omitted from the hypotheses.

PROOF. Since every weak neighborhood of 0 in X is an original neighborhood of 0, it is obvious from the definition of "bounded" that every originally bounded subset of X is weakly bounded. The converse is the nontrivial part of the theorem.

Suppose $E \subset X$ is weakly bounded and U is an original neighborhood of 0 in X.

Since X is locally convex, there is a convex, balanced, original neighborhood V of 0 in X such that $\bar{V} \subset U$. Let $K \subset X^*$ be the polar of V:

$$(1) \qquad K = \{\Lambda \in X^* : |\Lambda x| \leq 1 \quad \text{for all} \quad x \in V\}.$$

We claim that

$$(2) \qquad \bar{V} = \{x \in X : |\Lambda x| \leq 1 \quad \text{for all} \quad \Lambda \in K\}.$$

It is clear that V is a subset of the right side of (2) and hence so is $\bar{V}$, since the right side of (2) is closed. Suppose $x_0 \in X$ but $x_0 \notin \bar{V}$. Theorem 3.7 (with $\bar{V}$ in place of B) then shows that $\Lambda x_0 > 1$ for some $\Lambda \in K$. This proves (2).

Since E is weakly bounded, there corresponds to each $\Lambda \in X^*$ a number $\gamma(\Lambda) < \infty$ such that

$$(3) \qquad |\Lambda x| \leq \gamma(\Lambda) \qquad (x \in E).$$

Since K is convex and weak*-compact (Theorem 3.15) and since the functions $\Lambda \to \Lambda x$ are weak*-continuous, we can apply Theorem 2.9 (with X^* in place of X and the scalar field in place of Y) to conclude from (3) that there is a constant $\gamma < \infty$ such that

$$(4) \qquad |\Lambda x| \leq \gamma \qquad (x \in E, \Lambda \in K).$$

Now (2) and (4) show that $\gamma^{-1}x \in \bar{V} \subset U$ for all $x \in E$. Since V is balanced,

$$(5) \qquad E \subset t\bar{V} \subset tU \qquad (t > \gamma).$$

Thus E is originally bounded. ////

Corollary. *If X is a normed space, if $E \subset X$, and if*

$$(6) \qquad \sup_{x \in E} |\Lambda x| < \infty \qquad (\Lambda \in X^*)$$

then there exists $\gamma < \infty$ such that

$$(7) \qquad \|x\| \leq \gamma \qquad (x \in E).$$

PROOF. Normed spaces are locally convex; (6) says that E is weakly bounded, and (7) says that E is originally bounded. ////

We now turn to the question: What can one say about the convex hull H of a compact set K? Even in a Hilbert space, H need not be closed, and there are situations in which $\bar{H}$ is not compact (Exercises 20, 22). In

Fréchet spaces the latter pathology does not occur (Theorem 3.20). The proof of this will depend on the fact that a subset of a complete metric space is compact if and only if it is closed and totally bounded (Appendix A4).

3.19 Definitions (*a*) If X is a vector space and $E \subset X$, the *convex hull* of E will be denoted by $co(E)$. Recall that $co(E)$ is the intersection of all convex subsets of X which contain E. Equivalently, $co(E)$ is the set of all finite convex combinations of members of E.

(*b*) If X is a topological vector space and $E \subset X$, the *closed convex hull* of E, written $\overline{co}(E)$, is the closure of $co(E)$.

(*c*) A subset E of a metric space X is said to be *totally bounded* if E lies in the union of finitely many open balls of radius ε, for every $\varepsilon > 0$.

The same concept can be defined in any topological vector space, metrizable or not:

(*d*) A set E in a topological vector space X is said to be *totally bounded* if to every neighborhood V of 0 in X corresponds a *finite* set F such that $E \subset F + V$.

If X happens to be a metrizable topological vector space, then these two notions of total boundedness coincide, provided we restrict ourselves to *invariant* metrics that are compatible with the topology of X. (The proof of this is as in Section 1.25.)

3.20 Theorem

(*a*) *If $A_1, \ldots, A_n$ are compact convex sets in a topological vector space X, then $co(A_1 \cup \cdots \cup A_n)$ is compact.*

(*b*) *If X is a locally convex topological vector space and $E \subset X$ is totally bounded, then $co(E)$ is totally bounded.*

(*c*) *If X is a Fréchet space and $K \subset X$ is compact, then $\overline{co}(K)$ is compact.*

(*d*) *If K is a compact set in R^n, then $co(K)$ is compact.*

PROOF. (*a*) Let S be the simplex in R^n consisting of all $s = (s_1, \ldots, s_n)$ with $s_i \geq 0$, $s_1 + \cdots + s_n = 1$. Put $A = A_1 \times \cdots \times A_n$. Define $f: S \times A \to X$ by

$$(1) \qquad f(s, a) = s_1 a_1 + \cdots + s_n a_n$$

and put $K = f(S \times A)$.

It is clear that K is compact and that $K \subset co(A_1 \cup \cdots \cup A_n)$. We will see that this inclusion is actually an equality.

If (s, a) and (t, b) are in $S \times A$ and if $\alpha \geq 0$, $\beta \geq 0$, $\alpha + \beta = 1$, then

$$\alpha f(s, a) + \beta f(t, b) = f(u, c), \tag{2}$$

where $u = \alpha s + \beta t \in S$ and $c \in A$, because

$$c_i = \frac{\alpha s_i a_i + \beta t_i b_i}{\alpha s_i + \beta t_i} \in A_i \qquad (1 \leq i \leq n). \tag{3}$$

This shows that K is convex. Since $A_i \subset K$ for each i [take $s_i = 1$ in (1), $s_j = 0$ for $j \neq i$], the convexity of K implies that $co(A_1 \cup \cdots \cup A_n) \subset K$. This proves (*a*).

(*b*) Let U be a neighborhood of 0 in X. Choose a convex neighborhood V of 0 in X such that $V + V \subset U$. Then $E \subset F + V$ for some finite set $F \subset X$. Hence $E \subset co(F) + V$. The latter set is convex. It follows that

$$co(E) \subset co(F) + V. \tag{4}$$

But $co(F)$ is compact [a special case of (*a*)], and therefore $co(F) \subset F_1 + V$ for some finite set $F_1 \subset X$. Thus

$$co(E) \subset F_1 + V + V \subset F_1 + U. \tag{5}$$

Since U was arbitrary, $co(E)$ is totally bounded.

(*c*) Closures of totally bounded sets are totally bounded in every metric space, and hence are compact in every complete metric space (Appendix A4). So if K is compact in a Fréchet space, then K is obviously totally bounded; hence $co(K)$ is totally bounded, by (*b*), and therefore $\overline{co}(K)$ is compact.

(*d*) Let S be the simplex in R^{n+1} consisting of all $t = (t_1, \ldots, t_{n+1})$ with $t_i \geq 0$ and $\sum t_i = 1$. Let K be compact, $K \subset R^n$. By the proposition that follows, $x \in co(K)$ if and only if

$$x = t_1 x_1 + \cdots + t_{n+1} x_{n+1} \tag{6}$$

for some $t \in S$ and $x_i \in K$ $(1 \leq i \leq n + 1)$. In other words, $co(K)$ is the image of $S \times K^{n+1}$ under the continuous mapping

$$(t, x_1, \ldots, x_{n+1}) \to t_1 x_1 + \cdots + t_{n+1} x_{n+1}. \tag{7}$$

Hence $co(K)$ is compact. ////

Proposition. *If $E \subset R^n$ and $x \in co(E)$, then x lies in the convex hull of some subset of E which contains at most $n + 1$ points.*

PROOF. It is enough to show that if $k > n$ and $x = \sum_1^{k+1} t_i x_i$ is a convex combination of some $k + 1$ vectors $x_i \in R^n$, then x is actually a convex combination of some k of these vectors.

Assume, with no loss of generality, that $t_i > 0$ for $1 \le i \le k + 1$. The null space of the linear map

$$(8) \qquad (a_1, \ldots, a_{k+1}) \to \left(\sum_1^{k+1} a_i x_i, \sum_1^{k+1} a_i \right),$$

which sends R^{k+1} into $R^n \times R$, has positive dimension, since $k > n$. Hence there exists $(a_1, \ldots, a_{k+1})$, with some $a_i \neq 0$, so that $\sum a_i x_i = 0$ and $\sum a_i = 0$. Since $t_i > 0$ for all i, there is a constant λ such that $|\lambda a_i| \le t_i$ for all i and $\lambda a_j = t_j$ for at least one j. Setting $c_i = t_i - \lambda a_i$, we conclude that $x = \sum c_i x_i$ and that at least one c_j is 0; note also that $\sum c_i = \sum t_i = 1$ and that $c_i \ge 0$ for all i. ////

The following analogue of part (*b*) of the separation theorem 3.4 will be used in the proof of the Krein-Milman theorem.

3.21 Theorem *Suppose X is a topological vector space on which X^* separates points. Suppose A and B are disjoint, nonempty, compact, convex sets in X. Then there exists $\Lambda \in X^*$ such that*

$$(1) \qquad \sup_{x \in A} \operatorname{Re} \Lambda x < \inf_{y \in B} \operatorname{Re} \Lambda y.$$

Note that part of the hypothesis is weaker than in (*b*) of Theorem 3.4 (since local convexity of X implies that X^* separates points on X); to make up for this, it is now assumed that *both* A and B are compact.

PROOF. Let X_w be X with its weak topology. The sets A and B are evidently compact in X_w. They are also closed in X_w (because X_w is a Hausdorff space). Since X_w is locally convex, (*b*) of Theorem 3.4 can be applied to X_w in place of X; it gives us a $\Lambda \in (X_w)^*$ that satisfies (1). But we saw in Section 3.11 (as a consequence of Theorem 3.10) that $(X_w)^* = X^*$. ////

3.22 Extreme points Let K be a subset of a vector space X. A nonempty set $S \subset K$ is called an *extreme set* of K if no point of S is an internal point of any line interval whose end points are in K, except when both end points are in S. Analytically, the condition can be expressed as follows: If $x \in K$, $y \in K$, $0 < t < 1$, and

$$(1 - t)x + ty \in S,$$

then $x \in S$ and $y \in S$.

The *extreme points* of K are the extreme sets that consist of just one point.

The set of all extreme points of K will be denoted by $E(K)$.

The following two theorems show that under certain conditions $E(K)$ is quite a large set.

3.23 The Krein-Milman theorem *Suppose X is a topological vector space on which X^* separates points. If K is a nonempty compact convex set in X, then K is the closed convex hull of the set of its extreme points.*

In symbols, $K = \overline{co}(E(K))$.

PROOF. Let $\mathscr{P}$ be the collection of all compact extreme sets of K. Since $K \in \mathscr{P}$, $\mathscr{P} \neq \varnothing$. We shall use the following two properties of $\mathscr{P}$:

(*a*) *The intersection S of any nonempty subcollection of $\mathscr{P}$ is a member of $\mathscr{P}$, unless $S = \varnothing$.*

(*b*) *If $S \in \mathscr{P}$, $\Lambda \in X^*$, μ is the maximum of* Re Λ *on S, and*

$$S_\Lambda = \{x \in S: \text{Re } \Lambda x = \mu\},$$

then $S_\Lambda \in \mathscr{P}$.

The proof of (*a*) is immediate. To prove (*b*), suppose $tx + (1-t)y = z \in S_\Lambda$, $x \in K$, $y \in K$, $0 < t < 1$. Since $z \in S$ and $S \in \mathscr{P}$, we have $x \in S$ and $y \in S$. Hence Re $\Lambda x \leq \mu$, Re $\Lambda y \leq \mu$. Since Re $\Lambda z = \mu$ and Λ is linear, we conclude: Re $\Lambda x = \mu =$ Re Λy. Hence $x \in S_\Lambda$ and $y \in S_\Lambda$. This proves (*b*).

Choose some $S \in \mathscr{P}$. Let $\mathscr{P}'$ be the collection of all members of $\mathscr{P}$ that are subsets of S. Since $S \in \mathscr{P}'$, $\mathscr{P}'$ is not empty. Partially order $\mathscr{P}'$ by set inclusion, let Ω be a maximal totally ordered subcollection of $\mathscr{P}'$, and let M be the intersection of all members of Ω. Since Ω is a collection of compact sets with the finite intersection property, $M \neq \varnothing$. By (*a*), $M \in \mathscr{P}'$. The maximality of Ω implies that no proper subset of M belongs to $\mathscr{P}$. It now follows from (*b*) that every $\Lambda \in X^*$ is constant on M. Since X^* separates points on X, M has only one point. Therefore M is an extreme point of K.

We have now proved that

$$E(K) \cap S \neq \varnothing \tag{1}$$

for every $S \in \mathscr{P}$. In other words, *every compact extreme set of K contains on extreme point of K.*

Since K is compact and convex (the assumed convexity of K will now be used for the first time), we have

$$\overline{co}(E(K)) \subset K \tag{2}$$

and this shows that $\overline{co}(E(K))$ is compact.

Assume, to reach a contradiction, that some $x_0 \in K$ is not in $\overline{co}(E(K))$. Theorem 3.21 furnishes then a $\Lambda \in X^*$ such that

Re $\Lambda x <$ Re Λx_0 for every $x \in \overline{co}(E(K))$. If K_Λ is defined as in (*b*), then $K_\Lambda \in \mathscr{P}$. Our choice of Λ shows that K_Λ is disjoint from $\overline{co}(E(K))$, and this contradicts (1). ////

Remark. The convexity of K was used only to show that $\overline{co}(E(K))$ is compact. If X were assumed to be locally convex, the compactness of $\overline{co}(E(K))$ would not be needed, since one could use (*b*) of Theorem 3.4 in place of Theorem 3.21. The above argument proves then that $K \subset \overline{co}(E(K))$. The following version of the Krein-Milman theorem is thus obtained:

3.24 Theorem *If K is a compact subset of a locally convex space then $K \subset \overline{co}(E(K))$.*

Equivalently, $\overline{co}(K) = \overline{co}(E(K))$.

It may happen in this situation that $\overline{co}(K)$ has extreme points which are not in K. (See Exercise 33.) The next theorem shows that this pathology cannot occur if $\overline{co}(K)$ is compact. Therefore it occurs in no Fréchet space, by (*c*) of Theorem 3.20.

3.25 Milman's theorem *If K is a compact set in a locally convex space X, and if $\overline{co}(K)$ is also compact, then every extreme point of $\overline{co}(K)$ lies in K.*

PROOF. Assume that some extreme point p of $\overline{co}(K)$ is not in K. Then there is a convex balanced neighborhood V of 0 in X such that

$$(1) \qquad (p + \overline{V}) \cap K = \varnothing.$$

Choose $x_1, \ldots, x_n$ in K so that $K \subset \bigcup_1^n (x_i + V)$. Each set

$$(2) \qquad A_i = \overline{co}(K \cap (x_i + V)) \qquad (1 \le i \le n)$$

is convex and also compact, since $A_i \subset \overline{co}(K)$. Also, $K \subset A_1 \cup \cdots \cup A_n$. Part (*a*) of Theorem 3.20 shows therefore that

$$(3) \qquad \overline{co}(K) \subset \overline{co}(A_1 \cup \cdots \cup A_n) = co(A_1 \cup \cdots \cup A_n).$$

But the opposite inclusion holds also, because $A_i \subset \overline{co}(K)$ for each i. Thus

$$(4) \qquad \overline{co}(K) = co(A_1 \cup \cdots \cup A_n).$$

In particular, $p = t_1 y_1 + \cdots + t_N y_N$, where each y_j lies in some A_i, each t_j is positive, and $\sum t_j = 1$. The grouping

$$(5) \qquad p = t_1 y_1 + (1 - t_1) \frac{t_2 y_2 + \cdots + t_N y_N}{t_2 + \cdots + t_N}$$

exhibits p as a convex combination of two points of $\overline{co}(K)$, by (4). Since p is an extreme point of $\overline{co}(K)$, we conclude from (5) that $y_1 = p$. Thus, for some i,

(6) $$p \in A_i \subset x_i + \bar{V} \subset K + \bar{V},$$

which contradicts (1). [Note that $A_i \subset x_i + \bar{V}$, by (2), because V is convex.] ////

Vector-Valued Integration

Sometimes it is desirable to be able to integrate functions f that are defined on some measure space Q (with a real or complex measure μ) and whose values lie in some topological vector space X. The first problem is to associate with these data a vector in X that deserves to be called

$$\int_Q f\, d\mu,$$

i.e., which has at least some of the properties that integrals usually have. For instance, the equation

$$\Lambda\left(\int_Q f\, d\mu\right) = \int_Q (\Lambda f)\, d\mu$$

ought to hold for every $\Lambda \in X^*$, because it does hold for sums, and because integrals are (or ought to be) limits of sums in some sense or other. In fact, our definition will be based on this single requirement.

Many other approaches to vector-valued integration have been studied in great detail; in some of these, the integrals are defined more directly as limits of sums (see Exercise 23).

3.26 Definition Suppose μ is a measure on a measure space Q, X is a topological vector space on which X^* separates points, and f is a function from Q into X such that the scalar functions Λf are integrable with respect to μ, for every $\Lambda \in X^*$; note that Λf is defined by

(1) $$(\Lambda f)(q) = \Lambda(f(q)) \qquad (q \in Q).$$

If there exists a vector $y \in X$ such that

(2) $$\Lambda y = \int_Q (\Lambda f)\, d\mu$$

for every $\Lambda \in X^*$, then we define

(3) $$\int_Q f\, d\mu = y.$$

Remarks. It is clear that there is at most one such y, because X^* separates points on X. *Thus there is no uniqueness problem.*

Existence will be proved only in the rather special case (sufficient for many applications) in which Q is compact and f is continuous. In that case, $f(Q)$ is compact, and the only other requirement that will be imposed is that the closed convex hull of $f(Q)$ should be compact. By Theorem 3.20, *this additional requirement is automatically satisfied when X is a Fréchet space.*

Recall that a *Borel measure* on a compact (or locally compact) Hausdorff space Q is a measure defined on the σ-algebra of all Borel sets in Q; this is the smallest σ-algebra that contains all open subsets of Q. A *probability measure* is a positive measure of total mass 1.

3.27 Theorem *Suppose*

(*a*) *X is a topological vector space on which X^* separates points, and*

(*b*) *μ is a Borel probability measure on a compact Hausdorff space Q.*

If $f: Q \to X$ is continuous and if $\overline{co}(f(Q))$ is compact in X, then the integral

$$y = \int_Q f \, d\mu \tag{1}$$

exists, in the sense of Definition 3.26.

Moreover, $y \in \overline{co}(f(Q))$.

Remark. If ν is any positive Borel measure on Q, then some scalar multiple of ν is a probability measure. The theorem therefore holds (except for its last sentence) with ν in place of μ. It can then be extended to real-valued Borel measures (by the Jordan decomposition theorem) and (if the scalar field of X is $\mathbb{C}$) to complex ones.

Exercise 24 gives another generalization.

PROOF. Regard X as a real vector space. Put $H = co(f(Q))$. We have to prove that there exists $y \in H$ such that

$$\Lambda y = \int_Q (\Lambda f) \, d\mu \tag{2}$$

for *every* $\Lambda \in X^*$.

Let $L = \{\Lambda_1, \ldots, \Lambda_n\}$ be a finite subset of X^*. Let E_L be the set of all $y \in \bar{H}$ that satisfy (2) for every $\Lambda \in L$. Each E_L is closed (by the continuity of Λ) and is therefore *compact*, since $\bar{H}$ is compact. If no E_L

is empty, the collection of all E_L has the finite intersection property. The intersection of all E_L is therefore not empty, and any y in it satisfies (2) for every $\Lambda \in X^*$. It is therefore enough to prove $E_L \neq \varnothing$.

Regard $L = (\Lambda_1, \ldots, \Lambda_n)$ as a mapping from X into R^n, and put $K = L(f(Q))$. Define

$$m_i = \int_Q (\Lambda_i f)\, d\mu \qquad (1 \leq i \leq n). \tag{3}$$

We claim that the point $m = (m_1, \ldots, m_n)$ lies in the convex hull of K.

If $t = (t_1, \ldots, t_n) \in R^n$ is not in this hull, then [by Theorem 3.20 and (*b*) of Theorem 3.4 and the known form of the linear functionals on R^n] there are real numbers $c_1, \ldots, c_n$ such that

$$\sum_{i=1}^{n} c_i u_i < \sum_{i=1}^{n} c_i t_i \tag{4}$$

if $u = (u_1, \ldots, u_n) \in K$. Hence

$$\sum_{i=1}^{n} c_i \Lambda_i f(q) < \sum_{i=1}^{n} c_i t_i \qquad (q \in Q). \tag{5}$$

Since μ is a probability measure, integration of the left side of (5) gives $\sum c_i m_i < \sum c_i t_i$. Thus $t \neq m$.

This shows that m lies in the convex hull of K. Since $K = L(f(Q))$ and L is linear, it follows that $m = Ly$ for some y in the convex hull H of $f(Q)$. For this y we have

$$\Lambda_i y = m_i = \int_Q (\Lambda_i f)\, d\mu \qquad (1 \leq i \leq n). \tag{6}$$

Hence $y \in E_L$. This completes the proof. ////

3.28 Theorem *Suppose*

(*a*) *X is a topological vector space on which X^* separates points,*
(*b*) *Q is a compact subset of X, and*
(*c*) *the closed convex hull $\bar{H}$ of Q is compact.*

Then $y \in \bar{H}$ if and only if there is a regular Borel probability measure μ on Q such that

$$y = \int_Q x\, d\mu(x). \tag{1}$$

Remarks. The integral is to be understood as in Definition 3.26, with $f(x) = x$.

Recall that a positive Borel measure on Q is said to be *regular* if

(2) $$\mu(E) = \sup \{\mu(K): K \subset E\} = \inf \{\mu(G): E \subset G\}$$

for every Borel set $E \subset Q$, where K ranges over the compact subsets of E and G ranges over the open supersets of E.

The integral (1) represents every $y \in \bar{H}$ as a "weighted average" of Q, or as the "center of mass" of a certain unit mass distributed over Q.

We stress once more that (c) follows from (b) if X is a Fréchet space.

PROOF. Regard X again as a real vector space. Let $C(Q)$ be the Banach space of all real continuous functions on Q, with the supremum norm. The Riesz representation theorem identifies the dual space $C(Q)^*$ with the space of all real Borel measures on Q that are differences of regular positive ones. With this identification in mind, we define a mapping

(3) $$\phi: C(Q)^* \to X$$

by

(4) $$\phi(\mu) = \int_Q x \, d\mu(x).$$

Let P be the set of all regular Borel probability measures on Q. The theorem asserts that $\phi(P) = \bar{H}$.

For each $x \in Q$, the unit mass δ_x concentrated at x belongs to P. Since $\phi(\delta_x) = x$, we see that $Q \subset \phi(P)$. Since ϕ is linear and P is convex, it follows that $H \subset \phi(P)$, where H is the convex hull of Q. By Theorem 3.27, $\phi(P) \subset \bar{H}$. Therefore all that remains to be done is to show that $\phi(P)$ is closed in X.

This is a consequence of the following two facts:

(*i*) *P is weak*-compact in $C(Q)^*$.*

(*ii*) *The mapping ϕ defined by* (4) *is continuous if $C(Q)^*$ is given its weak*-topology and if X is given its weak topology.*

Once we have (*i*) and (*ii*), it follows that $\phi(P)$ is weakly compact, hence weakly closed, and since weakly closed sets are strongly closed, we have the desired conclusion.

To prove (*i*), note that

(5) $$P \subset \left\{\mu: \left| \int_Q h \, d\mu \right| \leq 1 \text{ if } \|h\| < 1 \right\}$$

and that this larger set is weak*-compact, by the Banach-Alaoglu theorem. It is therefore enough to show that P is weak*-closed.

If $h \in C(Q)$ and $h \geq 0$, put

$$E_h = \left\{\mu \colon \int_Q h\, d\mu \geq 0\right\}. \tag{6}$$

Since $\mu \to \int h\, d\mu$ is continuous, by the definition of the weak*-topology, each E_h is weak*-closed. So is the set

$$E = \left\{\mu \colon \int_Q 1\, d\mu = 1\right\}. \tag{7}$$

Since P is the intersection of E and the sets E_h, P is weak*-closed.

To prove (*ii*) it is enough to prove that ϕ is continuous at the origin, since ϕ is linear. Every weak neighborhood of 0 in X contains a set of the form

$$W = \{y \in X : |\Lambda_i y| < r_i \quad \text{for} \quad 1 \leq i \leq n\}, \tag{8}$$

where $\Lambda_i \in X^*$ and $r_i > 0$. The restrictions of the Λ_i to Q lie in $C(Q)$. Hence

$$V = \left\{\mu \in C(Q)^* \colon \left|\int_Q \Lambda_i\, d\mu\right| < r_i \quad \text{for} \quad 1 \leq i \leq n\right\} \tag{9}$$

is a weak*-neighborhood of 0 in $C(Q)^*$. But

$$\int_Q \Lambda_i\, d\mu = \Lambda_i\left(\int_Q x\, d\mu(x)\right) = \Lambda_i \phi(\mu), \tag{10}$$

by Definition 3.26. It follows from (8), (9), and (10) that $\phi(V) \subset W$. Hence ϕ is continuous. ////

The following simple inequality sharpens the last assertion in the statement of Theorem 3.27.

3.29 Theorem *Suppose Q is a compact Hausdorff space, X is a Banach space, $f : Q \to X$ is continuous, and μ is a positive Borel measure on Q. Then*

$$\left\|\int_Q f\, d\mu\right\| \leq \int_Q \|f\|\, d\mu.$$

PROOF. Put $y = \int f\, d\mu$. By the corollary to Theorem 3.3, there is a $\Lambda \in X^*$ such that $\Lambda y = \|y\|$ and $|\Lambda x| \leq \|x\|$ for all $x \in X$. In particular,

$$|\Lambda f(s)| \leq \|f(s)\|$$

for all $s \in Q$. By Theorem 3.27, it follows that

$$\|y\| = \Lambda y = \int_Q (\Lambda f)\, d\mu \leq \int_Q \|f\|\, d\mu. \qquad ////$$

Holomorphic Functions

In the study of Banach algebras, as well as in some other contexts, it is useful to enlarge the concept of holomorphic function from complex-valued ones to vector-valued ones. (Of course, one can also generalize the domains, by going from $\mathbb{C}$ to $\mathbb{C}^n$ and even beyond. But this is another story.) There are at least two very natural definitions of "holomorphic" available in this general setting, a "weak" one and a "strong" one. They turn out to define the same class of functions if the values are assumed to lie in a Fréchet space.

3.30 Definition Let Ω be an open set in $\mathbb{C}$ and let X be a complex topological vector space.

(*a*) A function $f: \Omega \to X$ is said to be *weakly holomorphic in* Ω if Λf is holomorphic in the ordinary sense for every $\Lambda \in X^*$.

(*b*) A function $f: \Omega \to X$ is said to be *strongly holomorphic in* Ω if

$$\lim_{w \to z} \frac{f(w) - f(z)}{w - z}$$

exists (in the topology of X) for every $z \in \Omega$.

Note that the above quotient is the product of the scalar $(w - z)^{-1}$ and the vector $f(w) - f(z)$ in X.

The continuity of the functionals Λ that occur in (*a*) makes it obvious that every strongly holomorphic function is weakly holomorphic. The converse is true when X is a Fréchet space, but it is far from obvious. (Recall that weakly convergent sequences may very well fail to converge originally.) The Cauchy theorem will play an important role in this proof, as will Theorem 3.18.

The *index* of a point $z \in \mathbb{C}$ with respect to a closed path Γ that does not pass through z will be denoted by $\text{Ind}_\Gamma(z)$. We recall that

$$\text{Ind}_\Gamma(z) = \frac{1}{2\pi i} \int_\Gamma \frac{d\zeta}{\zeta - z}.$$

All paths considered here and later are assumed to be piecewise continuously differentiable, or at least rectifiable.

3.31 Theorem *Let Ω be open in $\mathbb{C}$, let X be a complex Fréchet space, and assume that*

$$f: \Omega \to X$$

is weakly holomorphic. The following conclusions hold:

(*a*) *f is strongly continuous in Ω.*

(*b*) *The Cauchy theorem and the Cauchy formula hold: If* Γ *is a closed path in* Ω *such that* $\text{Ind}_\Gamma(w) = 0$ *for every* $w \notin \Omega$, *then*

$$\int_\Gamma f(\zeta)\, d\zeta = 0, \tag{1}$$

and

$$f(z) = \frac{1}{2\pi i} \int_\Gamma (\zeta - z)^{-1} f(\zeta)\, d\zeta \tag{2}$$

if $z \in \Omega$ *and* $\text{Ind}_\Gamma(z) = 1$. *If* Γ_1 *and* Γ_2 *are closed paths in* Ω *such that*

$$\text{Ind}_{\Gamma_1}(w) = \text{Ind}_{\Gamma_2}(w)$$

for every $w \notin \Omega$, *then*

$$\int_{\Gamma_1} f(\zeta)\, d\zeta = \int_{\Gamma_2} f(\zeta)\, d\zeta. \tag{3}$$

(*c*) f *is strongly holomorphic in* Ω.

The integrals in (*b*) are to be understood in the sense of Theorem 3.27. Either one can regard $d\zeta$ as a complex measure on the range of Γ (a compact subset of $\mathbb{C}$), or one can parametrize Γ and integrate with respect to Lebesgue measure on a compact interval in R.

PROOF. (*a*) Assume $0 \in \Omega$. We shall prove that f is strongly continuous at 0. Define

$$\Delta_r = \{z \in \mathbb{C}: |z| \le r\}. \tag{4}$$

Then $\Delta_{2r} \subset \Omega$ for some $r > 0$. Let Γ be the positively oriented boundary of Δ_{2r}.

Fix $\Lambda \in X^*$. Since Λf is holomorphic,

$$\frac{(\Lambda f)(z) - (\Lambda f)(0)}{z} = \frac{1}{2\pi i} \int_\Gamma \frac{(\Lambda f)(\zeta)}{(\zeta - z)\zeta}\, d\zeta \tag{5}$$

if $0 < |z| < 2r$. Let $M(\Lambda)$ be the maximum of $|\Lambda f|$ on Δ_{2r}. If $0 < |z| \le r$, it follows that

$$|z^{-1}\Lambda[f(z) - f(0)]| \le r^{-1} M(\Lambda). \tag{6}$$

The set of all quotients

$$\left\{\frac{f(z) - f(0)}{z} : 0 < |z| \le r\right\} \tag{7}$$

is therefore weakly bounded in X. By Theorem 3.18, this set is also strongly bounded. Thus if V is any (strong) neighborhood of 0 in X, there exists $t < \infty$ such that

$$f(z) - f(0) \in ztV \qquad (0 < |z| \leq r). \tag{8}$$

Consequently, $f(z) \to f(0)$ strongly, as $z \to 0$. [It may be of some interest to observe that the proof of (*a*) used only the local convexity of X. Neither metrizability nor completeness has played a role so far.]

This was the crux of the matter. The rest is now almost automatic.

(*b*) By (*a*) and Theorem 3.27, the integrals in (1) to (3) exist. These three formulas are correct (by the theory of ordinary holomorphic functions) if f is replaced in them by Λf, where Λ is any member of X^*. The formulas are therefore correct as stated, by Definition 3.26.

(*c*) Assume, as in the proof of (*a*), that $\Delta_{2r} \subset \Omega$, and choose Γ as in (*a*). Define

$$y = \frac{1}{2\pi i} \int_\Gamma \zeta^{-2} f(\zeta)\, d\zeta. \tag{9}$$

The Cauchy formula (2) shows, after a small computation, that

$$\frac{f(z) - f(0)}{z} = y + zg(z) \tag{10}$$

if $0 < |z| < r$, where

$$g(z) = \frac{1}{2\pi} \int_{-\pi}^{\pi} [2re^{i\theta}(2re^{i\theta} - z)]^{-1} f(2re^{i\theta})\, d\theta. \tag{11}$$

Let V be a convex balanced neighborhood of 0 in X. Put $K = \{f(\zeta) \colon |\zeta| = 2r\}$. Then K is compact, so that $K \subset tV$ for some $t < \infty$. If $s = tr^{-2}$ and $|z| \leq r$, it follows that the integrand (11) lies in sV for every θ. Thus $g(z) \in s\bar{V}$ if $|z| \leq r$. The left side of (10) therefore converges strongly to y, as $z \to 0$. ////

The following extension of Liouville's theorem concerning bounded entire functions does not even depend on Theorem 3.31. It can be used in the study of spectra in Banach algebras. (See Exercise 10, Chapter 10.)

3.32 Theorem *Suppose X is a complex topological vector space on which X^* separates points. Suppose $f \colon \mathbb{C} \to X$ is weakly holomorphic and $f(\mathbb{C})$ is a weakly bounded subset of X. Then f is constant.*

PROOF. For every $\Lambda \in X^*$, Λf is a bounded (complex-valued) entire function. If $z \in \mathcal{C}$, it follows from Liouville's theorem that

$$\Lambda f(z) = \Lambda f(0).$$

Since X^* separates points on X, this implies $f(z) = f(0)$, for every $z \in \mathcal{C}$. ////

Part (*d*) of Exercise 5 describes a weakly bounded set which is not originally bounded, in an F-space X on which X^* separates points. Compare with Theorem 3.18.

Exercises

1. Call a set $H \subset R^n$ a *hyperplane* if there exist real numbers $a_1, \ldots, a_n, c$ (with $a_i \neq 0$ for at least one i) such that H consists of all points $x = (x_1, \ldots, x_n)$ that satisfy $\sum a_i x_i = c$.

Suppose E is a convex set in R^n, with nonempty interior, and y is a boundary point of E. Prove that there is a hyperplane H such that $y \in H$ and E lies entirely on one side of H. (State the conclusion more precisely.) *Suggestion:* Suppose 0 is an interior point of E, let M be the one-dimensional subspace that contains y, and apply Theorem 3.2.

2. Suppose $L^2 = L^2([-1, 1])$, with respect to Lebesgue measure. For each scalar α, let E_α be the set of all continuous functions f on $[-1, 1]$ such that $f(0) = \alpha$. Show that each E_α is convex and that each is dense in L^2. Thus E_α and E_β are disjoint convex sets (if $\alpha \neq \beta$) which cannot be separated by any continuous linear functional Λ on L^2. *Hint:* What is $\Lambda(E_\alpha)$?

3. Suppose X is a real vector space (without topology). Call a point $x_0 \in A \subset X$ an *internal* point of A if $A - x_0$ is an absorbing set.

(*a*) Suppose A and B are disjoint convex sets in X, and A has an internal point. Prove that there is a nonconstant linear functional Λ on X such that $\Lambda(A) \cap \Lambda(B)$ contains at most one point. (The proof is similar to that of Theorem 3.4.)

(*b*) Show (with $X = R^2$, for example) that it may not be possible to have $\Lambda(A)$ and $\Lambda(B)$ disjoint, under the hypotheses of (*a*).

4. Let ℓ^∞ be the space of all real bounded functions x on the positive integers. Let τ be the translation operator defined on ℓ^∞ by the equation

$$(\tau x)(n) = x(n+1) \qquad (n = 1, 2, 3, \ldots).$$

Prove that there exists a linear functional Λ on ℓ^∞ (called a *Banach limit*) such that

(*a*) $\Lambda \tau x = \Lambda x$, and

(*b*) $\liminf_{n \to \infty} x(n) \le \Lambda x \le \limsup_{n \to \infty} x(n)$

for every $x \in \ell^\infty$.

Suggestion: Define

$$\Lambda_n x = \frac{x(1) + \cdots + x(n)}{n}$$

$$M = \{x \in \ell^\infty : \lim_{n\to\infty} \Lambda_n x = \Lambda x \text{ exists}\}$$

$$p(x) = \limsup_{n\to\infty} \Lambda_n x$$

and apply Theorem 3.2.

5. For $0 < p < \infty$, let ℓ^p be the space of all functions x (real or complex, as the case may be) on the positive integers, such that

$$\sum_{n=1}^{\infty} |x(n)|^p < \infty.$$

For $1 \le p < \infty$, define $\|x\|_p = \{\sum |x(n)|^p\}^{1/p}$, and define $\|x\|_\infty = \sup_n |x(n)|$.

(*a*) Assume $1 \le p < \infty$. Prove that $\|x\|_p$ and $\|x\|_\infty$ make ℓ^p and ℓ^∞ into Banach spaces. If $p^{-1} + q^{-1} = 1$, prove that $(\ell^p)^* = \ell^q$, in the following sense: There is a one-to-one correspondence $\Lambda \leftrightarrow y$ between $(\ell^p)^*$ and ℓ^q, given by

$$\Lambda x = \sum x(n)y(n) \qquad (x \in \ell^p).$$

(*b*) Assume $1 < p < \infty$ and prove that ℓ^p contains sequences that converge weakly but not strongly.

(*c*) On the other hand, prove that every weakly convergent sequence in ℓ^1 converges strongly, in spite of the fact that the weak topology of ℓ^1 is different from its strong topology (which is induced by the norm).

(*d*) If $0 < p < 1$, prove that ℓ^p, metrized by

$$d(x, y) = \sum_{n=1}^{\infty} |x(n) - y(n)|^p,$$

is a locally bounded F-space which is not locally convex but that $(\ell^p)^*$ nevertheless separates points on ℓ^p. (Thus there are many convex open sets in ℓ^p but not enough to form a base for its topology.) Show that $(\ell^p)^* = \ell^\infty$, in the same sense as in (*a*). Show also that the set of all x with $\Sigma|x(n)| < 1$ is weakly bounded but not originally bounded.

(*e*) For $0 < p \le 1$, let τ_p be the weak*-topology induced on ℓ^∞ by ℓ^p; see (*a*) and (*d*). If $0 < p < r \le 1$, show that τ_p and τ_r are *different* topologies (is one weaker than the other?) but that they induce the same topology on each norm-bounded subset of ℓ^∞. *Hint:* The norm-closed unit ball of ℓ^∞ is weak*-compact.

6. Put $f_n(t) = e^{int}$ $(-\pi \le t \le \pi)$; let $L^p = L^p(-\pi, \pi)$, with respect to Lebesgue measure. If $1 \le p < \infty$, prove that $f_n \to 0$ weakly in L^p, but not strongly.

7. $L^\infty([0, 1])$ has its norm topology ($\|f\|_\infty$ is the essential supremum of $|f|$) and its weak*-topology as the dual of L^1. Show that C, the space of all continuous functions on $[0, 1]$, is dense in L^∞ in one of these topologies but not in the other. (Compare with the corollaries to Theorem 3.12.) Show the same with "closed" in place of "dense."

8. Let C be the Banach space of all complex continuous functions on $[0, 1]$, with the supremum norm. Let B be the closed unit ball of C. Show that there exist continuous linear functionals Λ on C for which $\Lambda(B)$ is an *open* subset of the complex plane; in particular, $|\Lambda|$ attains no maximum on B.

9. Let $E \subset L^2(-\pi, \pi)$ be the set of all functions

$$f_{m,n}(t) = e^{imt} + me^{int},$$

where m, n are integers and $0 \le m < n$. Let E_1 be the set of all $g \in L^2$ such that some sequence in E converges weakly to g. (E_1 is called the *weak sequential closure* of E.)
   (*a*) Find all $g \in E_1$.
   (*b*) Find all g in the weak closure $\bar{E}_w$ of E.
   (*c*) Show that $0 \in \bar{E}_w$ but 0 is not in E_1, although 0 lies in the weak sequential closure of E_1.

   This example shows that a weak sequential closure need not be weakly sequentially closed. The passage from a set to its weak sequential closure is therefore not a closure operation, in the sense in which that term is usually used in topology. (See also Exercise 28.)

10. Represent ℓ^1 as the space of all real functions x on $S = \{(m, n): m \ge 1, n \ge 1\}$, such that

$$\|x\|_1 = \sum |x(m, n)| < \infty.$$

Let c_0 be the space of all real functions y on S such that $y(m, n) \to 0$ as $m + n \to \infty$, with norm $\|y\|_\infty = \sup |y(m, n)|$.

   Let M be the subspace of ℓ^1 consisting of all $x \in \ell^1$ that satisfy the equations

$$mx(m, 1) = \sum_{n=2}^{\infty} x(m, n) \qquad (m = 1, 2, 3, \ldots).$$

   (*a*) Prove that $\ell^1 = (c_0)^*$. (See also Exercise 24, Chapter 4.)
   (*b*) Prove that M is a norm-closed subspace of ℓ^1.
   (*c*) Prove that M is weak*-dense in ℓ^1 [relative to the weak*-topology given by (*a*)].
   (*d*) Let B be the norm-closed unit ball of ℓ^1. In spite of (*c*), prove that the weak*-closure of $M \cap B$ contains no ball. *Suggestion:* If $\delta > 0$ and $m > 2/\delta$, then

$$|x(m, 1)| \le \frac{\|x\|}{m} < \frac{\delta}{2}$$

   if $x \in M \cap B$, although $x(m, 1) = \delta$ for some $x \in \delta B$. Thus δB is not in the weak*-closure of $M \cap B$. Extend this to balls with other centers.
   (*e*) Put $x_0(m, 1) = m^{-2}$, $x_0(m, n) = 0$ when $n \ge 2$. Prove that no *sequence* in M is weak*-convergent to x_0, in spite of (*c*). *Hint:* Weak*-convergence of $\{x_j\}$ to x_0 implies that $x_j(m, n) \to x_0(m, n)$ for all m, n, as $j \to \infty$, and that $\{\|x_j\|_1\}$ is bounded.

11. Let X be an infinite-dimensional Fréchet space. Prove that X^*, with its weak*-topology, is of the first category in itself.

12. Show that the norm-closed unit ball of c_0 is not weakly compact; recall that $(c_0)^* = \ell^1$ (Exercise 10).

13. Put $f_N(t) = N^{-1} \sum_{n=1}^{N^2} e^{int}$. Prove that $f_N \to 0$ weakly in $L^2(-\pi, \pi)$.

By Theorem 3.13, some sequence of convex combinations of the f_N converges to 0 in the L^2-norm. Find such a sequence. Show that $g_N = N^{-1}(f_1 + \cdots + f_N)$ will not do.

14. (*a*) Suppose Ω is a locally compact Hausdorff space. For each compact $K \subset \Omega$ define a seminorm p_K on $C(\Omega)$, the space of all complex continuous functions on Ω, by

$$p_K(f) = \sup \{|f(x)| : x \in K\}.$$

Give $C(\Omega)$ the topology induced by this collection of seminorms. Prove that to every $\Lambda \in C(\Omega)^*$ correspond a compact $K \subset \Omega$ and a complex Borel measure μ on K such that

$$\Lambda f = \int_K f \, d\mu \qquad [f \in C(\Omega])].$$

(*b*) Suppose Ω is an open set in $\mathbb{C}$. Find a countable collection Γ of measures with compact support in Ω such that $H(\Omega)$ (the space of all holomorphic functions in Ω) consists of exactly those $f \in C(\Omega)$ which satisfy $\int f \, d\mu = 0$ for every $\mu \in \Gamma$.

15. Let X be a topological vector space on which X^* separates points. Prove that the weak*-topology of X^* is metrizable if and only if X has a finite or countable Hamel basis. (See Exercise 1, Chapter 2 for the definition.)

16. Prove that the closed unit ball of L^1 (relative to Lebesgue measure on the unit interval) has no extreme points but that every point on the "surface" of the unit ball in L^p $(1 < p < \infty)$ is an extreme point of the ball.

17. Determine the extreme points of the closed unit ball of C, the space of all continuous functions on the unit interval, with the supremum norm. (The answer depends on the choice of the scalar field.)

18. Let K be the smallest convex set in R^3 that contains the points (1, 0, 1), (1, 0, − 1), and ($\cos \theta$, $\sin \theta$, 0), for $0 \le \theta \le 2\pi$. Show that K is compact but that the set of all extreme points of K is not compact. Does such an example exist in R^2?

19. Suppose K is a compact convex set in R^n. Prove that every $x \in K$ is a convex combination of at most $n + 1$ extreme points of K. *Suggestion:* Use induction on n. Draw a line from some extreme point of K through x to where it leaves K. Use Exercise 1.

20. Let $\{u_1, u_2, u_3, \ldots\}$ be a sequence of pairwise orthogonal unit vectors in a Hilbert space. Let K consist of the vectors 0 and $n^{-1}u_n$ $(n \ge 1)$. Show that (*a*) K is compact; (*b*) $co(K)$ is bounded; (*c*) $co(K)$ is not closed. Find all extreme points of $\overline{co}(K)$.

21. If $0 < p < 1$, every $f \in L^p$ (except $f = 0$) is the arithmetic mean of two functions whose distance from 0 is less than that of f. (See Section 1.47.) Use this to construct an explicit example of a countable compact set K in L^p (with 0 as its only limit point) which has no extreme point.

22. If $0 < p < 1$, show that ℓ^p contains a compact set K whose convex hull is unbounded. This happens in spite of the fact that $(\ell^p)^*$ separates points on ℓ^p;

see Exercise 5. *Suggestion:* Define $x_n \in \ell^p$ by

$$x_n(n) = n^{p-1}, \qquad x_n(m) = 0 \qquad \text{if } m \neq n.$$

Let K consist of $0, x_1, x_2, x_3, \ldots$. If

$$y_N = N^{-1}(x_1 + \cdots + x_N),$$

show that $\{y_N\}$ is unbounded in ℓ^p.

23. Suppose μ is a Borel probability measure on a compact Hausdorff space Q, X is a Fréchet space, and $f: Q \to X$ is continuous. A *partition* of Q is, by definition, a finite collection of disjoint Borel subsets of Q whose union is Q. Prove that to every neighborhood V of 0 in X there corresponds a partition $\{E_i\}$ such that the difference

$$z = \int_Q f\, d\mu - \sum_i \mu(E_i) f(s_i)$$

lies in V for every choice of $s_i \in E_i$. (This exhibits the integral as a strong limit of "Riemann sums.") *Suggestion:* Take V convex and balanced. If $\Lambda \in X^*$ and if $|\Lambda x| \leq 1$ for every $x \in V$, then $|\Lambda z| \leq 1$, provided that the sets E_i are chosen so that $f(s) - f(t) \in V$ whenever s and t lie in the same E_i.

24. In addition to the hypotheses of Theorem 3.27, assume that T is a continuous linear mapping of X into a topological vector space Y on which Y^* separates points, and prove that

$$T \int_Q f\, d\mu = \int_Q (Tf)\, d\mu.$$

Hint: $\Lambda T \in X^*$ for every $\Lambda \in Y^*$.

25. Let E be the set of all extreme points of a compact set K in a topological vector space X on which X^* separates points. Prove that to every $y \in K$ corresponds a regular Borel probability measure μ on $Q = \bar{E}$ such that

$$y = \int_Q x\, d\mu(x).$$

26. Suppose Ω is a region in $\mathbb{C}$, X is a Fréchet space, and $f: \Omega \to X$ is holomorphic.

(*a*) State and prove a theorem concerning the power series representation of f, that is, concerning the formula $f(z) = \sum (z - a)^n c_n$, where $c_n \in X$.

(*b*) Generalize Morera's theorem to X-valued holomorphic functions.

(*c*) For a sequence of complex holomorphic functions in Ω, uniform convergence on compact subsets of Ω implies that the limit is holomorphic. Does this generalize to X-valued holomorphic functions?

27. Suppose $\{\alpha_i\}$ is a bounded set of distinct complex numbers, $f(z) = \sum_0^\infty c_n z^n$ is an entire function with every $c_n \neq 0$, and

$$g_i(z) = f(\alpha_i z).$$

Prove that the vector space generated by the functions g_i is dense in the Fréchet space $H(\mathbb{C})$ defined in Section 1.45.

Suggestion: Assume μ is a measure with compact support such that $\int g_i \, d\mu = 0$ for all i. Put

$$\phi(w) = \int f(wz) \, d\mu(z) \qquad (w \in \mathcal{C}).$$

Prove that $\phi(w) = 0$ for all w. Deduce that $\int z^n \, d\mu(z) = 0$ for $n = 1, 2, 3, \ldots$. Use Exercise 14.

Describe the closed subspace of $H(\mathcal{C})$ generated by the functions g_i if some of the c_n are 0.

28. Suppose X is a Fréchet space (or, more generally, a metrizable locally convex space). Prove the following statements:

(*a*) X^* is the union of countably many weak*-compact sets E_n.

(*b*) If X is separable, each E_n is metrizable. The weak*-topology of X^* is therefore separable, and some countable subset of X^* separates points on X. (Compare with Exercise 15.)

(*c*) If K is a weakly compact subset of X and if $x_0 \in K$ is a weak limit point of some countable set $E \subset K$, then there is a sequence $\{x_n\}$ in E which converges weakly to x_0. *Hint:* Let Y be the smallest closed subspace of X that contains E. Apply (*b*) to Y to conclude that the weak topology of $K \cap Y$ is metrizable.

Remark: The point of (*c*) is the existence of convergent *subsequences* rather than *subnets*. Note that there exist compact Hausdorff spaces in which no sequence of distinct points converges. For an example, see Exercise 18, Chapter 11.

29. Let $C(K)$ be the Banach space of all continuous complex functions on the compact Hausdorff space K, with the supremum norm. For $p \in K$, define $\Lambda_p \in C(K)^*$ by $\Lambda_p f = f(p)$. Show that $p \to \Lambda_p$ is a homeomorphism of K into $C(K)^*$, equipped with its weak*-topology. Part (*c*) of Exercise 28 can therefore not be extended to weak*-compact sets.

30. Suppose that p is an extreme point of some convex set K, and that $p = t_1 x_1 + \cdots + t_n x_n$, where $\sum t_i = 1$, $t_i > 0$ and $x_i \in K$ for all i. Prove that $x_i = p$ for all i.

31. Suppose that $A_1, \ldots, A_n$ are convex sets in a vector space X. Prove that every $x \in co(A_1 \cup \cdots \cup A_n)$ can be represented in the form

$$x = t_1 a_1 + \cdots + t_n a_n,$$

with $a_i \in A_i$ and $t_i \geq 0$ for all i, $\sum t_i = 1$.

32. Let X be an infinite-dimensional Banach space and let $S = \{x \in X : \|x\| = 1\}$ be the unit sphere of X. We want to cover S with finitely many closed balls, none of which contains the origin of X. Can this be done in (*a*) every X, (*b*) some X, (*c*) no X?

33. Let $C(I)$ be the Banach space of all continuous complex functions on the closed unit interval I, with the supremum norm. Let $M = C(I)^*$, the space of all complex Borel measures on I. Give M the weak*-topology induced by $C(I)$.

For each $t \in I$, let $e_t \in M$ be the "evaluation functional" defined by $e_t f = f(t)$, and define $\Lambda \in M$ by $\Lambda f = \int_0^1 f(s) \, ds$.

(*a*) Show that $t - e_t$ is a continuous map from I into M and that $K = \{e_t : t \in I\}$ is a compact set in M.

(*b*) Show that $\Lambda \in \overline{co}(K)$.

(*c*) Find all $\mu \in \overline{co}(K)$.

(*d*) Let X be the subspace of M consisting of all finite linear combinations

$$c_0 \Lambda + c_1 e_{t_1} + \cdots + c_n e_{t_n}$$

with complex coefficients c_j. Note that $co(K) \subset X$ and that $X \cap \overline{co}(K)$ is the closed convex hull of K *within* X. Prove that Λ is an extreme point of $X \cap \overline{co}(K)$, even though Λ is not in K.