CONVEXITY

§-1. The course presents the fundamentals of convex analysis (in finite dimensional spaces) under four headings:

1. Convex Sets
2. Separation Theorems
3. Extreme points and the representation of convex sets
4. Convex Functions.

The material is basic to the mathematical theory of optimization. Its importance can be appreciated fully only through a consideration of the many applications: Convex programming, of which linear programming is a special case; Control Theory etc. Unfortunately we can do little more than hint at these applications. Some are taken up in other courses, for example: Linear programming; Calculus of variations and optimization theory; Approximation theory.

There is no single book suitable as a text for the course. The following books are, however, valuable reference sources.

Valentine, Frederick, "Convex Sets", Krieger, 1976.

Rockafellar, R. Tyrell, "Convex Analysis", Princeton University Press, 1972.

Holmes, Richard B., "Geometric Functional Analysis", (Chapter I only), Springer, 1975.

Eggleston, H.G., "Convexity", Cambridge University Press, 1969.

Bazaraa, M.S. and Shetty, C.M., "Foundations of Optimization", Springer, 1976.

§0. Preliminaries (which may be assumed throughout the course)

Everything will take place in, or on, *n-dimensional Euclidean space* (for some n)*. That is, the space E^n of ordered n-tupples of real numbers (typical element $\underset{\sim}{x} = (x_1, x_2, \ldots, x_n)$) regarded as a linear (or vector) space over the scalar field of real numbers $\mathcal{R}$ with addition and scalar multiplication defined component-wise:

$$\underset{\sim}{x} + \underset{\sim}{y} = (x_1, \ldots, x_n) + (y_1, \ldots, y_n) = (x_1+y_1, \ldots, x_n+y_n)$$

and
$$\lambda\underset{\sim}{x} = \lambda(x_1, \ldots, x_n) = (\lambda x_1, \ldots, \lambda x_n);$$

and equipped with the natural inner-product

$$(\underset{\sim}{x}, \underset{\sim}{y}) = \underset{\sim}{x}.\underset{\sim}{y} = \sum_{i=1}^{n} x_i y_i .$$

We say $\underset{\sim}{x}$ is orthogonal to $\underset{\sim}{y}$ and write $\underset{\sim}{x} \perp \underset{\sim}{y}$ if $(\underset{\sim}{x}, \underset{\sim}{y}) = 0$.
While no familiarity with either metric spaces or Approximation Theory is assumed the following notions are basic.

The "length" of $\underset{\sim}{x} \in E^n$ is $\|x\| = \sqrt{(\underset{\sim}{x}, \underset{\sim}{x})} = \sqrt{\sum_{i=1}^{n} x_i^2}$.

EXERCISE: Show that, for $\underset{\sim}{x}, \underset{\sim}{y}, \underset{\sim}{z} \in E^n$ and $\lambda \in \mathcal{R}$

(i) $(\underset{\sim}{x}, \underset{\sim}{y}) = (\underset{\sim}{y}, \underset{\sim}{x})$

(ii) $(\lambda\underset{\sim}{x}, \underset{\sim}{y}) = \lambda(\underset{\sim}{x}, \underset{\sim}{y})$

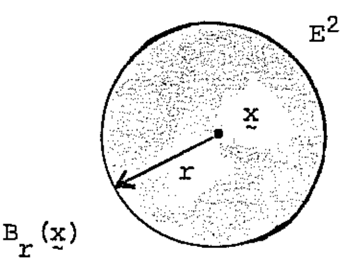
(iii) $(\underset{\sim}{x}+\underset{\sim}{y}, \underset{\sim}{z}) = (\underset{\sim}{x}, \underset{\sim}{z}) + (\underset{\sim}{y}, \underset{\sim}{z})$.

Also show that, $\|\lambda\underset{\sim}{x}\| = |\lambda|\|\underset{\sim}{x}\|$ and $\|\underset{\sim}{x}\| \geqslant 0$ with $\|\underset{\sim}{x}\| = 0$ if and only if $\underset{\sim}{x} = 0$.

The "distance between any two vectors" $\underset{\sim}{x}$ and $\underset{\sim}{y} \in E^n$ is

$$d(\underset{\sim}{x}, \underset{\sim}{y}) = \|\underset{\sim}{x}-\underset{\sim}{y}\| = \sqrt{\sum_{i=1}^{n} (x_i - y_i)^2}.$$

By the ball of radius r centre $\underset{\sim}{x}$ in E^n we mean
$B_r(\underset{\sim}{x}) = \{\underset{\sim}{y} \in E^n : \|\underset{\sim}{x}-\underset{\sim}{y}\| < r\}$.



* This leads to no loss of generality. Every n-dimensional normed linear space is "topologically isomorphic" to n-dimensional Euclidean space (see the proof of Theorem 2 of Approximation Theory, pp.13-14), so all our results apply in any finite dimensional space. Some remain true in infinite dimensional spaces.

$\underset{\sim}{x}$ is an <u>interior point</u> of $A \subseteq E^n$ if there exists $r > 0$ such that $B_r(\underset{\sim}{x}) \subseteq A$.

A <u>sequence</u> $(\underset{\sim}{x}_n) \equiv \underset{\sim}{x}_1, \underset{\sim}{x}_2, \ldots, \underset{\sim}{x}_n, \ldots$ of vectors in E^n is said to <u>converge</u> to $\underset{\sim}{x} \in E^n$ if $\|\underset{\sim}{x}_n - \underset{\sim}{x}\| \to 0$ as $n \to \infty$.

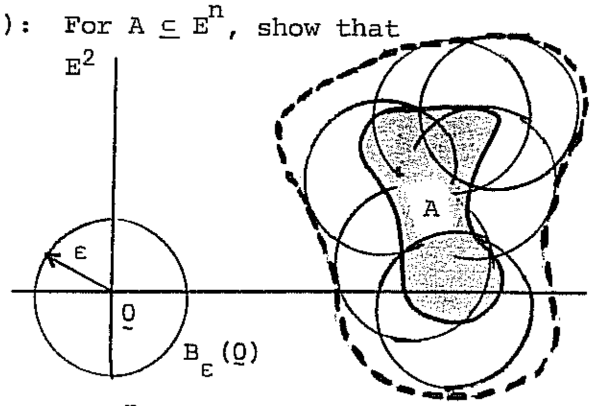
A subset A of E^n is <u>closed</u> if whenever $(\underset{\sim}{a}_n)$ is a sequence of elements of A which converges to $\underset{\sim}{x} \in E^n$ we have $\underset{\sim}{x} \in A$.

For $A \subseteq E^n$ the smallest closed set containing A is known as the <u>closure</u> of A and is denoted by $\overline{A}$ (not to be confused with the complement of A which will be written as $E^n \backslash A$).

Clearly A is closed if and only if $A = \overline{A}$.

EXERCISE (optional at this stage): For $A \subseteq E^n$, show that

$$\overline{A} = \bigcap_{\varepsilon>0} [A + B_\varepsilon(\underset{\sim}{0})]$$

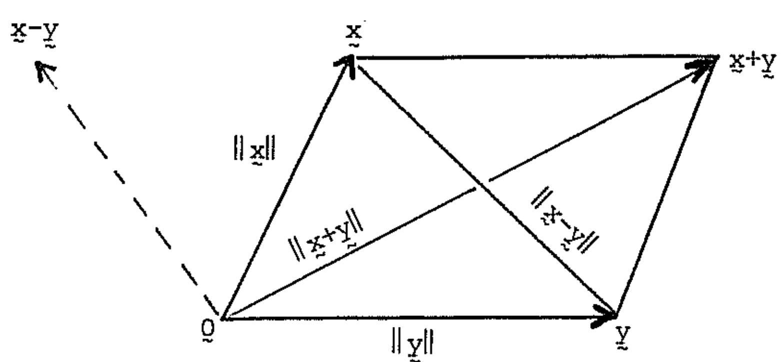


[Note: Here and elsewhere, for $A, B \subseteq E^n$

$A + B = \{\underset{\sim}{x} \in E^n : \underset{\sim}{x} = \underset{\sim}{a} + \underset{\sim}{b}$ for some $\underset{\sim}{a} \in A$ and some $\underset{\sim}{b} \in B\}$.]

The "norm" in E^n satisfies the following parallelogram law

$$\|\underset{\sim}{x} - \underset{\sim}{y}\|^2 + \|\underset{\sim}{x} + \underset{\sim}{y}\|^2 = 2\|\underset{\sim}{x}\|^2 + 2\|\underset{\sim}{y}\|^2.$$



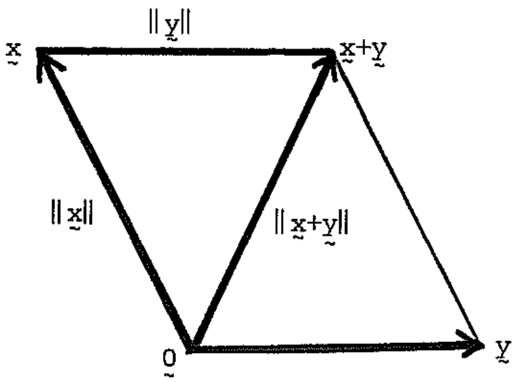
Proof. $\|\underset{\sim}{x} - \underset{\sim}{y}\|^2 = (\underset{\sim}{x} - \underset{\sim}{y}, \underset{\sim}{x} - \underset{\sim}{y})$

$= (\underset{\sim}{x}, \underset{\sim}{x}) - (\underset{\sim}{x}, \underset{\sim}{y}) - (\underset{\sim}{y}, \underset{\sim}{x}) + (\underset{\sim}{y}, \underset{\sim}{y})$

$= \|\underset{\sim}{x}\|^2 - 2(\underset{\sim}{x}, \underset{\sim}{y}) + \|\underset{\sim}{y}\|^2.$

Similarly $\|\underset{\sim}{x}+\underset{\sim}{y}\|^2 = \|\underset{\sim}{x}\|^2 + 2(\underset{\sim}{x},\underset{\sim}{y}) + \|\underset{\sim}{y}\|^2$.
Adding these two identities yields the result. □

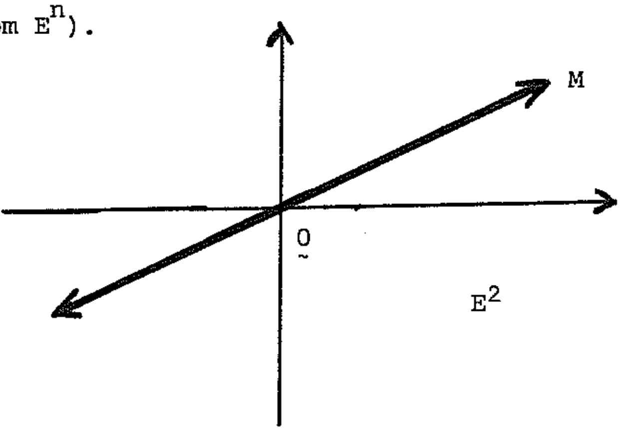
We will also have occasion to use the triangle inequality:
For $\underset{\sim}{x}, \underset{\sim}{y} \in E^n$, $\|\underset{\sim}{x} + \underset{\sim}{y}\| \leqslant \|\underset{\sim}{x}\| + \|\underset{\sim}{y}\|$.



(See either, Analysis in Metric spaces, p.10, or Approximation Theory, p.25, for a proof.)

Most of elementary linear algebra is important for our analysis. (Use your second year notes as a reference whenever necessary.) The following are of most immediate use.

$M \subseteq E^n$ is a <u>subspace</u> of E^n (written $M \leqslant E^n$) if and only if $\lambda\underset{\sim}{a} + \underset{\sim}{b} \in M$ whenever $\underset{\sim}{a}, \underset{\sim}{b} \in M$ and $\lambda \in \mathbb{R}$ (that is, whenever M is a linear space in its own right with the operations of vector addition and scalar multiplication inherited from E^n).



A sometimes useful result is that *every subspace* M *of* E^n *is closed*. (You might like to try and prove this.) It is also therefore <u>complete</u>, since E^n is itself complete. That is, if $(\underset{\sim}{m}_n) \subset M$ is a <u>Cauchy sequence</u>:
$\|\underset{\sim}{m}_n - \underset{\sim}{m}_k\| \to 0$ as $n,k \to \infty$; then $(\underset{\sim}{m}_n)$ converges to some element of M.
(For the completeness of E^n see either, Analysis in Metric Spaces, p.18, or Approximation Theory, Corollary 2 on p.14.)

The <u>span</u> of $A \subseteq E^n$, is

$$\langle A\rangle = \{\underset{\sim}{x} \in E^n : \underset{\sim}{x} = \sum_{i=1}^{m} \lambda_i \underset{\sim}{a}_i \text{ for some } m \in N,\ \lambda_1,\lambda_2,\ldots,\lambda_m \in \mathbb{R}$$

and $\underset{\sim}{a}_1, \underset{\sim}{a}_2, \ldots, \underset{\sim}{a}_m \in A\}$ the set of all finite linear combinations of elements of A.

$\langle A\rangle$ is the smallest subspace of E^n containing A, that is, *if* $M \leqslant E^n$ *and* $A \subseteq M$ *then* $\langle A\rangle \leqslant M$.

For latter purposes it will be convenient to re-christen the span of A as the <u>linear hull</u> of A.

If $M \leqslant E^n$ is the linear hull (span) of $B = \{\underset{\sim}{b}_1, \underset{\sim}{b}_2, \ldots, \underset{\sim}{b}_m\}$ and B is a linearly independent set of vectors, then B is a <u>basis</u> for M. Further the dimension of M, dim M = m the number of vectors in B.

If $B = \{\underset{\sim}{b}_1, \underset{\sim}{b}_2, \ldots, \underset{\sim}{b}_m\}$ is a basis for $M \leqslant E^n$, then it is possible to extend B to a basis for E^n. That is, there exists n-m vectors $\underset{\sim}{b}_{m+1}, \underset{\sim}{b}_{m+2}, \ldots, \underset{\sim}{b}_n$ such that $\{\underset{\sim}{b}_1, \underset{\sim}{b}_2, \ldots, \underset{\sim}{b}_m, \underset{\sim}{b}_{m+1}, \ldots, \underset{\sim}{b}_n\}$ is a basis for E^n.

Using the Gram-Schmidt orthogonalization procedure (refer to approximation theory notes, p.27 and exercise 32, or to linear algebra), if $B = \{\underset{\sim}{b}_1, \underset{\sim}{b}_2, \ldots, \underset{\sim}{b}_n\}$ is a basis for E^n we can construct a basis $D = \{\underset{\sim}{d}_1, \underset{\sim}{d}_2, \ldots, \underset{\sim}{d}_n\}$ such that D is <u>orthonormal</u>, that is

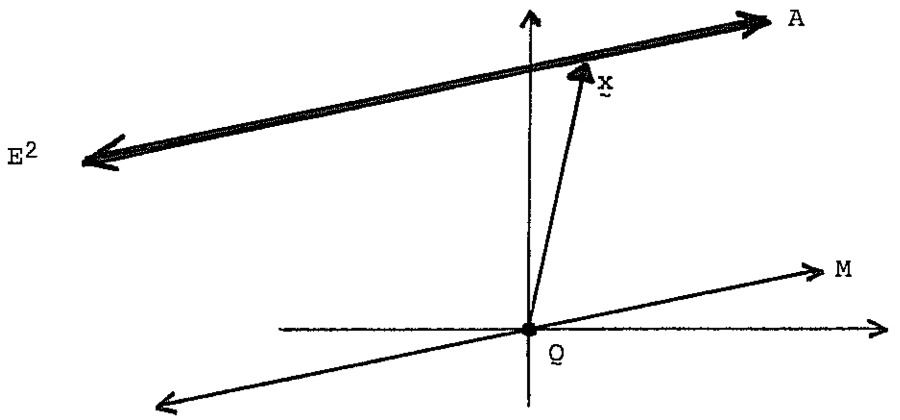
$$(\underset{\sim}{d}_i, \underset{\sim}{d}_j) = \begin{cases} 1 \text{ if } i = j \\ 0 \text{ if } i \neq j \end{cases}$$, and, for any $k \in \{1,2,\ldots,n\}$,

$$\langle b_1, b_2, \ldots, \underset{\sim}{b}_k\rangle = \langle \underset{\sim}{d}_1, \underset{\sim}{d}_2, \ldots, \underset{\sim}{d}_k\rangle.$$

Chapter 1 - Convex Sets

§1. *Affine Sets* (otherwise known as: "Affine Manifolds", "Linear Varieties" or "Flats")

DEFINITION: An affine subset of E^n is a translate of a subspace of E^n. That is, A is an affine subset of E^n if and only if $A = \underset{\sim}{x} + M$ for some $\underset{\sim}{x} \in E^n$ and subspace $M \subseteq E^n$. We say M is a subspace parallel to A.



Theorem 1: *If* A *is an affine set, then there exists a* *unique* *subspace of* E^n *parallel to* A.

Proof. Since A is affine $A = \underset{\sim}{x} + M$ for some $\underset{\sim}{x} \in A$ and $M \leqslant E^n$.

Now, assume $A = \underset{\sim}{x}' + M'$, then

$$\underset{\sim}{x} + M = \underset{\sim}{x}' + M'$$

or

$$M = (\underset{\sim}{x}' - \underset{\sim}{x}) + M'.$$

In particular, given any $\underset{\sim}{m} \in M$ there exists $\underset{\sim}{m}' \in M'$ such that $\underset{\sim}{m} = (\underset{\sim}{x}' - \underset{\sim}{x}) + \underset{\sim}{m}'$. Since $\underset{\sim}{0} \in M$, there exists $\underset{\sim}{m}' \in M'$ such that $\underset{\sim}{0} = (\underset{\sim}{x}' - \underset{\sim}{x}) + \underset{\sim}{m}'$. Thus $(\underset{\sim}{x}' - \underset{\sim}{x}) = -\underset{\sim}{m}' \in M'$ (as M' is a subspace) and so $M = (\underset{\sim}{x}' - \underset{\sim}{x}) + M' \subseteq M'$. Similarly $M' \subseteq M$, and $M = M'$. □

EXERCISES:
(1) Describe the affine subsets of E^2 and E^3.
(2) Let A be an affine subset of E^n and let $\underset{\sim}{a}$ be any element of A. Show that the subspace M parallel to A is given by $M = A - \underset{\sim}{a}$.
(3) Let $A = \underset{\sim}{x} + M$ be an affine subset of E^n where M is the subspace parallel to A. Show that $A = \underset{\sim}{x}' + M$ if and only if $\underset{\sim}{x}' = \underset{\sim}{x} + \underset{\sim}{m}$ for some element $\underset{\sim}{m}$ of M.

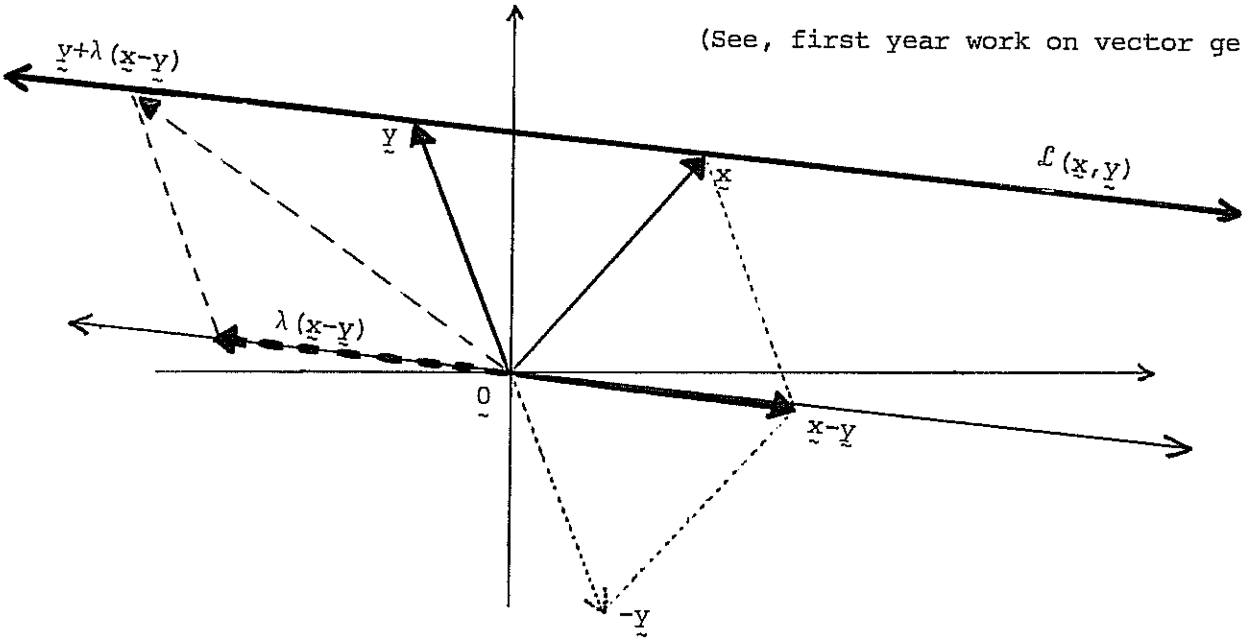
Structure of Affine Sets.

If $\underset{\sim}{x}, \underset{\sim}{y} \in E^n$ we denote by $\mathcal{L}(\underset{\sim}{x},\underset{\sim}{y})$ the line through $\underset{\sim}{x}$ and $\underset{\sim}{y}$.

Thus,

$$\mathcal{L}(\underset{\sim}{x},\underset{\sim}{y}) = \{\underset{\sim}{z} \in E^n : \underset{\sim}{z} = \underset{\sim}{y} + \lambda(\underset{\sim}{x} - \underset{\sim}{y}),\ \lambda \in \mathcal{R}\}$$
$$= \{\underset{\sim}{z} \in E^n : \underset{\sim}{z} = \lambda\underset{\sim}{x} + (1-\lambda)\underset{\sim}{y},\ \lambda \in \mathcal{R}\}.$$

(See, first year work on vector geometry.)



We now collect together a number of useful characterizations for affine sets.

Theorem 2: *For a non-empty subset* A *of* E^n *the following are equivalent.*

(i) A *is an affine set.*

(ii) *If* $\underset{\sim}{x}$, $\underset{\sim}{y} \in A$, *then* $\mathcal{L}(\underset{\sim}{x},\underset{\sim}{y}) \subseteq A$.

(iii) *If* $\underset{\sim}{x}$, $\underset{\sim}{y} \in A$ *and* $\lambda \in \mathcal{R}$, *then* $\lambda\underset{\sim}{x} + (1-\lambda)\underset{\sim}{y} \in A$.

(iv) *If* $\underset{\sim}{x}$, $\underset{\sim}{y} \in A$ *and* $\alpha, \beta \in \mathcal{R}$ *are such that* $\alpha + \beta = 1$, *then* $\alpha\underset{\sim}{x} + \beta\underset{\sim}{y} \in A$.

(v) *If* $\underset{\sim}{x}_1, \underset{\sim}{x}_2, \ldots, \underset{\sim}{x}_m \in A$ *and* $\alpha_1, \alpha_2, \ldots, \alpha_m \in \mathcal{R}$ *are such that* $\sum_{i=1}^{m} \alpha_i = 1$, *then* $\sum_{i=1}^{n} \alpha_i \underset{\sim}{x}_i \in A$.

(vi) *For any* $\underset{\sim}{a} \in A$, $A - \underset{\sim}{a}$ *is a subspace of* E^n.

REMARK: You should compare (iv) with the characterization of a subspace: M is a subspace if and only if $\alpha\underset{\sim}{x} + \beta\underset{\sim}{y} \in M$ whenever $\underset{\sim}{x}$, $\underset{\sim}{y} \in M$ and $\alpha, \beta \in \mathcal{R}$.

QUESTION: Which of the following two statements is correct:

(i) Every subspace is an affine set?;

(ii) Every affine set is a subspace?

Proof (of Theorem 2). (iii) and (iv) are simply restatements of (ii) using the definition of $\mathcal{L}(\underset{\sim}{x},\underset{\sim}{y})$, we therefore have (ii) $\Leftrightarrow$ (iii) $\Leftrightarrow$ (iv). It is also clear that (v) $\Rightarrow$ (iv). [(iv) is the special case of (v) for $m = 2$.] It is therefore sufficient to show (i) $\Rightarrow$ (ii), (iv) $\Rightarrow$ (v) and (iv) $\Rightarrow$ (vi) $\Rightarrow$ (i).

(i) $\Rightarrow$ (ii).

Let $A = \underset{\sim}{x}_0 + M$, where M is the subspace parallel to A, then

$$\underset{\sim}{x} = \underset{\sim}{x}_0 + \underset{\sim}{m}_1$$

and

$$\underset{\sim}{y} = \underset{\sim}{x}_0 + \underset{\sim}{m}_2$$

for some $\underset{\sim}{m}_1$ and $\underset{\sim}{m}_2 \in M$.

Now, let $\underset{\sim}{z} \in \mathcal{L}(\underset{\sim}{x},\underset{\sim}{y})$, then

$$\begin{aligned}\underset{\sim}{z} &= \lambda\underset{\sim}{x} + (1 - \lambda)\underset{\sim}{y}, \text{ for some } \lambda \in \mathcal{R}\\ &= \lambda(\underset{\sim}{x}_0 + \underset{\sim}{m}_1) + (1 - \lambda)(\underset{\sim}{x}_0 + \underset{\sim}{m}_2)\\ &= (\lambda + (1 - \lambda))\underset{\sim}{x}_0 + \lambda\underset{\sim}{m}_1 + (1 - \lambda)\underset{\sim}{m}_2\\ &= \underset{\sim}{x}_0 + \underset{\sim}{m}, \text{ where } \underset{\sim}{m} = \lambda\underset{\sim}{m}_1 + (1 - \lambda)\underset{\sim}{m}_2\\ &\in A \text{ as } \underset{\sim}{m} \in M \text{ (since, M is a subspace).}\end{aligned}$$

(iv) $\Rightarrow$ (v).

We use induction on m. (iv) is simply the statement that (v) is true for $m = 2$. Now, assume (iv) holds for $m = k - 1$. Let

$$\underset{\sim}{y} = \sum_{i=1}^{k} \alpha_i \underset{\sim}{x}_i \text{ where } \underset{\sim}{x}_i \in A,\ \alpha_i \in \mathcal{R} \text{ for } i = 1,2,\ldots,k \text{ and } \sum_{i=1}^{k} \alpha_i = 1.$$

Since $\sum_{i=1}^{k} \alpha_i = 1$ there must exist $(k - 1)$ of the α_i's whose sum is non-zero (why?). By rearranging the sum we may, without loss of generality, assume that $S = \sum_{i=1}^{k-1} \alpha_i \neq 0$.

Let $\alpha_i' = \alpha_i/S$ for $i = 1,2,\ldots,k-1$.

Then, $\sum_{i=1}^{k-1} \alpha_i' = 1$ and so by assumption

$$\underset{\sim}{y}_1 = \sum_{i=1}^{k-1} \alpha_i' \underset{\sim}{x}_i \in A.$$

The case $m = 2$ therefore yields that

$$S\underset{\sim}{y}_1 + (1 - S)\underset{\sim}{x}_k \in A,$$

however $1 - S = \alpha_k$ $\left(S = \sum_{i=1}^{k-1} \alpha_i \text{ so } 1 = \sum_{i=1}^{k} \alpha_i = S + \alpha_k\right)$,

and
$$Sy_1 + \alpha_k x_k = S\Big(\sum_{i=1}^{k-1} \frac{\alpha_i}{S} x_i\Big) + \alpha_k x_k$$
$$= \sum_{i=1}^{k-1} \alpha_i x_i + \alpha_k x_k = y.$$

So $y \in A$ as required.

(iv) $\Rightarrow$ (vi).

Let $M = A - a$, then:

(a) For $m \in M$ and $\lambda \in \mathbb{R}$ we have $m = a_1 - a$ for some $a_1 \in A$ and so

$$\lambda m = \lambda(a_1 - a) = \lambda a_1 + (1 - \lambda)a - a$$
$$= a' - a, \text{ where } a' = \lambda a_1 + (1 - \lambda)a.$$

Now, $a' \in A$ by (iv) ($\Leftrightarrow$ (iii)) as $a_1, a \in A$, so $\lambda m \in A - a = M$.

(b) For $m_1, m_2 \in M$ we have $m_1 = a_1 - a$ and $m_2 = a_2 - a$ for some a_1 and $a_2 \in A$, and so

$$\tfrac{1}{2}(m_1 + m_2) = (\tfrac{1}{2}a_1 + \tfrac{1}{2}a_2) - a$$
$$= a' - a.$$

Now, $a' = \frac{1}{2}a_1 + \frac{1}{2}a_2 \in A$ by (iv) again, as $a_1, a_2 \in A$, so $\frac{1}{2}(m_1 + m_2) \in M$. Thus $m_1 + m_2 = 2\ (\frac{1}{2}(m_1 + m_2)) \in M$ by (a).

(a) and (b) together show M is a subspace.

(vi) $\Rightarrow$ (i).

If $M = A - a$ is a subspace, it suffices to note that $A = a + M$, so A is a transtate of the subspace M and hence A is an affine set. □

EXERCISES (1) If A_1, A_2 are affine subsets of E^n with $A_1 \cap A_2 \neq \emptyset$, show that $A_1 \cap A_2$ is an affine subset of E^n

(2) A mapping $T: E^n \to E^m$ is an <u>affine mapping</u> if $T(x) = L(x) + a$ where L is a linear mapping from E^n to E^m and a is a fixed vector in E^m. That is, T is the composite of a linear mapping and a translation.

Let T be an affine mapping from E^n to E^m, if A is any affine subset of E^n, show that T(A) is also an affine set.

(REMARK: In fact it can be shown that a mapping is affine if and only if it maps affine sets to affine sets.)

DEFINITION: By an <u>affine combination</u> of elements of any subset $S \subseteq E^n$ we mean a vector of the form

$$\sum_{i=1}^{m} \alpha_i \underset{\sim}{s}_i \text{ where } \underset{\sim}{s}_1, \underset{\sim}{s}_2, \ldots, \underset{\sim}{s}_m \in S;\ \alpha_1, \alpha_2, \ldots, \alpha_m \in \mathcal{R} \text{ and } \sum_{i=1}^{m} \alpha_i = 1.$$

Thus, an affine combination is a finite linear combination in which the scalar coefficients sum to 1.

(v) of Theorem 2 could now be restated as: *Any affine combination of elements of A is itself an element of A* (or, "A is closed under affine combinations"). By the <u>affine hull</u> of any subset $S \subseteq E^n$ we mean the set of all possible affine combinations of elements of S. The affine hull of S will be denoted by <u>aff(S)</u>. Thus,

$$\text{aff}(S) = \left\{ \underset{\sim}{x} \in E^n : \underset{\sim}{x} = \sum_{i=1}^{m} \alpha_i \underset{\sim}{s}_i, \text{ for some } m \in N,\ \underset{\sim}{s}_i \in S,\ \alpha_i \in \mathcal{R}\ (i=1,\ldots,m) \text{ and } \sum_{i=1}^{m} \alpha_i = 1 \right\}.$$

[The affine hull of S should be compared with the linear hull of S - see p.5.]

Theorem 3: *Let* $S \subseteq E^n$, *then* aff(S) *is the smallest affine set in* E^n *which contains* S. That is, aff(S) is an affine set and if A is any other affine set in E^n with $S \subseteq A$ we have $\text{aff}(S) \subseteq A$.

Proof. We first show aff(S) is an affine set.

Let $\underset{\sim}{x}, \underset{\sim}{y} \in \text{aff}(S)$, that is

$\underset{\sim}{x}$ is an affine combination of some finite set F_x of elements of S and $\underset{\sim}{y}$ is an affine combination of some other finite set F_y of elements of S. Let $F = F_x \cup F_y$, then F is a finite set of elements of S, that is $F = \{\underset{\sim}{s}_1, \underset{\sim}{s}_2, \ldots, \underset{\sim}{s}_m\}$ for some $m \in N$ and $\underset{\sim}{s}_1, \underset{\sim}{s}_2, \ldots, \underset{\sim}{s}_m \in S$. Further $\underset{\sim}{x}$ is an affine combination of elements of F. (We already know $\underset{\sim}{x}$ is an affine combination of the elements of $F_x \subseteq F$, assign 0 as the coefficient for elements in $F \backslash F_x$.) That is

$$\underset{\sim}{x} = \sum_{i=1}^{m} \alpha_i \underset{\sim}{s}_i \text{ for some } \alpha_1, \alpha_2, \ldots, \alpha_m \in \mathcal{R} \text{ with } \sum_{i=1}^{m} \alpha_i = 1.$$

Similarly, $\underset{\sim}{y} = \sum_{i=1}^{m} \beta_i \underset{\sim}{s}_i$ for some $\beta_1, \beta_2, \ldots, \beta_m \in \mathcal{R}$ with $\sum_{i=1}^{m} \beta_i = 1$.

Hence, for any $\lambda \in \mathcal{R}$ we have

$$\lambda \underset{\sim}{x} + (1 - \lambda)\underset{\sim}{y} = \lambda\Big(\sum_{i=1}^{m} \alpha_i \underset{\sim}{s}_i\Big) + (1 - \lambda)\Big(\sum_{i=1}^{m} \beta_i \underset{\sim}{s}_i\Big)$$

$$= \sum_{i=1}^{m} \big(\lambda\alpha_i + (1 - \lambda)\beta_i\big)\underset{\sim}{s}_i$$

and $\sum_{i=1}^{m} \big(\lambda\alpha_i + (1 - \lambda)\beta_i\big) = \lambda \sum_{i=1}^{m} \alpha_i + (1 - \lambda)\sum_{i=1}^{m} \beta_i = \lambda + (1 - \lambda) = 1,$

so $\lambda \underset{\sim}{x} + (1 - \lambda)\underset{\sim}{y} \in \text{aff}(S)$.

Thus by (iii) of theorem 2, aff(S) is an affine set.

Now, let A be any affine set in E^n with $S \subseteq A$. If $\underset{\sim}{x} \in \text{aff}(S)$, then $\underset{\sim}{x}$ is an affine combination of elements of S, however elements of S are also elements of A, so $\underset{\sim}{x}$ is an affine combination of elements of A. Thus, by (v) of theorem 2, $\underset{\sim}{x} \in A$. We therefore have $\text{aff}(S) \subseteq A$ as required. □

Corollary: aff(S) *equals the intersection of all affine sets in* E^n *which contain* S. That is

$$\text{aff}(S) = \bigcap\{A : S \subseteq A \text{ and } A \text{ is an affine subset of } E^n\}.$$

EXERCISE: If $S \subseteq T \subseteq E^n$ show that $\text{aff}(S) \subseteq \text{aff}(T)$

REMARK: The affine hull of a set is in many ways its more "natural home". For example, the affine hull of a triangle in E^3 is the plane containing it (show this). We normally work with a triangle as a figure in the plane, rather than a figure orientated in three dimensional space.

Affine dimension

The <u>(affine) dimension</u> of an affine set A, dim A, is the dimension of the subspace parallel to A.

The <u>dimension of any subset</u> $S \subseteq E^n$, dim S, is the dimension of aff(S).

Thus, dim S = m if and only if $\text{aff}(S) = \underset{\sim}{x} + M$ for some $\underset{\sim}{x} \in E^n$ and some subspace $M \leqslant E^n$ of dimension m.

Theorem 4: $A \subseteq E^n$ *is an affine set of dimension* m *if and only if there exists* m + 1 *vectors* $\underset{\sim}{a}_0, \underset{\sim}{a}_1, \ldots, \underset{\sim}{a}_m \in A$ *such that*

(i) $A = \text{aff}\{\underset{\sim}{a}_0, \ldots, \underset{\sim}{a}_m\}$

and (ii) for any $k \in \{0,1,2,\ldots,m\}$ *the set of* m *vectors*

$\{\underset{\sim}{a}_0-\underset{\sim}{a}_k, \underset{\sim}{a}_1-\underset{\sim}{a}_k, \ldots, \underset{\sim}{a}_{k-1}-\underset{\sim}{a}_k, \underset{\sim}{a}_{k+1}-\underset{\sim}{a}_k, \ldots, \underset{\sim}{a}_m-\underset{\sim}{a}_k\}$
is linearly independent.

REMARK: The set $\{\underset{\sim}{a}_0, \underset{\sim}{a}_1, \ldots, \underset{\sim}{a}_m\}$ is referred to as an <u>affine basis</u> for A.

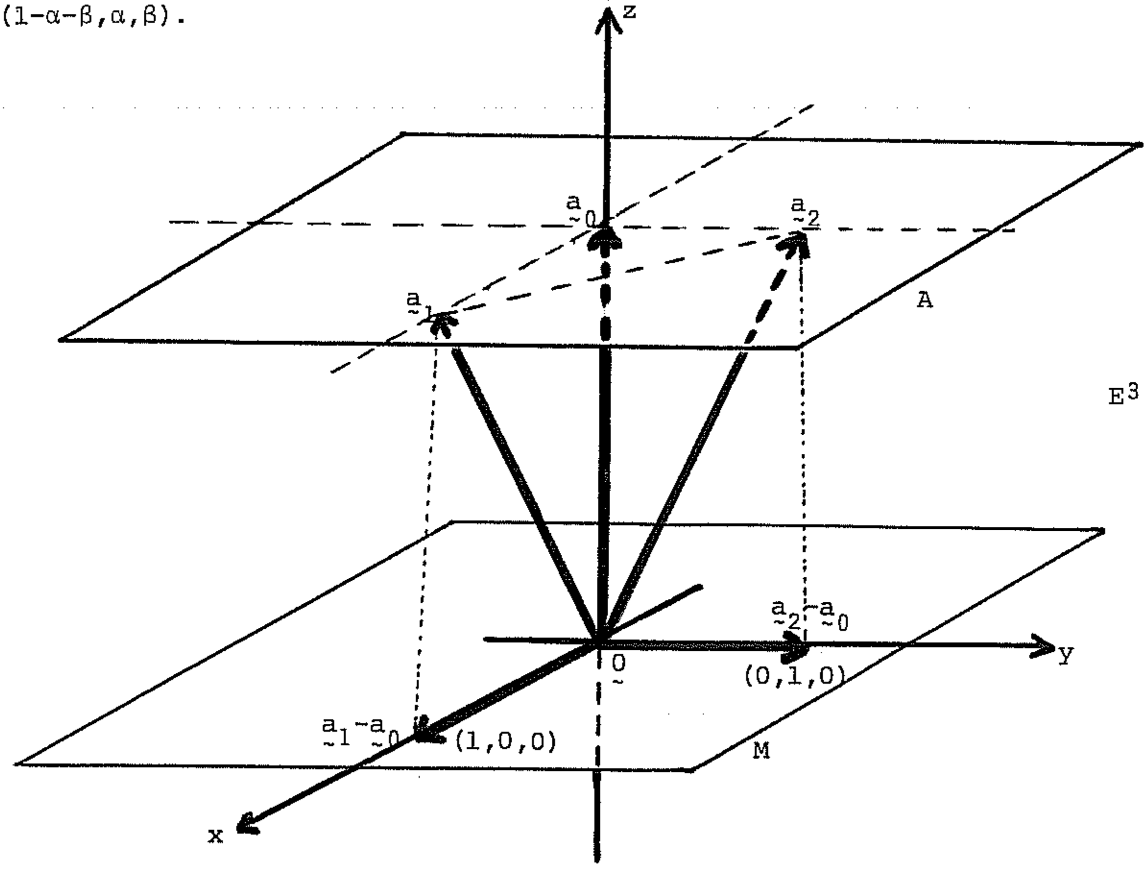
From the above theorem we can develop a theory of affine bases analogous to bases of linear spaces. For example consider the following results.

EXERCISES: (1) Show that any set of vectors $\{\underset{\sim}{a}_0, \underset{\sim}{a}_1, \ldots, \underset{\sim}{a}_m\}$ satisfying (ii) of theorem 4 is affinely independent in the sense that no vector in it is an affine combination of the other elements of the set.

(2) If $\{\underset{\sim}{a}_0, \underset{\sim}{a}_1, \ldots, \underset{\sim}{a}_m\}$ is an affine basis for A, show that for each point $\underset{\sim}{a} \in A$ there exists a unique set of scalars $\lambda_0, \lambda_1, \ldots, \lambda_m$ such that $\underset{\sim}{a} = \sum_{i=0}^{m} \lambda_i \underset{\sim}{a}_i$ and $\sum_{i=0}^{m} \lambda_i = 1$. $(\lambda_0, \lambda_1, \ldots, \lambda_m)$ are known as the barycentric coordinates of $\underset{\sim}{a}$ with respect to the given affine basis.

These ideas are illustrated in the following EXAMPLE.

In E^3 $S = \{\underset{\sim}{a}_0 = (0,0,1), \underset{\sim}{a}_1 = (1,0,1), \underset{\sim}{a}_2 = (0,1,1)\}$ is affinely independent. The affine hull of S is $A = \{(x,y,z): z = 1\}$ the plane through (0,0,1) parallel to the x-y plane. The unique subspace parallel to A is the x-y plane $M = \{(x,y,z): z = 0\}$. Since M is 2 dimensional, dim A = 2. The barycentric coordinates of the point $(\alpha,\beta,1) \in A$ with respect to S are $(1-\alpha-\beta, \alpha, \beta)$.



Proof (of theorem 4).

(⇒) Let A be an affine set of dimension m, then A = $\underset{\sim}{a}_0$ + M for some $\underset{\sim}{a}_0 \in A$ (as $\underset{\sim}{0} \in M$) and some m-dimensional subspace M.
Let $\{\underset{\sim}{b}_1, \underset{\sim}{b}_2, \ldots, \underset{\sim}{b}_m\}$ be a basis for M.
Then, for $\underset{\sim}{a} \in A$ we have $\underset{\sim}{a} = \underset{\sim}{a}_0 + \underset{\sim}{m}$ for some $\underset{\sim}{m} \in M$

$$= \underset{\sim}{a}_0 + \sum_{i=1}^{m} \lambda_i \underset{\sim}{b}_i \text{ for some } \lambda_1, \lambda_2, \ldots, \lambda_m \in \mathbb{R},$$

$$= \sum_{i=1}^{m} \lambda_i (\underset{\sim}{b}_i + \underset{\sim}{a}_0) + (1 - \sum_{i=1}^{m} \lambda_i)\underset{\sim}{a}_0.$$

This last expression is an affine combination of the m+1 vectors

$\underset{\sim}{a}_0$, $\underset{\sim}{a}_1 = \underset{\sim}{a}_0 + \underset{\sim}{b}_1$, $\underset{\sim}{a}_2 = \underset{\sim}{a}_0 + \underset{\sim}{b}_2, \ldots, \underset{\sim}{a}_m = \underset{\sim}{a}_0 + \underset{\sim}{b}_m$, so

$\underset{\sim}{a} \in \text{aff}\{\underset{\sim}{a}_0, \underset{\sim}{a}_1, \ldots, \underset{\sim}{a}_m\}$ or $A \subseteq \text{aff}\{\underset{\sim}{a}_0, \underset{\sim}{a}_1, \ldots, \underset{\sim}{a}_m\}$.

Further each $\underset{\sim}{a}_i \in A$ so $\text{aff}\{\underset{\sim}{a}_0, \underset{\sim}{a}_1, \ldots, \underset{\sim}{a}_m\} \subseteq A$. We therefore have $A = \text{aff}\{\underset{\sim}{a}_0, \underset{\sim}{a}_1, \ldots, \underset{\sim}{a}_m\}$, establishing (i).

To establish (ii) first note that if k = 0 we have

$\{\underset{\sim}{a}_1-\underset{\sim}{a}_0, \ldots, \underset{\sim}{a}_m-\underset{\sim}{a}_0\} = \{\underset{\sim}{b}_1, \ldots, \underset{\sim}{b}_m\}$ is linearly independent.
On the other hand, if $k \in \{1,2,\ldots,m\}$ we have

$$\{\underset{\sim}{a} - \underset{\sim}{a}_k, \underset{\sim}{a}_1-\underset{\sim}{a}_k, \ldots, \underset{\sim}{a}_{k-1} - \underset{\sim}{a}_k, \underset{\sim}{a}_{k+1} - \underset{\sim}{a}_k, \ldots, \underset{\sim}{a}_m - \underset{\sim}{a}_k\}$$
$$= \{-\underset{\sim}{b}_k, \underset{\sim}{b}_1-\underset{\sim}{b}_k, \ldots, \underset{\sim}{b}_{k-1}-\underset{\sim}{b}_k, \underset{\sim}{b}_{k+1}-\underset{\sim}{b}_k, \ldots, \underset{\sim}{b}_m-\underset{\sim}{b}_k\}$$

which is linearly independent since the $\underset{\sim}{b}_i$'s are. (Why?)

(⇐) From (i) we have that A is an affine set. Let M be the unique subspace parallel to A. By (vi) of theorem 2 M = A − $\underset{\sim}{a}_0$ or A = $\underset{\sim}{a}_0$ + M. Hence, if $\underset{\sim}{m}$ is any element of M $\underset{\sim}{a}_0 + \underset{\sim}{m} \in A$, so by (i),

$$\underset{\sim}{a}_0 + \underset{\sim}{m} = \sum_{i=0}^{m} \alpha_i \underset{\sim}{a}_i \text{ for some } \alpha_0, \alpha_1, \ldots, \alpha_m \in \mathbb{R} \text{ with } \sum_{i=0}^{m} \alpha_i = 1.$$

$$\text{Thus, } \underset{\sim}{m} = \left(\sum_{i=0}^{m} \alpha_i \underset{\sim}{a}_i\right) - \underset{\sim}{a}_0 = \left(\sum_{i=0}^{m} \alpha_i \underset{\sim}{a}_i - \sum_{i=0}^{m} \alpha_i \underset{\sim}{a}_0\right)$$

$$= \sum_{i=0}^{m} \alpha_i (\underset{\sim}{a}_i - \underset{\sim}{a}_0) = \sum_{i=1}^{m} \alpha_i (\underset{\sim}{a}_i - \underset{\sim}{a}_0).$$

This shows that $\underset{\sim}{m}$ is a linear combination of $\{\underset{\sim}{a}_1-\underset{\sim}{a}_0, \ldots, \underset{\sim}{a}_m-\underset{\sim}{a}_0\}$, so $\{\underset{\sim}{a}_1-\underset{\sim}{a}_0, \ldots, \underset{\sim}{a}_m-\underset{\sim}{a}_0\}$ spans M. Further by (ii) $\{\underset{\sim}{a}_1-\underset{\sim}{a}_0, \ldots, \underset{\sim}{a}_m-\underset{\sim}{a}_0\}$ is linearly independent. It follows that M has dimension m, and so dim A = m. □

The concluding problem introduces some valuable *topological properties* of an affine subset of E^n.

PROBLEM: *Let* A *be an affine subset of* E^n show that

(i) A *is closed*

and hence (ii) A *is complete.*

§2. Hyperplanes

DEFINITION: A <u>hyperplane</u> in n-dimensional Euclidean space E^n is an affine subset of (affine) dimension n - 1. That is, $H \subseteq E^n$ is a hyperplane if and only if $H = x_0 + M$ where M, the unique subspace parallel to H, has dimension n - 1.

For EXAMPLE: A hyperplane in E^2 is a line.

A hyperplane in E^3 is a plane.

We begin the study of hyperplanes by characterizing (n-1) dimensional subspaces of E^n.

DEFINITION: For any non-zero vector $\underset{\sim}{x}_0 \in E^n$ we define the <u>orthogonal complement</u> of $\underset{\sim}{x}_0$, $\underset{\sim}{x}_0^{\perp}$ (pronounced "x_0-perp.") to be

$$\underset{\sim}{x}_0^{\perp} = \{\underset{\sim}{y} \in E^n : (\underset{\sim}{y}, \underset{\sim}{x}_0) = 0\}$$

EXERCISE: Show that $\underset{\sim}{x}_0^{\perp}$ is a subspace of E^n for any $\underset{\sim}{x}_0 \in E^n$, $\underset{\sim}{x}_0 \neq 0$.
[Hint: See the definition and exercises on inner-products given on p.2.]

THEOREM 1: *M is an (n-1)-dimensional subspace of E^n if and only if $M = \underset{\sim}{x}_0^{\perp}$ for some non-zero $\underset{\sim}{x}_0 \in E^n$.*

Proof. (⇐) For any non-zero $\underset{\sim}{x}_0 \in E^n$ we must show $\underset{\sim}{x}_0^{\perp}$ has dimension n-1. We begin by showing that any vector $\underset{\sim}{y} \in E^n$ may be written as $\underset{\sim}{y} = \lambda\underset{\sim}{x}_0 + \underset{\sim}{m}$ for some scalar λ and some $\underset{\sim}{m} \in \underset{\sim}{x}_0^{\perp}$. Since $\underset{\sim}{x}_0 \neq 0$ we have $\|\underset{\sim}{x}_0\|^2 = (\underset{\sim}{x}_0, \underset{\sim}{x}_0) \neq 0$ and so the scalar $\frac{(\underset{\sim}{y}, \underset{\sim}{x}_0)}{(\underset{\sim}{x}_0, \underset{\sim}{x}_0)}$ may be formed. Let $\lambda = \frac{(\underset{\sim}{y}, \underset{\sim}{x}_0)}{(\underset{\sim}{x}_0, \underset{\sim}{x}_0)}$, then $\underset{\sim}{y} = \lambda\underset{\sim}{x}_0 + \underset{\sim}{m}$, where $\underset{\sim}{m} = \underset{\sim}{y} - \lambda\underset{\sim}{x}_0$ is such that

$$(\underset{\sim}{m}, \underset{\sim}{x}_0) = (\underset{\sim}{y}, \underset{\sim}{x}_0) - \lambda(\underset{\sim}{x}_0, \underset{\sim}{x}_0) = (\underset{\sim}{y}, \underset{\sim}{x}_0) - \frac{(\underset{\sim}{y}, \underset{\sim}{x}_0)}{(\underset{\sim}{x}_0, \underset{\sim}{x}_0)}(\underset{\sim}{x}_0, \underset{\sim}{x}_0) = 0,$$

that is, $\underset{\sim}{m} \in \underset{\sim}{x}_0^{\perp}$.

Now, if $\underset{\sim}{b}_1, \ldots, \underset{\sim}{b}_k$ is a basis for $\underset{\sim}{x}_0^{\perp}$ we have $\underset{\sim}{m} = \sum_{i=1}^{k} \lambda_i \underset{\sim}{b}_i$ and so any $\underset{\sim}{y} \in E^n$ may be written as $\underset{\sim}{y} = \lambda\underset{\sim}{x}_0 + \lambda_1\underset{\sim}{b}_1 + \ldots + \lambda_k\underset{\sim}{b}_k$. Thus the set of k + 1 vectors $\{\underset{\sim}{x}_0, \underset{\sim}{b}_1, \ldots, \underset{\sim}{b}_k\}$ is a basis for E^n and so n = k + 1. It now follows that $\dim \underset{\sim}{x}_0^{\perp} = k = n - 1$.

(⇒) Let M be an (n-1)-dimensional subspace of E^n. By the remarks made on p.5 (last paragraph), there is an orthonormal basis $\{\underset{\sim}{b}_1, \ldots, \underset{\sim}{b}_n\}$ of E^n with $M = \langle\underset{\sim}{b}_1, \ldots, \underset{\sim}{b}_{n-1}\rangle$. We show $M = \underset{\sim}{b}_n^{\perp}$. Since $(\underset{\sim}{b}_i, \underset{\sim}{b}_n) = 0$ for i = 1,2,...,n-1 and any $\underset{\sim}{m} \in M$ may be expressed as $\underset{\sim}{m} = \sum_{i=1}^{n-1} \lambda_i \underset{\sim}{b}_i$ for some $\lambda_1, \ldots, \lambda_{n-1}$ it follows that $(\underset{\sim}{m}, \underset{\sim}{b}_n) = 0$. Thus $M \leqslant \underset{\sim}{b}_n^{\perp}$.

In addition, $\dim \underset{\sim}{b}_n^{\perp} = n - 1$, by the (⇐) part of the proof

$= \dim M$

and so, $M = \underset{\sim}{b}_n^{\perp}$. The result now follows by setting $\underset{\sim}{x}_0 = \underset{\sim}{b}_n$. □

Corollary 2: *H is a hyperplane in E^n if and only if* *$H = \lambda \underset{\sim}{x}_0 + \underset{\sim}{x}_0^{\perp}$ for some λ and some non-zero vector $\underset{\sim}{x}_0 \in E^n$.*

Proof. (⇐) Follows immediately from the definition of hyperplane and the preceding theorem.

(⇒) If H is a hyperplane, then $H = \underset{\sim}{a}_0 + M$ for some $\underset{\sim}{a}_0 \in E^n$ and some (n-1)-dimensional subspace M of E^n. By theorem 1 $M = \underset{\sim}{x}_0^{\perp}$ for some non-zero $\underset{\sim}{x}_0 \in E^n$. Further $\underset{\sim}{a}_0 = \lambda \underset{\sim}{x}_0 + \underset{\sim}{m}$ for some λ and some $\underset{\sim}{m} \in M$ (see the (⇐) part of the proof to theorem 1).

Thus $H = (\lambda \underset{\sim}{x}_0 + \underset{\sim}{m}) + \underset{\sim}{x}_0^{\perp} = \lambda \underset{\sim}{x}_0 + \underset{\sim}{x}_0^{\perp}$ as $\underset{\sim}{m} \in \underset{\sim}{x}_0^{\perp}$ so $\underset{\sim}{m} + \underset{\sim}{x}_0^{\perp} = \underset{\sim}{x}_0^{\perp}$. □

EXERCISE: If H is a hyperplane with $\underset{\sim}{0} \notin H$, show that $H = \underset{\sim}{x}_0 + \underset{\sim}{x}_0^{\perp}$ for some non-zero $\underset{\sim}{x}_0 \in E^n$. [Hint: In this case show that the λ in corollary 2 is not equal to zero, then show that $\underset{\sim}{x}_0^{\perp} = (\lambda \underset{\sim}{x}_0)^{\perp}$ for any $\lambda \neq 0$.]

As a consequence of Corollary 2 we arrive at our final, and most useful, characterization of a hyperplane.

THEOREM 3. *H is a hyperplane in E^n if and only if for some constant $c \in \mathbb{R}$ and some non-zero $\underset{\sim}{x}_0 \in E^n$ we have $H = \{\underset{\sim}{y} \in E^n : (\underset{\sim}{y}, \underset{\sim}{x}_0) = c\}$.*

Proof. (⇒) If H is a hyperplane then $H = \lambda \underset{\sim}{x}_0 + \underset{\sim}{x}_0^{\perp}$ for some λ and some $\underset{\sim}{x}_0 \in E^n$, $\underset{\sim}{x}_0 \neq \underset{\sim}{0}$ (Corollary 2). Thus

$$\begin{aligned} \underset{\sim}{y} \in H &\Leftrightarrow \underset{\sim}{y} = \lambda \underset{\sim}{x}_0 + \underset{\sim}{x}_0^{\perp} \\ &\Leftrightarrow \underset{\sim}{y} - \lambda \underset{\sim}{x}_0 \in \underset{\sim}{x}_0^{\perp} \\ &\Leftrightarrow (\underset{\sim}{y} - \lambda \underset{\sim}{x}_0, \underset{\sim}{x}_0) = 0 \\ &\Leftrightarrow (\underset{\sim}{y}, \underset{\sim}{x}_0) = \lambda (\underset{\sim}{x}_0, \underset{\sim}{x}_0) \end{aligned}$$

The result now follows by taking $c = \lambda (\underset{\sim}{x}_0, \underset{\sim}{x}_0)$.

(⇐) If $H = \{\underset{\sim}{y} \in E^n : (\underset{\sim}{y}, \underset{\sim}{x}_0) = c\}$, then

$$\begin{aligned} \underset{\sim}{y} \in H &\Leftrightarrow (\underset{\sim}{y}, \underset{\sim}{x}_0) = c \\ &\Leftrightarrow (\underset{\sim}{y}, \underset{\sim}{x}_0) = \lambda (\underset{\sim}{x}_0, \underset{\sim}{x}_0) \text{ where } \lambda = \frac{c}{(\underset{\sim}{x}_0, \underset{\sim}{x}_0)} \\ &\Leftrightarrow (\underset{\sim}{y} - \lambda \underset{\sim}{x}_0, \underset{\sim}{x}_0) = 0 \\ &\Leftrightarrow \underset{\sim}{y} - \lambda \underset{\sim}{x}_0 \in \underset{\sim}{x}_0^{\perp} \\ &\Leftrightarrow \underset{\sim}{y} \in \lambda \underset{\sim}{x}_0 + \underset{\sim}{x}_0^{\perp} \end{aligned}$$

Thus, $H = \lambda \underset{\sim}{x}_0 + \underset{\sim}{x}_0^{\perp}$ and so H is a hyperplane. □

ILLUSTRATION: From theorem 3 we have that H is a hyperplane in E^3 if and only if $H = \{\underset{\sim}{y} \in E^3\colon (\underset{\sim}{y}, \underset{\sim}{x}_0) = c\}$ for some $c \in \mathcal{R}$ and some $\underset{\sim}{x}_0 \in E^n$, $\underset{\sim}{x}_0 \neq 0$.
Letting $\underset{\sim}{y} = (x, y, z)$ and $\underset{\sim}{x}_0 = (a_1, a_2, a_3)$ we therefore have

$$H = \{(x,y,z) \in E^3\colon a_1 x + a_2 y + a_3 z = c\} \text{ - a plane in } E^3!$$

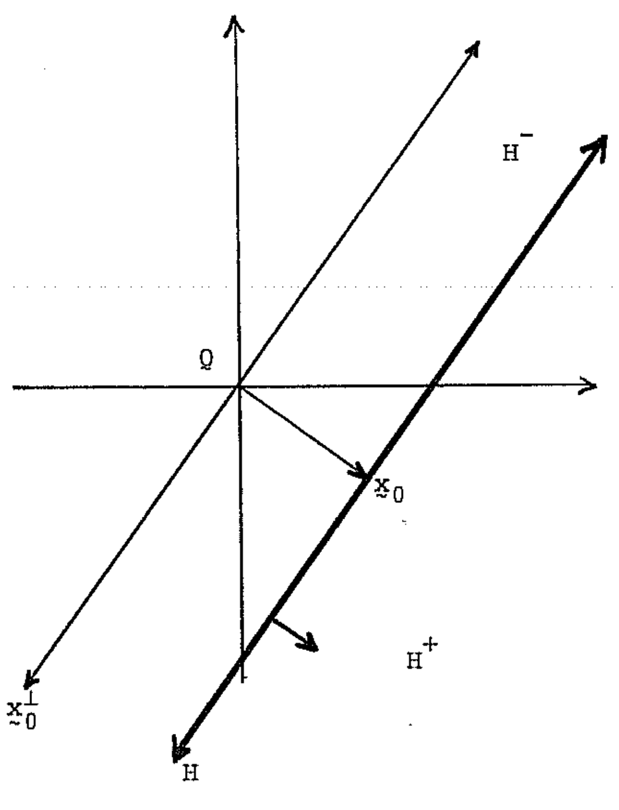
Just as a line divides E^2 into two half-spaces and a plane divides E^3 into two half-spaces, theorem 3 shows that in E^n a hyperplane $H = \{\underset{\sim}{y} \in E^n\colon (\underset{\sim}{y}, \underset{\sim}{x}_0) = c\}$ may be used to divide E^n into two closed "half-spaces"

$$H^+ = \{\underset{\sim}{y} \in E^n\colon (\underset{\sim}{y}, \underset{\sim}{x}_0) \geqslant c\}$$

and

$$H^- = \{\underset{\sim}{y} \in E^n\colon (\underset{\sim}{y}, \underset{\sim}{x}_0) \leqslant c\}$$

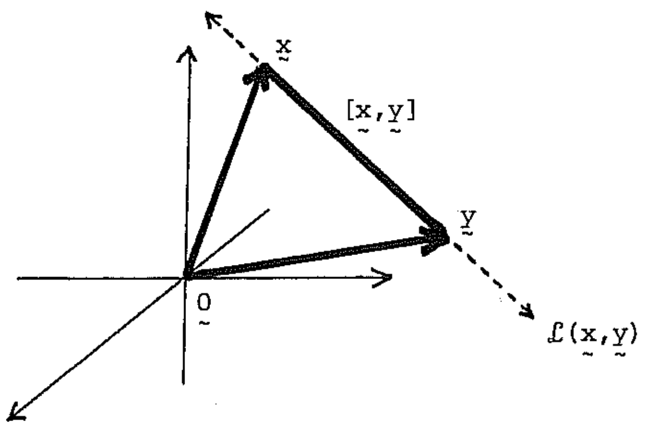
Note: The Open Half-spaces $\{\underset{\sim}{y} \in E^n\colon (\underset{\sim}{y}, \underset{\sim}{x}_0) > c\}$ and $\{\underset{\sim}{y} \in E^n\colon (\underset{\sim}{y}, \underset{\sim}{x}_0) < c\}$ can also be defined.



§3. Convex Sets

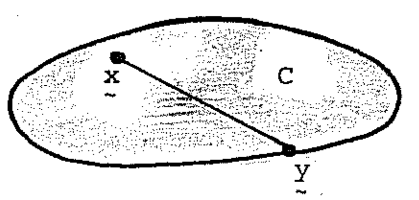
3.1 *Algebraic Structure for Convex Sets*

For $\underset{\sim}{x}, \underset{\sim}{y} \in E^n$ we denote by $[\underset{\sim}{x},\underset{\sim}{y}]$ the segment of line between $\underset{\sim}{x}$ and $\underset{\sim}{y}$.

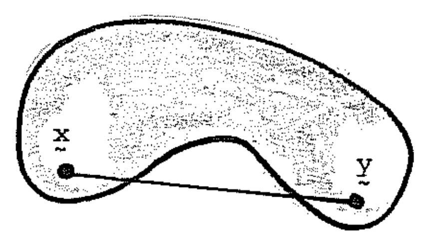


Thus $[\underset{\sim}{x},\underset{\sim}{y}] \subset \mathcal{L}(\underset{\sim}{x},\underset{\sim}{y})$ and $\underset{\sim}{z} \in [\underset{\sim}{x},\underset{\sim}{y}]$ if and only if $\underset{\sim}{z} = \lambda\underset{\sim}{x} + (1-\lambda)\underset{\sim}{y}$ for some λ with $0 \le \lambda \le 1$ (see diagram on page 7).

DEFINITION: A subset C of E^n is <u>CONVEX</u> if for every pair of points $\underset{\sim}{x}, \underset{\sim}{y} \in C$ we have $[\underset{\sim}{x},\underset{\sim}{y}] \subseteq C$.
That is, C *is convex if and only if, for every pair of points* $\underset{\sim}{x}, \underset{\sim}{y} \in C$ *and every* λ *with* $0 \le \lambda \le 1$ *we have* $\lambda\underset{\sim}{x} + (1-\lambda)\underset{\sim}{y} \in C$.

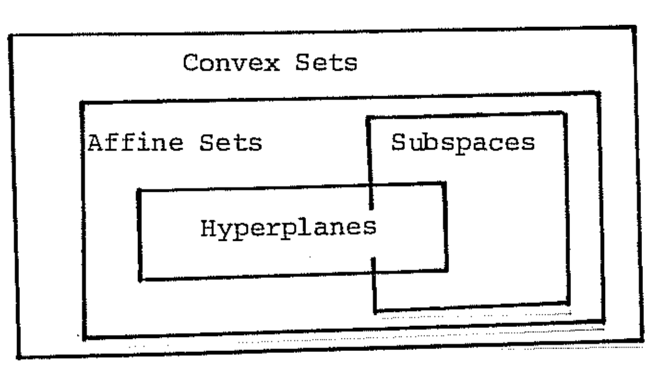


A convex subset of E^2.



A non-convex subset of E^2.

EXAMPLES: 1) For the type of sets studied so far we have the following inclusions.



(You should convince yourself of the truth of these statements, viz: Any affine set is a convex set etc. You should also give examples to show that each of the inclusions is proper.)

2) *Any half-space* (see page 17) *is convex.* For example, let

$$C = H^{+} = \{\underset{\sim}{y} \in E^n : (\underset{\sim}{y}, \underset{\sim}{x}_0) \geq c\} .$$

To show C is convex, let $\underset{\sim}{x}, \underset{\sim}{y} \in C$, then $(\underset{\sim}{x}, \underset{\sim}{x}_0) \geq c$ and $(\underset{\sim}{y}, \underset{\sim}{x}_0) \geq c$. For any λ with $0 \leq \lambda \leq 1$ we therefore have

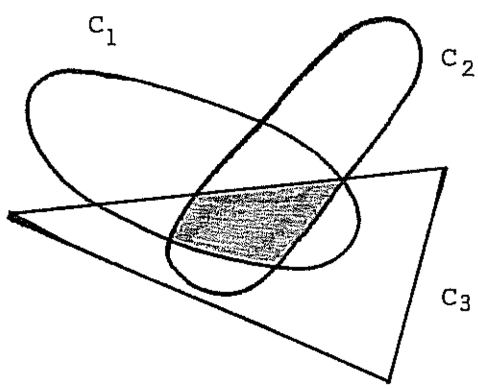
$$\begin{aligned}(\lambda\underset{\sim}{x} + (1-\lambda)\underset{\sim}{y}, \underset{\sim}{x}_0) &= \lambda(\underset{\sim}{x}, \underset{\sim}{x}_0) + (1-\lambda)(\underset{\sim}{y}, \underset{\sim}{x}_0) \\ &\geq \lambda c + (1-\lambda)c \quad \text{as both } \lambda \text{ and } (1-\lambda) \text{ are positive} \\ &= c.\end{aligned}$$

Thus $\lambda\underset{\sim}{x} + (1-\lambda)\underset{\sim}{y} \in C$, and so C is convex.

We will make extensive use of the following result.

THEOREM 1: *The intersection of any collection of convex sets in* E^n *is itself a convex set*[1].

Proof. EXERCISE.



EXERCISES: 1) Prove Theorem 1 above.

2) If A, B are convex subsets of E^n show that

i) $A + B = \{\underset{\sim}{x} \in E^n : \underset{\sim}{x} = \underset{\sim}{a} + \underset{\sim}{b}$ for some $\underset{\sim}{a} \in A$ and some $\underset{\sim}{b} \in B\}$ is a convex set;

(1) Here, and elsewhere, we take the empty set $\emptyset$ to be a convex set. (Indeed, how could $\emptyset$ be non-convex?)

and ii) $\mu A = \{\underset{\sim}{x} \in E^n : \underset{\sim}{x} = \mu\underset{\sim}{a} \text{ for some } \underset{\sim}{a} \in A\}$ is a convex set. Here μ is any fixed real number.

Hence deduce that $A + \mu B$ is a convex set.

3) Give an example to show that for two convex sets A and B, the union $A \cup B$ is not in general convex.

4) If A is a convex subset of E^n and $T : E^n \to E^m$ is an affine mapping (see Exercise 2 on p.9), show that T(A) is a convex subset of E^m.

5) Show that the ball of radius $r > 0$ and centre $\underset{\sim}{x} \in E^n$, $B_r(\underset{\sim}{x})$ - see p.2 - is a convex subset. [Hint: Use the triangle inequality given on p.4.]

6) For any pair of real numbers a, b with $a < b$ show that the closed interval $[a,b] = \{x \in \mathcal{R}: a \le x \le b\}$ is a convex subset of E^1.

The following corollary to Theorem 1 is of considerable importance for many applications, particularly linear programming.

THEOREM 2: *The solution set of a system of simultaneous linear inequalities (including equalities) in* n *variables is a convex subset of* E^n .

Proof. A given system of m simultaneous linear inequalities in the n real variables $x_1, x_2, \ldots, x_n$ has the form

$$\begin{array}{l} a_{11}x_1 + a_{12}x_2 + \ldots + a_{1n}x_n \;\; \mathcal{R}_1 \; b_1 \\ a_{21}x_1 + a_{22}x_2 + \ldots + a_{2n}x_n \;\; \mathcal{R}_2 \; b_2 \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ a_{m1}x_1 + a_{m2}x_2 + \ldots + a_{mn}x_n \;\; \mathcal{R}_m \; b_m \end{array}$$

where, for $i = 1,2,\ldots,m$, $\mathcal{R}_i$ stands for one of the five relationships $<, \le, =, \ge, >$.

This system may be rewritten as

$$(\underset{\sim}{x},\underset{\sim}{a}_1)\ \mathcal{R}_1\ b_1$$
$$(\underset{\sim}{x},\underset{\sim}{a}_2)\ \mathcal{R}_2\ b_2$$
$$\cdots\cdots\cdots\cdots$$
$$(\underset{\sim}{x},\underset{\sim}{a}_m)\ \mathcal{R}_m\ b_m$$

where $\underset{\sim}{x} = (x_1, x_2, \ldots, x_n)$ and $\underset{\sim}{a}_i = (a_{i1}, a_{i2}, \ldots, a_{in})$ for $i = 1,2,\ldots,m$.

The solution set of the simultaneous system is

$$C = C_1 \cap C_2 \cap \ldots \cap C_m$$

where, for $i = 1,2,\ldots,m$, C_i is the solution set of the i'th linear inequality, that is

$$C_i = \{\underset{\sim}{x} \in E^n: (\underset{\sim}{x},\underset{\sim}{a}_i)\ \mathcal{R}_i\ b_i\} .$$

From this we recognise that each C_i is either a half-space or a hyperplane and so, C_i is a convex set. The convexity of C now follows from Theorem 1. □

Convex Combinations

Let S be any non-empty subset of E^n. We say $\underset{\sim}{x}$ is a (finite) <u>convex combination</u> of element of S if

$$\underset{\sim}{x} = \lambda_1\underset{\sim}{s}_1 + \lambda_2\underset{\sim}{s}_2 + \ldots + \lambda_m\underset{\sim}{s}_m \quad \text{for some } m \in N ;$$

$\underset{\sim}{s}_1, \ldots, \underset{\sim}{s}_m \in S$, and scalars $\lambda_1, \lambda_2, \ldots, \lambda_m$ with $\lambda_i \geq 0$ for $i = 1,2,\ldots,m$ and $\sum_{i=1}^{m} \lambda_i = 1$ (clearly this requires that each $\lambda_i \in [0,1]$).

By the <u>convex hull</u> of S we understand the set of all possible convex combinations of elements of S, which we denote by <u>co(S)</u>.

Since any element $\underset{\sim}{s} \in S$ may be written as $\underset{\sim}{s} = 1\underset{\sim}{s} + 0\underset{\sim}{s}$, we see that $S \subseteq co(S)$.

We begin by establishing that a convex set is closed under convex combinations.

LEMMA 3: *If* C *is a convex subset of* E^n, *then*

$$co(C) = C.$$

Proof. By the preceeding remark it suffices to show that $co(C) \subseteq C$. Thus, if $\underset{\sim}{x} = \sum_{i=1}^{m} \lambda_i \underset{\sim}{c}_i$ where $\underset{\sim}{c}_1, \ldots, \underset{\sim}{c}_m \in C$; $\lambda_1, \ldots, \lambda_m \geq 0$ and $\sum_{i=1}^{m} \lambda_i = 1$, we must show $\underset{\sim}{x} \in C$.

The proof is by induction on m the length of the sum.

If m = 2 then $\underset{\sim}{x} = \lambda_1 \underset{\sim}{c}_1 + (1-\lambda_1)\underset{\sim}{c}_2 \in C$ by the definition of convexity.

Assume every convex combination of length (m - 1) or less is in C, and let $\underset{\sim}{x} = \sum_{i=1}^{m} \lambda_i \underset{\sim}{c}_i$, where $m \geq 3$ and without loss of generality we may assume that each $\lambda_i \gneq 0$ (otherwise $\underset{\sim}{x}$ is really a convex combination of length (m - 1) or less and so is in C).

Now, $\underset{\sim}{x} = \lambda_1 \underset{\sim}{c}_1 + \sum_{i=2}^{m} \lambda_i \underset{\sim}{c}_i$ and $\mu = \sum_{i=2}^{m} \lambda_i \gneq 0$ (since $\lambda_2, \lambda_3, \ldots > 0$).

Let $\underset{\sim}{y} = \sum_{i=2}^{m} \frac{\lambda_i}{\mu} \underset{\sim}{c}_i$, then $\frac{\lambda_i}{\mu} > 0$ for i = 2, ..., m and $\sum_{i=2}^{m} \frac{\lambda_i}{\mu} = \frac{1}{\mu} \sum_{i=2}^{m} \lambda_i = \frac{1}{\mu} \cdot \mu = 1$, so $\underset{\sim}{y}$ is a convex combination of elements of C of length (m - 1) and so $\underset{\sim}{y} \in C$.

Further, $\underset{\sim}{x} = \lambda_1 \underset{\sim}{c}_1 + \mu \underset{\sim}{y}$ and $\lambda_1, \mu > 0$ with $\lambda_1 + \mu = \lambda_1 + \sum_{i=2}^{m} \lambda_i = \sum_{i=1}^{m} \lambda_i = 1$, hence $\underset{\sim}{x}$ is a convex combination of elements of C of length 2 and so $\underset{\sim}{x} \in C$. □

THEOREM 4: *For any non-empty subset* S *of* E^n, co(S) *is the smallest convex subset of* E^n *containing* S.

Proof. We must show i) co(S) is convex

ii) if C is any convex subset of E^n with $S \subseteq C$, then $co(S) \subseteq C$.

i) For any $\underset{\sim}{x}, \underset{\sim}{y} \in co(S)$ we have

$$\underset{\sim}{x} = \sum_{i=1}^{m} \lambda_i \underset{\sim}{S}_i \text{ for some } m \in \mathbb{N};\ \underset{\sim}{S}_1, \ldots, \underset{\sim}{S}_m \in S \text{ and } \lambda_1, \lambda_2, \ldots, \lambda_m \geq 0 \text{ with } \sum_{i=1}^{m} \lambda_i = 1$$

and

$$\underset{\sim}{y} = \sum_{i=1}^{m'} \mu_i \underset{\sim}{S}_i' \text{ for some } m' \in \mathbb{N};\ \underset{\sim}{S}_1', \ldots, \underset{\sim}{S}_{m'}' \in S \text{ and } \mu_1, \mu_2, \ldots, \mu_{m'} \geq 0 \text{ with } \sum_{i=1}^{m'} \mu_i = 1.$$

Further, without loss of generality we may assume $m = m'$ and $\underset{\sim}{S}_i = \underset{\sim}{S}_i'$ for $i = 1,2,\ldots,m$ (assign 0 as the coefficient of any "unnecessary" $\underset{\sim}{S}_i$'s in either sum).

Now, for any λ with $0 \leq \lambda \leq 1$ we have

$$\lambda \underset{\sim}{x} + (1-\lambda)\underset{\sim}{y} = \lambda \sum_{i=1}^{m} \lambda_i \underset{\sim}{S}_i + (1-\lambda) \sum_{i=1}^{m} \lambda_i \underset{\sim}{S}_i$$

$$= \sum_{i=1}^{m} \left[\lambda\lambda_i + (1-\lambda)\mu_i\right] \underset{\sim}{S}_i$$

and, for each $i = 1,2,\ldots,m$ $\lambda\lambda_i + (1-\lambda)\mu_i \geq 0$, as each of λ, $(1-\lambda)$, λ_i and μ_i is positive, while

$$\sum_{i=1}^{m} \left[\lambda\lambda_i + (1-\lambda)\mu_i\right] = \lambda \sum_{i=1}^{m} \lambda_i + (1-\lambda) \sum_{i=1}^{m} \mu_i$$

$$= \lambda \cdot 1 + (1-\lambda) \cdot 1 = 1.$$

So, $\lambda \underset{\sim}{x} + (1-\lambda)\underset{\sim}{y}$ is a convex combination of elements of S, that is $\lambda \underset{\sim}{x} + (1-\lambda)\underset{\sim}{y} \in co(S)$. It follows that co(S) is convex.

ii) Since $S \subseteq C$ any convex combination of elements of S is also a convex combination of elements of C, that is, $co(S) \subseteq co(C)$.

By lemma 3, $co(C) = C$, so $co(S) \subseteq C$.

□

Corollaries: 1) *For any non-empty subset* S *of* E^n *we have* co(S) *equals the intersection of all convex subsets of* E^n *which contain* S.

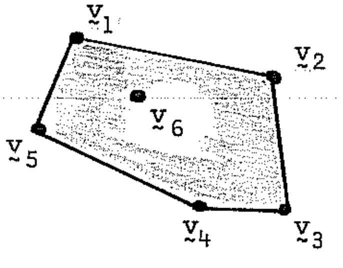
2) *A subset* C *of* E^n *is convex if and only if* C = co(C).

EXERCISES: 1) In E^2 find co({(0,0),(1,0),(0,1)}).

2) Show that co(A ∩ B) ⊆ co(A) ∩ co(B) for any two subsets A, B of E^n. Give an example to show that the reverse inclusion need not be true.

SIMPLICES

The convex hull of a finite set of vectors in E^n is known as a polytope. For example: in E^2 any convex polygon is the convex hull of its vertices and so is a polytope;



a polytope in E^2.

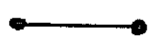
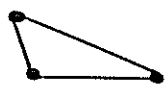
in E^3 any convex polyhedron is the convex hull of its vertices and so is a polytope.

In E^n, a polytope which is the convex hull of an affinely independent set of vectors (see p.12) is known as a simplex. That is, $S \subseteq E^n$ is a simplex if and only if

$$S = \text{co}\,\{\underset{\sim}{b}_0, \underset{\sim}{b}_1, \ldots, \underset{\sim}{b}_m\}, \text{ where}$$

$\underset{\sim}{b}_1 - \underset{\sim}{b}_0$, $\underset{\sim}{b}_2 - \underset{\sim}{b}_0$, ..., $\underset{\sim}{b}_m - \underset{\sim}{b}_0$ is a linearly independent set of m vectors. The vectors $\underset{\sim}{b}_0, \underset{\sim}{b}_1, \ldots, b_m$ are known as the vertices of S. The number of vertices, m, is the simplectic dimension of S. Since the number of vectors in a linearly independent subset of E^n is at most n, the types of simplices in E^n is strictly limited.

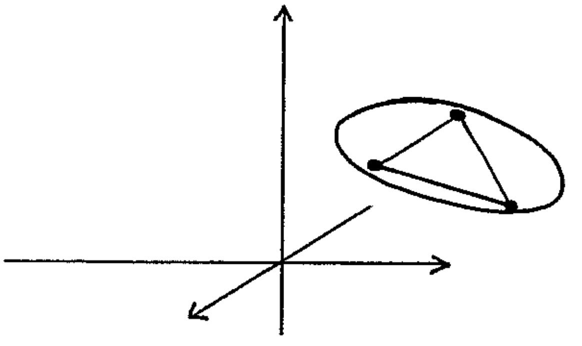
For EXAMPLE we have:

Space	Possible Simplices	Simplectic dimension (= Number of vertices)
E^2	a single point 	1
	a closed line segment 	2
	a triangle 	3
E^3	a single point 	1
	a closed line segment 	2
	a triangle 	3
	a tetrahedron 	4

It is of basic importance that every polytope in E^n is a union of simplices from E^n. We will not require this result and so will not attempt to prove it.

THEOREM 5: *The dimension of a convex set* $C \subseteq E^n$ (see the definition on p.11) *equals the maximum of the simplectic dimensions of the simplices contained in* C *minus* 1.

Thus, the dimension of a circular disc in E^3 is 2, since it contains a triangle (simplectic dimension 3) but does not contain any tetrahedron.



Proof. Let k+1 be the largest simplectic dimension of any simplex contained in C.

Let $S = \text{co}\,\{\underset{\sim}{b}_0, \underset{\sim}{b}_1, \ldots, \underset{\sim}{b}_k\}$ be a simplex in C of simplectic dimension k+1, and let $A = \text{aff}\,\{\underset{\sim}{b}_0, \underset{\sim}{b}_1, \ldots, \underset{\sim}{b}_k\}$. Then dim A = k, and since $\{\underset{\sim}{b}_0, \underset{\sim}{b}_1, \ldots, \underset{\sim}{b}_k\} \subseteq C$ we have $A \subseteq \text{aff}\, C$ (see exercise on p.11). We next show $C \subseteq A$. For, assume this were not the case, then there exists $\underset{\sim}{b} \in C \backslash A$. Thus, $\underset{\sim}{b} \notin \text{aff}\,\{\underset{\sim}{b}_0, \ldots, \underset{\sim}{b}_k\}$ so, by exercise 1) on p.12, $\{\underset{\sim}{b}_0, \underset{\sim}{b}_1, \ldots, \underset{\sim}{b}_k, \underset{\sim}{b}\}$ is an affinely independent set of k+2 vectors in C and $S' = \text{co}\,\{\underset{\sim}{b}_0, \underset{\sim}{b}_1, \ldots, \underset{\sim}{b}_k, \underset{\sim}{b}\}$ is a simplex in C of simplectic dimension k+2, contradicting the choice of k.

We therefore have $C \subseteq A$ and so $\text{aff}(C) \subseteq A$.

Thus $\text{aff}(C) = A$ and

$$\dim C = \dim \text{aff}(C) = \dim A = k. \qquad \square$$

As an immediate consequence of the above proof we have:

Corollary 6: *Let* C *be any non-empty convex subset of* E^n, *then an affine basis for* aff(C) *may be chosen from among the elements of* C .

Corollary 7: *Let* C *be a convex subset of* E^n *of dimension* k, *then* C *contains a simplex of simpletic dimension* k+1.

3.2 *Topological Properties of Convex Sets*

PROPOSITION 8: *Let* C *be a convex subset of* E^n, *then the closure of* C, $\bar{C}$ *(see p.3), is also a convex set.*

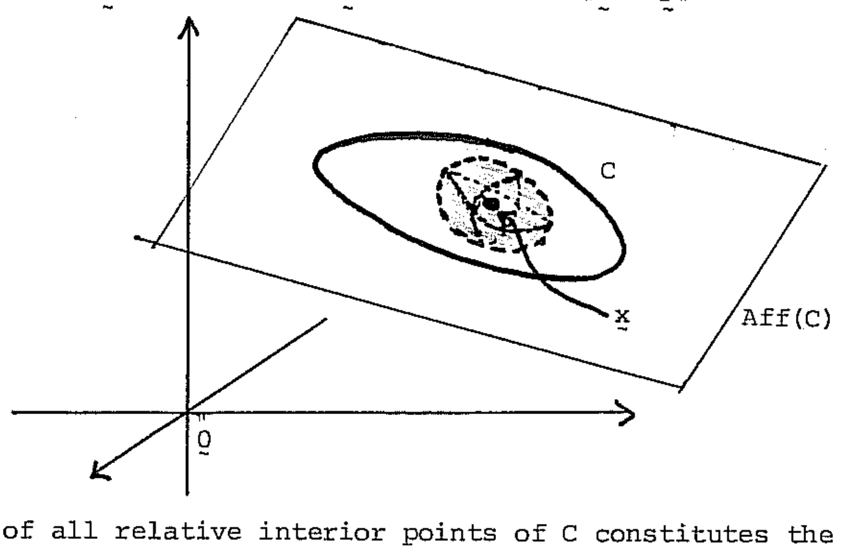
Proof. By the exercise on p.3 we have that $\bar{C} = \bigcap_{\varepsilon>0} [C + B_\varepsilon(\underset{\sim}{0})]$.

By Exercises 5) and 2) on p.20 and 19 respectively, we have that $C + B_\varepsilon(\underset{\sim}{0})$ is convex for each $\varepsilon > 0$. The result now follows from Theorem 1 (p.19). □

DEFINITIONS: Let C be a (convex) subset of E^n, we say $\underset{\sim}{x} \in C$ is a *relative interior point of* C if there exists $\varepsilon > 0$ such that

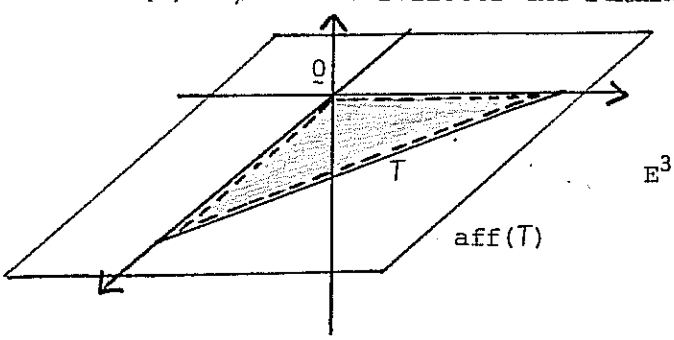
$$B_\varepsilon(\underset{\sim}{x}) \cap \text{Aff}(C) \subseteq C .$$

That is, $\underset{\sim}{x} \in C$ is a relative interior point of C if and only if for some $\varepsilon > 0$ we have $\underset{\sim}{y} \in C$ whenever $\underset{\sim}{y} \in \text{aff}(C)$ and $\|\underset{\sim}{x} - \underset{\sim}{y}\| < \varepsilon$.



The set of all relative interior points of C constitutes the *relative interior* of C, denoted by rel. int.(C) (We say C is *relatively open* if rel. int.(C) = C.)

REMARKS: 1) The definition of relative interior point should be constrasted with the usual topological definition of interior point (see p.3). For example the triangle $T = \text{co}\{(1,0,0), (0,1,0), (0,0,0)\}$ in E^3 has rel. int.$(T) = \{(x,y,z): z = 0, x > 0, y > 0, x+y < 1\}$, on the other hand $\mathrm{int}(T) = \emptyset$. This reflects the remark made on p.11.



2) It would be possible to similarly define the relative closure of C to be

$$\text{rel. cl.}(C) = \{\underset{\sim}{x} \in \text{aff } C: \text{ there exists } (\underset{\sim}{c}_n)_{n=1}^{\infty} \subset C \text{ with } \underset{\sim}{c}_n \to \underset{\sim}{x}\} .$$

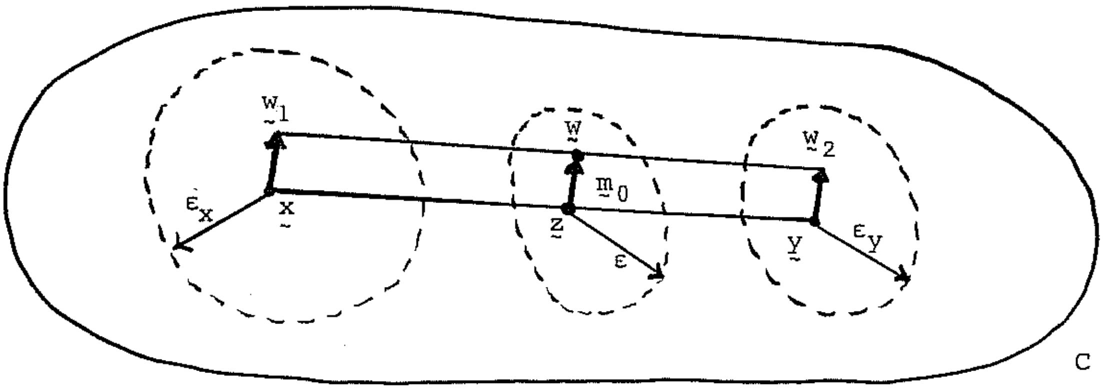
However, since aff(C) is closed and $\bar{C}$ is the smallest closed set containing C we have

$$\text{rel.cl.}(C) = \bar{C} \cap \text{aff}(C) = \bar{C} .$$

So the notion of relative closure is superfluous.

PROPOSITION 9: *Let* C *be a convex subset of* E^n, *then the relative interior of* C, rel. int.(C), *is a convex set.*

Proof. Let $\underset{\sim}{x}, \underset{\sim}{y} \in$ rel. int.(C) and let $\lambda \in (0,1)$ we must show $\underset{\sim}{z} = \lambda\underset{\sim}{x} + (1-\lambda)\underset{\sim}{y} \in$ rel. int.(C). Since $\underset{\sim}{x} \in$ rel. int.(C), there exists $\varepsilon_x > 0$ such that $\underset{\sim}{w} \in C$ whenever $\|\underset{\sim}{x} - \underset{\sim}{w}\| < \varepsilon_x$ and $\underset{\sim}{w} \in$ aff (C). Further, $\underset{\sim}{w} \in$ aff(C) if and only if $\underset{\sim}{m} = \underset{\sim}{w} - \underset{\sim}{x} \in M$ the unique subspace parallel to aff(C). So we have $\underset{\sim}{x} + \underset{\sim}{m} \in C$ whenever $\underset{\sim}{m} \in M$ and $\|\underset{\sim}{m}\| < \varepsilon_x$. Similarly, $\underset{\sim}{y} \in$ rel. int. C implies there exists $\varepsilon_y > 0$ such that $\underset{\sim}{y} + \underset{\sim}{m} \in C$ whenever $\underset{\sim}{m} \in M$ and $\|\underset{\sim}{m}\| < \varepsilon_y$. Let $\varepsilon = \text{Min}\{\varepsilon_x, \varepsilon_y\}$, we show that $B_\varepsilon(\underset{\sim}{z}) \cap \text{aff}(C) \subseteq C$, and so conclude that $\underset{\sim}{z} \in$ rel. int. (C).



Now, if $\underset{\sim}{w} \in B_\varepsilon(z) \cap \text{aff}(C)$ we have $\underset{\sim}{w} = \underset{\sim}{z} + \underset{\sim}{m}_0$ for some $\underset{\sim}{m}_0 \in M$ with $\|\underset{\sim}{m}_0\| < \varepsilon$.

Let $\underset{\sim}{w}_1 = \underset{\sim}{x} + \underset{\sim}{m}_0$ and $\underset{\sim}{w}_2 = \underset{\sim}{y} + \underset{\sim}{m}_0$, then, by above, $\underset{\sim}{w}_1$ and $\underset{\sim}{w}_2 \in C$ and so, by the convexity of C,

$$\lambda \underset{\sim}{w}_1 + (1-\lambda)\underset{\sim}{w}_2 \in C,$$

however

$$\begin{aligned}\lambda \underset{\sim}{w}_1 + (1-\lambda)\underset{\sim}{w}_2 &= \lambda(\underset{\sim}{x} + \underset{\sim}{m}_0) + (1-\lambda)(\underset{\sim}{y} + \underset{\sim}{m}_0) \\ &= \lambda \underset{\sim}{x} + (1-\lambda)\underset{\sim}{y} + \underset{\sim}{m}_0 \\ &= \underset{\sim}{z} + \underset{\sim}{m}_0 \\ &= \underset{\sim}{w}\ .\end{aligned}$$

So $\underset{\sim}{w} \in C$, and the result follows. □

THEOREM 10: *Let* C *be a non-empty convex subset of* E^n, *then* rel. int.(C) $\neq \emptyset$.

At first sight this result may appear somewhat surprising.

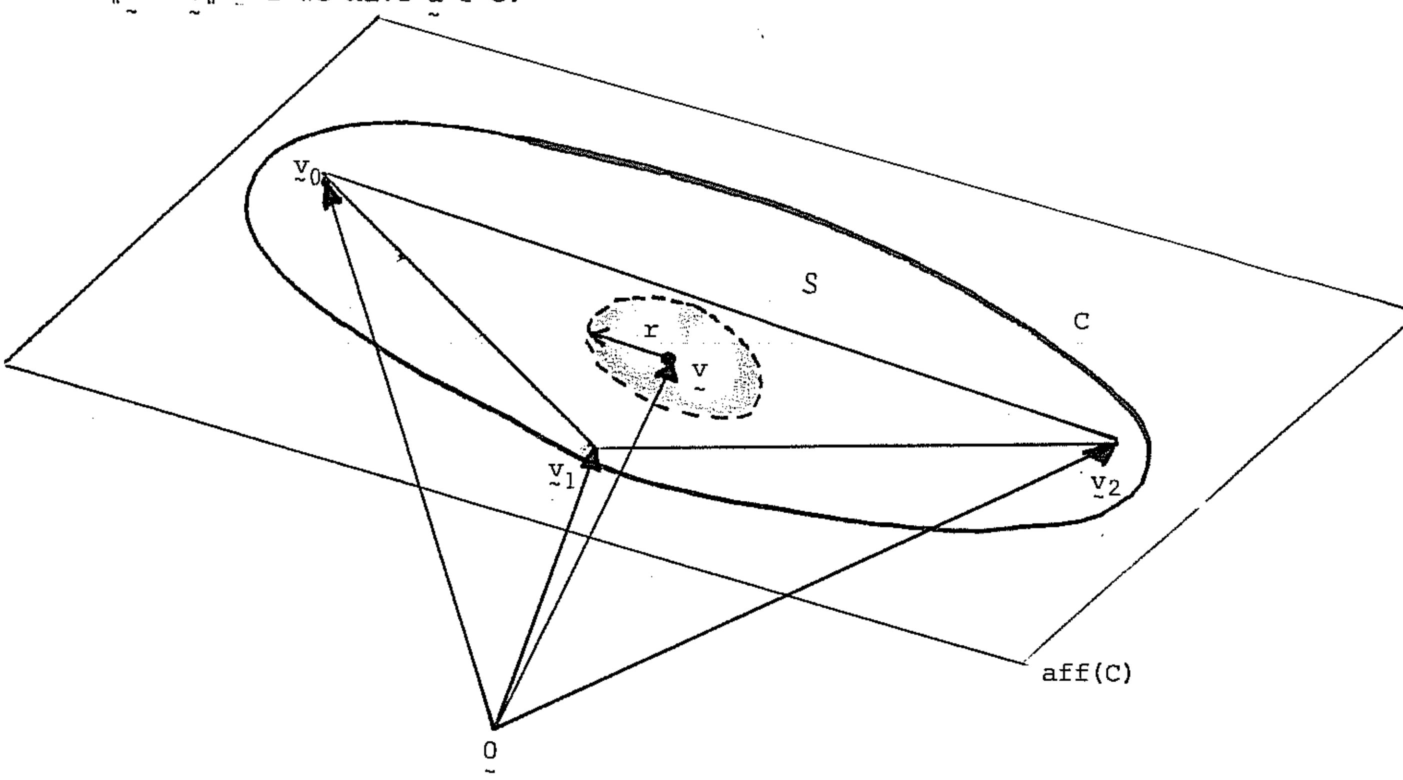
EXERCISE: Find the relative interior for each of the following subsets of E^2.

i) The closed line interval $\{(x,y): y = 0,\ 0 \le x \le 1\}$;

ii) The single point $\{(0,0)\}$.

Proof (of theorem 10). Let dim C = m (= dim aff(C) > 0), then by Corollary 7 on p.26 C contains a simplex (of simplectic dimension m+1), $S = \text{co}\,\{\underset{\sim}{v}_0, \underset{\sim}{v}_1, \ldots, \underset{\sim}{v}_m\} \subseteq C$, where $\underset{\sim}{v}_0, \underset{\sim}{v}_1, \ldots, \underset{\sim}{v}_m$ are affinely independent.

First note, that $\underset{\sim}{v}_0, \underset{\sim}{v}_1, \ldots, \underset{\sim}{v}_m$ is an affinely independent set of m+1 vectors in $C \subseteq \text{aff}(C)$ and so is an affine basis for aff(C). That is any vector $\underset{\sim}{a} \in \text{aff}(C)$ may be written as $\underset{\sim}{a} = \sum_{i=0}^{m} \lambda_i \underset{\sim}{v}_i$ where $\sum_{i=0}^{m} \lambda_i = 1$ and $\lambda_0, \lambda_1, \ldots, \lambda_m$ are the Barrycentric coordinates of $\underset{\sim}{a}$. The proof is completed by showing that $\underset{\sim}{v} = \frac{1}{m+1} \sum_{i=0}^{m} \underset{\sim}{v}_i$ is a point in the relative interior of C (clearly $\underset{\sim}{v} \in C$). That is, we show, there exists $r > 0$ such that whenever $\underset{\sim}{a} \in \text{aff}(C)$ and $\|\underset{\sim}{a} - \underset{\sim}{v}\| \le r$ we have $\underset{\sim}{a} \in C$.



We begin by noting, $\underset{\sim}{a} \in \text{aff}(C)$ if and only if $\underset{\sim}{a} = \sum_{i=0}^{m} \lambda_i \underset{\sim}{v}_i$ for some $\lambda_1, \lambda_2, \ldots, \lambda_m$ with $\sum_{i=0}^{m} \lambda_i = 1$. If, in addition, each $\lambda_i \ge 0$, then the above sum is a convex combination of elements of C and so $\underset{\sim}{a}$ is in C. This, it suffices to show that there exists an $r > 0$ such that whenever

$$\left\| \sum_{i=0}^{m} \lambda_i \underset{\sim}{v}_i - \left(\frac{1}{m+1}\right) \sum_{i=0}^{m} \underset{\sim}{v}_i \right\| < r , \quad \sum_{i=0}^{m} \lambda_i = 1$$

we have $\lambda_i \geq 0$ for each $i = 0, 1, \ldots, m$.

We make use of the fact that the barrycentric coordinates of a point vary continuously with the position of the point*. That is, given $\varepsilon > 0$ there exists, for each $i = 0,1,\ldots, m$ a $\delta_i > 0$ such that $|\lambda_i - \lambda_i'| < \varepsilon$ whenever $\|\underset{\sim}{a} - \underset{\sim}{a}'\| < \delta_i$, where

$$\underset{\sim}{a} = \sum_{i=0}^{m} \lambda_i \underset{\sim}{v}_i , \quad \underset{\sim}{a}' = \sum_{i=0}^{m} \lambda_i' \underset{\sim}{v}_i \quad \text{and} \quad \sum_{i=0}^{m} \lambda_i = \sum_{i=0}^{m} \lambda_i' = 1.$$

Now, consider the particular case when $\underset{\sim}{a}' = \underset{\sim}{v} = \sum_{i=0}^{m} \frac{1}{m+1} \underset{\sim}{v}_i$ and $\varepsilon = \frac{1}{m+1}$. Choose r to be the smallest of the corresponding δ_i, then $r > 0$ (as there are only a finite number of δ_i each of which is strictly positive) and if $\left\| \sum_{i=0}^{m} \lambda_i \underset{\sim}{v}_i - \sum_{i=0}^{m} \frac{1}{m+1} \underset{\sim}{v}_i \right\| < r$ ($< \delta_i$ each i) we have

$$|\lambda_i - \frac{1}{m+1}| < \frac{1}{m+1} \quad \text{so} \quad \lambda_i \geq 0. \quad \square$$

COROLLARY 11: *Let C be any non-empty convex subset of* E^n, *then*

$$\text{aff(rel. int.(C))} = \text{aff(C)} = \text{aff}(\bar{C}).$$

Proof. We have rel. int.(C) $\subseteq C \subseteq \bar{C} \subseteq$ aff(C) as aff(C) is closed. Hence

$$\text{aff rel. int.(C)} \subseteq \text{aff}(\bar{C}) \subseteq \text{aff(C)}.$$

---

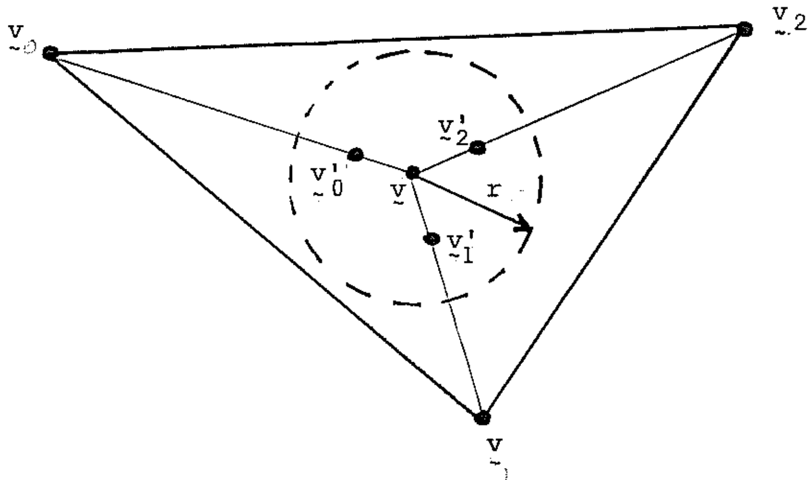
* A proof of this depends on compactness arguments similar to those used in "Approximation Theory". It suffices to show that if $\underset{\sim}{a}_n = \sum_{i=0}^{m} \lambda_i^{(n)} \underset{\sim}{v}_i$ are such that $\|\underset{\sim}{a}_n\| \to 0$ as $n \to \infty$ then $\lambda_i^{(n)} \to 0$ as $n \to \infty$ for $i = 0,1,\ldots,m$. Now by the theorem of Hine-Borel there exists subsequences $\lambda_i^{(n_k)}$ such that $\lambda_i^{(n_k)} \to \mu_i$ for each i as $k \to \infty$. Let $\underset{\sim}{u} = \sum_{i=0}^{m} \mu_i \underset{\sim}{v}_i$, then $\|\underset{\sim}{a}_n - \underset{\sim}{u}\| \leq \sum_{i=0}^{m} |\lambda_i^{(n)} - \mu_i| \|\underset{\sim}{v}_i\| \to 0$ however $\|\underset{\sim}{a}_n - \underset{\sim}{0}\| \to 0$, so by uniqueness of limits $\underset{\sim}{u} = \underset{\sim}{0}$ and then by the uniqueness of barrycentric coordinates we have $\mu_i = 0$ or $\lambda_i^{(n_k)} \to 0$.

It thus suffices to show $\text{aff}(C) = \text{aff}(\text{rel. int.}(C))$. To do this, note that

$$\text{aff}(C) = \text{aff}\{\underset{\sim}{v}_0, \underset{\sim}{v}_1, \ldots, \underset{\sim}{v}_m\}$$

$$= \text{aff}\{\underset{\sim}{v}'_0, \underset{\sim}{v}'_1, \ldots, \underset{\sim}{v}'_m\}$$

where $\underset{\sim}{v}'_i = \underset{\sim}{v} + \frac{r}{2\|\underset{\sim}{v} - \underset{\sim}{v}_i\|}(\underset{\sim}{v} - \underset{\sim}{v}_i)$ (see diagram below), and by the construction given in the above proof $\underset{\sim}{v}'_i \in \text{rel. int.}(C)$.



□

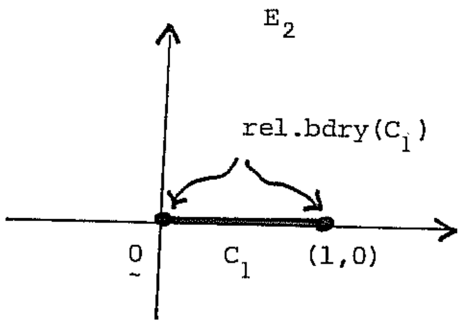
DEFINITION: Let C be a (convex) subset of E^n, by the <u>relative boundary</u> of C we mean

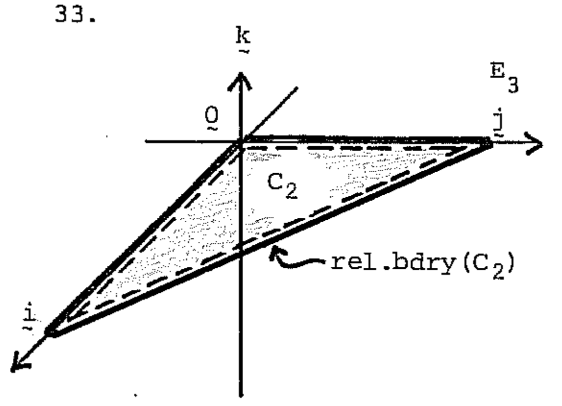
$$\text{rel. bdry}(C) = \bar{C}\backslash\text{rel. int.}(C).$$

Note: By the previous theorem $\text{rel. bdry}(C) \subsetneq \bar{C}$.

For EXAMPLE: If C_1 is the closed interval $[(0,0), (1,0)]$ in E^2 then $\text{rel. bdry}(C_1) = \{(0,0), (1,0)\}$; if C_2 is the "open" triangle $\{(x,y): z = 0, x > 0, y > 0 \text{ and } x + y < 1\}$ in E^3 then

$$\text{rel. bdry}(C_2) = [\underset{\sim}{0},\underset{\sim}{i}] \cup [\underset{\sim}{0},\underset{\sim}{j}] \cup [\underset{\sim}{i},\underset{\sim}{j}]$$





(Note: Using the usual topological definition of boundary, $\bar{C}\setminus\text{int.}(C)$, we would have bdry $C_1 = C_1$ and bdry $C_2 = \bar{C}_2$ the whole "solid" triangle.)

THEOREM 11: *Let* C *be a closed bounded convex subset of* E^n *containing at least two distinct points, then*

$$C = \text{co}(\text{rel. bdry}(C)).$$

Proof. Let $\underset{\sim}{x} \in C$ and choose any other point $\underset{\sim}{y} \in C$ (that is, $\underset{\sim}{y} \in C$ and $\underset{\sim}{y} \neq \underset{\sim}{x}$.)

Let $I = \mathcal{L}(\underset{\sim}{x},\underset{\sim}{y}) \cap C$, then I is a closed (as both C and $\mathcal{L}(\underset{\sim}{x},\underset{\sim}{y})$ are closed - see problem on p.14), convex (by theorem 1, as both $\mathcal{L}(\underset{\sim}{x},\underset{\sim}{y})$ and C are convex), bounded (as C is bounded) subset of E^n.

Further, from the definition of $\mathcal{L}(\underset{\sim}{x},\underset{\sim}{y})$ - see p.7 - any point in I has the form $\lambda\underset{\sim}{x} + (1-\lambda)\underset{\sim}{y}$ for some $\lambda \in \mathbb{R}$.

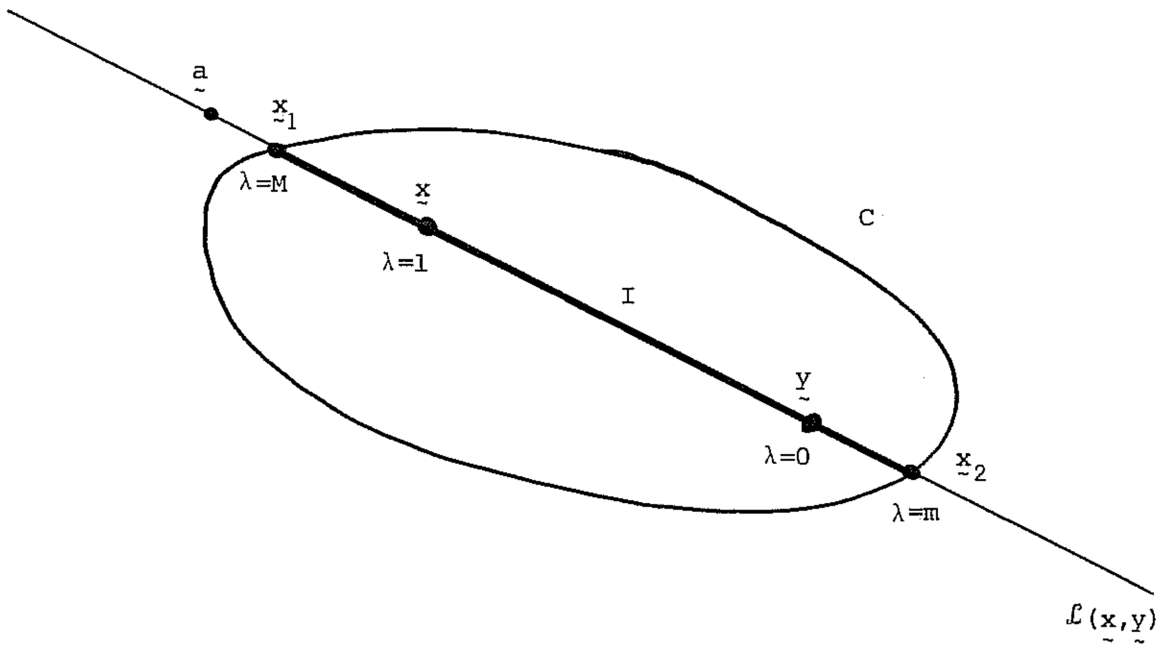
Let $\Lambda = \{\lambda \in \mathbb{R}: \lambda\underset{\sim}{x} + (1-\lambda)\underset{\sim}{y} \in I\}$, since I is bounded and $\|\lambda\underset{\sim}{x} + (1-\lambda)\underset{\sim}{y}\| \to \infty$ as $|\lambda| \to \infty$ (prove this) we have that both

$$M = \text{Sup}\,\Lambda \text{ and } m = \inf\,\Lambda$$

exist and are finite.

Further, since $\underset{\sim}{x}$ $(\lambda = 1)$ and $\underset{\sim}{y}$ $(\lambda = 0)$ are in I we have $M \geq 1$ and $m \leq 0$. Let

$$\underset{\sim}{x}_1 = M\,\underset{\sim}{x} + (1-M)\underset{\sim}{y} \quad \text{and} \quad \underset{\sim}{x}_2 = m\,\underset{\sim}{x} + (1-m)\underset{\sim}{y} \qquad \ldots(*)$$



Then, since I is closed we have x_1, $x_2 \in I$ (give details, hint: there exists a sequence $\lambda_1, \lambda_2, \ldots, \lambda_n, \ldots$ such that $\lambda_n \to M$ and $\lambda_n x + (1-\lambda_n) y \in I$).

Also, from (*) we have

$$(1-m)x_1 = (1-m)M\,x + (1-m)(1-M)y$$

and

$$(1-M)x_2 = (1-M)m\,x + (1-m)(1-M)y \quad ,$$

subtracting and re-arranging gives

$$x = \frac{(1-m)}{M-m}x_1 + \frac{(M-1)}{M-m}x_2 \; .$$

Thus, x is a convex combination of x_1 and x_2 ($M > m$ and $m \leq 0$ so $\frac{1-m}{M-m} > 0$, similarly $M > 1$ so $\frac{M-1}{M-m} > 0$ and $\frac{(1-m) + (M-1)}{M - m} = 1$).

It now suffices to show x_1, $x_2 \in$ rel. bdry C. Assume this were not the case, then x_1 (or x_2) must be in rel. int.(C). Assume $x_1 \in$ rel. int.(C), that is, there exists an $r > 0$ such that $a \in$ aff(C) and $\|a - x_1\| < r$ implies $a \in C$.

Let $\underset{\sim}{a} = (M+\varepsilon)\underset{\sim}{x} + (1-M-\varepsilon)\underset{\sim}{y}$, where $\varepsilon = \dfrac{r}{2\|\underset{\sim}{x} - \underset{\sim}{y}\|} > 0$. Then,

$\underset{\sim}{a} \in \mathcal{L}(\underset{\sim}{x},\underset{\sim}{y}) \subset \text{aff}(C)$ and

$$\|\underset{\sim}{a} - \underset{\sim}{x}_1\| = \|(M+\varepsilon)\underset{\sim}{x} + (1-M-\varepsilon)\underset{\sim}{y} - M\underset{\sim}{x} - (1-M)\underset{\sim}{y}\|$$

$$= \varepsilon\|\underset{\sim}{x} - \underset{\sim}{y}\| < r.$$

So $\underset{\sim}{a} \in C$. Hence $\underset{\sim}{a} \in I$ and $\text{Sup}\,\Lambda \geq M + \varepsilon > M$, contradicting the definition of M. Therefore $\underset{\sim}{x}_1 \in$ rel. bdry C. A similar argument shows that $\underset{\sim}{x}_2 \notin$ rel. int.(C) and the result is established.

□

CHAPTER 2 - SEPARATION THEOREMS

Let A, B be two subsets of E^n, we say the hyperplane

$$H = \{\underset{\sim}{x} \in E^n : (\underset{\sim}{x}, \underset{\sim}{x}_0) = c\} :$$

<u>separates</u> A and B if $A \subseteq H^+$ and $B \subseteq H^-$, that is

$$(\underset{\sim}{a}, \underset{\sim}{x}_0) \geq c \quad \text{for all } \underset{\sim}{a} \in A$$

and

$$(\underset{\sim}{b}, \underset{\sim}{x}_0) \leq c \quad \text{for all } \underset{\sim}{b} \in B;$$

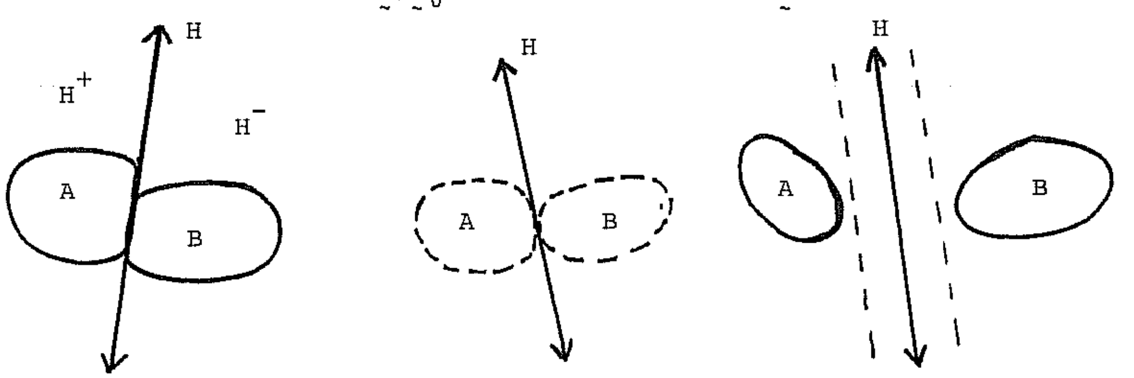
<u>strictly separates</u> A and B if $(\underset{\sim}{a}, \underset{\sim}{x}_0) > c$ for all $\underset{\sim}{a} \in A$
and
$(\underset{\sim}{b}, \underset{\sim}{x}_0) < c$ for all $\underset{\sim}{b} \in B$;

<u>strongly separates</u> A and B if there exists an $\varepsilon > 0$ such that

$$(\underset{\sim}{a}, \underset{\sim}{x}_0) > c + \varepsilon \quad \text{for all } \underset{\sim}{a} \in A$$

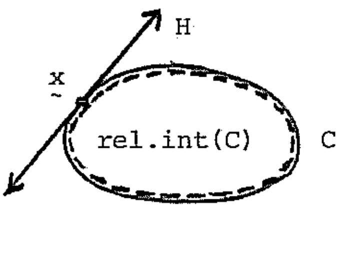
and

$$(\underset{\sim}{b}, \underset{\sim}{x}_0) < c - \varepsilon \quad \text{for all } \underset{\sim}{b} \in B.$$

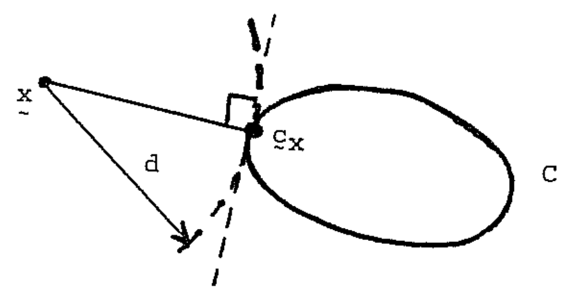


separation strict separation strong separation

Let C be a convex subset of E^n. We say $\underset{\sim}{x} \in C$ is a <u>support point</u> of C or C *is supported by* H *at* $\underset{\sim}{x}$, if there exists a hyperplane H such that $\underset{\sim}{x} \in H$ and $H \cap \text{rel. int.}(C) = \emptyset$. - *intuitively H is a "tangent plane" to C at* $\underset{\sim}{x}$.



LEMMA 1: *Let* C *be a non-empty closed convex subset of* E^n *and let* $\underset{\sim}{x}$ *be any point of* E^n *not in* C, *then there exists a unique point* $\underset{\sim}{c}_x$ *of* C *which is closest to* $\underset{\sim}{x}$. That is, $\|\underset{\sim}{c}_x - \underset{\sim}{x}\| < \|\underset{\sim}{c} - \underset{\sim}{x}\|$ for all $\underset{\sim}{c} \in C$, $\underset{\sim}{c} \neq \underset{\sim}{c}_x$.



Proof. Let $d = \inf\{\|\underset{\sim}{c} - \underset{\sim}{x}\| : \underset{\sim}{c} \in C\}$, then there exists a sequence $(\underset{\sim}{c}_n) \subset C$ with $\|\underset{\sim}{c}_n - \underset{\sim}{x}\| \to d$ as $n \to \infty$. Now, since C is convex $\frac{\underset{\sim}{c}_n + \underset{\sim}{c}_m}{2} \in C$ and so $d \leq \left\|\frac{\underset{\sim}{c}_n + \underset{\sim}{c}_m}{2} - \underset{\sim}{x}\right\|$ or $2d \leq \left\|(\underset{\sim}{c}_n - \underset{\sim}{x}) + (\underset{\sim}{c}_m - \underset{\sim}{x})\right\|$.

We therefore have

$$0 \leq \|\underset{\sim}{c}_n - \underset{\sim}{c}_m\|^2 = \|(\underset{\sim}{c}_n - \underset{\sim}{x}) - (\underset{\sim}{c}_m - \underset{\sim}{x})\|^2$$

$$= -\|(\underset{\sim}{c}_n - \underset{\sim}{x}) + (\underset{\sim}{c}_m - \underset{\sim}{x})\|^2 + 2\|\underset{\sim}{c}_n - \underset{\sim}{x}\|^2 + 2\|\underset{\sim}{c}_m - \underset{\sim}{x}\|^2$$ - by the parallelogram law (p.3)

$$\leq -(2d)^2 + 2\|\underset{\sim}{c}_n - \underset{\sim}{x}\|^2 + 2\|\underset{\sim}{c}_m - \underset{\sim}{x}\|^2$$

and so taking the limit as $n, m \to \infty$ we have

$$0 \leq \|\underset{\sim}{c}_n - \underset{\sim}{c}_m\|^2 \leq -(2d)^2 + 2d^2 + 2d^2 = 0.$$

That is, $\|\underset{\sim}{c}_n - \underset{\sim}{c}_m\| \to 0$ as $m, n \to \infty$, or $(\underset{\sim}{c}_n)$ is a Cauchy sequence. Hence, since C is closed and E^n is complete, $(\underset{\sim}{c}_n)$ converges to some point, $\underset{\sim}{c}_x$ say, of C. Further, $\|\underset{\sim}{c}_x - \underset{\sim}{x}\| = \lim_{n\to\infty} \|\underset{\sim}{c}_n - \underset{\sim}{x}\| = d$, so $\underset{\sim}{c}_x$ is a closest point of C to $\underset{\sim}{x}$.

Now assume $\underset{\sim}{c}'$ is another closest point of C to $\underset{\sim}{x}$, that is $\|\underset{\sim}{c}' - \underset{\sim}{x}\| = d$, then since $\frac{\underset{\sim}{c}_x + \underset{\sim}{c}'}{2} \in C$, as C is convex, we have

$$d \leq \left\|\frac{\underset{\sim}{c}_x + \underset{\sim}{c}'}{2} - \underset{\sim}{x}\right\|$$

$$\leq \left\|\frac{\underset{\sim}{c}_x - \underset{\sim}{x}}{2}\right\| + \left\|\frac{\underset{\sim}{c}' - \underset{\sim}{x}}{2}\right\| , \quad \text{by the triangle inequality.}$$

$$= \frac{d}{2} + \frac{d}{2} = d.$$

That is, $\|\underset{\sim}{c}_x + \underset{\sim}{c}' - 2\underset{\sim}{x}\| = 2d$ and so, using the parallelogram law

$$\|\underset{\sim}{c}_x - \underset{\sim}{c}'\|^2 = \|(\underset{\sim}{c}_x - \underset{\sim}{x}) - (\underset{\sim}{c}' - \underset{\sim}{x})\|^2$$

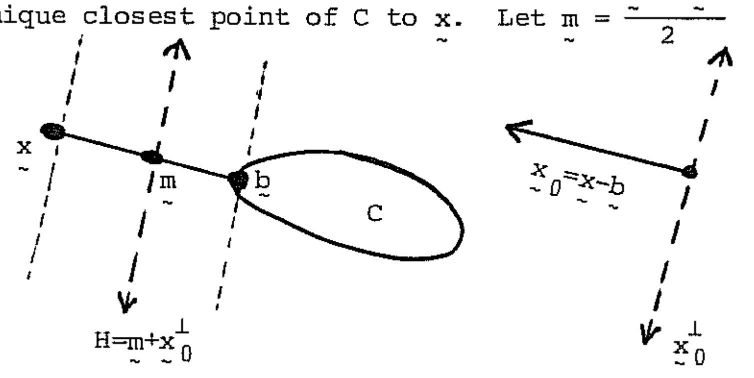
$$= -\|(\underset{\sim}{c}_x - \underset{\sim}{x}) + (\underset{\sim}{c}' - \underset{\sim}{x})\|^2 + 2\|\underset{\sim}{c}_x - \underset{\sim}{x}\|^2 + 2\|\underset{\sim}{c}' - \underset{\sim}{x}\|^2$$

$$= -(2d)^2 + 2d^2 + 2d^2 = 0,$$

or $\|\underset{\sim}{c}_x - \underset{\sim}{c}'\| = 0$ and so $\underset{\sim}{c}' = \underset{\sim}{c}_x$. □

THEOREM 2: *Let* C *be a non-empty closed convex subset of* E^n *and let* $\underset{\sim}{x} \in E^n$ *with* $\underset{\sim}{x} \notin C$ *then there exists a hyperplane which strongly separates* $\{\underset{\sim}{x}\}$ *from* C.

Proof. Let $\underset{\sim}{b}$ be the unique closest point of C to $\underset{\sim}{x}$. Let $\underset{\sim}{m} = \frac{\underset{\sim}{x} + \underset{\sim}{b}}{2}$ and $\underset{\sim}{x}_0 = \underset{\sim}{x} - \underset{\sim}{b}$ ($\neq 0$).



Take $H = \underset{\sim}{m} + \underset{\sim}{x}_0^{\perp}$,

that is

$H = \{\underset{\sim}{y} \in E^n : (\underset{\sim}{y}, \underset{\sim}{x}_0) = c\}$

where $c = (\underset{\sim}{m}, \underset{\sim}{x}_0)$.

First note that $\underset{\sim}{x} = \underset{\sim}{m} + \underset{\sim}{x}_0/2$ and $\underset{\sim}{b} = \underset{\sim}{m} - \underset{\sim}{x}_0/2$. So $(\underset{\sim}{x}, \underset{\sim}{x}_0) = c + \|x_0\|^2/2$ and $(\underset{\sim}{b}, \underset{\sim}{x}_0) = c - \|x_0\|^2/2$. The proof that H strongly separates $\{\underset{\sim}{x}\}$

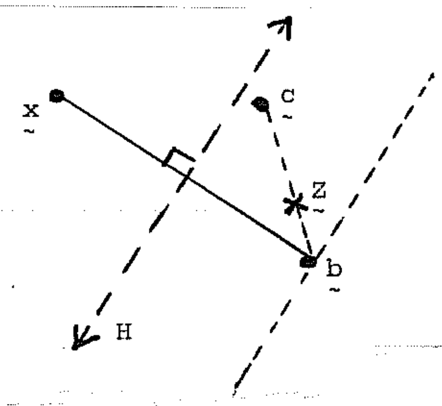
from C is completed by showing $(\underset{\sim}{c},\underset{\sim}{x}_0) \leq (\underset{\sim}{b},\underset{\sim}{x}_0)$ for any $\underset{\sim}{c} \in C$, for then we have $(\underset{\sim}{x},\underset{\sim}{x}_0) \geq c + \varepsilon$ while $(\underset{\sim}{c},\underset{\sim}{x}_0) < c - \varepsilon$ for all $\underset{\sim}{c} \in C$ where $\varepsilon = \|x_0\|^2/4$.

Now assume there exists $\underset{\sim}{c} \in C$ with $(\underset{\sim}{c},\underset{\sim}{x}_0) > (\underset{\sim}{b},\underset{\sim}{x}_0)$, then, since C is convex $\underset{\sim}{Z} = \lambda\underset{\sim}{c} + (1-\lambda)\underset{\sim}{b} \in C$ for all $\lambda \in [0,1]$. Further,

$$
\begin{aligned}
\|\underset{\sim}{Z} - \underset{\sim}{x}\|^2 &= (\lambda\underset{\sim}{c} + (1-\lambda)\underset{\sim}{b} - \underset{\sim}{x}, \lambda\underset{\sim}{c} + (1-\lambda)\underset{\sim}{b} - \underset{\sim}{x}) \\
&= (\underset{\sim}{b} - \underset{\sim}{x} + \lambda(\underset{\sim}{c} - \underset{\sim}{b}), \underset{\sim}{b} - \underset{\sim}{x} + \lambda(\underset{\sim}{c} - \underset{\sim}{b})) \\
&= \|\underset{\sim}{b} - \underset{\sim}{x}\|^2 + 2\lambda(\underset{\sim}{c} - \underset{\sim}{b}, \underset{\sim}{b} - \underset{\sim}{x}) + \lambda^2\|\underset{\sim}{c} - \underset{\sim}{b}\|^2 .
\end{aligned}
$$

Now $\underset{\sim}{x}_0 = \underset{\sim}{x} - \underset{\sim}{b}$ so $(\underset{\sim}{c},\underset{\sim}{x}_0) > (\underset{\sim}{b},\underset{\sim}{x}_0)$ gives $(\underset{\sim}{c},\underset{\sim}{x}-\underset{\sim}{b}) > (\underset{\sim}{b},\underset{\sim}{x}-\underset{\sim}{b})$ and so $(\underset{\sim}{c} - \underset{\sim}{b}, \underset{\sim}{x} - \underset{\sim}{b}) > 0$

or $\quad (\underset{\sim}{c} - \underset{\sim}{b}, \underset{\sim}{b} - \underset{\sim}{x}) < 0.$



We therefore have

$$\|\underset{\sim}{Z} - \underset{\sim}{x}\|^2 = \|\underset{\sim}{b} - \underset{\sim}{x}\|^2 - 2\lambda p + \lambda^2 q$$

where $p = -(\underset{\sim}{c} - \underset{\sim}{b}, \underset{\sim}{b} - \underset{\sim}{x}) > 0$

and $\quad q = \|c - b\|^2 > 0.$

So, for sufficiently small positive λ, indeed for any λ with $0 < \lambda < \frac{2p}{q}$ we have

$$\|\underset{\sim}{Z} - \underset{\sim}{x}\| < \|\underset{\sim}{b} - \underset{\sim}{x}\| .$$

Choosing λ to be less than 1 we also have $\underset{\sim}{Z} \in C$, contradicting the fact that $\underset{\sim}{b}$ is the closest point of C to $\underset{\sim}{x}$. We therefore conclude that such a $\underset{\sim}{c}$ cannot exist and so $(\underset{\sim}{c},\underset{\sim}{x}_0) \leq (\underset{\sim}{b},\underset{\sim}{x}_0)$ for all $\underset{\sim}{c} \in C$.

$\square$