Topological vector spaces

Harald Hanche-Olsen
hanche@math.ntnu.no

Topological vector spaces

A *topological vector space* is a (real or complex) vector space equipped with a Hausdorff topology, so that the vector space operations are continuous. That is, if the space is called X then vector addition is a continuous map $X \times X \to X$, and multiplication by scalars is a continuous map $\mathbb{C} \times X \to X$ (or $\mathbb{R} \times X \to X$).

One simple and common way to define a topological vector space, is by starting with a vector space X and a family $\mathscr{P}$ of *seminorms* on X, separating the points of X. (For details see the previous note on topology.) The topology generated by $\mathscr{P}$ does in fact make X a topological vector space. (*Exercise*: Prove this.)

It is useful to note that the topology of a topological vector space is completely specified by the neighbourhoods of 0. For addition by any constant x is continuous, so the neighbourhoods of x are the sets of the form $x + U$ where U is a neighbourhood of 0.

A topological vector space is called *locally convex* if every neighourhood of 0 contains a convex neighbourhood of 0. If the topology is generated by a family of seminorms, the space is locally convex, since every neighbourhood of 0 contains a neighbourhood of the form

$$\{x \in X \colon p_j(x) < \varepsilon,\ j = 1, \dots, n\} \qquad (p_1, \dots, p_n \in \mathscr{P}),$$

which is clearly convex.

Examples include the *weak* topology on a normed space X, and the weak* topology on its dual.

In fact, the topology of any locally convex space is generated by seminorms. For any neighbourhood of 0 contains a neighbourhood U that is not only convex but *balanced*, in the sense that $\alpha x \in U$ whenever $x \in U$ and α is a scalar with $|\alpha| = 1$. For such a neighbourhood U we can create a seminorm p by

$$p(x) = \big(\inf\{t > 0 \colon x/t \in U\}\big),$$

and the set of all such seminorms generate the topology. (p is a *semi*norm because the supremum may in fact be infinite, so that $p(x) = 0$ for some x.) Since we shall not need this result, we omit the proof. (But proving it could be a useful exercise.)

We already know that bounded linear functionals on a Banach space are continuous. Below is the corresponding result for topological vector spaces.

1 Proposition. *For a linear functional f on a topological vector space X, the following are equivalent:*

1. *f is continous,*
2. *f is bounded on some neighbourhood of 0,*
3. *$\operatorname{Re} f$ is bounded below or above on some nonempty open set,*

Proof: Clearly, if f is continuous then f is bounded on the neighbourhood of 0 $\{x \in X \colon |f(x)| < 1\}$, which implies the second condition.

If f is bounded on the neighbourhood U of 0, say $|f| < M$ on U, then $|f(x)| < \varepsilon$ if $x \in (\varepsilon/M)U$. But $(\varepsilon/M)U$ is also a neighbourhood of 0, so f is continuous.

Obviously, the second condition implies the third.

Next, if $\operatorname{Re} f$ is bounded below on some nonempty open set, say $\operatorname{Re} f > a$ on U, let $u \in U$. Then $\operatorname{Re} f$ is bounded below on the neighbourhood $U - u$ of 0, and above on the neighbourhood $u - U$ of 0. Hence $\operatorname{Re} f$ is bounded on a neighbourhood $(U - u) \cap (u - U)$ of 0. So if X is a real space, f is bounded on this neighbourhood. If X is a complex space, it follows that f is bounded on the same neighbourhood, since $\operatorname{Im} f(x) = -\operatorname{Re} f(ix)$. ∎

For a linear functional f on X, we define its *kernel* of f to be

$$\ker f = \{x \in X \colon f(x) = 0\}.$$

2 Proposition. *A linear functional on a locally convex topological vector space is continuous if and only if its kernel is closed.*

Proof: If f is continuous, then $\ker f = f^{-1}(\{0\})$ is closed, since $\{0\}$ is closed.

Next, assume that $\ker f$ is closed. If $f = 0$, then f is continuous. Assume therefore $f \neq 0$. Pick $w \in X$ with $f(w) = 1$. Since $\ker f$ is closed, f is nonzero on a neighbourhood W of w. Since X is locally convex, W can be taken to be convex. If X is a real space, we must have $f|_W > 0$, so f is bounded below on W, and f must be continuous.

If X is a complex space, we note that $\{f(x) \colon x \in W\}$ is a convex subset of $\mathbb{C}$ which does not contain 0. Then that set lies in some halfplane, meaning that there is some real number θ with $\operatorname{Re}(e^{i\theta} f(x)) \ge 0$ for all $x \in W$. But then $\operatorname{Re}(e^{i\theta} f)$ is bounded below on W, so $e^{i\theta} f$ is continuous. But then so is f. ∎

3 Proposition. *Let X be a normed vector space. A linear functional on X^* is of the form $f \mapsto f(x)$ where $x \in X$ if and only if it is weakly* continuous.*

Proof: The functionals $f \mapsto f(x)$ are weakly* continuous by the construction of the weak* topology on X^*.

So we only need to consider a weakly* continuous functional ξ on X^*, and must show it can be written $f \mapsto f(x)$.

By the construction of the weak* topology, there are $x_1, \ldots, x_n \in X$ and $\delta > 0$ with $|\xi(f)| < 1$ whenever $|f(x_k)| < \delta$ for $k = 1, \ldots, n$. In particular,

$$\xi(f) = 0 \text{ whenever } \big(f(x_1), \ldots, f(x_n)\big) = (0, \ldots, 0)$$

(for then $|\xi(tf)| < 1$ for all t).

It follows that we can define a linear functional ζ on $\mathbb{R}^n$ by setting

$$\zeta\big(f(x_1), \ldots, f(x_n)\big) = \xi(f), \qquad f \in X^*.$$

This is well defined, for if $\big(f(x_1), \ldots, f(x_n)\big) = \big(g(x_1), \ldots, g(x_n)\big)$ then $\xi(f - g) = 0$ by the above. Strictly speaking, this may define ζ only on a subspace of $\mathbb{R}^n$, but the definition can be expanded to all of $\mathbb{R}^n$. Write $\zeta(y) = c_1 y_1 + \cdots + c_n y_n$. Then

$$\xi(f) = c_1 f(x_1) + \cdots + c_n f(x_n) = f(x), \qquad \text{where } x = c_1 + \cdots + c_n.$$

This completes the proof. ∎

The weak topology on a Banach space truly is weaker than the norm topology, at least if the space is infinite-dimensional. For any weak neighbourhood of 0 will contain a set of the form $\{x : |f_k(x)| < \varepsilon,\ k = 1, \ldots, n\}$ where $f_k \in X^*$. In particular, it contains $\{x : |f_k(x)| = 0,\ k = 1, \ldots, n\}$, which is an infinite-dimensional space.

A word on notation: It is common to write $\rightharpoonup$ for weak convergence and $\overset{*}{\rightharpoonup}$ for weak* convergence. In the case of *reflexive* Banach spaces, such as Hilbert spaces and L^p spaces for $1 < p < \infty$, there is of course no essential difference, and the notation $\rightharpoonup$ is generally preferred.

In a weaker topology, convergence is easier to achieve because there are fewer neighbourhoods. Here is a simple example. Let $1 < p \le \infty$ and let $e_k \in \ell^p$ be the sequence which has 1 at position k and 0 at all other positions. Then, because $\|e_j - e_k\|_p = 2^{1/p}$ when $j \ne k$, the sequence $(e_k)_{k=1}^\infty$ is not Cauchy, and therefore not convergent in norm. But still, $e_k \rightharpoonup 0$ (weakly) as $k \to \infty$, because whenever $x \in \ell^q$ with $1 \le q < \infty$ then $x_k \to 0$.

In some important cases, however, weak convergence with an added condition does imply norm convergence. Recall that L^p spaces are uniformly convex for $1 < p < \infty$. (We have proved this only for $2 \le p < \infty$, but nevertheless the more general result holds.) In the existence theory of partial differential equations, one commonly proves weak convergence of approximate solutions in some L^p space, and one then needs norm convergence to complete the existence proof. The following result is useful in such situations.

4 Proposition. *Let X be a uniformly convex normed space, and let $(x_k)_{k=1}^\infty$ be a weakly convergent sequence in X with weak limit x. Then $x_k \to x$ (in norm) if and only if $\|x_k\| \to \|x\|$.*

Proof: If $x_k \to x$ in norm then $\|x_k\| \to \|x\|$. We need to prove the converse. If $x = 0$, the converse is nothing but the definition of norm convergence. So we may assume that $x \ne 0$, and (dividing everything by $\|x\|$) we may as well assume that $\|x\| = 1$ and $\|x_k\| \to 1$. We may even replace x_k by $x_k / \|x_k\|$.

So now our assumptions are $\|x_k\| = \|x\| = 1$ and $x_k \rightharpoonup x$, and we need to prove that $\|x - x_k\| \to 0$.

By (a corollary to) the Hahn–Banach theorem there is a linear functional $f \in X^*$ with $\|f\| = f(x) = 1$. So $f(x_k) \to 1$. Let $\varepsilon > 0$, and pick $\delta > 0$, thanks to the uniform convexity of X, so that $\|u - v\| < \varepsilon$ whenever $\|u\| = \|v\| = 1$ and $\|u + v\| < 2 - \delta$. Pick a number N so that $\operatorname{Re} f(x_k) > 1 - \delta$ whenever $k \ge N$. For such a k, then, $\operatorname{Re} f(x + x_k) = f(x) + \operatorname{Re} f(x_k) > 2 - \delta$, so $\|x + x_k\| > 2 - \delta$. Thus $\|x - x_k\| < \varepsilon$ for such k. ∎

The Banach–Alaoglu theorem

The following theorem is also known by Alaouglu's name alone.[1]

5 Theorem. (Banach–Alaoglu) *Let X be a normed space. The closed unit ball in the dual space X^* is compact in the weak* topology.*

Proof: Basically, we can identify the closed unit ball $\overline{B}$ of X^* with a closed subset of the space

$$\Xi = \prod_{x \in X} \big\{ z \in \mathbb{C} : |z| \le \|x\| \big\},$$

with the product topology. More precisely, if $f \in X^*$ and $\|f\| \le 1$, we write f_x in place of $f(x)$, and so $f = (f_x)_{x \in X}$ *is* the wanted element of Ξ. A general element $f \in \Xi$ belongs to $\overline{B}$ if and only if it is linear, i.e.,

$$f(x + y) = f(x) + f(y), \quad f(\alpha x) = \alpha f(x), \qquad x, y \in X, \alpha \in \mathbb{C}.$$

(The bound $\|f\| \le 1$ is already built into Ξ.) But for given $x, y \in X$ and $\alpha \in \mathbb{C}$, the quantities $f(x + y)$, $f(x)$, $f(y)$, $f(\alpha x)$ are continuous functions of f with respect to

[1] This theorem was also featured in the Norwegian edition of Donald Duck magazine quite a long time ago (seventies?). In the story, Gyro Gearloose (Petter Smart in Norwegian) has created a machine that makes its users very smart. He tries it on Donald's nephews, and the first thing they say when they emerge from the machine is the statement of Alaouglu's theorem. Well, almost. The exact text (as far as I remember) was: "Enhetsballen er kompakt i den svarte stjernetypologien." Close enough. Apparently, in the American edition, they said something of a rather more trivial nature.

the product topology (in an earlier notation, they are the projections $\pi_{x+y}(f)$, etc.), which proves that X^* is indeed closed in Ξ. Also, the weak* topology on X^* is just the topology which X^* inherits as a subspace of Ξ. Thus X^* is a closed subspace of a compact space, and therefore itself compact. ∎

This is well and good, but compactness only guarantees the existence of convergent *filters*. We wish to have convergent *sequences*. The following lemma will help.

6 Lemma. *Assume that X is a separable normed space. Then the weak* topology on the closed unit ball of X^* is metrizable.*

Proof: Let $\{x_k\colon k = 1,2,\ldots\}$ be a dense subset of X. We define a metric d on X^* by

$$d(f,g) = \sum_{k=1}^{\infty} 2^{-k} \wedge |f(x_k) - g(x_k)|.$$

(Here $a \wedge b$ is short for the minimum of a and b.) That d satisfies the axioms for a metric is obvious, with the possible exception of the requirement that $d(f,g) > 0$ when $f \neq g$. But if $f \neq g$ then there is some $x \in X$ with $f(x) \neq g(x)$. If k is chosen so that $\|x_k - x\|$ is small enough, we should get $f(x_k) \neq g(x_k)$, so that one of the terms in the sum defining $d(f,g)$ is nonzero. In fact

$$\begin{aligned}\bigl|f(x_k) - g(x_k) - \bigl(f(x) - g(x)\bigr)\bigr| &= \bigl|(f-g)(x_k - x)\bigr| \\ &\le \|f-g\|\,\|x - x_k\| < \bigl|\bigl(f(x) - g(x)\bigr)\bigr|\end{aligned}$$

if $\|x - x_k\| \le \bigl|\bigl(f(x) - g(x)\bigr)\bigr| / \|f - g\|$, and then $f(x_k) - g(x_k)$ must be nonzero.

We might hope that d generates the weak* topology on X^*, but that turns out not to be so. But it does generate the relative topology inherited to the closed unit ball B of X^*, as we now show.

First consider a d-neighbourhood of $g \in B$. It will contain an ε-ball, that is, $\{f \in B\colon d(f,g) < \varepsilon\}$. Pick n so that $\sum_{k=n+1}^{\infty} 2^{-k} < \frac{1}{2}\varepsilon$. Then the given d-neighbourhood contains the weak*-neighbourhood

$$\Bigl\{f \in B\colon |f(x_k) - g(x_k)| < \frac{\varepsilon}{2n} \text{ for } k = 1,\ldots,n\Bigr\}.$$

On the other hand, consider a weak*-neighbourhood of $g \in B$. It contains a set of the form

$$V = \bigl\{f \in B\colon |f(z_j) - g(z_j)| < \varepsilon \text{ for } j = 1,\ldots,m\bigr\},$$

where $z_1,\ldots,z_m \in X$. Now pick some n so that, for each $j = 1,\ldots,n$, there is some $k \le n$ with $\|x_k - z_j\| < \frac{1}{4}\varepsilon$. Let $\delta = 2^{-n} \wedge \frac{1}{2}\varepsilon$. We claim that V contains the d-neighbourhood $\{f \in B\colon d(f,g) < \delta\}$.

To see this, note that if $d(f,g) < \delta$ then $|f(x_k) - g(x_k| < \frac{1}{2}\varepsilon$ for $k = 1,\ldots,n$. For a given j, pick some k with $\|x_k - z_j\| < \frac{1}{4}\varepsilon$. Then, since $\|f - g\| \le \|f\| + \|g\| \le 2$,

$$|f(z_j) - g(z_j)| \le |(f-g)(z_j - x_k)| + |f(x_k) - g(x_k)| < \tfrac{1}{2}\varepsilon + \tfrac{1}{2}\varepsilon = \varepsilon,$$

where have used that $\|f - g\| \le 2$. (This is where we need to restrict our attention to a *bounded* subset of X^*.)

This proves the second half. ∎

7 Proposition. *Let X be a separable normed space. Then any bounded sequence in X^* has a weakly* convergent subsequence.*

Proof: Let $(f_k)_{k=1}^{\infty}$ be a bounded sequence in X^*. We may assume, without loss of generality, that $\|f_k\| \le 1$ for all k. The unit ball of X^* with the weak* topology is compact and metrizable, but it is well known that any sequence in a compact metric space has a convergent subsequence.

We could pretend that we do not know this fact, and proceed directly: This is, if nothing else, a handy exercise to learn to translate back and forth between filters and sequences. The set of tails $\{f_k\colon k \ge n\}$ of the sequence generates a filter $\mathscr{F}$ on B. By compactness, there is a refinment $\mathscr{G}$ which converges to some $g \in B$. Then for each $\varepsilon > 0$, the ball $\{f \in B\colon d(f,g) < \varepsilon\}$ belongs to $\mathscr{G}$, and so does every tail of the original sequence. Hence the intersection of the ball and the tail belongs to $\mathscr{G}$, and this intersection is therefore not empty. In plain language, this means that for each $\varepsilon > 0$ and each n there is some $k > n$ with $d(f_k, g) < \varepsilon$. From this knowledge, building a convergent subsequence is easy: Pick k_1 with $d(f_{k_1}, g) < 1$, then pick $k_2 > k_1$ with $d(f_{k_2}, g) < \frac{1}{2}$, then pick $k_3 > k_2$ with $d(f_{k_3}, g) < \frac{1}{3}$, and so forth. ∎

It is useful to know a sort of converse to the above result: To have any reasonable hope of getting a weakly* convergent subsequence, we had better start with a bounded sequence.

8 Proposition. *Any weakly convergent sequence on a normed space, or weakly* convergent sequence on the dual of a Banach space, is necessarily bounded.*

Proof: Let X be a Banach space, and $(f_k)_{k=1}^{\infty}$ a weakly* convergent sequence. For each $x \in X$, the sequence $\bigl(f_k(x)\bigr)_{k=1}^{\infty}$ is convergent, and hence bounded. By the Banach–Steinhaus theorem (uniform boundedness principle), the sequence $(f_k)_{k=1}^{\infty}$ is bounded.

The corresponding result for weakly convergent sequences on a normed space is proved the same way, but now using the Banach–Steinhaus theorem for functionals on the dual space, which as we know is complete, so the theorem is applicable. ∎

It is important in the second part of the previous result that the space be complete. To see why, let $\ell^1_{\mathrm{c}} \subset \ell^1$ be the set of sequences all of whose entries except a finite number are zero. We use the ℓ^1 norm on this space. Let f_k be the functional $f_k(x) = kx_k$. Then $f_k \xrightarrow{*} 0$ on ℓ^1_{c}, but $\|f_k\| = k \to \infty$.

The geometric Hahn–Banach theorem and its immediate consequences

Let X be a real vector space. Recall that a *sublinear functional* on X is a function $p\colon X \to \mathbb{R}$ so that $p(\alpha x) = \alpha p(x)$ for $\alpha \in \mathbb{R}$, $x \in X$ and $p(x+y) \le p(x) + p(y)$ for $x, y \in X$. Let us say that p *dominates* a linear functional f if $f(x) \le p(x)$ for all x. Finally, recall that the *Hahn–Banach theorem* states that, if p is a sublinear functional on X and f is a linear functional on a subspace of X, dominated by p, then f has an extension to all of X which again is dominated by p.

We shall be interested in the geometric implications of the Hahn–Banach theorem. We shall call x an *interior* point of a set $C \subseteq X$ for every $y \in X$ there is some $\varepsilon > 0$ so that $x + ty \in X$ whenever $|t| < \varepsilon$.[2] It is easy to show that, if p is a sublinear functional on X, then $\{x \in X\colon\ p(x) < 1\}$ and $\{x \in X\colon\ p(x) \le 1\}$ are convex sets with 0 an interior point. The converse is also true:

9 Lemma. *Let C be a convex subset of a real vector space X, and assume that* 0 *is an interior point of C. Then the function p defined by*

$$p(x) = \inf\bigl\{t > 0\colon\ x/t \in C\bigr\}$$

is a sublinear functional on X. Moreover, x is an interior point of C if and only if $p(x) < 1$, and $x \notin C$ if $p(x) > 1$.

The functional p is sometimes called the *gauge* of C.

Proof: The homogeneity condition $p(\alpha x) = \alpha p(x)$ when $\alpha > 0$ is fairly obvious.

For subadditivity, if $x/s \in C$ and $y/t \in C$ then we can form the convex combination

$$\frac{s}{s+t}\,\frac{x}{s} + \frac{t}{s+t}\,\frac{y}{t} = \frac{x+y}{s+t}$$

so that $(x+y)/(s+t) \in C$, and $p(x+y) \le s+t$. Taking the infimum over all s and t satisfying the conditions, we find $p(x+y) \le p(x) + p(y)$.

[2] There is potential for confusion here, as *interior* is usually a *topological* concept. If X is a topological vector space then an interior point of C (in the topological sense, i.e., C is a neighbourhood of x) is also an interior point in the sense defined here. The converse is not true.

If $p(x) > 1$ then $x/t \notin C$ for some $t > 1$. Equivalently, $sx \notin C$ for some $s < 1$. But then $x \notin C$, since $sx = sx + (1-s)0$ is a convex combination of x and 0 and C is convex.

If $p(x) = 1$ then x is not an interior point of C. For then $x + tx = (1+t)x \notin C$ for any $t > 0$.

If $p(x) < 1$, then $x \in C$ first of all, for then there is some $t < 1$ so that $x/t \in C$. And so $x = t(x/t) + (1-t)0 \in C$ because C is convex. Next, if $y \in X$ and $t \in \mathbb{R}$ then $p(x+ty) \le p(x) + |t|\bigl(p(y) \vee p(-y)\bigr) < 1$ when $|t|$ is small enough, so $x + ty \in C$ when $|t|$ is small enough. Thus x is an interior point of C. (Here $a \vee b$ is the maximum of a and b.) ∎

10 Theorem. (Hahn–Banach separation theorem I)
Let X be a real vector space, and C a convex subset of X with at least one interior point. Let $x \in X \setminus C$. Then there exists a linear functional f on X so that $f(z) \le f(x)$ for every $z \in C$, and $f(z) < f(x)$ when z is an interior point of C.

Proof: We may, without loss of generality, assume that 0 is an interior point of C. (Otherwise, replace C by $C - w$ and x by $x - w$ where w is an interior point of C.)

Let p be the gauge of C. Define $f(tx) = tp(x)$ for $t \in \mathbb{R}$; then f is a linear functional on the one-dimensional space spanned by x which is dominated by p. By the Hahn–Banach theorem, we can extend f to all of X, with the extension still dominated by p. For $z \in C$ we find $f(z) \le p(z) \le 1 \le p(x) = f(x)$, with the middle inequality being strict when z is interior in C. ∎

11 Theorem. (Hahn–Banach separation theorem II)
Let X be a real vector space, and U and V two nonempty disjoint convex subsets of X, at least one of which contains an interior point. Then there is a nonzero linear functional f on X and a constant c so that $f(u) \le c \le f(v)$ for all $u \in U$, $v \in V$. If u is an interior point of U then $f(u) < c$. Similarly, if v is an interior point of V then $f(v) > c$.

Proof: Let $C = U - V = \{u - v\colon\ u \in U, v \in V\}$. Then $0 \notin C$ because $U \cap V = \emptyset$. Moreover, C is convex and has an interior point. Thus there is a linear functional f so that $f(x) \le f(0) = 0$ for $x \in C$, and $f(x) < 0$ for any interior point x of C. Thus, if $u \in U$ and $v \in V$, we get $f(u - v) \le 0$, and so $f(u) \le f(v)$. It follows that $\sup_{u\in U} \le f(v)$ for all $v \in V$, and then also $\sup_{u\in U} \le \inf_{v\in V} f(v)$. We can pick c between these two numbers to finish the proof. If u is an interior point of U and $v \in V$, let $z \in X$ with $f(z) > 0$. Then $u + tz \in U$ when t is small and positive, so $f(u) < f(u + tz) \le c$. The proof of the corresponding strict inequality for an interior point v of V is proved the same way. ∎

We now investigate the consequences of the Hahn–Banach theorem in locally convex spaces.

12 Corollary. *Let X be a locally convex topological vector space, C a closed, convex subset of X, and $w \in X \setminus C$. Then there is a continuous linear functional f on X and a constant c so that $\operatorname{Re} f(x) \le c < \operatorname{Re} f(w)$ for all $x \in C$.*

Proof: Let V be a convex neighbourhood of w with $V \cap C = \emptyset$. Apply the previous theorem to C and V, noting that w is an interior point of V. If X is a complex space, apply this result to X as a real space, then use the fact that any real linear functional on X is the real part of a complex linear functional on X. ∎

13 Corollary. *Let X be a locally convex topological vector space, and let $w \in X$ be a nonzero vector. Then there is a continuous linear functional f on X with $f(w) \neq 0$.*

Proof: Apply the previous corollary with $C = \{0\}$. ∎

14 Corollary. *Let C be a (norm) closed convex subset of a normed space X. Then C is weakly closed.*

Proof: Let $w \in X \setminus C$, and pick a linear functional f as in Corollary 12. Then the weak neighbourhood $\{x \colon f(x) > c\}$ of w is disjoint from C. ∎

This has interesting consequences for weakly convergent sequences. From a weakly convergent sequence one can form a new sequence by taking convex combinations of members of the original sequence to obtain a norm convergent sequence. The details follow these definitions.

Given a set of points $A \subset X$, a *convex combination* of points in A is a point which can be written

$$\sum_{k=1}^{n} t_k a_k, \qquad a_k \in A, \quad t_k \ge 0, \quad \sum_{k=1}^{n} t_k = 1.$$

The set of all convex combinations of points in A forms a convex set, called the *convex hull* of A, and written $\operatorname{co} A$. If X is a topological vector space, the closure of $\operatorname{co} A$ is also convex. It is called the *closed convex hull* of A, and written $\overline{\operatorname{co}}\, A$.

15 Corollary. *Let X be a normed space, assume $x_k \in X$ and $x_k \rightharpoonup x$ (weakly). Then, given $\varepsilon > 0$ and a number N, there exists some convex combination z of $x_N, x_{N+1}, \ldots$ with $\|x - z\| < \varepsilon$.*

Proof: Let $A = \{x_N, x_{N+1}, \ldots\}$, and consider $\overline{\operatorname{co}}\, A$ (with closure in the norm topology). By Corollary 14, $\overline{\operatorname{co}}\, A$ is weakly closed as well. Since $x_k \rightharpoonup x$ and $x_k \in A$ for $k \ge N$, $x \in \overline{\operatorname{co}}\, A$. The conclusion is immediate. ∎

The *Banach–Saks* theorem is a special case of the above corollary for L^p spaces. In this case, the convex combinations can be chosen to be of the special form $\left(x_{n_1} + \cdots + x_{n_m}\right) / m$.

Uniform convexity and reflexivity

We are now ready to prove the theorem that will finish our proof on the dual space of L^p for $1 < p < \infty$. The proof presented here is essentially due to J. R. Ringrose (J. London Math. Soc. **34** (1959), 92.)

We need a density result first.

16 Lemma. *The closed unit ball in the dual space of a normed space is weakly* closed.*

Proof: Let X be a normed space, and B the closed unit ball of X^*, Then B is defined by inequalities $|f(x)| \le 1$ where $x \in X$ and $\|x\| \le 1$, and the maps $x \mapsto f(x)$ are weakly* continuous. ∎

17 Proposition. *Let X be a normed space. Identify X with its canonical image in X^{**}. Then the closed unit ball of X^{**} is the weak*-closure of the unit ball of X.*

Proof: By the above lemma applied to X^* instead of X, the closed unit ball of X^{**} is weakly* closed.

Let B be the closed unit ball of X, and $\overline{B}$ its weak*-closure. If $\overline{B}$ is not the entire unit ball of X^{**} there is some $\xi \in X^{**}$ with $\|\xi\| \le 1$ and $\xi \notin \overline{B}$. We can find a weakly* continuous linear functional F on X^{**}, and a constant c, so that $\operatorname{Re} F(x) \le c < \operatorname{Re} F(\xi)$ whenever $x \in \overline{B}$. But this functional must be of the form $F(\zeta) = \zeta(f)$ where $f \in X^*$. Bearing in mind that we have identified X with its canonical image in X^{**}, we then have $\operatorname{Re} f(x) \le c < \operatorname{Re} \xi(f)$ for every $x \in B$.

The first inequality implies $\|f\| \le c$. The other implies $c < \|\xi\| \cdot \|f\|$, and combining these two inequalities we obtain $\|\xi\| > 1$. But this contradicts the assumption $\|\xi\| \le 1$, so we are done. ∎

The following lemma will perhaps make Ringrose's extremely short proof clearer. Recall that the *diameter* of a subset A of a normed space is defined to be $\operatorname{diam} A = \inf\{\|x - y\| \colon x, y \in A\}$.

18 Lemma. *Let $A \subseteq X^*$ where X is a normed space. Then the diameter of the weak*-closure of A equals the diameter of A.*

Proof: Now let $d = \operatorname{diam} A$, and write $\overline{A}$ for the weak*-closure of A.

Let $x \in A$. Then $A \subseteq x + dB$. Since $x + dB$ is weakly* closed, $\overline{A} \subseteq x + dB$. Thus $\|x - y\| \le d$ for all $x \in A$ and $y \in \overline{A}$.

Next, fix $y \in \overline{A}$. The previous paragraph implies $A \subseteq y + dB$. As before, this implies $\overline{A} \subseteq y + dB$. Thus $\|x - y\| \le d$ for all $x, y \in \overline{A}$, and hence $\operatorname{diam} \overline{A} \le d$.

On the other hand $\overline{A} \supseteq A$, so we must have $\operatorname{diam} \overline{A} \ge d$ as well. ∎

19 Theorem. (Milman–Pettis) *A uniformly convex Banach space is reflexive.*

Proof: Let X be a uniformly convex Banach space. Let $\xi \in X^{**}$. We must show that $\xi \in X$ (where we have again identified X with its canonical image in X^{**}). We may assume that $\|\xi\| = 1$.

Let $\varepsilon > 0$, and pick $\delta > 0$ so that $\|x - y\| < \varepsilon$ whenever $x, y \in X$, $\|x\| \le 1$, $\|y\| \le 1$ and $\|x + y\| > 2 - \delta$.

Let B be the closed unit ball of X (as a subset of X^{**}), and let $\overline{B}$ be its weak*-closure, which as we know is the unit ball in X^{**}. In particular, $\xi \in \overline{B}$.

From the definition of norm in X^{**}, we can find $f \in X^*$ with $\|f\| = 1$ and $\xi(f) > 1 - \frac{1}{2}\delta$. Let $V = \{\zeta \in X^{**} : \zeta(f) > 1 - \frac{1}{2}\delta\}$. Then V is weakly* open.

Now $\xi \in V \cap \overline{B} \subseteq \overline{V \cap B}$, where the bar denotes weak*-closure. (See below.)

If $x, y \in V \cap B$ then $\|x + y\| > 2 - \delta$ because $f(x + y) = f(x) + f(y) > 2 - \delta$. Thus $\|x - y\| < \varepsilon$, and so $\operatorname{diam} V \cap B \le \varepsilon$. By the previous lemma, $\operatorname{diam} \overline{V \cap B} \le \varepsilon$ as well. In particular, $\|\xi - x\| \le \varepsilon$, where $x \in V \cap B$.

We have shown that ξ lies in the norm closure of X. But since X is a Banach space, it is complete and therefore norm closed in X^{**}. Thus $\xi \in X$, and the proof is complete.

Note: The inclusion $V \cap \overline{B} \subseteq \overline{V \cap B}$ used above holds in any topological space, where V is open. For then the complement W of $\overline{V \cap B}$ is open, and $W \cap V \cap B = \emptyset$. And since $W \cap V$ is open, this implies $W \cap V \cap \overline{B} = \emptyset$, which is another way to state the desired inclusion. ∎

The Krein–Milman theorem

Any point in a cube, or an octahedron, or any other convex polygon, can be written as a convex combination of the corners of the polygon. The Krein–Milman theorem is an infinite-dimensional generalization of this fact.

Before getting to the theorem, however, we need another easy conseqence of the Hahn–Banach separation theorems.

20 Proposition. *Let X be a locally convex topological vector space, and $A \subset X$. Then $w \in \overline{\operatorname{co}} A$ if and only if* $\operatorname{Re} f(w) \le c$ *for every continuous linear functional f on X and every scalar c satisfying* $\operatorname{Re} f(x) \le c$ *for all $x \in A$.*

Proof: Let C be the set of points satisfying the stated condition. Clearly, C is closed and convex, and $A \subseteq C$. Thus $\overline{\operatorname{co}} A \subseteq C$.

If $w \notin \overline{\operatorname{co}} A$, then by Corollary 12 (with $\overline{\operatorname{co}} A$ in the place of C) there is a continuous linear functional f and a constant c so that $\operatorname{Re} f(x) \le c < \operatorname{Re} f(w)$ for all $x \in \overline{\operatorname{co}} A$. In particular this holds for all $x \in A$, and so $w \notin C$. Thus $C \subseteq \overline{\operatorname{co}} A$. ∎

Let K be a convex set in some vector space. A *face* of K is a nonempty convex subset $F \subseteq K$ so that, whenever $u, v \in K$, $0 < t < 1$, and $ty + (1 - t)v \in F$, it follows that $u, v \in F$. An *extreme point* of K is a point x so that $\{x\}$ is a face of K. In other words, whenever $u, v \in K$, $0 < t < 1$, and $ty + (1 - t)v = x$, it follows that $u = v = x$. The set of extreme points of K is also called the *extreme boundary* of K, and written $\partial_e K$. (Make a drawing illustrating these concepts!)

As an example, assume real scalars (for simplicity) and let f be a linear functional so that $f(x) \le c$ for all $x \in K$. Then $\{x \in K : f(x) = c\}$ is a face of K, if it is nonempty.

As another example, consider an ordinary closed cube in $\mathbb{R}^3$. The faces of this cube are: The cube itself, its sides (what we think of as its faces in everyday language), its edges, and its corners (or rather the singleton sets made up of its corners). In particular, the corners are the extreme points of the cube.

21 Theorem. (Krein–Milman) *Any compact convex subset of a locally convex vector space is the closed convex hull of its extreme boundary.*

More briefly, if K is such a set then $K = \overline{\operatorname{co}}\, \partial_e K$.

Proof: We prove this for the case of real scalars. The case for complex scalars follows in the usual way.

We shall begin by proving a much weaker statement, namely that *any compact convex set* (in a locally convex space) *contains at least one extreme point.*

To this end, let K be a compact convex set, and consider the collection Φ of all its closed faces. We shall prove that there exists a *minimal* closed face, and then we shall prove that a minimal closed face contains only one point.

For the first part, we shall use Zorn's lemma. We need to show that Φ is inductively ordered by *inverse* inclusion. If $\mathscr{C}$ is a chain in Φ, then $\bigcap \mathscr{C} \neq \emptyset$ because $\mathscr{C}$ consists of closed sets, it has the finite intersection property (for it is totally ordered, so any intersection of a finite number of members of $\mathscr{C}$ is simply the smallest of them), and K is compact. It is also a face (exercise: show this). Thus Zorn's lemma proves the existence of a *minimal* member of Φ.

Now let F be a minimal closed face of K. Assume that F contains two distinct points x and y. Let f be a continous linear functional with $f(x) \neq f(y)$. Then f attains its maximum on F, because F is compact (it is a closed subset of the compact set K). Writing c for this maximum, then $\{z \in F\colon f(z) = c\}$ is a face of F, and therefore a face of K as well. Since F was a *minimal* face, this new face must be all of F, so that $f(z) = c$ for all $z \in F$. But this contradicts $f(x) \neq f(y)$. Thus F cannot contain two distinct points, so it consists of a single point. Thus K contains at least one extreme point.

Now, assume that $K \neq \overline{\text{co}}\,\partial_e K$. Since obviosly $\overline{\text{co}}\,\partial_e K \subseteq K$, we can now use Corollary 12 to find a continuous linear functional f and a constant c so that $f(x) \le c$ for all $x \in \overline{\text{co}}\,\partial_e K$, but $f(x) > c$ for at least some points in K. Let m be the maximum value of f over K. It exists because K is compact. Then $\{x \in X\colon f(x) = m\}$ is a closed face of K. In particular, this set is compact and convex, so it contains at least one extreme point (of itself, and therefore of K). But now if this extreme point is called x then $f(x) = m$, but also $f(x) \le c < m$ because $x \in \partial_e K \subseteq \overline{\text{co}}\,\partial_e K$. This contradiction completes the proof. ∎

If we throw away a corner of a polyhedron such as a cube or octahedron in three dimensional space, the remaining corners will have a smaller convex hull. On the other hand, we can trow away single points from the extreme boundary of a disk, and the closed convex hull of the remaining points will still be the entire disk. The next theorem, known as Milman's converse to the Krein–Milman theorem, tells us the precise story of what can be thrown away and what must be kept.

22 Theorem. (Milman) *Let K be a compact convex subset of a locally convex topological vector space X. Let $A \subseteq K$ so that $K = \overline{\text{co}}\,A$. Then $\partial_e K \subseteq \overline{A}$.*

Proof: Assuming the conclusion is wrong, we let $w \in \partial_e K$ while $w \notin \overline{A}$.

There is an open convex neighbourhood U of 0 with $(w + U - U) \cap A = \emptyset$. Then $(w + U) \cap (\overline{A} + U) = \emptyset$ as well. Since $\overline{A}$ is a closed subset of the compact set K, it is compact, so there are $x_1, \dots, x_n \in \overline{A}$ so that the n sets $x_k + U$ cover $\overline{A}$.

Let $K_k = \overline{\text{co}}\big((x_k + U) \cap \overline{A}\big)$. Each K_k is a compact convex subset of K, and $K = \overline{\text{co}}(K_1 \cup \ldots \cup K_n)$ because $A \subseteq K_1 \cup \ldots \cup K_n$.

Even better: Any point $x \in K$ can be written as a convex combination

$$x = \sum_{k=1}^{n} t_k x_k, \qquad t_k \ge 0,\ x_k \in K_k,\ k = 1, \dots, n,\ \sum_{k=1}^{n} t_k = 1.$$

For the set of all sums as given above is a compact set (it is a continous image of the compact set $T \times K_1 \times \cdots \times K_n$, where $T = \{t \in \mathbb{R}^n\colon t_k \ge 0,\ \sum_{k=1}^{n} t_k = 1\}$) and also convex, so that set *is* the closed convex hull of $K_1, \dots, K_n$.

In particular, w can be written that way. But $w \notin K_k$ for any k, for $K_k \subseteq \overline{x_k + U}$ which is disjoint from $w + U$. Thus, when we write $w = \sum_{k=1}^{n} t_k x_k$ at least two of the coefficients, say t_k and $t_{k'}$, are nonzero. But then, varying t_k and $t_{k'}$ while keeping their sum constant, we get a line segment lying in K with w in its interior. This is impossible, since w is assumed to be an extreme point of K.

There is a tiny hole in the above argument, for we could have $x_k = x_{k'}$ (the sets K_k will overlap), in which case the line segment degenerates to a point. But this hole is easy to plug: We cannot have *all* those x_k for which $t_k > 0$ being identical, for then we would have $w \in K_k$, and that is not the case. ∎