Graduate Texts in Mathematics

Richard B. Holmes

Geometric Functional Analysis and its Applications

Springer-Verlag
New York Heidelberg Berlin
World Publishing Corporation Beijing China

Richard B. Holmes

Geometric Functional Analysis and its Applications

Springer-Verlag New York Heidelberg Berlin
World Publishing Corporation, Beijing, China

Richard B. Holmes

Purdue University
Division of Mathematical Sciences
West Lafayette, Indiana 47907

Reprinted in China by World Publishing Corporation
For distribution and sale in the People's Republic of China only
华人民共和国发行

AMS Subject Classifications
Primary: 46.01, 46N05
Secondary: 46A05, 46B10, 52A05, 41A65

Library of Congress Cataloging in Publication Data

Holmes, Richard B.
Geometric functional analysis and its applications.
(Graduate texts in mathematics; v. 24)
Bibliography: p. 237
Includes index.
1. Functional analysis. I. Title. II. Series.
QA320.H63 515'.7 75 6803

Reprinted in China by World Publishing Corporation
For distribution and sale in the People's Republic of China only
只限在中华人民共和国发行
ISBN 0-387-90136-1 Springer-Verlag New York Heidelberg Berlin
ISBN 3-540-90136-1 Springer-Verlag Berlin Heidelberg New York
ISBN 7-5062-0054-6 World Publishing Corporation China

Preface

This book has evolved from my experience over the past decade in teaching and doing research in functional analysis and certain of its applications. These applications are to optimization theory in general and to best approximation theory in particular. The geometric nature of the subjects has greatly influenced the approach to functional analysis presented herein, especially its basis on the unifying concept of convexity. Most of the major theorems either concern or depend on properties of convex sets; the others generally pertain to conjugate spaces or compactness properties, both of which topics are important for the proper setting and resolution of optimization problems. In consequence, and in contrast to most other treatments of functional analysis, there is no discussion of spectral theory, and only the most basic and general properties of linear operators are established.

Some of the theoretical highlights of the book are the Banach space theorems associated with the names of Dixmier, Krein, James, Smulian, Bishop-Phelps, Brondsted-Rockafellar, and Bessaga-Pelczynski. Prior to these (and others) we establish to two most important principles of geometric functional analysis: the extended Krein-Milman theorem and the Hahn-Banach principle, the latter appearing in ten different but equivalent formulations (some of which are optimality criteria for convex programs). In addition, a good deal of attention is paid to properties and characterizations of conjugate spaces, especially reflexive spaces. On the other hand, the following (incomplete) list provides a sample of the type of applications discussed:

Systems of linear equations and inequalities;
Existence and uniqueness of best approximations;
Simultaneous approximation and interpolation;
Lyapunov convexity theorem;
Bang-bang principle of control theory;
Solutions of convex programs;
Moment problems;
Error estimation in numerical analysis;
Splines;
Michael selection theorem;
Complementarity problems;
Variational inequalities;
Uniqueness of Hahn-Banach extensions.

Also, "geometric" proofs of the Borsuk-Dugundji extension theorem, the Stone-Weierstrass density theorem, the Dieudonne separation theorem, and the fixed point theorems of Schauder and Fan-Kakutani are given as further applications of the theory.

Over 200 problems appear at the ends of the various chapters. Some are intended to be of a rather routine nature, such as supplying the details to a deliberately sketchy or omitted argument in the text. Many others, however, constitute significant further results, converses, or counter-examples. The problems of this type are usually non-trivial and I have taken some pains to include substantial hints. (The design of such hints is an interesting exercise for an author: he hopes to keep the student on course without completely giving everything away in the process.) In any event, readers are strongly urged to at least peruse all the problems. Otherwise, I fear, a good deal of the total value of the book may be lost.

The presentation is intended to be accessible to students whose mathematical background includes basic courses in linear algebra, measure theory, and general topology. The requisite linear algebra is reviewed in §1, while the measure theory is needed mainly for examples. Thus the most essential background is the topological one, and it is freely assumed. Hence, with the exception of a few results concerning dispersed topological spaces (such as the Cantor-Bendixson lemma) needed in §25, no purely topological theorems are proved in this book. Such exclusions are warranted, I feel, because of the availability of many excellent texts on general topology. In particular, the union of the well-known books by J. Dugundji and J. Kelley contains all the necessary topological prerequisites (along with much additional material). Actually the present book can probably be read concurrently with courses in topology and measure theory, since Chapter I, which might be considered a brief second course on linear algebra with convexity, employs no topological concepts beyond standard properties of Euclidean spaces (the single exception to this assertion being the use of Ascoli's theorem in **7C**).

This book owes a great deal to numerous mathematicians who have produced over the last few years substantial simplifications of the proofs of virtually all the major results presented herein. Indeed, most of the proofs we give have now reached a stage of such conciseness and elegance that I consider their collective availability to be an important justification for a new book on functional analysis. But as has already been indicated, my primary intent has been to produce a source of functional analytic information for workers in the broad areas of modern optimization and approximation theory. However, it is also my hope that the book may serve the needs of students who intend to specialize in the very active and exciting ongoing research in Banach space theory.

I am grateful to Professor Paul Halmos for his invitation to contribute the book to this series, and for his interest and encouragement along the way to its completion. Also my thanks go to Professors Philip Smith and Joseph Ward for reading the manuscript and providing numerous corrections. As usual, Nancy Eberle and Judy Snider provided expert clerical assistance in the preparation of the manuscript.

Table of Contents

Chapter I

Convexity in Linear Spaces

Our purpose in this first chapter is to establish the basic terminology and properties of convex sets and functions, and of the associated geometry. All concepts are "primitive", in the sense that no topological notions are involved beyond the natural (Euclidean) topology of the scalar field. The latter will always be either the real number field $\mathbb{R}$, or the complex number field $\mathbb{C}$. The most important result is the "basic separation theorem", which asserts that under certain conditions two disjoint convex sets lie on opposite sides of a hyperplane. Such a result, providing both an analytic and a geometric description of a common underlying phenomenon, is absolutely indispensible for the further development of the subject. It depends implicitly on the axiom of choice which is invoked in the form of Zorn's lemma to prove the key lemma of Stone. Several other equally fundamental results (the "support theorem", the "subdifferentiability theorem", and two extension theorems) are established as equivalent formulations of the basic separation theorem. After indicating a few applications of these ideas we conclude the chapter with an introduction to the important notion of extremal sets (in particular extreme points) of convex sets.

§1. Linear Spaces

In this section we review briefly and without proofs some elementary results from linear algebra, with which the reader is assumed to be familiar. The main purpose is to establish some terminology and notation.

A. Let X be a linear space over the real or complex number field. The zero-vector in X is always denoted by θ. If $\{x_i\}$ is a subset of X, a *linear combination* of $\{x_i\}$ is a vector $x \in X$ expressible as $x = \Sigma\lambda_i x_i$, for certain scalars λ_i, only finitely many of which are non-zero. A subset of X is a (*linear*) *subspace* if it contains every possible linear combination of its members. The *linear hull* (*span*) of a subset S of X, consists of all linear combinations of its members, and thus span(S) is the smallest subspace of X that contains S. The subset S is *linearly independent* if no vector in S lies in the linear hull of the remaining vectors in S. Finally, the subset S is a (*Hamel*) *basis* for X if S is linearly independent and span(S) $= X$.

Lemma. *S is a basis for X if and only if S is a maximal linearly independent subset of S.*

Theorem. *Any non-trivial linear space has a basis; in fact, each non-empty linearly independent subset is contained in a basis.*

B. As the preceding theorem suggests, there is no unique choice of basis possible for a linear space. Nevertheless, all is not chaos: it is a remarkable fact that all bases for a given linear space contain the same number of elements.

Theorem. *Any two bases for a linear space have the same cardinality.*

It is thus consistent to define the (*Hamel*) *dimension* $\dim(X)$ of a linear space X as the cardinal number of an arbitrary basis for X. Let us now recall that if X and Y are linear spaces over the same field then a map $T: X \to Y$ is *linear* provided that

$$T(x + z) = T(x) + T(z), \qquad x, z \in X,$$
$$T(\alpha x) = \alpha T(x), \qquad x \in X, \qquad \alpha \text{ scalar}.$$

It follows that X and Y have the same dimension exactly when they are *isomorphic*, that is, when there exists a bijective linear map between X and Y.

C. We next review some constructions which yield new linear spaces from given ones. First, let $\{X_\alpha\}$ be a family of linear spaces over the same scalar field. Then the Cartesian product $\Pi_\alpha X_\alpha$ becomes a linear space (the *product* of the spaces X_α) if addition and scalar multiplication are defined component-wise. On the other hand, let $M_1, \ldots, M_n$ be subspaces of a linear space X and suppose they are *independent* in the sense that each is disjoint from the span of the others. Then their linear hull (in X) is called the *direct sum* of the subspaces $M_1, \ldots, M_n$ and written $M_1 \oplus \cdots \oplus M_n$ or simply $\bigoplus_{i=1}^{n} M_i$. The point of this definition is that if $M = \bigoplus_{i=1}^{n} M_i$, then each $x \in M$ can be uniquely expressed as $x = \sum_{i=1}^{n} m_i$, where $m_i \in M_i$, $i = 1, \ldots, n$.

Now let M be a subspace of X. For fixed $x \in X$, the subset $x + M \equiv \{x + y : y \in M\}$ is called an *affine subspace* (*flat*) *parallel to* M. Clearly, $x_1 + M = x_2 + M$ if and only if $x_1 - x_2 \in M$, so that the affine subspaces parallel to M are exactly the equivalence classes for the equivalence relation "$\sim_M$" defined by $x_1 \sim_M x_2$ if and only if $x_1 - x_2 \in M$. Now, if we define

$$(x + M) + (y + M) = (x + y) + M,$$
$$\alpha(x + M) = \alpha x + M, \qquad \alpha \text{ scalar}$$

then the collection of all affine subspaces parallel to M becomes a linear space X/M called the *quotient space* of X by M.

Theorem. *Let M be a subspace of the linear space X. Then there exist subspaces N such that $M \oplus N = X$, and any such subspace is isomorphic to the quotient space X/M.*

Any subspace N for which $M \oplus N = X$ is called a *complementary subspace* (*complement*) of M in X. Its dimension is by definition the *codimension* of M in X. The theorem also allows us to state that symbolically

$$\operatorname{codim}_X(M) = \dim(X/M),$$

where the subscript may be dropped provided the ambient linear space X is clearly specified. In fact, this theorem seems to suggest that there is not a great need for the construct X/M, and this is so in the purely algebraic case. However, later when we must deal with Banach spaces X and closed subspaces M, we shall see that generally there will be no closed complementary subspace. In this case the quotient space X/M becomes a Banach space and serves as a valuable substitute for the missing complement.

Now let M be a subspace of X, and choose a complementary subspace $N: M \oplus N = X$. Then we can define a linear map $P: X \to M$ by $P(m + n) = m$, $m \in M$, $n \in N$. P is called the *projection* of X on M (along N). We have similarly that $I - P$ is the projection of X on N (along M), where I is the identity map on X. The existence of such projections allows us the luxury of extending linear maps defined initially on a subspace of X: if $T: M \to Y$ is linear, then $\bar{T} \equiv T \circ P$ is a linear map from X to Y that agrees with T on M. Such a map $\bar{T}$ is an *extension* of T.

D. Let X be a linear space over the scalar field $\mathbb{F}$. The set of all linear maps $\phi: X \to \mathbb{F}$ becomes a new linear space X' with linear space operations defined by

$$(\phi + \psi)(x) \equiv \phi(X) + \psi(x),$$
$$(\alpha\phi)(x) \equiv \alpha\phi(x), \qquad \alpha \in \mathbb{F}, \qquad x \in X.$$

X' is called the *algebraic conjugate* (*dual*) *space* of X and its elements are called *linear functionals* on X. Observe that if $\dim(X) = n$ (a cardinal number) then X' is isomorphic to the product of n copies of the scalar field. As we shall see many times, it is often convenient to write

$$\phi(x) = \langle x, \phi \rangle,$$

for $x \in X$, $\phi \in X'$. The reason for this is that often the vector x and/or the linear functional ϕ may be given in a notation already containing parentheses or other complications.

Since X' is a linear space in a natural fashion, we can construct its algebraic conjugate space $(X')'$, which we write simply as X''. We call X'' the *second algebraic conjugate space* of X. We then have a map $J_X: X \to X''$ defined by

$$\langle \phi, J_X(x) \rangle = \langle x, \phi \rangle, \qquad x \in X, \qquad \phi \in X'.$$

This map is clearly linear; it is called the *canonical embedding* of X into X''. This terminology is justified by the next theorem.

Theorem. *The map J_X just defined is always injective, and is surjective exactly when* $\dim(X)$ *is finite.*

Thus, under the canonical embedding J_X, the linear space X is isomorphic to a subspace of its second algebraic dual space, and this subspace is proper (not all of X'') unless X is of finite dimension. In either case, we see that if it suits our purposes, we can consider that a given linear space consists of linear functionals acting on some other linear space (namely, X').

E. The proper affine subspaces of a linear space X can be partially ordered by inclusion. Any maximal element of this partially ordered set is a *hyperplane* in X.

Lemma. *An affine subspace V in X is a hyperplane if and only if there is a non-zero $\phi \in X'$ and a scalar α such that $V = \{x \in X : \phi(x) = \alpha\} \equiv [\phi; \alpha]$.*

Thus the hyperplanes in X correspond to the *level sets* of non-zero linear functionals on X. We can alternatively say that the hyperplanes in X consist of the elements of all possible quotient spaces $X/\ker(\phi)$, where $\phi \in X'$, $\phi \neq \theta$, and $\ker(\phi) \equiv [\phi; 0]$, the *kernel* (*null-space*) of ϕ. The hyperplanes in X which contain the zero-vector are in particular seen to coincide with the subspaces of codimension one. More generally, the subspaces of codimension n (n a positive integer) are exactly the kernels of linear maps on X of rank n (that is, with n-dimensional image).

F. Suppose that X is a complex linear space. Then in particular X is a real linear space if we admit only multiplication by real scalars. This underlying real vector space X_R is called the real restriction of X. Suppose that $\phi \in X'$. Then the maps

$$x \mapsto \operatorname{re} \phi(x),$$
$$x \mapsto \operatorname{im} \phi(x), \qquad x \in X,$$

are clearly linear functionals on X_R, that is, they belong to X'_R. On the other hand, since $\phi(ix) = i\phi(x)$, $x \in X$, we see that

$$\operatorname{im} \phi(x) = -\operatorname{re} \phi(ix)$$

so that ϕ is completely determined by its real part. Similarly, if we start with $\psi \in X'_R$, and define

$$\phi(x) = \psi(x) - i\psi(ix),$$

we find that $\phi \in X'$. To sum up, the correspondence $\psi \mapsto \phi$ just defined is an isomorphism between $X'_R \equiv (X_R)'$ and $(X')_R$.

This correspondence will be important in our later work with convex sets and functions. The separation, support, subdifferentiability, etc. results all concern various inequalities involving linear functionals; it is thus necessary that these linear functionals assume only real values. Consequently, in the sequel, linear spaces will often be assumed real. The preceding remarks then allow the results under discussion to be applied to complex linear spaces also, by passage to the real restriction, the associated linear functionals being simply the real parts of the complex linear functionals.

G. We give next a primitive version of the "quotient theorem", which allows us intuitively to "divide" one linear map by another. The more substantial result involving continuity questions appears in Chapter III.

Let X, Y, Z be linear spaces and let $S: X \to Y$, $T: X \to Z$ be linear maps. We ask whether there exists a linear map $R: Y \to Z$ such that $T = R \circ S$. An obvious necessary condition for this to occur is that $\ker(S) \subset \ker(T)$; it is more useful to note that this condition is also sufficient.

Theorem. *Let the linear maps S and T be prescribed as above, and assume that* $\ker(S) \subset \ker(T)$. *Then there exists a linear map R, uniquely specified on* range(S), *such that* $T = R \circ S$.

One consequence of this theorem, important for later work on weak topologies, is the following.

Corollary. *Let X be a linear space and let* $\phi_1, \ldots, \phi_n, \psi \in X'$. *Then* $\psi \in \operatorname{span}\{\phi_1, \ldots, \phi_n\}$ *if and only if*

$$\bigcap_{i=1}^{n} \ker(\phi_i) \subset \ker(\psi).$$

H. Let M be a subspace of the linear space X. The *annihilator* M° of M consists of those linear functionals in X' that vanish at each point of M. It is clearly a subspace of X'. Similarly, if N is a subspace of X', its *pre-annihilator* ${}^\circ N$ consists of all vectors in X at which every functional in N vanishes. Thus:

$$M^\circ = \bigcap_{x \in M} \ker(J_X(x)),$$

$${}^\circ N = J_X^{-1}(\operatorname{range}(J_X) \cap N^\circ).$$

Let $T: X \to Y$ be a linear map. The *transpose* T' is the linear map from Y' to X' defined by

$$\langle x, T'(\psi) \rangle = \langle T(x), \psi \rangle, \qquad x \in X, \qquad \psi \in Y'.$$

It may be recalled that when X and Y are (real) finite dimensional Euclidean spaces, and T is represented by a matrix (with respect to the standard unit vector bases in X and Y), then T' is represented by the transposed matrix, whence the above terminology.

Lemma. *Let* $T: X \to Y$ *be a linear map. Then* $\ker(T') = \operatorname{range}(T)^\circ$ *and* $\operatorname{range}(T') = \ker(T)^\circ$.

Thus we see that T is surjective (resp., injective) if and only if T' is injective (resp., surjective). The various constructs in the preceding sub-sections can now all be tied together in the following way. Let us say that the linear spaces X and Y are *canonically isomorphic*, written $X \cong Y$, if an isomorphism between them can be constructed without the use of bases in either space. For example, we clearly have $X \cong J_X(X)$. On the other hand, it may be recalled that none of the usual isomorphisms between a finite dimensional space and its algebraic conjugate space is canonical.

Theorem. *Let M be a subspace of the linear space X. Then*

a) $M^\circ \cong (X/M)'$;

b) $M' \cong X'/M^\circ$.

The proof of a) follows from an application of the lemma to the *quotient map* $Q_M: X \to X/M$, defined by $Q_M(x) \equiv x + M$. Since Q_M is clearly surjective, its transpose $Q'_M: (X/M)' \to X'$ is an isomorphism onto its range, which is $(\ker(Q_M))^\circ = M^\circ$. The proof of b) proceeds similarly by applying the lemma to the identity injection of M into X.

§2. Convex Sets

In this section we establish the most basic properties of convex sets in linear spaces, and prove the crucial lemma of Stone. This lemma is, in effect, the cornerstone of our entire subject, as we shall see shortly. Throughout this section, X is an arbitrary linear space.

A. Let $x, y \in X$ with $x \neq y$. The *line segment* joining x and y is the set $[x, y] = \{\alpha x + (1 - \alpha)y : 0 \leqslant \alpha \leqslant 1\}$. Similarly we put $[x, y) = [x, y]\backslash\{y\}$, and $(x, y) = [x, y)\backslash\{x\}$. If $A \subset X$, then A is *star-shaped* with respect to $p \in A$ if $[p, x] \subset A$, for all $x \in A$, and A is *convex* if it is star-shaped with respect to each of its elements. Clearly a translate of a convex set is convex, hence each affine subspace of X is convex.

Since the intersection of a family of convex sets is again convex, we can define, for any $A \subset X$, the *convex hull* of A, written co(A), to be the intersection of all convex sets in X that contain S. Thus co(A) is the smallest convex set in X that contains A. This set admits an alternative description, namely

$$\mathrm{co}(A) = \{\Sigma\alpha_i x_i : 0 \leqslant \alpha_i \leqslant 1, \Sigma\alpha_i = 1, x_i \in A\},$$

the set of all *convex combinations* of points in A. (We emphasize again that all linear combinations of vectors involve only finitely many non-zero terms.) We have, for instance, that $\mathrm{co}(\{x, y\}) = [x, y]$. More generally, if we define the *join* of two sets A and B in X to be $\cup \{[x, y] : x \in A, y \in B\}$, then

(2.1) $$\mathrm{co}(A \cup B) = \mathrm{join}(\mathrm{co}(A), \mathrm{co}(B)),$$

so that if A and B are convex, then their join is convex and is, in fact, the convex hull of their union.

Let us define addition and scalar multiplication on the family $P(X)$ of non-empty subsets of X by

$$\alpha A + \beta B \equiv \{\alpha a + \beta b : a \in A, b \in B\},$$

where $A, B \subset X$ and α, β are scalars. This definition does not define a linear space structure on $P(X)$; nevertheless, it proves to be quite convenient. For instance, we can state

(2.2) $$\mathrm{co}(\alpha A + \beta B) = \alpha\,\mathrm{co}(A) + \beta\,\mathrm{co}(B).$$

A set $A \subset X$ is *balanced* (*equilibrated*) if $\alpha A \subset A$ whenever $|\alpha| \leqslant 1$. The *balanced hull* of A, bal(A), is the intersection of all balanced subsets of X that contain A, and is therefore the smallest balanced set in X that contains A. Alternatively:

$$\mathrm{bal}(A) = \cup\{\alpha A : |\alpha| \leqslant 1\}.$$

Finally, a set which is both convex and balanced is called *absolutely convex*. The smallest such set containing a given set A is the *absolute convex* *hull* of A, written $\text{aco}(A)$. For example, $\text{aco}(\{x\}) = [-x, x]$, if X is a real linear space. In general, we have

$$\text{aco}(A) = \text{co}(\text{bal}(A))$$
$$= \{\Sigma\alpha_i x_i : \Sigma|\alpha_i| \leqslant 1, x_i \in A\},$$

the set of all *absolute convex combinations* of points in A. In particular, we see that A is absolutely convex if and only if $a, b \in A$ and $|\alpha| + |\beta| \leqslant 1$ implies $\alpha a + \beta b \in A$.

B. We come now to the celebrated result of Stone. Two non-empty convex sets C and D in X are *complementary* if they form a partition of X, that is, $C \cap D = \varnothing$, $C \cup D = X$. An evident example of a pair of complementary convex sets occurs when X is real: choose a non-zero $\phi \in X'$ and put $C = \{x \in X : \phi(x) \geqslant 0\}$, $D = X \backslash C$.

Lemma. *Let A and B be disjoint convex subsets of X. Then there exist complementary convex sets C and D in X such that $A \subset C$, $B \subset D$.*

Proof. Let $\mathscr{C}$ be the class of all convex sets in X disjoint from B and containing A; certainly $A \in \mathscr{C}$. After partially ordering $\mathscr{C}$ by inclusion, we apply Zorn's lemma to obtain a maximal element $C \in \mathscr{C}$. It now suffices to put $D \equiv X \backslash C$ and prove that D is convex. If D were not convex, there would be $x, z \in D$ and $y \in (x, z) \cap C$. Because C is a maximal element of $\mathscr{C}$, there must be points $p, q \in C$ such that both (p, x) and (q, z) intersect B, say at points u, v, resp. (Reason by contradiction; if the last statement were false, then the following assertion (*) would hold: for all pairs $\{p, q\} \subset C$, either $(p, x) \cap B = \varnothing$ or $(q, z) \cap B = \varnothing$. Now if $(q, z) \cap B = \varnothing$, for all $q \in C$, then $C \subset \text{co}(\{z, C\})$ and C is not maximal. Consequently, there is some $\bar{q} \in C$ for which $(\bar{q}, z) \cap B \neq \varnothing$. But then, if there were a point $\bar{p} \in C$ such that $(\bar{p}, x) \cap B \neq \varnothing$, the pair $\{\bar{p}, \bar{q}\}$ would violate (*). Thus, for all $p \in C$, $(p, x) \cap B \neq \varnothing$, $C \subset \text{co}(\{x, C\})$, and C is not maximal.) Now, however, we find that $[u, v] \cap \text{co}(\{p, q, y\}) \neq \varnothing$, which contradicts the disjointness of B and C. □

C. Let A and B be subsets of X. The *core* of A *relative* to B, written $\text{cor}_B(A)$, consists of all points $a \in A$ such that for each $b \in B \backslash \{a\}$ there exists $x \in (a, b)$ for which $[a, x] \subset A$. Intuitively, it is possible to move from each $a \in \text{cor}_B(A)$ towards any point of B while staying in A. The core of A relative to X is called simply the *core* (*algebraic interior*) of A and written $\text{cor}(A)$. Sets $A \subset X$ for which $A = \text{cor}(A)$ are called *algebraically open*, while points neither in $\text{cor}(A)$ nor in $\text{cor}(X \backslash A)$ are called *bounding points* of A; they constitute the *algebraic boundary* of A. It is easy to see that the core of any (absolutely) convex set is again (absolutely) convex.

A second important instance of the relative core concept occurs when B is the smallest affine subspace that contains A. This subspace, $\text{aff}(A)$ (the *affine hull* of A), can be described as $\{\Sigma\alpha_i x_i : \Sigma\alpha_i = 1, x_i \in A\}$ or, equivalently, as $x + \text{span}(A - A)$, for any fixed $x \in A$. Now the set $\text{cor}_{\text{aff}(A)}(A)$ is called the *intrinsic core* of A and written $\text{icr}(A)$. In particular, when A is convex, $a \in \text{icr}(A)$ if and only if for each $x \in A \setminus \{a\}$, there exists $y \in A$ such that $a \in (x, y)$; intuitively, given $a \in \text{icr}(A)$, it is possible to move linearly from any point in A past a and remain in A.

In general, $\text{icr}(A)$ will be empty; but in a variety of special cases we can show $\text{icr}(A)$ and even $\text{cor}(A)$ are not empty. For example, it should be clear that if X is a finite dimensional Euclidean space and $A \subset X$ is convex, then $\text{cor}(A)$ is just the topological interior of A. But this last assertion fails in the infinite dimensional case as we shall see later, after introducing the necessary topological notions. We now work towards a sufficient condition for a convex set to have non-empty intrinsic core.

A finite set $\{x_0, x_1, \ldots, x_n\} \subset X$ is *affinely independent* (*in general position*) if the set $\{x_1 - x_0, \ldots, x_n - x_0\}$ is linearly independent. The convex hull of such a set is called an *n-simplex* with *vertices* $x_0, x_1, \ldots, x_n$. In this case, each point in the n-simplex can be uniquely expressed as a convex combination of the vertices; the coefficients in this convex combination are the *barycentric coordinates* of the point.

Lemma. *Let A be an n-simplex in X. Then* $\text{icr}(A)$ *consists of all points in A each of whose barycentric coordinates is positive. In particular,* $\text{icr}(A) \neq \emptyset$.

Proof. Let the vertices of A be $\{x_0, x_1, \ldots, x_n\}$. Let $a = \Sigma \alpha_i x_i$ and $b = \Sigma \beta_i x_i$ be points of A with all $\alpha_i > 0$. To show $a \in \text{icr}(A)$, it is sufficient to show that $b + \lambda(a - b) \in A$ for some $\lambda > 1$. If we put $\lambda = 1 + \varepsilon$, the condition on ε becomes

$$\alpha_i + \varepsilon(\alpha_i - \beta_i) \geqslant 0, \qquad i = 0, 1, \ldots, n,$$

$$\sum_{i=0}^{n} \alpha_i + \varepsilon(\alpha_i - \beta_i) = 1.$$

Since $\sum_{i=0}^{n} (\alpha_i - \beta_i) = 1 - 1 = 0$, the second condition always holds, and since all $\alpha_i > 0$, the first condition holds for all sufficiently small positive ε. Conversely, let $a = \Sigma \alpha_i x_i$ have a zero coefficient, say $\alpha_k = 0$. Then we claim that $x_k + \lambda(a - x_k) \notin A$, for any $\lambda > 1$. For otherwise, for some $\lambda \geqslant 1$ we would have

$$x_k + \lambda(a - x_k) = \sum_{i=0}^{n} \beta_i x_i \in A.$$

It would follow that

$$a = \frac{\beta_k + \lambda - 1}{\lambda} x_k + \sum_{i \neq k} \gamma_i x_i,$$

for certain coefficients γ_i. But in this representation of a, the x_k-coefficient is clearly positive (since $\beta_k \geqslant 0$). This leads us to a contradiction, since the barycentric coordinates of a are uniquely determined, and the x_k-coefficient of a was assumed to vanish. □

The *dimension* of an affine subspace $x + M$ of X is by definition the dimension of the subspace M. The *dimension* of an arbitrary convex set A in X is the dimension of $\text{aff}(A)$. A nice way of writing this definition symbolically is

$$\dim(A) \equiv \dim(\text{span}(A - A)).$$

It follows from the preceding lemma that every non-empty finite dimensional convex set A has a non-empty intrinsic core. Indeed, if $\dim(A) = n$ (finite), then A must contain an affinely independent set $\{x_0, x_1, \ldots, x_n\}$ and hence the n-simplex $\text{co}(\{x_0, x_1, \ldots, x_n\})$.

Theorem. *Let A be a convex subset of the finite dimensional linear space X. Then* $\text{cor}(A) \neq \varnothing$ *if and only if* $\text{aff}(A) = X$.

Proof. If $\text{aff}(A) = X$, the last remark shows that $\text{cor}(A) = \text{icr}(A) \neq \varnothing$. Conversely, if $p \in \text{cor}(A)$, and $x \in X$, there is some positive ε for which $[p, p + \varepsilon(x - p)] \subset A$. Then with $\lambda \equiv (\varepsilon - 1)/\varepsilon$, we have

$$x = \lambda p + (1 - \lambda)(p + \varepsilon(x - p)) \in \text{aff}(A). \qquad \square$$

Remark. The conclusion of this theorem fails in any infinite dimensional space. More precisely, in any such space X we can find a convex set A with empty core such that $\text{aff}(A) = X$. To do this we simply let A consist of all vectors in X whose coordinates wrt some given basis for X are non-negative. Clearly $A - A = X$, while $\text{cor}(A) = \varnothing$.

D. Let $A \subset X$. A point $x \in X$ is *linearly accessible* from A if there exists $a \in A$, $a \neq x$, such that $(a, x) \subset A$. We write $\text{lina}(A)$ for the set of all such x, and put $\text{lin}(A) = A \cup \text{lina}(A)$. For example, when A is the open unit disc in the Euclidean plane, and B is its boundary the unit circle, we have that $\text{lina}(B) = \varnothing$ while $\text{lin}(A) = \text{lina}(A) = A \cup B$. In general, one suspects (correctly) that when X is a finite dimensional Euclidean space, and $A \subset X$ is convex then $\text{lin}(A)$ is the topological closure of A. But we have to go a bit further to be able to prove this.

The "lin" operation can be used to characterize finite dimensional spaces. We give one such result next and another in the exercises. Let us say that a subset of A of X is *ubiquitous* if $\text{lin}(A) = X$.

Theorem. *The linear space X is infinite dimensional if and only if X contains a proper convex ubiquitous subset.*

Proof. Assume first that X is finite dimensional, and let A be a convex ubiquitous set in X. Now clearly A cannot belong to any proper affine subspace of X. Hence $\text{aff}(A) = X$ and thus, by **2C**, $\text{cor}(A)$ is non-empty. Without loss of generality, we can suppose that $\theta \in \text{cor}(A)$. Now, given any $x \in X$, there is some $y \in X$ such that $[y, 2x) \subset A$, and there is a positive number t such that $t(2x - y) \in A$. It is easy to see that the half-line $\{\lambda x + (1 - \lambda)t(2x - y) : \lambda \geqslant 0\}$ will intersect the segment $[y, 2x)$; but this of course means that x is a convex combination of two points in A, hence $x \in A$ also.

Conversely, assume that X is infinite dimensional. We can select a well-ordered basis for X (since any set can be well-ordered, according to Zermelo's theorem). Now we define A to be the set of all vectors in X whose last coordinate (wrt this basis) is positive. A is evidently a proper convex subset of X, and we claim that it is ubiquitous. Indeed, given any $x \in X$, we can choose a basis vector y "beyond" any of the finitely many basis vectors used to represent x. But then, if $t > 0$, we have $x + ty \in A$; in particular, $x \in \text{lina}(A)$. □

E. We give one further result involving the notions of core and "lina" which will be needed shortly to establish the basic separation theorem of **4B**. It is convenient to first isolate a special case as a lemma.

Lemma. *Let A be a convex subset of the linear space X, and let $p \in \text{cor}(A)$. For any $x \in A$, we have $[p, x) \subset \text{cor}(A)$, and hence*

$$\text{cor}(A) = \cup\{[p, x): x \in A\}.$$

Proof. Choose any $y \in [p, x)$, say $y = tx + (1 - t)p$, where $0 < t < 1$. Then given any $z \in X$, there is some $\lambda > 0$ so that $p + \lambda z \in A$. Hence $y + (1 - t)\lambda z = (1 - t)(p + \lambda z) + tx \in A$, proving that $y \in \text{cor}(A)$. Finally, given any $q \in \text{cor}(A)$, $q \neq p$, there exists some $\delta > 0$ such that $x \equiv q + \delta(q - p) \in A$. It follows that $q = (\delta p + x)/(1 + \delta) \in [p, x)$. □

Theorem. *Let A be a convex subset of the linear space X, and $p \in \text{cor}(A)$. Then for any $x \in \text{lina}(A)$ we have $[p, x) \subset \text{cor}(A)$.*

Proof. We can assume that $p = \theta$. Since $x \in \text{lina}(A)$, there is some $z \in A$ such that $[z, x) \subset A$, and since $\theta \in \text{cor}(A)$, there is some $\delta > 0$ such that $-\delta z \in A$. Arguing as in **2D**, given any point tx, $0 < t < 1$, the line $\{\lambda tx + (1 - \lambda)(-\delta z): \lambda \geqslant 0\}$ will intersect the segment $[z, x)$ if δ is taken sufficiently small. Consequently, the segment $[\theta, x)$ lies in A. But now the preceding lemma allows us to conclude that in fact $[\theta, x)$ lies in $\text{cor}(A)$. □

§3. Convex Functions

In this section we introduce the notion of convex function and its most important special case, the "sublinear" function. With such functions we can associate in a natural fashion certain convex sets. The geometric analysis of such sets developed in subsequent sections makes possible many non-trivial conclusions about the given functions.

A. Intuitively, a real-valued function defined on an interval is convex if its graph never "dents inward" or, more precisely, if the chord joining any two points on the graph always lies on or above the graph. In general, we say that if A is a convex set in a linear space X then a real-valued function f defined on A is *convex* on A if the subset of $X \times \mathbb{R}^1$ defined as $\{(x, t): x \in A, f(x) \leqslant t\}$ is convex. This set is called the *epigraph* of f, written $\text{epi}(f)$.