LECTURES ON FUNCTIONAL ANALYSIS

R. SHVYDKOY

1. Normed Spaces

1.1. Basic concepts and notation. Let us consider a vector space $X$ over the field $\mathbb{K} = \mathbb{R}$ or $\mathbb{C}$. We say that $X$ has finite dimension $n$ is there is a system of $n$ linearly independent vectors $\{x_1, \ldots, x_n\}$ in $X$ which spans $X$. We denote the linear span of a set $S \subset X$ by $[S]$. If $X$ has no finite dimension, $X$ is called infinite dimensional. A function $\| \cdot \|$ is said to define a norm on $X$ if the following axioms hold:

(i) $\|x\| = 0$ iff $x = 0$,
(ii) $\|\alpha x\| = |\alpha|\|x\|$ for all $x \in X$, $\alpha \in \mathbb{K}$,
(iii) $\|x + y\| \leq \|x\| + \|y\|$ (triangle inequality).

A function $\| \cdot \|$ satisfying only (ii) and (iii) is called a pseudo-norm. We write $(X, \| \cdot \|)$ to indicate that $X$ is equipped with the norm $\| \cdot \|$ if it is not clear from the context.

Exercise 1.1. Show that (iii) is equivalent to

$$\|x - y\| \leq \|x \pm y\|.$$ 

Let us notice that a norm generates a metric, called norm-metric, on the space $X$ via

$$d(x, y) = \|x - y\|.$$ 

The corresponding topology is called norm-topology. The norm-topology naturally gives rise to the concept of convergence and continuity. A sequence $\{x_n\}_{n=1}^{\infty} \subset X$ is said to converge to $x$ in the norm-metric, or strongly, if $\|x_n - x\| \to 0$ as $n \to \infty$. A function $f : X \to Y$, where $Y$ is a topological space, is continuous if $f(x_n) \to f(x)$ whenever $x_n \to x$, or equivalently, if $f^{-1}(U)$ is open for any open $U \subset Y$.

Exercise 1.2. Prove that the norm $\| \cdot \| : X \to \mathbb{R}$ is a continuous function on $X$.

A series $\sum_{n=1}^{\infty} x_n$ is said to converge to its sum $x \in X$ if the sequence of partial sums $\sum_{n=1}^{N} x_n$ tends to $x$. A series converges absolutely if the numerical series $\sum_{n=1}^{\infty} \|x_n\|$ converges. The natural fact that an absolutely convergent series itself converges does not necessarily hold in general normed spaces. To ensure this property one has to impose an additional completeness assumption. A metric space $(M, d)$ is complete if its every Cauchy sequence has a limit, i.e. if $d(x_n, x_m) \to 0$ implies $\exists x \in M$ such that $d(x_n, x) \to 0$. A complete normed space $(X, \| \cdot \|)$ is called a Banach space.

Exercise 1.3. Prove that a normed space $(X, \| \cdot \|)$ is Banach if and only if every absolutely convergent series in $X$ is convergent.
For two sets \( A, B \subseteq X \) we denote by \( A + B \) their algebraic sum \( \{ x + y : x \in A, y \in B \} \), and constant multiple by \( \alpha A = \{ \alpha x : x \in A \} \). Next, \[ B(X) = \{ x \in X : \| x \| \leq 1 \}, \quad \text{the closed unit ball} \]
\[ D(X) = \{ x \in X : \| x \| < 1 \}, \quad \text{the open unit ball} \]
\[ S(X) = \{ x \in X : \| x \| = 1 \}, \quad \text{the unit sphere} \]
\[ B_r(x_0) = \{ x \in X : \| x - x_0 \| \leq r \} = x_0 + rB(X) \]
\[ D_r(x_0) = \{ x \in X : \| x - x_0 \| < r \} = x_0 + rD(X). \]
The family of open ball forms a basis for the norm-topology.

1.2. **Classical examples.** The simplest example of a normed space is the Euclidean space \( \ell_2^n = (\mathbb{K}^n, \| \cdot \|_2) \) with the norm given by
\[
\| x \|_2 = \left( \sum_{i=1}^{n} |x_i|^2 \right)^{1/2}.
\]
The Euclidean norm is generated by the inner product \( \langle x, y \rangle = \sum x_i\bar{y}_i \) via \( \| x \|_2 = \langle x, x \rangle^{1/2} \).

The triangle inequality in this case is a consequence of the Cauchy-Schwartz inequality:
\[
\| x + y \|_2^2 = \| x \|_2^2 + 2 \text{Re} \langle x, y \rangle + \| y \|_2^2 \leq \| x \|_2^2 + 2 \| x \| \| y \| + \| y \|_2^2 \leq (\| x \| + \| y \|)^2.
\]

There is a range of other natural norms on \( \mathbb{K}^n \), of which \( \| \cdot \|_2 \) is a part, and for which the triangle inequality is not so straightforward. We introduce them next.

Let \( 1 \leq p < \infty \). We define \( \ell_p \) as the space of sequences \( x = (x_1, x_2, \ldots) \) such that \( \sum_i |x_i|^p < \infty \) and endow it with the norm
\[
\| x \|_p = \left( \sum_{i=1}^{\infty} |x_i|^p \right)^{1/p}.
\]
For \( p = \infty \), \( \ell_\infty \) is the space of bounded sequences endowed with the supremum-norm
\[
\| x \|_\infty = \sup_i |x_i|.
\]

The corresponding \( n \)-dimensional analogue is denoted \( \ell_p^n \). At this point, other than in cases \( p = 1, \infty \), it not clear whether \( \ell_p \) is a linear space and \( \| \cdot \|_p \) defines a norm on it. We will show it next and establish several very important inequalities as we go along the proof.

**Lemma 1.4.** \((\ell_p, \| \cdot \|_p) \) is a Banach space for all \( 1 \leq p \leq \infty \).

**Proof.** First, let \( p = \infty \). The axioms of norm in this case are trivial. To show that \( \ell_\infty \) is complete let \( x_n = \{ x_n(j) \}_{j=1}^{\infty} \) be a Cauchy sequence. Hence, every numerical sequence \( \{ x_n(j) \} \) is Cauchy. This implies that \( x_n(j) \rightarrow x(j) \) as \( n \rightarrow \infty \). Since \( \{ x_n \} \) is Cauchy, we have \( \| x_n - x_m \|_\infty < \varepsilon \) for all \( n, m > N \). Thus, \( |x_n(j) - x_m(j)| < \varepsilon \) for all \( j \in \mathbb{N} \) as well. Let us fix \( n \) and \( j \) and let \( m \rightarrow \infty \) in the last inequality. We obtain \( |x_n(j) - x(j)| \leq \varepsilon \) for all \( j \), and hence \( \| x_n - x \| \leq \varepsilon \), for all \( n > N \). We have shown that \( x_n \rightarrow x \).

Now let \( p < \infty \). Let us prove the triangle inequality first. By concavity of \( \ln(x) \), we have
\[
\ln(\lambda a + \mu b) \geq \lambda \ln(a) + \mu \ln(b),
\]
for all \( \lambda + \mu = 1 \), \( \lambda, \mu \geq 0 \), and \( a, b > 0 \). Exponentiating the above inequality we obtain
\[
ab \leq \frac{a^p}{p} + \frac{b^q}{q}, \quad \text{(Young’s Inequality)}
\]
whenever $\frac{1}{p} + \frac{1}{q} = 1$, $p \geq 1$. Next, consider finite sequences $x = \{x_i\}_{i=1}^n$, $y = \{y_i\}_{i=1}^n$ and observe by (2)

$$\left| \sum_{i=1}^n x_i y_i \right| \leq \frac{1}{p} \sum_{i=1}^n |x_i|^p + \frac{1}{q} \sum_{i=1}^n |y_i|^q = \frac{1}{p} \|x\|_p^p + \frac{1}{q} \|y\|_q^q.$$ 

Thus, if $\|x\|_p = \|y\|_q = 1$, then $\sum_i x_i y_i \leq 1$. For general $x$ and $y$, after normalization we obtain

$$\sum_{i=1}^n x_i y_i \leq \|x\|_p \|y\|_q, \quad \text{(Hölder Inequality)}.$$ 

Finally,

$$\sum_{i=1}^n |x_i + y_i|^p \leq \sum_{i=1}^n |x_i + y_i|^p |x_i| + \sum_{i=1}^n |x_i + y_i|^{p-1} |y_i|$$

$$\leq \left( \sum_{i=1}^n |x_i + y_i|^{(p-1)q} \right)^{1/q} \left[ \|x\|_p + \|y\|_p \right] = \|x + y\|_p^{p/q} \left[ \|x\|_p + \|y\|_p \right].$$

Thus,

$$\|x + y\|_p^p \leq \|x + y\|_p^{p/q} \left[ \|x\|_p + \|y\|_p \right],$$

and this implies

$$\|x + y\|_p \leq \|x\|_p + \|y\|_p, \quad \text{(Minkowski’s inequality)}$$

which is what we need only for finite sequences. It remains to notice that if $x, y \in \ell_p$ are arbitrary, then the above inequality shows that the partial sums of the $p$-series of $x + y$ are uniformly bounded, which in turn implies that $x + y \in \ell_p$ and the triangle inequality (iii) holds as desired.

To prove that $\ell_p$ is complete, let $x_n = \{x_n(j)\}_{j=1}^\infty$ be Cauchy. Then, as before we can pass to the limit in every coordinate $x_n(j) \to x(j)$. For a fixed $J \in \mathbb{N}$, $\varepsilon > 0$ and $n, m$ large enough, we have

$$\sum_{j=1}^J |x_n(j) - x_m(j)|^p < \varepsilon.$$ 

Letting $m \to \infty$, we obtain

$$\sum_{j=1}^J |x_n(j) - x(j)|^p < \varepsilon.$$ 

In particular, this implies that all partial sums of the series $\sum_j |x(j)|^p$ are bounded, and hence $x = \{x(j)\}_j \in \ell_p$. Now, let us let $J \to \infty$ in the estimate above. We obtain then

$$\|x_n - x\|_p \leq \varepsilon^{1/p}, \quad \text{thus} \quad x_n \to x.$$ 

A normed space $X$ is called **separable** if it contains a countable dense subset, i.e. if there is $S \subset X$, $\text{card } S = \omega_0$ such that for every $x \in X$ and any $\varepsilon > 0$ there is $y \in S$ with $\|x - y\| < \varepsilon$.

**Exercise 1.5.** Show that $\ell_p$ is a separable space for all $1 \leq p < \infty$.

**Exercise 1.6.** Show that $\ell_\infty$ is not a separable space. Hint: consider the set of vectors $\{x_A\}_{A \in \mathbb{N}}$, where $x_A$ is the characteristic function of $A$.

Our next classical example is $c_0$. This is the space of sequences $\{x_j\}_{j=1}^\infty$ such that $\lim_{j \to \infty} x_j = 0$ endowed with the uniform $\| \cdot \|_\infty$ norm.
1.3. Constructing new spaces from old ones. There are many ways to construct new spaces from existing examples. A direct product of two linear spaces \( X \) and \( Y \), denoted \( X \times Y \) is the set of pairs \( \{(x, y) : x \in X, y \in Y \} \) endowed with the coordinate-wise operation of summation and multiplication by a scalar. This makes \( X \times Y \) into a linear space. Identifying elements of the product \( (x, 0) \) with \( x \), and \( (0, y) \) with \( y \) arranges a natural embedding of \( X \) and \( Y \) into \( X \times Y \). We thus can write \( x + y = (x, y) \). Let now \( (X, \| \cdot \|_X) \) and \( (Y, \| \cdot \|_Y) \) be normed and let \( 1 \leq p < \infty \). We can define a new norm on the product \( X \times Y \) by

\[
\|x + y\|_p = (\|x\|_X^p + \|y\|_Y^p)^{1/p}.
\]

The verification that this rule defined a norm is immediate from Minkowski’s inequality established above. The obtained normed space is called the \( \ell_p \)-sum of the \( X \) and \( Y \) and denoted \( X \oplus_p Y \). For \( p = \infty \) we naturally define \( X \oplus_\infty Y \) equipped with the norm \( \|x + y\|_\infty = \max\{\|x\|_X, \|y\|_Y\} \). Similarly, we define \( \ell_p \)-sums of any number of spaces and even countably many spaces by requiring a member of \( X_1 \oplus_p X_2 \oplus_p \ldots \) to be a sequence of vectors \( x = \{x_1, \ldots\} \) such that \( \|x\|_p = \left( \sum_{j=1}^\infty \|x_j\|_{X_j}^p \right)^{1/p} < \infty \), or bounded in the case \( p = \infty \).

Exercise 1.7. Verify that all of the newly introduced spaces are Banach if the original spaces are Banach.

Let us notice that for any pair of vectors \( x \in S(X) \) and \( y \in S(Y) \), the span \([x, y]\) will be identical to \( \ell_p^2 \) in the \( \ell_p \)-product of spaces. So, for example, the unit ball of the \( X \oplus_1 \mathbb{R} \) will look like a symmetric tent with \( B(X) \) being the base and \((0, 1)\) the top point. The ball of \( X \oplus_\infty \mathbb{R} \) would be the cylinder with base \( B(X) \) and height 1.

A subspace \( Y \) of a linear space \( X \) is a subset which is closed under the linear operations. \( Y \) is closed if it is closed in the norm-topology of \( X \). If in addition \( X \) is complete, then so is every closed subspace. Thus, any closed subspace of a Banach space is Banach. Let us now fix a closed subspace \( Y \subset X \) and consider the equivalence relation \( x_1 \sim x_2 \) off \( x_1 - x_2 \in Y \). This defines a conjugacy class \([x]\), for every \( x \in X \). The space of all conjugacy classes is called the factor-space of \( X \) by \( Y \), denoted \( X/Y \), with the natural linear operations inherited from \( X \). We can endow \( X/Y \) with a norm too, called the factor-norm:

\[
\| [x] \| = \inf \{ \|x + y\| : y \in Y \} = \operatorname{dist}(x, Y).
\]

Exercise 1.8. Show that the above defines a norm on \( X/Y \). Show that if \( X \) is complete, then \( X/Y \) is complete as well in the factor-norm.

If \( X \) is endowed with a pseudo-norm, \( \| \cdot \| \), consider \( X_0 = \{ x \in X : \|x\| = 0 \} \). This is a closed linear subspace of \( X \), and moreover, \( \|x + y\| = \|x\| \) for all \( x \in X, y \in X_0 \). It is easy to show that (4) defines a norm on \( X/X_0 \), i.e. axiom (i) holds.

1.4. Norm comparison and equivalence. Let \( (X, \| \cdot \|) \) be a normed space and \( Y \subset X \) is a subspace with another norm \( \| \cdot \|'. \) We say that the norm \( \| \cdot \| \) is stronger than \( \| \cdot \|' \) if there exists a constant \( C > 0 \) such that

\[
\|y\| \leq C \|y\|', \text{ for all } y \in Y.
\]

The two norms are equivalent if there are \( c, C > 0 \) for which

\[
c \|y\|' \leq \|y\| \leq C \|y\|', \text{ for all } y \in Y.
\]

Geometrically, (5) means that \( B \|\cdot\|'(Y) \subset CB \|\cdot\|'(Y) \), while (6) means that there is embedding in both sides, \( cB \|\cdot\|'(Y) \subset B \|\cdot\|'(Y) \subset CB \|\cdot\|'(Y) \). The stronger norm, therefore, defines a finer topology on \( Y \), while equivalent norms define the same topology.
Example 1.9. We have $\ell_p \subset \ell_q$, for all $1 \leq p \leq q \leq \infty$, and
\begin{equation}
\|x\|_q \leq \|x\|_p. 
\end{equation}
Indeed, assuming that $x = (x_1, \ldots) \in S(\ell_p)$ implies that all $|x_i| \leq 1$. Hence, $|x_i|^q \leq |x_i|^p$, and thus, $x \in \ell_q$. Moreover, $\|x\|_q \leq 1$. The general inequality (7) follows by homogeneity.

Example 1.10. Let $(\Omega, \Sigma, \mu)$ be a finite measure space, $\mu(\Omega) < \infty$. We have the opposite embeddings for the Lebesgue spaces: $L^q(d\mu) \subset L^p(d\mu)$, for all $1 \leq p \leq q \leq \infty$, and
\begin{equation}
\|f\|_p \leq \mu(\Omega)^{\frac{1}{p} - \frac{1}{q}} \|f\|_q, \text{ for all } f \in L^q(d\mu).
\end{equation}
It readily follows from the Hölder inequality,
\[
\int_\Omega |f|^p d\mu \leq \left( \int_\Omega |f|^q d\mu \right)^{\frac{p}{q}} \mu(\Omega)^{1 - \frac{p}{q}}.
\]
Thus, $\|f\|_p \leq \mu(\Omega)^{1 - \frac{p}{q}} \|f\|_q^p$, from which (8) follows.

Exercise 1.11. Show that in all of the examples above the norms are not equivalent on the subspace in question.

Exercise 1.12. Verify the inequality
\[
\|x\|_q \leq \|x\|_p \leq n^{\frac{1}{p} - \frac{1}{q}} \|x\|_q,
\]
for vectors $x \in \mathbb{R}^n$. Can you interpret the upper bound as a particular case of (8)?

Exercise 1.13. Show that if $\mu(\Omega) = \infty$, then the $L^p$-norms are not comparable on $L^p \cap L^q$, i.e. neither is stronger than the other.

Theorem 1.14. On a finite dimensional linear space $X$ all norms are equivalent.

Proof. By transitivity, it suffices to show that all norms on $\mathbb{R}^n$ are equivalent to the norm of $\ell_1^n$. So, let $\| \cdot \|$ be a norm on $\mathbb{R}^n$, and let $\{e_i\}_{i=1}^n$ be the vectors of the standard unit basis. Then for $x = \sum x_i e_i$,
\[
\|x\| \leq \sum |x_i| \|e_i\| \leq \max\{\|e_i\|\} \sum |x_i| = M\|x\|_1.
\]
To establish an inequality from below, let us consider the norm-function $N(x) = \|x\|$. By compactness of $S(\ell_1^n)$ and continuity of $N$, $N$ attains its minimum on $S(\ell_1^n)$ at $x_0$. Then $N(x_0) = c > 0$, since $N$ never vanishes on a non-zero vector. So, $\|x\| \geq c$, for all $x \in S(\ell_1^n)$, and hence $\|x\| \geq c\|x\|_1$, by homogeneity.

1.5. Linear bounded operators. A map $T$ between two linear spaces $X \to Y$ is called a linear operator if $T(\alpha x + \beta y) = \alpha T(x) + \beta T(y)$. We usually drop the parentheses, $T(x) = Tx$, when a linear operator is in question. Suppose $(X, \| \cdot \|_X)$ and $(Y, \| \cdot \|_Y)$ are normed. A linear operator $T : X \to Y$ is called bounded or continuous if there exists a constant $C > 0$ such that
\begin{equation}
\|Tx\|_Y \leq C\|x\|_X,
\end{equation}
holds for all $x \in X$. We denote the set of all linear bounded operators between $X$ and $Y$ by $\mathcal{L}(X, Y)$. The following theorem justifies the terminology.

Theorem 1.15. Let $T : X \to Y$ be a linear operator. The following are equivalent:
(i) $T \in \mathcal{L}(X, Y)$;
(ii) $T$ maps bounded sets into bounded sets;
(iii) $T$ is continuous as a map between $X$ and $Y$ endowed with their norm topologies;
(iv) $T$ is continuous at the origin.

Proof. The implication $(i) \Rightarrow (ii)$ is clear from (9). Conversely, $T$, in particular, is bounded on the unit ball of $X$, i.e. there exists a $C > 0$ such that, $\|Tx\|_Y \leq 1$, for all $x \in B(X)$. If $x \in X$ is arbitrary, then $\bar{x} = x/\|x\| \in B(X)$, and hence $\|T\bar{x}\|_Y \leq C$. So, by linearity, we obtain (9).

The implication $(i) \Rightarrow (iv)$ is also clear directly from (9). If (iv) holds, and $x_0 \in X$ is arbitrary, then for $y \to 0$, by linearity and continuity at the origin, we have

$$T(x_0 + y) = Tx_0 + Ty \to Tx_0,$$

showing that $T$ is continuous at $x_0$. Thus, $(iii)$ holds. Finally, if (iv) holds, then there is a $\delta > 0$ such that $\|x\| < \delta$ implies $\|Tx\|_Y < 1$. So, if $x$ is arbitrary, consider $x' = \delta x/\|x\|$.

Then $\|Tx'\| \leq 1$ implies $\|Tx\| \leq \|x\|/\delta$ giving us (9). □

If $T \in \mathcal{L}(X, Y)$ we define the norm $T$ as follows:

$$\|T\| = \inf\{C > 0 : (9) \text{ holds}\}.$$

In particular, for any $x \in X$, $\|Tx\| \leq C\|x\|$ holds for all $C > 0$ for which (9) holds. This shows that the infimum is in fact attained,

$$\|Tx\|_Y \leq \|T\|\|x\|_X, \quad \forall x \in X.$$

Exercise 1.16. Show that

$$\|T\| = \sup_{x \in B(X)} \|Tx\|_Y = \sup_{x \in S(X)} \|Tx\|_Y.$$

These identities say that the norm of an operator is the measure of deformation of the unit ball of $X$ under $T$. In particular, if $\|T\| \leq 1$, then $T$ is called a contraction.

Exercise 1.17. Suppose that $T, S \in \mathcal{L}(X, Y)$. Then

$$\|T + S\| \leq \|T\| + \|S\|.$$

The set $\mathcal{L}(X, Y)$ of all bounded linear operators between $X$ and $Y$ clearly forms a linear space, and the above exercise shows that the operator norm endows it with a norm.

**Theorem 1.18.** If $(X, \| \cdot \|_X)$ is normed and $(Y, \| \cdot \|_Y)$ is Banach, then the space $\mathcal{L}(X, Y)$ is Banach in its operator norm.

**Proof.** Suppose $\{T_n\} \subset \mathcal{L}(X, Y)$ is Cauchy. Then, in particular, for any fixed $x \in X$,

$$\|T_n x - T_m x\|_Y \leq \|T_n - T_m\|\|x\|_X \to 0,$$

as $n, m \to \infty$. Thus, $\{T_n x\}$ is Cauchy in $Y$. We, therefore can define the limit

$$T(x) = \lim_{n \to \infty} T_n x,$$

for each $x \in X$. By linearity of $T_n$’s the limit is a linear operator as well. To show that it is bounded, observe that the original sequence of operators is bounded, thus $\|T_n\| \leq M$ for some $M$ and all $n$. So, $\|T x\|_Y \leq M\|x\|_X$ for all $X$, which proves boundedness. Finally, to show that $T_n \to T$ in operator norm, let us fix $\varepsilon > 0$, then for all $n, m$ large and all $x \in B(X)$ we have

$$\|T_n x - T_m x\|_Y < \varepsilon.$$
Let us keep $n$ fixed and let $m \to \infty$. We already know that $T_m x \to Tx$, thus,
\[ \|T_n x - Tx\| \leq \varepsilon \]
holds for all $x \in B(X)$ and all $n$ large. This gives $\|T_n - T\| \leq \varepsilon$ for all $n$ large, which completes the proof.

\begin{proof}
    \[ \varepsilon \]
    \end{proof}

Exercise 1.19. Suppose that $T : X \to Y$ and $S : Y \to Z$ are bounded. Prove that
\[ \|S \circ T\| \leq \|S\| \|T\| . \]

Let us introduce some terminology associated with operators. Let $T \in \mathcal{L}(X,Y)$. The kernel is $T$ is defined by Ker $T = T^{-1}(0) = \{x \in X : Tx = 0\}$, image by Im $T = \{y \in Y : \exists x \in X, Tx = y\}$. Note that for a bounded operator the kernel is always closed, while the image may not be. We say that $T \in \mathcal{L}(X,Y)$ is an isomorphism if $T$ is bijective and the inverse $T^{-1}$ is bounded. This is equivalent to requiring that $T$ is surjective and there are constants $c,C > 0$ such that
\begin{equation}
    c\|x\| \leq \|Tx\| \leq C\|x\|, \text{ for all } x \in X.
\end{equation}
Note that the sharp constants in (10) are in fact $c = \|T^{-1}\|$ and $C = \|T\|$. If an isomorphism exists between two spaces we call the spaces isomorphic, and denote $X \approx Y$. $T$ is called an isometry, or isometric isomorphism, if $\|Tx\| = \|x\|$ for all $x \in X$, and we denote $X \simeq Y$ for isometrically isomorphic spaces. We generally don’t distinguish such spaces, and simply refer to them as equal, although sometimes specify the identification rule, i.e. $T : X \to Y$, between their elements.

Exercise 1.20. Show that $\ell_1^2 \simeq \ell_2^2$, but $\ell_1^n \not\simeq \ell_\infty^n$ for all $n \geq 3$.

Let us observe that equivalence between two norms introduced in Section 1.4 is in the new terms equivalent to the identity $i : (X,\|\cdot\|) \to (X,\|\cdot\|)$ being an isomorphism.

We say that $T : X \to Y$ is an isomorphic embedding if (10) holds without subjectivity assumption. In this case, the image of $T$ is closed, and Im $T \approx X$.

1.6. The factor-map and criterion of compactness of $B(X)$. Let $Y$ be a closed proper subspace of a normed space $X$. Let us consider the factor-map $J : X \to X/Y$ defined by the rule
\[ Jx = [x] . \]
From the definition of the factor-norm, it is clear that $\|Jx\| \leq \|x\|$, thus making $J$ a contraction map. To show that in fact $\|J\| = 1$, let us fix one $x \in X$ not in $Y$. Then $\|[x]\| > 0$. For a fixed $\varepsilon > 0$, let us find $y \in Y$ such that
\[ \|[x]\| \leq \|x + y\| \leq \|[x]\| + \varepsilon . \]
Consider the normalized unit vector $\frac{x + y}{\|x + y\|}$. Then
\[ \|J(\frac{x + y}{\|x + y\|})\| = \frac{\|[x]\|}{\|x + y\|} \geq 1 - \frac{\varepsilon}{\|[x]\|} . \]
Since, $x$ is fixed and $\varepsilon$ is arbitrary, we obtain $\|J\| = 1$. This observation shows that for any closed proper subspace $Y$ we can find a vector on the unit sphere $S(X)$ which is almost a distance 1 away from $Y$. This is more than enough to prove the following theorem.

Theorem 1.21. The unit ball of a normed space $(X,\|\cdot\|_X)$ is compact if and only if $\dim X < \infty$. 
Proof. We have already seen in Section 1.4 that the unit ball of a finite dimensional spaces is compact. Let us assume now that dim $X = \infty$ and show that the ball is not compact. It suffices to construct a separated sequence of vectors $x_1, x_2, \ldots$ so that all $\|x_n\| = 1$ and $\|x_n - x_m\| \geq 1/2$. Indeed, any such sequence would not contain a convergent subsequence. To this end, let us fix an arbitrary first vector $x_1 \in S(X)$. Consider the space $Y_1 = [x_1]$, and find $x_2 \in S(X)$ such that dist$(x_2, Y_1) > 1/2$. Then consider $Y_2 = [x_1, x_2]$ and find $x_3 \in S(X)$ with dist$(x_3, Y_2) > 1/2$, and so on. The process will never terminate since $X$ is not a span of finitely many vectors. It is easy to see that the obtained sequence is as desired.

If $T : X \to Y$ is a bounded operator and $X_0 = \ker T$, we construct a new operator $\tilde{T} : X/X_0 \to Y$ by the rule $\tilde{T} \circ J = T$. One can easily check that this definition is not ambiguous. Moreover, one has $\|T\| \leq \|\tilde{T}\|$. And on the other hand, if $\|\tilde{T}[x]\| > \|\tilde{T}\| - \varepsilon$ and $\|x\| < 1$, then for some $x_0 \in X_0$, $\|x + x_0\| < 1$, and yet $\tilde{T}[x + x_0] = T(x + x_0)$. This shows the opposite inequality $\|T\| \leq \|\tilde{T}\|$. Thus, $\|\tilde{T}\| = \|T\|$. Notice that the new operator has trivial kernel.

1.7. Direct sums. Suppose $Z$ is a linear space, $X, Y \subset Z$ are subspaces such that $Z = X + Y$ and $X \cap Y$. This is equivalent to the statement that for every $z \in Z$ there exists a unique couple of vectors $x \in X$, $y \in Y$ such that $z = x + y$. We thus have two linear maps $Pz = x$, $Qz = y$, so that $P + Q = Id$, called projections. Suppose now, that $Z$ is normed. It is not given that the operators $P, Q$ are bounded. We say that $Z$ is a direct sum of $X$ and $Y$, and write $Z = X \oplus Y$, if $P$, or equivalently, $Q$, is bounded. The summands of the direct sum are necessarily closed subspaces since $X = \ker Q$ and $Y = \ker P$. A closed subspace $X \subset Z$ is called complemented if there is a $Y \subset Z$ such that $Z = X \oplus Y$. In this case $Y$ is called a complement of $X$, and a complement is never unique. What defines $Y$ uniquely is the projection $P : Z \to X$. We say that $P$ is the projection onto $X$ along $Y$.

**Theorem 1.22.** An algebraic sum $Z = X + Y$ is direct if and only if dist$(S(X), S(Y)) > 0$.

Proof. Suppose he projection $P : Z \to X$ is bounded. Let $x \in S(X)$ and $y \in S(Y)$. Then $\|x - y\| \geq \|P^{-1}\|P(x - y)\| = \|P^{-1}\|\|x\| = \|P\|^{-1}\|P\|\|x\|$. Thus, dist$(S(X), S(Y)) \geq \|P\|^{-1}$.

Suppose now that dist$(S(X), S(Y)) > 0$, and yet $P$ is not bounded. It implies that there exists a sequence $x_n + y_n \in S(Z)$ such that $\|x_n\| \to \infty$. By the triangle inequality, $\|x_n\| \leq \|y_n\| \leq \|x_n\| + 1$. Thus, $\frac{y_n}{\|x_n\|} \to 1$. We have $\left|\frac{x_n}{\|x_n\|} + \frac{y_n}{\|x_n\|}\right| \to 0$. On the other hand,

\[
\left|\frac{x_n}{\|x_n\|} + \frac{y_n}{\|x_n\|}\right| = \frac{x_n}{\|x_n\|} + \frac{y_n}{\|y_n\|} = \frac{y_n}{\|x_n\|} \left(\frac{\|y_n\|}{\|x_n\|} - 1\right)
\]

In view of all of the above, $\left|\frac{x_n}{\|x_n\|} + \frac{y_n}{\|y_n\|}\right| \to 0$, in contradiction with our assumption.

A subspace $X \subset Z$ is called finite co-dimensional if $Z = X + Y$ for some $Y \subset Z$ with dim $Y < \infty$.

**Corollary 1.23.** Every finite co-dimensional closed subspace is complemented.

Indeed, let $Z = X + Y$, $X \cap Y$, $X$ closed, and dim $Y < \infty$. If the spheres of $X$ and $Y$ are not separated, then there exist sequences $y_n \in S(Y)$ and $x_n \in S(X)$ such that $\|x_n - y_n\| \to 0$. By compactness we can choose a subsequence $y_{n_k} \to y \in S(Y)$. Thus, $x_n \to y$ as well, which by the closedness of $S(X)$ implies $y \in X$, a contradiction.
Exercise 1.24. Show that the norm of any projection operator is at least 1.

Exercise 1.25. Prove that a bounded operator $P : X \to X$ is a projection onto a subspace if and only if it is idempotent, $P^2 = P$.

Exercise 1.26. Show that if $Z = X \oplus Y$, then $Z/X \approx Y$. Hint: consider the projector $P : Z \to Y$ along $X$, and its factor by the kernel $\tilde{P} : Z/X \to Y$.

1.8. Completion and extension by continuity.

1.9. The dual space. If the target space of a linear operator is the scalar field, the operator is called a linear functional. The space of linear functionals on $X$ is denoted $X^*$, while the space of all linear bounded functionals by $X^*$, called the dual space. It is often possible to identify the dual of a Banach space up to an isometry.

Example 1.27. sdfdfsdfs

Exercise 1.28. Show that if $X = X_1 \oplus_p \ldots \oplus_p X_n$, then $X^* = X_1^* \oplus_q \ldots \oplus_q X_n^*$, where $p,q$ are conjugates. Show that the same is true for infinite $\ell_p$-sums if $1 \leq p < \infty$.

1.10. Structure of linear functionals. Suppose $f \in X^\setminus \{0\}$, and let $x_0 \in X$ be a vector such that $f(x_0) \neq 0$. Then let $x \in X$ an arbitrary vector. Define $y = x - \frac{f(x)}{f(x_0)} x_0$. Then clearly, $f(y) = 0$. It shows that for any $x$ there exist unique $y \in \ker(f)$ and $\lambda \in \mathbb{R}$ such that

$$x = \lambda x_0 + y.$$ 

In particular it shows that the kernel of $f$ is one co-dimensional. We will show that the distinction between bounded and unbounded functionals comes in the condition of closeness of the kernel, or even less restrictively, its density.

Lemma 1.29. Let $f \in X^\setminus \{0\}$. The following are equivalent:

(a) $f \in X^*$;
(b) $\ker(f)$ is closed;
(c) $\ker(f)$ is not dense in $X$.

Proof. The implications $(a) \Rightarrow (b) \Rightarrow (c)$ are trivial. Suppose that $\ker(f)$ is not dense. Then for some ball $B_r(x_0) \cap \ker(f) = \emptyset$. Let $y \in S(X)$. Then $g(t) = f(x_0 + ty) = f(x_0) + tf(y)$ is a continuous non-vanishing function on $(-r,r)$. This implies that the sign of it has to agree with that of $f(x_0)$. Assuming $f(x_0) > 0$ we then have $f(x_0) + tf(y) > 0$, and so, $|f(y)| \leq r^{-1} f(x_0)$. This shows that $f$ is bounded on the unit sphere and completes the proof. □

Geometrically linear bounded functionals can be identified with affine hyperplanes. Thus, if $f \in X^*$, then $H(f) = \{ x \in X : f(x) = 1 \}$, defines $f$ uniquely. If $f \in S(X^*)$, then the hyperplane is in some sense tangent to the unit sphere of $X$, namely, it does not deep inside the interior of the unit ball and it approaches arbitrarily close to $S(X)$. Note that a functional may not necessarily attain its highest values on the sphere, i.e. the norm. For example, let $X = \ell_1$, and $f = (1/2, 2/3, 3/4, \ldots) \in S(\ell_\infty)$. There is no sequence $x \in S(\ell_1)$ for which $f(x) = 1$. If $x \in S(X)$ and $f \in S(X^*)$ with $f(x) = 1$, then $f$ is called a supporting functional of $x$. Existence of supporting functionals is not immediately obvious, and it brings us to an even more fundamental question – does there exist at least one non-zero bounded linear functional on a given normed space?
2. Fundamental Principles

2.1. The Hahn-Banach extension theorem. The essence of the Hahn-Banach extension theorem is to show that a given bounded functional defined on a linear subspace of $X$ can be extended boundedly to the whole space $X$ retaining not only its boundedness but also its norm. The boundedness can be expressed as the condition of domination by the norm-function, i.e. if $Y \subset X$ and $f \in Y'$ then $f \in Y^*$ if and only if

$$f(y) \leq C\|y\|,$$

for some $C > 0$. We will in fact need a more general extension result that will be useful when establishing separation theorems later in Section 2.4. We thus consider a positively homogeneous convex functional $p : X \to \mathbb{R} \cup \{\infty\}$, which means that $p(tx) = tp(x)$ for all $x \in X$ and $t \geq 0$, and $p(\lambda x + (1 - \lambda)y) \leq \lambda p(x) + (1 - \lambda)p(y)$, for all $0 < \lambda < 1$, $x, y \in X$. The latter is equivalent to the triangle inequality, $p(x + y) \leq p(x) + p(y)$. Note that a norm, or a quasi-norm, is an example of such a functional. We say that $p$ dominates $f$ on $Y$ if $f(y) \leq p(y)$ for all $y \in Y$.

Theorem 2.1 (Hahn-Banach extension theorem). Suppose $Y \subset X$, and $p$ is a positively homogeneous convex functional defined on $X$. Then every linear functional $f \in Y'$ dominated by $p$ on $Y$ can be extended to a linear functional $f \in X'$ dominated by the same $p$ on all of $X$.

In the core of the proof lies Zorn’s Lemma, which we recall briefly. Let $P$ be a partially ordered set. A subset $C$ of $P$ is called a chain if its every two elements are comparable, i.e. $\forall a, b \in C$ either $a \leq b$ or $b \leq a$. An upper bound for a set $A \subset P$ is an element $b \in P$ such that $b \geq a$, for all $a \in A$. A maximal element $m$ is an element with the property that if $a \geq m$, then $a = m$. Generally, it may not be unique.

Lemma 2.2 (Zorn’s Lemma). If every chain of $P$ has an upper bound, then $P$ contains a maximal element.

Zorn’s lemma is equivalent to the Axiom of Choice.

Proof of Theorem 2.1. Let us introduce the set of pairs $P = \{(f, Y) : f \in Y', f \leq p\}$ ordered by $(f_1, Y_1) \leq (f_2, Y_2)$ iff $Y_1 \subset Y_2$ and $f_2|_{Y_1} = f_1$. Let $C \subset P$ be a chain. Define $Y = \cup_{(f, Y) \in C} Y$, and let $\tilde{f}(y) = f(y)$ if $y \in Y$. This defines an upper bound of $C$. By Zorn’s Lemma there exists a maximal element $m = (f_0, Y_0) \in P$. Let us show that $Y_0 = X$. Indeed, if not, then there is a vector $x_0 \in X \setminus Y_0$. Thus, for every element $x \in Z = [x_0, Y]$ there exist unique $\lambda \in \mathbb{R}$ and $y \in Y_0$ such that $x = \lambda x_0 + y$. We construct an extension $f$ of $f_0$ to $Z$ so that $(f, Z) \in P$ and run into contradiction with the maximality of $m$. In order to do that it suffices to find a value of $f$ only on $x_0$. Let $c = f(x_0)$, then to ensure that $f$ is still dominated by $p$ we need to satisfy

$$\lambda c + f_0(y) \leq p(\lambda x_0 + y).$$

For $\lambda > 0$ this is equivalent to

$$(11) \quad c \leq p(x_0 + y') - f_0(y'),$$

and for $\lambda < 0$ to

$$(12) \quad c \geq f_0(y'') - p(-x_0 + y'').$$
In order for such a $c$ to exist one has to make sure that any number on the right hand side of (11) is no less than any number on the right hand side of (12), i.e.

$$p(x_0 + y') - f_0(y') \geq f_0(y'') - p(-x_0 + y''),$$

for all $y', y'' \in Y$. This is true indeed, since in view of the convexity and dominance, we have

$$p(x_0 + y') + p(-x_0 + y'') \geq p(y' + y'') \geq f(y' + y'') = f(y') + f(y'').$$

Let us discuss some immediate consequences of the Hahn-Banach theorem. First, every vector on a normed space $(X, \| \cdot \|)$ has a supporting functional. Indeed, let $x \in X$, define $f \in [x]^*$ by $f(\lambda x) = \lambda \| x \|$. Then $\| f \| = 1$, which means $f$ is dominated by the norm. The extension then has the same norm 1 and supports $x$.

For a normed space $X$, one can consider the dual of the dual space, $X^{**}$, called second dual. There is a canonical isometric embedding $i : X \hookrightarrow X^{**}$ defined as follows: $i(x)(x^*) = x^*(x)$. It is convenient to use parentheses to indicate action of a functional: $x^*(x) = (x^*, x)$. In this notation $(i(x), x^*) = (x^*, x)$ or simply, $(x, x^*) = (x^*, x)$. To show that $i$ is an isometry, notice that $\| (i(x), x^*) \| \leq \| x^* \|_X \| x \|_X$, thus $\| i(x) \|_{X^{**}} \leq \| x \|_X$. On the other hand, let $x^*$ be a supporting functional. Then $\| x^* \|_X = 1$, and $(i(x), x^*) = x^*(x) = \| x \|_X$. We will think of $X$ is a subspace of $X^{**}$ with the natural identification of elements described above. If the embedding $X \hookrightarrow X^{**}$ exhaust all elements of $X^{**}$, i.e. $X = X^{**}$, then $X$ is called reflexive. We will return to a discussion of reflexive spaces later as they possess very important compactness properties.

Let $T : X \rightarrow Y$ be a bounded operator. We can define the adjoint or dual operator $T^* : Y^* \rightarrow X^*$ by the rule $(T^*y^*, x) = (y^*, Tx)$. Again, using the Hahn-Banach theorem we show that

$$\| T \| = \| T^* \|.$$  

First, $| (T^*y^*, x) | \leq \| y^* \|_Y \| Tx \| \leq \| y^* \|_Y \| T \| \| x \|_Y$. This shows $\| T^* \| \leq \| T \|$. Let now $\varepsilon > 0$ be given. Find $x \in S(X)$ such that $\| Tx \| \geq \| T \| - \varepsilon$. Then let $y^* \in S(Y^*)$ be a supporting functional for $Tx$. We have $(T^*y^*, x) = (y^*, Tx) = \| Tx \| \geq \| T \| - \varepsilon$. This shows the opposite inequality.

**Exercise 2.3.** Let $T^{**} : X^{**} \rightarrow Y^{**}$ be the second adjoint operator, i.e. $T^{**} = (T^*)^*$. Show that $T^{**}|_X = T$.

**Exercise 2.4.** Show that $X^*$ is always complemented in $X^{***}$. Hint: consider the adjoint $i^* : X^{***} \rightarrow X^*$.

**Exercise 2.5.** Prove that $T : X \rightarrow Y$ is an isomorphism if and only if $T^* : Y^* \rightarrow X^*$ is.

**Exercise 2.6.** Show that $X$ is reflexive if and only if $X^*$ is reflexive.

**Exercise 2.7.** Let $Y \subset X$ be a closed subspace, and $X$ is Banach. Define $Y^\perp = \{ f \in X^* : f|_Y = 0 \}$. This is a closed subspace of $X^*$, called the annihilator of $Y$. Show that $Y^* \cong X^*/Y^\perp$, and $(X/Y)^* \cong Y^\perp$.

### 2.2. Convex sets.
We say that a set $A \subset X$ is convex if $x, y \in A$ implies $\lambda x + (1 - \lambda)y \in A$ for all $0 < \lambda < 1$, i.e. with every pair of points $A$ contains the interval connecting them. A direct consequence of homogeneity and triangle inequality of the norm is that any ball is
a convex set. For an arbitrary set $A \subset X$ we define the convex hull of $A$ as the set of all convex combinations of elements from $A$:

$$\text{conv } A = \left\{ \sum_{i=1}^{n} \lambda_i a_i : a_i \in A, \sum_{i=1}^{n} \lambda_i = 1, \lambda_i \geq 0, n \in \mathbb{N} \right\}. $$

It is the smallest convex subset of $X$ containing $A$, or equivalently, the intersection

$$\text{conv } A = \bigcap_{A \subset C, C \text{ convex}} C.$$ 

The topological closure of the convex hull $\overline{\text{conv } A}$ is the same as the smallest closed convex set containing $A$, or the intersection of such sets.

**Theorem 2.8 (Caratheodori).** Let $A \subset \mathbb{R}^n$, then every point $a \in \text{conv } A$ can be represented as a convex combination of at most $n+1$ elements in $A$.

**Proof.** Suppose $x = \sum_{i=1}^{N} \lambda_i a_i$, all $\lambda_i > 0$, $\sum \lambda_i = 1$, and $N > n + 1$. We will find a way to introduce a correction into the convex combination above as to reduce the number of elements in the sum by 1. Then the proof follows by iteration.

First, let us observe that since $N > n + 1$, the number of elements in the family $a_2 - a_1, a_3 - a_1, \ldots, a_N - a_1$ is larger than the dimension and hence they are not linearly independent. So, we can find constants $t_i \in \mathbb{R}$, not all of which are zero, such that $\sum_{i=2}^{N} t_i(a_i - a_1) = 0$. Denoting $t_1 = -\sum t_i$, we can write

$$\sum_{i=1}^{N} t_i a_i = 0.$$

By reversing the sign of all the $t_i$’s if necessary, we can assume that at least one of them is positive. We will now adjust the original convex combination by a constant multiple of the zero-sum above, thus not changing the $x$:

$$x = \sum_{i=1}^{N} \lambda_i a_i - \varepsilon \sum_{i=1}^{N} t_i a_i = \sum_{i=1}^{N} (\lambda_i - \varepsilon t_i) a_i.$$

Letting $\varepsilon = \min_{i>0} \{\lambda_i/t_i\}$ ensures that $\mu_i = \lambda_i - \varepsilon t_i \geq 0$ for all $i$, and that for some $i_0$, $\mu_{i_0} = 0$. Yet, clearly, $\sum \mu_i = 1$. Thus, the new representation

$$x = \sum_{i=1}^{N} \mu_i a_i,$$

is at least one term shorter. \qed

**Corollary 2.9.** If $A \subset \mathbb{R}^n$ is closed, then $\text{conv } A$ is closed too.

Indeed, simply use the previous theorem and pass to nested subsequences in all $n + 1$ terms by compactness. In the infinite dimensions closeness or even compactness of $A$ is not sufficient to conclude that $\text{conv } A$ is automatically closed. Let us consider the following example. Let $X = \ell_2$, and $A = \{ \frac{1}{n} e_n \} \cup \{0\}$. It is easy to see that $A$ is compact. Any element of $\text{conv } A$ has only finitely many non-zero entries, yet $x = \sum_{n=1}^{\infty} \frac{1}{2^n} e_n \in \text{conv } A$. 
2.3. Minkowski’s functionals. Let us recall from Section 2.4 that a function \( p : X \rightarrow \mathbb{R} \cup \{\infty\} \) is called convex if for any \( x, y \in X \) one has \( p(\lambda x + (1 - \lambda)y) \leq \lambda p(x) + (1 - \lambda)p(y) \) for all \( 0 < \lambda < 1 \). A function \( q : X \rightarrow \mathbb{R} \cup \{-\infty\} \) is called concave if for any \( x, y \in X \) one has \( q(\lambda x + (1 - \lambda)y) \geq \lambda q(x) + (1 - \lambda)q(y) \) for all \( 0 < \lambda < 1 \). Is it easy to see that if \( p \) is convex then the sub-level sets \( \{ p \leq p_0 \} \) are convex, and if \( q \) is concave then the super-level sets \( \{ q \geq q_0 \} \) are convex.

Suppose \( A \subset X \) is convex and \( 0 \in A \). We associate to \( A \) a convex function, called Minkowski’s functional, \( p_A \) so that \( A \) is “almost” given as a sub level set of \( p_A \). We define \( p_A(x) \) as follows. Suppose there is no \( t \geq 0 \) for which \( x \in tA \), then \( p_A(x) = \infty \). If \( x \in tA \) for some \( t \geq 0 \), we set

\[
p_A(x) = \inf\{t \geq 0 : x \in tA\}.
\]

We list the basic properties of the Minkowski’s functional.

(a) \( p_A \) is positively homogeneous and convex;
(b) \( \{ p_A < 1 \} \subset A \subset \{ p_A \leq 1 \} \).

For \( \alpha > 0 \), \( x \in tA \) if and only if \( \alpha x \in \alpha tA \). This readily implies (a). Notice that for positively homogeneous functionals convexity is equivalent to triangle inequality, \( p_A(x + y) \leq p_A(x) + p_A(y) \). So, let \( x, y \in X \). If any of \( p_A(x) \) or \( p_A(y) \) equal \( \infty \), the inequality becomes trivial. If both are finite, then for every \( \varepsilon > 0 \) we can find \( t_1 < p_A(x) + \varepsilon \) and \( t_2 < p_A(y) + \varepsilon \) such that \( x \in t_1A \) and \( y \in t_2A \). Then

\[
x + y \in t_1A + t_2A = (t_1 + t_2)\left( \frac{t_1}{t_1 + t_2}A + \frac{t_2}{t_1 + t_2}A \right) \subset (t_1 + t_2)A.
\]

This shows that \( p_A(x + y) \leq p_A(x) + p_A(y) + \varepsilon \), for all \( \varepsilon > 0 \). Finally, (c) follows directly from the definition and that \( 0 \in A \).

Suppose now \( B \) is another convex set not containing a small ball around the origin, i.e. there is \( \delta > 0 \) such that \( \delta B(X) \cap B = \emptyset \). We can associate a similar, but now concave functional to \( B \) as follows. If \( x \in tB \), for no \( t \geq 0 \), then \( q_B(x) = -\infty \). Otherwise, we define

\[
q_B(x) = \sup\{t \geq 0 : x \in tB\}.
\]

Condition \( \delta B(X) \cap B = \emptyset \) warrants that the supremum is finite for any \( x \in X \). The following list of properties can be established in a similar fashion:

(a) \( q_B \) is positively homogeneous and concave;
(b) \( \{ q_B > 1 \} \subset B \subset \{ q_B \geq 1 \} \).

Suppose now that we have two disjoint convex sets \( A \) and \( B \) satisfying all the assumptions above, and let \( p_A \) and \( q_B \) be the corresponding Minkowski’s functionals. If \( p_A(x) < \infty \), let \( t \geq 0 \) be such that \( x \in tA \). Since \( 0 \in A \), the whole interval \( [0, x] \) is in \( tA \) and therefore not in \( tB \). This in turn implies that \( x \notin sB \) for any \( s \geq t \), for if such \( s \) existed, then \( \frac{t}{s}x \in tB \) contradicting the previous. As a consequence, \( q_B(x) \leq t \). We have shown that

\[
q_B(x) \leq p_A(x), \text{ for all } x \in X.
\]

2.4. Separation theorems.

**Theorem 2.10** (Generalized Hahn-Banach theorem). Let \( p \) be convex, and \( q \) concave functionals defined on \( X \). Let \( Y \subset X \), \( f \in Y' \) such that

\[
f(y) \leq p(x + y) - q(x), \text{ for all } y \in Y, x \in X.
\]
Then \( f \) can be extended to all of \( X \), \( \tilde{f} \in X' \), satisfying

\[
q(x) \leq \tilde{f}(x) \leq p(x), \quad \text{for all } x \in X.
\]

**Proof.** The proof goes exactly the same way as before. We only need to check that if \( Y \subset X \), and \( x_0 \in X \setminus Y \), then we can extend \( Y \) to \( Z = [x_0, Y] \) preserving the domination property (14). If \( c = f(x_0) \), then we need

\[
\lambda c + f(y) \leq p(x + \lambda x_0 + y) - q(x),
\]

for all \( x \in X \) and \( y \in Y \) and \( \lambda \in \mathbb{R} \). Again, for \( \lambda > 0 \) this is equivalent to

\[
c \leq p(x' + x_0 + y') - q(x') - f(y'),
\]

while for \( \lambda < 0 \),

\[
c \geq f(y'') - p(x'' - x_0 + y'') + q(x'').
\]

The existence of \( c \) is ensured if

\[
p(x' + x_0 + y') - q(x') - f(y') \geq f(y'') - p(x'' - x_0 + y'') + q(x''),
\]

which is true since

\[
p(x' + x_0 + y') + p(x'' - x_0 + y'') - q(x') - q(x'') \geq p(x' + x'' + y' + y'') - q(x' + x'') \geq f(y' + y'').
\]

\[\square\]

**Theorem 2.11** (Separation Theorems). Let \( A, B \) be two disjoint convex subsets of a normed space \( X \).

(i) If \( A \not\subset \overline{B} \), then there exists \( f \in X^* \setminus \{0\} \) such that

\[
\sup f(A) \leq \inf f(B).
\]

(ii) If \( A \) has a non-empty interior, then there exists \( f \in X^* \setminus \{0\} \) such that (16) holds.

(iii) If \( A = \{x_0\} \) and \( B \) is closed then there exists \( f \in X^* \setminus \{0\} \) such that

\[
f(x_0) < \inf f(B).
\]

**Proof.** Suppose \( A \not\subset \overline{B} \), then there exists \( x_0 \in A \) and \( \delta > 0 \) such that \( B_\delta(x_0) \cap B = \emptyset \). By moving \( x_0 \) to the origin we satisfy all the conditions on \( A \) and \( B \) as above, which allows us to define Minkowski’s functions, \( q_B \leq p_A \). Thus, we have (14) for \( f = 0 \), and \( Y = \{0\} \). By Theorem 2.10, there exists and extension \( f \) for which (15) holds, and thus, in view of (13), \( f \) separates \( A \) and \( B \).

If \( A \) has a non-empty interior then clearly (i) holds. Let us assume that \( \varepsilon B(X) \subset A \) and let \( f \) be the functional constructed above. Since \( f(x) \leq p_A(x) \), we conclude that whenever \( \|x\| \leq \varepsilon \), then \( f(x) \leq 1 \). This shows \( \|f\| \leq \varepsilon^{-1} \).

Finally if \( A = \{x_0\} \) and \( B \) is closed, we apply case (ii) to \( A = B_\delta(x_0) \) for small \( \delta > 0 \). Since \( f \neq 0 \), there is \( y \in S(X) \) for which \( f(y) > 0 \). Thus, \( f(x_0) < f(x_0) + \varepsilon f(y) \leq \inf f(B) \). \[\square\]

**Exercise 2.12.** If a strict inequality holds in (16), then \( A \) and \( B \) are called **strictly separated.** Show that if \( A \) is compact and \( B \) closed convex disjoint sets, then they can be strictly separated.

The condition of (i) is not sufficient to guarantee that a bounded separator would exist. Let us consider the following example. Let \( X = \ell_{2,0} \) be the linear space of finite sequences
Lectures on Functional Analysis 15

endowed with the $\ell_2$-norm, $A = \text{conv}\{\vec{e}_n\}_n$, and $B = \frac{1}{2}A$. These are convex disjoint sets. Notice that for any $b \in B$, let $b = \frac{1}{2}\sum \lambda_i \vec{e}_i$, we have

$$
\|b\|_2^2 = \frac{1}{4} \sum |\lambda_i|^2 \leq \frac{1}{4} \sum \lambda_i = \frac{1}{4}.
$$

Thus, $B \subset \frac{1}{2}B(X)$. Since any $\vec{e}_n \notin \frac{1}{2}B(X)$, the conditions of Theorem 2.11 (i) are satisfied. Next observe that $0 \in \overline{A} \cap \overline{B}$, by considering the sequences $x_n = (\vec{e}_1 + \cdots + \vec{e}_n)/n$ and $y_n = \frac{1}{2}x_n$. Suppose $f \in \ell_2/\{0\}$ separates the two sets, i.e. $\sup f(A) \leq c \leq \inf f(B)$. Since $0$ is in the closure of both sets, we conclude that $c = 0$. Then $f$ as a sequence is positive on the one hand and negative on the other hand. Thus, $f = 0$, a contradiction. It is easy to construct an unbounded separator though by taking $f = (1, 1, \ldots)$.

**Corollary 2.13.** Let $S \subset X$ be a subset of a normed space. Then

$$
\text{conv } S = \bigcap_{f \in X^*} \{x : f(x) \leq \sup f(S)\}.
$$

Indeed, the inclusion $\subset$ is obvious. If however $x_0 \notin \text{conv } S$, then Theorem 2.11 (iii) provides a functional such that $f(x_0) < \inf f(\text{conv } S) \leq \inf f(S)$. Reversing the sign of $f$ shows that $x_0$ is not in one of the sets in the intersection.

2.5. **Baire Category Theorem.** Let us consider for a moment a general complete metric space $X$, without necessarily a given linear structure. Let us ask ourselves how "big" such a space can be. One may say, a singleton $X = \{x_0\}$ is an obvious example of a "small" complete metric space. Well, it is not actually that small compared to its own standards. After all that one point is closed and open, and it is in fact a ball of any radius centered around itself. It therefore is rather "big", in a sense. To make this discussion more precise let us agree on what we mean by a "small" subset of $X$. We say that $F$ is nowhere dense in $X$ is for any open set $U$ there exists an open subset $V \subset U$ with no intersection with $F$. In other words, $\overline{F}$ has empty interior. A subset $F \subset X$ is called of 1st Baire category if $F = \bigcup_{i=1}^\infty F_i$, where all $F_i$’s are nowhere dense. A subset is of 2nd Baire category if it is not of the 1st category. Thus, in the example above $X$ itself is clearly of the 2nd category. This in fact holds in general.

**Theorem 2.14** (Baire Category Theorem). *Any complete metric space is a 2nd Baire category set.*

**Proof.** Let us suppose, on the contrary, that $X = \bigcup F_i$, and all $F_i$’s are nowhere dense. Then, there is a closed ball $B_{\varepsilon_1}(x_1)$ with $\varepsilon_1 < 1$ disjoint from $F_1$. Since $F_2$ is however dense, there is a ball $B_{\varepsilon_2}(x_2) \subset D_{\varepsilon_1}(x_1)$, with $\varepsilon_2 < 1/2$, disjoint from $F_2$, continuing in the same manner we find a sequence of nested closed balls $B_{\varepsilon_n}(x_n)$, with $\varepsilon_n < 1/n$. Clearly, $d(x_n, x_m) < 1/m$, for all $n > m$. Thus, the sequence $\{x_n\}$ is Cauchy. By completeness, there exists a limit $x$, which belongs to all the balls, and hence not in any $F_j$’s, a contradiction.

There are many consequences of the Baire category theorem, some of which are given in the exercises below.

**Theorem 2.15** (Banach-Steinhauss uniform boundedness principle). *Let $\mathcal{F} \subset \mathcal{L}(X,Y)$ be a family of bounded operators, and $X$ is a Banach space. Suppose for any $x \in X$, $\sup_{T \in \mathcal{F}} \|Tx\| < \infty$. Then, the family is uniformly bounded, i.e. $\sup_{T \in \mathcal{F}} \|T\| < \infty$. 
Proof. Let \( F_n = \{ x \in X : \sup_{T \in \mathcal{F}} \|Tx\| \leq n \} \). Not that each set \( F_n \) is closed, and by assumption \( X = \bigcup_n F_n \). Hence, by the Baire Category Theorem, one of \( F_n \)'s contains a ball \( B_r(x_0) \). This implies that for all \( x \in B(X), \|T(x_0 + rx)\| \leq n, \) for all \( T \in \mathcal{F} \). Thus, \( \|Tx\| \leq r^{-1}(n + \sup_{T \in \mathcal{F}} \|Tx_0\|) \), implying that desired result.

\( \square \)

**Corollary 2.16.** Let \( S \subset X \) be a subset such that for every \( x^* \in X^* \), sup \( |x^*(S)| < \infty \). Then \( S \) is bounded.

Indeed, if viewed as a subset of \( X^{**} \), \( S \) is a pointwise bounded family of operators. Hence, it is norm-bounded by the Banach-Steinhauss theorem.

2.6. Open mapping theorem.

**Lemma 2.17.** Suppose \( T : X \to Y \) is bounded, and \( X \) is a Banach space. Suppose that \( D(Y) \subset \overline{T(D(X))} \), then \( \frac{1}{2}D(Y) \subset T(D(X)) \).

**Proof.** Let us note, by linearity of \( T \), that

\[
(18) \quad rD(Y) \subset \overline{T(rD(X))}
\]

for any \( r > 0 \). Let us fix \( y \in D_{1/2}(Y) \) and let us fix a small \( \varepsilon > 0 \) to be specified later. By (18) we can find \( x_1 \in 3/4D(X) \) such that \( \|y - y_1\| < \varepsilon \), where \( Tx_1 = y_1 \). Since \( y - y_1 \in \varepsilon D(Y) \), one finds \( x_2 \in \varepsilon D(Y) \) such that \( \|y - y_1 - y_2\| < \varepsilon/2 \), where \( y_2 = Tx_2 \). Continuing this way, we construct a sequence \( \{x_n\} \) with \( \|x_n\| \leq \varepsilon/2^n \). Let \( x = \sum_n x_n \). Then by construction \( Tx = y \), and \( \|x\| \leq 3/4 + 2\varepsilon < 1 \), provided \( \varepsilon \) is small. \( \square \)

**Theorem 2.18** (Open mapping theorem). Suppose \( T \in \mathcal{L}(X,Y) \), and both spaces are Banach. If, in addition, \( T \) is surjective, then \( T \) is an open mapping, i.e. \( T(U) \) is open for every open \( U \).

**Proof.** Suppose, \( U \) is open. Let \( x_0 \in U \), and let \( D_\varepsilon(x_0) \subset U \). We prove the theorem if we show that \( T(D_\varepsilon(x_0)) \) contains an open neighborhood of \( Tx_0 \). Since, \( T(D_\varepsilon(x_0)) = Tx_0 + \varepsilon T(D(X)) \), it amounts to showing that \( T(D(X)) \) contains a ball centered at the origin. Since \( T \) is surjective, we have \( Y = \bigcup_n nT(D(X)) \). By the Baire Category Theorem, one of the sets \( nT(D(X)) \) is dense in some ball \( D_\delta(y_0) \), and hence, by linearity, so is \( T(D(X)) \). Since \( T(D(X)) \) is convex and symmetric with respect to the origin,

\[
\delta D(Y) \subset \text{conv}\{D_\delta(y_0), D_\delta(-y_0)\} \subset \overline{T(D(X))}
\]

Applying Lemma 2.17 to the operator \( \frac{1}{\delta}T \), we conclude \( \frac{\delta}{2}D(Y) \subset T(D(X)) \) and the proof is finished. \( \square \)

As a direct consequence of the open mapping theorem we obtain the following.

**Corollary 2.19.** Suppose \( T \in \mathcal{L}(X,Y) \) is bijective. Then \( T^{-1} \) is automatically bounded.

**Corollary 2.20.** Two norms on a Banach space are either equivalent or incomparable.

Indeed, if \( C\|\cdot\|_1 \geq \|\cdot\|_2 \), then the identity operator \( i : (X,\|\cdot\|_1) \to (X,\|\cdot\|_2) \) is bounded. Hence, the inverse is bounded, which establishes the equivalence.

**Theorem 2.21** (Closed graph theorem). Suppose \( T : X \to Y \) is a linear operator such that if \( x_n \to x \) and \( Tx_n \to y \), then \( Tx = y \). Then \( T \) is bounded.
Proof. Let $\Gamma = \{(x, Tx) : x \in X\}$ denote the graph of $T$. Thus, by the assumption, $\Gamma$ is closed in the product topology of the product $X \times Y$. Since it is also a linear subset, $\Gamma$ as a closed subspace of the $l_1$-sum, $X \otimes_1 Y$ is a Banach spaces. Let us consider the projection onto $X$ restricted to $\Gamma$, $P : \Gamma \rightarrow X$, given by $P(x, Tx) = x$. The operator is clearly bounded and easy to show that $P$ is bijective. Hence, $P^{-1}$ is bound, which implies
\[ \|x\| + \|Tx\| \leq C\|x\|, \]
and thus $T$ is bounded. \qed

3. Weak topologies

3.1. Topology of not necessarily metrizable spaces. Let $(X, \tau)$ be a topological space with topology of open sets $\tau$. A subset $\{x_\alpha\}_{\alpha \in A} \subset X$ is called a net, if the index set $A$ is partially ordered and directed, i.e. for every pair $\alpha, \beta \in A$ there is $\gamma \in A$ with $\gamma \geq \alpha, \gamma \geq \beta$. A subnet is a net $\{y_\beta\}_{\beta \in B}$ with a map $n : A \rightarrow B$ such that $y_\beta = x_{n(\beta)}$, $n$ is monotone, and for every $\alpha \in A$ there is $\beta \in B$ with $n(\beta) \geq \alpha$. A net $\{x_\alpha\}_{\alpha \in A}$ is said to be convergent to $x \in X$ if for every open neighborhood $U$ of $x$, there is $\alpha_0 \in A$ such that $x_\alpha \in U$ for all $\alpha \geq \alpha_0$. A function $f : X \rightarrow Y$, where $Y$ is another topological space, is cooled continuous if $f^{-1}(G) \in \tau$ for any open $G \subset Y$.

Lemma 3.1. A function $f : X \rightarrow Y$, is continuous if and only if for any convergent net $\lim_{\alpha \in A} x_\alpha = x$, $\lim_{\alpha \in A} f(x_\alpha) = f(x)$.

Proof. Suppose $f$ is continuous, and let $\lim_{\alpha \in A} x_\alpha = x$. For any open $G$ containing $f(x)$, $f^{-1}(G)$ is open and contains $x$. Since eventually all $x_\alpha$ are in $f^{-1}(G)$, then all $f(x_\alpha)$ will be in $G$.

Conversely, suppose there is open $G \subset Y$ such that $f^{-1}(G)$ is not open. Thus, there is a point $x_0 \in f^{-1}(G)$ such that any open neighborhood $U$ of $x$ contains a point outside $f^{-1}(G)$. Let us fix one such point $x_U$ for every $U$. Let $A = \{U \in \tau : U \text{ open, } x \in U\}$. It is a net ordered by inclusion. Clearly, $x_U \rightarrow x$, since for every $U$ containing $x$, all elements of the net, namely starting from $x_U$, will fall inside $U$. Yet, $f(x_U) \notin G$, and thus $f(x_U) \not\rightarrow f(x)$. \qed

Exercise 3.2. Show that a subset $F \subset X$ is closed if and only if the limit of every convergent net inside $F$ is contained in $F$.

We say that $X$ is compact if every open cover of $X$ contains a finite sub cover.

Lemma 3.3. $X$ is compact if and only if every net contains a convergent subnet.

Proof. Suppose $X$ is compact, and let $\{x_\alpha\}_{\alpha \in A} \subset X$ be a net. First we let us establish existence of a cluster point. A point $y \in X$ is a cluster point of a net if for every $U \in \tau$ containing $y$ and every $\alpha_0$, there is $\alpha \geq \alpha_0$ such that $x_\alpha \in U$. Suppose that our net does not have cluster points. Thus, for every $y \in X$ there is $U_y$ and $\alpha_y \in A$ such that $x_\alpha \notin U_y$ for all $\alpha \geq \alpha_y$. Consider the open cover $\{U_y\}_{y \in X}$. By compactness there is a finite sub cover $U_{y_1}, \ldots, U_{y_n}$. Since $A$ is a net, there is a $\alpha \geq \alpha_{y_i}$ for all $i = 1, \ldots, n$. Then $x_\alpha$ is in none of the open sets above, which shows that they don’t form a cover.

So, let $y$ be a cluster point. Let $B = \{(U, \alpha) : y \in U, U \in \tau, x_\alpha \in U\}$ be ordered by reverse inclusion on the first component, and by the order of $A$ on the second. For $\beta = (U, \alpha)$, let $y_\beta = x_\alpha$, and let $n(\beta) = \alpha$. It is routine to show that $\{y_\beta\}_{\beta \in B}$ is a subnet converging to $y$.

Conversely, suppose every net has a converging subnet, and yet on the contrary, $X$ is not compact. This implies that there is an open cover $\mathcal{U}$ which has no finite subcover. Let us
define \( A = \{ \alpha = (U_1, \ldots, U_n) : U_i \in \mathcal{U}, n \in \mathbb{N} \} \) ordered by \( \alpha \succeq \beta \) if \( \beta \subset \alpha \). Clearly \( A \) is also directed. By assumption, for any \( \alpha = (U_1, \ldots, U_n) \) there is \( x_\alpha \notin \bigcup_i U_i \). The net \( \{ x_\alpha \}_{\alpha \in A} \) has a converging subnet \( \{ y_\beta \}_{\beta \in B} \), and \( y = \lim y_\beta \). Since \( \mathcal{U} \) is a cover, there is \( U \in \mathcal{U} \) with \( y \in U \). Let \( \alpha = (U) \). By the definition of a subnet, there is \( \beta' \in B \) such that \( n(\beta') \succeq \alpha \) and \( y_{\beta'} = x_{n(\beta')} \), and there is another \( \beta'' \succeq \beta' \) such that \( y_{\beta''} \in U \). By monotonicity of \( n \), \( n(\beta'') \succeq \alpha \), and yet \( x_{n(\beta'')} \in U \), in contradiction with the construction.

\[ \square \]

Exercise 3.4. A topology \( \tau_1 \) on \( X \) is said to be stronger than another topology \( \tau_2 \) on \( X \) if for any point \( x \in X \) any open neighborhood of \( x \) in \( \tau_2 \) contains an open neighborhood of \( x \) in \( \tau_1 \). We denote it \( \tau_1 \succeq \tau_2 \). If \( \tau_1 \succeq \tau_2 \) and \( \tau_2 \succeq \tau_1 \), then the topologies are called equivalent. For example, equivalent norms on a normed space \( X \) define equivalent norm-topologies. Show that in general, \( \tau_1 \succeq \tau_2 \) if and only if a net converging in \( \tau_1 \) also converges in \( \tau_2 \).

Let \( X \) be a set. A family of subsets \( \mathcal{F} \subset 2^X \) is called a filter if

1. \( \emptyset \notin \mathcal{F} \);
2. if \( F_1, \ldots, F_n \) are elements of \( \mathcal{F} \), then \( \bigcap_{j=1}^n F_j \in \mathcal{F} \);
3. if \( F \in \mathcal{F} \) and \( F \subseteq S \), then \( S \in \mathcal{F} \).

Let \( P \) be the set of all filters in \( X \) ordered by inclusion. A routine verification shows that \( P \) satisfies the conditions of Zorn’s Lemma. Every maximal element of \( P \) is called an ultrafilter. In fact, for any filter \( \mathcal{F} \) there is an ultrafilter containing \( \mathcal{F} \), for the subset of \( P \) of filters containing the given one satisfies Zorn’s Lemma as well. Ultrafilters can be characterized by adding one more condition to the three above: \( \mathcal{U} \) is an ultrafilter if and only if it is a filter and

4. for any subset \( A \subseteq X \) either \( A \in \mathcal{U} \) or \( X \setminus A \in \mathcal{U} \).

Indeed, if (4) holds and \( \mathcal{U}' \) is another filter containing \( \mathcal{U} \), then any set \( A \in \mathcal{U}' \) should be in \( \mathcal{U} \), for otherwise, \( X \setminus A \in \mathcal{U} \), and then \( \emptyset = A \cap (X \setminus A) \in \mathcal{U}' \). Conversely, if \( A \subseteq X \) is such that \( X \setminus A \notin \mathcal{U} \), then by (3) every \( F \in \mathcal{U} \) must intersect with \( A \). Define a new family \( \mathcal{U}' = \{ S : F \cap A \subseteq S, F \in \mathcal{U} \} \). Clearly, \( \mathcal{U} \subseteq \mathcal{U}' \), \( A \in \mathcal{U}' \), and one can easily check that \( \mathcal{U}' \) is a filter. By maximality of \( \mathcal{U}, \mathcal{U} = \mathcal{U}' \), and hence \( A \in \mathcal{U} \). An alternative to (4) is a formally stronger, but equivalent condition:

4’ if \( A_1 \cup \ldots \cup A_n \in \mathcal{U} \), then some \( A_i \in \mathcal{U} \).

Indeed, if non of \( A_i \)'s belongs to \( \mathcal{U} \), then all the complements do, and hence their intersection, which is \( X \setminus (A_1 \cup \ldots \cup A_n) \). This is incompatible with (1).

The compactness of a topological space can be restated in terms of convergence of ultrafilters. So, let \((X, \tau)\) be a topological space and \( \mathcal{F} \) be a filter on it. We say that \( \lim \mathcal{F} = x \) if every neighborhood of \( x \) has a non-empty intersection with any element of the filter. If \( \mathcal{F} \) is an ultrafilter, which will be our standard assumption, we showed above that every set that intersects every element of \( \mathcal{F} \) must lie in \( \mathcal{F} \). Thus, in this case \( \lim \mathcal{F} = x \) iff every open neighborhood of \( x \) is contained in \( \mathcal{F} \). If every two distinct points in \( X \) can be separated by disjoint open neighborhoods, and such a space is called Hausdorff, then clearly, the limit is unique. What follows, however, does not require this assumption.

Lemma 3.5. \( X \) is compact if and only if every ultrafilter in \( X \) converges to a point in \( X \).

Proof. Suppose \( X \) is compact, and let \( \mathcal{U} \) be an ultrafilter on \( X \). If \( \mathcal{U} \) does not converge to any point in \( X \), then any point \( x \in X \) is contained in \( U_x \in \tau \) with \( U_x \notin \mathcal{U} \). Thus, \( \{ U_x \}_{x \in X} \) form an open cover of \( X \), which must contain a finite sub cover, \( U_1, \ldots, U_n \). But \( \cup_j U_j = X \subset \mathcal{U} \), so by (4’) one of the sets must be in \( \mathcal{U} \), a contradiction.
Conversely, let $C = \{U\}$ be an open cover of $X$. Suppose that it contains no finite subcover. Thus, any finite intersection $(X \setminus U_1) \cap \ldots \cap (X \setminus U_n)$ in non-empty. This shows that the family $\mathcal{F} = \{F \subset X : X \setminus U \subset F, \text{ for some } U \in C\}$ is a filter, and let $U$ be an ultrafilter containing $\mathcal{F}$. By assumption, let $x = \lim U$. Since $C$ is a cover, there is $U \in C$ with $x \in U$. But, $X \setminus U \in \mathcal{U}$, so $U$ cannot be in $\mathcal{U}$, a contradiction. \hfill $\square$

Ultrafilters are useful for many purposes. In particular, they provide an economical proof of Tihonov’s compactness theorem, which will be used later to establish the main result of this section, the Alaoglu Theorem.

Let $\{(X_\gamma, \tau_\gamma)\}_{\gamma \in \Gamma}$ be a collection of topological spaces. The product space, $X = \prod_{\gamma \in \Gamma} X_\gamma$ is the set of functions $x : \Gamma \to \bigcup_{\gamma \in \Gamma} X_\gamma$ such that $x(\gamma) \in X_\gamma$. We usually denote $x(\gamma) = x_\gamma$ and write $x = \{x_\gamma\}_{\gamma \in \Gamma}$. Let $\pi_\gamma : X \to X_\gamma$ be the usual projection map. We define the product topology on $X$ to be the topology generated by the sets $\pi^{-1}(U_\gamma)$, where $U_\gamma \in \tau_\gamma$. This is also the minimal topology on $X$ in which all the projection maps are continuous.

**Exercise 3.6.** Show that $x_\alpha \to x$ in the product topology if and only if $\pi_\gamma(x_\alpha) \to \pi_\gamma(x)$ for every $\gamma \in \Gamma$.

**Theorem 3.7** (Tihonov’s compactness theorem). If all $X_\gamma, \gamma \in \Gamma$, are compact, then the product $X = \prod_{\gamma \in \Gamma} X_\gamma$ is compact in the product topology.

**Proof.** Let $U$ be an ultrafilter on $X$. For every $\gamma \in \Gamma$, consider $U_\gamma = \pi_\gamma(U)$. Then $U_\gamma$ is an ultrafilter. Since $X_\gamma$ is compact, there exists a limit $x_\gamma = \lim U_\gamma$. Let us show that $x = \{x_\gamma\}_{\gamma \in \Gamma} = \lim U$. Let $U$ be open neighborhood of $x$ in the product topology. Then with $x, U$ contains a finite intersection of basis sets $\cap_{i=1}^n \pi^{-1}_i(U_i)$. We have $x_\gamma \in U_\gamma$, and hence $U_i \subseteq U_{\gamma_i}$, which means there exist $A_i \in U$ such that $\pi_{\gamma_i}(A_i) = U_i$. Then $A_i \subseteq \pi_{\gamma_i}^{-1}(U_i)$, which implies that $\pi_{\gamma_i}^{-1}(U_i) \in U$, for each $i$, and hence $\cap_{i=1}^n \pi_{\gamma_i}^{-1}(U_i) \in U$. Since $U$ contains that intersection, it must itself be in the ultrafilter $U$. \hfill $\square$

3.2. **Weak topology.** Let $X$ be a Banach space. We define the weak topology on $X$ as the topology with the following base of neighborhoods: for $x \in X$, $\varepsilon_1, \ldots, \varepsilon_n > 0$ and $f_1, \ldots, f_n \in X^*$, let

$$U_{\varepsilon_1, \ldots, \varepsilon_n}^{f_1, \ldots, f_n}(x) = \{y \in X : |f_i(y) - f_i(x)| < \varepsilon_i, \forall i = \frac{1}{n}\}.$$ 

(19)

Thus, a neighborhood is an intersection of open slabs of finite widths. We say that a sequence, or a net $\{x_\alpha\}_{\alpha \in A}$ converges weakly to $x$, and denote $x_\alpha \overset{w}{\to} x$, if it converges in the sense of the weak topology.

**Exercise 3.8.** Show that $x_\alpha \overset{w}{\to} x$ if and only if $f(x_\alpha) \to f(x)$ for every functional $f \in X^*$.

**Exercise 3.9.** Show that if a sequence $x_n$ converges weakly, then it is bounded in the norm-topology.

**Exercise 3.10.** Suppose dim $X = \infty$. Construct a net $\{x_\alpha\}_{\alpha \in A}$ in $X$ such that $x_\alpha \to 0$ weakly, yet for every $\alpha_0 \in A$ and $N > 0$, there is $\alpha \geq \alpha_0$ such that $\|x_\alpha\| \geq N$. Hint: let $A = \{(f_1, \ldots, f_n; N) : f_j \in X^*, n \in \mathbb{N}, N > 0\}$.

Define a directed partial order on $A$, and for every $\alpha \in A$ pick an $x_\alpha \in \cap_j \text{Ker} f_j$, with $\|x_\alpha\| > N$. Show that $x_\alpha \to 0$ weakly and is frequently unbounded.
From now on we will use the term ”strong” with respect to anything related to the norm-topology, as opposed to ”weak” that refers to anything related to the weak topology. For example, ”strongly compact set” v.s. ”weakly compact set”, or ”strong convergence” v.s. ”weak convergence”.

It is clear that the weak topology is weaker than the norm-topology on any normed space. In fact, on an infinite dimensional space it is strictly weaker. To see this, we show that any neighborhood (19) is unbounded. Indeed, let $H = \cap_i \text{Ker } f_i$. This is a non-empty space, for otherwise $X$ would have been a span of $x_i$’s, with $x_i \notin \text{Ker } f_i$. Then $U_{f_1, \ldots, f_n}(x)$ contains all of $x + H$.

**Lemma 3.11.** The weak topology is not metrizable on an infinite dimensional space.

**Proof.** Suppose, on the contrary that there is a metric $d(\cdot, \cdot)$ that defines the weak topology. Consider the sequence of balls $\{d(x, 0) < 1/n\}$. Each contains a weak neighborhood of the origin. We have shown every weak neighborhood is unbounded. Thus, we can find a $x_n$ within the $n$th ball with $\|x_n\| > n$. So, on the one hand, $x_n \to 0$ weakly, and yet $\{x_n\}$ is unbounded, in contradiction with Exercise 3.2.

The weak topology is still ”fine” enough to separate points, and even larger sets, by the Separation Theorems.

**Exercise 3.12.** Show that if $A$ is closed and $B$ is strongly compact convex sets, then there are two disjoint weakly open neighborhoods of $A$ and $B$.

**Lemma 3.13.** Let $\{x_n\}_{n=1}^\infty \subset X$ be a sequence in any of the spaces $X = c_0, \ell_p$, for $1 < p < \infty$. Then $x_n \overset{w}{\to} x$ if and only if $\{x_n\}$ is norm bounded and converges to $x$ pointwise, i.e. $x_n(j) \to x(j)$, for all $j \in \mathbb{N}$.

**Proof.** We present the proof for $X = c_0$ and leave the $\ell_p$-case as an exercise. So, suppose $x_n \overset{w}{\to} x$. Then the sequence is bounded by Exercise 3.2. Moreover, taking $f = e_j$, we obtain the pointwise convergence. Conversely, if $x_n \to x$ pointwise, and is bounded, let $M = \sup_n \|x_n\|$, and let $f \in \ell_1 = c_0^*$. Given $\varepsilon > 0$, let $N \in \mathbb{N}$ be such that $\sum_{j>N} |f(j)| < \varepsilon/(2M)$. Then, for large $n$ we have

$$\left| \sum_{j \leq N} x_n(j)f(j) - \sum_{j \leq N} x(j)f(j) \right| < \varepsilon/2.$$ 

Thus,

$$\left| \sum_j x_n(j)f(j) - \sum_j x(j)f(j) \right| \leq \left| \sum_{j \leq N} x_n(j)f(j) - \sum_{j \leq N} x(j)f(j) \right| + \left| \sum_{j>N} x_n(j)f(j) - \sum_{j>N} x(j)f(j) \right| < \varepsilon.$$

**Lemma 3.14.** In $\ell_1$, a sequence $\{x_n\}_{n=1}^\infty$ converges weakly to $x$ if and only if it converges to $x$ strongly.

Let us note that in spite of Exercise 3.4, the lemma above does not imply that the norm and weak topologies are equivalent, because it deals only with sequences. Loosely speaking, the reason why on $\ell_1$ weak and strong convergences for sequences are equivalent is because the dual $\ell_\infty$ is ”very large”, so large that weak convergence is just as hard to arrange as
strong convergence. Banach spaces with the property stated in Lemma 3.14 are sometimes called Kadets-Klee spaces.

As a weaker topology, the weak topology provides a smaller family of open sets than the norm topology. Thus any weakly closed set is strongly closed as well. As a consequence of the Separation Theorem 2.11, it turns out that among the class of convex sets the property of being closed is the same in weak and strong topologies.

**Lemma 3.15.** A convex set \( C \subset X \) is strongly closed if and only if it is weakly closed.

**Proof.** Clearly, if \( C \) is strongly closed, and yet there is a point \( x_0 \in \overline{C}^w \) not in \( C \). Then by Theorem 2.11, there is a functional \( f \in X^* \) so that \( f(x_0) > c > \sup f(C) \). Thus \( x_0 \) belongs to the open set \( U = \{ f > c \} \) disjoint from \( C \), a contradiction. \( \square \)

The exact same argument shows that the weak and strong closures of a convex set \( C \) coincide. This has a rather interesting consequence for relationship between weak and strong convergence of sequences.

**Exercise 3.16.** Show that if \( x_n \to x \) weakly, then there is a sequence of convex combinations made of \( x_n \)'s that converge to \( x \) strongly.

**Exercise 3.17 (Weak lower-semi-continuity of norm).** Show that if \( x_n \to x \) weakly, then \( \liminf_{n \to \infty} \| x_n \| \geq \| x \| \). More generally, if a net \( x_\alpha \rightharpoonup x \), then for any \( \varepsilon > 0 \) there is \( \alpha_0 \) so that for all \( \alpha \geq \alpha_0 \), \( \| x_\alpha \| \geq \| x \| - \varepsilon \).

3.3. **Weak* topology.** Consider now the dual space \( X^* \). As any Banach space it has its own weak topology determined by the functionals from "upstairs", i.e. \( X^{**} \). However, one can define a weaker Hausdorff topology on \( X^* \) determined by the functionals from "downstairs", i.e. \( X \). Let \( x_1, \ldots, x_n \in X \) and \( \varepsilon_1, \ldots, \varepsilon_n > 0 \), and \( f \in X^* \). We define a weak*-open neighborhood of \( f \) to be

\[
U_{\varepsilon_1, \ldots, \varepsilon_n}^{x_1, \ldots, x_n}(f) = \{ g \in X^* : |g(x_i) - f(x_i)| < \varepsilon_i, \forall i = 1, n \}.
\]

Identifying element of \( X \) as vectors in \( X^{**} \) we see it is just a special subclass of neighborhoods defined earlier in (19). It is still a Hausdorff topology as pairs of distinct functionals in \( X^* \) can be separated by elements of \( X \). A sequence \( x_n \) converging weak* to \( x \) is necessarily bounded by the Banach-Steinhauss Theorem.

**Exercise 3.18.** Show that a net \( f_\alpha \rightharpoonup f \) if and only if \( f_\alpha(x) \to f(x) \) for every \( x \in X \).

**Exercise 3.19.** Show that any weakly* convergent sequence in \( X^* \) is strongly bounded.

**Exercise 3.20 (Weak* lower-semi-continuity of norm).** Show that if \( f_n \rightharpoonup f \), then

\[
\liminf_{n \to \infty} \| f_n \| \geq \| f \|.
\]

More generally, if a net \( \{ f_\alpha \}_{\alpha \in A} \) converges weakly* to \( f \), then for any \( \varepsilon > 0 \) there is \( \alpha_0 \in A \) so that for all \( \alpha \geq \alpha_0 \), \( \| f_\alpha \| \geq \| f \| - \varepsilon \).

How much the weak* topology may be weaker than the weak topology is illustrated by the following example (c.f. Lemma 3.14).

\[1\] It sounds a bit counterintuitive from the linguistic point. But if you think what it takes for a set to be closed it becomes clear. If a set "survives" weak limits from within itself, then it should definitely survive strong limits.
Exercise 3.21. Show that in $\ell_1$, $x_n \overset{w^*}{\to} x$ if and only if $\{x_n\}$ is bounded and $x_n \to x$ pointwise.

Theorem 3.22 (Alaoglu). The unit ball of a dual space is compact in the weak*-topology.

Proof. Notice that for any $f \in B(X^*)$, and $x \in X$, $f(x) \in [-\|x\|, \|x\|]$. This naturally suggests to consider $B(X^*)$ as a subset of the product space $T = \prod_{x \in X} [-\|x\|, \|x\|]$. By Tihonov’s theorem, this product space is compact in the product topology. It suffices to show that $B(X^*)$ is closed in $T$, because convergence of nets in the product topology is equivalent to pointwise convergence, which for elements of $B(X^*)$ amounts to weak* convergence.

To this end, let $\{f_\alpha\}_{\alpha \in A}$ be a net in $B(X^*)$ with $\lim_{\alpha} f_\alpha = f \in T$. By linearity of $f_\alpha$’s and the “pointwise” sense of the limit above, we conclude that

$$f(\lambda x + \mu y) = f_\alpha(\lambda x + \mu y) = \lambda f_\alpha(x) + \mu f_\alpha(y) \to \lambda f(x) + \mu f(y).$$

Thus, $f$ is linear, and since $|f_\alpha(x)| \leq \|x\|$, we also have $|f(x)| \leq \|x\|$ for all $x \in X$, which identifies $f$ as an element of $B(X^*)$.

As an immediate consequence we see that for a reflexive Banach space $X$, the unit ball is weakly compact. This the property actually characterizes reflexiveness.

Theorem 3.23 (Kakutani). A Banach space $X$ is reflexive if and only if its unit ball is weakly compact.

The theorem will follow from the next lemma.

Lemma 3.24. The unit ball $B(X)$ is weakly* dense in $B(X^{**})$. In particular, $B(X^{**})^{w^*} = B(X^{**})$.

Indeed, let $B(X)$ be weakly compact. In view of Lemma 3.24, for any $x^{**} \in B(X^{**})$ we find a net $x_\alpha \overset{w^*}{\to} x^{**}$, $x_\alpha \in B(X)$. By compactness, there is a subnet $y_\beta \overset{w^*}{\to} x$ for some $x \in B(X)$, and yet that same subnet converges to $x^{**}$ in the weak* sense of $X^{**}$. Thus, for every $x^*$ we have $(x^{**}, x^*) \leftarrow (y_\beta, x^*) = (x^*, y_\beta) \to (x^*, x)$, which identifies $x^{**}$ as $x$, and thus $B(X) = B(X^{**})$, implying $X = X^{**}$.

In order to prove Lemma 3.24, we have to go back to the separation theorem and make one adjustment to it in the context of weak* topology.

Lemma 3.25. Suppose that $f, f_1, \ldots, f_n \in X'$, and $\cap_{j=1}^n \ker f_j \subset \ker f$. Then $f \in \{f_1, \ldots, f_n\}$.

Proof. We will prove the lemma by induction. Suppose $n = 1$, and $f_1 \neq 0$ (otherwise the statement is trivial). By the structure of the linear functionals discussed in Section 1.10, there is $x_1 \in X$ with $f_1(x_1) = 1$, such that for every $x \in X$ we have $x = \lambda x_1 + y$, where $y \in \ker f_1$. Since $f(y) = 0$ we have $f(x) = \lambda f(x_1) = f(x_1) f_1(x)$, as desired.

Suppose the statement is true for $n$. Let us assume $\cap_{j=1}^{n+1} \ker f_j \subset \ker f$. Consider the space $Y = \ker f_{n+1}$. Then $\cap_{j=1}^n \ker f_j |_Y \subset \ker f |_Y$. By the induction hypothesis, $f |_Y = \sum_{j=1}^n a_j f_j |_Y$. By the structure of $f_{n+1}$ we have for any $x \in X$, $x = \lambda x_{n+1} + y$ for some
Proof of Lemma 3.24. Let 

\[ f(x) = \lambda f(x_{n+1}) + \sum_{j=1}^{n} a_j f_j(y) = f(x_{n+1}) f_{n+1}(x) + \sum_{j=1}^{n} a_j f_j(y) \]

\[ = f(x_{n+1}) f_{n+1}(x) + \sum_{j=1}^{n} a_j f_j(x) - \lambda \sum_{j=1}^{n} a_j f_j(x_{n+1}) \]

\[ = \left( f(x_{n+1}) - \sum_{j=1}^{n} a_j f_j(x_{n+1}) \right) f_{n+1}(x) + \sum_{j=1}^{n} a_j f_j(x) \]

\[ = \sum_{j=1}^{n+1} a_j f_j(x). \tag{21} \]

\[ \square \]

Lemma 3.26. If \( x^{**} \in X^{**} \) is continuous in the weak* topology, then \( x^{**} \in X \).

Proof. By the assumption \( (x^{**})^{-1}(-1,1) \) contains a weak* neighborhood of the origin, say, \( U_{\varepsilon_1, \ldots, \varepsilon_n}(0) \). In particular, \( x^{**} \in (-1,1) \) on \( \cap_j \text{ Ker } x_j \). Since the latter is a linear space, \( x^{**} \) must in fact vanish on it. Thus, by Lemma 3.25, \( x^{**} \in [x_1, \ldots, x_n] \subset X \).

\[ \square \]

Theorem 3.27 (Separation Theorem for weak* topology). Suppose \( B \) is a weakly*-closed convex subset of \( X^{*} \), and \( f \notin B \). Then, there is \( x \in X \) such that \( sup B(x) < 1 < f(x) \).

Proof. Let us follow the proof of Theorem 2.11. Let us assume \( f = 0 \), and let \( A = U_{\varepsilon_1, \ldots, \varepsilon_n}(0) \) be a weak* neighborhood of 0 disjoint from \( B \). We then associate Minkowski’s functionals \( p_A \) and \( q_B \) to \( A \) and \( B \) respectfully. As a result of Hahn-Banach Theorem, we find a separating functional \( F \) so that \( q_B \leq F \leq p_A \). Since \( f \in A \) implies \( p_A(f) \leq 1 \), we see that \( F \) is bounded from above on \( A \). Like in the proof of Lemma 3.26 we conclude that \( F \) vanishes on the intersection of the kernels of the \( x_1, \ldots, x_n \), and hence, \( F \in X \). By rescaling \( F \) if necessary we can arrange the constant of separation to be 1.

\[ \square \]

Proof of Lemma 3.24. Let \( B = \overline{B(X)}^{w^*} \). By Exercise 3.20, \( B \subset B(X^{**}) \). Suppose there is \( F \in B(X^{**}) \setminus B \). By Theorem 3.27, we find a \( x^* \in X^* \) such that \( B(x^*) < 1 < F(x^*) \). The first inequality holds, in particular, on \( B(X) \), which shows that \( \|x^*\| \leq 1 \). This runs into contradiction with the second inequality.

\[ \square \]

Exercise 3.28. Recall that \( (c_0)^{**} = (\ell_1)^* = \ell_\infty \). Show that \( B(c_0) \) is weakly* sequentially dense in \( B(\ell_\infty) \), i.e. for every \( F \in B(\ell_\infty) \) there is a sequence \( x_n \in B(c_0) \) converging weakly* to \( F \).

Corollary 3.29. Let \( Y \subset X \) be a closed subspace of a reflexive space \( X \). Then \( Y \) and \( X/Y \) are reflexive.

Proof. Indeed, \( B(Y) = Y \cap B(X) \). Since the this set if convex and closed, it is weakly closed in \( B(X) \) and hence compact in the weak topology of \( X \). However, by the Hahn-Banach extension theorem the topology induced on \( Y \) by the weak topology of \( X \) is exactly the weak topology of \( Y \). Thus, \( B(Y) \) is weakly compact in \( Y \). Now, by Exercise 2.7 and by the previous, \( (X/Y)^* \) is a subspace of a reflexive space, which makes it reflexive. Then by Exercise 2.6 \( (X/Y) \) itself is reflexive.