

Anthony W. Knapp

Advanced Real Analysis

Along with a companion volume
Basic Real Analysis

Selected Pages from Chapter IV: pp. 105–140

Birkhäuser
Boston • Basel • Berlin

Anthony W. Knapp
81 Upper Sheep Pasture Road
East Setauket, NY, 11733-1729
U.S.A.
e-mail to: aknapp@math.sunysb.edu
<http://www.math.sunysb.edu/~aknapp/books/advanced.html>

Cover design by Mary Burgess.

Mathematics Subject Classification (2000): 46-01, 42-01, 43-01, 35-01,
34-01, 47-01, 58-01, 60A99, 28C10

Library of Congress Cataloging-in-Publication Data

Knapp, Anthony W.

Advanced real analysis: along with a companion volume Basic real analysis / Anthony
W. Knapp

p. cm. – (Cornerstones)

Includes bibliographical references and index.

ISBN 0-8176-4382-6 (alk. paper)

I. Mathematical analysis. I. Title. II. Cornerstones (Birkhäuser)

QA300.K56 2005

515—dc22

2005048070

ISBN-10 0-8176-4382-6

eISBN 0-8176-4442-3

Printed on acid-free paper.

ISBN-13 978-0-8176-4382-9

Basic Real Analysis

ISBN 0-8176-3250-6

Basic Real Analysis and Advanced Real Analysis (Set)

ISBN 0-8176-4407-5

©2005 Anthony W. Knapp

All rights reserved. This work may not be translated or copied in whole or in part without the written permission of the publisher (Birkhäuser Boston, c/o Springer Science+Business Media Inc., 233 Spring Street, New York, NY 10013, USA) and the author, except for brief excerpts in connection with reviews or scholarly analysis. Use in connection with any form of information storage and retrieval, electronic adaptation, computer software, or by similar or dissimilar methodology now known or hereafter developed is forbidden.

The use in this publication of trade names, trademarks, service marks and similar terms, even if they are not identified as such, is not to be taken as an expression of opinion as to whether or not they are subject to proprietary rights.

Printed in the United States of America. (MP)

9 8 7 6 5 4 3 2 1

SPIN 11372219

Birkhäuser is a part of *Springer Science+Business Media*

www.birkhauser.com

CHAPTER IV

Topics in Functional Analysis

Abstract. This chapter pursues three lines of investigation in the subject of functional analysis—one involving smooth functions and distributions, one involving fixed-point theorems, and one involving spectral theory.

Section 1 introduces topological vector spaces. These are real or complex vector spaces with a Hausdorff topology in which addition and scalar multiplication are continuous. Examples include normed linear spaces, spaces given by a separating family of countably many seminorms, and weak and weak-star topologies in the context of Banach spaces. Various general properties of topological vector spaces are proved, and it is proved that the quotient of a topological vector space by a closed vector subspace is Hausdorff and is therefore a topological vector space.

Section 2 introduces a topology on the space $C^\infty(U)$ of smooth functions on an open subset of \mathbb{R}^N . The support of a continuous linear functional on $C^\infty(U)$ is defined and shown to be a compact subset of U . Accordingly, the continuous linear functionals are called distributions of compact support.

Section 3 studies weak and weak-star topologies in more detail. The main result is Alaoglu's Theorem, which says that the closed unit ball in the weak-star topology on the dual of a normed linear space is compact. In an earlier chapter a preliminary form of this theorem was used to construct elements in a dual space as limits of weak-star convergent subsequences.

Section 4 follows Alaoglu's Theorem along a particular path, giving what amounts to a first example of the Gelfand theory of Banach algebras. The relevant theorem, known as the Stone Representation Theorem, says that conjugate-closed uniformly closed subalgebras containing the constants in $B(S)$ are isomorphic via a norm-preserving algebra isomorphism to the space of all continuous functions on some compact Hausdorff space. The compact space in question is the space of multiplicative linear functionals on the subalgebra, and the proof of compactness uses Alaoglu's Theorem.

Sections 5–6 return to the lines of study toward distributions and fixed-point theorems. Section 5 studies the relationship between convexity and the existence of separating linear functionals. The main theorem makes use of the Hahn–Banach Theorem. Section 6 introduces locally convex topological vector spaces. Application of the basic separation theorem from the previous section shows the existence of many continuous linear functionals on such a space.

Section 7 specializes to the line of study via smooth functions and distributions. The topic is the introduction of a certain locally convex topology on the space $C_{\text{com}}^\infty(U)$ of smooth functions of compact support on U . This is best characterized by a universal mapping property introduced in the section.

Sections 8–9 pursue locally convex spaces along the other line of study that split off in Section 5. Section 8 gives the Krein–Milman Theorem, which asserts the existence of a supply of extreme points for any nonempty compact convex set in a locally convex topological vector space. Section 9 relates compact convex sets to the subject of fixed-point theorems.

Section 10 takes up the abstract theory of Banach algebras, with particular attention to commutative C^* algebras with identity. Three examples are the algebras characterized by the Stone Representation Theorem, any L^∞ space, and any adjoint-closed commutative Banach algebra consisting of bounded linear operators on a Hilbert space and containing the identity.

Section 11 continues the investigation of the last of the examples in the previous section and derives the Spectral Theorem for bounded self-adjoint operators and certain related families of operators. Powerful applications follow from a functional calculus implied by the Spectral Theorem. The section concludes with remarks about the Spectral Theorem for unbounded self-adjoint operators.

1. Topological Vector Spaces

In this section we shall work with vector spaces over \mathbb{R} or \mathbb{C} , and the distinction between the two fields will not be very important. We write \mathbb{F} for this field of scalars. A **topological vector space** or **linear topological space** is a vector space X over \mathbb{F} with a Hausdorff topology such that addition, as a mapping $X \times X \rightarrow X$, and scalar multiplication, as a mapping $\mathbb{F} \times X \rightarrow X$, are continuous. The mappings that we study between topological vector spaces are the continuous linear functions, which may be referred to as “continuous linear operators.” An **isomorphism** of topological vector spaces over \mathbb{F} is a continuous linear operator with a continuous inverse.

The simplest examples of topological vector spaces are the spaces \mathbb{F}^N of column vectors with the usual metric topology. Since the topologies of \mathbb{F}^N , $\mathbb{F}^N \times \mathbb{F}^N$, and $\mathbb{F} \times \mathbb{F}^N$ are given by metrics, continuity of functions defined on any of these spaces may be tested by sequences. In particular, continuity of the vector-space operations on \mathbb{F}^N reduces to the familiar results about limits of sums of vectors and limits of scalars times vectors. Moreover, if $L : \mathbb{F}^N \rightarrow Y$ is any linear function from \mathbb{F}^N into a topological vector space over \mathbb{F} , then L is continuous. To see this, let $\{e_1, \dots, e_N\}$ be the standard basis of column vectors, and let (\cdot, \cdot) be the standard inner product on \mathbb{F}^N , namely the dot product if $\mathbb{F} = \mathbb{R}$ and the usual Hermitian inner product if $\mathbb{F} = \mathbb{C}$. Write $y_j = L(e_j)$. For any x in \mathbb{F}^N , we have

$$L(x) = \sum_{j=1}^N (x, e_j)L(e_j) = \sum_{j=1}^N (x, e_j)y_j.$$

If $\{x_n\}$ is a sequence converging to x in \mathbb{F}^N , then the continuity of the inner product forces $(x_n, e_j) \rightarrow (x, e_j)$ for each j . Then $L(x_n)$ tends to $L(x)$ in Y since the vector space operations are continuous in Y . Hence L is continuous.

A second class of examples is the class of normed linear spaces. These were defined in *Basic*, and the continuity of the operations was established there.¹ The spaces \mathbb{F}^N of column vectors are examples. Further examples include the space $B(S)$ of all bounded scalar-valued functions on a nonempty set S with the supremum norm, the vector subspace $C(S)$ of continuous members of $B(S)$ when S is a topological space, the vector subspaces $C_{\text{com}}(S)$ and $C_0(S)$ of continuous functions of compact support and of continuous functions vanishing at infinity when S is locally compact Hausdorff, the space $L^p(X, \mu)$ for $1 \leq p \leq \infty$ when (X, μ) is a measure space, and the space $M(S)$ of finite regular Borel complex measures on a locally compact Hausdorff space with the total variation norm.

A wider class of examples, which includes the normed linear spaces, is the class of topological vector spaces defined by seminorms. Seminorms were defined in Section III.1. If we have a family $\{\|\cdot\|_s\}$ of seminorms on a vector space X over \mathbb{F} , with indexing given by s in some nonempty set S , the corresponding topology on X is defined as the weak topology determined by all functions $x \mapsto \|x - y\|_s$ for $s \in S$ and $y \in X$. A base for the open sets of X is obtained as follows: For each triple (y, s, r) , with y in X , with s one of the seminorm indices, and with $r > 0$, the set $\{x \mid \|x - y\|_s < r\}$ is to be in the base, and the base consists of all finite intersections of these sets as (y, s, r) varies.

In order to obtain a topological vector space from a system of seminorms, we must ensure the Hausdorff property, and we do so by insisting that the only f in X with $\|f\|_s = 0$ for all s is $f = 0$. In this case the family of seminorms is called a **separating family**. Let us go through the argument that a space defined by a separating family of seminorms is a topological vector space.

Proposition 4.1. Let X be a vector space over \mathbb{F} endowed with a separating family $\{\|\cdot\|_s\}$ of seminorms. Then the weak topology determined by all functions $x \mapsto \|x - y\|_s$ makes X into a topological vector space.

PROOF. To see that X is Hausdorff, let x_0 and y_0 be distinct points of X . By assumption, there exists some s such that $\|x_0 - y_0\|_s$ is a positive number r . The sets $\{x \mid \|x - x_0\|_s < r/2\}$ and $\{y \mid \|y - y_0\|_s < r/2\}$ are disjoint and open, and they contain x_0 and y_0 , respectively. Hence X is Hausdorff.

To see that addition is continuous, we are to show that if a net $\{(x_\alpha, y_\alpha)\}$ is convergent in $X \times X$ to (x_0, y_0) , then $\{x_\alpha + y_\alpha\}$ converges to $x_0 + y_0$. This means that if $\|x_\alpha - x_0\|_s + \|y_\alpha - y_0\|_s$ tends to 0 for each s , then $\|(x_\alpha + y_\alpha) - (x_0 + y_0)\|_s$ tends to 0 for each s . This is immediate from the triangle inequality for the seminorm $\|\cdot\|_s$, and hence addition is continuous. The proof that scalar multiplication is continuous is similar.

¹The definition appears in Section V.9 of *Basic*, and the continuity of the operations is proved in Proposition 5.55.

We have encountered two distinctly different kinds of examples of topological vector spaces defined by families of seminorms. In the first kind a countable family of seminorms suffices to define the topology. Normed linear spaces are examples. So is the Schwartz space $\mathcal{S}(\mathbb{R}^N)$, consisting of all smooth scalar-valued functions on \mathbb{R}^N such that the product of any polynomial with any iterated partial derivative of the function is bounded. The defining seminorms for the Schwartz space are

$$\|f\|_{P,Q} = \sup_{x \in \mathbb{R}^N} |P(x)(Q(D)f)(x)|,$$

where P and Q are arbitrary polynomials. We saw in Section III.1 that the same topology arises if we use only the countably many seminorms for which P is some monomial x^α and Q is some monomial x^β . This family of seminorms is a separating family because if $\|f\|_{1,1} = 0$, then $f = 0$.

Another example of a topological vector space whose topology can be defined by countably many seminorms is the space $C^\infty(U)$ of smooth scalar-valued functions on a nonempty open set U of \mathbb{R}^N with the topology of uniform convergence on compact sets of all derivatives. The family of seminorms is indexed by pairs (K, P) with K a compact subset of U and with P a polynomial, the corresponding seminorm being $\|f\|_{K,P} = \sup_{x \in K} |(P(D)f)(x)|$. The Hausdorff condition is satisfied because if $\|f\|_{K,1} = 0$ for all K , then $f = 0$. We shall see in the next section that the topology can be defined by a countable subfamily of these seminorms.

Still a third space of smooth scalar-valued functions, besides $\mathcal{S}(\mathbb{R}^N)$ and $C^\infty(U)$, will be of interest to us. This is the space $C_{\text{com}}^\infty(U)$ of smooth functions on a nonempty open U with compact support contained in U . The useful topology on this space is more complicated than the topologies considered so far. In particular, it cannot be given by countably many seminorms. Describing the topology requires some preparation, and we come back to the details in Section 7.

The examples we have encountered of topological vector spaces defined by an uncountable family of seminorms, but not definable by a countable family, are qualitatively different from the examples above. Indeed, they lead along a different theoretical path, as we shall see—one that takes us in the direction of spectral theory rather than distribution theory.

The first class of such examples is the class of normed linear spaces X with the “weak topology,” as contrasted with the norm topology. Let X^* be the set of linear functionals of X that are continuous in the norm topology. The **weak topology** on X was defined in Chapter X of *Basic* as the weakest topology that makes all members of X^* continuous. Of course, any set that is open in the weak topology on X is open in the norm topology. A base for the open sets in the weak topology on X is obtained as follows: For each triple (x_0, x^*, r) , with x_0 in X , x^* in X^* , and $r > 0$, the set $\{x \mid |x^*(x - x_0)| < r\}$ is to be in the base, and the base

consists of all finite intersections of these sets as (x_0, x^*, r) varies. The weak topology is given by the family of seminorms $\| \cdot \|_{x^*} = |x^*(\cdot)|$. The proof that the weak topology is Hausdorff requires the fact, for each $x \neq 0$ in X , that there is some member x^* with $x^*(x) \neq 0$; this fact is one of the standard corollaries of the Hahn–Banach Theorem. Examples of weak topologies will be discussed in Section 3.

Similarly the weak-star topology on X^* , when X is a normed linear space, was defined in *Basic* as the weakest topology on X^* that makes all members of X continuous. This is given by the family of seminorms $\| \cdot \|_x = | \cdot (x) |$. Here the relevant fact for seeing that the topology is Hausdorff is that for each $x^* \neq 0$ in X^* , there is some x in X with $x^*(x) \neq 0$. This is just a matter of the definition of $x^* \neq 0$ and depends on no theorem. Examples of weak-star topologies will be discussed in Section 3.

The above classes of examples by no means exhaust the possibilities for topological vector spaces. Let us mention briefly one example that is not even close to being definable by seminorms. It is the space $L^p([0, 1])$ with $0 < p < 1$. This is the vector space of all real-valued Borel functions on $[0, 1]$ with $\int_{[0,1]} |f|^p dx$ finite, except that we identify two functions if they differ only on a set of measure 0. Let us see that $d(f, g) = \int_{[0,1]} |f - g|^p dx$ is a metric. We need only verify the triangle inequality in the form $\int_{[0,1]} |f + g|^p dx \leq \int_{[0,1]} |f|^p dx + \int_{[0,1]} |g|^p dx$. To check this, we observe for nonnegative r that $(1 + r)^p - (1 + r^p)$ is 0 at $r = 0$ and has negative derivative $p((1 + r)^{p-1} - r^{p-1})$ since $p - 1$ is negative. Thus $(1 + r)^p \leq 1 + r^p$ for $r \geq 0$, and consequently $|a + b|^p \leq (|a| + |b|)^p \leq |a|^p + |b|^p$ for all real a and b . Taking $a = f(x)$ and $b = g(x)$ and integrating, we obtain the desired triangle inequality. One readily shows that $L^p([0, 1])$ with this metric is a topological vector space. On the other hand, this topological vector space is rather pathological, as is shown in Problem 8 at the end of the chapter. For example it has no nonzero continuous linear functionals, whereas nonzero topological vector spaces whose topologies are given by seminorms always have enough continuous linear functionals to separate points.²

Now we turn our attention to a few results valid for arbitrary topological vector spaces.

Proposition 4.2. In any topological vector space, the closure of any vector subspace is a vector subspace.

PROOF. Let V be a vector subspace of the topological vector space X . If x and y are in V^{cl} , then (x, y) is in $V^{\text{cl}} \times V^{\text{cl}} = (V \times V)^{\text{cl}}$. Any continuous function

²More precisely it will be observed in Section 6 that topological vector spaces whose topologies are given by seminorms are “locally convex,” and it will be proved in that same section that locally convex spaces always have enough continuous linear functionals to separate points.

f has the property for any set S that $f(S^{\text{cl}}) \subseteq f(S)^{\text{cl}}$. Applying this fact to the addition function, we see that $x + y$ is in V^{cl} since V is the image of $V \times V$ under addition. Thus V^{cl} is closed under addition. Similarly V^{cl} is closed under scalar multiplication.

Lemma 4.3. If X is a real or complex vector space in which addition and scalar multiplication are continuous and if $\{0\}$ is a closed subset of X , then X is Hausdorff and hence is a topological vector space.

PROOF. Since translations are homeomorphisms, it is enough to separate 0 and an arbitrary $x \neq 0$ by disjoint open neighborhoods. Since $X - \{0\}$ is open, so is $V = X - \{x\}$. By continuity of subtraction, choose an open neighborhood U of 0 such that the set of differences satisfies $U - U \subseteq V$. Then U and $x + U$ are open neighborhoods of 0 and x . If y is in their intersection, then y is in U , and y is of the form $x + u$ for some u in U . Hence $x = y - u$ exhibits x as in $U - U \subseteq V = X - \{x\}$, contradiction. Thus we can take U and $x + U$ as the required disjoint open neighborhoods of 0 and x .

Proposition 4.4. If X is a topological vector space, if Y is a closed vector subspace, and if the quotient vector space X/Y is given the quotient topology,³ then X/Y is a topological vector space, and the quotient map $q : X \rightarrow X/Y$ carries open sets to open sets.

PROOF. If U is open in X , then $q^{-1}(q(U)) = \bigcup_{y \in Y} (y + U)$ exhibits $q^{-1}(q(U))$ as the union of open sets and hence as an open set. By definition of the topology on X/Y , $q(U)$ is open in X/Y . Hence q carries open sets in X to open sets in X/Y .

To see that addition is continuous in X/Y , let x_1 and x_2 be in X , and let E be an open neighborhood of the member $x_1 + x_2 + Y$ of X/Y . Then $q^{-1}(E)$ is an open neighborhood of $x_1 + x_2$ in X . By continuity of addition in X , there exist open neighborhoods U_1 of x_1 and U_2 of x_2 such that $U_1 + U_2 \subseteq q^{-1}(E)$. The map q is open and linear, and hence $q(U_1)$ and $q(U_2)$ are open subsets of X/Y with $q(U_1) + q(U_2) \subseteq q(q^{-1}(E)) = E$. Thus addition is continuous in X/Y .

To see that scalar multiplication is continuous in X/Y , let c be a scalar, let x be in X , and let E be an open neighborhood of cx in X/Y . Then $q^{-1}(E)$ is an open neighborhood of cx in X . By continuity of scalar multiplication in X , there exist open neighborhoods A of c in the scalars and U of x in X such that $AU \subseteq q^{-1}(E)$. Then $q(U)$ is an open subset of X/Y such that $Aq(U) \subseteq q(q^{-1}(E)) = E$. Hence scalar multiplication is continuous in X/Y .

Applying Lemma 4.3, we see that X/Y is Hausdorff. Therefore X/Y is a topological vector space.

³If $q : X \rightarrow X/Y$ is the quotient mapping, the open sets E of X/Y are defined as all subsets such that $q^{-1}(E)$ is open in X .

Proposition 4.5. If Y is an n -dimensional topological vector space over \mathbb{F} , then Y is isomorphic to \mathbb{F}^n .

PROOF. Let y_1, \dots, y_n be a vector-space basis of Y , and let (\cdot, \cdot) and $|\cdot|$ be the usual inner product and norm on \mathbb{F}^n . If e_1, \dots, e_n is the standard basis of \mathbb{F}^n , define $L(\sum_{j=1}^n c_j e_j) = \sum_{j=1}^n c_j y_j$. Then L is one-one and hence is onto Y . We saw earlier in this section that L is continuous. We shall prove that L^{-1} is continuous, and it is enough to do so at 0 in Y .

Assuming on the contrary that L^{-1} is not continuous at 0, we can find some $\epsilon > 0$ such that no open neighborhood U of 0 in Y maps under L^{-1} into the open neighborhood $\{|x| < \epsilon\}$ of 0 in \mathbb{F}^n . For each such U , find y_U in U with $|L^{-1}(y_U)| \geq \epsilon$. Define $z_U = |L^{-1}(y_U)|^{-1} y_U$. The net $\{y_U\}$ tends to 0 in Y by construction, and the numbers $|L^{-1}(y_U)|^{-1}$ are bounded by ϵ^{-1} . By continuity of scalar multiplication in Y , z_U has limit 0 in Y . On the other hand, the members of \mathbb{F}^n defined by $x_U = L^{-1}(z_U) = |L^{-1}(y_U)|^{-1} L^{-1}(y_U)$ have $|x_U| = 1$ for all U . The unit sphere in \mathbb{F}^n is compact, and it follows that $\{x_U\}$ has a convergent subnet, say $\{x_{U_\mu}\}$, with some limit x_0 such that $|x_0| = 1$. We have $L(x_{U_\mu}) = z_{U_\mu}$, and passage to the limit gives $L(x_0) = \lim_\mu L(x_{U_\mu}) = \lim_\mu z_{U_\mu} = 0$. On the other hand, L is one-one, and hence the equality $L(x_0) = 0$ for some x_0 with $|x_0| = 1$ is a contradiction. We conclude that L^{-1} is continuous.

Corollary 4.6. Every finite-dimensional vector subspace of a topological vector space is closed.

PROOF. Let V be an n -dimensional subspace of a topological vector space X , and suppose that V^{cl} properly contains V . Choose x_0 in $V^{\text{cl}} - V$, and form the vector subspace $W = V + \mathbb{F}x_0$. Then the closure of V in W , being a vector subspace (Proposition 4.2), is W . The vector subspace W has dimension $n + 1$, and Proposition 4.5 shows that W is isomorphic to \mathbb{F}^{n+1} . All vector subspaces of \mathbb{F}^{n+1} are closed in \mathbb{F}^{n+1} , and hence V is closed in W , contradiction.

Lemma 4.7. If X is a topological vector space, K is a compact subset of X , and V is an open neighborhood of 0, then there exists $\epsilon > 0$ such that $\delta K \subseteq V$ whenever $|\delta| \leq \epsilon$.

PROOF. For each $k \in K$, choose $\epsilon_k > 0$ and an open neighborhood U_k of k such that $\delta U_k \subseteq V$ whenever $|\delta| \leq \epsilon_k$; this is possible since scalar multiplication is continuous at the point where the scalar is 0 and the vector is k . The open sets U_k cover K , and the compactness of K implies that there is a finite subcover: $K \subseteq U_{k_1} \cup \dots \cup U_{k_m}$. Then $\delta K \subseteq V$ whenever $|\delta| \leq \min_{1 \leq j \leq m} \epsilon_{k_j}$.

Proposition 4.8. Every locally compact topological vector space is finite dimensional.

PROOF. Let X be a locally compact topological vector space, let K be a compact neighborhood of 0 , and let U be its interior. Suppose that we have a sequence $\{y_m\}$ in X with the property that for any $\delta > 0$, there is an integer M such that $m \geq M$ implies y_m lies in δK . Then the result of Lemma 4.7 implies that $\{y_m\}$ tends to 0 .

The sets $\{k + \frac{1}{2}U \mid k \in K\}$ form an open cover of K . If $\{k_1 + \frac{1}{2}U, \dots, k_n + \frac{1}{2}U\}$ is a finite subcover, we prove that $\{k_1, \dots, k_n\}$ spans X . It is enough to prove that $S = \{k_1, \dots, k_n\}$ spans U . If x is in U , then x is in one of the sets of the finite subcover, say $k_{j_1} + \frac{1}{2}U$. Write $x = k_{j_1} + \frac{1}{2}u_1$ accordingly. The finite subcover covers K and hence its interior U , and thus $\frac{1}{2}U$ is covered by $\frac{1}{2}(k_1 + \frac{1}{2}U), \dots, \frac{1}{2}(k_n + \frac{1}{2}U)$. Applying this observation to the element $\frac{1}{2}u_1$ of $\frac{1}{2}U$, we see that x is in $k_{j_1} + \frac{1}{2}(k_{j_2} + \frac{1}{2}U)$ for some k_{j_2} . Write $x = k_{j_1} + \frac{1}{2}k_{j_2} + \frac{1}{4}u_2$ accordingly. Continuing in this way, we see that

$$x \quad \text{is in} \quad k_{j_1} + \frac{1}{2}k_{j_2} + \dots + \frac{1}{2^{r-1}}k_{j_r} + \frac{1}{2^r}U \quad \text{for each } r.$$

Put $x_r = k_{j_1} + \frac{1}{2}k_{j_2} + \dots + \frac{1}{2^{r-1}}k_{j_r}$. This is an element of the finite-dimensional subspace spanned by S , which is closed by Corollary 4.6; thus if $\{x_r\}$ converges, it must converge to a member x_0 of this subspace. Using the result of the previous paragraph, we shall show that $x - x_r$ converges to 0 . Then we can conclude that x_r converges to x , hence that x is in the span of S . To see that $x - x_r$ converges to 0 , choose l such that $|\delta_0| \leq 2^{-l}$ implies $\delta_0 K \subseteq U$. Applying the criterion of the previous paragraph, let $\delta > 0$ be given. Choose M such that $2^{-M}\delta^{-1} \leq 2^{-l}$. Then $m \geq M$ implies that $2^{-m}\delta^{-1} \leq 2^{-M}\delta^{-1} \leq 2^{-l}$. Thus $2^{-m}\delta^{-1}$ is an allowable choice of δ_0 , and we therefore obtain $2^{-m}\delta^{-1}K \subseteq U$ and $2^{-m}K \subseteq \delta U$. For $m \geq M$, the element $x - x_m$ lies in $2^{-m}U \subseteq 2^{-m}K$, and we have just proved that $2^{-m}K \subseteq \delta U$. Thus $x - x_m$ lies in δU , and the criterion of the previous paragraph applies. Hence $x - x_m$ tends to 0 . This completes the proof.

2. $C^\infty(U)$, Distributions, and Support

As was mentioned in Section III.1, distributions are continuous linear functionals on vector spaces of smooth functions. Their properties are deceptively simple-looking and enormously helpful in working with linear partial differential equations. We considered tempered distributions in Section III.1; these are the continuous linear functionals on the space $\mathcal{S}(\mathbb{R}^N)$ of Schwartz functions on \mathbb{R}^N . In this section we study the topology on the space $C^\infty(U)$ of arbitrary scalar-valued smooth functions on an open subset U of \mathbb{R}^N , together with the associated space of distributions.

To topologize $C^\infty(U)$, we use the family of seminorms indexed by pairs (K, P) with K a compact subset of U and with P a polynomial, the $(K, P)^{\text{th}}$ seminorm

being $\|f\|_{K,P} = \sup_{x \in K} |(P(D)f)(x)|$. The resulting topology is Hausdorff, and $C^\infty(U)$ becomes a topological vector space.

Let us see that this topology is given by a countable subfamily of these seminorms and is therefore implemented by a metric. It is certainly sufficient to consider only the monomials D^α instead of all polynomials $P(D)$, and thus the P index of (K, P) can be assumed to run through a countable set. We make use of a notion already used in Section III.2. An **exhausting sequence** of compact subsets of U is an increasing sequence of compact sets with union U such that each set is contained in the interior of the next set. An exhausting sequence exists in any locally compact separable metric space. If $\{K_n\}$ is an exhausting sequence for U and if K is a compact subset of U , then the interiors K_n^o of the K_n 's form an open cover of K , and there is a finite subcover; since the members of the open cover are nested, K is contained in some single K_n^o and hence in K_n . Therefore $\|f\|_{K,P} \leq \|f\|_{K_n,P}$ for every P , and we can discard all the seminorms except the ones from some K_n . In short, the countably many seminorms $\|f\|_{K_n, x^\alpha} = \sup_{x \in K_n} |(D^\alpha f)(x)|$ suffice to determine the topology of $C^\infty(U)$. In particular, the topology is independent of the choice of exhausting sequence.

After the statement of Theorem 3.9, we constructed a smooth partition of unity $\{\psi_n\}_{n \geq 1}$ associated to an exhausting sequence $\{K_n\}_{n \geq 1}$ of an open subset U of \mathbb{R}^N . Such a partition of unity is sometimes useful, and Problem 9 at the end of the chapter illustrates this fact. The functions ψ_n are in $C^\infty(U)$ and have the properties that $\sum_{n=1}^{\infty} \psi_n(x) = 1$ on U , $\psi_1(x) > 0$ on K_3 , $\psi_1(x) = 0$ on $(K_4^o)^c$, and for $n \geq 2$,

$$\psi_n(x) \begin{cases} > 0 & \text{for } x \in K_{n+2} - K_{n+1}^o, \\ = 0 & \text{for } x \in (K_{n+3}^o)^c \cup K_n. \end{cases}$$

Since $C^\infty(U)$ is a metric space, its topology may be characterized in terms of convergence of sequences: a sequence of functions converges in $C^\infty(U)$ if and only if the functions converge uniformly on each compact subset of U and so do each of their iterated partial derivatives

If a particular metric for $C^\infty(U)$ is specified as constructed in Section III.1 from an enumeration of some determining countable family of seminorms, then it is apparent that a sequence of functions is Cauchy in $C^\infty(U)$ if and only if the functions and all their iterated partial derivatives are uniformly Cauchy on each compact subset of U . As a consequence we can see that $C^\infty(U)$ is complete as a metric space: in fact, let us extract limits from each uniformly Cauchy sequence of derivatives and use the standard theorem on derivatives of convergent sequences whose derivatives converge uniformly; the result is that we obtain a member of $C^\infty(U)$ to which the Cauchy sequence converges.

It is unimportant which particular metric is used for this completeness argument. The relevant consequence is that the Baire Category Theorem⁴ is applicable to $C^\infty(U)$, and the statement of the Baire Category Theorem makes no reference to a particular metric.

In similar fashion one checks that $\mathcal{S}(\mathbb{R}^N)$, whose topology is likewise given by countably many seminorms, is complete as a metric space.

The vector space of continuous linear functionals on $C^\infty(U)$, i.e., its **continuous dual**, is called the space of all **distributions of compact support** on U and is traditionally⁵ denoted by $\mathcal{E}'(U)$. The words “of compact support” require some explanation and justification, which we come back to after giving an example.

EXAMPLE. Take finitely many complex Borel measures ρ_α of compact support on U , the indexing being by the set of n -tuples α of nonnegative integers with $|\alpha| \leq m$, and define

$$T(\varphi) = \sum_{|\alpha| \leq m} \int_U D^\alpha \varphi(x) d\rho_\alpha(x).$$

It is easy to check that T is a distribution of compact support on U . A theorem in Chapter V will provide a converse, saying essentially that every continuous linear functional on $C^\infty(U)$ is of this form.

Let us observe that the vector subspace $C_{\text{com}}^\infty(U)$ is dense in $C^\infty(U)$. In fact, let $\{K_j\}$ be an exhausting sequence of compact sets in U , and choose $\psi_j \in C_{\text{com}}^\infty(\mathbb{R}^n)$ by Proposition 3.5f to be 1 on K_j and 0 off K_{j+1} . If f is in $C^\infty(U)$, then $\psi_j f$ is in $C_{\text{com}}^\infty(U)$ and tends to f in every seminorm on $C^\infty(U)$.

To obtain a useful notion of “support” for a distribution, we need the following lemma.

Lemma 4.9. If U_1 and U_2 are nonempty open sets in \mathbb{R}^N and if φ is in $C_{\text{com}}^\infty(U_1 \cup U_2)$, then there exist $\varphi_1 \in C_{\text{com}}^\infty(U_1)$ and $\varphi_2 \in C_{\text{com}}^\infty(U_2)$ such that $\varphi = \varphi_1 + \varphi_2$.

PROOF. Let L be the compact support of φ , and choose a compact set K such that $L \subseteq K^\circ \subseteq K \subseteq U_1 \cup U_2$. Then $\{U_1, U_2\}$ is a finite open cover of K , and Lemma 3.15b of *Basic* produces an open cover $\{V_1, V_2\}$ of K such that V_1^{cl} is a compact subset of U_1 and V_2^{cl} is a compact subset of U_2 . Proposition 3.5f produces functions $g_1 \in C_{\text{com}}^\infty(U_1)$ and $g_2 \in C_{\text{com}}^\infty(U_2)$ with values in $[0, 1]$ such that g_1 is 1 on V_1^{cl} and g_2 is 1 on V_2^{cl} . Then $g = g_1 + g_2$ is in $C_{\text{com}}^\infty(U_1 \cup U_2)$ and

⁴Theorem 2.53 of *Basic*.

⁵The tradition dates back to Laurent Schwartz’s work, in which $\mathcal{E}(U)$ was the notation for $C^\infty(U)$ and $\mathcal{E}'(U)$ was the space of continuous linear functionals.

is 1 on K . If W is the open set where $g \neq 0$, then Proposition 3.5f produces a function h in $C_{\text{com}}^\infty(W)$ with values in $[0, 1]$ such that h is 1 on K . The function $1 - h$ is smooth, has values in $[0, 1]$, is 1 where $g \neq 0$, and is 0 on K . Hence $g + (1 - h)$ is a smooth function that is everywhere positive on \mathbb{R}^N and equals g on K . Therefore the functions $g_1/(g + 1 - h)$ and $g_2/(g + 1 - h)$ are smooth functions on \mathbb{R}^N compactly supported in U_1 and U_2 , respectively, with sum equal to 1 on K . If we define $\varphi_1 = g_1\varphi$ and $\varphi_2 = g_2\varphi$, then φ_1 and φ_2 have the required properties.

Proposition 4.10. If T is an arbitrary linear functional on $C_{\text{com}}^\infty(U)$ and if U' is the union of all open subsets U_γ of U such that T vanishes on $C_{\text{com}}^\infty(U_\gamma)$, then T vanishes on $C_{\text{com}}^\infty(U')$.

PROOF. Let φ be in $C_{\text{com}}^\infty(U')$, and let K be the support of φ . The open sets U_γ form an open cover of K , and some finite subcollection must have $K \subseteq U_{\gamma_1} \cup \cdots \cup U_{\gamma_p}$. Lemma 4.9 applied inductively shows that φ is the sum of functions in $C_{\text{com}}^\infty(U_j)$, $1 \leq j \leq p$. Since T is 0 on each of these, it is 0 on the sum.

If T is in $\mathcal{E}'(U)$, the **support** of T is the complement of the set U' in Proposition 4.10, i.e., the complement of the union of all open sets U_γ such that T vanishes on $C_{\text{com}}^\infty(U_\gamma)$. If T has empty support, then $T = 0$ because T vanishes on $C_{\text{com}}^\infty(U)$ and $C_{\text{com}}^\infty(U)$ is dense in $C^\infty(U)$.

Proposition 4.11. Every member T of $\mathcal{E}'(U)$ has compact support.

REMARKS. For the moment this proposition justifies using the name “distributions of compact support” for the continuous linear functionals on $C^\infty(U)$. After we define general distributions in Section V.1, we shall have to return to this matter.

PROOF. Let $\{K_n\}$ be an exhausting sequence of compact sets in U . If T is not supported in any K_n , then there is some f_n in $C_{\text{com}}^\infty(U - K_n)$ with $T(f_n) \neq 0$. Put $g_n = f_n/T(f_n)$, so that $T(g_n) = 1$. If K is any compact subset of U , then $K \subseteq K_n$ for large n , and $g_n|_K = 0$ for such n . Thus g_n tends to 0 in $C^\infty(U)$ while $T(g_n)$ tends to $1 \neq 0 = T(0)$, in contradiction to continuity of T .

Similarly we can use Proposition 4.10 to define the **support** of a tempered distribution T in $\mathcal{S}'(\mathbb{R}^N)$ as the complement of the union of all open sets U_γ such that T vanishes on $C_{\text{com}}^\infty(U_\gamma)$. Tempered distributions need not have compact support; for example, the function 1 defines a tempered distribution whose support is \mathbb{R}^N .

In the case of tempered distributions, a little argument is required to show that the only tempered distribution with empty support is the 0 distribution. What is needed is the following fact.

Proposition 4.12. $C_{\text{com}}^{\infty}(\mathbb{R}^N)$ is dense in $\mathcal{S}(\mathbb{R}^N)$.

REMARKS. If T in $\mathcal{S}'(\mathbb{R}^N)$ has empty support, then T vanishes on $C_{\text{com}}^{\infty}(\mathbb{R}^N)$. Proposition 4.12 and the continuity of T imply that $T = 0$ on $\mathcal{S}(\mathbb{R}^N)$. Thus the only tempered distribution with empty support is the 0 distribution.

PROOF. Fix h in $C_{\text{com}}^{\infty}(\mathbb{R}^N)$ with values in $[0, 1]$ such that $h(x)$ is 1 for $|x| \leq 1$ and is 0 for $|x| \geq 2$. Define $h_R(x) = h(R^{-1}x)$. If φ is in $\mathcal{S}(\mathbb{R}^N)$, we shall show that $\lim_{R \rightarrow \infty} h_R \varphi = \varphi$ in the metric space $\mathcal{S}(\mathbb{R}^N)$, and then the proposition will follow. Thus we want $\lim_{R \rightarrow \infty} \sup_{x \in \mathbb{R}^N} |x^\gamma D^\alpha(\varphi - h_R \varphi)(x)| = 0$. By the Leibniz rule, $D^\alpha(h_R \varphi) = h_R D^\alpha \varphi + \sum_{\beta < \alpha} c_\beta (D^{\alpha-\beta} h_R)(D^\beta \varphi)$. Hence it is enough to prove that

$$\lim_{R \rightarrow \infty} \sup_{x \in \mathbb{R}^N} |x^\gamma (1 - h_R) D^\alpha \varphi| = 0$$

$$\text{and} \quad \lim_{R \rightarrow \infty} \sup_{x \in \mathbb{R}^N} |x^\gamma (D^{\alpha-\beta} h_R)(D^\beta \varphi)| = 0 \quad \text{for } \beta < \alpha.$$

The first of these limit formulas is a consequence of the fact that $x^\gamma D^\alpha \varphi$ vanishes at infinity, which in turn follows from the fact that $x^\gamma (1 + |x|^2) D^\alpha \varphi$ is bounded, i.e., that $\|\varphi\|_{x^\gamma(1+|x|^2), x^\alpha}$ is finite. For the second of these limit formulas, we observe from the chain rule that $D^{\alpha-\beta} h_R(x) = R^{-|\alpha-\beta|} D^{\alpha-\beta} h(R^{-1}x)$. For $\beta < \alpha$, this function is dominated in absolute value by $c_\alpha R^{-1}$. Hence $\sup_{x \in \mathbb{R}^N} |x^\gamma (D^{\alpha-\beta} h_R)(D^\beta \varphi)| \leq c_\alpha R^{-1} \sum_{\beta < \alpha} \|\varphi\|_{x^\gamma, x^\beta}$, and the limit on R is 0.

3. Weak and Weak-Star Topologies, Alaoglu's Theorem

Let X be a normed linear space, and let X^* be its dual, which we know to be a Banach space. We have defined the **weak topology** on X to be the weakest topology on X making all members of X^* continuous, i.e., making $x \mapsto x^*(x)$ continuous for each x^* in X^* . This topology is given by the family of seminorms $\|x\|_{x^*} = |x^*(x)|$ indexed by X^* . The **weak-star topology** on X^* relative to X is the weakest topology on X^* making all members of $\iota(X)$ continuous,⁶ i.e., making $x^* \mapsto x^*(x)$ continuous for each x in X . This topology is given by the family of seminorms $\|x^*\|_x = |x^*(x)|$ indexed by X . In this section we

⁶The symbol ι denotes the canonical map $X \rightarrow X^{**}$ given by $\iota(x)(x^*) = x^*(x)$.

study these topologies⁷ in more detail, proving an important theorem about the weak-star topology.

We shall discuss some examples in a moment. The space X^* is a normed linear space in its own right, and therefore it has a well-defined weak topology. The definitions make the weak topology on X^* the same as the weak-star topology on X^* relative to X if X is reflexive, but we cannot draw this conclusion in general.

The weak topology on X is of less importance to real analysis than the weak-star topology on X^* , and thus the main interest in the weak topology on X will be in the case that X is reflexive. It is also true that exact conditions that interpret the weak or weak-star topology in a particular example tend not to be useful. Nevertheless, it may still be helpful to consider examples in order to get a better sense of what these topologies do.

We shall discuss the examples in terms of convergence. However, the convergence will involve only convergence of sequences, not convergence of general nets. A difficulty with nets is that one cannot draw familiar conclusions from convergence of nets even in the case of nets in the real numbers; for example, a convergent net of real numbers need not be bounded, just eventually bounded.

In order to have it available in the discussion, we prove one fact about convergence of sequences in weak and weak-star topologies before coming to the examples.

Proposition 4.13. Let X be a normed linear space, and let X^* be its dual space.

(a) If $\{x_n\}$ is a sequence in X converging to x_0 in the weak topology on X , then $\{\|x_n\|\}$ is a bounded sequence in \mathbb{R} and $\|x_0\| \leq \liminf_n \|x_n\|$.

(b) If X is a Banach space and if $\{x_n^*\}$ is a sequence in X^* converging to x_0^* in the weak-star topology on X^* relative to X , then $\{\|x_n^*\|\}$ is a bounded sequence in \mathbb{R} and $\|x_0^*\| \leq \liminf_n \|x_n^*\|$.

PROOF. For the first half of (a), let $\iota : X \rightarrow X^{**}$ be the canonical map. Since the sequence $\{\iota(x_n)(x^*)\}$ converges to $x^*(x_0)$ for each x^* in X^* , $\{\iota(x_n)\}$ is a set of bounded linear functionals on the Banach space X^* with $\{\iota(x_n)(x^*)\}$ bounded for each x^* in X^* . By the Uniform Boundedness Theorem the norms $\|\iota(x_n)\|$ are bounded. Since ι preserves norms as a consequence of the Hahn–Banach Theorem, the norms $\|x_n\|$ are bounded. For the second half of (a), let x^* be arbitrary in X^* with $\|x^*\| \leq 1$. Then

$$|x^*(x_0)| = \lim |x^*(x_n)| \leq \liminf \|x^*\| \|x_n\| \leq \liminf \|x_n\|.$$

Taking the supremum over x^* with $\|x^*\| \leq 1$ and applying the formula $\|x_0\| = \sup_{\|x^*\| \leq 1} |x^*(x_0)|$, which is known from the Hahn–Banach Theorem, we obtain $\|x_0\| \leq \liminf \|x_n\|$.

⁷The weak topology on X is also called the X^* topology of X , and the weak-star topology on X^* is also called the X topology of X^* .

For the first half of (b), $\{x_n^*\}$ is a set of bounded linear functionals on the Banach space X with $\{x_n^*(x)\}$ bounded for each x in X . Then the Uniform Boundedness Theorem shows that the norms $\|x_n^*\|$ are bounded. For the second half of (b), let x be arbitrary in X with $\|x\| \leq 1$. Then

$$|x_0^*(x)| = \lim |x_n^*(x)| \leq \liminf \|x_n^*\| \|x\| \leq \liminf \|x_n^*\|.$$

Taking the supremum over x and applying the definition of $\|x_0^*\|$, we obtain $\|x_0^*\| \leq \liminf \|x_n^*\|$.

EXAMPLES OF CONVERGENCE IN WEAK TOPOLOGIES.

(1) $X = L^p(S, \mu)$ when $1 < p < \infty$. Then $X^* \cong L^{p'}(X, \mu)$, where p' is the dual index⁸ of p . The assertion is that a sequence $\{f_n\}$ tends weakly to f in L^p if and only if $\{\|f_n\|_p\}$ is bounded and $\lim \int_E f_n d\mu = \int_E f d\mu$ for every measurable subset E of S of finite measure. The necessity is immediate from Proposition 4.13a and from taking the member of X^* to be the indicator function of E . Let us prove the sufficiency. From $\lim \int_E f_n d\mu = \int_E f d\mu$, we see that $\lim \int_S f_n t d\mu = \int_S f t d\mu$ for t simple if t is 0 off a set of finite measure. Let g be given in $L^{p'}(S, \mu)$, and choose a sequence $\{t_m\}$ of simple functions equal to 0 off sets of finite measure such that $\lim_m t_m = g$ in the norm topology of $L^{p'}$. For all m and n , we have

$$\begin{aligned} & \left| \int_S f_n g d\mu - \int_S f g d\mu \right| \\ & \leq \left| \int_S f_n (g - t_m) d\mu \right| + \left| \int_S f_n t_m d\mu - \int_S f t_m d\mu \right| \\ & \quad + \left| \int_S f (t_m - g) d\mu \right| \\ & \leq \|f_n\|_p \|g - t_m\|_{p'} + \left| \int_S f_n t_m d\mu - \int_S f t_m d\mu \right| + \|f\|_p \|g - t_m\|_{p'}. \end{aligned}$$

The first and third terms on the right tend to 0 as m tends to infinity, uniformly in n . If $\epsilon > 0$ is given, choose m such that those two terms are $< \epsilon$, and then, with m fixed, choose n large enough to make the middle term $< \epsilon$.

(2) $X = C(S)$ with S compact Hausdorff, $C(S)$ being the space of continuous scalar-valued functions on S . Then X^* may be identified with the space $M(S)$ of (signed or) complex regular Borel measures on S , with the total-variation norm.⁹ The assertion is that a sequence $\{f_n\}$ tends weakly to f in $C(S)$ if and only if $\{\|f_n\|\}$ is bounded and $\lim f_n = f$ pointwise. The necessity is immediate from Proposition 4.13a and from taking the member of X^* to be any point mass at a point

⁸The index p' is defined by $\frac{1}{p} + \frac{1}{p'} = 1$. This duality was proved in Theorem 9.19 of *Basic* when μ is σ -finite, but it holds without this restrictive assumption on μ .

⁹This identification was obtained in *Basic* in Theorem 11.24 for real scalars and in Theorem 11.26 for complex scalars. The starting point for the identification is the Riesz Representation Theorem.

of S . For the sufficiency we simply observe that any member of $M(S)$ is a finite linear combination of regular Borel measures μ on S and $\lim \int_S f_n d\mu = \int_S f d\mu$ for any Borel measure μ by dominated convergence.

(3) $X = C_0(S)$ with S locally compact separable metric, $C_0(S)$ being the space of continuous scalar-valued functions vanishing at infinity. Again the dual X^* may be identified with the space $M(S)$ of complex regular Borel measures on S , with the total-variation norm. This example can be handled by applying the previous example to the one-point compactification of S . All signed or complex Borel measures are automatically regular in this case. A sequence $\{f_n\}$ tends weakly to f in $C_0(S)$ if and only if $\{\|f_n\|\}$ is bounded and $\lim f_n = f$ pointwise.

EXAMPLES OF CONVERGENCE IN WEAK-STAR TOPOLOGIES.

(1) $X = L^p(S, \mu)$ and $X^* \cong L^{p'}(S, \mu)$ when $1 < p < \infty$, p' being the dual index of p . This X is reflexive. Therefore the first example of convergence in weak topologies shows that $\{f_n\}$ converges weak-star in $L^{p'}(S, \mu)$ relative to $L^p(S, \mu)$ if and only if $\{\|f_n\|_{p'}\}$ is bounded and $\lim \int_E f_n d\mu = \int_E f d\mu$ for every measurable subset E of S of finite measure.

(2) $X = L^1(S, \mu)$ and $X^* \cong L^\infty(S, \mu)$ when μ is σ -finite. This X is usually not reflexive. However, the condition for weak-star convergence is the same as in the previous example: $\{f_n\}$ converges weak-star in $L^\infty(S, \mu)$ relative to $L^1(S, \mu)$ if and only if $\{\|f_n\|_\infty\}$ is bounded and $\lim \int_E f_n d\mu = \int_E f d\mu$ for every measurable subset E of S of finite measure. The argument in the first example of convergence in weak topologies can easily be modified to prove this.

(3) $X = C(S)$ with S compact Hausdorff, and $X = C_0(S)$ with S locally compact separable metric. Weak-star convergence of complex regular Borel measures does not have a useful necessary and sufficient condition beyond the definition. The notion of weak-star convergence in this situation is, nevertheless, quite helpful as a device for producing new complex measures out of old ones.¹⁰

A theorem about the weak topology, due to Banach, is that the vector subspaces that are closed in the weak topology are the same as the vector subspaces that are closed in the norm topology. More generally the closed convex sets coincide in the weak and norm topologies. We shall not have occasion to use this theorem or mention any of its applications, and we therefore omit the proof.

The weak-star topology has results of more immediate interest, and we turn our attention to those. Theorem 5.58 of *Basic* established for any separable normed linear space X that any bounded sequence in the dual X^* has a weak-star convergent subsequence; this was called a "preliminary form of Alaoglu's Theorem."

¹⁰*Warning.* Many probabilists and some other people use the unfortunate term "weak convergence" for this instance of weak-star convergence.

Theorem 4.14 Let X be a normed linear space with dual X^* .

(a) (**Alaoglu's Theorem**) The closed unit ball of X^* is compact in the weak-star topology relative to X .

(b) If X is separable, then the closed unit ball of X^* is a separable metric space in the weak-star topology.

REMARKS. By (a), any net $\{x_\alpha^*\}$ in X^* with $\|x_\alpha^*\|$ bounded has a subnet $\{x_{\alpha_\mu}^*\}$ and an element x_0^* in X^* such that $x_{\alpha_\mu}^*(x) \rightarrow x_0^*(x)$ for every x in X . By (b), this conclusion about nets can be replaced by a conclusion about sequences if X is separable. Thus we recover the “preliminary form” of Alaoglu's Theorem. The results of Section III.4 give an example of the utility of the two parts of this theorem; together they lead to a proof that harmonic functions in $\mathcal{H}^p(\mathbb{R}_+^{N+1})$ are automatically Poisson integrals of functions if $p > 1$ or of complex measures if $p = 1$.

PROOF. Let B be the closed unit ball in X^* , let $D(r)$ be the closed disk in \mathbb{C} with radius r and center 0, and let $C = \prod_{x \in X} D(\|x\|)$. Define $F : B \rightarrow C$ by $F(x^*) = \prod_{x \in X} x^*(x)$. The function F is well defined since $|x^*(x)| \leq \|x\|$ for all x^* in B and all x in X . It is continuous as a map into the product space since $x^* \mapsto x^*(x)$ is continuous for each x , it is one-one since x^* is determined by its values on each x , and it is a homeomorphism with its image by definition of weak topology. Since C is compact by the Tychonoff Product Theorem, (a) will follow if it is shown that $F(B)$ is closed in C . Let p_x denote the projection of C to its x^{th} coordinate. If x and x' are in X and if $\{f_\alpha\}$ is a net in C convergent to f_0 in C , then an equality $p_{x+x'}(f_\alpha) = p_x(f_\alpha) + p_{x'}(f_\alpha)$ for all α implies that $p_{x+x'}(f_0) = p_x(f_0) + p_{x'}(f_0)$ by continuity of $p_{x+x'}$, p_x , and $p_{x'}$. Thus the set

$$S(x, x') = \{f \in C \mid p_{x+x'}(f) = p_x(f) + p_{x'}(f)\}$$

is closed, and similarly the set

$$T(x, c) = \{f \in C \mid cp_x(f) = p_x(cf)\}$$

is closed. The intersection of all $S(x, x')$'s and all $T(x, c)$'s is the set of linear members of C , hence is exactly $F(B)$. Thus $F(B)$ is closed.

For (b), we continue with B and $D(r)$ as above, but we change C and F slightly. Let $\{x_n\}$ be a countable dense set in the norm topology of X , let $C = \prod_{x_n} D(\|x_n\|)$, and define $F : B \rightarrow C$ by $F(x^*) = \prod_{x_n} x^*(x_n)$. As in the proof of (a), F is continuous. It is one-one since any x^* , being continuous, is determined by its values on the dense set $\{x_n\}$. The domain is compact by (a). The range space C is a separable metric space and is in particular Hausdorff. Hence B is exhibited as homeomorphic to $F(B)$, which is a subspace of the separable metric space C and is therefore separable.

4. Stone Representation Theorem

In this section we begin to follow Alaoglu's Theorem along paths different from its use for creating limit functions and measures out of sequences that are bounded in a weak-star topology. We shall work in this section with what amounts to an example—one of the motivating examples behind a stunning idea of I. M. Gelfand around 1940 that brings algebra, real analysis, and complex analysis together in a profound way. The example gives a view of subalgebras of the algebra $B(S)$ of all bounded functions on a set S in terms of compactness. The stunning idea that came out, on which we shall elaborate shortly, is that the mechanism in the proof is the same mechanism that lies behind the Fourier transform on \mathbb{R}^N , that this mechanism can be cast in abstract form as a theory of commutative Banach algebras, and that the theory gives a new perspective about spectra. In particular, it leads directly to the full Spectral Theorem for bounded and unbounded self-adjoint operators, extending the theorem for compact self-adjoint operators that was proved as Theorem 2.3. In turn, the Spectral Theorem has many applications to the study of particular operators.

Let us first state the theorem about $B(S)$, then discuss Gelfand's stunning idea about the mechanism, and finally give the proof of the theorem. We shall pursue the Gelfand idea in Sections 10–11 later in this chapter.

We have discussed $B(S)$ as the Banach space of bounded complex-valued functions on a nonempty set S , the norm being the supremum norm. In this Banach space pointwise multiplication makes $B(S)$ into a complex associative algebra¹¹ with identity (namely the function 1), there is an operation of complex conjugation, and there is a notion of positivity (namely pointwise positivity of a function). The theorem concerns subalgebras of $B(S)$ containing 1, closed under conjugation, and closed under uniform limits.

Theorem 4.15 (Stone Representation Theorem). Let S be a nonempty set, and let \mathcal{A} be a uniformly closed subalgebra of $B(S)$ with the properties that \mathcal{A} is stable under complex conjugation and contains 1. Then there exist a compact Hausdorff space S_1 , a function $p : S \rightarrow S_1$ with dense image, and a norm-preserving algebra isomorphism U of \mathcal{A} onto $C(S_1)$ preserving conjugation and positivity, mapping 1 to 1, and having the property that $U(f)(p(s)) = f(s)$ for all s in S . If S is a Hausdorff topological space and \mathcal{A} consists of continuous functions, then p is continuous.

¹¹An **associative algebra** \mathcal{A} over \mathbb{C} is a vector space with a \mathbb{C} bilinear associative multiplication, i.e., with an operation $\mathcal{A} \times \mathcal{A} \rightarrow \mathcal{A}$ satisfying $(ab)c = a(bc)$, $a(b+c) = ab+ac$, $(a+b)c = ac+bc$, and $a(\lambda c) = (\lambda a)c = \lambda(ac)$ if λ is in \mathbb{C} and a, b, c are in \mathcal{A} . This definition does not assume the existence of an identity element.

The idea of the proof is to consider the Banach-space dual \mathcal{A}^* and focus on those members of \mathcal{A}^* that are nonzero and respect multiplication—the nonzero continuous multiplicative linear functionals on \mathcal{A} . The ones that come immediately to mind are the evaluations at each point: for a point s of S , the evaluation at s is given by $e_s(f) = f(s)$, and it is a multiplicative linear functional, certainly of norm 1. The set S_1 in the theorem will be the set of all such continuous multiplicative linear functionals, the function p will be given by $p(s) = e_s$ for $s \in S$, and the mapping U will be given by $U(f)(\ell) = \ell(f)$ for each multiplicative linear functional ℓ .

The Banach space $\mathcal{A} \subseteq B(S)$, with its multiplication, is a **Banach algebra** in the sense that it is an associative algebra over \mathbb{C} , with or without identity, such that $\|fg\| \leq \|f\|\|g\|$ for all f and g in \mathcal{A} . Another well-known Banach algebra is $L^1(\mathbb{R}^N)$. The norm in this case is the usual L^1 norm, and the multiplication is convolution, which satisfies $\|f * g\|_1 \leq \|f\|_1 \|g\|_1$ for all f and g in $L^1(\mathbb{R}^N)$.

The stunning idea of Gelfand's is that the formula that defines U in the Stone theorem is the same formula that gives the Fourier transform in the case of $L^1(\mathbb{R}^N)$. Specifically the nonzero multiplicative linear functionals in the case of $L^1(\mathbb{R}^N)$ are the evaluations at points of the Fourier transform, i.e., the mappings $f \mapsto \widehat{f}(y) = \int_{\mathbb{R}^N} f(x) e^{-2\pi i x \cdot y} dx$. These linear functionals are multiplicative because convolution goes into pointwise product under the Fourier transform.

What $\mathcal{A} \subseteq B(S)$ and $L^1(\mathbb{R}^N)$ have in common is, in the first place, that they are commutative Banach algebras. In addition, each has a conjugate-linear mapping $f \mapsto f^*$ that respects multiplication: complex conjugation in the case of \mathcal{A} and the map $f \mapsto f^*$ with $f^*(x) = \overline{f(-x)}$ in the case of $L^1(\mathbb{R}^N)$. These conjugate-linear mappings interact well with the norm. The subalgebra \mathcal{A} of $B(S)$ satisfies

- (i) $\|ff^*\| = \|f\|\|f^*\|$ for all f ,
- (ii) $\|f^*\| = \|f\|$ for all f ,

while $L^1(\mathbb{R}^N)$ satisfies just (ii). The theory that Gelfand developed applies best when both (i) and (ii) are satisfied, as is the case with \mathcal{A} and also any L^∞ space, and it works somewhat when just (ii) holds, as with $L^1(\mathbb{R}^N)$.

Another example of a Banach algebra is the algebra $\mathcal{B}(H, H)$ of bounded linear operators from a Hilbert space H to itself, with the operator norm. The conjugate-linear mapping on $\mathcal{B}(H, H)$ is passage to the adjoint, and (i) and (ii) both hold. The thing that is missing is commutativity for $\mathcal{B}(H, H)$. However, if we take a single operator A and its adjoint A^* , assume that A commutes with A^* , and take the Banach algebra generated by A and A^* , then we have another example to which the Gelfand theory applies well. The Spectral Theorem for bounded self-adjoint operators is the eventual consequence.

The idea of considering the Banach subalgebra generated by A is a natural one because of one's experience in the subject of modern algebra: the study of

all complex polynomials in a square matrix A is a useful tool in understanding a single linear transformation, including obtaining canonical forms for it like the Jordan form. Thus the use of an analogy with a topic in algebra leads one to a better understanding of a topic in analysis.

In this case ideas flowed in the reverse direction as well. The multiplicative linear functionals correspond, by passage to their kernels, to those ideals in the algebra that are maximal.¹² In effect the Banach algebra was being studied through its space of maximal ideals. About 1960, no doubt partly because of the success of the idea of considering the maximal ideals of a Banach algebra, the consideration of the totality of prime ideals of a commutative ring as a space began to play an important role in algebraic number theory and algebraic geometry.

PROOF OF THEOREM 4.15. Let S_1 be the set of all nonzero continuous multiplicative linear functionals ℓ on \mathcal{A} with $\ell(\bar{f}) = \overline{\ell(f)}$. Let us see that each such has norm 1. In fact, choose f with $\ell(f) \neq 0$. Then $\ell(f) = \ell(f1) = \ell(f)\ell(1)$ shows that $\ell(1) = 1$, and hence $\|\ell\| \geq 1$. For any f with $\|f\|_{\text{sup}} \leq 1$, if we had $|\ell(f)| > 1$, then $|\ell(f)|^n = |\ell(f^n)| \leq \|\ell\|$ for all n would give a contradiction as soon as n is sufficiently large. We conclude that $\|\ell\| \leq 1$.

Therefore S_1 is a subset of the unit ball of the Banach-space dual \mathcal{A}^* . We give S_1 the relative topology from the weak-star topology on \mathcal{A}^* . Let us define the function $p : S \rightarrow S_1$, and in the process we shall have proved that S_1 is not empty. Every s in S defines an evaluation linear functional e_s in S_1 by $e_s(f) = f(s)$, and the function p is defined by $p(s) = e_s$ for s in S . To see that S_1 is a closed subset of the unit ball of \mathcal{A}^* in the weak-star topology, let $\{\ell_\alpha\}$ be a net in S_1 converging to some $\ell \in \mathcal{A}^*$, the convergence being in the weak-star topology. Then we have $\ell_\alpha(fg) = \ell_\alpha(f)\ell_\alpha(g)$ and $\ell_\alpha(\bar{f}) = \overline{\ell_\alpha(f)}$ for all f and g in \mathcal{A} . Passing to the limit, we obtain $\ell(fg) = \ell(f)\ell(g)$ and $\ell(\bar{f}) = \overline{\ell(f)}$. Hence S_1 is closed. By Alaoglu's Theorem (Theorem 4.14a), S_1 is compact. It is Hausdorff since \mathcal{A}^* is Hausdorff in the weak-star topology.

Certainly we have $\sup_{s \in S} |e_s(f)| = \|f\|_{\text{sup}}$. Since any ℓ in S_1 has $\|\ell\| \leq 1$, we obtain

$$\sup_{\ell \in S_1} |\ell(f)| = \|f\|_{\text{sup}}. \quad (*)$$

The definition of $U : \mathcal{A} \rightarrow C(S_1)$ is $U(f)(\ell) = \ell(f)$, and this makes $U(f)(p(s)) = U(f)(e_s) = e_s(f) = f(s)$. The function $U(f)$ on S_1 is continuous by definition of the weak-star topology. Because of the definition of S_1 , U is an algebra homomorphism respecting complex conjugation and mapping 1 to 1.

¹²Checking that there are no other maximal ideals than the kernels of multiplicative linear functionals requires proving that every complex "Banach field" is 1-dimensional, an early result in the subject of Banach algebras and one that uses complex analysis in its proof. Details appear in Section 10.

Also, (*) shows that U is an isometry. Since \mathcal{A} is Cauchy complete, so is $U(\mathcal{A})$. Therefore $U(\mathcal{A})$ is a uniformly closed subalgebra of $C(S_1)$ stable under complex conjugation and containing the constants. It separates points of S_1 by the definition of equality of linear functionals. By the Stone–Weierstrass Theorem, $U(\mathcal{A}) = C(S_1)$. Since U is an isometry, U is one-one. Thus U is an algebra isomorphism of \mathcal{A} onto $C(S_1)$.

If $p(S)$ were not dense in $C(S_1)$, then Urysohn’s Lemma would allow us to find a nonzero continuous function F on $C(S_1)$ with values in $[0, 1]$ such that F is 0 everywhere on $p(S)$. Since U is onto $C(S_1)$, choose $f \in \mathcal{A}$ with $U(f) = F$. If s is in S , then $0 = F(p(s)) = U(f)(p(s)) = f(s)$. Hence $\|f\|_{\text{sup}} = 0$. By (*), $\ell(f) = 0$ for all $\ell \in S_1$. Then every $\ell \in S_1$ has $0 = \ell(f) = U(f)(\ell) = F(\ell)$, and $F = 0$, contradiction. We conclude that $p(S)$ is dense.

To see that U carries functions ≥ 0 to functions ≥ 0 , we observe first that the identity $\ell(\bar{f}) = \overline{\ell(f)}$ for $\ell \in S_1$ and the equality $\bar{\bar{f}} = f$ for f real together imply that $\ell(f) = \overline{\ell(\bar{f})} = \overline{\overline{\ell(f)}}$ for f real. Hence f real implies $\ell(f)$ real. If $f \geq 0$, then $\|\|f\|_{\text{sup}} - f\|_{\text{sup}} \leq \|f\|_{\text{sup}}$. Since $\|\ell\| \leq 1$, we therefore have $\ell(\|f\|_{\text{sup}} - f) \leq \|\|f\|_{\text{sup}} - f\|_{\text{sup}} \leq \|f\|_{\text{sup}}$. Since $\ell(1) = 1$, this says that $\ell(f) \geq 0$. This inequality for all ℓ implies that $U(f) \geq 0$.

Finally suppose that S is a Hausdorff topological space and that $\mathcal{A} \subseteq C(S)$. We are to show that $p : S \rightarrow S_1$ is continuous. If $s_\alpha \rightarrow s_0$ for a net in S , we want $p(s_\alpha) \rightarrow p(s_0)$, i.e., $e_{s_\alpha} \rightarrow e_{s_0}$. According to the definition of the weak-star topology, we are thus to show that $f(s_\alpha) \rightarrow f(s_0)$ for every f in \mathcal{A} . But this is immediate from the continuity of f on S .

We give three examples. A fourth example, concerning “almost periodic functions,” will be considered in the problems at the end of Chapter VI. For this fourth example the compact Hausdorff space of Theorem 4.15 admits the structure of a compact group, and the representation theory of Chapter VI is applicable to describe the structure of the space of almost periodic functions.

Problems 21–25 at the end of the chapter develop the theory of Theorem 4.15 further.

EXAMPLES.

(1) $\mathcal{A} = C(S)$ with S compact Hausdorff. Then p is a homeomorphism of S onto S_1 . In fact, $p(S)$ is always dense in S_1 . Here p is continuous and S is compact. Thus $p(S)$ is closed and must equal S_1 . The map p is one-one because Urysohn’s Lemma produces functions taking different values at two distinct points s and s' of S and thus exhibiting $e_{s'}$ and e_s as distinct linear functionals. Since p is continuous and one-one from a compact space onto a Hausdorff space, it is a homeomorphism.

(2) One-point compactification. Let S be a locally compact Hausdorff space, and let \mathcal{A} be the subalgebra of $C(S)$ consisting of all continuous functions having

limits at infinity. For a function f , this condition means that there is some number c such that for each $\epsilon > 0$, some compact subset K of S has the property that $|f(s) - c| \leq \epsilon$ for all s not in K . Then S_1 may be identified with the one-point compactification of S .

(3) Stone–Čech compactification. Let S be a topological space, and let $\mathcal{A} = C(S)$. The resulting compact Hausdorff space S_1 is called the **Stone–Čech compactification** of S . This space tends to be huge. For example, if $S = [0, +\infty)$, the corresponding S_1 has cardinality greater than the cardinality of \mathbb{R} .

5. Linear Functionals and Convex Sets

For this section and the next we discuss aspects of functional analysis that lead toward the theory of distributions and toward the use of fixed-point theorems. The topic is the role of convex sets in real and complex vector spaces—first without any topology and then with an overlay of topology consistent with convex sets. Sections 7–9 will then explore the consequences of this development, first in connection with smooth functions and then in connection with fixed-point theorems.

Let X be a real or complex vector space. A subset E of X is **convex** if for each x and y in E , all points $(1 - t)x + ty$ are in E for $0 \leq t \leq 1$.

Proposition 4.16. Convex sets in a real or complex vector space have the following elementary properties:

- (a) the arbitrary intersection of convex sets is convex,
- (b) if E is convex and x_1, \dots, x_n are in E and t_1, \dots, t_n are nonnegative reals with $t_1 + \dots + t_n = 1$, then $t_1x_1 + \dots + t_nx_n$ is in E ,
- (c) if E_1 and E_2 are convex, then so are $E_1 + E_2$, $E_1 - E_2$, and cE for any scalar c ,
- (d) if $L : X \rightarrow Y$ is linear between two vector spaces with the same scalars and if E is a convex subset of X , then $L(E)$ is convex in Y ,
- (e) if $L : X \rightarrow Y$ is linear between two vector spaces with the same scalars and if E is a convex subset of Y , then $L^{-1}(E)$ is convex in X .

PROOF. Conclusions (a), (c), (d), and (e) are completely straightforward. For (b), we induct on n , the case $n = 2$ being the definition of “convex.” Suppose that the result is known for n and that members x_1, \dots, x_{n+1} of X and nonnegative reals t_1, \dots, t_{n+1} with sum 1 are given. We may assume that $t_1 \neq 1$. Put $s = t_2 + \dots + t_{n+1}$ and $y = (1 - t_1)^{-1}(t_2x_2 + \dots + t_{n+1}x_{n+1})$. Since the reals $(1 - t_1)^{-1}t_2, \dots, (1 - t_1)^{-1}t_{n+1}$ are nonnegative and have sum 1, the inductive hypothesis shows that y is in E . Since t_1 and s are nonnegative and have sum 1, $t_1x_1 + sy = t_1x_1 + \dots + t_{n+1}x_{n+1}$ is in E . This completes the induction.

Let E be a subset of our vector space X . We say that a point p in E is an **internal point** of E if for each x in X , there is an $\epsilon > 0$ such that $p + \delta x$ is in E for all scalars¹³ δ with $|\delta| \leq \epsilon$. If p in X is neither an internal point of E nor an internal point of E^c , we say that p is a **bounding point** of E . These notions make no use of any topology on X .

Let K be a convex subset of X , and suppose that 0 is an internal point of K . For each x in X , let

$$\rho(x) = \inf\{a > 0 \mid a^{-1}x \in K\}.$$

The function $\rho(x)$ is called the **support function** of K . For an example let X be a normed linear space, and let K be the unit ball; then $\rho(x) = \|x\|$.

We are going to see that $\rho(x)$ has some bearing on controlling the linear functionals on X , as a consequence of the Hahn–Banach Theorem. By the “Hahn–Banach Theorem” here, we mean not the usual theorem for normed linear spaces¹⁴ but the more primitive statement¹⁵ from which that is derived:

HAHN–BANACH THEOREM. Let X be a *real* vector space, and let p be a real-valued function on X with

$$p(x + x') \leq p(x) + p(x') \quad \text{and} \quad p(tx) = tp(x)$$

for all x and x' in X and all real $t \geq 0$. If f is a linear functional on a vector subspace Y of X with $f(y) \leq p(y)$ for all y in Y , then there exists a linear functional F on X with $F(y) = f(y)$ for all $y \in Y$ and $F(x) \leq p(x)$ for all $x \in X$.

Before discussing linear functionals in our present context, let us observe some properties of the support function $\rho(x)$. Properties (b), (c), and (e) in the next lemma are the properties of the dominating function p in the Hahn–Banach Theorem as stated above.

Lemma 4.17. Let K be a convex subset of a vector space X , and suppose that 0 is an internal point. Then the support function $\rho(x)$ of K satisfies

- (a) $\rho(x) \geq 0$,
- (b) $\rho(x) < \infty$,
- (c) $\rho(ax) = a\rho(x)$ for $a \geq 0$,
- (d) $\rho(x) \leq 1$ for all x in K ,
- (e) $\rho(x + y) \leq \rho(x) + \rho(y)$,
- (f) $\rho(x) < 1$ if and only if x is an internal point of K ,
- (g) $\rho(x) = 1$ characterizes the bounding points of K .

¹³The scalars are complex numbers if X is complex, real numbers if X is real.

¹⁴As in Theorem 12.13 of *Basic*.

¹⁵As in Lemma 12.14 of *Basic*.

PROOF. Conclusions (a), (c), and (d) are immediate, and (b) follows since 0 is an internal point of K .

For (e), let c be arbitrary with $c > \rho(x) + \rho(y)$. We show that $c^{-1}(x + y)$ is in K . Since c is arbitrary, it follows that the infimum of all numbers d with $d^{-1}(x + y)$ in K is $\leq \rho(x) + \rho(y)$; consequently $\rho(x + y)$ will have to be $\leq \rho(x) + \rho(y)$, and (e) will be proved. Thus write $c = a + b$ with $a > \rho(x)$ and $b > \rho(y)$. Since K is convex,

$$c^{-1}(x + y) = (a + b)^{-1}(x + y) = \frac{a}{a+b} a^{-1}x + \frac{b}{a+b} b^{-1}y$$

is in K , as required.

For (f), let x be an internal point of K . Then $x + \epsilon x = (1 + \epsilon)x$ is in K for some $\epsilon > 0$, and hence $\rho(x) \leq (1 + \epsilon)^{-1} < 1$.

Conversely suppose that $\rho(x) < 1$, and put $\epsilon = 1 - \rho(x)$. Fix y . Since 0 is an internal point of K , we can find $\mu > 0$ such that δy is in K for $|\delta| \leq \mu$. If c is any scalar of absolute value 1, then $c\mu y$ is in K , and hence $\rho(cy) \leq \mu^{-1}$. If δ is a scalar with $|\delta| < \epsilon\mu$, write $\delta = c'|\delta|$ with $|c'| = 1$. Then $\rho(\delta y) = |\delta|\rho(c'y) \leq |\delta|\mu^{-1} < \epsilon$. Applying (e) gives

$$\rho(x + \delta y) \leq \rho(x) + \rho(\delta y) = (1 - \epsilon) + \rho(\delta y) < (1 - \epsilon) + \epsilon = 1.$$

By definition of ρ , $1^{-1}(x + \delta y)$ is in K , i.e., $x + \delta y$ is in K . Thus x is an internal point of K .

For (g), we can argue in the same way as with (f) to see that $\rho(x) > 1$ characterizes the internal points of K^c . Therefore $\rho(x) = 1$ characterizes the bounding points of K .

We shall now apply the Hahn–Banach Theorem to prove the basic separation theorem.

Theorem 4.18. Let M and N be disjoint nonempty convex subsets of a real or complex vector space X , and suppose that M has an internal point. Then there exists a nonzero linear functional F on X such that for some real c , $\operatorname{Re} F \leq c$ on M and $\operatorname{Re} F \geq c$ on N .

PROOF. First suppose that X is real. If m is an internal point of M , then 0 is an internal point of $M - m$, and we can replace M and N by $M - m$ and $N - m$. Changing notation, we may assume from the outset that 0 is an internal point of M .

If x_0 is in N , then $-x_0$ is an internal point of $M - N$, and 0 is an internal point of $K = M - N + x_0$. Since M and N are assumed disjoint, $M - N$ does not contain 0; thus K does not contain x_0 . Let ρ be the support function

of K ; this function satisfies the properties of the function p in the Hahn–Banach Theorem, according to Lemma 4.17. Moreover, $\rho(x_0) \geq 1$ by Lemma 4.17f. Define $f(ax_0) = a\rho(x_0)$ for all (real) scalars a . Then f is a nonzero linear functional on the 1-dimensional space of real multiples of x_0 , and it satisfies

$$\begin{aligned} a \geq 0 & \quad \text{implies} & \quad f(ax_0) = a\rho(x_0) = \rho(ax_0), \\ a < 0 & \quad \text{implies} & \quad f(ax_0) = af(x_0) < 0 \leq \rho(ax_0). \end{aligned}$$

The Hahn–Banach Theorem shows that f extends to a linear functional F on X with $F(x) \leq \rho(x)$ for all x . Then $F(x_0) \geq 1$, and Lemma 4.17 shows that $\rho(K) \leq 1$. Hence

$$F(x_0) \geq 1 \quad \text{and} \quad F(M - N + x_0) \leq 1.$$

Thus we have $F(M - N + x_0) \leq F(x_0)$, $F(M - N) \leq 0$, $F(m - n) \leq 0$ for all m in M and n in N , and $F(m) \leq F(n)$ for all m and n . Taking the supremum over m in M and the infimum over n in N gives the conclusion of the theorem for X real.

Now suppose that the vector space X is complex. We can initially regard X as a real vector space by forgetting about complex scalars, and then the previous case allows us to construct a real-linear F such that $F(M) \leq c \leq F(N)$. Put $G(x) = F(x) - iF(ix)$. Since $G(ix) = F(ix) - iF(i^2x) = F(ix) - iF(-x) = F(ix) + iF(x) = i(F(x) - iF(ix)) = iG(x)$, G is complex linear. The real part of G equals F , and therefore G satisfies the conclusion of the theorem.

6. Locally Convex Spaces

In this section we shall apply the discussion of convex sets and linear functionals in the context of topological vector spaces. A topological vector space X is said to be **locally convex** if there is a base for its topology that consists of convex sets.

Let us see that any topological vector space X whose topology is given by a family of seminorms $\|\cdot\|_s$ is locally convex. A base for the open sets consists of all finite intersections of sets $U(y, s, r) = \{x \mid \|x - y\|_s < r\}$ with y in X , s equal to one of the seminorm indices, and $r > 0$. If x and x' are in $U(y, s, r)$ and if $0 \leq t \leq 1$, then

$$\begin{aligned} \|((1-t)x + tx') - y\|_s &= \|(1-t)(x - y) + t(x' - y)\|_s \\ &\leq \|(1-t)(x - y)\|_s + \|t(x' - y)\|_s \\ &= (1-t)\|x - y\|_s + t\|x' - y\|_s \\ &< (1-t)r + tr = r. \end{aligned}$$

Hence $((1-t)x + tx')$ is in $U(y, s, r)$, and $U(y, s, r)$ is convex. Since the arbitrary intersection of convex sets is convex by Proposition 4.16a, every member of the base for the topology is convex. Thus X is locally convex.

We are going to show that every locally convex topological vector space has many continuous linear functionals, enough to distinguish any two disjoint closed convex sets when one of them is compact. This result will in particular be applicable to the spaces $\mathcal{S}(\mathbb{R}^N)$ and $C^\infty(U)$ since their topologies are given by seminorms.

We begin with two lemmas that do not need an assumption of local convexity on the topological vector space.

Lemma 4.19. In any topological vector space if K_1 and K_2 are closed sets with K_1 compact, then the set $K_1 - K_2$ of differences is closed.

PROOF. It is simplest to use nets. Thus let x be a limit point of $K_1 - K_2$, and let $\{x_n\}$ be any net in $K_1 - K_2$ converging to x . Since each x_n is in $K_1 - K_2$, we can write it as $x_n = k_n^{(1)} - k_n^{(2)}$ with $k_n^{(1)}$ in K_1 and $k_n^{(2)}$ in K_2 . Since K_1 is compact, $\{k_n^{(1)}\}$ has a convergent subnet, say $\{k_{n_j}^{(1)}\}$. Let $k^{(1)}$ be the limit of $\{k_{n_j}^{(1)}\}$ in K_1 . Both $\{x_{n_j}\}$ and $\{k_{n_j}^{(1)}\}$ are convergent, and $\{k_{n_j}^{(2)}\}$ must be convergent because $k_{n_j}^{(2)} = k_{n_j}^{(1)} - x_{n_j}$ and subtraction is continuous. Let k_2 be its limit. This limit has to be in K_2 since K_2 is closed, and then the equation $x = k^{(1)} - k^{(2)}$ exhibits x as in $K_1 - K_2$. Hence $K_1 - K_2$ is closed.

Lemma 4.20. Let X be any topological vector space, let K_1 and K_2 be disjoint convex sets, and suppose that K_1 has nonempty interior. Then there exists a nonzero continuous linear functional F on X with $\operatorname{Re} F(K_1) \leq c$ and $c \leq \operatorname{Re} F(K_2)$ for some real number c .

PROOF. The key observation is that any interior point of a subset E of X is internal. In fact, if p is in E° and x is in X , then $p + \delta x$ is in E° for $\delta = 0$. By continuity of the vector-space operations and openness of E° , $p + \delta x$ is in E° for $|\delta|$ sufficiently small. Therefore p is an internal point.

Since K_1 consequently has an internal point, Theorem 4.18 produces a nonzero linear functional F such that

$$\operatorname{Re} F(K_1) \leq c \quad \text{and} \quad c \leq \operatorname{Re} F(K_2) \quad (*)$$

for some real number c . We complete the proof of the lemma by showing that F is continuous. Let f and g be the real and imaginary parts of F . Then $g(x) = -if(ix)$, and it is enough to show that f is continuous. Fix an interior point p of K_1 , and choose an open neighborhood U of 0 such that $p + U \subseteq K_1$. Then

$f(U) \subseteq f(K_1) - f(p)$ since f is real linear, and (*) shows that $f(U) \leq c - f(p)$. So $f(U) \leq a$ for some $a > 0$. If $V = U \cap (-U)$, then

$$f(V) = f(U \cap (-U)) \subseteq f(U) \cap f(-U) = f(U) \cap (-f(U)) \subseteq [-a, a],$$

and therefore $f(\epsilon a^{-1}V) \subseteq [-\epsilon, \epsilon]$. In other words, f is continuous at 0. Then $f(x + \epsilon a^{-1}V) \subseteq f(x) + [-\epsilon, \epsilon]$, and f is continuous everywhere.

Theorem 4.21. Let X be a locally convex topological vector space, let K_1 and K_2 be disjoint closed convex subsets of X , and suppose that K_1 is compact. Then there exist $\epsilon > 0$, a real constant c , and a continuous linear functional F on X such that

$$\operatorname{Re} F(K_2) \leq c - \epsilon \quad \text{and} \quad c \leq \operatorname{Re} F(K_1).$$

PROOF. Lemma 4.19 shows that $K_1 - K_2$ is closed, and $K_1 - K_2$ does not contain 0 because K_1 and K_2 are disjoint. Since X is locally convex, we can choose a convex open neighborhood U of 0 disjoint from $K_1 - K_2$. Proposition 4.16c shows that $K_1 - K_2$ is convex, and Lemma 4.20 therefore applies to the sets U and $K_1 - K_2$ and yields a nonzero continuous linear functional F such that

$$\operatorname{Re} F(U) \leq d \quad \text{and} \quad d \leq \operatorname{Re} F(K_1 - K_2)$$

for some real d . Since F is not zero, we can find x_0 in X with $F(x_0) = 1$. Choose $\epsilon > 0$ such that $|a| < \epsilon$ implies ax_0 is in U . Then

$$d \geq \operatorname{Re} F(U) \supseteq \operatorname{Re} F(\{ax_0 \mid |a| < \epsilon\}) = (-\epsilon, \epsilon),$$

and hence $d \geq \epsilon$. Therefore all k_1 in K_1 and k_2 in K_2 have

$$\operatorname{Re} F(k_1) - \operatorname{Re} F(k_2) = \operatorname{Re} F(k_1 - k_2) \geq d \geq \epsilon,$$

so that $\operatorname{Re} F(k_1) \geq \epsilon + \operatorname{Re} F(k_2)$. Taking $c = \inf_{k_1 \in K_1} \operatorname{Re} F(k_1)$ now yields the conclusion of the theorem.

Corollary 4.22. Let X be a locally convex topological vector space, let K be a closed convex subset of X , and let p be a point of X not in K . Then there exists a continuous linear functional F on X such that

$$\sup_{k \in K} \operatorname{Re} F(k) < \operatorname{Re} F(p).$$

PROOF. This is the special case of Theorem 4.21 in which the given compact set is a singleton set.

Corollary 4.23. If X is a locally convex topological vector space and if p and q are distinct points of X , then there exists a continuous linear functional F on X such that $F(p) \neq F(q)$.

PROOF. This is the special case of Corollary 4.22 in which the given closed convex set is a singleton set.

We conclude this section with a simple result about locally convex topological vector spaces that we shall need in the next section.

Proposition 4.24. If X is a locally convex topological vector space and Y is a closed vector subspace, then the topological vector space X/Y is locally convex.

REMARK. X/Y is a topological vector space by Proposition 4.4.

PROOF. Let E be an open neighborhood of a given point of X/Y . Without loss of generality, we may take the given point to be the 0 coset. If $q : X \rightarrow X/Y$ is the quotient map, $q^{-1}(E)$ is an open neighborhood of 0 in X . Since X is locally convex, there is a convex open neighborhood U of 0 in X with $U \subseteq q^{-1}(E)$. The map q carries open sets to open sets by Proposition 4.4 and carries convex sets to convex sets by Proposition 4.16d, and thus $q(U)$ is an open convex neighborhood of the 0 coset in X/Y contained in E .

7. Topology on $C_{\text{com}}^{\infty}(U)$

In this section we carry the discussion of local convexity in Sections 5–6 along the path toward applications to smooth functions. Our objective will be to topologize the space $C_{\text{com}}^{\infty}(U)$ of smooth functions of compact support on the open set U of \mathbb{R}^N . The members of $C_{\text{com}}^{\infty}(U)$ extend to functions in $C_{\text{com}}^{\infty}(\mathbb{R}^N)$ by defining them to be 0 outside U , and we often make this identification without special comment.

The important thing about the topology will be what it accomplishes, rather than what the open sets are, and we shall therefore work toward a characterization of the topology, together with an existence proof. The characterization will be in terms of a universal mapping property, and local convexity will be part of that property. Ultimately it is possible to give an explicit description of the open sets, but we leave such a description for Problem 9 at the end of the chapter. The explicit description will show in particular that the topology is given by an uncountable family of seminorms that cannot be reduced to a countable family except when U is empty.

Let us state the universal mapping property informally now, so that the ingredients become clear. Let K be any compact subset of the given open set U of \mathbb{R}^N ,

and define C_K^∞ to be the vector space of all smooth functions of compact support on \mathbb{R}^N with support contained in K . The space C_K^∞ becomes a locally convex topological vector space when we impose the countable family of seminorms $\|f\|_\alpha = \sup_{x \in K} |D^\alpha f(x)|$, with α running over all differentiation multi-indices. Set-theoretically, $C_{\text{com}}^\infty(U)$ is the union of all C_K^∞ as K runs through the compact subsets of U . The topology on $C_{\text{com}}^\infty(U)$ will be arranged so that

- (i) every inclusion $C_K^\infty \subseteq C_{\text{com}}^\infty(U)$ is continuous,
- (ii) whenever a linear mapping $C_{\text{com}}^\infty(U) \rightarrow X$ is given into a locally convex linear topological space X and the composition $C_K^\infty \rightarrow C_{\text{com}}^\infty(U) \rightarrow X$ is continuous for every K , then the given mapping $C_{\text{com}}^\infty(U) \rightarrow X$ is continuous.

It will automatically have the additional property

- (iii) every inclusion $C_K^\infty \subseteq C_{\text{com}}^\infty(U)$ is a homeomorphism with its image.

We shall proceed somewhat abstractly, so as to be able to construct the topology of a locally convex topological vector space out of simpler data. If (X, \mathcal{T}) is a topological space and p is in X , we define a **local neighborhood base** for \mathcal{T} at p to be a collection \mathcal{N}_p of neighborhoods of p , not necessarily open, such that if V is any open set containing p , then there exists N in \mathcal{N}_p with $N \subseteq V$. If X is a topological vector space with topology \mathcal{T} and if \mathcal{N}_0 is a local neighborhood base at 0, then $\{p + N \mid N \in \mathcal{N}_0\}$ is a local neighborhood base at p because translation by x is a homeomorphism. A set is open if and only if it is a neighborhood of each of its points. Consequently we can recover \mathcal{T} from a local neighborhood base \mathcal{N}_0 at 0 by this description: a subset V of X is open if and only if for each p in V , there exists N_p in \mathcal{N}_0 such that $p + N_p \subseteq V$.

Let us observe two properties of a local neighborhood base \mathcal{N}_0 at 0 for a topological vector space X . The first follows from the fact that X is Hausdorff, more particularly that each one-point subset of X is closed. The property is that for each $x \neq 0$ in X , there is some M_x in \mathcal{N}_0 with x not in M_x .

The second follows from the fact that 0 is an interior point of each member N of \mathcal{N}_0 . The property is that 0 is an internal point of N in the sense of Section 5. The fact that interior implies internal was proved in the first paragraph of the proof of Lemma 4.20.

We shall show in Lemma 4.25 that we can arrange in the locally convex case for each member N of a local neighborhood base \mathcal{N}_0 at 0 to have the additional property of being **circled** in the sense that $zN \subseteq N$ for all scalars z with $|z| \leq 1$.

Then we shall see in Proposition 4.26 that we can formulate a tidy necessary and sufficient condition for a system of sets containing 0 in a real or complex vector space X to be a local neighborhood base for a topology on X that makes X into a locally convex topological vector space.

Lemma 4.25. Any locally convex topological vector space has a local neighborhood base at 0 consisting of convex circled sets.

PROOF. It is enough to show that if V is an open neighborhood of 0, then there is an open subneighborhood U of 0 that is convex and circled. Since the underlying topological vector space is locally convex, we may assume that V is convex. Replacing V by $V \cap (-V)$, we may assume by parts (a) and (c) of Proposition 4.16 that V is stable under multiplication by -1 . Since V is convex, it follows that $cV \subseteq V$ for any real c with $|c| \leq 1$. If the field of scalars is \mathbb{R} , then the proof of the lemma is complete at this point.

Thus suppose that the field of scalars is \mathbb{C} . If V is a convex open neighborhood of 0, put

$$W = \{u \in V \mid zu \in V \text{ for all } z \in \mathbb{C} \text{ with } |z| \leq 1\}.$$

Then W is convex by Proposition 4.16a, and it is circled. Let us show that $W \supseteq \frac{1}{2}V \cap \frac{1}{2}iV$. Thus let u be an element of $\frac{1}{2}V \cap \frac{1}{2}iV$, and write it as $u = \frac{1}{2}v_1 = \frac{1}{2}iv_2$ with v_1 and v_2 in V . Let $z \in \mathbb{C}$ be given with $|z| \leq 1$, and let x and y be the real and imaginary parts of z . The vectors $\pm v_1$ and 0 are in V , and V is convex; since $|x| \leq 1$, xv_1 is in V . Similarly $-yv_2$ is in V . We can write $zu = \frac{1}{2}(x + iy)v_1 = \frac{1}{2}(xv_1) + \frac{1}{2}(-yv_2)$, and this is in V since V is convex. Therefore zu is in V , and u is in U . Hence $W \supseteq \frac{1}{2}V \cap \frac{1}{2}iV$, as asserted.

Let U be the interior W° of W . Then U is an open neighborhood of 0, and we show that it is convex and circled; this will complete the proof. Let u and v be in U . Since U is open, we can find an open neighborhood N of 0 such that $u + N \subseteq U$ and $v + N \subseteq U$. If n is in N and if t satisfies $0 \leq t \leq 1$, then $(1 - t)u + tv + n = (1 - t)(u + n) + t(v + n)$ exhibits $(1 - t)u + tv + n$ as a convex combination of a member of $u + N \subseteq W$ and a member of $v + N \subseteq W$, hence as a member of W . Therefore every member of $(1 - t)u + tv + N$ lies in W , and U is convex.

To see that U is circled, let u and N be as in the previous paragraph with $u + N \subseteq U$. If $|z| \leq 1$, then $u + N \subseteq W$ implies $z(u + N) \subseteq W$ since W is circled. Hence $zu + zN \subseteq W$. Since zN is open, $zu + zN$ is an open neighborhood of zu contained in W , and we must have $zu + zN \subseteq W^\circ = U$. Therefore U is circled.

Proposition 4.26. Let X be a real or complex vector space. If X has a topology making it into a locally convex topological vector space, then X has a local neighborhood base \mathcal{N}_0 at 0 for that topology such that

- (a) each N in \mathcal{N}_0 is convex and circled with 0 as an internal point,
- (b) whenever M and N are in \mathcal{N}_0 , there is some P in \mathcal{N}_0 with $P \subseteq M \cap N$,
- (c) whenever N is in \mathcal{N}_0 and a is a nonzero scalar, then aN is in \mathcal{N}_0 ,
- (d) each $x \neq 0$ in X has some associated M_x in \mathcal{N}_0 such that x is not in M_x .

Conversely if \mathcal{N}_0 is any family of subsets of the vector space X such that (a), (b), (c), and (d) hold, then there exists one and only one topology on X making X into a locally convex topological vector space in such a way that \mathcal{N}_0 is a local neighborhood base at 0.

PROOF. For the direct part of the proof, Lemma 4.25 shows that there is some local neighborhood base at 0 consisting of convex circled sets. To such a local neighborhood base we are free to add any additional neighborhoods of 0. Thus we may take \mathcal{N}_0 to consist of *all* convex circled neighborhoods of 0. Then (b) and (c) hold, and (d) holds since the topology is Hausdorff. Since 0 is an internal point of any neighborhood of 0, (a) holds. This proves existence.

For the converse there is only one possibility for the topology \mathcal{T} : V is open if for each x in V , there is some N_x in \mathcal{N}_0 with $x + N_x \subseteq V$. This proves the uniqueness of \mathcal{T} , and we are to prove existence. For existence we define open sets in this way and define \mathcal{T} to be the collection of all open sets. The definition makes \emptyset open and the arbitrary union of open sets open, and (b) makes the intersection of two open sets open.

We shall show that the complement of any $\{x_0\}$ is open. Then it follows by taking unions that X is open, so that \mathcal{T} is a topology; also we will have proved that every one-point set is closed. If $x_1 \neq x_0$, we use (d) to choose $M_{x_0-x_1}$ in \mathcal{N}_0 with $x_0 - x_1$ not in $M_{x_0-x_1}$. Then $x_1 + M_{x_0-x_1} \subseteq X - \{x_0\}$. Since x_1 is arbitrary, $X - \{x_0\}$ is open.

With \mathcal{T} established as a topology, let us see that every member of \mathcal{N}_0 is a neighborhood of 0. This step involves considering the family of sets aN for fixed N in \mathcal{N}_0 and for arbitrary positive a . If $0 < t < 1$ and if n_1 and n_2 are in N , then $(1-t)n_1 + tn_2$ is in N since (a) says that N is convex. Hence $(1-t)N + tN \subseteq N$. If $a > 0$ and $b > 0$, then we can take $t = b(a+b)^{-1}$ and obtain $a(a+b)^{-1}N + b(a+b)^{-1}N \subseteq N$. Multiplying by $a+b$ gives

$$aN + bN \subseteq (a+b)N \quad \text{for all positive } a \text{ and } b. \quad (*)$$

In particular the sets aN are nested for $a > 0$, i.e., $0 < a < a'$ implies $aN \subseteq a'N$.

From these facts we can show that each N in \mathcal{N}_0 is a neighborhood of 0. Given N , define $U = \bigcup_{0 < a < 1} aN$. This is a subset of N by the nesting property, and we shall prove that it is open. If x is in U , then x is in aN for some a with $0 < a < 1$, and (*) shows that $x + \frac{1}{2}(1-a)N \subseteq U$. By (c), $\frac{1}{2}(1-a)N$ is in \mathcal{N}_0 , and therefore $\frac{1}{2}(1-a)N$ can serve as a member N_x of \mathcal{N}_0 such that $x + N_x \subseteq U$. We conclude that U is open. Therefore N is a neighborhood of 0.

Next let us see that translations are homeomorphisms. If V is open and if x_0 is given, we know that each x in V has an associated N_x such that $x + N_x \subseteq V$. If y is in $x_0 + V$, then $x = y - x_0$ is in V and we see that $(y - x_0) + N_{y-x_0} \subseteq V$ and $y + N_{y-x_0} \subseteq x_0 + V$. Hence $x_0 + V$ is open, and every translation is a homeomorphism.

Let us see that addition is continuous at $(0, 0)$, and then the fact that translations are homeomorphisms implies that addition is continuous everywhere. If V is an open neighborhood of 0 , then the definition of open set says that there is some N in \mathcal{N}_0 with $0 + N \subseteq V$. By (c), $\frac{1}{2}N$ is in \mathcal{N}_0 . It is enough to prove that $(\frac{1}{2}N, \frac{1}{2}N)$ maps into V under addition. But this is immediate from (*) since $\frac{1}{2}N + \frac{1}{2}N \subseteq N \subseteq V$.

Next we investigate continuity of the mapping $x \mapsto ax$ for $a \neq 0$. It is enough to show that if V is open, then so is $a^{-1}V$. Since V is open, every x in V has an associated N_x in \mathcal{N}_0 such that $x + N_x \subseteq V$. The most general element of $a^{-1}V$ is of the form $a^{-1}x$ with x in V , and we have $a^{-1}x + a^{-1}N_x \subseteq a^{-1}V$. Since (c) shows $a^{-1}N_x$ to be in \mathcal{N}_0 , we conclude that $a^{-1}V$ is open.

Let us see that scalar multiplication is continuous at $(1, x)$, and then the fact that $x \mapsto ax$ is continuous for $a \neq 0$ implies that scalar multiplication is continuous everywhere except possibly at $(0, x)$. Let V be an open neighborhood of x , and choose N in \mathcal{N}_0 with $x + N \subseteq V$. Since N is in \mathcal{N}_0 , (c) shows that $\frac{1}{3}N$ is in \mathcal{N}_0 . Then 0 is an internal point of $\frac{1}{3}N$ by (a), and there exists $\epsilon > 0$ such that $-\epsilon \leq c \leq \epsilon$ implies that cx is in $\frac{1}{3}N$. There is no loss of generality in taking $\epsilon < 1$. Since $\frac{1}{3}N$ is circled by (a), cx is in $\frac{1}{3}N$ for $|c| \leq \epsilon$. Let A be the set of scalars with $|a - 1| < \epsilon$. We show that scalar multiplication carries $A \times (x + \frac{1}{3}N)$ into V . In fact, if a is in A and $\frac{1}{3}n_1$ is in $\frac{1}{3}N$, then $|a| < 2$, $\frac{1}{3}an_1$ is in $\frac{2}{3}N$, and (*) gives

$$a(x + \frac{1}{3}n_1) = (ax - x) + (x + \frac{1}{3}an_1) \in \frac{1}{3}N + (x + \frac{2}{3}N) \subseteq x + N \subseteq V.$$

To complete the proof of continuity of scalar multiplication, we show continuity at all points $(0, x)$. Let V be an open neighborhood of 0 in X , and choose N in \mathcal{N}_0 with $0 + N \subseteq V$. Since 0 is an internal point of N , there is some $\epsilon > 0$ such that cx is in N for real c with $|c| \leq \epsilon$. For this ϵ , $\frac{1}{2}\epsilon x$ is in $\frac{1}{2}N$. If $|z| < 1$ and y is in $\frac{1}{2}N$, then $(z, \frac{1}{2}\epsilon x + y)$ maps to $\frac{1}{2}z\epsilon x + zy$, which lies in $\frac{1}{2}N + \frac{1}{2}N$ since N is circled. In turn, this is contained in N by (*) and therefore is contained in V . So $(\frac{1}{2}\epsilon z, x + 2\epsilon^{-1}y)$ maps into V if $|z| < 1$ and y is in $\frac{1}{2}N$. Altering the definitions of z and y , we conclude that $(z, x + y)$ maps into V if $|z| < \frac{1}{2}\epsilon$ and y is in $\epsilon^{-1}N$. This proves the continuity.

Since $\{0\}$ is a closed set, Lemma 4.3 is applicable and shows that X is Hausdorff, hence is a topological vector space. Inside any open neighborhood V of 0 lies some set N in \mathcal{U}_0 , and $\bigcup_{0 < a < 1} aN$ is a convex open subneighborhood of V . Therefore the topology is locally convex.

We are almost in a position to topologize $C_{\text{com}}^{\infty}(U)$. If i_K denotes the inclusion of C_K^{∞} into $C_{\text{com}}^{\infty}(U)$, we shall define a convex circled subset N in $C_{\text{com}}^{\infty}(U)$

having 0 as an internal point to be in a local neighborhood base at 0 if $i_K^{-1}(N)$ is a neighborhood of 0 in C_K^∞ for every compact subset K of U . Then conditions (a), (b), and (c) in Proposition 4.26 will be met, and an examination of the proof of that proposition shows that we obtain a topology for $C_{\text{com}}^\infty(U)$ in which addition and scalar multiplication are continuous. What is lacking is the Hausdorff property, which follows once (d) holds in Proposition 4.26. Verifying (d) requires a construction, whose main step is given in the following lemma.

Lemma 4.27. Let X be a locally convex topological vector space, let Y be a closed vector subspace, and let Y be given the relative topology, which is locally convex. If N is a convex circled neighborhood of 0 in Y and x_0 is a point in X not in N , then there exists a convex circled neighborhood M of 0 in X such that $M \cap Y = N$ and such that x_0 is not in M .

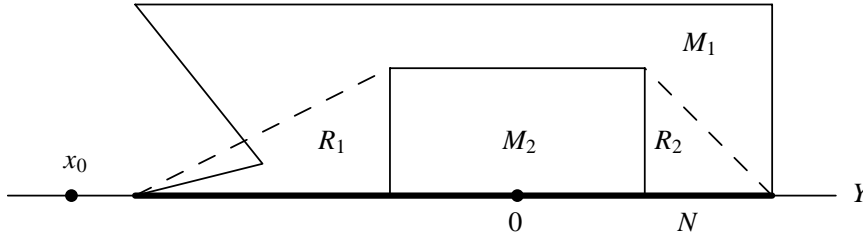


FIGURE 4.1. Extension of convex circled neighborhood of 0.
The lemma extends N to the set given in the figure
by $M_3 = R_1 \cup M_2 \cup R_2$.

PROOF. Since N is a neighborhood of 0 in Y and since Y has the relative topology, there exists a neighborhood M_1 of 0 in X such that $M_1 \cap Y = N$. We shall adjust M_1 to make it convex circled and to arrange that x_0 is not in it. Since X is locally convex, we can find a convex circled neighborhood M_2 of 0 contained in M_1 . Taking a cue from Figure 4.1, define

$$M_3 = \{(1-t)n + tm_2 \mid n \in N, m_2 \in M_2, 0 \leq t \leq 1\}.$$

This is a neighborhood of 0 since it contains M_2 , and it is convex circled since N and M_2 are convex circled.

We shall prove that

$$M_3 \cap Y = N.$$

Certainly $M_3 \cap Y \supseteq N$. For the reverse inclusion let m_3 be in $M_3 \cap Y$, and write $m_3 = (1-t)n + tm_2$ with $n \in N$, $m_2 \in M_2$, and $0 \leq t \leq 1$. If $t = 0$, then $m_3 = n$ is already in N . If $t > 0$, then $m_2 = t^{-1}(m_3 - (1-t)n)$ exhibits m_2 as a

linear combination of members of Y , hence as a member of Y . Since $M_2 \subseteq M_1$, m_2 is in $M_1 \cap Y = N$. Therefore m_3 is a convex combination of the members n and m_2 of N and must lie in N since N is convex. Consequently $M_3 \cap Y = N$.

If x_0 lies in Y , then we can take $M = M_3$ since x_0 is by assumption not in N and cannot therefore be in the larger set M_3 . If x_0 is not in Y , then Proposition 4.24 says that X/Y is a locally convex topological vector space. Since $x_0 + Y$ is not the 0 coset, we can find a convex circled neighborhood P of the 0 coset that does not contain $x_0 + Y$. If $q : X \rightarrow X/Y$ is the quotient map, then $q^{-1}(P)$ by Proposition 4.16e is a convex circled neighborhood of 0 in X that does not contain x_0 and satisfies $q^{-1}(P) \cap Y = Y$. Therefore $M = M_3 \cap q^{-1}(P)$ is a convex circled neighborhood of 0 in X that does not contain x_0 and satisfies $M \cap Y = N$.

Proposition 4.28. Let X be a real or complex vector space, and suppose that X is the increasing union $X = \bigcup_{p=1}^{\infty} X_p$ of a sequence of locally convex topological vector spaces such that for each p , X_p is a closed vector subspace of X_{p+1} and has the relative topology. Then there exists a unique topology on X making it into a locally convex topological vector space in such a way that

- (a) each inclusion $i_p : X_p \rightarrow X$ is continuous,
- (b) whenever $L : X \rightarrow Y$ is a linear function from X into a locally convex topological vector space Y such that $L \circ i_p : X_p \rightarrow Y$ is continuous for all p , then L is continuous.

This unique topology has the property that

- (c) each inclusion $i_p : X_p \rightarrow X$ is a homeomorphism with its image.

PROOF. Let \mathcal{N}_0 be the family of all convex circled subsets N of X having 0 as an internal point such that $i_p^{-1}(N)$ is a neighborhood of 0 in X_p for all p . We shall prove that \mathcal{N}_0 satisfies the four conditions (a) through (d) of Proposition 4.26, so that X has a unique topology making it into a locally convex topological vector space in such a way that \mathcal{N}_0 is a local neighborhood base at 0. Condition (a) holds by definition. Condition (b) holds because the intersection of two convex circled subsets with 0 as an internal point is again a convex circled set with 0 as an internal point and because the intersection of two neighborhoods is a neighborhood. Condition (c) holds because multiplication by a nonzero scalar sends convex circled sets with 0 as an internal point into convex circled sets with 0 as an internal point and because multiplication by a nonzero scalar sends neighborhoods of 0 to neighborhoods of 0.

We have to prove (d) in Proposition 4.26, namely that each $x_0 \neq 0$ in X has some associated M in \mathcal{N}_0 such that x_0 is not in M . Since $X = \bigcup_{p=1}^{\infty} X_p$, choose p_0 as small as possible so that x_0 is in X_{p_0} . Since X_{p_0} satisfies (a) through (d) and since $x_0 \neq 0$, we can find some convex circled neighborhood M_{p_0} of 0 in X_{p_0} that

does not contain x_0 . Proceeding inductively by means of Lemma 4.27, we can find, for each $p > p_0$, a convex circled neighborhood M_p of 0 in X_p that does not contain x_0 such that $M_p \cap X_{p-1} = M_{p-1}$. Define $M = \bigcup_{p \geq p_0} M_p$. Then M is convex circled since each M_p has this property. To see that 0 is an internal point of M , we argue as follows: for each x in X , x lies in some X_p , the set M_p has 0 as an internal point since M_p is a neighborhood of 0, M_p contains all cx for c real and small, and the larger set M contains all cx for c real and small. For each $p \geq p_0$, the set $i_p^{-1}(M)$ equals M_p , which was constructed as a neighborhood of 0 in X_p . The intersection $i_k^{-1}(M) = M_p \cap X_k$ has to be a neighborhood of 0 in X_k for $k < p$ since M_p is a neighborhood of 0 in X_p , and the set M is therefore in \mathcal{N}_0 . Thus M meets the requirement of being a member of \mathcal{N}_0 that does not contain x_0 , and (d) holds in Proposition 4.26.

We are left with proving (a) through (c) in the present proposition and with proving that no other topology meets these conditions. For (a), since i_p is linear, it is enough to prove continuity at 0. Hence we are to see that if N is in \mathcal{N}_0 , then $i_p^{-1}(N)$ is a neighborhood of 0 in X_p . But this is just one of the defining conditions for the set N to be in \mathcal{N}_0 .

For (b), since L is linear, it is enough to prove continuity at 0. Since Y is locally convex, the convex circled neighborhoods of 0 in Y form a local neighborhood base. If E is such a neighborhood, we are to show that $N = L^{-1}(E)$ is a neighborhood of 0 in X . The set E is convex and circled with 0 as an internal point, and hence the same thing is true of N . Also, $i_p^{-1}(N) = i_p^{-1}L^{-1}(E) = (L \circ i_p)^{-1}(E)$ is a neighborhood of 0 in X_p since $L \circ i_p$ is by assumption continuous. Therefore $N = L^{-1}(E)$ is in \mathcal{N}_0 , and then $L^{-1}(E)$ is a neighborhood of 0 in the topology imposed on X . Hence L is continuous at 0 and is continuous.

For (c), we again use Lemma 4.27, except that this time we do not need a point x_0 . We are to show that if N_{p_0} is a neighborhood of 0 in X_{p_0} , then $i(N_{p_0})$ is a neighborhood of 0 in the relative topology that X defines on X_{p_0} . Since X_{p_0} is locally convex, there is no loss of generality in assuming that N_{p_0} is convex circled. Proceeding inductively for $p > p_0$, we use the lemma to construct a convex circled neighborhood N_p of 0 in X_p such that $N_p \cap X_{p-1} = N_{p-1}$. Put $N = \bigcup_{p \geq p_0} N_p$. Arguing in the same way as earlier in the proof, we see that N is in \mathcal{N}_0 . Then $i(N_{p_0}) = X_{p_0} \cap N$, and $i(N_{p_0})$ is exhibited as the intersection of X_{p_0} with a neighborhood of 0 in X . This proves (c).

Finally suppose that the constructed topology on X is \mathcal{T} and that \mathcal{T}' is a second topology making X into a locally convex topological vector space in such a way that (a) and (b) hold. Let $1_{\mathcal{T}}$ be the identity map from (X, \mathcal{T}) to (X, \mathcal{T}') . By (a) for \mathcal{T}' , the composition $1_{\mathcal{T}} \circ i_p : X_p \rightarrow X$ is continuous. By (b) for \mathcal{T} , $1_{\mathcal{T}}$ is continuous from (X, \mathcal{T}) to (X, \mathcal{T}') . Reversing the roles of \mathcal{T} and \mathcal{T}' , we see that the identity map is continuous from (X, \mathcal{T}') to (X, \mathcal{T}) . Therefore $1_{\mathcal{T}}$ is a homeomorphism.

In the terminology of abstract functional analysis, one says that X in Proposition 4.28 is a **strict inductive limit**¹⁶ of the spaces X_p . With extra hypotheses that are satisfied in our case of interest, one says that X acquires the **LF topology**¹⁷ from the X_p 's.

Now let us apply the abstract theory to $C_{\text{com}}^\infty(U)$. If $\{K_p\}$ is any exhausting sequence of compact subsets of U , then we apply Proposition 4.28 with $X = C_{\text{com}}^\infty(U)$ and $X_p = C_{K_p}^\infty$. For the inclusion $X_p \subseteq X_{p+1}$, the restriction to $C_{K_p}^\infty$ of the seminorms on $C_{K_{p+1}}^\infty$ yields the seminorms for $C_{K_p}^\infty$, and therefore X_p has the relative topology as a vector subspace of X_{p+1} . The space X_p is a closed subspace because $C_{K_p}^\infty$ is Cauchy complete and because complete subsets of a metric space are closed. Thus the hypotheses are satisfied, and $C_{\text{com}}^\infty(U)$ acquires a unique topology as a locally convex topological vector space such that

- (i) each inclusion $C_{K_p}^\infty \subseteq C_{\text{com}}^\infty(U)$ is continuous,
- (ii) whenever a linear mapping $C_{\text{com}}^\infty(U) \rightarrow X$ is given into a locally convex linear topological space X and the composition $C_{K_p}^\infty \rightarrow C_{\text{com}}^\infty(U) \rightarrow X$ is continuous for every p , then the given mapping $C_{\text{com}}^\infty(U) \rightarrow X$ is continuous.

Furthermore

- (iii) each inclusion $C_{K_p}^\infty \subseteq C_{\text{com}}^\infty(U)$ is a homeomorphism with its image.

To complete our construction, all we have to do is show that the resulting topology on $C_{\text{com}}^\infty(U)$ does not depend on the choice of exhausting sequence.

Proposition 4.29. The inductive limit topology on $C_{\text{com}}^\infty(U)$ is independent of the choice of exhausting sequence. Consequently

- (a) each inclusion $C_K^\infty \subseteq C_{\text{com}}^\infty(U)$ is a homeomorphism with its image,
- (b) whenever a linear mapping $C_{\text{com}}^\infty(U) \rightarrow X$ is given into a locally convex linear topological space X and the composition $C_K^\infty \rightarrow C_{\text{com}}^\infty(U) \rightarrow X$ is continuous for every compact subset K of U , then the given mapping $C_{\text{com}}^\infty(U) \rightarrow X$ is continuous.

¹⁶The words “direct limit” mean the same thing as “inductive limit,” but “inductive” is more common in this situation. The term “strict” refers to the fact that the successive inclusions $i_{p+1,p} : X_p \rightarrow X_{p+1}$ are one-one with $i_{p+1,p}(X_p)$ homeomorphic to X_p . The notion of “direct limit” is a construction in category theory that is useful within several different categories. Uniqueness of the direct limit up to canonical isomorphism is a formality built into the definition; existence depends on the particular category. For this situation the construction is taking place within the category of locally convex topological vector spaces (and continuous linear maps). A direct-limit construction within a different category plays a role in Problems 26–30 at the end of the chapter, and those problems are continued at the end of Chapter VI.

¹⁷“ LF ” refers to “Fréchet limit.” In the usual situation the spaces X_p are assumed to be locally convex complete metric topological vector spaces, i.e., “Fréchet spaces.” The X_p 's have this property in the application to $C_{\text{com}}^\infty(U)$.

PROOF. Write X for $C_{\text{com}}^{\infty}(U)$ with its topology defined relative to an exhausting sequence $\{K_p\}$ of compact subsets of U , and write Y for $C_{\text{com}}^{\infty}(U)$ with its topology defined relative to an exhausting sequence $\{K'_p\}$. If K_k is a member of the sequence $\{K_p\}$, then $K_k \subseteq K'_p$ for $p \geq$ some index p_0 depending on k since the interiors of the sets K'_p cover the compact set K_k . The inclusion $K_k \subseteq K'_p$ is continuous for $p \geq p_0$, and therefore the composition $K_k \rightarrow K'_{p_0} \rightarrow Y$ is continuous. This continuity for all k implies that the identity map from X into Y is continuous. Reversing the roles of X and Y , we see that the identity map is a homeomorphism.

8. Krein–Milman Theorem

In this section we carry the discussion of local convexity in Sections 5–6 along the path toward fixed-point theorems. Our objective will be to prove a fundamental existence theorem about “extreme points.”

If K is a convex set in a real or complex vector space and if x_0 is in K , we say that x_0 is an **extreme point** of K if x_0 is not in the interior of any **line segment** belonging to K , i.e., if

$$x_0 = (1-t)x + ty \text{ with } 0 < t < 1 \text{ and } x, y \in K \quad \text{implies} \quad x_0 = x = y.$$

Let X be a topological vector space, and let K be a closed convex subset of X . A nonempty closed convex subset S of K is called a **face** if whenever ℓ is a line segment belonging to K , in the above sense, and ℓ has an interior point in S , then the whole line segment belongs to S . With this definition, x_0 is an extreme point of K if and only if the singleton set $\{x_0\}$ is a face.

If E is a subset of X , then the **closed convex hull** of E is defined to be the intersection of all closed convex subsets of X that contain E . It may be described explicitly as the closure of the set of all convex combinations of members of E .

Theorem 4.30 (Krein–Milman Theorem). If K is a compact convex set in a locally convex topological vector space, then K is the closed convex hull of the set of extreme points of K . In particular, if K is nonempty, then K has an extreme point.

PROOF. Let X be the underlying topological vector space. We may assume, without loss of generality, that K is nonempty. Let us see that if f is any continuous linear functional on X , then the subset of K on which $\text{Re} f$ assumes its maximum value is a face. In fact, let S be the subset where $g = \text{Re} f$ assumes its maximum value m . Then S is nonempty since K is compact and g is continuous, and the continuity and real linearity of g imply that S is closed and convex. To