Functional Analysis, a graduate course

Michael Papadimitrakis

Department of Mathematics and Applied Mathematics

University of Crete

Contents

Chapter 1

Normed spaces

1.1 Norms.

Let X be a linear space over the field F, where $F = \mathbb{R}$ or $F = \mathbb{C}$.
We recall the operations:

$$A + B = \{a + b \,|\, a \in A, b \in B\}, \quad x + A = \{x + a \,|\, a \in A\}, \quad \lambda A = \{\lambda a \,|\, a \in A\}$$

for every $A, B \subseteq X$ and every $\lambda \in F$.

Definition. *We say that the function*

$$\|\cdot\| : X \to \mathbb{R}$$

is a **norm** *on X, if*

(i) $\|x\| \geq 0$,

(ii) $\|x\| = 0 \;\Leftrightarrow\; x = 0$,

(iii) $\|\lambda x\| = |\lambda|\,\|x\|$,

(iv) $\|x + y\| \leq \|x\| + \|y\|$,

for every $x, y \in X$ and every $\lambda \in F$.

From the properties of the norm $\|\cdot\|$ we easily get

$$\|-x\| = \|x\|, \quad \big|\|x\| - \|y\|\big| \leq \|x \pm y\| \leq \|x\| + \|y\|$$

for every $x, y \in X$.

Example 1.1.1. A trivial example of a normed space over F is the field F itself with the absolute value $|\cdot| : F \to \mathbb{R}$ as a norm.

The norm $\|\cdot\|$ on X induces a **metric** on X, i.e. the function $d : X \times X \to \mathbb{R}$ defined by

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

for every $x, y \in X$. This metric has the usual properties:

(i) $d(x, y) \geq 0$,

(ii) $d(x, y) = 0 \;\Leftrightarrow\; x = y$,

(iii) $d(y, x) = d(x, y)$,

(iv) $d(x,z) \le d(x,y) + d(y,z)$

for every $x, y, z \in X$.

The metric d induced by a norm as above has the additional properties

(v) $d(x+y, x+z) = d(y,z)$, i.e. the metric is **translation invariant**,

(vi) $d(\lambda x, \lambda y) = |\lambda|\, d(x,y)$, i.e. the metric is **positive homogenuous**.

Now as in any metric space we define neighborhoods of points, i.e. the **open balls** and the **closed balls**

$$B(a;r) = \{x \in X \mid \|x-a\| < r\}, \quad \overline{B}(a;r) = \{x \in X \mid \|x-a\| \le r\},$$

with center $a \in X$ and radius $r > 0$.

Two easily proved identities are

$$x + B(a;r) = B(x+a;r), \quad \lambda\, B(a;r) = B(\lambda a; |\lambda| r).$$

As in any metric space, we define the **open subsets** and the **closed subsets** of the normed space X. The set $A \subseteq X$ is open if for every $a \in A$ there is some $r > 0$ so that $B(a;r) \subseteq A$, i.e. if every $a \in A$ is an interior point of A. The set $A \subseteq X$ is closed if its complement $A^c = X \setminus A$ is open.

It is well known that the special sets $\emptyset$ and X are open and closed, that the union of open sets is open and the intersection of closed sets is closed, and that the intersection of finitely many open sets is open and the union of finitely many closed sets is closed.

We also have the notion of **convergence** of sequences in the normed space X. We say that the sequence (x_n) in X converges to $x \in X$, and we denote this by $x_n \to x$, if for every $\epsilon > 0$ there is $n_0 \in \mathbb{N}$ so that $\|x_n - x\| < \epsilon$ for every $n \ge n_0$. Of course

$$x_n \to x \text{ in } X \quad \Leftrightarrow \quad d(x_n, x) = \|x_n - x\| \to 0 \text{ in } \mathbb{R}.$$

Proposition 1.1. *The linear space operations of a normed space X are continuous, i.e.*
(i) if $x_n \to x$ and $y_n \to y$ in X, then $x_n + y_n \to x + y$ in X.
(ii) if $x_n \to x$ in X and $\lambda_n \to \lambda$ in F, then $\lambda_n x_n \to \lambda x$ in X.

Proof. These two properties are implied by the inequalities

$$\|(x_n + y_n) - (x+y)\| \le \|x_n - x\| + \|y_n - y\|,$$

$$\|\lambda_n x_n - \lambda x\| \le |\lambda_n|\, \|x_n - x\| + |\lambda_n - \lambda|\, \|x\|.$$

□

Proposition 1.2. *The norm of a normed space X is continuous, i.e. if $x_n \to x$ in X, then $\|x_n\| \to \|x\|$ in $\mathbb{R}$.*

Proof. This is implied by the inequality $\big|\|x_n\| - \|x\|\big| \le \|x_n - x\|$. □

Proposition 1.3. *Let X be a normed space and Y be a linear subspace of X. Then* $\mathrm{cl}(Y)$, *the closure of Y in X, is a linear subspace of X.*

Proof. Let $a, b \in \mathrm{cl}(Y)$ and $\lambda, \kappa \in F$.
There are sequences $(a_n), (b_n) \in Y$ so that $a_n \to a$ and $b_n \to b$. By the continuity of the linear space operations, we have that $\lambda a_n + \kappa b_n \to \lambda a + \kappa b$. Since $\lambda a_n + \kappa b_n \in Y$ for every n, we get that $\lambda a + \kappa b \in \mathrm{cl}(Y)$.
So $\mathrm{cl}(Y)$ is a linear subspace of X. □

Proposition 1.4. *Let X be a normed space.*
(i) If $A \subseteq X$ is open (closed), then $x + A$ is open (closed).
(ii) If $A \subseteq X$ is open (closed) and $\lambda \neq 0$, then λA is open (closed).

Proof. (i) Let A be open. We take any $b \in x + A$ and then $b - x \in A$. So there is $r > 0$ such that $B(b - x; r) \subseteq A$, and then

$$B(b; r) = x + B(b - x; r) \subseteq x + A.$$

Therefore $x + A$ is open.
Now let A be closed. We take any sequence (b_n) in $x + A$ and we assume that $b_n \to b$. Then $(b_n - x)$ is a sequence in A and $b_n - x \to b - x$. Thus $b - x \in A$ and so $b \in x + A$.
Therefore $x + A$ is closed.
(ii) The proof is similar. □

As in any metric space we have the notion of **compactness** for subsets of a normed space X. We say that $K \subseteq X$ is compact if every open covering of K has a finite subcovering of K. This means that if $K \subseteq \bigcup_{A \in \mathcal{A}} A$, where every $A \in \mathcal{A}$ is an open subset of X, then there are $A_1, \ldots, A_n \in \mathcal{A}$ so that $K \subseteq \bigcup_{k=1}^{n} A_k$.

We know that $K \subseteq X$ is compact if and only if every sequence in K has some subsequence which converges to an element of K.

We also know that every compact subset of X is closed and bounded (i.e. it is contained in some ball). In general, the converse is not correct.

1.2 Hölder and Minkowski inequalities.

Hölder's inequality for sums. *Let $p, q > 1$ and $\frac{1}{p} + \frac{1}{q} = 1$, and let $\lambda_k, \kappa_k \geq 0$ for every $k \in \mathbb{N}$.*
(i) If $\sum_{k=1}^{+\infty} \lambda_k^p < +\infty$ and $\sum_{k=1}^{+\infty} \kappa_k^q < +\infty$, then

$$\sum_{k=1}^{+\infty} \lambda_k \kappa_k \leq \Big(\sum_{k=1}^{+\infty} \lambda_k^p \Big)^{1/p} \Big(\sum_{k=1}^{+\infty} \kappa_k^q \Big)^{1/q}.$$

(ii) If $\sum_{k=1}^{+\infty} \lambda_k < +\infty$ and $\sup_k \kappa_k < +\infty$, then

$$\sum_{k=1}^{+\infty} \lambda_k \kappa_k \leq \Big(\sum_{k=1}^{+\infty} \lambda_k \Big) \sup_k \kappa_k.$$

Proof. (i) We observe that the function $f(t) = \frac{1}{p} t^p + \frac{1}{q} - t$ has minimum value $f(1) = 0$ in $[0, +\infty)$. I.e. $t \leq \frac{1}{p} t^p + \frac{1}{q}$ for every $t \geq 0$. We use $t = \frac{\lambda}{\kappa^{q/p}}$ and we get

$$\lambda\kappa \leq \frac{1}{p} \lambda^p + \frac{1}{q} \kappa^q$$

for every $\lambda, \kappa \geq 0$.
If $\sum_{k=1}^{+\infty} \lambda_k^p = \sum_{k=1}^{+\infty} \kappa_k^q = 1$, then, using the last inequality, we get

$$\sum_{k=1}^{+\infty} \lambda_k \kappa_k \leq \frac{1}{p} \sum_{k=1}^{+\infty} \lambda_k^p + \frac{1}{q} \sum_{k=1}^{+\infty} \kappa_k^q = \frac{1}{p} + \frac{1}{q} = 1.$$

If $0 < \sum_{k=1}^{+\infty} \lambda_k^p < +\infty$ and $0 < \sum_{k=1}^{+\infty} \kappa_k^q < +\infty$, then we set

$$A = \Big(\sum_{k=1}^{+\infty} \lambda_k^p \Big)^{1/p}, \quad B = \Big(\sum_{k=1}^{+\infty} \kappa_k^q \Big)^{1/q}$$

and we observe that $\sum_{k=1}^{+\infty}\left(\frac{\lambda_k}{A}\right)^p=\sum_{k=1}^{+\infty}\left(\frac{\kappa_k}{B}\right)^q=1$. Hence

$$\sum_{k=1}^{+\infty}\lambda_k\kappa_k=AB\sum_{k=1}^{+\infty}\frac{\lambda_k}{A}\frac{\kappa_k}{B}\leq AB.$$

If one of $\sum_{k=1}^{+\infty}\lambda_k^p$, $\sum_{k=1}^{+\infty}\kappa_k^q$ is equal to 0, the inequality of (i) is obvious: it becomes $0\leq 0$.
(ii) This is trivial. $\square$

If $p=q=2$, Hölder's inequality for sums is usually called **Cauchy's inequality**.

If $p,q>1$ and $\frac{1}{p}+\frac{1}{q}=1$, then p,q are called **dual** exponents. Since $\frac{1}{+\infty}=0$, the $1,+\infty$ are also dual exponents.

Minkowski's inequality for sums. *Let* $p\geq 1$*, and let* $\lambda_k,\kappa_k\geq 0$ *for every* $k\in\mathbb{N}$.
(i) If $\sum_{k=1}^{+\infty}\lambda_k^p<+\infty$ *and* $\sum_{k=1}^{+\infty}\kappa_k^p<+\infty$*, then*

$$\Big(\sum_{k=1}^{+\infty}(\lambda_k+\kappa_k)^p\Big)^{1/p}\leq\Big(\sum_{k=1}^{+\infty}\lambda_k^p\Big)^{1/p}+\Big(\sum_{k=1}^{+\infty}\kappa_k^p\Big)^{1/p}.$$

(ii) If $\sup_k\lambda_k<+\infty$ *and* $\sup_k\kappa_k<+\infty$*, then*

$$\sup_k(\lambda_k+\kappa_k)\leq\sup_k\lambda_k+\sup_k\kappa_k.$$

Proof. (i) The inequality of (i) is an obvious equality when $p=1$.
Now we take $p>1$ and $q=\frac{p}{p-1}$, and so $\frac{1}{p}+\frac{1}{q}=1$.
Since $\sum_{k=1}^{+\infty}\lambda_k^p<+\infty$ and $\sum_{k=1}^{+\infty}\kappa_k^p<+\infty$, we get

$$\sum_{k=1}^{+\infty}(\lambda_k+\kappa_k)^p\leq 2^{p-1}\sum_{k=1}^{+\infty}\lambda_k^p+2^{p-1}\sum_{k=1}^{+\infty}\kappa_k^p<+\infty.$$

For the last inequality we used the trivial inequality

$$(\lambda+\kappa)^p\leq 2^{p-1}(\lambda^p+\kappa^p)$$

for $\lambda,\kappa\geq 0$, which can be proved using the convexity of the function t^p in $[0,+\infty)$. Then

$$\sum_{k=1}^{+\infty}(\lambda_k+\kappa_k)^p=\sum_{k=1}^{+\infty}(\lambda_k+\kappa_k)(\lambda_k+\kappa_k)^{p-1}=\sum_{k=1}^{+\infty}\lambda_k(\lambda_k+\kappa_k)^{p-1}+\sum_{k=1}^{+\infty}\kappa_k(\lambda_k+\kappa_k)^{p-1}$$

and, using Hölder's inequality for sums,

$$\sum_{k=1}^{+\infty}(\lambda_k+\kappa_k)^p\leq\Big(\sum_{k=1}^{+\infty}\lambda_k^p\Big)^{1/p}\Big(\sum_{k=1}^{+\infty}(\lambda_k+\kappa_k)^p\Big)^{1/q}+\Big(\sum_{k=1}^{+\infty}\kappa_k^p\Big)^{1/p}\Big(\sum_{k=1}^{+\infty}(\lambda_k+\kappa_k)^p\Big)^{1/q}.$$

If $\sum_{k=1}^{+\infty}(\lambda_k+\kappa_k)^p>0$, we divide the last inequality with $\big(\sum_{k=1}^{+\infty}(\lambda_k+\kappa_k)^p\big)^{1/q}$ and we finish the proof. If $\sum_{k=1}^{+\infty}(\lambda_k+\kappa_k)^p=0$, then the inequality of (i) is trivial: $0\leq 0+0$.
(ii) Trivial. $\square$

Hölder's inequality for integrals. *Let* $p,q>1$ *and* $\frac{1}{p}+\frac{1}{q}=1$*. Let* (Ω,Σ,μ) *be a measure space, and* $f,g\in\mathcal{M}(\Omega,\Sigma)$ *(i.e.* f,g *are* Σ*-measurable in* Ω*) with* $f,g\geq 0$ μ*-a.e. in* Ω.
(i) If $\int_\Omega f^p\,d\mu<+\infty$ *and* $\int_\Omega g^q\,d\mu<+\infty$*, then*

$$\int_\Omega fg\,d\mu\leq\Big(\int_\Omega f^p\,d\mu\Big)^{1/p}\Big(\int_\Omega g^q\,d\mu\Big)^{1/q}.$$

(ii) If $\int_\Omega f\,d\mu<+\infty$ *and* $\text{ess-sup}_\Omega\,g<+\infty$*, then*

$$\int_\Omega fg\,d\mu\leq\Big(\int_\Omega f\,d\mu\Big)\,\text{ess-sup}_\Omega\,g.$$

Proof. If $\int_\Omega f^p\,d\mu = \int_\Omega g^q\,d\mu = 1$, then, using the inequality $\lambda\kappa \le \frac{1}{p}\,\lambda^p + \frac{1}{q}\,\kappa^q$, we get

$$\int_\Omega fg\,d\mu \le \frac{1}{p}\int_\Omega f^p\,d\mu + \frac{1}{q}\int_\Omega g^q\,d\mu = \frac{1}{p} + \frac{1}{q} = 1.$$

If $0 < \int_\Omega f^p\,d\mu < +\infty$ and $0 < \int_\Omega g^q\,d\mu < +\infty$, then we set

$$A = \Big(\int_\Omega f^p\,d\mu\Big)^{1/p}, \quad B = \Big(\int_\Omega g^q\,d\mu\Big)^{1/q}$$

and we observe that $\int_\Omega \big(\frac{f}{A}\big)^p\,d\mu = \int_\Omega \big(\frac{g}{B}\big)^q\,d\mu = 1$. Hence

$$\int_\Omega fg\,d\mu = AB\int_\Omega \frac{f}{A}\,\frac{g}{B}\,d\mu \le AB.$$

If one of $\int_\Omega f^p\,d\mu$, $\int_\Omega g^q\,d\mu$ is equal to 0, the inequality of (i) is obvious: it becomes $0 \le 0$.
(ii) Trivial. $\square$

If $p = q = 2$, Hölder's inequality for integrals is usually called **Schwarz's inequality** or **Buniakowsky's inequality**.

Minkowski's inequality for integrals. *Let $p \ge 1$. Let (Ω, Σ, μ) be a measure space, and $f, g \in \mathcal{M}(\Omega, \Sigma)$ (i.e. f, g are Σ-measurable in Ω) with $f, g \ge 0$ μ-a.e. in Ω.*
(i) If $\int_\Omega f^p\,d\mu < +\infty$ and $\int_\Omega g^p\,d\mu < +\infty$, then

$$\Big(\int_\Omega (f+g)^p\,d\mu\Big)^{1/p} \le \Big(\int_\Omega f^p\,d\mu\Big)^{1/p} + \Big(\int_\Omega g^p\,d\mu\Big)^{1/p}.$$

(ii) If ess-sup$_\Omega\, f < +\infty$ *and* ess-sup$_\Omega\, g < +\infty$*, then*

$$\text{ess-sup}_\Omega(f+g) \le \text{ess-sup}_\Omega\, f + \text{ess-sup}_\Omega\, g.$$

Proof. (i) The inequality of (i) is an obvious equality when $p = 1$.
Now let $p > 1$ and $q = \frac{p}{p-1}$, so that $\frac{1}{p} + \frac{1}{q} = 1$.
Since $\int_\Omega f^p\,d\mu < +\infty$ and $\int_\Omega g^p\,d\mu < +\infty$, using the inequality $(\lambda+\kappa)^p \le 2^{p-1}(\lambda^p + \kappa^p)$, we get

$$\int_\Omega (f+g)^p\,d\mu \le 2^{p-1}\int_\Omega f^p\,d\mu + 2^{p-1}\int_\Omega g^p\,d\mu < +\infty.$$

Then

$$\int_\Omega (f+g)^p\,d\mu = \int_\Omega (f+g)(f+g)^{p-1}\,d\mu = \int_\Omega f(f+g)^{p-1}\,d\mu + \int_\Omega g(f+g)^{p-1}\,d\mu$$

and, using Hölder's inequality for integrals,

$$\int_\Omega (f+g)^p\,d\mu \le \Big(\int_\Omega f^p\,d\mu\Big)^{1/p}\Big(\int_\Omega (f+g)^p\,d\mu\Big)^{1/q} + \Big(\int_\Omega g^p\,d\mu\Big)^{1/p}\Big(\int_\Omega (f+g)^p\,d\mu\Big)^{1/q}.$$

If $\int_\Omega (f+g)^p\,d\mu > 0$, we divide the last inequality with $\big(\int_\Omega (f+g)^p\,d\mu\big)^{1/q}$ and we finish the proof. If $\int_\Omega (f+g)^p\,d\mu = 0$, then the inequality of (i) is trivial: $0 \le 0 + 0$.
(ii) Trivial. $\square$

1.3 Subspaces, cartesian products, quotient spaces.

We shall now see three ways to produce new normed spaces from old ones. The first is to consider a *subspace* of a normed space. The second is to consider the *cartesian product* of normed spaces. And the third way is to consider the *quotient space* of a normed space over any subspace of it.

Proposition 1.5. *Let X be a normed space with norm $\|\cdot\| : X \to \mathbb{R}$, and let Y be a linear subspace of X. Then the restriction $\|\cdot\| : Y \to \mathbb{R}$ is a norm on Y.*

Proof. This is obvious. □

Definition. *The linear subspace Y of the normed space X, equipped with the restriction on Y of the norm on X, is called* **subspace** *of X.*

We assume that $X_1, \ldots, X_m$ are normed spaces (over the same F) with norms $\|\cdot\|_1, \ldots, \|\cdot\|_m$. We consider the *cartesian product* $X = X_1 \times \cdots \times X_m$ and for every $x = (x_1, \ldots, x_m) \in X = X_1 \times \cdots \times X_m$ we define

$$\|x\|_p = \begin{cases} (\|x_1\|_1^p + \cdots + \|x_m\|_m^p)^{1/p}, & 1 \le p < +\infty, \\ \max\{\|x_1\|_1, \ldots, \|x_m\|_m\}, & p = \infty. \end{cases}$$

Proposition 1.6. *For every $p \in [1, +\infty]$ the function $\|\cdot\|_p : X \to \mathbb{R}$ just defined is a norm on $X = X_1 \times \cdots \times X_m$.*

Proof. All properties of the norm are trivially satisfied by $\|\cdot\|_p$ except for the last one, the triangle inequality, which is implied by Minkowski's inequality for sums. □

Definition. *The norm $\|\cdot\|_p$ just defined on the cartesian product $X = X_1 \times \cdots \times X_m$ of normed spaces is called* **p-norm** *on X.*

Example 1.3.1. We consider $X_1 = \ldots = X_m = F$ with $\|\cdot\|_1 = \ldots = \|\cdot\|_m = |\cdot|$ and then we get the cartesian product $X = F \times \cdots \times F = F^m$ with the p-norm, which is defined for every $x = (\lambda_1, \ldots, \lambda_m) \in F \times \cdots \times F = F^m$ by

$$\|x\|_p = \begin{cases} (|\lambda_1|^p + \cdots + |\lambda_m|^p)^{1/p}, & 1 \le p < +\infty, \\ \max\{|\lambda_1|, \ldots, |\lambda_m|\}, & p = +\infty. \end{cases}$$

The case $p = 2$ gives the usual **euclidean norm** on F^m.

Finally, we consider a normed space X and a linear subspace Z of X. We also consider the *quotient space*

$$X/Z = \{x + Z \,|\, x \in X\}.$$

The elements of X/Z are subsets of X: they are the *parallel translations* of Z.

We know from Linear Algebra (and we can easily prove) the following facts:

1. If $\xi \in X/Z$, then: $\xi = x + Z \;\Leftrightarrow\; x \in \xi$.

2. If $\xi, \eta \in X/Z$, then: $\xi \cap \eta \neq \emptyset \;\Rightarrow\; \xi = \eta$.

3. $x + Z = y + Z \;\Leftrightarrow\; x - y \in Z$.

It is well known from Linear Algebra that the quotient space X/Z is a linear space with addition and multiplication defined by

$$(x+Z)+(y+Z)=(x+y)+Z, \quad \lambda\,(x+Z)=(\lambda x)+Z.$$

The zero element of X/Z is $0+Z=Z$.

It is easy to show that the equality $(x+Z)+(y+Z)=(x+y)+Z$ is not just a formal definition; it is a true equality between subsets of X. If $\lambda\neq 0$, then the same is true for the equality $\lambda\,(x+Z)=(\lambda x)+Z$. If $\lambda=0$, then $0\,(x+Z)=(0x)+Z$ is not true as an equality between sets: we have $0\,(x+Z)=\{0\}$ and $(0x)+Z=Z$.

Now we define the function

$$\|\cdot\|_{X/Z}:X/Z\to\mathbb{R}$$

by

$$\|\xi\|_{X/Z}=\inf\{\|x\|\,|\,x\in\xi\}$$

for every $\xi\in X/Z$.

Proposition 1.7. *If Z is a closed linear subspace of the normed space X, then the function $\|\cdot\|_{X/Z}$: $X/Z\to\mathbb{R}$ just defined is a norm on X/Z.*

Proof. (i) It is obvious that $\|\xi\|_{X/Z}\geq 0$ for every $\xi\in X/Z$.
(ii) If $\xi=Z$, i.e. if ξ is the zero element of X/Z, then $0\in\xi$ and so $0\leq\|\xi\|_{X/Z}\leq\|0\|=0$ and hence $\|\xi\|_{X/Z}=0$.
Conversely, let $\|\xi\|_{X/Z}=0$. Then there are $x_n\in\xi$ so that $\|x_n\|\to 0$, i.e. $x_n\to 0$. But ξ is a closed subset of X, since it is a translation of the closed set Z. Hence $0\in\xi$ and so $\xi=0+Z=Z$, the zero element of X/Z.
(iii) If $\lambda=0$, then $0\xi=Z$ and so $\|0\xi\|_{X/Z}=\|Z\|_{X/Z}=0$. Also, trivially $|0|\|\xi\|_{X/Z}=0$. So the equality $\|0\xi\|_{X/Z}=|0|\|\xi\|_{X/Z}$ is correct.
Now let $\lambda\neq 0$ and $\xi\in X/Z$. We take any $x\in\xi$ and then we have $\lambda x\in\lambda\xi$. Therefore

$$\|\lambda\xi\|_{X/Z}\leq\|\lambda x\|=|\lambda|\|x\|.$$

Taking the infimum over all $x\in\xi$, we find

$$\|\lambda\xi\|_{X/Z}\leq|\lambda|\|\xi\|_{X/Z}.$$

We apply this to $\frac{1}{\lambda}$ in the place of λ and to $\lambda\xi$ in the place of ξ, and we get

$$\|\xi\|_{X/Z}\leq\frac{1}{|\lambda|}\|\lambda\xi\|_{X/Z}$$

and so

$$|\lambda|\|\xi\|_{X/Z}\leq\|\lambda\xi\|_{X/Z}.$$

The two inequalities imply

$$\|\lambda\xi\|_{X/Z}=|\lambda|\|\xi\|_{X/Z}.$$

(iv) Let $\xi,\eta\in X/Z$. We take any $x\in\xi$ and any $y\in\eta$. Then $x+y\in\xi+\eta$ and hence

$$\|\xi+\eta\|_{X/Z}\leq\|x+y\|\leq\|x\|+\|y\|.$$

Taking the infimum over all $x\in\xi$ and, independently, over all $y\in\eta$, we find

$$\|\xi+\eta\|_{X/Z}\leq\|\xi\|_{X/Z}+\|\eta\|_{X/Z}.$$

□

1.4 Banach spaces.

Definition. *The normed space X is called* **Banach space** *if it is complete, i.e. if every Cauchy sequence in X converges to an element of X.*

Proposition 1.8. *Let X be a normed space with norm $\|\cdot\|$. Then the following are equivalent.*
(i) X is a Banach space.
(ii) For every sequence (x_n) in X: if $\sum_{n=1}^{+\infty}\|x_n\| < +\infty$, then $\sum_{n=1}^{+\infty} x_n$ converges to an element of X.

Proof. Let X be complete and let $\sum_{n=1}^{+\infty}\|x_n\| < +\infty$. We consider the partial sums $s_n = x_1 + \cdots + x_n$ and then for $n < m$ we have

$$\|s_m - s_n\| = \|x_{n+1} + \cdots + x_m\| \leq \|x_{n+1}\| + \cdots + \|x_m\| \to 0$$

when $m, n \to +\infty$. Thus (s_n) is a Cauchy sequence and so it converges to an element of X.
Conversely, we assume that (ii) holds. We take any Cauchy sequence (x_n) in X. Then for every k there is $n_k \in \mathbb{N}$ so that $\|x_n - x_m\| < \frac{1}{k^2}$ when $n, m \geq n_k$. We may choose n_k so that (n_k) is strictly increasing and then $\|x_{n_{k+1}} - x_{n_k}\| < \frac{1}{k^2}$ for every k. Thus

$$\|x_{n_1}\| + \sum_{k=1}^{+\infty}\|x_{n_{k+1}} - x_{n_k}\| < +\infty.$$

By our assumption, the series $x_{n_1} + \sum_{k=1}^{+\infty}(x_{n_{k+1}} - x_{n_k})$ converges to some element $x \in X$. Observing the telescoping partial sums of the last series we get that $x_{n_k} \to x$. Since (x_n) is Cauchy,

$$x_k = (x_k - x_{n_k}) + x_{n_k} \to 0 + x = x$$

when $k \to +\infty$.
Therefore X is complete. $\square$

Proposition 1.9. *Let X be a Banach space and Y be a subspace of X. Then Y is a Banach space if and only if Y is closed.*

Proof. Let Y be a Banach space. We take any sequence (y_n) in Y so that $y_n \to y \in X$. Since (y_n) converges, it is a Cauchy sequence. But Y is complete, so (y_n) converges to an element of Y. Since the limit of a sequence is unique, we get that $y \in Y$.
Therefore Y is closed.
Conversely, let Y be closed. We take any Cauchy sequence (y_n) in Y. Since X is complete, $y_n \to y$ for some $y \in X$. Since Y is closed, we get $y \in Y$.
Hence Y is complete. $\square$

Observe that in the first part of the last proof the assumption of the completeness of X was *not* used. Therefore,
If Y is a complete subspace of a normed space X, then Y is closed in X.

Proposition 1.10. *Let $X_1, \ldots, X_m$ be Banach spaces with norms $\|\cdot\|_1, \ldots, \|\cdot\|_m$. Then the product space $X = X_1 \times \cdots \times X_m$ equipped with any of the p-norms is a Banach space.*

Proof. Let (x_n) be a Cauchy sequence in X, where $x_n = (x_{n,1}, \ldots, x_{n,m})$ for every n. Clearly, for every $j = 1, \ldots, m$ we have

$$\|x_{n,j} - x_{k,j}\|_j \leq \|x_n - x_k\|_p \to 0$$

as $n, k \to +\infty$ and so $(x_{n,j})$ is a Cauchy sequence in X_j. Thus $x_{n,j} \to x_j$ for some $x_j \in X_j$. We consider the element $x = (x_1, \ldots, x_m) \in X$ and then

$$\|x_n - x\|_p = (\|x_{n,1} - x_1\|_1^p + \cdots + \|x_{n,m} - x_m\|_m^p)^{1/p} \to 0$$

when $1 \leq p < +\infty$, and

$$\|x_n - x\|_\infty = \max\{\|x_{n,1} - x_1\|_1, \ldots, \|x_{n,m} - x_m\|_m\} \to 0$$

when $p = +\infty$.
So X is complete. □

Proposition 1.11. *Let X be a Banach space and Z be a closed subspace of X. Then X/Z is a Banach space.*

Proof. We consider $\xi_n \in X/Z$ so that

$$\sum_{n=1}^{+\infty} \|\xi_n\|_{X/Z} < +\infty.$$

Since $\|\xi_n\|_{X/Z} = \inf\{\|x\| \mid x \in \xi_n\}$, there is some $x_n \in \xi_n$ so that

$$\|x_n\| < \|\xi_n\|_{X/Z} + \frac{1}{n^2}.$$

Therefore

$$\sum_{n=1}^{+\infty} \|x_n\| < +\infty.$$

Since X is a Banach space, the series $\sum_{n=1}^{+\infty} x_n$ converges to an element $s \in X$, i.e.

$$x_1 + \cdots + x_n \to s$$

when $n \to +\infty$. We consider $\eta = s + Z \in X/Z$. Then $x_1 + \cdots + x_n \in \xi_1 + \cdots + \xi_n$ and $s \in \eta$ and hence $(x_1 + \cdots + x_n) - s \in (\xi_1 + \cdots + \xi_n) - \eta$. Thus

$$\|(\xi_1 + \cdots + \xi_n) - \eta\|_{X/Z} \leq \|(x_1 + \cdots + x_n) - s\| \to 0$$

when $n \to +\infty$. We conclude that the series $\sum_{n=1}^{+\infty} \xi_n$ converges to an element of X/Z.
So X/Z is a Banach space. □

1.5 Linear isometries.

Definition. *Let X, Y be normed spaces with norms $\|\cdot\|_X, \|\cdot\|_Y$, and let $T : X \to Y$ be a linear operator with the property*

$$\|T(x)\|_Y = \|x\|_X$$

for every $x \in X$. Then we say that T is a **linear isometry** *of X* **into** *Y. It is clear that $T(x) = 0$ implies $x = 0$, and so T is one-to-one.*
If T is onto Y, i.e. if $T(X) = Y$, then we say that T is a **linear isometry** *of X* **onto** *Y. We also say that X is* **linearly isometric** *to Y.*

It is easy to see that the relation between normed spaces of being linearly isometric is an equivalence relation.

A linear isometry $T : X \to Y$ is continuous. Indeed, if $x_n \to x$ in X, then

$$\|T(x_n) - T(x)\|_Y = \|T(x_n - x)\|_Y = \|x_n - x\|_X \to 0$$

and thus $T(x_n) \to T(x)$ in Y.

If T is a linear isometry of X into Y, then we may "identify" every $x \in X$ with the corresponding $T(x) \in Y$ and so we may "identify" X with the subspace $T(X)$ of Y.

Proposition 1.12. *Let X be a normed space with norm $\|\cdot\|_X$, let Y be a linear space and let $T : X \to Y$ be a linear operator which is one-to-one in X and onto Y. Then there is a norm on Y so that T becomes a linear isometry of X onto Y.*

Proof. We take any $y \in Y$, we consider the unique $x \in X$ so that $T(x) = y$, and we define

$$\|y\|_Y = \|x\|_X.$$

We can easily prove that the function $\|\cdot\|_Y : Y \to \mathbb{R}$ just defined is a norm on Y.
Of course, since $T(x) = y$, the equality $\|y\|_Y = \|x\|_X$ can be written $\|T(x)\|_Y = \|x\|_X$ and so T is a linear isometry of X onto Y. □

In other words, when we have two isomorphic linear spaces and one of them has a norm, then we can transfer this norm to the other linear space so that the two spaces become linearly isometric.

Example 1.5.1. Let X be a linear space of finite dimension and let $\{b_1, \dots, b_m\}$ be a basis of X. We consider the normed space F^m with any of the p-norms $\|\cdot\|_p$, $1 \le p \le +\infty$. We also consider the linear operator $T : F^m \to X$ defined for every $(\lambda_1, \dots, \lambda_m) \in F^m$ by

$$T(\lambda_1, \dots, \lambda_m) = \lambda_1 b_1 + \cdots + \lambda_m b_m.$$

Then T is one-to-one in F^m and onto X, and so the p-norm on F^m can be transfered to a norm $\|\cdot\|_p : X \to \mathbb{R}$. This norm is defined for every $x = \lambda_1 b_1 + \cdots + \lambda_m b_m \in X$ by the formula

$$\begin{aligned}\|x\|_p &= \|\lambda_1 b_1 + \cdots + \lambda_m b_m\|_p = \|T(\lambda_1, \dots, \lambda_m)\|_p = \|(\lambda_1, \dots, \lambda_m)\|_p \\ &= \begin{cases} (|\lambda_1|^p + \cdots + |\lambda_m|^p)^{1/p}, & 1 \le p < +\infty, \\ \max\{|\lambda_1|, \dots, |\lambda_m|\}, & p = +\infty. \end{cases}\end{aligned}$$

The norm $\|\cdot\|_p$ on X just defined is called **p-norm** on X with respect to the basis $\{b_1, \dots, b_m\}$. Of course, if we change the basis of X, then we shall get a different norm on X: the coefficient m-tuple $(\lambda_1, \dots, \lambda_m)$ of any $x \in X$ depends on the basis.

1.6 Equivalent norms.

Definition. *Two norms $\|\cdot\|$ and $|||\cdot|||$ on the same linear space X are called* **equivalent** *if there are constatns $c, C > 0$ so that*

$$c\|x\| \le |||x||| \le C\|x\|$$

for every $x \in X$.

Proposition 1.13. *Let $\|\cdot\|$, $|||\cdot|||$ be two norms on the linear space X. The following are equivalent.*
(i) The norms are equivalent.
(ii) For every sequence (x_n) in X: $|||x_n||| \to 0$ if and only if $\|x_n\| \to 0$.

Proof. Assume that the two norms are equivalent, i.e. that

$$c\|x\| \le |||x||| \le C\|x\|$$

for every $x \in X$, and let $\|x_n\| \to 0$. Then

$$|||x_n||| \le C\|x_n\| \to 0$$

and hence $|||x_n||| \to 0$. In the same manner, if $|||x_n||| \to 0$ we get that $\|x_n\| \to 0$.
For the converse we assume that there is no $c > 0$ so that $c\|x\| \le |||x|||$ for every $x \in X$. Therefore, for every $n \in \mathbb{N}$ there is $x_n \in X$ so that $\frac{1}{n}\|x_n\| > |||x_n|||$. We consider the elements

$$y_n = \frac{1}{\|x_n\|}\, x_n$$

for which we have

$$\|y_n\| = 1, \quad |||y_n||| < \frac{1}{n}.$$

Then $|||y_n||| \to 0$ but $\|y_n\| \not\to 0$, and we get a contradiction to (ii). In the same manner we get a contradiction to (ii) if there is no $C > 0$ so that $|||x||| \leq C\|x\|$ for every $x \in X$. $\square$

So we see that, if two norms on the same linear space are equivalent, then a sequence (x_n) converges to x with respect to one of the norms if and only if (x_n) converges to x with respect to other norm.

Assume again that the norms $\|\cdot\|$ and $|||\cdot|||$ on the same linear space X are equivalent, i.e. they satisfy $c\|x\| \leq |||x||| \leq C\|x\|$ for every $x \in X$. If $B(a; r)$ is a ball with respect to the norm $\|\cdot\|$ and $D(a; r)$ is a ball with respect to the norm $|||\cdot|||$, then

$$D(a; cr) \subseteq B(a; r) \subseteq D(a; Cr).$$

Therefore, if a set $A \subseteq X$ is open with respect to one of the norms, then A is open with respect to the other norm. Since the closed sets are the complements of the open sets, the same is true for closed subsets of X. And, since the notion of compact set depends solely on the notion of open set, the same is true for compact subsets of X. Finally, if a set A is contained in a ball with respect to one of the norms, then it is contained in a ball with respect to the other norm.

In other words, equivalent norms define the same convergent sequences (with the same limits) and the same open, closed, compact, and bounded sets.

1.7 Finite dimensional normed spaces.

Proposition 1.14. *Let X be a linear space of finite dimension. Then every two norms on X are equivalent.*

Proof. Let $\{b_1, \ldots, b_m\}$ be a basis of X. We consider the 2-norm on X defined for every $x = \lambda_1 b_1 + \cdots + \lambda_m b_m \in X$ by

$$\|x\|_2 = (|\lambda_1|^2 + \cdots + |\lambda_m|^2)^{1/2}.$$

We shall prove that every other norm $\|\cdot\|$ on X is equivalent to $\|\cdot\|_2$.
Initially, for every $x = \lambda_1 b_1 + \cdots + \lambda_m b_m \in X$ we get

$$\|x\| \leq |\lambda_1|\,\|b_1\| + \cdots + |\lambda_m|\,\|b_m\| \leq (\|b_1\|^2 + \cdots + \|b_m\|^2)^{1/2}(|\lambda_1|^2 + \cdots + |\lambda_m|^2)^{1/2} = C\|x\|_2,$$

where $C = (\|b_1\|^2 + \cdots + \|b_m\|^2)^{1/2}$. The second inequality above is Cauchy's inequality.
Now assume that there is no $c > 0$ so that $c\|x\|_2 \leq \|x\|$ for every $x \in X$. Then, as in the proof of proposition 1.13, we see that there is a sequence (x_n) in X so that

$$\|x_n\|_2 = 1, \quad \|x_n\| \to 0.$$

If $x_n = \lambda_{n,1} b_1 + \cdots + \lambda_{n,m} b_m$, then from the last equality we get that

$$|\lambda_{n,1}|^2 + \cdots + |\lambda_{n,m}|^2 = 1$$

for every n.
Since the unit sphere of F^m is a compact set, there is a subsequence (x_{n_k}) of (x_n) so that

$$(\lambda_{n_k,1}, \ldots, \lambda_{n_k,m}) \to (\lambda_1, \ldots, \lambda_m)$$

for some $(\lambda_1, \ldots, \lambda_m) \in F^m$ satisfying

$$|\lambda_1|^2 + \cdots + |\lambda_m|^2 = 1.$$

We consider the element

$$x = \lambda_1 b_1 + \cdots + \lambda_m b_m$$

of X and then we have

$$\|x_{n_k} - x\| \le |\lambda_{n_k,1} - \lambda_1| \, \|b_1\| + \cdots + |\lambda_{n_k,m} - \lambda_m| \, \|b_m\| \to 0$$

and

$$\|x_{n_k} - x\|_2 = (|\lambda_{n_k,1} - \lambda_1|^2 + \cdots + |\lambda_{n_k,m} - \lambda_m|^2)^{1/2} \to 0.$$

Hence $x_{n_k} \to x$ with respect to both norms, and so

$$\|x\|_2 = 1, \quad \|x\| = 0.$$

This is impossible. □

Proposition 1.15. *Let X be a normed space of finite dimension. Then,*
(i) every closed and bounded subset of X is compact.
(ii) X is a Banach space.

Proof. (i) Let $\{b_1, \ldots, b_m\}$ be a basis of X. Besides the norm $\| \cdot \|$ on X, we also consider the 2-norm on X defined for every $x = \lambda_1 b_1 + \cdots + \lambda_m b_m \in X$ by

$$\|x\|_2 = (|\lambda_1|^2 + \cdots + |\lambda_m|^2)^{1/2}.$$

We also consider the linear operator

$$T : F^m \to X$$

defined for every $(\lambda_1, \ldots, \lambda_m) \in F^m$ by

$$T(\lambda_1, \ldots, \lambda_m) = \lambda_1 b_1 + \cdots + \lambda_m b_m.$$

As we have already observed, T is one-to-one in F^m and onto X. Moreover, T and T^{-1} are linear isometries between F^m and X, if we consider the two spaces equipped with their 2-norms.
Now, let $K \subseteq X$ be closed and bounded (with respect to the norm $\| \cdot \|$). Since every two norms on X are equivalent, K is closed and bounded with respect to the 2-norm on X. Now, since $T : F^m \to X$ is a linear isometry with respect to the 2-norms on F^m and X, $T^{-1}(K)$ is closed and bounded in F^m. But F^m with its 2-norm is the standard euclidean space and so $T^{-1}(K)$ is compact. Therefore $K = T(T^{-1}(K))$ is compact in X with respect to the 2-norm on X. Finally, since the norm $\| \cdot \|$ and the 2-norm on X are equivalent, K is compact in X with respect to its original norm $\| \cdot \|$.
(ii) Let (x_n) be a Cauchy sequence in X. Then (x_n) is bounded, i.e. it is contained in some closed ball $\overline{B}(0; r)$. By (i), this closed ball is compact and so (x_n) has a convergent subsequence. Since (x_n) is a Cauchy sequence, it is convergent. □

Proposition 1.16. *Let X be a normed space and let Y be a subspace of X of finite dimension. Then Y is closed.*

Proof. Since Y is a normed space of finite dimension, it is a complete subspace of X and hence closed in X. □

1.8 Completion.

Definition. *Let X be a normed space. We say that the normed space $\overline{X}$ is a* **completion** *of X if $\overline{X}$ is complete, i.e. a Banach space, and there is a linear isometry $T : X \to \overline{X}$ so that $T(X)$ is a dense subspace of $\overline{X}$.*

In other words, a Banach space $\overline{X}$ is a completion of X if X is linearly isometric to a dense subspace of $\overline{X}$.

A trivial case is when X itself is complete. Then we may consider $\overline{X} = X$, i.e. X is a completion of itself.

Theorem 1.1. *Let X be a normed space. Then there is at least one completion of X. Moreover, every two completions of X are linearly isometric.*

Proof. We shall construct a completion $\overline{X}$ of X.
We first consider the set of all Cauchy sequences in X:

$$\widehat{X} = \{(x_n) \,|\, (x_n) \text{ is a Cauchy sequence in } X\}.$$

Then we consider a relation between Cauchy sequences:

$$(x_n) \equiv (y_n) \quad \text{if} \quad x_n - y_n \to 0 \text{ in } X.$$

This is obviously an equivalence relation in $\widehat{X}$, and so we may consider the quotient space consisting of all equivalence classes:

$$\overline{X} = \{[(x_n)] \,|\, (x_n) \in \widehat{X}\} = \widehat{X}/\equiv$$

We define algebraic operations in $\overline{X}$:

$$[(x_n)] + [(y_n)] = [(x_n + y_n)], \quad \lambda\,[(x_n)] = [(\lambda x_n)].$$

(It is easy to check that these are well defined.) Thus, $\overline{X}$ is a linear space over F.
If (x_n) is a Cauchy sequence in X, then it is easy to see that $(\|x_n\|)$ is a Cauchy sequence in $\mathbb{R}$ and so it converges to some real number. Hence we may define a norm on $\overline{X}$ by:

$$\|[(x_n)]\|_{\overline{X}} = \lim_{n\to+\infty} \|x_n\|.$$

(Again, it is easy to check that this is well defined and that it has the properties of a norm.) So $\overline{X}$ is a normed space.
Next we consider the linear operator

$$T : X \to \overline{X}$$

defined for every $x \in X$ by

$$T(x) = [(x)],$$

where (x) is, of course, the constant sequence $x, x, x, \ldots$. It is easy to see that T is a linear operator, and that T is a linear isometry of X into $\overline{X}$:

$$\|T(x)\|_{\overline{X}} = \|[(x)]\|_{\overline{X}} = \lim_{n\to+\infty} \|x\| = \|x\|.$$

Now, take any $[(x_n)] \in \overline{X}$ and any $\epsilon > 0$. Since (x_n) is a Cauchy sequence in X, there is n_0 so that $\|x_n - x_m\| < \epsilon$ for every $n, m \geq n_0$. Therefore,

$$\|[(x_n)] - T(x_{n_0})\|_{\overline{X}} = \|[(x_n)] - [(x_{n_0})]\|_{\overline{X}} = \|[(x_n - x_{n_0})]\|_{\overline{X}} = \lim_{n\to+\infty} \|x_n - x_{n_0}\| \leq \epsilon.$$

This means that $T(X)$ is dense in $\overline{X}$.
Finally, let (ξ_n) be a Cauchy sequence in $\overline{X}$. Since $T(X)$ is dense in $\overline{X}$, for every n there is some $x_n \in X$ so that

$$\|\xi_n - T(x_n)\|_{\overline{X}} < \frac{1}{n}.$$

Then we get

$$\begin{aligned}\|x_n - x_m\| &= \|T(x_n - x_m)\|_{\overline{X}} = \|T(x_n) - T(x_m)\|_{\overline{X}} \\ &\le \|T(x_n) - \xi_n\|_{\overline{X}} + \|\xi_n - \xi_m\|_{\overline{X}} + \|\xi_m - T(x_m)\|_{\overline{X}} \to 0\end{aligned}$$

when $n, m \to +\infty$ and so (x_n) is a Cauchy sequence in X. We now consider the element $\xi = [(x_n)]$ of $\overline{X}$ and we get

$$\|\xi_m - \xi\|_{\overline{X}} \le \|\xi_m - T(x_m)\|_{\overline{X}} + \|T(x_m) - \xi\|_{\overline{X}} < \frac{1}{m} + \|[(x_m)] - [(x_n)]\|_{\overline{X}}$$

where (x_m) is the constant sequence $x_m, x_m, \ldots$. So

$$\|\xi_m - \xi\|_{\overline{X}} \le \frac{1}{m} + \|[(x_m - x_n)]\|_{\overline{X}} = \frac{1}{m} + \lim_{n\to+\infty} \|x_m - x_n\| \to 0$$

when $m \to +\infty$.
We conclude that every Cauchy sequence in $\overline{X}$ converges to an element of $\overline{X}$.
Now, assume that $\overline{X}_1$ and $\overline{X}_2$ are two completions of X. Thus there are linear isometries $T_1 : X \to \overline{X}_1$ and $T_2 : X \to \overline{X}_2$ so that $T_1(X)$ and $T_2(X)$ are dense subspaces of $\overline{X}_1$ and $\overline{X}_2$.
We take any $\xi_1 \in \overline{X}_1$. Then there is a sequence (x_n) in X so that $T_1(x_n) \to \xi_1$ in $\overline{X}_1$. So $(T_1(x_n))$ is a Cauchy sequence in $\overline{X}_1$, and since T_1 is a linear isometry, (x_n) is a Cauchy sequence in X. Now, since T_2 is a linear isometry, $(T_2(x_n))$ is a Cauchy sequence in $\overline{X}_2$. But $\overline{X}_2$ is complete, and so there is some $\xi_2 \in \overline{X}_2$ so that $T_2(x_n) \to \xi_2$ in $\overline{X}_2$. Now, this procedure defines a function

$$T : \overline{X}_1 \to \overline{X}_2$$

so that for every $\xi_1 \in \overline{X}_1$ we have

$$T(\xi_1) = \xi_2.$$

Taking this procedure backwards from an arbitrary $\xi_2 \in \overline{X}_2$ to $\xi_1 \in \overline{X}_1$ we see that T is one-to-one in $\overline{X}_1$ and onto $\overline{X}_2$. It is also easy to check that T is a linear operator, and that it is a linear isometry

$$\|T(\xi_1)\|_{\overline{X}_2} = \|\xi_2\|_{\overline{X}_2} = \lim_{n\to+\infty} \|T_2(x_n)\|_{\overline{X}_2} = \lim_{n\to+\infty} \|x_n\| = \lim_{n\to+\infty} \|T_1(x_n)\|_{\overline{X}_1} = \|\xi_1\|_{\overline{X}_1}.$$

Therefore, $\overline{X}_1$ and $\overline{X}_2$ are linearly isometric. □

1.9 Sequence spaces.

Definition. *We define the following spaces whose elements are sequences in F:*

$$\begin{aligned} c &= \{(\lambda_k) \,|\, (\lambda_k) \text{ converges in } F\} \\ c_0 &= \{(\lambda_k) \,|\, \lambda_k \to 0 \text{ in } F\} \\ l^\infty &= \{(\lambda_k) \,|\, (\lambda_k) \text{ is bounded}\} \\ l^p &= \Big\{(\lambda_k) \,\Big|\, \sum_{k=1}^{+\infty} |\lambda_k|^p < +\infty\Big\}, \quad 1 \le p < +\infty. \end{aligned}$$

The algebraic operations in all these spaces are defined component-wise as usual:

$$(\lambda_k)+(\kappa_k)=(\lambda_k+\kappa_k),\quad \lambda\,(\lambda_k)=(\lambda\lambda_k).$$

These operations are well defined in these spaces, since if $(\lambda_k),(\kappa_k)$ are convergent, or convergent to 0, or bounded, then $(\lambda_k+\kappa_k),(\lambda\lambda_k)$ are also convergent, or convergent to 0, or bounded. Regarding the last space, we observe that if $(\lambda_k)\in l^p$, i.e. if $\sum_{k=1}^{\infty}|\lambda_k|^p<+\infty$, then

$$\sum_{k=1}^{+\infty}|\lambda\lambda_k|^p=|\lambda|^p\sum_{k=1}^{+\infty}|\lambda_k|^p<+\infty$$

and hence $(\lambda\lambda_k)\in l^p$. Also if $(\lambda_k),(\kappa_k)\in l^p$, i.e. if $\sum_{k=1}^{+\infty}|\lambda_k|^p<+\infty$ and $\sum_{k=1}^{+\infty}|\kappa_k|^p<+\infty$, then, as we saw in the proof of Minkowski's inequality for sums,

$$\sum_{k=1}^{+\infty}|\lambda_k+\kappa_k|^p\le\sum_{k=1}^{+\infty}(|\lambda_k|+|\kappa_k|)^p\le 2^{p-1}\sum_{k=1}^{+\infty}|\lambda_k|^p+2^{p-1}\sum_{k=1}^{+\infty}|\kappa_k|^p<+\infty,$$

and hence $(\lambda_k+\kappa_k)\in l^p$.

Thus all these sequence spaces are linear spaces over F.

We have the obvious inclusions

$$l^p\subseteq c_0\subseteq c\subseteq l^\infty.$$

We can also prove that

$$l^p\subseteq l^q,\quad \text{if } 1\le p<q<+\infty.$$

Indeed, if $(\lambda_k)\in l^p$, then $\lambda_k\to 0$, and so there is k_0 so that $|\lambda_k|\le 1$ for every $k\ge k_0$. Then

$$\sum_{k=k_0}^{+\infty}|\lambda_k|^q\le\sum_{k=k_0}^{+\infty}|\lambda_k|^p<+\infty,$$

and so $(\lambda_k)\in l^q$.

Definition. *If $1\le p\le+\infty$, we consider the function $\|\cdot\|_p:l^p\to\mathbb{R}$ defined for every $x=(\lambda_k)\in l^p$ by*

$$\|x\|_p=\begin{cases}\big(\sum_{k=1}^{+\infty}|\lambda_k|^p\big)^{1/p}, & 1\le p<+\infty,\\ \sup_k|\lambda_k|, & p=+\infty.\end{cases}$$

Minkowski's inequality for sums implies that $\|\cdot\|_p$ is a norm on l^p; it is the **p-norm** of l^p.

Theorem 1.2. *If $1\le p\le+\infty$, then l^p with the norm $\|\cdot\|_p$ is a Banach space.*

Proof. We consider the case $1\le p<+\infty$.
We take a Cauchy sequence (x_n) in l^p. If $x_n=(\lambda_{n,k})$ for every n, then for every k we have

$$|\lambda_{n,k}-\lambda_{m,k}|\le\|x_n-x_m\|_p\to 0$$

when $n,m\to+\infty$. Since F is complete, for every k there is $\lambda_k\in F$ so that $\lambda_{n,k}\to\lambda_k$ when $n\to+\infty$, and we consider the sequence

$$x=(\lambda_k).$$

We take n_0 so that $\|x_n-x_m\|_p<1$ for every $n,m\ge n_0$. Then for every K and every $n\ge n_0$ we get

$$\Big(\sum_{k=1}^{K}|\lambda_{n,k}|^p\Big)^{1/p}\le\Big(\sum_{k=1}^{+\infty}|\lambda_{n,k}|^p\Big)^{1/p}=\|x_n\|_p\le\|x_n-x_{n_0}\|_p+\|x_{n_0}\|_p<1+\|x_{n_0}\|_p.$$

Taking the limit first when $n \to +\infty$ and next when $K \to +\infty$, we find

$$\Big(\sum_{k=1}^{+\infty} |\lambda_k|^p\Big)^{1/p} \le 1 + \|x_{n_0}\|_p < +\infty$$

and so $x \in l^p$.
Now we take any $\epsilon > 0$ and a coresponding n_0 so that $\|x_n - x_m\|_p < \epsilon$ for every $n, m \ge n_0$. Then for every K and every $n, m \ge n_0$ we get

$$\Big(\sum_{k=1}^{K} |\lambda_{n,k} - \lambda_{m,k}|^p\Big)^{1/p} \le \Big(\sum_{k=1}^{+\infty} |\lambda_{n,k} - \lambda_{m,k}|^p\Big)^{1/p} = \|x_n - x_m\|_p < \epsilon.$$

Taking the limit first when $m \to +\infty$ and next when $K \to +\infty$, we find

$$\|x_n - x\|_p = \Big(\sum_{k=1}^{+\infty} |\lambda_{n,k} - \lambda_k|^p\Big)^{1/p} \le \epsilon$$

for every $n \ge n_0$. Thus $x_n \to x$ in l^p.
Finally, let $p = +\infty$ and consider a Cauchy sequence (x_n) in l^∞. If $x_n = (\lambda_{n,k})$ for every n, then for every k we have

$$|\lambda_{n,k} - \lambda_{m,k}| \le \|x_n - x_m\|_\infty \to 0$$

when $n, m \to +\infty$. Again, since F is complete, for every k there is $\lambda_k \in F$ so that $\lambda_{n,k} \to \lambda_k$ when $n \to +\infty$. Define

$$x = (\lambda_k).$$

We take n_0 so that $\|x_n - x_m\|_\infty < 1$ for every $n, m \ge n_0$. Then for every k and every $n \ge n_0$ we have

$$|\lambda_{n,k}| \le \|x_n\|_\infty \le \|x_n - x_{n_0}\|_\infty + \|x_{n_0}\|_\infty < 1 + \|x_{n_0}\|_\infty.$$

Taking the limit when $n \to +\infty$, we find

$$|\lambda_k| \le 1 + \|x_{n_0}\|_\infty < +\infty$$

for every k, and so $x \in l^\infty$.
Now we take any $\epsilon > 0$ and a coresponding n_0 so that $\|x_n - x_m\|_\infty < \epsilon$ for every $n, m \ge n_0$. Then for every k and every $n, m \ge n_0$ we get

$$|\lambda_{n,k} - \lambda_{m,k}| \le \|x_n - x_m\|_\infty < \epsilon.$$

Taking the limit when $m \to +\infty$, we find

$$|\lambda_{n,k} - \lambda_k| \le \epsilon$$

for every k and every $n \ge n_0$. Thus $\|x_n - x\|_\infty \le \epsilon$ for every $n \ge n_0$ and so $x_n \to x$ in l^∞. $\square$

Now c and c_0 are linear subspaces of l^∞, and so they are normed spaces equipped with the restriction of the norm $\|\cdot\|_\infty$ on each of them.

Theorem 1.3. *The spaces c, c_0 with the norm $\|\cdot\|_\infty$ are Banach spaces.*

Proof. Since l^∞ is a Banach space, it is enough to prove that c, c_0 are closed in l^∞.
Let (x_n) be a sequence in c and $x_n \to x$ in l^∞. Let $x_n = (\lambda_{n,k})$ for every n, and $x = (\lambda_k)$.
For any $\epsilon > 0$ there is n_0 so that $\|x_n - x\|_\infty < \epsilon$ for every $n \ge n_0$. Then for every k we have

$$|\lambda_{n_0,k} - \lambda_k| \le \|x_{n_0} - x\|_\infty < \epsilon.$$

Since $(\lambda_{n_0,k})$ is a Cauchy sequence in F, there is k_0 so that $|\lambda_{n_0,k} - \lambda_{n_0,l}| < \epsilon$ for every $k, l \geq k_0$. Then

$$|\lambda_k - \lambda_l| \leq |\lambda_k - \lambda_{n_0,k}| + |\lambda_{n_0,k} - \lambda_{n_0,l}| + |\lambda_{n_0,l} - \lambda_l| < 3\epsilon$$

for every $k, l \geq k_0$. Therefore $x = (\lambda_k)$ is a Cauchy sequence in F and so it belongs to c.
Now let (x_n) be a sequence in c_0 and $x_n \to x$ in l^∞. Let $x_n = (\lambda_{n,k})$ for every n, and $x = (\lambda_k)$. For any $\epsilon > 0$ there is n_0 so that $\|x_n - x\|_\infty < \epsilon$ for every $n \geq n_0$. Then for every k we have

$$|\lambda_{n_0,k} - \lambda_k| \leq \|x_{n_0} - x\|_\infty < \epsilon.$$

Since $\lambda_{n_0,k} \to 0$ when $k \to +\infty$, there is k_0 so that $|\lambda_{n_0,k}| < \epsilon$ for every $k \geq k_0$. Therefore,

$$|\lambda_k| \leq |\lambda_k - \lambda_{n_0,k}| + |\lambda_{n_0,k}| < 2\epsilon$$

for every $k \geq k_0$, and so $x \in c_0$. □

Definition. *We define the sequence space*

$$c_{00} = \{(\lambda_k) \mid \lambda_k = 0 \text{ after some value of } k\}.$$

It is obvious that c_{00} is a linear subspace of all previous spaces c, c_0, and l^p, $1 \leq p \leq +\infty$. So in each of these spaces c_{00} is a subspace when we consider the norm of the space restricted to c_{00}.

Proposition 1.17. *c_{00} is a dense subspace of each of the spaces c_0, and l^p, $1 \leq p < +\infty$.*

Proof. We take any $x \in c_0$, with $x = (\lambda_k)$, and any $\epsilon > 0$. Since $\lambda_k \to 0$, there is k_0 so that $|\lambda_k| < \epsilon$ for every $k \geq k_0$. Now we take the sequence $y = (\kappa_k)$, where $\kappa_k = \lambda_k$ for $k < k_0$, and $\kappa_k = 0$ for $k \geq k_0$. Then $y \in c_{00}$. Also, $\kappa_k - \lambda_k = 0$ for $k < k_0$, and $\kappa_k - \lambda_k = -\lambda_k$ for $k \geq k_0$. Then

$$\|y - x\|_\infty = \sup_k |\kappa_k - \lambda_k| = \sup_{k \geq k_0} |\lambda_k| \leq \epsilon.$$

So c_{00} is dense in c_0.
Now let $1 \leq p < +\infty$ and take any $x \in l^p$, with $x = (\lambda_k)$, and any $\epsilon > 0$. Since $\sum_{k=1}^{+\infty} |\lambda_k|^p < +\infty$, there is k_0 so that $\sum_{k=k_0}^{+\infty} |\lambda_k|^p < \epsilon^p$. Now, as before, we take the sequence $y = (\kappa_k)$, where $\kappa_k = \lambda_k$ for $k < k_0$, and $\kappa_k = 0$ for $k \geq k_0$. Then $y \in c_{00}$. Also, $\kappa_k - \lambda_k = 0$ for $k < k_0$, and $\kappa_k - \lambda_k = -\lambda_k$ for $k \geq k_0$. Therefore,

$$\|y - x\|_p = \Big(\sum_{k=1}^{+\infty} |\kappa_k - \lambda_k|^p \Big)^{1/p} = \Big(\sum_{k=k_0}^{+\infty} |\lambda_k|^p \Big)^{1/p} < \epsilon.$$

So c_{00} is dense in l^p. □

The space c_{00} with the norm $\|\cdot\|_\infty$ is certainly not complete. To see this we consider any element x of $c_0 \setminus c_{00}$. For example, we may take x to be any sequence in F which converges to 0 and whose terms are all $\neq 0$. Since c_{00} is dence in c_0, there is a sequence (x_n) in c_{00} so that $x_n \to x$ in c_0. Then (x_n) is a Cauchy sequence in c_0 and hence in c_{00} (since the two spaces have the same norm) but it does not converge to an element of c_{00}.

In this case, c_0 is a completion of c_{00} with the norm $\|\cdot\|_\infty$.

With exactly the same argument, we see that c_{00} with the norm $\|\cdot\|_p$ is not complete, and that, in this case, l^p is a completion of c_{00} with the norm $\|\cdot\|_p$.

1.10 Function spaces.

1.10.1 Bounded continuous functions.

Definition. *We consider the space of all bounded functions $f : A \to F$, where A is any non-empty set:*

$$B(A) = \{f \mid f : A \to F \text{ is bounded in } A\}$$

If f, g are bounded in A and $\lambda \in F$, then $f + g, \lambda f$ are also bounded in A. So $B(A)$ is a linear space over F.

Definition. *We consider the function $\|\cdot\|_u : B(A) \to \mathbb{R}$ defined for every $f \in B(A)$ by*

$$\|f\|_u = \sup\{|f(a)| \mid a \in A\}$$

It is easy to see that $\|\cdot\|_u$ is a norm on $B(A)$; it is called **uniform norm** *on $B(A)$.*
If $f_n \to f$ in $B(A)$, we say that (f_n) **converges** *to f* **uniformly** *in A.*

Theorem 1.4. *$B(A)$ with the uniform norm is a Banach space.*

Proof. Take (f_n) in $B(A)$ so that $\|f_n - f_m\|_u \to 0$ when $n, m \to +\infty$. Then for every $a \in A$ we have

$$|f_n(a) - f_m(a)| \le \|f_n - f_m\|_u \to 0$$

when $n, m \to +\infty$, and so the sequence $(f_n(a))$ converges to some element of F. We consider $f : A \to F$ defined for every $a \in A$ by

$$f(a) = \lim_{n \to +\infty} f_n(a).$$

We consider n_0 so that $\|f_n - f_m\|_u < 1$ for every $n, m \ge n_0$. Then for every $a \in A$ and every $n \ge n_0$ we have

$$|f_n(a)| \le \|f_n\|_u \le \|f_n - f_{n_0}\|_u + \|f_{n_0}\|_u < 1 + \|f_{n_0}\|_u.$$

Taking the limit when $n \to +\infty$, we find

$$|f(a)| \le 1 + \|f_{n_0}\|_u$$

for every $a \in A$ and so $f \in B(A)$.
Now we take any $\epsilon > 0$ and then there is n_0 so that $\|f_n - f_m\|_u < \epsilon$ for every $n, m \ge n_0$. Then for every $n, m \ge n_0$ and every $a \in A$ we have

$$|f_n(a) - f_m(a)| \le \|f_n - f_m\|_u < \epsilon.$$

Taking the limit when $m \to +\infty$, we get

$$|f_n(a) - f(a)| \le \epsilon$$

for every $n \ge n_0$ and every $a \in A$. So $\|f_n - f\|_u \le \epsilon$ for every $n \ge n_0$, i.e. $f_n \to f$ in $B(A)$. $\square$

Definition. *We consider the spaces of all continuous and of all bounded and continuous functions $f : A \to F$, where A is any non-empty subset of a metric space or, more generally, of a topological space:*

$$C(A) = \{f \mid f : A \to F \text{ is continuous in } A\},$$
$$BC(A) = B(A) \cap C(A) = \{f \mid f : A \to F \text{ is bounded and continuous in } A\}.$$

It is clear that both spaces are linear spaces and that $BC(A)$ is a linear subspace of $B(A)$. Therefore, we may consider the restriction on $BC(A)$ of the uniform norm $\|\cdot\|_u$ on $B(A)$, and then $BC(A)$ becomes a subspace of $B(A)$.

Theorem 1.5. *Let A be a topological space. Then $BC(A)$ with the uniform norm is a Banach space.*

Proof. It is enough to prove that $BC(A)$ is a closed subspace of $B(A)$.
We take any sequence (f_n) in $BC(A)$ so that $f_n \to f$ for some $f \in B(A)$. We take any $a \in A$ and any $\epsilon > 0$. Then there is n_0 so that

$$|f_n(b) - f(b)| \le \|f_n - f\|_u \le \frac{\epsilon}{3}$$

for every $n \ge n_0$ and every $b \in A$. Since f_{n_0} is continuous at a, there is some open $U \subseteq A$ so that $a \in U$ and

$$|f_{n_0}(b) - f_{n_0}(a)| \le \frac{\epsilon}{3}$$

for every $b \in U$. Then

$$|f(b) - f(a)| \le |f(b) - f_{n_0}(b)| + |f_{n_0}(b) - f_{n_0}(a)| + |f_{n_0}(a) - f(a)| \le \frac{\epsilon}{3} + \frac{\epsilon}{3} + \frac{\epsilon}{3} = \epsilon$$

for every $b \in U$. So f is continuous at any $a \in A$. $\square$

1.10.2 Measurable functions.

Now we consider any measurable space (Ω, Σ), i.e. any non-empty set Ω and a σ-algebra Σ of subsets of Ω. We also consider the set of all functions $f : \Omega \to F$ which are measurable with respect to Σ:

$$\mathcal{M}(\Omega) = \mathcal{M}(\Omega, \Sigma) = \{f \mid f : \Omega \to F \text{ is measurable with respect to } \Sigma\}.$$

Since the sum of measurable functions and the product of a number and a measurable function are measurable functions we see that $\mathcal{M}(\Omega)$ is a linear space over F.

Now we also consider a measure μ on Σ, i.e. a measure space (Ω, Σ, μ). As in the basic theory of Measure and Integration, we consider equal every two functions in $\mathcal{M}(\Omega)$ which differ only in a set of μ-measure equal to 0.

Definition. *We define the function spaces*

$$L^\infty(\Omega) = L^\infty(\Omega, \Sigma, \mu) = \{f \in \mathcal{M}(\Omega) \mid f \text{ is essentially bounded in } \Omega\},$$
$$L^p(\Omega) = L^p(\Omega, \Sigma, \mu) = \Big\{f \in \mathcal{M}(\Omega) \,\Big|\, \int_\Omega |f|^p \, d\mu < +\infty\Big\}, \quad 1 \le p < +\infty.$$

It is easy to see that the sum of essentially bounded functions and the product of a number and an essentially bounded function are essentially bounded functions. Hence $L^\infty(\Omega)$ is a linear space over F.

Regarding the space $L^p(\Omega)$ with $1 \le p < +\infty$, we have that, if $f \in L^p(\Omega)$, then

$$\int_\Omega |\lambda f|^p \, d\mu = |\lambda|^p \int_\Omega |f|^p \, d\mu < +\infty$$

and hence $\lambda f \in L^p(\Omega)$. Also, if $f, g \in L^p(\Omega)$, then, as we saw in the proof of Minkowski's inequality for integrals,

$$\int_\Omega |f + g|^p \, d\mu \le 2^{p-1} \int_\Omega |f|^p \, d\mu + 2^{p-1} \int_\Omega |g|^p \, d\mu < +\infty,$$

and so $f + g \in L^p(\Omega)$.

Therefore, the space $L^p(\Omega)$ is a linear space over F.

Definition. *If* $1 \leq p \leq +\infty$, *we consider the function* $\|\cdot\|_p : L^p(\Omega) \to \mathbb{R}$ *defined for every* $f \in L^p(\Omega)$ *by*

$$\|f\|_p = \begin{cases} \left(\int_\Omega |f|^p \, d\mu\right)^{1/p}, & 1 \leq p < +\infty, \\ \text{ess-sup}_\Omega |f|, & p = +\infty. \end{cases}$$

Minkowski's inequality for integrals shows that $\|\cdot\|_p$ is a norm on $L^p(\Omega)$; it is the **p-norm** of $L^p(\Omega)$.

Theorem 1.6. $L^p(\Omega)$ *with the* p*-norm is a Banach space.*

Proof. We first consider the case $1 \leq p < +\infty$.
We take a sequence (f_n) in $L^p(\Omega)$ so that $\|f_n - f_m\|_p \to 0$ when $n, m \to +\infty$.
For every $k \in \mathbb{N}$ there is n_k so that $\|f_n - f_m\|_p < \frac{1}{2^k}$ for every $n, m \geq n_k$. We may assume that (n_k) is strictly increasing and so we have $\|f_{n_{k+1}} - f_{n_k}\| < \frac{1}{2^k}$ for every k.
We consider the function

$$s_k = |f_{n_1}| + |f_{n_2} - f_{n_1}| + \cdots + |f_{n_k} - f_{n_{k-1}}| \in L^p(\Omega),$$

and we have that

$$\begin{aligned}\left(\int_\Omega s_k^p \, d\mu\right)^{1/p} &= \|s_k\|_p \leq \|f_{n_1}\|_p + \|f_{n_2} - f_{n_1}\|_p + \cdots + \|f_{n_k} - f_{n_{k-1}}\|_p \\ &\leq \|f_{n_1}\|_p + \frac{1}{2} + \cdots + \frac{1}{2^{k-1}} \leq \|f_{n_1}\|_p + 1.\end{aligned}$$

Since (s_k) is an increasing sequence of non-negative functions, the monotone convergence theorem implies that the function $S = \lim_{k\to+\infty} s_k : \Omega \to [0, +\infty]$ satisfies $\int_\Omega S^p \, d\mu < +\infty$. Hence, $S(x) < +\infty$ for μ-a.e. $x \in \Omega$ and so the series $f_{n_1}(x) + \sum_{k=2}^{+\infty}(f_{n_k}(x) - f_{n_{k-1}}(x))$ converges absolutely for μ-a.e. $x \in \Omega$. Therefore the limit

$$\lim_{k\to+\infty} \big(f_{n_1}(x) + (f_{n_2}(x) - f_{n_1}(x)) + \cdots + (f_{n_k}(x) - f_{n_{k-1}}(x))\big) = \lim_{k\to+\infty} f_{n_k}(x)$$

exists in F for μ-a.e. $x \in \Omega$.
Now we consider the function $f : \Omega \to F$ defined for μ-a.e. $x \in \Omega$ by

$$f(x) = \lim_{k\to+\infty} f_{n_k}(x).$$

We have

$$|f_{n_k}(x)|^p = \big|f_{n_1}(x) + (f_{n_2}(x) - f_{n_1}(x)) + \cdots + (f_{n_k}(x) - f_{n_{k-1}}(x))\big|^p \leq s_k(x)^p \leq S^p(x)$$

for μ-a.e. $x \in \Omega$, and hence $|f|^p \leq S^p$ μ-a.e. in Ω. So

$$\int_\Omega |f|^p \, d\mu \leq \int_\Omega S^p \, d\mu < +\infty,$$

and thus $f \in L^p(\Omega)$. Moreover, the dominated convergence theorem implies that

$$\|f_{n_k} - f\|_p = \left(\int_\Omega |f_{n_k} - f|^p \, d\mu\right)^{1/p} \to 0$$

when $k \to +\infty$. Finally, since (f_n) is a Cauchy sequence,

$$\|f_k - f\|_p \leq \|f_k - f_{n_k}\|_p + \|f_{n_k} - f\|_p \to 0$$

when $k \to +\infty$.
Now let $p = +\infty$ and take a Cauchy sequence (f_n) in $L^\infty(\Omega)$.
Considering the union of countably many appropriate sets of μ-measure equal to 0, we see that

$$|f_n(x) - f_m(x)| \leq \|f_n - f_m\|_\infty$$

for μ-a.e. $x \in \Omega$ and every n, m. Therefore, $\lim_{n\to+\infty} f_n(x)$ exists in F for μ-a.e. $x \in \Omega$. Now we consider the function f defined for μ-a.e. $x \in \Omega$ by

$$f(x) = \lim_{n\to+\infty} f_n(x).$$

For any $\epsilon > 0$ there is n_0 so that

$$|f_n(x) - f_m(x)| \leq \|f_n - f_m\|_\infty < \epsilon$$

for μ-a.e. $x \in \Omega$ and every $n, m \geq n_0$. Taking the limit when $m \to +\infty$, we find that

$$|f_n(x) - f(x)| \leq \epsilon$$

for μ-a.e. $x \in \Omega$ and every $n \geq n_0$. Therefore $\|f_n - f\|_\infty \leq \epsilon$ for every $n \geq n_0$, and so $f_n \to f$ in $L^\infty(\Omega)$. $\square$

A special case of the above is when Ω is a topological space. In this special case we may consider Σ to be the smallest σ-algebra which contains all open subsets of Ω. This σ-algebra is denoted $\mathcal{B}(\Omega)$ and it is called **σ-algebra of the Borel subsets** of Ω. Since $\mathcal{B}(\Omega)$ is a σ-algebra which contains all open subsets of Ω, it also contains all closed subsets of Ω, as well as all countable intersections of open subsets of Ω and all countable unions of closed subsets of Ω. The elements of $\mathcal{B}(\Omega)$ are called **Borel subsets** of Ω.

A measure μ on $\mathcal{B}(\Omega)$ is called **Borel measure** in Ω. If μ also satisfies $\mu(K) < +\infty$ for every compact $K \subseteq \Omega$, then it is called **locally finite** Borel measure.

Every continuous function $f : \Omega \to F$ is measurable with respect to $\mathcal{B}(\Omega)$. So if μ is a Borel measure in Ω, we may consider the subset

$$C(\Omega) \cap L^p(\Omega) = \Big\{ f \,\Big|\, f : \Omega \to F \text{ continuous in } \Omega \text{ with } \int_\Omega |f|^p \, d\mu < +\infty \Big\}$$

of $L^p(\Omega)$, which consists of all functions which are continuous and p-integrable in Ω. Then $C(\Omega) \cap L^p(\Omega)$ is a subspace of $L^p(\Omega)$, and it is well known that $C(\Omega) \cap L^p(\Omega)$ is dense in $L^p(\Omega)$. In other words, $L^p(\Omega)$ is a completion of $C(\Omega) \cap L^p(\Omega)$.

1.10.3 Differentiable functions.

Let U be an open subset of $\mathbb{R}^d$ and let $f : U \to F$. We take any multi-index $\alpha = (\alpha_1, \ldots, \alpha_d) \in \mathbb{N}_0^d$ with *length* $|\alpha| = \alpha_1 + \cdots + \alpha_d$, and we consider the derivative of order $|\alpha|$ at any $x \in U$:

$$D^\alpha f(x) = \frac{\partial^{|\alpha|} f}{\partial x^\alpha}(x) = \frac{\partial^{|\alpha|} f}{\partial x_1^{\alpha_1} \cdots \partial x_d^{\alpha_d}}(x).$$

Definition. *For every $k \in \mathbb{N} \cup \{+\infty\}$ we define the space*

$$C^k(U) = \{ f \mid f : U \to F \text{ has continuous derivatives of order } \leq k \text{ in } U \}.$$

We also define $C^0(U) = C(U)$.

It is clear that all $C^k(U)$ are linear spaces over F, and that

$$C^\infty(U) \subseteq C^k(U) \subseteq C^l(U) \subseteq C(U)$$

for every $k, l \in \mathbb{N}_0$ with $k \geq l$.

In the following we consider m to be the Lebesgue measure in $U \subseteq \mathbb{R}^d$ and when we write $L^p(U)$ we consider U with the Lebesgue measure m.

Definition. *If $k \in \mathbb{N}_0$ and $1 \leq p < +\infty$, we define the space*

$$C^{k,p}(U) = \{f \in C^k(U) \mid D^\alpha f \in L^p(U) \text{ when } |\alpha| \leq k\}.$$

If $k \in \mathbb{N}_0$ and $p = +\infty$, we define the space

$$C^{k,\infty}(U) = \{f \in C^k(U) \mid D^\alpha f \in BC(U) \text{ when } |\alpha| \leq k\}.$$

Definition. *Let $k \in \mathbb{N}_0$ and $1 \leq p \leq +\infty$. We consider the function $\|\cdot\|_{k,p} : C^{k,p}(U) \to \mathbb{R}$ defined for every $f \in C^{k,p}(U)$ by*

$$\|f\|_{k,p} = \begin{cases} \big(\sum_{|\alpha|\leq k} \int_U |D^\alpha f|^p \, dm\big)^{1/p}, & 1 \leq p < +\infty, \\ \sum_{|\alpha|\leq k} \|D^\alpha f\|_u, & p = +\infty. \end{cases}$$

Of course, when we write $\|\cdot\|_u$ we mean the uniform norm on $BC(U)$.

Proposition 1.18. *The function $\|\cdot\|_{k,p} : C^{k,p}(U) \to \mathbb{R}$ is a norm on $C^{k,p}(U)$.*

Proof. Let $1 \leq p < +\infty$. Then, using Minkowski's inequalities for sums and integrals, we get

$$\begin{aligned} \|f + g\|_{k,p} &= \Big(\sum_{|\alpha|\leq k} \int_U |D^\alpha f + D^\alpha g|^p \, dm \Big)^{1/p} \\ &\leq \Big(\sum_{|\alpha|\leq k} \Big[\Big(\int_U |D^\alpha f|^p \, dm \Big)^{1/p} + \Big(\int_U |D^\alpha g|^p \, dm \Big)^{1/p} \Big]^p \Big)^{1/p} \\ &\leq \Big(\sum_{|\alpha|\leq k} \int_U |D^\alpha f|^p \, dm \Big)^{1/p} + \Big(\sum_{|\alpha|\leq k} \int_U |D^\alpha g|^p \, dm \Big)^{1/p} \\ &= \|f\|_{k,p} + \|g\|_{k,p}. \end{aligned}$$

All other properties of the norm, as well as the case $p = +\infty$, are straightforward. □

Theorem 1.7. *$C^{k,\infty}(U)$ is a Banach space.*

Proof. If $k = 0$, then $C^{0,\infty}(U) = BC(U)$ and we already know that this is a Banach space.
So we take $k \geq 1$ and we consider a Cauchy sequence (f_n) in $C^{k,\infty}(U)$, i.e.

$$\sum_{|\alpha|\leq k} \|D^\alpha f_n - D^\alpha f_m\|_u \to 0$$

when $n, m \to +\infty$. Then for every α with $|\alpha| \leq k$ we have that $\|D^\alpha f_n - D^\alpha f_m\|_u \to 0$ when $n, m \to +\infty$, and so $(D^\alpha f_n)$ is a Cauchy sequence in $BC(U)$. Therefore, there is $f_\alpha \in BC(U)$ so that $D^\alpha f_n \to f_\alpha$ uniformly in U.
We take any $x = (x_1, \ldots, x_j, \ldots, x_d) \in U$ and a small $h \in \mathbb{R}$ so that the linear segment with endpoints x and $x + he_j = (x_1, \ldots, x_j + h, \ldots, x_d)$ is contained in U. Then for every n we have

$$f_n(x_1, \ldots, x_j + h, \ldots, x_d) - f_n(x_1, \ldots, x_j, \ldots, x_d) = \int_0^h \frac{\partial f_n}{\partial x_j}(x_1, \ldots, x_j + t, \ldots, x_d) \, dt.$$

Because of uniform convegrence, when $n \to +\infty$ the left side of this equality converges to

$$f_{(0,\dots,0)}(x_1,\dots,x_j+h,\dots,x_d) - f_{(0,\dots,0)}(x_1,\dots,x_j,\dots,x_d)$$

and the right side converges to

$$\int_0^h f_{(0,\dots,1,\dots,0)}(x_1,\dots,x_j+t,\dots,x_d)\,dt,$$

where the 1 in the last multi-index appears at the j-th place. Thus

$$\begin{aligned} f_{(0,\dots,0)}(x_1,\dots,x_j+h,\dots,x_d) - & f_{(0,\dots,0)}(x_1,\dots,x_j,\dots,x_d) \\ &= \int_0^h f_{(0,\dots,1,\dots,0)}(x_1,\dots,x_j+t,\dots,x_d)\,dt. \end{aligned}$$

Since the integrated function is continuous in t, we may differentiate the integral with respect to h at $h=0$ and we get

$$\frac{\partial f_{(0,\dots,0)}}{\partial x_j}(x) = f_{(0,\dots,1,\dots,0)}(x).$$

Therefore, if we define $f = f_{(0,\dots,0)}$, then $f_{(0,\dots,1,\dots,0)} = \frac{\partial f}{\partial x_j}$ in U.
In the same way, we can show inductively that for every α with $|\alpha| \le k$ we have $f_\alpha = D^\alpha f$ in U.
Thus for every α with $|\alpha| \le k$ we have $\|D^\alpha f_n - D^\alpha f\|_u \to 0$ when $n \to +\infty$, and so

$$\|f_n - f\|_{k,\infty} = \sum_{|\alpha|\le k} \|D^\alpha f_n - D^\alpha f\|_u \to 0$$

when $n \to +\infty$. $\square$

If $1 \le p < +\infty$, then the normed spaces $C^{k,p}(U)$ are *not* complete. We shall now say a few things about the completion of each of these spaces.

Definition. *Let X be a topological space and let $f : X \to F$ be continuous in X. The set*

$$\mathrm{supp}(f) = \mathrm{cl}(\{x \in X \mid f(x) \ne 0\})$$

is called **support** *of f.*
If $\mathrm{supp}(f)$ *is compact, then we say that f* **has compact support**.

It is easy to see that $X \setminus \mathrm{supp}(f)$ is the largest open subset of X in which f is identically 0.

Definition. *Let $U \subseteq \mathbb{R}^d$ be open. We define the space*

$$C_c^\infty(U) = \{f \in C^\infty(U) \mid f \textit{ has compact support } \subseteq U\}.$$

Lemma 1.1. *Let X be a topological space, $\lambda \in F$ and $f, g : X \to F$ be continuous in X. Then* $\mathrm{supp}(\lambda f) \subseteq \mathrm{supp}(f)$ *and* $\mathrm{supp}(f+g) \subseteq \mathrm{supp}(f) \cup \mathrm{supp}(g)$.

Proof. Trivial. $\square$

Proposition 1.19. *$C_c^\infty(U)$ is a linear subspace of $C^\infty(U)$.*

Proof. From the last lemma it is obvious that, if $f, g \in C_c^\infty(U)$, then $\mathrm{supp}(\lambda f)$, being a closed subset of the compact $\mathrm{supp}(f)$, is a compact subset of U, and also $\mathrm{supp}(f+g)$, being a closed subset of the compact $\mathrm{supp}(f) \cup \mathrm{supp}(g)$, is a compact subset of U. Therefore $\lambda f, f+g \in C_c^\infty(U)$. $\square$

If $f : U \to F$ has continuous derivatives $D^\alpha f$ in U for every α with $|\alpha| \le k$, then integration by parts implies that

$$\int_U D^\alpha f\,\phi\,dm = (-1)^{|\alpha|}\int_U f\,D^\alpha\phi\,dm$$

for every α with $|\alpha| \le k$, and every $\phi \in C_c^\infty(U)$.

Definition. *Let $1 \le p < +\infty$ and $f \in L^p(U)$. We say that the function $f_\alpha \in L^p(U)$ is a* **weak α-derivative** *of f in U if*

$$\int_U f_\alpha\,\phi\,dm = (-1)^{|\alpha|}\int_U f\,D^\alpha\phi\,dm$$

for every $\phi \in C_c^\infty(U)$.

If a weak α-derivative of f exists, then it is unique. Indeed, if both $f'_\alpha, f''_\alpha \in L^p(U)$ are weak α-derivatives of f, then

$$\int_U f'_\alpha\,\phi\,dm = (-1)^{|\alpha|}\int_U f\,D^\alpha\phi\,dm = \int_U f''_\alpha\,\phi\,dm$$

and hence

$$\int_U (f'_\alpha - f''_\alpha)\,\phi\,dm = 0$$

for every $\phi \in C_c^\infty(U)$. This implies that $f'_\alpha = f''_\alpha$ m-a.e. in U[1]. Therefore, if f has derivative $D^\alpha f$ in the usual sense and $D^\alpha f \in L^p(U)$, then a function $f_\alpha \in L^p(U)$ is a weak α-derivative of f if and only if $f_\alpha = D^\alpha f$ m-a.e. in U.

The weak derivatives f_α are substitutes for the usual derivatives $D^\alpha f$ whenever the function f is not differentiable in the usual sense. We agree to denote $D^\alpha f$ the weak α-derivative f_α, even when $D^\alpha f$ does not exist in the usual sense.

Definition. *Let $k \in \mathbb{N}$ and $1 \le p < +\infty$. We call* **Sobolev space** *and denote $W^{k,p}(U)$ the set of all functions $f \in L^p(U)$ which have weak α-derivatives $D^\alpha f$ in $L^p(U)$ for every α with $|\alpha| \le k$. We consider the function $\|\cdot\|_{k,p} : W^{k,p}(U) \to \mathbb{R}$ defined for every $f \in W^{k,p}(U)$ by*

$$\|f\|_{k,p} = \Big(\sum_{|\alpha|\le k}\int_U |D^\alpha f|^p\,dm\Big)^{1/p}.$$

It is clear that $\|\cdot\|_{k,p}$ is a norm on $W^{k,p}(U)$ and that $C^{k,p}(U)$ is a subspace of $W^{k,p}(U)$.

Proposition 1.20. *$W^{k,p}(U)$ with the norm $\|\cdot\|_{k,p}$ is a Banach space.*

Proof. Let (f_n) be a Cauchy sequence in $W^{k,p}(U)$, i.e. $\|f_n - f_m\|_{k,p} \to 0$ when $n, m \to +\infty$. Then for every α with $|\alpha| \le k$ we have

$$\|D^\alpha f_n - D^\alpha f_m\|_p \le \|f_n - f_m\|_{k,p} \to 0$$

when $n, m \to +\infty$. So, for every α with $|\alpha| \le k$ we have that $(D^\alpha f_n)$ is a Cauchy sequence in $L^p(U)$ and hence there is some $f_\alpha \in L^p(U)$ so that $D^\alpha f_n \to f_\alpha$ in $L^p(U)$.
In particular, when $\alpha = (0, \ldots, 0)$ we have a function $f = f_{(0,\ldots,0)}$ so that $f_n \to f$ in $L^p(U)$.
Now we consider any $\phi \in C_c^\infty(U)$ and then we have the equality

$$\int_U D^\alpha f_n\,\phi\,dm = (-1)^{|\alpha|}\int_U f_n\,D^\alpha\phi\,dm$$

for every n. Using Hölder's inequality, with the dual exponents p and q, we have that

$$\Big|\int_U D^\alpha f_n\,\phi\,dm - \int_U f_\alpha\,\phi\,dm\Big| = \Big|\int_U (D^\alpha f_n - f_\alpha)\,\phi\,dm\Big| \le \|D^\alpha f_n - f_\alpha\|_p\|\phi\|_q \to 0$$

[1]This conclusion together with the fact that the space $C_c^\infty(U)$ is non-empty are taken for granted.

and

$$\Big|\int_U f_n\, D^\alpha \phi\, dm - \int_U f\, D^\alpha \phi\, dm\Big| = \Big|\int_U (f_n - f)\, D^\alpha \phi\, dm\Big| \le \|f_n - f\|_p \|D^\alpha \phi\|_q \to 0$$

when $n \to +\infty$. Thus,

$$\int_U f_\alpha\, \phi\, dm = (-1)^{|\alpha|} \int_U f\, D^\alpha \phi\, dm.$$

Since this holds for every $\phi \in C_c^\infty(U)$, we have that f_α is a weak α-derivative of f in $L^p(U)$, i.e. $f_\alpha = D^\alpha f$.

So the function $f \in L^p(U)$ has weak derivatives $D^\alpha f$ in $L^p(U)$ for every α with $|\alpha| \le k$. In other words, $f \in W^{k,p}(U)$.

Finally, since $D^\alpha f_n \to f_\alpha = D^\alpha f$ in $L^p(U)$ for every α with $|\alpha| \le k$, we find that

$$\|f_n - f\|_{k,p} = \Big(\sum_{|\alpha|\le k} \int_U |D^\alpha f_n - D^\alpha f|^p\, dm\Big)^{1/p} = \Big(\sum_{|\alpha|\le k} \|D^\alpha f_n - D^\alpha f\|_p^p\Big)^{1/p} \to 0$$

when $n \to +\infty$. Thus $f_n \to f$ in $W^{k,p}(U)$. □

Proposition 1.21. [2] *$C^{k,p}(U)$ is a dense subspace of $W^{k,p}(U)$.*

Therefore the Sobolev space $W^{k,p}(U)$ is a completion of $C^{k,p}(U)$.

1.11 Measure spaces.

Let Ω be a non-empty set and Σ be a σ-algebra of subsets of Ω. We recall from the basic course on Measure and Integration that a function $\mu : \Sigma \to F$ is called **real** (if $F = \mathbb{R}$) or **complex** (if $F = \mathbb{C}$) **measure** on Σ, if

$$\mu(\emptyset) = 0, \quad \mu\Big(\bigcup_{j=1}^{+\infty} A_j\Big) = \sum_{j=1}^{+\infty} \mu(A_j)$$

for every pairwise disjoint $A_j \in \Sigma$, $j \in \mathbb{N}$. (In particular, the last series converges.) Note that a real or complex measure does not take the values $\pm\infty$ and ∞.

We define

$$\mathcal{A}(\Omega) = \mathcal{A}(\Omega, \Sigma) = \{\mu \,|\, \mu \text{ is a real or complex measure on } \Sigma\}.$$

It is easy to see that, if $\mu, \nu \in \mathcal{A}(\Omega)$ and $\lambda \in F$, then $\mu + \nu, \lambda\mu \in \mathcal{A}(\Omega)$. So $\mathcal{A}(\Omega)$ is a linear space over F.

If a real or complex measure μ satisfies $\mu(A) \ge 0$ for every $A \in \Sigma$, then we say that μ is a **non-negative real measure**, and then μ is a special case of a *measure*: a measure on Σ is a function $\mu : \Sigma \to [0, +\infty]$ which satisfies $\mu(\emptyset) = 0$ and $\mu(\bigcup_{j=1}^{+\infty} A_j) = \sum_{j=1}^{+\infty} \mu(A_j)$ for every pairwise disjoint $A_j \in \Sigma$, $j \in \mathbb{N}$. Therefore, a measure μ is a non-negative real measure if and only if $\mu(\Omega) < +\infty$.

Definition. *Let (Ω, Σ) be a measurable space, and μ be a real or complex measure on Σ. For every $A \in \Sigma$ we define*

$$|\mu|(A) = \sup\Big\{\sum_{m=1}^{n} |\mu(A_m)| \,\Big|\, n \in \mathbb{N}, A_1, \dots, A_n \in \Sigma \textit{ are pairwise disjoint}, \bigcup_{m=1}^{n} A_m \subseteq A\Big\}.$$

Then $|\mu|(A)$ is called **total variation** *of μ in A.*

[2] We shall not prove (now) proposition 1.21. For the proof see the book "*Sobolev Spaces*" by Adams and Fournier.

Lemma 1.2. *Let $K \subseteq \mathbb{C}$ be finite. Then there is $M \subseteq K$, so that $|\sum_{\lambda \in M} \lambda| \geq \frac{1}{6} \sum_{\lambda \in K} |\lambda|$.*

Proof. $\mathbb{C}$ is the union of

$$Q_1 = \{\lambda \mid \operatorname{Re}\lambda \geq |\operatorname{Im}\lambda|\}, \quad Q_2 = \{\lambda \mid \operatorname{Re}\lambda \leq -|\operatorname{Im}\lambda|\},$$

$$Q_3 = \{\lambda \mid \operatorname{Im}\lambda \geq |\operatorname{Re}\lambda|\}, \quad Q_4 = \{\lambda \mid \operatorname{Im}\lambda \leq -|\operatorname{Re}\lambda|\}.$$

If $\lambda_1, \dots, \lambda_n \in Q_1$, then

$$|\lambda_1 + \dots + \lambda_n| \geq \operatorname{Re}(\lambda_1 + \dots + \lambda_n) = \operatorname{Re}\lambda_1 + \dots + \operatorname{Re}\lambda_n \geq \frac{1}{\sqrt{2}}(|\lambda_1| + \dots + |\lambda_n|).$$

The same is true if $\lambda_1, \dots, \lambda_n$ all belong to one of Q_2, Q_3, Q_4.
Now, we split K in four pairwise disjoint subsets K_1, K_2, K_3, K_4, so that each contains elements of K in Q_1, Q_2, Q_3, Q_4, respectively. Then at least one of them, say M, satisfies

$$\sum_{\lambda \in M} |\lambda| \geq \frac{1}{4} \sum_{\lambda \in K} |\lambda|.$$

and so

$$\Big|\sum_{\lambda \in M} \lambda\Big| \geq \frac{1}{\sqrt{2}} \sum_{\lambda \in M} |\lambda| \geq \frac{1}{4\sqrt{2}} \sum_{\lambda \in K} |\lambda| \geq \frac{1}{6} \sum_{\lambda \in K} |\lambda|.$$

□

Theorem 1.8. *If μ is a complex measure on Σ, then $|\mu|$ is a non-negative real measure on Σ. In particular, $|\mu|(\Omega) < +\infty$.*

Proof. It is obvious that $|\mu|(A) \geq 0$ for every $A \in \Sigma$, and that $|\mu|(\emptyset) = 0$.
Now let $A^1, A^2, \ldots \in \Sigma$ be pairwise disjoint, and $A = \bigcup_{j=1}^{+\infty} A^j$.
We take pairwise disjoint $A_1, \dots, A_n \in \Sigma$ with $\bigcup_{m=1}^n A_m \subseteq A$. We consider the $A_m^j = A^j \cap A_m$ and then

$$A_m = \bigcup_{j=1}^{+\infty} A_m^j, \quad \bigcup_{m=1}^{n} A_m^j \subseteq A^j.$$

Therefore,

$$\sum_{m=1}^n |\mu(A_m)| = \sum_{m=1}^n \Big|\sum_{j=1}^{+\infty} \mu(A_m^j)\Big| \leq \sum_{m=1}^n \sum_{j=1}^{+\infty} |\mu(A_m^j)| = \sum_{j=1}^{+\infty} \sum_{m=1}^n |\mu(A_m^j)| \leq \sum_{j=1}^{+\infty} |\mu|(A^j).$$

Taking the supremum of the left side, we get $|\mu|(A) \leq \sum_{j=1}^{+\infty} |\mu|(A^j)$.
We take any J, and for every $j = 1, \dots, J$ we take any $\lambda_j < |\mu|(A^j)$. Then there are pairwise disjoint $A_1^j, \dots, A_{n_j}^j \in \Sigma$ so that

$$\bigcup_{m=1}^{n_j} A_m^j \subseteq A^j, \quad \lambda_j < \sum_{m=1}^{n_j} |\mu(A_m^j)|.$$

Then $A_1^1, \dots, A_{n_J}^J$ are pairwise disjoint and their union is contained in A. Hence

$$\sum_{j=1}^J \lambda_j < \sum_{j=1}^J \sum_{m=1}^{n_j} |\mu(A_m^j)| \leq |\mu|(A).$$

Taking first the supremum over the $\lambda_1, \dots, \lambda_J$ and then the limit when $J \to +\infty$, we get that $\sum_{j=1}^{+\infty} |\mu|(A^j) \leq |\mu|(A)$.

We conclude that $\sum_{j=1}^{+\infty} |\mu|(A^j) = |\mu|(A)$, i.e. that $|\mu|$ is a measure, and we still have to prove that $|\mu|(\Omega) < +\infty$.
We assume that $|\mu|(\Omega) = +\infty$, and we claim that there are $B_1, B_2, \ldots \in \Sigma$ so that

$$B_1 \supseteq B_2 \supseteq B_3 \supseteq \ldots, \quad |\mu|(B_k) = +\infty, \quad |\mu(B_k)| \geq k - 1$$

for every k. We take $B_1 = \Omega$ and we assume that we have proven the existence of the first $B_1, \ldots, B_k$. Since $|\mu|(B_k) = +\infty$, there are pairwise disjoint $A_1, \ldots, A_n \in \Sigma$ so that

$$\bigcup_{m=1}^{n} A_m \subseteq B_k, \quad \sum_{m=1}^{n} |\mu(A_m)| \geq 6(|\mu(B_k)| + k).$$

According to lemma 1.2, there are some of the $A_1, \ldots, A_n$, which we may assume that they are the $A_1, \ldots, A_l$, so that

$$\Big|\sum_{m=1}^{l} \mu(A_m)\Big| \geq \frac{1}{6} \sum_{m=1}^{n} |\mu(A_m)| \geq |\mu(B_k)| + k.$$

We set $S = \bigcup_{m=1}^{l} A_m \subseteq B_k$, and then

$$|\mu(S)| \geq |\mu(B_k)| + k.$$

Since $|\mu|(S) + |\mu|(B_k \setminus S) = |\mu|(B_k) = +\infty$, we have that either $|\mu|(S) = +\infty$ or $|\mu|(B_k \setminus S) = +\infty$. In the first case we set $B_{k+1} = S \subseteq B_k$, and then $|\mu(B_{k+1})| \geq |\mu(B_k)| + k \geq k$. In the second case we set $B_{k+1} = B_k \setminus S \subseteq B_k$, and then $|\mu(B_{k+1})| \geq |\mu(S)| - |\mu(B_k)| \geq k$.
In any case we have proven the existence of an appropriate B_{k+1} and hence the claim.
Now we consider the pairwise disjoint $A_1 = B_1 \setminus B_2, A_2 = B_2 \setminus B_3, \ldots$ and the $B_\infty = \bigcap_{k=1}^{+\infty} B_k$. Then

$$\begin{aligned}
\mu(B_1) - \mu(B_\infty) = \mu(B_1 \setminus B_\infty) = \mu\Big(\bigcup_{m=1}^{+\infty} A_m\Big) &= \sum_{m=1}^{+\infty} \mu(A_m) \\
&= \lim_{k\to+\infty} \sum_{m=1}^{k-1} \mu(A_m) = \lim_{k\to+\infty} (\mu(B_1) - \mu(B_k)).
\end{aligned}$$

Therefore $\lim_{k\to+\infty} \mu(B_k) = \mu(B_\infty)$, i.e. $|\mu(B_\infty)| = +\infty$, and we arrive at a contradiction. □

Definition. *If μ is a complex measure on Σ, then the non-negative real measure $|\mu|$ on Σ is called* **absolute variation** *of μ and the number $|\mu|(\Omega)$ is called* **total variation** *of μ.*

Definition. *We consider the function $\|\cdot\| : \mathcal{A}(\Omega, \Sigma) \to \mathbb{R}$ defined for every $\mu \in \mathcal{A}(\Omega, \Sigma)$ by*

$$\|\mu\| = |\mu|(\Omega).$$

Proposition 1.22. *$\|\cdot\| : \mathcal{A}(\Omega, \Sigma) \to \mathbb{R}$ is a norm on $\mathcal{A}(\Omega, \Sigma)$.*

Proof. Let $\|\mu\| = 0$. Then for every $A \in \Sigma$ we have $|\mu(A)| \leq |\mu|(\Omega) = 0$, and hence $\mu(A) = 0$. So $\mu = 0$.
Let $\mu \in \mathcal{A}(\Omega, \Sigma)$ and $\lambda \in F$. We take pairwise disjoint $A_1, \ldots, A_n \in \Sigma$ and we have

$$\sum_{m=1}^{n} |(\lambda\mu)(A_m)| = |\lambda| \sum_{m=1}^{n} |\mu(A_m)|.$$

Taking the supremum of both sides, we find $\|\lambda\mu\| = |\lambda|\|\mu\|$.
Now let $\mu, \nu \in \mathcal{A}(\Omega, \Sigma)$. For every pairwise disjoint $A_1, \ldots, A_n \in \Sigma$ we have

$$\sum_{m=1}^{n} |(\mu + \nu)(A_m)| \leq \sum_{m=1}^{n} |\mu(A_m)| + \sum_{m=1}^{n} |\nu(A_m)| \leq \|\mu\| + \|\nu\|.$$

Taking the supremum of the left side, we get $\|\mu + \nu\| \leq \|\mu\| + \|\nu\|$. □

Theorem 1.9. *$\mathcal{A}(\Omega, \Sigma)$ with the norm $\|\cdot\|$ is a Banach space.*

Proof. Let (μ_n) be a Cauchy sequence in $\mathcal{A}(\Omega, \Sigma)$. Then for every $A \in \Sigma$ we have

$$|\mu_n(A) - \mu_m(A)| \le \|\mu_n - \mu_m\| \to 0$$

when $n, m \to +\infty$ and so the limit $\lim_{n\to+\infty} \mu_n(A)$ exists in F. Thus, we may consider the function $\mu : \Sigma \to F$ defined for every $A \in \Sigma$ by

$$\mu(A) = \lim_{n\to+\infty} \mu_n(A).$$

Clearly, $\mu(\emptyset) = \lim_{n\to+\infty} \mu_n(\emptyset) = 0$.
Now we take pairwise disjoint $A_1, A_2, \ldots \in \Sigma$ and $A = \bigcup_{j=1}^{+\infty} A_j$. For every $\epsilon > 0$ there is n_0 so that $\|\mu_n - \mu_m\| < \epsilon$ for every $n, m \ge n_0$. Since $\sum_{j=1}^{+\infty} \mu_{n_0}(A_j) = \mu_{n_0}(A)$, there is J_0 so that

$$\Big|\mu_{n_0}(A) - \sum_{j=1}^{J} \mu_{n_0}(A_j)\Big| < \epsilon$$

for every $J \ge J_0$. If $m \ge n_0$, then

$$\begin{aligned}\Big|(\mu_{n_0}(A) - \mu_m(A)) - \sum_{j=1}^{J}(\mu_{n_0}(A_j) - \mu_m(A_j))\Big| &= \Big|\sum_{j=J+1}^{+\infty}(\mu_{n_0}(A_j) - \mu_m(A_j))\Big| \\ &\le \sum_{j=J+1}^{+\infty} |\mu_{n_0}(A_j) - \mu_m(A_j)| \le \|\mu_{n_0} - \mu_m\| < \epsilon.\end{aligned}$$

Taking the limit when $m \to +\infty$ we find

$$\Big|(\mu_{n_0}(A) - \mu(A)) - \sum_{j=1}^{J}(\mu_{n_0}(A_j) - \mu(A_j))\Big| \le \epsilon$$

and hence

$$\Big|\mu(A) - \sum_{j=1}^{J} \mu(A_j)\Big| \le 2\epsilon$$

for every $J \ge J_0$. Thus $\sum_{j=1}^{+\infty} \mu(A_j) = \mu(A)$ and so $\mu \in \mathcal{A}(\Omega, \Sigma)$.
Finally, for any $\epsilon > 0$ we choose n_0 as before. We take pairwise disjoint $A_1, \ldots, A_k \in \Sigma$, and then for every $n, m \ge n_0$ we have

$$\sum_{j=1}^{k} |\mu_n(A_j) - \mu_m(A_j)| \le \|\mu_n - \mu_m\| < \epsilon.$$

We take the limit when $m \to +\infty$ and we get

$$\sum_{j=1}^{k} |\mu_n(A_j) - \mu(A_j)| \le \epsilon.$$

Considering the supremum of the left side, we get $\|\mu_n - \mu\| \le \epsilon$ for every $n \ge n_0$. Therefore, $\mu_n \to \mu$ in $\mathcal{A}(\Omega, \Sigma)$. □

Definition. *If μ is a real measure on Σ, then the non-negative real measures $\mu^+ = \frac{1}{2}(|\mu| + \mu)$ and $\mu^- = \frac{1}{2}(|\mu| - \mu)$ are called* **positive variation** *of μ and* **negative variation** *of μ.*

That these two real measures are non-negative follows from the definition of $|\mu|(A)$: we have $|\mu|(A) \geq |\mu(A)|$ for every $A \in \Sigma$. The two identities

$$\mu = \mu^+ - \mu^-, \quad |\mu| = \mu^+ + \mu^-$$

are clear.

It is easy to prove for any complex measure $\mu : \Sigma \to \mathbb{C}$ that the functions $\operatorname{Re}\mu, \operatorname{Im}\mu : \Sigma \to \mathbb{R}$ defined for every $A \in \Sigma$ by

$$\operatorname{Re}\mu(A) = \operatorname{Re}(\mu(A)), \quad \operatorname{Im}\mu(A) = \operatorname{Im}(\mu(A)),$$

are real measures on Σ. Also the function $\overline{\mu} : \Sigma \to \mathbb{C}$ defined for every $A \in \Sigma$ by

$$\overline{\mu}(A) = \overline{\mu(A)},$$

is a complex measure on Σ. Moreover,

$$\mu = \operatorname{Re}\mu + i\operatorname{Im}\mu, \quad \overline{\mu} = \operatorname{Re}\mu - i\operatorname{Im}\mu,$$

$$|\operatorname{Re}\mu| \leq |\mu|, \quad |\operatorname{Im}\mu| \leq |\mu|, \quad |\mu| \leq |\operatorname{Re}\mu| + |\operatorname{Im}\mu|,$$
$$|\mu_1 + \mu_2| \leq |\mu_1| + |\mu_2|, \quad |\overline{\mu}| = |\mu|, \quad |\lambda\mu| = |\lambda||\mu|.$$

A special case of the above is when Ω is a topological space, and Σ is the σ-algebra $\mathcal{B}(\Omega)$ of the Borel subsets of Ω. Then every real or complex measure on $\mathcal{B}(\Omega)$ is called **real** or **complex Borel measure** in Ω.

Definition. *A real or complex Borel measure μ on $\mathcal{B}(\Omega)$, i.e. an element of $\mathcal{A}(\Omega, \mathcal{B}(\Omega))$ is called* **regular***, if for every $A \in \mathcal{B}(\Omega)$ and every $\epsilon > 0$ there are $K, U \subseteq \Omega$ so that K is compact, U is open, and*

$$K \subseteq A \subseteq U, \quad |\mu|(U \setminus K) < \epsilon.$$

The set of all regular Borel measures on $\mathcal{B}(\Omega)$ is denoted

$$\mathcal{A}_r(\Omega, \mathcal{B}(\Omega)) = \{\mu \in \mathcal{A}(\Omega, \mathcal{B}(\Omega)) \mid \mu \textit{ is regular}\}.$$

Proposition 1.23. *$\mathcal{A}_r(\Omega, \mathcal{B}(\Omega))$ is a linear subspace of $\mathcal{A}(\Omega, \mathcal{B}(\Omega))$.*
If $\mu \in \mathcal{A}_r(\Omega, \mathcal{B}(\Omega))$, then $|\mu| \in \mathcal{A}_r(\Omega, \mathcal{B}(\Omega))$.

Proof. Trivial. □

So we may consider $\mathcal{A}_r(\Omega, \mathcal{B}(\Omega))$ to be a subspace of $\mathcal{A}(\Omega, \mathcal{B}(\Omega))$ with the total variation $\|\cdot\|$ as norm.

Proposition 1.24. *$\mathcal{A}_r(\Omega, \mathcal{B}(\Omega))$ is a closed subspace of $\mathcal{A}(\Omega, \mathcal{B}(\Omega))$ and so it is a Banach space.*

Proof. Let (μ_n) be a sequence in $\mathcal{A}_r(\Omega, \mathcal{B}(\Omega))$, and let $\mu_n \to \mu$ in $\mathcal{A}(\Omega, \mathcal{B}(\Omega))$.
We take any $A \in \mathcal{B}(\Omega)$ and any $\epsilon > 0$. Then there is n_0 so that $\|\mu_{n_0} - \mu\| < \epsilon$. Moreover, since μ_{n_0} is regular, there are $K, U \subseteq \Omega$ so that K is compact, U is open, and

$$K \subseteq A \subseteq U, \quad |\mu_{n_0}|(U \setminus K) < \epsilon.$$

Then

$$\begin{aligned}|\mu|(U \setminus K) &\leq |\mu - \mu_{n_0}|(U \setminus K) + |\mu_{n_0}|(U \setminus K) \leq |\mu - \mu_{n_0}|(\Omega) + |\mu_{n_0}|(U \setminus K) \\ &= \|\mu - \mu_{n_0}\| + |\mu_{n_0}|(U \setminus K) < 2\epsilon.\end{aligned}$$

Thus, $\mu \in \mathcal{A}_r(\Omega, \mathcal{B}(\Omega))$. □

1.12 Compact sets in infinite dimensional normed spaces.

Riesz's lemma. *Let X be a normed space, $Y \subsetneq X$ be a closed subspace of X, and $0 < t < 1$. Then there exists $x \in X$ so that $\|x\| = 1$ and $\inf_{y\in Y} \|x - y\| \geq t$.*

Proof. To begin with, we observe that if $\|x\| = 1$ then, since $0 \in Y$, we have

$$\inf_{y\in Y} \|x - y\| \leq \|x - 0\| = \|x\| = 1.$$

Now, we take any $x_0 \in X \setminus Y$ and then, since Y is closed, there is $r > 0$ so that $B(x_0, r) \cap Y = \emptyset$. This implies that

$$\inf_{y\in Y} \|x_0 - y\| \geq r > 0.$$

We denote $d_0 = \inf_{y\in Y} \|x_0 - y\|$ and then there is $y_0 \in Y$ so that

$$\frac{d_0}{t} > \|x_0 - y_0\| \geq d_0.$$

We set $x = \frac{x_0 - y_0}{\|x_0 - y_0\|}$ and then $\|x\| = 1$. Also

$$\|x - y\| = \left\| \frac{x_0 - y_0}{\|x_0 - y_0\|} - y \right\| = \frac{\|x_0 - (y_0 + \|x_0 - y_0\| y)\|}{\|x_0 - y_0\|} \geq \frac{d_0}{\|x_0 - y_0\|} > t$$

for every $y \in Y$. Thus, $\inf_{y\in Y} \|x - y\| \geq t$. □

We recall that for a subset A of a linear space X the set $\mathrm{span}(A)$, the *linear span* of A, is the linear subspace of X generated by A or, equivalently, the smallest linear subspace of X containing A or, equivalently, the set of all linear combinations of elements of A.

Proposition 1.25. *Let X be a normed space with $\dim(X) = +\infty$. Then the closed ball $\overline{B}(0; 1)$ is not compact.*

Proof. We take any $x_1 \in X$ with $\|x_1\| = 1$. Then the subspace $Y_1 = \mathrm{span}(\{x_1\})$ is a one-dimensional, and hence closed, subspace of X. By Riesz's lemma there is $x_2 \in X$ with $\|x_2\| = 1$ and $\inf_{y\in Y_1} \|x_2 - y\| \geq \frac{1}{2}$ and hence

$$\|x_2 - x_1\| \geq \frac{1}{2}.$$

Then the subspace $Y_2 = \mathrm{span}(\{x_1, x_2\})$ is a two-dimensional, and hence closed, subspace of X. By Riesz's lemma there is $x_3 \in X$ with $\|x_3\| = 1$ and $\inf_{y\in Y_2} \|x_3 - y\| \geq \frac{1}{2}$ and hence

$$\|x_3 - x_1\| \geq \frac{1}{2}, \quad \|x_3 - x_2\| \geq \frac{1}{2}.$$

Continuing inductively, we generate a sequence (x_n) in $\overline{B}(0; 1)$ so that

$$\|x_n - x_m\| \geq \frac{1}{2}$$

for every n, m with $n \neq m$. Obviously, this sequence has no convergent subsequence and so $\overline{B}(0; 1)$ is not compact. □

Proposition 1.26. *Let X be a normed space with $\dim(X) = +\infty$. Then every compact subset of X has empty interior.*

Proof. Let $K \subseteq X$ be compact and assume that a is an interior point of K.
Then $\overline{B}(a; r) \subseteq K$ for some $r > 0$ and so $\overline{B}(a; r)$ is compact. But the function $f : X \to X$ defined for every $x \in X$ by

$$f(x) = \frac{1}{r}\,(x - a)$$

is continuous in X and $f(\overline{B}(a; r)) = \overline{B}(0; 1)$. Therefore, $\overline{B}(0; 1)$ is compact and we arrive at a contradiction. So K has no interior points. □

1.13 Series.

Definition. *Let I be a non-empty set of indices and $\{\alpha_i \,|\, i \in I\}$ a set of non-negative real numbers, i.e. $\alpha_i \geq 0$ for every $i \in I$. We say that the* **series** $\sum_{i\in I} \alpha_i$ **converges** *if*

$$S = \sup\left\{ \sum_{i\in J} \alpha_i \,\middle|\, J \textit{ finite } \subseteq I \right\} < +\infty.$$

Then we also say that S is the **sum** *of the α_i, $i \in I$, and we write*

$$\sum_{i\in I} \alpha_i = S.$$

Lemma 1.3. *Let I be a non-empty set of indices and $\alpha_i \geq 0$ for every $i \in I$. If $\sum_{i\in I} \alpha_i$ converges, then the set $I_0 = \{i \in I \,|\, a_i > 0\}$ is countable.*

Proof. We consider $I_n = \{i \in I \,|\, \alpha_i \geq \frac{1}{n}\}$, and then we have $I_0 = \bigcup_{n=1}^{+\infty} I_n$. We take any finite $J \subseteq I_n$, and then

$$\frac{1}{n}\,\mathrm{card}(J) \leq \sum_{i\in J} \alpha_i \leq S.$$

Thus, $\mathrm{card}(J) \leq nS$, and so I_n is finite with $\mathrm{card}(I_n) \leq nS$. Therefore, I_0 is countable. $\square$

Definition. *Let X be a normed space with norm $\|\cdot\|$ and let (x_n) be a sequence in X. We say that the series $\sum_{n=1}^{+\infty} x_n$* **converges** *to $s \in X$ if $x_1 + \cdots + x_n \to s$. Then we also say that s is the* **sum** *of $\sum_{n=1}^{+\infty} x_n$ and we write*

$$\sum_{n=1}^{+\infty} x_n = s.$$

Theorem 1.10. *Let X be a Banach space with norm $\|\cdot\|$ and let $\{x_i \,|\, i \in I\} \subseteq X$, where I is a non-empty set of indices. If the series $\sum_{i\in I} \|x_i\|$ converges, then $I_0 = \{i \in I \,|\, x_i \neq 0\}$ is countable.*
(i) If I_0 is finite, then $\sum_{i\in I_0} x_i$ is just a finite sum.
(ii) If $\mathrm{card}(I_0) = +\infty$ and if $\{i_1, i_2, \ldots\}$ is any enumeration of I_0, then the series $\sum_{k=1}^{+\infty} x_{i_k}$ converges and the sum $s = \sum_{k=1}^{+\infty} x_{i_k}$ does not depend on the particular enumeration of I_0.

Proof. Lemma 1.3 implies that I_0 is countable.
(i) This is trivial.
(ii) We assume that $\mathrm{card}(I_0) = +\infty$ and that $\{i_1, i_2, \ldots\}$ is any enumeration of I_0. We also consider the partial sums

$$s_n = \sum_{k=1}^{n} x_{i_k}.$$

Let

$$\sum_{i\in I} \|x_i\| = S.$$

We take any $\epsilon > 0$ and then there is a finite $J \subseteq I$ so that

$$S - \epsilon < \sum_{i\in J} \|x_i\| \leq S.$$

If $J' \subseteq I$ is finite and $J \cap J' = \emptyset$, then

$$S - \epsilon + \sum_{i\in J'} \|x_i\| < \sum_{i\in J} \|x_i\| + \sum_{i\in J'} \|x_i\| = \sum_{i\in J\cup J'} \|x_i\| \leq S,$$

and so

$$\sum_{i\in J'} \|x_i\| < \epsilon.$$

Now, there is n_0 large enough so that $J \cap \{i_{n_0}, i_{n_0+1}, \ldots\} = \emptyset$. So if $n_0 \le n < m$, then the sets J and $J' = \{i_{n+1}, \ldots, i_m\}$ are disjoint and hence

$$\|s_m - s_n\| = \left\| \sum_{k=n+1}^{m} x_{i_k} \right\| \le \sum_{k=n+1}^{m} \|x_{i_k}\| = \sum_{i\in J'} \|x_i\| < \epsilon.$$

Therefore, (s_n) is a Cauchy sequence in X and so it converges to some $s \in X$. I.e.

$$\sum_{k=1}^{+\infty} x_{i_k} = s.$$

Finally, we consider any other enumeration $\{j_1, j_2, \ldots\}$ of I_0 and we consider the corresponding partial sums

$$t_n = \sum_{k=1}^{n} x_{j_k}.$$

Now, there is n_0 large enough so that $J \cap \{i_{n_0}, i_{n_0+1}, \ldots\} = \emptyset$ and $J \cap \{j_{n_0}, j_{n_0+1}, \ldots\} = \emptyset$. So if $n \ge n_0$, then the difference

$$s_n - t_n = \sum_{k=1}^{n} x_{i_k} - \sum_{k=1}^{n} x_{j_k}$$

contains only terms $\pm x_i$ with indices $i \in J'$, where $J' \subseteq I$ is finite and $J \cap J' = \emptyset$. Therefore, if $n \ge n_0$, then

$$\|s_n - t_n\| = \left\| \sum_{i\in J'} \pm x_i \right\| \le \sum_{i\in J'} \|x_i\| < \epsilon.$$

Thus, $s_n - t_n \to 0$, and since $s_n \to s$, we also get $t_n \to s$. $\square$

Definition. *Let X be a normed space with norm $\|\cdot\|$ and let $\{x_i \,|\, i \in I\} \subseteq X$, where I is a non-empty set of indices.*
(i) If the series $\sum_{i\in I} \|x_i\|$ converges, we say that the series $\sum_{i\in I} x_i$ **converges absolutely**.
(ii) If $I_0 = \{i \in I \,|\, x_i \ne 0\}$ is countable and infinite, and if the series $\sum_{k=1}^{+\infty} x_{i_k}$ converges in X for every enumeration $\{i_1, i_2, \ldots\}$ of I_0, and if the sum s of the last series does not depend on the enumeration of I_0, then we say that the series $\sum_{i\in I} x_i$ **converges unconditionally** *and that s is its* **sum**, *and we write*

$$\sum_{i\in I} x_i = s.$$

So theorem 1.10 says that, in a Banach space, if a series converges absolutely then it converges unconditionally.

1.14 Separable normed spaces.

Definition. *Let X be a normed space. We say that X is* **separable** *if there is a countable subset of X which is dense in X.*

Proposition 1.27. *All spaces l^p, $1 \le p < +\infty$, and c, c_0 are separable, but l^∞ is not separable.*

Proof. We say that $\lambda \in \mathbb{C}$ is rational if $\operatorname{Re}\lambda, \operatorname{Im}\lambda \in \mathbb{Q}$. It is obvious that the set of rational complex numbers is countable and dense in $\mathbb{C}$.
We consider the set

$$A = \{(\kappa_1, \ldots, \kappa_k, 0, 0, \ldots) \mid k \in \mathbb{N}, \kappa_1, \ldots, \kappa_k \text{ are rational in } F\}.$$

Then A is countable and it is a subset of every l^p, $1 \leq p < +\infty$, and of c_0.
Let $1 \leq p < +\infty$, and take any $x = (\lambda_k) \in l^p$ and any $\epsilon > 0$. Then there is k_0 so that

$$\sum_{k=k_0+1}^{+\infty} |\lambda_k|^p \leq \frac{\epsilon^p}{2}.$$

Also, for every $k = 1, \ldots, k_0$ there is a rational $\kappa_k \in F$ so that $|\lambda_k - \kappa_k| \leq \frac{\epsilon}{2^{1/p} k_0^{1/p}}$. We consider the element $y = (\kappa_1, \ldots, \kappa_{k_0}, 0, 0, \ldots) \in A$, and then

$$\|x - y\|_p^p = \sum_{k=1}^{k_0} |\lambda_k - \kappa_k|^p + \sum_{k=k_0+1}^{+\infty} |\lambda_k|^p \leq \epsilon^p,$$

and hence $\|x - y\|_p \leq \epsilon$. Thus A is dense in l^p and so l^p is separable.
Now, take any $x = (\lambda_k) \in c_0$ and any $\epsilon > 0$. Then there is k_0 so that $|\lambda_k| \leq \epsilon$ for every $k \geq k_0 + 1$. Also, for every $k = 1, \ldots, k_0$ there is a rational $\kappa_k \in F$ so that $|\lambda_k - \kappa_k| \leq \epsilon$. Then $y = (\kappa_1, \ldots, \kappa_{k_0}, 0, 0, \ldots) \in A$ and

$$\|x - y\|_\infty = \sup\{|\lambda_1 - \kappa_1|, \ldots, |\lambda_{k_0} - \kappa_{k_0}|, |\lambda_{k_0+1}|, |\lambda_{k_0+2}|, \ldots\} \leq \epsilon.$$

Therefore A is dense in c_0 and so c_0 is separable.
For the space c we consider the set

$$B = \{(\kappa_1, \ldots, \kappa_k, \kappa, \kappa, \ldots) \mid k \in \mathbb{N}, \kappa, \kappa_1, \ldots, \kappa_k \text{ are rational in } F\}.$$

Then B is a countable subset of c.
Now we take any $x = (\lambda_k) \in c$ and any $\epsilon > 0$. If $\lambda = \lim_{k \to +\infty} \lambda_k$, then there is k_0 so that $|\lambda_k - \lambda| \leq \frac{\epsilon}{2}$ for every $k \geq k_0 + 1$. Now, for every $k = 1, \ldots, k_0$ there is a rational $\kappa_k \in F$ so that $|\lambda_k - \kappa_k| \leq \epsilon$. Also, there is a rational $\kappa \in F$ so that $|\lambda - \kappa| < \frac{\epsilon}{2}$ and hence

$$|\lambda_k - \kappa| \leq |\lambda_k - \lambda| + |\lambda - \kappa| < \frac{\epsilon}{2} + \frac{\epsilon}{2} = \epsilon$$

for every $k \geq k_0$. Then $y = (\kappa_1, \ldots, \kappa_{k_0}, \kappa, \kappa, \ldots) \in B$ and

$$\|x - y\|_\infty = \sup\{|\lambda_1 - \kappa_1|, \ldots, |\lambda_{k_0} - \kappa_{k_0}|, |\lambda_{k_0+1} - \kappa|, |\lambda_{k_0+2} - \kappa|, \ldots\} \leq \epsilon.$$

Thus B is dense in c, and so c is separable.
Finally, assume that l^∞ has a countable and dense subset

$$C = \{x_1, x_2, \ldots\},$$

where $x_n = (\lambda_{n,k})$ for every n.
For each k we consider $\lambda_k \in F$ so that $|\lambda_k| \leq 1$ and $|\lambda_k - \lambda_{k,k}| \geq 1$, and we form the element $x = (\lambda_k) \in l^\infty$. Then

$$\|x - x_n\|_\infty \geq |\lambda_n - \lambda_{n,n}| \geq 1$$

for every n. So there is no element of C at a distance from x less than 1 and we arrive at a contradiction. $\square$

Proposition 1.28. *Let (Ω, Σ, μ) be a measure space, and assume that there is a countable $\Xi \subseteq \Sigma$ so that for every $A \in \Sigma$ with $\mu(A) < +\infty$ and every $\epsilon > 0$ there is $B \in \Xi$ with $\mu(B \triangle A) \leq \epsilon$. Then every $L^p(\Omega, \Sigma, \mu)$, $1 \leq p < +\infty$, is separable.*

Proof. We take any $f \in L^p(\Omega, \Sigma, \mu)$ and any $\epsilon > 0$. We know that there is a simple function

$$g = \sum_{k=1}^{n} \lambda_k \chi_{A_k} \in L^p(\Omega, \Sigma, \mu)$$

so that $\lambda_k \in F$, $A_k \in \Sigma$ and $\mu(A_k) < +\infty$ for every $k = 1, \ldots, n$ and

$$\Big(\int_\Omega |f - g|^p \, d\mu \Big)^{1/p} < \frac{1}{2}\, \epsilon.$$

We select $\eta > 0$ depending on ϵ in a way to be made precise in a moment.
For every $k = 1, \ldots, n$ there is $B_k \in \Xi$ so that $\mu(B_k \triangle A_k) \leq \eta$ and there is a rational $\kappa_k \in F$ so that $|\lambda_k - \kappa_k| \leq \eta$. We consider the function

$$h = \sum_{k=1}^{n} \kappa_k \chi_{B_k}$$

and we get

$$\begin{aligned}
\Big(\int_\Omega |f - h|^p \, d\mu \Big)^{1/p} &\leq \Big(\int_\Omega |f - g|^p \, d\mu \Big)^{1/p} + \Big(\int_\Omega |g - h|^p \, d\mu \Big)^{1/p} \\
&\leq \frac{1}{2}\, \epsilon + \sum_{k=1}^{n} |\lambda_k - \kappa_k| \Big(\int_\Omega |\chi_{A_k}|^p \, d\mu \Big)^{1/p} \\
&\quad + \sum_{k=1}^{n} |\kappa_k| \Big(\int_\Omega |\chi_{A_k} - \chi_{B_k}|^p \, d\mu \Big)^{1/p} \\
&= \frac{1}{2}\, \epsilon + \sum_{k=1}^{n} |\lambda_k - \kappa_k| (\mu(A_k))^{1/p} + \sum_{k=1}^{n} |\kappa_k| (\mu(B_k \triangle A_k))^{1/p} \\
&\leq \frac{1}{2}\, \epsilon + \eta \sum_{k=1}^{n} (\mu(A_k))^{1/p} + \eta^{1/p} \sum_{k=1}^{n} (|\lambda_k| + \eta).
\end{aligned}$$

Since

$$\eta \sum_{k=1}^{n} (\mu(A_k))^{1/p} + \eta^{1/p} \sum_{k=1}^{n} (|\lambda_k| + \eta) \to 0$$

when $\eta \to 0+$, we may select η so that the last sum is $\leq \frac{1}{2}\, \epsilon$ and hence $\big(\int_\Omega |f - h|^p \, d\mu \big)^{1/p} \leq \epsilon$. So the set

$$Q = \Big\{ \sum_{k=1}^{n} \kappa_k \chi_{B_k} \, \Big| \, n \in \mathbb{N}, \kappa_1, \ldots, \kappa_n \text{ are rational in } F, B_1, \ldots, B_n \in \Xi \Big\}$$

is countable and dense in $L^p(\Omega, \Sigma, \mu)$. $\square$

It is known that if Ω is a Borel set in $\mathbb{R}^d$, if $\Sigma = \mathcal{B}(\Omega)$, and if $\mu = m$ is the Lebesgue measure, then the collection Ξ of the sets of the form $B = P \cap \Omega$, where P is any finite union of parallelepipeds with rational vertices, has the assumed property in the last proposition. So the corresponding spaces $L^p(\Omega, \mathcal{B}(\Omega), m)$, $1 \leq p < +\infty$, are separable.

Chapter 2

Inner product spaces

2.1 Inner products.

Let X be a linear space over the field F, where $F = \mathbb{R}$ or $F = \mathbb{C}$.

Definition. *We say that the function*

$$\langle \cdot, \cdot \rangle : X \times X \to F$$

is an **inner product** *on* X*, if*

(i) $\langle x, x \rangle \geq 0$,

(ii) $\langle x, x \rangle = 0 \;\Leftrightarrow\; x = 0$,

(iii) $\langle y, x \rangle = \overline{\langle x, y \rangle}$,

(iv) $\langle \lambda x, y \rangle = \lambda \, \langle x, y \rangle$,

(v) $\langle x_1 + x_2, y \rangle = \langle x_1, y \rangle + \langle x_2, y \rangle$,

for every $x, x_1, x_2, y \in X$ *and every* $\lambda \in F$.

If $F = \mathbb{R}$, then of course (iii) becomes $\langle y, x \rangle = \langle x, y \rangle$.

Properties (iv), (v) say that $\langle \cdot, \cdot \rangle$ is linear in the first variable. If we combine these two properties with (iii) we get that $\langle \cdot, \cdot \rangle$ is *conjugate-linear* in the second variable:

(vi) $\langle x, \lambda y \rangle = \overline{\lambda} \, \langle x, y \rangle$,

(vii) $\langle x, y_1 + y_2 \rangle = \langle x, y_1 \rangle + \langle x, y_2 \rangle$,

for every $x, y, y_1, y_2 \in X$ and every $\lambda \in F$.

Again, if $F = \mathbb{R}$, then $\langle \cdot, \cdot \rangle$ is linear in the second variable.

Using $\lambda = 0$ in (iv) and (vi), we get

$$\langle 0, y \rangle = \langle x, 0 \rangle = 0$$

for every $x, y \in X$.

A very useful identity which results easily from (iv)-(vii) is

$$\langle \lambda x + \kappa y, \lambda x + \kappa y \rangle = |\lambda|^2 \langle x, x \rangle + 2 \, \mathrm{Re}(\lambda \, \overline{\kappa} \, \langle x, y \rangle) + |\kappa|^2 \langle y, y \rangle \tag{2.1}$$

for every $x, y \in X$ and every $\lambda, \kappa \in F$.

Schwarz's inequality. *Let X be a linear space with an inner product $\langle \cdot, \cdot \rangle$. Then*

$$|\langle x, y\rangle|^2 \leq \langle x, x\rangle \langle y, y\rangle$$

for every $x, y \in X$.

Proof. Let $\langle x, x\rangle = 0$. Then $x = 0$, hence $\langle x, y\rangle = 0$ and so Schwarz's inequality becomes $0 \leq 0$. Now let $\langle x, x\rangle > 0$. Then

$$|\lambda|^2\langle x, x\rangle + 2\,\mathrm{Re}(\lambda\,\langle x, y\rangle) + \langle y, y\rangle = \langle \lambda x + y, \lambda x + y\rangle \geq 0$$

for every $\lambda \in F$.
There is $\mu \in F$ so that $|\mu| = 1$ and $\mu\langle x, y\rangle = |\langle x, y\rangle|$. Taking $\lambda = t\mu$ with $t \in \mathbb{R}$ we get

$$t^2\langle x, x\rangle + 2t|\langle x, y\rangle| + \langle y, y\rangle \geq 0$$

for every $t \in \mathbb{R}$. If we use $t = -\frac{|\langle x,y\rangle|}{\langle x,x\rangle}$ in the last inequality, we get Schwarz's inequality. □

Proposition 2.1. *Let X be a linear space with an inner product $\langle \cdot, \cdot \rangle$. Then the function $\|\cdot\| : X \to \mathbb{R}$ defined for every $x \in X$ by*

$$\|x\| = \sqrt{\langle x, x\rangle}$$

is a norm on X.

Proof. All properties of the norm are easy to prove. For example, for the last property:

$$\begin{aligned}\|x + y\|^2 &= \langle x + y, x + y\rangle = \langle x, x\rangle + 2\,\mathrm{Re}(\langle x, y\rangle) + \langle y, y\rangle \leq \langle x, x\rangle + 2|\langle x, y\rangle| + \langle y, y\rangle \\ &\leq \langle x, x\rangle + 2\sqrt{\langle x, x\rangle}\,\sqrt{\langle y, y\rangle} + \langle y, y\rangle = \|x\|^2 + 2\|x\|\|y\| + \|y\|^2 = (\|x\| + \|y\|)^2,\end{aligned}$$

and hence $\|x + y\| \leq \|x\| + \|y\|$. □

Definition. *We say that the norm $\|\cdot\|$, which is defined by the inner product $\langle \cdot, \cdot \rangle$ as above, is the norm* **induced** *by the inner product.*

Now, Schwarz's inequality takes the form

$$|\langle x, y\rangle| \leq \|x\|\,\|y\|.$$

Also, identity (2.1) becomes

$$\|\lambda x + \kappa y\|^2 = |\lambda|^2\|x\|^2 + 2\,\mathrm{Re}(\lambda\,\overline{\kappa}\,\langle x, y\rangle) + |\kappa|^2\|y\|^2.$$

Moreover, taking $\lambda = 1, \kappa = 1$ and also $\lambda = 1, \kappa = -1$, and then adding the two resulting identities, we get the **parallelogram law**:

$$\|x + y\|^2 + \|x - y\|^2 = 2\|x\|^2 + 2\|y\|^2.$$

Example 2.1.1. A trivial example of a normed space over F is the field F itself with the inner product $\langle \lambda, \kappa\rangle = \lambda\,\overline{\kappa}$.

Proposition 2.2. *The inner product of an inner product space X is continuous, i.e. if $x_n \to x$ and $y_n \to y$ in X, then $\langle x_n, y_n\rangle \to \langle x, y\rangle$ in $\mathbb{R}$.*

Proof. This is implied by

$$\begin{aligned}|\langle x_n, y_n\rangle - \langle x, y\rangle| &\leq |\langle x_n, y_n\rangle - \langle x_n, y\rangle| + |\langle x_n, y\rangle - \langle x, y\rangle| = |\langle x_n, y_n - y\rangle| + |\langle x_n - x, y\rangle| \\ &\leq \|x_n\|\|y_n - y\| + \|x_n - x\|\|y\|.\end{aligned}$$

□

Definition. *Let X be an inner product space. If X with the norm induced by the inner product is complete, then we say that X is a* **Hilbert space***.*

Hence, a Hilbert space is Banach space with a norm induced by an inner product.

2.2 Subspaces, cartesian products.

Proposition 2.3. *Let X be an inner product space with inner product $\langle \cdot, \cdot \rangle : X \times X \to F$, and let Y be a linear subspace of X. Then the restriction $\langle \cdot, \cdot \rangle : Y \times Y \to F$ is an inner product on Y.*

Proof. Obvious. $\square$

Definition. *The linear subspace Y of an inner product space X, equipped with the restriction on Y of the inner product on X, is called* **subspace** *of X.*

Let $X_1, \ldots, X_m$ be inner product spaces with inner products $\langle \cdot, \cdot \rangle_1, \ldots, \langle \cdot, \cdot \rangle_m$. We consider the *cartesian product* $X = X_1 \times \cdots \times X_m$ and for every $x = (x_1, \ldots, x_m), y = (y_1, \ldots, y_m) \in X = X_1 \times \cdots \times X_m$ we define

$$\langle x, y \rangle = \langle x_1, y_1 \rangle_1 + \cdots + \langle x_m, y_m \rangle_m.$$

Proposition 2.4. *The function $\langle \cdot, \cdot \rangle : X \times X \to F$ just defined is an inner product on $X = X_1 \times \cdots \times X_m$.*

Proof. Trivial. $\square$

Example 2.2.1. We consider $X_1 = \ldots = X_m = F$ with $\langle \lambda, \kappa \rangle_1 = \ldots = \langle \lambda, \kappa \rangle_m = \lambda \, \overline{\kappa}$ and then we get the cartesian product $X = F \times \cdots \times F = F^m$ with the inner product which is defined for every $x = (\lambda_1, \ldots, \lambda_m), y = (\kappa_1, \ldots, \kappa_m) \in F \times \cdots \times F = F^m$ by

$$\langle x, y \rangle = \lambda_1 \, \overline{\kappa_1} + \cdots + \lambda_m \, \overline{\kappa_m}.$$

This is the standard **euclidean inner product** on F^m. Obviously, the norm induced by this inner product is the euclidean norm on F^m:

$$\langle x, x \rangle = \lambda_1 \, \overline{\lambda_1} + \cdots + \lambda_m \, \overline{\lambda_m} = |\lambda_1|^2 + \cdots + |\lambda_m|^2 = \|x\|_2^2.$$

2.3 Linear isometries.

Definition. *Let X, Y be inner product spaces with inner products $\langle \cdot, \cdot \rangle_X, \langle \cdot, \cdot \rangle_Y$, and let $T : X \to Y$ be a linear operator with the property*

$$\langle T(x_1), T(x_2) \rangle_Y = \langle x_1, x_2 \rangle_X$$

for every $x_1, x_2 \in X$. Then T is called **linear isometry** *of X* **into** *Y.*
If T is onto Y, i.e. if $T(X) = Y$, then T is called **linear isometry** *of X* **onto** *Y. We also say that X is* **linearly isometric** *to Y.*

Taking $x_1 = x_2 = x \in X$, we see that if $T : X \to Y$ is a linear isometry, then

$$\|T(x)\|_Y = \|x\|_X$$

for every $x \in X$, where the two norms are those which are induced by the inner products. In other words, an "inner product" linear isometry is also a "norm" linear isometry. We shall immediately see that the converse is also true. Indeed, assume that

$$\|T(x)\|_Y = \|x\|_X$$

for every $x \in X$. Then, taking $x = x_1 + x_2$, we get

$$\|T(x_1)\|_Y^2 + 2\,\mathrm{Re}(\langle T(x_1), T(x_2) \rangle_Y) + \|T(x_2)\|_Y^2 = \|x_1\|_X^2 + 2\,\mathrm{Re}(\langle x_1, x_2 \rangle_X) + \|x_2\|_X^2$$

and hence

$$\text{Re}(\langle T(x_1), T(x_2)\rangle_Y) = \text{Re}(\langle x_1, x_2\rangle_X)$$

for every $x_1, x_2 \in X$. If $F = \mathbb{R}$, then of course we get

$$\langle T(x_1), T(x_2)\rangle_Y = \langle x_1, x_2\rangle_X$$

for every $x_1, x_2 \in X$. If $F = \mathbb{C}$, then we use ix_2 in the place of x_2 and we get

$$\text{Im}(\langle T(x_1), T(x_2)\rangle_Y) = \text{Im}(\langle x_1, x_2\rangle_X)$$

for every $x_1, x_2 \in X$. Therefore

$$\langle T(x_1), T(x_2)\rangle_Y = \langle x_1, x_2\rangle_X$$

for every $x_1, x_2 \in X$.

Proposition 2.5. *Let X be an inner product space with inner product $\langle \cdot, \cdot\rangle_X$, let Y be a linear space and let $T : X \to Y$ be a linear operator which is one-to-one in X and onto Y. Then there is an inner product on Y so that T becomes a linear isometry of X onto Y.*

Proof. We take any $y_1, y_2 \in Y$, we consider the unique $x_1, x_2 \in X$ so that $T(x_1) = y_1$ and $T(x_2) = y_2$ and we define

$$\langle y_1, y_2\rangle_Y = \langle x_1, x_2\rangle_X.$$

It is easy to prove that the function $\langle \cdot, \cdot\rangle_Y : Y \times Y \to F$ just defined is an inner product on Y. Then, since $T(x_1) = y_1$, $T(x_2) = y_2$, the equality $\langle y_1, y_2\rangle_Y = \langle x_1, x_2\rangle_X$ can be written

$$\langle T(x_1), T(x_2)\rangle_Y = \langle x_1, x_2\rangle_X$$

and so T is a linear isometry of X onto Y. □

Thus, when we have two isomorphic linear spaces and one of them has an inner product, then we can transfer this inner product to the other linear space so that the two spaces become linearly isometric.

Example 2.3.1. Let X be a linear space of finite dimension and let $\{b_1, \ldots, b_m\}$ be a basis of X. We consider the inner product space F^m with the euclidean inner product. We also consider the linear operator $T : F^m \to X$ defined for every $(\lambda_1, \ldots, \lambda_m) \in F^m$ by

$$T(\lambda_1, \ldots, \lambda_m) = \lambda_1 b_1 + \cdots + \lambda_m b_m.$$

Then T is one-to-one in F^m and onto X, and so the euclidean inner product on F^m can be transfered to an inner product $\langle \cdot, \cdot\rangle : X \times X \to F$. This is defined for every $x = \lambda_1 b_1 + \cdots + \lambda_m b_m$ and $y = \kappa_1 b_1 + \cdots + \kappa_m b_m$ in X by the formula

$$\begin{aligned}
\langle x, y\rangle &= \langle \lambda_1 b_1 + \cdots + \lambda_m b_m, \kappa_1 b_1 + \cdots + \kappa_m b_m\rangle \\
&= \langle T(\lambda_1, \ldots, \lambda_m), T(\kappa_1, \ldots, \kappa_m)\rangle \\
&= \langle (\lambda_1, \ldots, \lambda_m), (\kappa_1, \ldots, \kappa_m)\rangle \\
&= \lambda_1 \overline{\kappa_1} + \cdots + \lambda_m \overline{\kappa_m}.
\end{aligned}$$

The inner product on X just defined is called **euclidean inner product** on X with respect to the basis $\{b_1, \ldots, b_m\}$.

2.4 Completion.

Definition. *Let X be an inner product space. We say that the inner product space $\overline{X}$ is a* **completion** *of X if $\overline{X}$ is complete, i.e. a Hilbert space, and there is a linear isometry $T : X \to \overline{X}$ so that $T(X)$ is a dense subspace of $\overline{X}$.*

Theorem 2.1. *Let X be an inner product space. Then there is at least one completion of X. Moreover, every two completions of X are linearly isometric.*

Proof. This is just a variant of the proof of theorem 1.1. Again we consider the set $\widehat{X}$ of all Cauchy sequences of X and then the same linear space $\overline{X}$ of the equivalence classes of Cauchy sequences. Now, instead of defining the norm on $\overline{X}$, we define the inner product by

$$\langle [(x_n)], [(y_n)] \rangle_{\overline{X}} = \lim_{n \to +\infty} \langle x_n, y_n \rangle.$$

It is obvious that the norm on $\overline{X}$ which is induced by the inner product just defined on $\overline{X}$ is the same as the norm defined in the proof of theorem 1.1. Then the rest of the proof is the same as the proof of theorem 1.1. The details are left to the interested reader. $\square$

2.5 Examples.

Besides the finite dimensional Hilbert spaces with their euclidean inner products, we have the following examples.

1. We have the sequence space

$$l^2 = \Big\{ (\lambda_k) \, \Big| \, \sum_{k=1}^{+\infty} |\lambda_k|^2 < +\infty \Big\}.$$

The inner product on l^2 is defined by

$$\langle x, y \rangle = \sum_{k=1}^{+\infty} \lambda_k \, \overline{\kappa_k}$$

for every $x = (\lambda_k), y = (\kappa_k) \in l^2$. Of course, the norm induced by this inner product is the 2-norm of l^2 which we know from the previous chapter:

$$\sqrt{\langle x, x \rangle} = \Big(\sum_{k=1}^{+\infty} \lambda_k \, \overline{\lambda_k} \Big)^{1/2} = \Big(\sum_{k=1}^{+\infty} |\lambda_k|^2 \Big)^{1/2} = \|x\|_2.$$

Schwarz's inequality in this case is a special case of Hölder's inequality:

$$\Big| \sum_{k=1}^{+\infty} \lambda_k \, \overline{\kappa_k} \Big| \leq \Big(\sum_{k=1}^{+\infty} |\lambda_k|^2 \Big)^{1/2} \Big(\sum_{k=1}^{+\infty} |\kappa_k|^2 \Big)^{1/2}.$$

Of course, with this inner product l^2 is a Hilbert space.

2. Then we have the function space

$$L^2(\Omega, \Sigma, \mu) = \Big\{ f \in \mathcal{M}(\Omega) \, \Big| \, \int_\Omega |f|^2 \, d\mu \Big\}.$$

The inner product on $L^2(\Omega, \Sigma, \mu)$ is defined by

$$\langle f, g \rangle = \int_\Omega f \, \overline{g} \, d\mu$$

for every $f, g \in L^2(\Omega, \Sigma, \mu)$. Again, the norm induced by this inner product is the 2-norm of $L^2(\Omega, \Sigma, \mu)$:

$$\sqrt{\langle f, f\rangle} = \Big(\int_\Omega f\,\overline{f}\,d\mu\Big)^{1/2} = \Big(\int_\Omega |f|^2\,d\mu\Big)^{1/2} = \|f\|_2.$$

As in the previous example, Schwarz's inequality is a special case of Hölder's inequality:

$$\Big|\int_\Omega f\,\overline{g}\,d\mu\Big| \le \Big(\int_\Omega |f|^2\,d\mu\Big)^{1/2}\Big(\int_\Omega |g|^2\,d\mu\Big)^{1/2}.$$

Moreover, $L^2(\Omega, \Sigma, \mu)$ with this inner product is a Hilbert space.
3. Finally, we have the Sobolev space $W^{k,2}(U)$, which is also denoted $H^k(U)$, i.e.

$$H^k(U) = W^{k,2}(U).$$

We recall that $H^k(U)$ is the set of all functions $f \in L^2(U)$ which have weak α-derivatives $D^\alpha f$ in $L^2(U)$ for every α with $|\alpha| \le k$. The inner product on $H^k(U)$ is defined by

$$\langle f, g\rangle_k = \sum_{|\alpha|\le k}\int_U D^\alpha f\,\overline{D^\alpha g}\,dm.$$

With this inner product, $H^k(U)$ is a Hilbert space.

2.6 Convex sets.

We know that a set K in a linear space X is **convex** if

$$a, b \in K,\ 0 \le t \le 1 \quad\Rightarrow\quad (1-t)a + tb \in K.$$

The set

$$[a, b] = \{(1-t)a + tb \,|\, 0 \le t \le 1\}$$

is considered as the *linear segment* with endpoints a, b.

Proposition 2.6. *Let X be an inner product space with inner product $\langle\cdot,\cdot\rangle$ and norm $\|\cdot\|$, let $K \subseteq X$ be convex and complete, and let $x_0 \in X$. Then there is a unique $y_0 \in K$ so that*

$$\|x_0 - y_0\| = \inf_{y\in K}\|x_0 - y\|.$$

Moreover,

$$\mathrm{Re}(\langle x_0 - y_0, y - y_0\rangle) \le 0$$

for every $y \in K$.
If X is a Hilbert space, we may only assume that K is convex and closed.

Proof. We denote

$$d = \inf_{y\in K}\|x_0 - y\|.$$

Then there is a sequence (y_n) in K so that

$$\|x_0 - y_n\| \to d$$

when $n \to +\infty$. Now, the parallelogram law implies

$$\begin{aligned}2\|x_0 - y_n\|^2 + 2\|x_0 - y_m\|^2 &= \|(x_0 - y_n) - (x_0 - y_m)\|^2 + \|(x_0 - y_n) + (x_0 - y_m)\|^2\\ &= \|y_n - y_m\|^2 + 4\Big\|x_0 - \frac{y_n + y_m}{2}\Big\|^2\\ &\ge \|y_n - y_m\|^2 + 2d^2\end{aligned}$$

for every n, m. The last inequality is implied by $\frac{y_n+y_m}{2} \in K$ which is due to the convexity of K. Taking the limit when $n, m \to +\infty$, we find that

$$\|y_n - y_m\| \to 0.$$

Thus, (y_n) is a Cauchy sequence in K and, since K is complete, there is $y_0 \in K$ so that $y_n \to y_0$. Now $\|x_0 - y_n\| \to \|x_0 - y_0\|$ and hence $\|x_0 - y_0\| = d$.
If we assume that $y_0' \in K$ and $\|x_0 - y_0'\| = d$, then exactly as before we have

$$4d^2 = 2\|x_0 - y_0\|^2 + 2\|x_0 - y_0'\|^2 = \|y_0 - y_0'\|^2 + 4\Big\|x_0 - \frac{y_0 + y_0'}{2}\Big\|^2 \geq \|y_0 - y_0'\|^2 + 4d^2.$$

Therefore, $\|y_0 - y_0'\|^2 \leq 0$ and so $y_0 = y_0'$. This proves the uniqueness of y_0.
Finally, we take any $y \in K$ and then for $0 \leq t \leq 1$ we have that $(1-t)y_0 + ty \in K$ and hence

$$\begin{aligned} d^2 &\leq \|x_0 - ((1-t)y_0 + ty)\|^2 = \|(x_0 - y_0) - t(y - y_0)\|^2 \\ &= \|x_0 - y_0\|^2 - 2t\,\mathrm{Re}(\langle x_0 - y_0, y - y_0\rangle) + t^2\|y - y_0\|^2 \\ &= d^2 - 2t\,\mathrm{Re}(\langle x_0 - y_0, y - y_0\rangle) + t^2\|y - y_0\|^2. \end{aligned}$$

When $0 < t \leq 1$ we get

$$2\,\mathrm{Re}(\langle x_0 - y_0, y - y_0\rangle) \leq t\|y - y_0\|^2,$$

and taking the limit when $t \to 0+$ we conclude that $\mathrm{Re}(\langle x_0 - y_0, y - y_0\rangle) \leq 0$. □

2.7 Orthogonality.

Definition. *Let X be an inner product space with inner product $\langle\cdot,\cdot\rangle$. Let $x, y \in X$, and $A, B \subseteq X$.*
(i) If $\langle x, y\rangle = 0$, we say that x, y are **orthogonal** *and we write*

$$x \perp y.$$

(ii) If $\langle x, a\rangle = 0$ for every $a \in A$, we say that x, A are **orthogonal** *and we write*

$$x \perp A.$$

(iii) If $\langle a, b\rangle = 0$ for every $a \in A$ and every $b \in B$, we say that A, B are **orthogonal** *and we write*

$$A \perp B.$$

It is obvious that

$$x \perp x \quad \Rightarrow \quad x = 0.$$

Therefore,

$$x \perp A,\ \ x \in A \quad \Rightarrow \quad x = 0,$$

$$A \perp B,\ \ A \cap B \neq \emptyset \quad \Rightarrow \quad A \cap B = \{0\}.$$

Proposition 2.7. *Let X be an inner product space, let $x, y, z \in X$ and (y_n) be a sequence in X.*
(i) If $x \perp y$ and $x \perp z$, then $x \perp (\lambda y + \kappa z)$ for every $\lambda, \kappa \in F$.
(ii) If $x \perp y_n$ for every n and $y_n \to y$, then $x \perp y$.

Proof. Trivial. □

For a subset A of a normed space X the set $\mathrm{clspan}(A)$, the *closed linear span* of A, is the closure of the linear span of A in X or, equivalently, the smallest closed subspace of X containing A or, equivalently, the set of the limits of the linear combinations of elements of A.

Proposition 2.8. *Let X be an inner product space, and let $x \in X$ and $A, B \subseteq X$.*
(i) If $x \perp A$, then $x \perp \text{clspan}(A)$.
(ii) If $A \perp B$, then $\text{clspan}(A) \perp \text{clspan}(B)$.

Proof. This is an easy corollary of proposition 2.7. □

Definition. *Let X be an inner product space and $A \subseteq X$. We denote*

$$A^{\perp} = \{x \in X \,|\, x \perp A\}.$$

We say that $A^{\perp}$ is the subspace which is **orthogonal** *to A.*

Proposition 2.9. *Let X be an inner product space and $A, B \subseteq X$.*
(i) $A^{\perp}$ is a closed subspace of X.
(ii) $\text{clspan}(A) \subseteq (A^{\perp})^{\perp}$.
(iii) $A \subseteq B \;\Rightarrow\; B^{\perp} \subseteq A^{\perp}$.
(iv) $(\text{clspan}(A))^{\perp} = A^{\perp}$.

Proof. Trivial. □

2.8 Otrhogonal complements.

Definition. *Let X be an inner product space. If Y, Z are subspaces of X and*

$$Y + Z = X, \quad Y \perp Z,$$

we say that each of Y, Z is the **orthogonal complement** *of the other.*

Proposition 2.10. *Let X be an inner product space, and let Y, Z be subspaces of X. If each of Y, Z is the orthogonal complement of the other, then $Z = Y^{\perp}$ and $Y = Z^{\perp}$. In particular, Y, Z are closed.*

Proof. Obviously, $Y \perp Z$ implies $Z \subseteq Y^{\perp}$.
Now, let $x \in Y^{\perp}$. Since $X = Y + Z$, there are $y \in Y$, $z \in Z$ so that $x = y + z$. From $x \in Y^{\perp}$ and $z \in Y^{\perp}$ we get $y = x - z \in Y^{\perp}$. Hence $y = 0$ and so $x = z \in Z$. Thus, $Y^{\perp} \subseteq Z$.
The proof of $Y = Z^{\perp}$ is symmetric. □

Proposition 2.11. *Let X be an inner product space, and let Y be a subspace of X.*
(i) Y has an orthogonal complement in X if and only if $Y + Y^{\perp} = X$.
(ii) If Y has an orthogonal complement in X, then Y is closed, its orthogonal complement is $Y^{\perp}$, and it is the orthogonal complement of $Y^{\perp}$, i.e. $Y = (Y^{\perp})^{\perp}$.

Proof. Clear from the definition and proposition 2.10. □

Theorem 2.2. *Let X be an inner product space with inner product $\langle \cdot, \cdot \rangle$ and norm $\|\cdot\|$, let Y be a complete subspace of X, and let $x_0 \in X$. Then there is a unique $y_0 \in Y$ so that*

$$\|x_0 - y_0\| = \inf_{y \in Y} \|x_0 - y\|.$$

Moreover,

$$x_0 - y_0 \perp Y.$$

Thus

$$X = Y + Y^{\perp}$$

and so each of $Y, Y^{\perp}$ is the orthogonal complement of the other. In particular, $Y = (Y^{\perp})^{\perp}$.
If X is a Hilbert space, we may only assume that Y is a closed subspace of X.

Proof. Every linear subspace is a convex set. So proposition 2.6 implies the existence of y_0 and also that

$$\mathrm{Re}(\langle x_0 - y_0, y - y_0\rangle) \leq 0$$

for every $y \in Y$. Since Y is a linear subspace, we have that $y \in Y$ if and only if $y - y_0 \in Y$, and so, replacing $y - y_0$ with y in the last inequality, we get

$$\mathrm{Re}(\langle x_0 - y_0, y\rangle) \leq 0$$

for every $y \in Y$. Now, replacing y with $-y$, we get $\mathrm{Re}(\langle x_0 - y_0, y\rangle) \geq 0$ for every $y \in Y$. So we have that

$$\mathrm{Re}(\langle x_0 - y_0, y\rangle) = 0$$

for every $y \in Y$.
If $F = \mathbb{R}$, then we get $\langle x_0 - y_0, y\rangle = 0$ for every $y \in Y$ and so $x_0 - y_0 \perp Y$.
If $F = \mathbb{C}$, then we replace y with iy and we get

$$\mathrm{Im}(\langle x_0 - y_0, y\rangle) = 0$$

for every $y \in Y$. Thus $\langle x_0 - y_0, y\rangle = 0$ for every $y \in Y$ and so $x_0 - y_0 \perp Y$.
If we set $z_0 = x_0 - y_0$, then we have $x_0 = y_0 + z_0$ with $y_0 \in Y$ and $z_0 \in Y^\perp$.
We conclude that $X = Y + Y^\perp$. All the rest are implied by proposition 2.11. □

Thus, every complete subspace of an inner product space (and hence every closed subspace of a Hilbert space) has an orthogonal complement.

Proposition 2.12. *Let X be an inner product space and $A \subseteq X$.*
(i) If $\mathrm{clspan}(A)$ *is complete, then* $\mathrm{clspan}(A) = (A^\perp)^\perp$.
(ii) If X is a Hilbert space, then $\mathrm{clspan}(A) = (A^\perp)^\perp$.

Proof. (i) Since $\mathrm{clspan}(A)$ is a complete subspace of X, theorem 2.2 implies that $(\mathrm{clspan}(A))^\perp = A^\perp$ is an orthogonal complement of $\mathrm{clspan}(A)$, and hence $\mathrm{clspan}(A) = (A^\perp)^\perp$.
(ii) Immediate from (i). □

2.9 Orders.

The content of this section is very general and belongs to the Foundations of Set Theory.

Definition. *Let A be a non-empty set and let $\prec \subseteq A \times A$. We say that the set $\prec$ is an* **order relation** *in A, if for every $a, a_1, a_2, a_3 \in A$:*
(i) $(a, a) \in \prec$,
(ii) if $(a_1, a_2) \in \prec$ and $(a_2, a_1) \in \prec$, then $a_1 = a_2$,
(iii) if $(a_1, a_2) \in \prec$ and $(a_2, a_3) \in \prec$, then $(a_1, a_3) \in \prec$.
If $\prec$ is an order relation in A, we say that A is **ordered** *by $\prec$.*
Finally, if $\prec$ is an order relation in A, we prefer to write

$$a \prec a'$$

instead of $(a, a') \in \prec$.

Thus, (i)-(iii) of the definition take the form
(i) $a \prec a$,
(ii) if $a_1 \prec a_2$ and $a_2 \prec a_1$, then $a_1 = a_2$,
(iii) if $a_1 \prec a_2$ and $a_2 \prec a_3$, then $a_1 \prec a_3$.

Example 2.9.1. $\mathbb{R}$ with the usual order relation $\leq$.

Example 2.9.2. $\mathbb{N}$ with the relation of divisibility $/$. I.e. a/b if a divides b.

Example 2.9.3. If Q is any non-empty set, we consider $\mathcal{P}(Q)$, the set of all subsets of Q, and as an order relation in $\mathcal{P}(Q)$ we consider the relation of inclusion $\subseteq$.

In the first example, for every $x, y \in \mathbb{R}$ we have either $x \leq y$ or $y \leq x$. In the second example, though, we have neither $2/3$ nor $3/2$. Similarly, in the third example, if Q contains at least two elements q_1, q_2, then the elements $\{q_1\}, \{q_2\}$ of $\mathcal{P}(Q)$ satisfy neither $\{q_1\} \subseteq \{q_2\}$ nor $\{q_2\} \subseteq \{q_1\}$.

Definition. *Let A be ordered by $\prec$, and $B \subseteq A$. Then we say that B is* **totally ordered** *if every $b_1, b_2 \in B$ satisfy either $b_1 \prec b_2$ or $b_2 \prec b_1$.*

Definition. *Let A be ordered by $\prec$, $B \subseteq A$, and $a \in A$. Then a is called* **upper bound** *of B, if $b \prec a$ for every $b \in B$.*

Definition. *Let A be ordered by $\prec$, and $a \in A$. Then a is called* **maximal element** *of A, if there is no $a' \in A$ such that $a \prec a'$ and $a \neq a'$.*

It is fairly standard to accept as an *axiom* the following statement.

Zorn's Lemma. *Let A be ordered by some order relation. If every totally ordered subset of A has an upper bound in A, then A has at least one maximal element.*

2.10 Orthonormal bases.

Definition. *Let X be an inner product space with inner product $\langle \cdot, \cdot \rangle$ and norm $\|\cdot\|$, and let $A \subseteq X$.*
(i) We say that A is **orthogonal** *if $a \neq 0$ for every $a \in A$ and $a_1 \perp a_2$ for every $a_1, a_2 \in A$, $a_1 \neq a_2$.*
(ii) We say that A is **orthonormal** *if $\|a\| = 1$ for every $a \in A$ and $a_1 \perp a_2$ for every $a_1, a_2 \in A$, $a_1 \neq a_2$.*

Of course, if A is orthonormal then it is orthogonal. Also, if A is orthogonal, then the set $A' = \{\frac{1}{\|a\|}\, a \,|\, a \in A\}$ is orthonormal.

Proposition 2.13. *Let X be an inner product space, and $A \subseteq X$. If A is orthogonal, then it is linearly independent.*

Proof. Assume that $a_1, \ldots, a_n \in A$ and $\lambda_1, \ldots, \lambda_n \in F$ so that

$$\lambda_1 a_1 + \cdots + \lambda_n a_n = 0.$$

If we take the inner product of both sides with a_k we find $\lambda_k = 0$. □

Definition. *Let X be an inner product space, and $A \subseteq X$.*
(i) We say that A is a **maximal orthonormal set** *of X, if A is orthonormal and there is no orthonormal set A' so that $A \subsetneqq A'$.*
(ii) We say that A is an **orthonormal basis** *of X, if A is orthonormal and* $\operatorname{clspan}(A) = X$.

It is easy to see that A is a maximal orthonormal set if and only if it is orthonormal and there is no $x \neq 0$ so that $x \perp A$. Also, A is an orthonormal basis if and only if it is orthonormal and every x is the limit of linear combinations of elements of A.

Proposition 2.14. *Let X be an inner product space, and $A \subseteq X$.*
(i) If A is an orthonormal basis of X, then A is a maximal orthonormal set of X.
(ii) If A is a maximal orthonormal set of X and X is a Hilbert space, then A is an orthonormal basis of X.

Proof. (i) Assume that A is an orthonormal basis of X and let $x \perp A$. Then $x \perp \text{clspan}(A) = X$ and hence $x = 0$. So A is a maximal orthonormal set of X.
(ii) Assume that A is a maximal orthonormal set of X and take any $x \in X$. Since $\text{clspan}(A)$ is a closed subspace of the Hilbert space X, there are $y, z \in X$ so that $y \in \text{clspan}(A)$, $z \perp \text{clspan}(A)$ and $x = y + z$. From $z \perp \text{clspan}(A)$ we get $z \perp A$ and, since A is a maximal orthonormal set of X, we find $z = 0$. Therefore, $x = y \in \text{clspan}(A)$.
We conclude that $X = \text{clspan}(A)$, i.e. A is an orthonormal basis of X. □

Theorem 2.3. *Let $X \neq \{0\}$ be an inner product space.*
(i) There exists a maximal orthonormal set A in X.
(ii) If A_0 is any orthonormal set in X, then there exists a maximal orthonormal set A in X so that $A_0 \subseteq A$.
If X is a Hilbert space, the maximal orthonormal set A in (i-ii) is an orthonormal basis of X.

Proof. Let $\langle \cdot, \cdot \rangle$ and $\| \cdot \|$ be the inner product and the norm of X.
(i) We consider the collection $\mathcal{A}$ of all orthonormal sets of X. If $a \neq 0$ is any element of X, then $\{\frac{a}{\|a\|}\}$ is an element of $\mathcal{A}$ and so $\mathcal{A}$ is non-empty. We also consider $\mathcal{A}$ ordered by set inclusion.
Now, let $\mathcal{B}$ be any totally ordered subcollection of $\mathcal{A}$. We define

$$A = \bigcup_{B \in \mathcal{B}} B.$$

Every $a \in A$ belongs to some $B \in \mathcal{B}$ and so $\|a\| = 1$. Also, if $a_1, a_2 \in A$ and $a_1 \neq a_2$, then there are $B_1, B_2 \in \mathcal{B}$ so that $a_1 \in B_1$ and $a_2 \in B_2$. Since $\mathcal{B}$ is totally ordered, we have that either $B_1 \subseteq B_2$ or $B_2 \subseteq B_1$, and hence both a_1, a_2 belong to one of B_1, B_2. Thus $a_1 \perp a_2$. Therefore, $A \in \mathcal{A}$ and A is obviously an upper bound of $\mathcal{B}$.
Now, Zorn's lemma implies that $\mathcal{A}$ has a maximal element.
(ii) We consider the collection $\mathcal{A}$ of all orthonormal sets of X which contain A_0. Then A_0 is an element of $\mathcal{A}$ and so $\mathcal{A}$ is non-empty. Now we repeat the proof of (i). □

Bessel's inequality. *Let X be an inner product space with inner product $\langle \cdot, \cdot \rangle$ and norm $\| \cdot \|$, and let A be an orthonormal set in X. Then*

$$\sum_{a \in A} |\langle x, a \rangle|^2 \leq \|x\|^2$$

for every $x \in X$.

Proof. We take any finite $B \subseteq A$ and we consider the element $z = x - \sum_{a \in B} \langle x, a \rangle a$. Then for every $a' \in B$ we get

$$\langle z, a' \rangle = \langle x, a' \rangle - \sum_{a \in B} \langle x, a \rangle \langle a, a' \rangle = \langle x, a' \rangle - \langle x, a' \rangle = 0.$$

So $z \perp a'$ for every $a' \in B$ and hence $z \perp \sum_{a \in B} \langle x, a \rangle a$. This implies

$$\|x\|^2 = \Big\| z + \sum_{a \in B} \langle x, a \rangle a \Big\|^2 = \|z\|^2 + \Big\| \sum_{a \in B} \langle x, a \rangle a \Big\|^2 \geq \Big\| \sum_{a \in B} \langle x, a \rangle a \Big\|^2 = \sum_{a \in B} |\langle x, a \rangle|^2.$$

Since this holds for every finite $B \subseteq A$, we conclude that $\sum_{a \in A} |\langle x, a \rangle|^2 \leq \|x\|^2$. □

The theorem of F.Riesz and Fischer. *Let X be a Hilbert space with inner product $\langle \cdot, \cdot \rangle$ and norm $\| \cdot \|$, let A be an orthonormal set in X, and let $\lambda_a \in F$, $a \in A$. If $\sum_{a \in A} |\lambda_a|^2 < +\infty$, then the series $\sum_{a \in A} \lambda_a \, a$ converges unconditionally in X. If $x = \sum_{a \in A} \lambda_a \, a$ is the sum of the series, then $x \in \text{clspan}(A)$, and*
(i) $\langle x, a \rangle = \lambda_a$ for every $a \in A$,
(ii) $\|x\|^2 = \sum_{a \in A} |\lambda_a|^2$,
(iii) $\langle x, y \rangle = \sum_{a \in A} \lambda_a \, \overline{\langle y, a \rangle}$ for every $y \in X$.

Proof. Due to lemma 1.3, from $\sum_{a\in A}|\lambda_a|^2 < +\infty$ we get that the set $A_0 = \{a \in A \,|\, \lambda_a \neq 0\}$ is countable. Since the case of A_0 being finite is trivial, we assume that A_0 is infinite, and we consider any enumeration $\{a_1, a_2, \ldots\}$ of A_0. Then the set $A_n = \{a_1, \ldots, a_n\}$ is a finite subset of A and hence

$$\sum_{k=1}^{n}|\lambda_{a_k}|^2 = \sum_{a\in A_n}|\lambda_a|^2 \leq \sum_{a\in A}|\lambda_a|^2.$$

This is true for every n and so

$$\sum_{k=1}^{+\infty}|\lambda_{a_k}|^2 \leq \sum_{a\in A}|\lambda_a|^2 < +\infty.$$

We set $s_n = \sum_{k=1}^{n}\lambda_{a_k}a_k$ for every n. Then for every n, m with $n < m$ we get

$$\|s_m - s_n\|^2 = \Big\|\sum_{k=n+1}^{m}\lambda_{a_k}a_k\Big\|^2 = \sum_{k=n+1}^{m}|\lambda_{a_k}|^2 \to 0$$

when $n, m \to +\infty$. Since X is a Hilbert space, there is $x \in X$ so that $s_n \to x$, i.e.

$$x = \sum_{k=1}^{+\infty}\lambda_{a_k}a_k.$$

Obviously, $s_n \in \text{span}(A)$ for every n, and so $x \in \text{clspan}(A)$. Moreover, for every $a \in A$ we have

$$\langle x, a\rangle = \lim_{n\to+\infty}\langle s_n, a\rangle = \lim_{n\to+\infty}\sum_{k=1}^{n}\lambda_{a_k}\langle a_k, a\rangle = \sum_{k=1}^{+\infty}\lambda_{a_k}\langle a_k, a\rangle = \lambda_a.$$

If we consider any other enumeration $\{a_1', a_2', \ldots\}$ of A_0, then again we have $x' = \sum_{k=1}^{+\infty}\lambda_{a_k'}a_k'$ for some $x' \in \text{clspan}(A)$, satisfying $\langle x', a\rangle = \lambda_a$ for every $a \in A$. Then

$$\langle x' - x, a\rangle = \langle x', a\rangle - \langle x, a\rangle = \lambda_a - \lambda_a = 0$$

for every $a \in A$. Thus $x' - x \perp A$ and so $x' - x \perp \text{clspan}(A)$. Since $x' - x \in \text{clspan}(A)$, we conclude that $x' - x = 0$, i.e. $x' = x$, and so the sum of the series $\sum_{k=1}^{+\infty}\lambda_{a_k}a_k$ does not depend on the enumeration of A_0. Thus, the series $\sum_{a\in A}\lambda_a\, a$ converges unconditionally in X and

$$\sum_{a\in A}\lambda_a a = x = \sum_{k=1}^{+\infty}\lambda_{a_k}a_k.$$

Now, for every $y \in X$ we get

$$\langle x, y\rangle = \lim_{n\to+\infty}\langle s_n, y\rangle = \lim_{n\to+\infty}\sum_{k=1}^{n}\lambda_{a_k}\langle a_k, y\rangle = \sum_{k=1}^{+\infty}\lambda_{a_k}\langle a_k, y\rangle = \sum_{k=1}^{+\infty}\lambda_{a_k}\,\overline{\langle y, a_k\rangle}.$$

Since the sum $\langle x, y\rangle$ of $\sum_{k=1}^{+\infty}\lambda_{a_k}\,\overline{\langle y, a_k\rangle}$ does not depend on the enumeration of A_0, we get that the series $\sum_{a\in A}\lambda_a\,\overline{\langle y, a\rangle}$ converges unconditionally, and

$$\langle x, y\rangle = \sum_{a\in A}\lambda_a\,\overline{\langle y, a\rangle}.$$

This is the equality of (iii) and, setting $y = x$, we get the equality of (ii). $\square$

Definition. *Let X be an inner product space with inner product $\langle\cdot,\cdot\rangle$, let A be an orthonormal set in X and $x \in X$. The numbers $\langle x, a\rangle$, $a \in A$, are called* **Fourier coefficients** *of x with respect to A, and the series $\sum_{a\in A}\langle x, a\rangle a$ is called* **Fourier series** *of x with respect to A.*

Theorem 2.4. *Let X be a Hilbert space with inner product $\langle\cdot,\cdot\rangle$ and norm $\|\cdot\|$, and let A be an orthonormal basis of X. Then the Fourier series $\sum_{a\in A}\langle x, a\rangle\, a$ of every $x \in X$ converges unconditionally in X and its sum is x, i.e.*

$$\sum_{a\in A}\langle x, a\rangle\, a = x.$$

Also,
(i) $\|x\|^2 = \sum_{a\in A}|\langle x, a\rangle|^2$,
(ii) $\langle x, y\rangle = \sum_{a\in A}\langle x, a\rangle\,\overline{\langle y, a\rangle}$ for every $y \in X$.
The last two equalities are called **Parseval's identities***.*

Proof. Bessel's inequality and then the theorem of F.Riesz and Fischer imply that $\sum_{a\in A}\langle x, a\rangle\, a$ converges unconditionally in X. If

$$x' = \sum_{a\in A}\langle x, a\rangle\, a$$

is the sum of the series, then $\langle x', a\rangle = \langle x, a\rangle$ for every $a \in A$. So $x' - x \perp A$ and, since A is a maximal orthonormal set, we get $x' = x$. Thus $\sum_{a\in A}\langle x, a\rangle\, a = x$, and then we get (i),(ii) from the theorem of F.Riesz and Fischer. $\square$

It is worth seeing that the two Parseval's identities are equivalent. Indeed, if (ii) holds for every $x, y \in X$, then, setting $y = x$, we see that (i) holds for every $x \in X$. Conversely, assume that (i) holds for every $x \in X$. Then it holds for $x, y, x + y$, i.e.

$$\|x\|^2 = \sum_{a\in A}|\langle x, a\rangle|^2, \quad \|y\|^2 = \sum_{a\in A}|\langle y, a\rangle|^2, \quad \|x+y\|^2 = \sum_{a\in A}|\langle x+y, a\rangle|^2.$$

The third equality implies

$$\|x\|^2 + 2\,\mathrm{Re}(\langle x, y\rangle) + \|y\|^2 = \sum_{a\in A}|\langle x, a\rangle|^2 + 2\sum_{a\in A}\mathrm{Re}(\langle x, a\rangle\,\overline{\langle y, a\rangle}) + \sum_{a\in A}|\langle y, a\rangle|^2.$$

Therefore,

$$\mathrm{Re}(\langle x, y\rangle) = \sum_{a\in A}\mathrm{Re}(\langle x, a\rangle\,\overline{\langle y, a\rangle})$$

for every $x, y \in X$. Now, if $F = \mathbb{R}$, then we have got (ii). If $F = \mathbb{C}$, then we replace y with iy and we get

$$\mathrm{Im}(\langle x, y\rangle) = \sum_{a\in A}\mathrm{Im}(\langle x, a\rangle\,\overline{\langle y, a\rangle})$$

for every $x, y \in X$. From the last two equalities we get (ii).

Example 2.10.1. In the space l^2 we consider the elements

$$e_n = (0, \ldots, 0, 1, 0, \ldots), \quad n \in \mathbb{N},$$

where e_n has all its coordinates equal to 0 except for the n-th coefficient which is equal to 1. It is trivial to see that

$$\langle e_n, e_m\rangle = \begin{cases} 1, & n = m, \\ 0, & n \neq m. \end{cases}$$

So the set $A = \{e_n \,|\, n \in \mathbb{N}\}$ is orthonormal in l^2.
If $x = (\lambda_n) \in l^2$ is orthogonal to A, then we have

$$\lambda_n = \langle x, e_n \rangle = 0$$

for every n and hence $x = 0$. Thus, A is a maximal orthonormal set in l^2 and, since l^2 is a Hilbert space, A is an orthonormal basis of l^2. So for every $x = (\lambda_n) \in l^2$ we can write

$$x = \sum_{n=1}^{+\infty} \langle x, e_n \rangle \, e_n = \sum_{n=1}^{+\infty} \lambda_n e_n.$$

Also, Parseval's identities for this particular orthonormal basis $A = \{e_n \,|\, n \in \mathbb{N}\}$ of l^2 take, for every $x = (\lambda_n), y = (\kappa_n) \in l^2$, the form

$$\|x\|^2 = \sum_{n=1}^{+\infty} |\langle x, e_n \rangle|^2 = \sum_{n=1}^{+\infty} |\lambda_n|^2,$$

$$\langle x, y \rangle = \sum_{n=1}^{+\infty} \langle x, e_n \rangle \, \overline{\langle y, e_n \rangle} = \sum_{n=1}^{+\infty} \lambda_n \, \overline{\kappa_n}.$$

In fact these identities are just the defining equalities for the norm and the inner product of l^2.

Example 2.10.2. In the space $L^2([0,1])$ with the Lebesgue measure of $[0,1]$, we consider the elements

$$e_n(t) = e^{2\pi i n t}, \quad n \in \mathbb{Z}.$$

Then we have

$$\langle e_n, e_m \rangle = \int_0^1 e_n(t) \, \overline{e_m(t)} \, dt = \int_0^1 e^{2\pi i (n-m) t} \, dt = \begin{cases} 1, & n = m, \\ 0, & n \neq m. \end{cases}$$

Therefore, the set $A = \{e_n \,|\, n \in \mathbb{Z}\}$ is orthonormal in $L^2([0,1])$.
If $f \in L^2([0,1])$, then the Fourier coefficient of f with respect to every e_n is denoted $\widehat{f}(n)$ and it is equal to

$$\widehat{f}(n) = \langle f, e_n \rangle = \int_0^1 f(t) \, \overline{e_n(t)} \, dt = \int_0^1 f(t) e^{-2\pi i n t} \, dt, \quad n \in \mathbb{Z}.$$

It is known that $A = \{e_n \,|\, n \in \mathbb{Z}\}$ is an orthonormal basis of $L^2([0,1])$. So every $f \in L^2([0,1])$ is equal to its Fourier series with respect to A, i.e.

$$f = \sum_{n \in \mathbb{Z}} \langle f, e_n \rangle \, e_n = \sum_{n \in \mathbb{Z}} \widehat{f}(n) \, e_n.$$

Also, Parseval's identities take, for every $f, g \in L^2([0,1])$, the form

$$\int_0^1 |f(t)|^2 \, dt = \|f\|^2 = \sum_{n \in \mathbb{Z}} |\langle f, e_n \rangle|^2 = \sum_{n \in \mathbb{Z}} |\widehat{f}(n)|^2,$$

$$\int_0^1 f(t) \, \overline{g(t)} \, dt = \langle f, g \rangle = \sum_{n \in \mathbb{Z}} \langle f, e_n \rangle \, \overline{\langle g, e_n \rangle} = \sum_{n \in \mathbb{Z}} \widehat{f}(n) \, \overline{\widehat{g}(n)}.$$

2.11 Orthogonal projections.

Definition. *Let X be an inner product space, and let Y be a subspace of X with an orthogonal complement in X. Then we know from proposition 2.11 that $Y + Y^\perp = X$, that Y is closed, that its orthogonal complement is $Y^\perp$ and that Y is the orthogonal complement of $Y^\perp$ in X.*
We consider the function

$$P_Y : X \to X$$

defined for every $x \in X$ by

$$P_Y(x) = y,$$

where $x = y + z$, with $y \in Y$ and $z \in Y^\perp$.
The function P_Y is called **orthogonal projection** *of X onto Y.*

Proposition 2.15. *Let X be an inner product space with inner product $\langle \cdot, \cdot \rangle$ and norm $\| \cdot \|$, and let Y be a subspace of X with an orthogonal complement in X. Then the orthogonal projection $P_Y : X \to X$ has the following properties.*
(i) P_Y is linear.
(ii) The range of P_Y is Y, i.e. $\mathrm{R}(P_Y) = Y$, and its null space is $Y^\perp$, i.e. $\mathrm{N}(P_Y) = Y^\perp$.
(iii) $P_Y \circ P_Y = P_Y$.
(iv) $\langle P_Y(x_1), x_2 \rangle = \langle x_1, P_Y(x_2) \rangle$ for every $x_1, x_2 \in X$.
(v) $\|P_Y(x)\| \leq \|x\|$ for every $x \in X$.

Proof. (i) Take $x_1, x_2 \in X$. Then there are $y_1, y_2 \in Y$ and $z_1, z_2 \in Y^\perp$ so that $x_1 = y_1 + z_1$ and $x_2 = y_2 + z_2$. Now, $y_1 + y_2 \in Y$ and $z_1 + z_2 \in Y^\perp$, and $x_1 + x_2 = (y_1 + y_2) + (z_1 + z_2)$. Therefore,

$$P_Y(x_1 + x_2) = y_1 + y_2 = P_Y(x_1) + P_Y(x_2).$$

Similarly, take $x \in X$ and $\lambda \in F$. Then there are $y \in Y$ and $z \in Y^\perp$ so that $x = y + z$. Now, $\lambda y \in Y$ and $\lambda z \in Y^\perp$, and $\lambda x = \lambda y + \lambda z$. Therefore,

$$P_Y(\lambda x) = \lambda y = \lambda P_Y(x).$$

(ii) It is clear that $\mathrm{R}(P_Y) \subseteq Y$. Now, take any $y \in Y$. Then $y = y + 0$ and $y \in Y$, $0 \in Y^\perp$. So $P_Y(y) = y$ and hence $y \in \mathrm{R}(P_Y)$. Therefore $Y \subseteq \mathrm{R}(P_Y)$.
Take any $z \in Y^\perp$. Then $z = 0 + z$ and $0 \in Y$, $z \in Y^\perp$. So $P_Y(z) = 0$ and hence $z \in \mathrm{N}(P_Y)$. Therefore, $Y^\perp \subseteq \mathrm{N}(P_Y)$.
Conversely, let $x \in \mathrm{N}(P_Y)$, i.e. $P_Y(x) = 0$. Then $x = 0 + z$ and $z \in Y^\perp$ and hence $x \in Y^\perp$. Therefore, $\mathrm{N}(P_Y) \subseteq Y^\perp$.
(iii) We saw in the proof of (ii) that $P_Y(y) = y$ for every $y \in Y$. Now, for any $x \in X$ we have that $P_Y(x) \in Y$ and hence $P_Y(P_Y(x)) = P_Y(x)$.
(iv) Take $x_1, x_2 \in X$. Then there are $y_1, y_2 \in Y$ and $z_1, z_2 \in Y^\perp$ so that $x_1 = y_1 + z_1$ and $x_2 = y_2 + z_2$. Now,

$$\langle P_Y(x_1), x_2 \rangle = \langle y_1, y_2 + z_2 \rangle = \langle y_1, y_2 \rangle = \langle y_1 + z_1, y_2 \rangle = \langle x_1, P_Y(x_2) \rangle.$$

(v) Take any $x \in X$. Then there are $y \in Y$ and $z \in Y^\perp$ so that $x = y + z$. Then

$$\|P_Y(x)\|^2 = \|y\|^2 \leq \|y\|^2 + \|z\|^2 = \|y + z\|^2 = \|x\|^2.$$

□

It is clear from the proof of (ii) of proposition 2.15 that if we restrict P_Y on Y then it is equal to the identity operator of Y:

$$P_Y(y) = y, \quad y \in Y.$$

Also, if we restrict P_Y on $Y^\perp$ then it is equal to the null operator of $Y^\perp$:

$$P_Y(z) = 0, \quad z \in Y^\perp.$$

Note that an orthogonal projection P_Y corresponds to a subspace Y which has an orthogonal complement in X (i.e. $Y^\perp$). In this case $Y^\perp$ also has an orthogonal complement in X (i.e. Y) and so the orthogonal projection $P_{Y^\perp}$ is also defined. Proposition 2.15 describes the properties of any orthogonal projection and hence of $P_{Y^\perp}$. The following proposition describes some extra properties of the pair of orthogonal projections P_Y and $P_{Y^\perp}$.

Proposition 2.16. *Let X be an inner product space, and let Y be a subspace of X with an orthogonal complement in X. Then:*
(i) $P_Y + P_{Y^\perp} = I$, the identity operator of X.
(ii) $P_{Y^\perp} \circ P_Y = P_Y \circ P_{Y^\perp} = 0$, the null operator of X.

Proof. Take any $x \in X$. Then there are $y \in Y$ and $z \in Y^\perp$ so that $x = y + z$.
(i) $P_Y(x) + P_{Y^\perp}(x) = y + z = x$.
(ii) $P_{Y^\perp}(P_Y(x)) = P_{Y^\perp}(y) = 0$ and $P_Y(P_{Y^\perp}(x)) = P_Y(z) = 0$. □

The following proposition describes the properties which characterize orthogonal projections among linear operators on an inner product space.

Proposition 2.17. *Let X be an inner product space with inner product $\langle \cdot, \cdot \rangle$, and let $P : X \to X$ be a linear operator. If $P \circ P = P$ and $\langle P(x_1), x_2 \rangle = \langle x_1, P(x_2) \rangle$ for every $x_1, x_2 \in X$, then there is a subspace Y of X, which has an orthogonal complement in X, so that $P = P_Y$.*

Proof. We consider the linear subspaces $Y = \mathrm{R}(P)$ and $Z = \mathrm{N}(P)$ of X.
Clearly, $P(z) = 0$ for every $z \in Z$. Also, if $y \in Y$, then $y = P(x)$ for some $x \in X$ and so

$$P(y) = P(P(x)) = P(x) = y.$$

Hence, for every $y \in Y$ and every $z \in Z$ we have

$$\langle y, z \rangle = \langle P(y), z \rangle = \langle y, P(z) \rangle = \langle y, 0 \rangle = 0,$$

and so $Y \perp Z$.
Now take any $x \in X$ and consider $y = P(x)$ and $z = x - P(x)$. Then $y \in Y$ and

$$P(z) = P(x) - P(P(x)) = P(x) - P(x) = 0,$$

i.e. $z \in Z$. Obviously: $x = y + z$ and we conclude that $X = Y + Z$ and $Y \perp Z$. Therefore, Y, Z are orthogonal complements of each other.
We just saw that for any $x \in X$ we have $x = P(x) + z$, where $P(x) \in Y$ and $z \in Z$. Hence $P_Y(x) = P(x)$. □

We know from theorem 2.2 that every complete subspace of an inner product space has an orthogonal complement and so defines a corresponding othogonal projection. Also, theorem 2.3 implies that every complete subspace of an inner product space has an orthonormal basis. Now we shall describe the orthogonal projection on a complete subspace in terms of an orthonormal basis of the subspace.

Proposition 2.18. *Let X be an inner product space with inner product $\langle \cdot, \cdot \rangle$, let Y be a complete subspace of X, and let A be any orthonormal basis of Y. Then for every $x \in X$ we have*

$$P_Y(x) = \sum_{a \in A} \langle x, a \rangle \, a.$$

Proof. Take any $x \in X$. Then there are $y \in Y$ and $z \in Y^\perp$ so that $x = y + z$.
Since $a \in Y$ for every $a \in A$, we have that

$$\langle x, a \rangle = \langle y + z, a \rangle = \langle y, a \rangle + \langle z, a \rangle = \langle y, a \rangle$$

for every $a \in A$. Hence

$$P_Y(x) = y = \sum_{a \in A} \langle y, a \rangle \, a = \sum_{a \in A} \langle x, a \rangle \, a.$$

□

2.12 Separable inner product spaces.

The theorem of Schmidt. *Let X be a separable inner product space with* $\dim(X) = +\infty$.
(i) Every orthonormal basis of X is countable and infinite.
(ii) X has an orthonormal basis.
(iii) If X is complete, then X and l^2 are linearly isometric.

Proof. Let M be a countable and dense subset of X.
(i) If A is any orthonormal basis of X, then for every $a \in A$ we consider the ball $B(a; \frac{\sqrt{2}}{2})$ and we observe that these balls are pairwise disjoint.
Now, for every $a \in A$ there is $x_a \in M \cap B(a; \frac{\sqrt{2}}{2})$, and we may consider the function $A \ni a \mapsto x_a \in M$. This function is one-to-one in A, and so A is countable.
If A is finite, i.e. $A = \{a_1, \dots, a_n\}$, then $X = \text{clspan}(\{a_1, \dots, a_n\}) = \text{span}(\{a_1, \dots, a_n\})$ has finite dimension. (We used that any subspace of finite dimension is closed.)
(ii) Now, let $M = \{x_1, x_2, x_3, \dots\}$.
Let n_1 be the least natural number so that $x_{n_1} \neq 0$. Then let n_2 be the least natural number so that x_{n_2} is not a multiple of x_{n_1}. We continue inductively: if we have found $n_1, \dots, n_{k-1}$, we let n_k be the natural number so that x_{n_k} is not a linear combination of $x_{n_1}, \dots, x_{n_{k-1}}$. If this process stops at some point, then there is N so that all $x_{N+1}, x_{N+2}, \dots$ are linear combinations of $x_1, \dots, x_N$. But then $X = \text{clspan}(\{x_1, \dots, x_N\}) = \text{span}(\{x_1, \dots, x_N\})$ and so X is finite dimensional. Thus, the above process does not end, and so we get the countable and infinite set

$$N = \{x_{n_1}, x_{n_2}, \dots\} \subseteq M.$$

Since every x_{n_k} is not a linear combination of $x_{n_1}, \dots, x_{n_{k-1}}$, the set N is linearly independent.
Now, take any $x \in X$ and any $\epsilon > 0$. Then there is $x_j \in M$ so that $\|x - x_j\| < \epsilon$. Then there is k so that $j < n_k$, and this implies that x_j is a linear combination of $x_{n_1}, \dots, x_{n_{k-1}}$. Therefore, $X = \text{clspan}(N)$.
For simplicity, we denote $y_k = x_{n_k}$, i.e.

$$N = \{y_1, y_2, \dots\}.$$

We define

$$a_1 = \frac{1}{\|y_1\|} \, y_1,$$

and then

$$\{a_1\} \text{ is orthonormal}, \quad \text{span}(\{a_1\}) = \text{span}(\{y_1\}).$$

Now assume that we have defined $a_1, \dots, a_{k-1}$ so that

$$\{a_1, \dots, a_{k-1}\} \text{ is orthonormal}, \quad \text{span}(\{a_1, \dots, a_{k-1}\}) = \text{span}(\{y_1, \dots, y_{k-1}\}).$$

We denote

$$M_{k-1} = \text{span}(\{a_1, \ldots, a_{k-1}\}) = \text{span}(\{y_1, \ldots, y_{k-1}\}).$$

Then $y_k \notin M_{k-1}$, and so $y_k - P_{M_{k-1}}(y_k) \neq 0$, where $P_{M_{k-1}}$ is the orthogonal projection on the finite dimensional subspace M_{k-1}. We define

$$a_k = \frac{1}{\|y_k - P_{M_{k-1}}(y_k)\|} \left(y_k - P_{M_{k-1}}(y_k)\right)$$

and then a_k is orthogonal to M_{k-1} with $\|a_k\| = 1$. Moreover, a_k is a linear combination of $y_1, \ldots, y_k$ and also y_k is a linear combination of $a_1, \ldots, a_k$. Thus

$$\{a_1, \ldots, a_{k-1}, a_k\} \text{ is orthonormal}, \quad \text{span}(\{a_1, \ldots, a_{k-1}, a_k\}) = \text{span}(\{y_1, \ldots, y_{k-1}, y_k\}).$$

Continuing inductively, we construct the set

$$A = \{a_1, a_2, \ldots\},$$

which is orthonormal and satisfies:

$$\text{clspan}(A) = \text{clspan}(N) = X.$$

Therefore, A is an orthonormal basis of X.
(iii) Let $A = \{a_1, a_2, \ldots\}$ be any orthonormal basis of X. If $x \in X$, then

$$\sum_{k=1}^{+\infty} |\langle x, a_k\rangle|^2 = \|x\|^2 < +\infty,$$

and so we may consider the function $T : X \to l^2$ defined for every $x \in X$ by

$$T(x) = \big(\langle x, a_1\rangle, \langle x, a_2\rangle, \ldots\big).$$

It is easy to see that T is linear. Also, T is a linear isometry, since

$$\|T(x)\|_2^2 = \sum_{k=1}^{+\infty} |\langle x, a_k\rangle|^2 = \|x\|^2$$

for every $x \in X$. If $(\lambda_k) \in l^2$, then there is $x \in X$ so that $\langle x, a_k\rangle = \lambda_k$ for every k. Thus $T(x) = (\lambda_k)$ and so T is onto l^2.

$\square$

This theorem is useful, because many classical Hilbert spaces are separable. For example $L^2(\Omega, \mathcal{B}(\Omega), m)$ with a Borel set $\Omega \subseteq \mathbb{R}^d$ and the Lebesgue measure m.

Chapter 3

The dual of a normed space

3.1 Bounded linear functionals.

Definition. *Let X be a normed space with norm $\|\cdot\|$, and let $l : X \to F$ be a linear functional on X. Then l is called* **bounded** *if there is $C \geq 0$ so that*

$$|l(x)| \leq C\|x\|$$

for every $x \in X$.

Proposition 3.1. *Let X be a normed space with norm $\|\cdot\|$, and let $l : X \to F$ be a linear functional on X. The following are equivalent:*
(i) l is continuous in X.
(ii) $\mathrm{N}(l)$ is closed in X.
(iii) l is bounded.
(iv) l is continuous at $0 \in X$.

Proof. $\mathrm{N}(l) = l^{-1}(\{0\})$ is the inverse image of a closed set, and so, if l is continuous in X, then $\mathrm{N}(l)$ is closed in X.
Now, assume that $\mathrm{N}(l)$ is closed in X. If $l = 0$, then l is obviously bounded. So assume that $l \neq 0$. Then there is $x_0 \in X$ so that $l(x_0) = 1$. Since $\mathrm{N}(l)$ is closed and $x_0 \notin \mathrm{N}(l)$, there is $r > 0$ so that $B(x_0; r) \cap \mathrm{N}(l) = \emptyset$. Now take any $x \in X$ with $l(x) \neq 0$. Then

$$l\Big(x_0 - \frac{x}{l(x)}\Big) = l(x_0) - \frac{l(x)}{l(x)} = 1 - 1 = 0.$$

Hence $x_0 - \frac{x}{l(x)} \in \mathrm{N}(l)$ and so $x_0 - \frac{x}{l(x)} \notin B(x_0; r)$. Thus $\|\frac{x}{l(x)}\| \geq r$, i.e. $|l(x)| \leq \frac{1}{r}\,\|x\|$. This is obviously true when $l(x) = 0$, and we conclude that

$$|l(x)| \leq \frac{1}{r}\,\|x\|$$

for every $x \in X$. Therefore, l is bounded.
If l is bounded, then there is $C \geq 0$ so that $|l(x)| \leq C\|x\|$ for every $x \in X$. If $x_n \to 0$ in X, then

$$|l(x_n)| \leq C\|x_n\| \to 0,$$

and so $l(x_n) \to 0$ in $\mathbb{R}$. Hence l is continuous at 0.
Finally, assume that l is continuous at 0. If $x_n \to x$ in X, then $x_n - x \to 0$ in X, and then $l(x_n) - l(x) = l(x_n - x) \to 0$ in $\mathbb{R}$, and then $l(x_n) \to l(x)$ in $\mathbb{R}$. So l is continuous in X. $\square$

Definition. *Let X be a normed space. The set of all continuous or, equivalently, bounded linear functionals on X is called* **dual space** *of X, and it is denoted X'.*

Proposition 3.2. *Let X be a normed space with norm $\|\cdot\|$. Then X' as a function space, with the usual addition of functions and the usual multiplication of numbers and functions, is a linear space.*

Proof. If $l, l_1, l_2 : X \to F$ and $\lambda \in F$, we consider the functions $l_1+l_2 : X \to F$ and $\lambda l : X \to F$ defined for every $x \in X$ by

$$(l_1 + l_2)(x) = l_1(x) + l_2(x), \quad (\lambda l)(x) = \lambda l(x).$$

It is known from Linear Algebra (and it is very easy to prove) that, if l, l_1, l_2 are linear functionals, then $l_1 + l_2$ and λl are also linear functionals. It is also clear that, if l, l_1, l_2 are continuous, then $l_1 + l_2$ and λl are also continuous. $\square$

Usually we denote the elements of X' with symbols like x', y' etc.

Definition. *Let X be a normed space with norm $\|\cdot\|$. For every $x' \in X'$ we define*

$$\|x'\| = \sup_{x\in X, \|x\|\leq 1} |x'(x)|.$$

Proposition 3.3. *Let X be a normed space with norm $\|\cdot\|$ and let $x' \in X'$. Then $\|x'\|$ is the smallest constant C which satifies the inequality $|x'(x)| \leq C\|x\|$ for every $x \in X$.*

Proof. For every $x \in X$, $x \neq 0$, we have $\left\|\frac{x}{\|x\|}\right\| = 1$, and then, by the definition of $\|x'\|$ we get

$$|x'(x)| = \left|x'\Big(\frac{x}{\|x\|}\Big)\right|\|x\| \leq \|x'\|\|x\|.$$

The inequality $|x'(x)| \leq \|x'\|\|x\|$ is obviously satistied if $x = 0$, and so $C = \|x'\|$ satisfies the inequality $|x'(x)| \leq C\|x\|$ for every $x \in X$.
Conversely, let C satisfy the inequality $|x'(x)| \leq C\|x\|$ for every $x \in X$. Then we have $|x'(x)| \leq C$ for every $x \in X$ with $\|x\| \leq 1$, and so $\|x'\| \leq C$. $\square$

So, if $x' \in X'$, then

$$|x'(x)| \leq \|x'\|\|x\| \quad \text{for every } x \in X.$$

Also,

$$|x'(x)| \leq C\|x\| \text{ for every } x \in X \quad \Rightarrow \quad \|x'\| \leq C.$$

Proposition 3.4. *Let X be a normed space with norm $\|\cdot\|$. The function $\|\cdot\| : X' \to \mathbb{R}$ defined above is a norm on X', and X' with this norm is a Banach space.*

Proof. Obviously, $\|x'\| \geq 0$ for every $x' \in X'$. It is also clear that $\|x'\| = 0$ if $x' = 0$.
If $x' \in X'$ and $\|x'\| = 0$, then $x'(x) = 0$ for every $x \in X$, and so $x' = 0$.
For every $x \in X$ and every $x_1', x_2' \in X'$ we have

$$|(x_1' + x_2')(x)| \leq |x_1'(x)| + |x_2'(x)| \leq \|x_1'\|\|x\| + \|x_2'\|\|x\| = (\|x_1'\| + \|x_2'\|)\|x\|.$$

Hence $\|x_1' + x_2'\| \leq \|x_1'\| + \|x_2'\|$.
For every $x' \in X'$ and every $\lambda \in F$ we have

$$\|\lambda x'\| = \sup_{x\in X, \|x\|\leq 1} |(\lambda x')(x)| = \sup_{x\in X, \|x\|\leq 1} |\lambda||x'(x)| = |\lambda| \sup_{x\in X, \|x\|\leq 1} |x'(x)| = |\lambda|\|x'\|.$$

Therefore, $\|\cdot\| : X' \to \mathbb{R}$ is a norm on X'.
Now take a sequence (x_n') in X' so that $\|x_n' - x_m'\| \to 0$ when $n, m \to +\infty$. For every $x \in X$ we have

$$|x_n'(x) - x_m'(x)| = |(x_n' - x_m')(x)| \leq \|x_n' - x_m'\|\|x\| \to 0$$

when $n, m \to +\infty$, and so $(x'_n(x))$ is a Cauchy sequence in F.
We consider the function $x' : X \to F$ defined for every $x \in X$ by

$$x'(x) = \lim_{n \to +\infty} x'_n(x).$$

Since each x'_n is a linear functional, we have for every $x, y \in X$ and $\lambda \in F$ that

$$x'(x + y) = \lim_{n \to +\infty} x'_n(x + y) = \lim_{n \to +\infty} x'_n(x) + \lim_{n \to +\infty} x'_n(y) = x'(x) + x'(y),$$

$$x'(\lambda x) = \lim_{n \to +\infty} x'_n(\lambda x) = \lambda \lim_{n \to +\infty} x'_n(x) = \lambda x'(x).$$

So x' is a linear functional on X.
Now, there is n_0 so that $\|x'_n - x'_m\| \leq 1$ for every $n, m \geq n_0$. Hence

$$|x'_n(x)| \leq |x'_n(x) - x'_{n_0}(x)| + |x'_{n_0}(x)| \leq \|x'_n - x'_{n_0}\|\|x\| + \|x'_{n_0}\|\|x\| \leq (1 + \|x'_{n_0}\|)\|x\|$$

for every $n \geq n_0$ and every $x \in X$. Taking the limit when $n \to +\infty$, we find

$$|x'(x)| \leq (1 + \|x'_{n_0}\|)\|x\|$$

for every $x \in X$. So x' is bounded, i.e. $x' \in X'$.
Finally, we take any $\epsilon > 0$ and then there is n_0 so that $\|x'_n - x'_m\| \leq \epsilon$ for every $n, m \geq n_0$. Then

$$|x'_n(x) - x'_m(x)| \leq \|x'_n - x'_m\|\|x\| \leq \epsilon\|x\|$$

for every $n, m \geq n_0$ and every $x \in X$. Taking the limit when $m \to +\infty$, we find

$$|x'_n(x) - x'(x)| \leq \epsilon\|x\|$$

for every $n \geq n_0$ and every $x \in X$. Therefore, $\|x'_n - x'\| \leq \epsilon$ for every $n \geq n_0$, and so $x'_n \to x'$ in X'. □

3.2 Finite dimensional spaces.

Theorem 3.1. *Let X be a finite dimensional normed space. Then X' is also finite dimensional with the same dimension as X.*

Proof. Let $\{b_1, \ldots, b_n\}$ be a basis of X. Since all norms on X are equivalent, a linear functional on X is continuous or not independently of the norm we are considering on X. So we may consider X equipped with its 2-norm with respect to the basis $\{b_1, \ldots, b_n\}$, i.e.

$$\|x\|_2 = (|\lambda_1|^2 + \cdots + |\lambda_n|^2)^{1/2}$$

for every $x = \lambda_1 b_1 + \cdots + \lambda_n b_n$ in X.
Now, we take any $z = \mu_1 b_1 + \cdots + \mu_n b_n$ in X, and we consider the function $l_z : X \to F$ defined for every $x = \lambda_1 b_1 + \cdots + \lambda_n b_n$ in X by

$$l_z(x) = \mu_1 \lambda_1 + \cdots + \mu_n \lambda_n.$$

It is very easy to show that l_z is a linear functional on X. We also have

$$|l_z(x)| \leq \|z\|_2 \|x\|_2$$

for every $x \in X$ and hence $l_z \in X'$ with $\|l_z\| \leq \|z\|_2$.
Now we consider the particular $x = \overline{\mu_1}\, b_1 + \cdots + \overline{\mu_n}\, b_n$ in X, and we get

$$\|z\|_2^2 = |\mu_1 \overline{\mu_1} + \cdots + \mu_n \overline{\mu_n}| = |l_z(x)| \leq \|l_z\|\|x\|_2 = \|l_z\|\|z\|_2,$$

and so $\|z\|_2 \leq \|l_z\|$. Therefore,

$$\|l_z\| = \|z\|_2$$

for every $z \in X$. Now we consider the function $T : X \to X'$ defined for every $z \in X$ by

$$T(z) = l_z.$$

If $z = \mu_1 b_1 + \cdots + \mu_n b_n$ and $w = \nu_1 b_1 + \cdots + \nu_n b_n$, then

$$l_{z+w}(x) = (\mu_1 + \nu_1)\lambda_1 + \cdots + (\mu_n + \nu_n)\lambda_n = l_z(x) + l_w(x)$$

for every $x = \lambda_1 b_1 + \cdots + \lambda_n b_n$ in X. Thus $l_{z+w} = l_z + l_w$, i.e. $T(z+w) = T(z) + T(w)$.
If $z = \mu_1 b_1 + \cdots + \mu_n b_n$ and $\mu \in F$, then

$$l_{\mu z}(x) = (\mu\mu_1)\lambda_1 + \cdots + (\mu\mu_n)\lambda_n = \mu l_z(x)$$

for every $x = \lambda_1 b_1 + \cdots + \lambda_n b_n$ in X. Thus $l_{\mu z} = \mu l_z$, i.e. $T(\mu z) = \mu T(z)$.
We conclude that $T : X \to X'$ is a linear operator.
We have already proven that $\|T(z)\| = \|l_z\| = \|z\|_2$ for every $z \in X$, and so T is a linear isometry of X into X'. Now we shall prove that T is onto X', i.e. that X and X' are linearly isometric.
We take any $l \in X'$ and we define

$$z = l(b_1)b_1 + \cdots + l(b_n)b_n \in X.$$

Then for every $x = \lambda_1 b_1 + \cdots + \lambda_n b_n$ in X we have

$$l_z(x) = l(b_1)\lambda_1 + \cdots + l(b_n)\lambda_n = l(\lambda_1 b_1 + \cdots + \lambda_n b_n) = l(x),$$

and hence $T(z) = l_z = l$. Therefore, T is onto X'. □

If $\{b_1, \ldots, b_n\}$ is the basis of X and $T : X \to X'$ is the linear isometry which appeared in the proof of theorem 3.1, we may define

$$b'_j = T(b_j) = l_{b_j}, \quad j = 1, \ldots, n.$$

Then $\{b'_1, \ldots, b'_n\}$ is a basis of X', and in Linear Algebra this basis is called **dual** to the basis $\{b_1, \ldots, b_n\}$ of X. It is easy to see that:

$$b'_j(b_i) = \begin{cases} 1, & i = j, \\ 0, & i \neq j. \end{cases}$$

3.3 Sequence spaces.

Theorem 3.2. *Let* $1 \leq p \leq +\infty$ *and* $\frac{1}{p} + \frac{1}{q} = 1$.
(i) If $1 \leq p < +\infty$*, then there is a linear isometry of* l^q *onto* $(l^p)'$.
(ii) If $p = +\infty$*, then there is a linear isometry of* l^1 *into* $(l^\infty)'$.

Proof. We take any $z = (\mu_k) \in l^q$ and we consider the function $l_z : l^p \to F$ defined for every $x = (\lambda_k) \in l^p$ by

$$l_z(x) = \sum_{k=1}^{+\infty} \mu_k \lambda_k.$$

Hölder's inequality implies that the series defining $l_z(x)$ converges absolutely. It is easy to see that l_z is a linear functional on l^p. Also, Hölder's inequality says that

$$|l_z(x)| \leq \|z\|_q \|x\|_p$$

for every $x \in l^p$ and hence $l_z \in (l^p)'$ and $\|l_z\| \leq \|z\|_q$.
If $1 < p < +\infty$ (and so $1 < q < +\infty$), we consider the numbers

$$\lambda_k = \begin{cases} \overline{\mu_k}\,|\mu_k|^{q-2}, & \mu_k \neq 0, \\ 0, & \mu_k = 0. \end{cases} \tag{3.1}$$

Then

$$\sum_{k=1}^{+\infty} |\lambda_k|^p = \sum_{k=1}^{+\infty} |\mu_k|^q < +\infty,$$

and so $x = (\lambda_k) \in l^p$ with $\|x\|_p = \|z\|_q^{q/p}$. Also

$$\sum_{k=1}^{+\infty} \mu_k \lambda_k = \sum_{k=1}^{+\infty} |\mu_k|^q.$$

Hence,

$$\|z\|_q^q = |l_z(x)| \leq \|l_z\| \|x\|_p = \|l_z\| \|z\|_q^{q/p},$$

and so $\|z\|_q \leq \|l_z\|$.
If $p = +\infty$ (and so $q = 1$), we select again the $x = (\lambda_k)$ given by (3.1). Then $|\lambda_k| \leq 1$ for every k and so $\|x\|_\infty \leq 1$. Also, $\sum_{k=1}^{+\infty} \mu_k \lambda_k = \sum_{k=1}^{+\infty} |\mu_k|$. Thus,

$$\|z\|_1 = \sum_{k=1}^{+\infty} \mu_k \lambda_k = |l_z(x)| \leq \|l_z\| \|x\|_\infty \leq \|l_z\|.$$

If $p = 1$ (and so $q = +\infty$), then

$$|\mu_k| = |l_z(e_k)| \leq \|l_z\| \|e_k\|_1 = \|l_z\|$$

for every k, and so $\|z\|_\infty \leq \|l_z\|$.
So, in any case we get

$$\|l_z\| = \|z\|_q.$$

We consider, now, the function $T : l^q \to (l^p)'$ defined for every $z \in l^q$ by

$$T(z) = l_z.$$

As in the proof of theorem 3.1 we see that $T : l^q \to (l^p)'$ is a linear operator. The equality $\|T(z)\| = \|l_z\| = \|z\|_q$ says that T is a linear isometry of l^q into $(l^p)'$.
Now, we take any $l \in (l^p)'$.
Let $1 < p < +\infty$. We consider $\mu_k = l(e_k)$ for every k, and λ_k as in (3.1). Then for every n we have

$$\sum_{k=1}^{n} |\mu_k|^q = \sum_{k=1}^{n} \mu_k \lambda_k = l\Big(\sum_{k=1}^{n} \lambda_k e_k\Big) \leq \|l\| \Big(\sum_{k=1}^{n} |\lambda_k|^p\Big)^{1/p} = \|l\| \Big(\sum_{k=1}^{n} |\mu_k|^q\Big)^{1/p}.$$

This implies $\sum_{k=1}^{n} |\mu_k|^q \leq \|l\|^q$ for every n and hence $\sum_{k=1}^{+\infty} |\mu_k|^q \leq \|l\|^q$. So if we define $z = (\mu_k)$, then $z \in l^q$ and $\|z\|_q \leq \|l\| < +\infty$.
If $p = 1$, we consider again $\mu_k = l(e_k)$, and then

$$|\mu_k| = |l(e_k)| \leq \|l\| \|e_k\|_1 = \|l\|$$

for every k. So if we define $z = (\mu_k)$, then $z \in l^\infty$ and $\|z\|_\infty \leq \|l\| < +\infty$.
So if $1 \leq p < +\infty$, we have $z \in l^q$.
Now for any $x = (\lambda_k) \in l^p$ we take $x_n = (\lambda_1, \ldots, \lambda_n, 0, 0, \ldots) = \sum_{k=1}^{n} \lambda_k e_k$ and then

$$l_z(x_n) = \sum_{k=1}^{n} \mu_k \lambda_k = \sum_{k=1}^{n} l(e_k) \lambda_k = l(x_n).$$

Since l_z, l are continuous and $x_n \to x$ in l^p, we get $l_z(x) = l(x)$. Thus $T(z) = l_z = l$ and so T is onto $(l^p)'$. $\square$

In fact the main result of theorem 3.2 is the "onto" part:
Let $1 \leq p < +\infty$ *and* $\frac{1}{p} + \frac{1}{q} = 1$. *Then for every* $l \in (l^p)'$ *there is a unique* $z = (\mu_k) \in l^q$ *so that*

$$\|l\| = \|z\|_q, \qquad l(x) = \sum_{k=1}^{+\infty} \mu_k \lambda_k \quad \text{for every } x = (\lambda_k) \in l^p.$$

In the case $p = +\infty, q = 1$ it is worth finding the point at which the last proof fails to show that the operator T is onto $(l^p)'$: the problem is that for the general $x \in l^\infty$ it is not always true that $x_n \to x$ in l^∞ !

Theorem 3.3. *There is a linear isometry of* l^1 *onto* $(c_0)'$.

Proof. We take any $z = (\mu_k) \in l^1$ and we consider the function $l_z : c_0 \to F$ defined for every $x = (\lambda_k) \in c_0$ by

$$l_z(x) = \sum_{k=1}^{+\infty} \mu_k \lambda_k.$$

The series converges absolutely, and it is clear that l_z is a linear functional on c_0. Also,

$$|l_z(x)| \leq \|z\|_1 \|x\|_\infty$$

for every $x \in c_0$ and hence $l_z \in (c_0)'$ with $\|l_z\| \leq \|z\|_1$.
We consider the λ_k defined in (3.1) (with $q = 1$) and then for every n we have

$$\sum_{k=1}^{n} |\mu_k| = \sum_{k=1}^{n} \mu_k \lambda_k = l_z\Big(\sum_{k=1}^{n} \lambda_k e_k\Big) \leq \|l_z\| \Big\|\sum_{k=1}^{n} \lambda_k e_k\Big\|_\infty \leq \|l_z\|,$$

and so $\|z\|_1 = \sum_{k=1}^{+\infty} |\mu_k| \leq \|l_z\|$.
Therefore, $\|l_z\| = \|z\|_1$.
We consider the function $T : l^1 \to (c_0)'$ defined for every $z \in l^1$ by

$$T(z) = l_z.$$

It is easy to see that T is a linear operator. Since we have proved that $\|T(z)\| = \|l_z\| = \|z\|_1$ for every $z \in l^1$, we have that T is a linear isometry of l^1 into $(c_0)'$.
Now we take any $l \in (c_0)'$. We define $\mu_k = l(e_k)$ for every k, and the same λ_k as above. Then for every n we get

$$\sum_{k=1}^{n} |\mu_k| = \sum_{k=1}^{n} \mu_k \lambda_k = l\Big(\sum_{k=1}^{n} \lambda_k e_k\Big) \leq \|l\| \Big\|\sum_{k=1}^{n} \lambda_k e_k\Big\|_\infty \leq \|l\|.$$

So, if we consider $z = (\mu_k)$, then $z \in l^1$ and $\|z\|_1 = \sum_{k=1}^{+\infty} |\mu_k| \leq \|l\|$.
Now, for every $x = (\lambda_k) \in c_0$ we take $x_n = (\lambda_1, \ldots, \lambda_n, 0, 0, \ldots) = \sum_{k=1}^{n} \lambda_k e_k$, and then

$$l_z(x_n) = \sum_{k=1}^{n} \mu_k \lambda_k = \sum_{k=1}^{n} l(e_k) \lambda_k = l(x_n).$$

Since l_z, l are continuous and $x_n \to x$ in c_0, we get that $l_z(x) = l(x)$. Thus $T(z) = l$ and so T is onto $(c_0)'$. $\square$

The main result of theorem 3.3 is its "onto" part:
For every $l \in (c_0)'$ *there is a unique* $z = (\mu_k) \in l^1$ *so that*

$$\|l\| = \|z\|_1, \qquad l(x) = \sum_{k=1}^{+\infty} \mu_k \lambda_k \quad \text{for every } x = (\lambda_k) \in c_0.$$

3.4 Inner product spaces.

Definition. *Let X, Y be linear spaces over F and let $T : X \to Y$. Then T is called* **conjugate-linear operator** *if*

$$T(x_1 + x_2) = T(x_1) + T(x_2), \quad T(\lambda x) = \overline{\lambda}\, T(x)$$

for every $x, x_1, x_2 \in X$ and every $\lambda \in F$.
If X, Y are normed spaces with norms $\| \cdot \|_X, \| \cdot \|_Y$, and $T : X \to Y$ is conjugate-linear and satisfies $\|T(x)\|_Y = \|x\|_X$ for every $x \in X$, then T is called **conjugate-linear isometry** *of X into Y.*

Of course, if $F = \mathbb{R}$, then a conjugate-linear operator is just a linear operator.

The theorem of F.Riesz. *Let X be an inner product space. Then there is a conjugate-linear isometry of X into X'. If X is a Hilbert space, then this conjugate-linear isometry is onto X'.*

Proof. For every $z \in X$ we consider the function $l_z : X \to F$ defined for every $x \in X$ by

$$l_z(x) = \langle x, z \rangle.$$

It is obvious that l_z is a linear functional on X. Also,

$$|l_z(x)| \leq \|z\| \|x\|$$

for every $x \in X$. Hence $l_z \in X'$ and $\|l_z\| \leq \|z\|$.
Moreover,

$$\|z\|^2 = \langle z, z \rangle = l_z(z) \leq \|l_z\| \|z\|.$$

So $\|z\| \leq \|l_z\|$ and hence

$$\|l_z\| = \|z\|.$$

We consider $T : X \to X'$ defined for every $z \in X$ by

$$T(z) = l_z.$$

It is easy to see that

$$T(z_1 + z_2) = T(z_1) + T(z_2), \quad T(\lambda z) = \overline{\lambda}\, T(z)$$

for every $z, z_1, z_2 \in X$ and every $\lambda \in F$. We have already proven that $\|T(z)\| = \|l_z\| = \|z\|$ for every $z \in X$ and so T is a conjugate-linear isometry of X into X'.
Now we assume that X is complete, and we take any $l \in X'$.
If $l = 0$, then, taking $z = 0$, we obviously have $T(z) = l_z = l$. So we assume that $l \neq 0$ and then $\mathrm{N}(l)$ is a *proper* closed subspace of X. We take any $x_0 \notin \mathrm{N}(l)$, and then there are $y_0 \in \mathrm{N}(l)$ and $z_0 \perp \mathrm{N}(l)$ so that $x_0 = y_0 + z_0$. Then $l(z_0) = l(x_0) \neq 0$.
Now we take any $x \in X$. Then there are $y \in \mathrm{N}(l)$ and $w \perp \mathrm{N}(l)$ so that $x = y + w$. Now we have

$$l\Big(w - \frac{l(w)}{l(z_0)}\, z_0\Big) = l(w) - \frac{l(w)}{l(z_0)}\, l(z_0) = 0,$$

and so $w - \frac{l(w)}{l(z_0)}\, z_0 \in \mathrm{N}(l)$. Since also $w - \frac{l(w)}{l(z_0)}\, z_0 \perp \mathrm{N}(l)$, we get $w - \frac{l(w)}{l(z_0)}\, z_0 = 0$ and hence $w = \frac{l(w)}{l(z_0)}\, z_0$. Therefore,

$$\langle x, z_0 \rangle = \langle y, z_0 \rangle + \langle w, z_0 \rangle = \langle w, z_0 \rangle = \frac{l(w)}{l(z_0)}\, \|z_0\|^2 = \frac{l(x)}{l(z_0)}\, \|z_0\|^2.$$

We define $z = \frac{\overline{l(z_0)}}{\|z_0\|^2}\, z_0$, and then we have $l_z(x) = \langle x, z \rangle = l(x)$ for every $x \in X$. I.e. $T(z) = l_z = l$, and so T is onto X'. $\square$

The main result of the theorem of F. Riesz is the "onto" part:
If X is a Hilbert space with inner product $\langle\cdot,\cdot\rangle$, then for every $l \in X'$ there is a unique $z \in X$ so that

$$\|l\| = \|z\|, \qquad l(x) = \langle x, z\rangle \quad \text{for every } x \in X.$$

3.5 Function spaces.

Definition. *Let (Ω, Σ, μ) be a measure space and let $\nu \in \mathcal{A}(\Omega, \Sigma)$.*
*(i) μ is called σ-**finite** if there are $A_1, A_2, \ldots \in \Sigma$ so that $\Omega = \bigcup_{j=1}^{+\infty} A_j$ and $\mu(A_j) < +\infty$ for every j.*
*(ii) ν is called **absolutely continuous** with respect to μ, if $\nu(A) = 0$ for every $A \in \Sigma$ with $\mu(A) = 0$.*

The next theorem is a well known result of Measure Theory. We shall see its proof by von Neumann with Hilbert space methods.

The theorem of Radon-Nikodym. *Let (Ω, Σ, μ) be a measure space and let $\nu \in \mathcal{A}(\Omega, \Sigma)$. If μ is σ-finite and ν is absolutely continuous with respect to μ, then there is a unique $h \in L^1(\Omega, \Sigma, \mu)$ so that $\nu(A) = \int_A h\, d\mu$ for every $A \in \Sigma$.*

Proof. We assume that ν is a non-negative real measure and that $\mu(\Omega) < +\infty$.
We consider the non-negative real measure $\lambda = \mu + \nu$, and the function $l : L^2(\Omega, \Sigma, \lambda) \to F$ defined for every $f \in L^2(\Omega, \Sigma, \lambda)$ by

$$l(f) = \int_\Omega f\, d\nu.$$

Using Schwartz's inequality, we get

$$|l(f)| \le \int_\Omega |f|\, d\nu \le \int_\Omega |f|\, d\lambda \le (\lambda(\Omega))^{1/2} \Big(\int_\Omega |f|^2\, d\lambda\Big)^{1/2} = (\lambda(\Omega))^{1/2} \|f\|_2$$

for every $f \in L^2(\Omega, \Sigma, \lambda)$. It is clear that l is a linear functional and hence $l \in \big(L^2(\Omega, \Sigma, \lambda)\big)'$. Since $L^2(\Omega, \Sigma, \lambda)$ is a Hilbert space, the theorem of F. Riesz implies that there is $g \in L^2(\Omega, \Sigma, \lambda)$ so that

$$\int_\Omega f\, d\nu = l(f) = \int_\Omega fg\, d\lambda \quad \text{for every } f \in L^2(\Omega, \Sigma, \lambda). \tag{3.2}$$

Now we consider the set $A = \{a \in \Omega \,|\, \mathrm{Im}(g(a)) > 0\} \in \Sigma$. If we use $f = \chi_A \in L^2(\Omega, \Sigma, \lambda)$ in (3.2), and if we equate the imaginary parts of both sides, we get $0 = \int_A \mathrm{Im}(g)\, d\lambda$, and hence $\lambda(A) = 0$. In the same manner we find $\lambda(A) = 0$ for the set $A = \{a \in \Omega \,|\, \mathrm{Im}(g(a)) < 0\}) = 0$. We conclude that $g(a) \in \mathbb{R}$ for λ-a.e. $a \in \Omega$.
Next we consider the set $A = \{a \in \Omega \,|\, \mathrm{Re}(g(a)) > 1\}$, we use $f = \chi_A$ in (3.2), and we equate tha real parts of both sides. Then we get $\int_A (1 - \mathrm{Re}(g))\, d\lambda \ge 0$ and hence $\lambda(A) = 0$. In the same way we get $\lambda(A) = 0$ for the set $A = \{a \in \Omega \,|\, \mathrm{Re}(g(a)) < 0\}$.
Hence, $0 \le g(a) \le 1$ for λ-a.e. $a \in \Omega$.
Our equality (3.2) is equivalent to

$$\int_\Omega f(1-g)\, d\nu = \int_\Omega fg\, d\mu \quad \text{for every } f \in L^2(\Omega, \Sigma, \lambda). \tag{3.3}$$

If we take $B = \{a \in \Omega \,|\, g(a) = 1\}$ and use $f = \chi_B$ in (3.3), we get $0 = \mu(B)$ and so $\nu(B) = 0$. Hence $\lambda(B) = 0$ and so $0 \le g(a) < 1$ for λ-a.e. $a \in \Omega$.
Now, for any $A \in \Sigma$ we consider the function $f = (1 + g + g^2 + \cdots + g^n)\chi_A$, and from (3.3) we get

$$\int_A (1 - g^{n+1})\, d\nu = \int_A (g + g^2 + \cdots + g^{n+1})\, d\mu.$$

The monotone convergence theorem implies $\nu(A) = \int_A \frac{g}{1-g}\, d\mu$. We set $h = \frac{g}{1-g}$, and so we have

$$\nu(A) = \int_A h\, d\mu$$

for every $A \in \Sigma$. Clearly, $0 \le h(a) < +\infty$ for λ-a.e. $a \in \Omega$. With $A = \Omega$ we get $\int_\Omega h\, d\mu = \nu(\Omega) < +\infty$, from which $h \in L^1(\Omega, \Sigma, \mu)$.
The general case of a real or complex measure ν and of a σ-finite measure μ can be derived from the particular case we just studied, using standard measure-theoretic techniques, and it is left as an exercise. The uniqueness of h is also left as an exercise. $\square$

Theorem 3.4. *Let* (Ω, Σ, μ) *be a measure space, and* $1 \le p \le +\infty$, $\frac{1}{p} + \frac{1}{q} = 1$.
(i) If $1 < p < +\infty$*, then there is a linear isometry of* $L^q(\Omega, \Sigma, \mu)$ *onto* $\big(L^p(\Omega, \Sigma, \mu)\big)'$.
If μ *is* σ*-finite, then the same result is true when* $p = 1$.
(ii) If $p = +\infty$*, then there is a linear isometry of* $L^1(\Omega, \Sigma, \mu)$ *into* $\big(L^\infty(\Omega, \Sigma, \mu)\big)'$.

Proof. For each $h \in L^q(\Omega, \Sigma, \mu)$ we consider the function $l_h : L^p(\Omega, \Sigma, \mu) \to F$ defined for every $f \in L^p(\Omega, \Sigma, \mu)$ by

$$l_h(f) = \int_\Omega f h\, d\mu.$$

Hölder's inequality implies that the function fh is integrable and so the integral defining $l_h(f)$ exists. Hölder's inequality implies the more precise inequality

$$|l_h(f)| \le \|h\|_q \|f\|_p$$

for every $f \in L^p(\Omega, \Sigma, \mu)$. It is easy to see that l_h is a linear functional, and so $l_h \in \big(L^p(\Omega, \Sigma, \mu)\big)'$ and $\|l_h\| \le \|h\|_q$.
If $1 < p < +\infty$ (and so $1 < q < +\infty$), we define

$$f(a) = \begin{cases} \overline{h(a)}\, |h(a)|^{q-2}, & h(a) \ne 0, \\ 0, & h(a) = 0. \end{cases} \tag{3.4}$$

Then $\int_\Omega |f|^p\, d\mu = \int_\Omega |h|^q\, d\mu < +\infty$ and so $f \in L^p(\Omega, \Sigma, \mu)$. Also, $\int_\Omega f h\, d\mu = \int_\Omega |h|^q\, d\mu$ and hence

$$\|h\|_q^q = |l_h(f)| \le \|l_h\| \|f\|_p = \|l_h\| \|h\|_q^{q/p}.$$

This implies $\|h\|_q \le \|l_h\|$.
If $p = +\infty, q = 1$, then with the f defined by (3.4) we get $\|f\|_\infty \le 1$ and $\int_\Omega f h\, d\mu = \int_\Omega |h|\, d\mu$ and hence

$$\|h\|_1 = |l_h(f)| \le \|l_h\| \|f\|_\infty \le \|l_h\|.$$

If $p = 1, q = +\infty$ and if μ is σ-finite, then there are pairwise disjoint $B_1, B_2, \ldots \in \Sigma$ so that $\Omega = \bigcup_{j=1}^{+\infty} B_j$ and $\mu(B_j) < +\infty$ for every j. If $\|h\|_\infty = 0$, then $h = 0$, and $l_h = 0$ and so $\|l_h\| = \|h\|_\infty = 0$. If $\|h\|_\infty > 0$, then for any j and any t such that $0 < t < \|h\|_\infty$ we consider the set $B_{j,t} = \{a \in B_j \,|\, |h(a)| \ge t\}$ and the function $f = \overline{h}\, |h|^{-1} \chi_{B_{j,t}}$. Then $\|f\|_1 = \mu(B_{j,t})$ and $\int_\Omega f h\, d\mu = \int_{B_{j,t}} |h|\, d\mu$ and hence

$$t\mu(B_{j,t}) \le |l_h(f)| \le \|l_h\| \|f\|_1 = \|l_h\| \mu(B_{j,t}).$$

Since $0 < \mu(B_{j,t}) < +\infty$ for at least one j, we get $t \le \|l_h\|$. Taking the limit when $t \to \|h\|_\infty$, we find $\|h\|_\infty \le \|l_h\|$.
So in any case we have that $\|l_h\| = \|h\|_q$ for every $h \in L^q(\Omega, \Sigma, \mu)$.
We consider the function $T : L^q(\Omega, \Sigma, \mu) \to \big(L^p(\Omega, \Sigma, \mu)\big)'$ defined for every $h \in L^q(\Omega, \Sigma, \mu)$ by

$$T(h) = l_h.$$

It is easy to see that T is a linear operator, and we have already proved that $\|T(h)\| = \|h\|_q$ for every $h \in L^q(\Omega, \Sigma, \mu)$. This says that T is a linear isometry of $L^q(\Omega, \Sigma, \mu)$ into $\big(L^p(\Omega, \Sigma, \mu)\big)'$.
For the rest of the proof we assume that $\mu(\Omega) < +\infty$ and $1 \leq p < +\infty$.
We take any $l \in \big(L^p(\Omega, \Sigma, \mu)\big)'$, and we consider $\nu : \Sigma \to F$ defined for every $A \in \Sigma$ by

$$\nu(A) = l(\chi_A).$$

Clearly $\nu(\emptyset) = l(0) = 0$.
If $A_1, A_2, \ldots \in \Sigma$ are pairwise disjoint, and $A = \bigcup_{k=1}^{+\infty} A_k$, we set $C_n = \bigcup_{k=1}^{n} A_k$, and then

$$\begin{aligned} |\nu(A) - \nu(C_n)| &= |l(\chi_A) - l(\chi_{C_n})| = |l(\chi_A - \chi_{C_n})| = |l(\chi_{A\setminus C_n})| \leq \|l\| \|\chi_{A\setminus C_n}\|_p \\ &= \|l\| (\mu(A \setminus C_n))^{1/p} \to 0, \end{aligned}$$

since $1 \leq p < +\infty$, $A \setminus C_n \downarrow \emptyset$ and μ is finite. Hence $\nu(C_n) \to \nu(A)$ and so

$$\sum_{k=1}^{n} \nu(A_k) = \sum_{k=1}^{n} l(\chi_{A_k}) = l\Big(\sum_{k=1}^{n} \chi_{A_k}\Big) = l(\chi_{C_n}) = \nu(C_n) \to \nu(A).$$

Therefore, $\nu(A) = \sum_{k=1}^{+\infty} \nu(A_k)$, and we conclude that ν is a complex measure on Σ.
If $\mu(A) = 0$, then

$$|\nu(A)| = |l(\chi_A)| \leq \|l\| \|\chi_A\|_p = \|l\| (\mu(A))^{1/p} = 0,$$

and so ν is absolutely continuous with respect to μ.
The theorem of Radon-Nikodym implies that there is $h \in L^1(\Omega, \Sigma, \mu)$ so that

$$\nu(A) = \int_A h \, d\mu$$

for every $A \in \Sigma$. Now, if $f = \sum_{k=1}^{n} \lambda_k \chi_{A_k}$ is any simple function, then

$$l(f) = \sum_{k=1}^{n} \lambda_k l(\chi_{A_k}) = \sum_{k=1}^{n} \lambda_k \nu(A_k) = \sum_{k=1}^{n} \lambda_k \int_{A_k} h \, d\mu = \int_\Omega f h \, d\mu.$$

If $f \in L^\infty(\Omega, \Sigma, \mu)$, there is a sequence (f_k) of simple functions so that $f_k \to f$ in $L^\infty(\Omega, \Sigma, \mu)$. Since μ is finite, we have that $f_k \to f$ in $L^p(\Omega, \Sigma, \mu)$. Now, l is continuous and so $l(f_k) \to l(f)$. Also,

$$\Big| \int_\Omega f_k h \, d\mu - \int_\Omega f h \, d\mu \Big| \leq \|f_k - f\|_\infty \|h\|_1 \to 0.$$

Hence

$$l(f) = \lim_{k\to+\infty} l(f_k) = \lim_{k\to+\infty} \int_\Omega f_k h \, d\mu = \int_\Omega f h \, d\mu \tag{3.5}$$

for every $f \in L^\infty(\Omega, \Sigma, \mu)$.
If $1 < p < +\infty$, then for every n we consider the set $A_n = \{a \in \Omega \,|\, |h(a)| \leq n\}$ and we define

$$f(a) = \begin{cases} \overline{h(a)}\, |h(a)|^{q-2} \chi_{A_n}(a), & h(a) \neq 0, \\ 0, & h(a) = 0. \end{cases}$$

Then $f \in L^\infty(\Omega, \Sigma, \mu)$ and also $\int_\Omega |f|^p \, d\mu = \int_{A_n} |h|^q \, d\mu$ and $\int_\Omega f h \, d\mu = \int_{A_n} |h|^q \, d\mu$. So, using (3.5), we find

$$\int_{A_n} |h|^q \, d\mu = \int_\Omega h f \, d\mu = l(f) \leq \|l\| \|f\|_p = \|l\| \Big(\int_{A_n} |h|^q \, d\mu \Big)^{1/p}.$$

Therefore $\int_{A_n} |h|^q \, d\mu \leq \|l\|^q$ for every n, and from the monotone convergence theorem we conclude that $\|h\|_q \leq \|l\| < +\infty$.
If $p = 1$, then for every $n > \|l\|$ we consider the set $A_n = \{a \in \Omega \,|\, \|l\| < |h(a)| \leq n\}$ and the function $f = \overline{h}\, |h|^{-1} \chi_{A_n}$. Then $\|f\|_1 = \mu(A_n)$ and $\int_\Omega fh \, d\mu = \int_{A_n} |h| \, d\mu$. Also $f \in L^\infty(\Omega, \Sigma, \mu)$, and, using (3.5),

$$\int_{A_n} |h| \, d\mu = \int_\Omega fh \, d\mu = l(f) \leq \|l\| \|f\|_1 = \|l\| \mu(A_n).$$

Therefore, $\mu(A_n) = 0$ for every n and hence $|h(a)| \leq \|l\|$ for μ-a.e. $a \in \Omega$, and we conclude that $\|h\|_\infty \leq \|l\| < +\infty$.
So in any case, $h \in L^q(\Omega, \Sigma, \mu)$. Then for every $f \in L^p(\Omega, \Sigma, \mu)$ we take a sequence (f_k) of simple functions so that $f_k \to f$ in $L^p(\Omega, \Sigma, \mu)$. Now, the continuity of l and Hölder's inequality together with (3.5) for each f_k imply

$$l(f) = \lim_{k \to +\infty} l(f_k) = \lim_{k \to +\infty} \int_\Omega f_k h \, d\mu = \int_\Omega fh \, d\mu = l_h(f).$$

Therefore, $l = l_h = T(h)$, and so T is onto $\big(L^p(\Omega, \Sigma, \mu)\big)'$.
The general case of a measure μ which is not necessarily finite can be derived from the particular case of a finite μ using standard measure-theoretic arguments and it is left as an exercise. $\square$

The main result of theorem 3.4 is its "onto" part:
Let (Ω, Σ, μ) *be a measure space, and* $1 < p < +\infty$, $\frac{1}{p} + \frac{1}{q} = 1$. *Then for every* $l \in \big(L^p(\Omega, \Sigma, \mu)\big)'$ *there is a unique* $h \in L^q(\Omega, \Sigma, \mu)$ *so that*

$$\|l\| = \|h\|_q, \qquad l(f) = \int_\Omega fh \, d\mu \quad \text{for every } f \in L^p(\Omega, \Sigma, \mu).$$

If μ *is* σ*-finite, then the same result is true when* $p = 1$.

Lemma 3.1. *Let* Ω *be a Hausdorff topological space, and let* $K, L \subseteq \Omega$ *be compact and disjoint. Then there are disjoint open* $U, V \subseteq \Omega$ *so that* $K \subseteq U$ *and* $L \subseteq V$.

Proof. Take any $x \in K$. For every $y \in L$ we consider disjoint open U_y, V_y so that $x \in U_y$ and $y \in V_y$. Then the collection $\{V_y \,|\, y \in L\}$ is an open covering of L, and so there are $y_1, \dots, y_n \in L$ so that $L \subseteq V_{y_1} \cup \dots \cup V_{y_n}$. Then the open sets $U_x = U_{y_1} \cap \dots \cap U_{y_n}$ and $V_x = V_{y_1} \cup \dots \cup V_{y_n}$ are disjoint, and $x \in U_x$ and $L \subseteq V_x$. Then the collection $\{U_x \,|\, x \in K\}$ is an open covering of K, and so there are $x_1, \dots, x_m \in K$ so that $K \subseteq U_{x_1} \cup \dots \cup U_{x_m}$. Then the open sets $U = U_{x_1} \cup \dots \cup U_{x_m}$ and $V = V_{x_1} \cap \dots \cap V_{x_m}$ are disjoint, and $K \subseteq U$ and $L \subseteq V$. $\square$

Urysohn's lemma. *Let* Ω *be a compact, Hausdorff topological space, and* $K, L \subseteq \Omega$ *be closed and disjoint. Then there is a continuous* $f : \Omega \to [0, 1]$ *so that* $f = 0$ *in* K*, and* $f = 1$ *in* L.

Proof. Lemma 3.1 implies that if $A \subseteq \Omega$ is closed and $B \subseteq \Omega$ is open, and if $A \subseteq B$, then there is an open U so that $A \subseteq U \subseteq \mathrm{cl}(U) \subseteq B$.
We consider $A_0 = K$ and $B_1 = \Omega \setminus L$. Then there is an open $B_{1/2}$ so that

$$A_0 \subseteq B_{1/2} \subseteq \mathrm{cl}(B_{1/2}) \subseteq B_1.$$

Then there are open $B_{1/4}$ and $B_{3/4}$ so that

$$A_0 \subseteq B_{1/4} \subseteq \mathrm{cl}(B_{1/4}) \subseteq B_{1/2} \subseteq \mathrm{cl}(B_{1/2}) \subseteq B_{3/4} \subseteq \mathrm{cl}(B_{3/4}) \subseteq B_1.$$

Let $\mathbb{Q}_d$ be the set of all rational numbers of the form $r = k/2^n$ with $0 < k \leq 2^n$.
Continuing inductively, we see that to every $r \in \mathbb{Q}_d$ corresponds an open set B_r, so that

$$A_0 \subseteq B_r \subseteq \mathrm{cl}(B_r) \subseteq B_s$$

for every $r, s \in \mathbb{Q}_d$ with $r < s$.
We consider $f : \Omega \to \mathbb{R}$ defined so that

$$f(x) = \begin{cases} \inf\{r \in \mathbb{Q}_d \,|\, x \in B_r\}, & x \in B_1, \\ 1, & x \in \Omega \setminus B_1. \end{cases}$$

Then $f = 0$ in K, and $f = 1$ in L, and $f : \Omega \to [0, 1]$. It remains to show that f is continuous.
Take any $x \in \Omega$ and any $\epsilon > 0$. If $0 < f(x) < 1$, then there are $r, r', s \in \mathbb{Q}_d$ so that

$$f(x) - \epsilon < r < r' < f(x) < s < f(x) + \epsilon.$$

If $y \in B_s$, then $f(y) \leq s < f(x)+\epsilon$. If $y \in \Omega \setminus \mathrm{cl}(B_r)$, then $y \notin B_r$, and then $f(y) \geq r > f(x)-\epsilon$.
Also, $x \in B_s$ and $x \notin B_{r'}$, and hence $x \in \Omega \setminus \mathrm{cl}(B_r)$. Therefore the open $V = B_s \cap (\Omega \setminus \mathrm{cl}(B_r))$ contains x, and $f(x) - \epsilon < f(y) < f(x) + \epsilon$ for every $y \in V$. So f is continuous at x.
If $f(x) = 1$, we consider, as above, $r, r' \in \mathbb{Q}_d$ so that $1 - \epsilon < r < r' < 1$. Then the open $V = \Omega \setminus \mathrm{cl}(B_r)$ contains x, and $1 - \epsilon < f(y) \leq 1 < 1 + \epsilon$ for every $y \in V$.
Similarly, if $f(x) = 0$, we consider $s \in \mathbb{Q}_d$ so that $0 < s < \epsilon$. Then the open $V = B_s$ contains x, and $-\epsilon < 0 \leq f(y) < \epsilon$ for every $y \in V$.
In any case, f is continuous at x. □

We should remark that Urysohn's lemma holds, more generally, for *normal* topological spaces Ω, i.e. Hausdorff topological spaces with the property: for every two disjoint closed $K, L \subseteq \Omega$ there are disjoint open $U, V \subseteq \Omega$ so that $K \subseteq U$ and $L \subseteq V$. This is the only property of Ω which was used in the proof of Urysohn's lemma. Lemma 3.1 says that compact, Hausdorff topological spaces are normal. Another class of normal spaces are the metric spaces. In fact, for a metric space Ω, Urysohn's lemma has a simple proof: we consider the function $f(x) = \frac{d(x,K)}{d(x,K)+d(x,L)}$ for every $x \in \Omega$, where $d(x, A) = \inf_{y\in A} d(x, y)$ for every $A \subseteq \Omega$.

Lemma 3.2. *Le Ω be a compact, Hausdorff topological space, let $K \subseteq \Omega$ be compact, and let $U_1, \ldots, U_n \subseteq \Omega$ be open, so that $K \subseteq U_1 \cup \cdots \cup U_n$. Then there are continuous $f_1, \ldots, f_n : \Omega \to [0, 1]$ so that $\mathrm{supp}(f_j) \subseteq U_j$ for every j and $f_1 + \cdots + f_n = 1$ in K.*

Proof. We have that $K \setminus (U_2 \cup \cdots \cup U_n) \subseteq U_1$ and so there is an open V_1 so that

$$K \setminus (U_2 \cup \cdots \cup U_n) \subseteq V_1 \subseteq \mathrm{cl}(V_1) \subseteq U_1.$$

Then $K \subseteq V_1 \cup U_2 \cup \cdots \cup U_n$ and hence $K \setminus (V_1 \cup U_3 \cup \cdots \cup U_n) \subseteq U_2$. So there is an open V_2 so that

$$K \setminus (V_1 \cup U_3 \cup \cdots \cup U_n) \subseteq V_2 \subseteq \mathrm{cl}(V_2) \subseteq U_2.$$

Then $K \subseteq V_1 \cup V_2 \cup U_3 \cup \cdots \cup U_n$. We continue inductively replacing the open $U_1, \ldots, U_n$ with the open $V_1, \ldots, V_n$ so that $K \subseteq V_1 \cup \cdots \cup V_n$, and $\mathrm{cl}(V_j) \subseteq U_j$ for every j.
We repeat this process, and we find open $W_1, \ldots, W_n$ so that $K \subseteq W_1 \cup \cdots \cup W_n$ and $\mathrm{cl}(W_j) \subseteq V_j \subseteq \mathrm{cl}(V_j) \subseteq U_j$ for every j.
Urysohn's lemma implies that there are continuous $g_1, \ldots, g_n : \Omega \to [0, 1]$ so that $g_j = 1$ in $\mathrm{cl}(W_j)$ and $g_j = 0$ in $\Omega \setminus V_j$. There is also a continuous $g_0 : \Omega \to [0, 1]$ so that $g_0 = 0$ in K and $g_0 = 1$ in $\Omega \setminus (W_1 \cup \cdots \cup W_n)$. Now we define

$$f_j = \frac{g_j}{g_0 + g_1 + \cdots + g_n}$$

for every $j = 1, \ldots, n$.
If $g_0(x) \neq 1$, then $x \in W_1 \cup \cdots \cup W_n$, and hence $g_j(x) = 1$ for some $j = 1, \ldots, n$. Therefore, $g_0 + g_1 + \cdots + g_n \geq 1$ in Ω, and so $f_1, \ldots, f_n : \Omega \to [0, 1]$ are continuous in Ω.
If $x \notin V_j$, then $g_j(x) = 0$, hence $f_j(x) = 0$. So $\mathrm{supp}(f_j) \subseteq \mathrm{cl}(V_j) \subseteq U_j$.
Also $f_1 + \cdots + f_n = \frac{g_1+\cdots+g_n}{g_0+g_1+\cdots+g_n} = 1$ in K, since $g_0 = 0$ in K. □

Definition. *Let Ω be a topological space, let $K \subseteq \Omega$ be compact, let $U_1, \dots, U_n \subseteq \Omega$ be open and $K \subseteq U_1 \cup \dots \cup U_n$. If $f_1, \dots, f_n : \Omega \to [0,1]$ are continuous,* $\mathrm{supp}(f_j) \subseteq U_j$ *for every j, and $f_1 + \dots + f_n = 1$ in K, then the collection $\{f_1, \dots, f_n\}$ is called* **partition of unity** *for K with respect to its open covering $\{U_1, \dots, U_n\}$.*

Thus, lemma 3.2 says that in a compact, Hausdorff topological space every compact set has a partition of unity with respect to any of its finite open coverings.

Lemma 3.3. *Let Ω be a topological space, and let $\mu \in \mathcal{A}(\Omega, \mathcal{B}(\Omega))$. Then for every $f \in C(\Omega)$ we have $\left| \int_\Omega f\, d\mu \right| \leq \int_\Omega |f|\, d|\mu| \leq \|f\|_u \|\mu\|$.*

Proof. It is enough to prove the left inequality. This is well known if f is real and μ is non-negative.
If f is real and μ is real, then $\mu = \mu^+ - \mu^-$, where μ^+, μ^- are the positive and the negative variations of μ, and so

$$\left| \int_\Omega f\, d\mu \right| \leq \left| \int_\Omega f\, d\mu^+ \right| + \left| \int_\Omega f\, d\mu^- \right| \leq \int_\Omega |f|\, d\mu^+ + \int_\Omega |f|\, d\mu^- = \int_\Omega |f|\, d|\mu|.$$

If f is complex and μ is complex, then

$$\begin{aligned}
\left| \int_\Omega f\, d\mu \right| &\leq \left| \int_\Omega \mathrm{Re}(f)\, d\,\mathrm{Re}(\mu) \right| + \left| \int_\Omega \mathrm{Re}(f)\, d\,\mathrm{Im}(\mu) \right| \\
&\quad + \left| \int_\Omega \mathrm{Im}(f)\, d\,\mathrm{Re}(\mu) \right| + \left| \int_\Omega \mathrm{Im}(f)\, d\,\mathrm{Im}(\mu) \right| \\
&\leq \int_\Omega |\mathrm{Re}(f)|\, d|\mathrm{Re}(\mu)| + \int_\Omega |\mathrm{Re}(f)|\, d|\mathrm{Im}(\mu)| \\
&\quad + \int_\Omega |\mathrm{Im}(f)|\, d|\mathrm{Re}(\mu)| + \int_\Omega |\mathrm{Im}(f)|\, d|\mathrm{Im}(\mu)| \\
&\leq 4 \int_\Omega |f|\, d|\mu|.
\end{aligned}$$

Now we decompose the disc $\{\lambda \in \mathbb{C} \,|\, |\lambda| \leq \|f\|_u\}$ in pairwise disjoint Borel sets $Q_1, \dots, Q_n$ where each of them has diameter $\leq \epsilon$, and we consider the $A_j = \{x \in \Omega \,|\, f(x) \in Q_j\}$. We also take one $\lambda_j \in Q_j$ for every j, and then

$$\begin{aligned}
\left| \int_\Omega f\, d\mu \right| &\leq \sum_{j=1}^n \left| \int_{A_j} f\, d\mu \right| \leq \sum_{j=1}^n \left| \int_{A_j} (f - \lambda_j)\, d\mu \right| + \sum_{j=1}^n |\lambda_j| |\mu(A_j)| \\
&\leq 4 \sum_{j=1}^n \epsilon |\mu|(A_j) + \sum_{j=1}^n |\lambda_j| |\mu|(A_j) \\
&\leq 4\epsilon |\mu|(\Omega) + \sum_{j=1}^n \int_{A_j} |f|\, d|\mu| + \sum_{j=1}^n \int_{A_j} |f - \lambda_j|\, d|\mu| \\
&\leq 5\epsilon |\mu|(\Omega) + \int_\Omega |f|\, d|\mu|.
\end{aligned}$$

Since $\epsilon > 0$ is arbitrary, we get $\left| \int_\Omega f\, d\mu \right| \leq \int_\Omega |f|\, d|\mu|$. □

Definition. *Let $l : C(\Omega) \to F$ be a linear functional.*
(i) We say that l is **real**, *if $l(f) \in \mathbb{R}$ for every real $f \in C(\Omega)$.*
(i) We say that l is **non-negative**, *if $l(f) \geq 0$ for every non-negative $f \in C(\Omega)$.*

It is easy to see that, if l is real, then $\mathrm{Re}(l(f)) = l(\mathrm{Re}(f))$ and $\mathrm{Im}(l(f)) = l(\mathrm{Im}(f))$ and also $\overline{l(f)} = l(\overline{f})$ for every $f \in C(\Omega)$. Similarly, if l is non-negative, then $l(f) \leq l(g)$ for all real $f, g \in C(\Omega)$ with $f \leq g$ in Ω.

The theorem of F.Riesz-Radon-Banach-Kakutani. *Let Ω be a compact, Hausdorff topological space. Then there is a linear isometry of $\mathcal{A}_r(\Omega, \mathcal{B}(\Omega))$ onto $C(\Omega)'$.*

Proof. For every $\mu \in \mathcal{A}_r(\Omega, \mathcal{B}(\Omega))$ we consider the functon $l_\mu : C(\Omega) \to F$ defined for every $f \in C(\Omega)$ by

$$l_\mu(f) = \int_\Omega f \, d\mu.$$

Then l_μ is a linear functional on $C(\Omega)$ and

$$|l_\mu(f)| = \Big| \int_\Omega f \, d\mu \Big| \leq \|\mu\| \|f\|_u$$

for every $f \in C(\Omega)$. Therefore, $l_\mu \in C(\Omega)'$ and $\|l_\mu\| \leq \|\mu\|$.
We take any $\epsilon > 0$. The definition of $\|\mu\|$ implies that there are pairwise disjoint Borel sets $A_1, \dots, A_n \subseteq \Omega$ so that

$$\|\mu\| - \epsilon < |\mu(A_1)| + \cdots + |\mu(A_n)|.$$

Since μ is regular, for every j there is a compact $K_j \subseteq A_j$ so that $|\mu|(A_j \setminus K_j) < \frac{1}{n} \, \epsilon$ and so

$$\|\mu\| - 2\epsilon < |\mu(K_1)| + \cdots + |\mu(K_n)|.$$

Since $K_1, \dots, K_n$ are pairwise disjoint, there are pairwise disjoint open $U_1, \dots, U_n$ so that $K_j \subseteq U_j$ for every j, and $|\mu|(U_j \setminus K_j) < \frac{1}{n} \, \epsilon$ for every j. Urysohn's lemma implies that for every j there is a continuous $f_j : \Omega \to [0, 1]$ so that $f_j = 1$ in K_j and $f_j = 0$ in $\Omega \setminus U_j$. Finally, we consider

$$\lambda_j = \begin{cases} \overline{\int_{U_j} f_j \, d\mu} \, \big| \int_{U_j} f_j \, d\mu \big|^{-1}, & \int_{U_j} f_j \, d\mu \neq 0, \\ 0, & \int_{U_j} f_j \, d\mu = 0, \end{cases}$$

and $f = \sum_{j=1}^n \lambda_j f_j$. Now, $|f| \leq \sum_{j=1}^n |\lambda_j| f_j \leq \sum_{j=1}^n f_j \leq 1$ in Ω. Thus,

$$\begin{aligned} \|l_\mu\| \geq \|f\|_u \|l_\mu\| \geq \Big| \int_\Omega f \, d\mu \Big| &= \Big| \sum_{j=1}^n \lambda_j \int_{U_j} f_j \, d\mu \Big| = \sum_{j=1}^n \Big| \int_{U_j} f_j \, d\mu \Big| \\ &\geq \sum_{j=1}^n |\mu(K_j)| - \sum_{j=1}^n \Big| \int_{U_j \setminus K_j} f_j \, d\mu \Big| > \|\mu\| - 2\epsilon - \sum_{j=1}^n |\mu|(U_j \setminus K_j) > \|\mu\| - 3\epsilon. \end{aligned}$$

Since $\epsilon > 0$ is arbitrary, we conclude that $\|l_\mu\| \geq \|\mu\|$ and hence $\|l_\mu\| = \|\mu\|$.
Assume that l_μ is real. We consider any Borel set A, and then a compact $K \subseteq A$ and an open $U \supseteq A$ so that $|\mu|(U \setminus K) < \epsilon$. There is a continuous $f : \Omega \to [0, 1]$ so that $f = 1$ in K and $f = 0$ in $\Omega \setminus U$. Then

$$\begin{aligned} 0 = \Big| \mathrm{Im}\Big(\int_\Omega f \, d\mu \Big) \Big| &\geq |\,\mathrm{Im}(\mu(K))| - \Big| \mathrm{Im}\Big(\int_{U \setminus K} f \, d\mu \Big) \Big| \\ &\geq |\,\mathrm{Im}(\mu(A))| - |\,\mathrm{Im}(\mu(A \setminus K))| - |\mu|(U \setminus K) \\ &\geq |\,\mathrm{Im}(\mu(A))| - 2|\mu|(U \setminus K) \geq |\,\mathrm{Im}(\mu(A))| - 2\epsilon. \end{aligned}$$

Since $\epsilon > 0$ is arbitrary, we get $\mathrm{Im}(\mu(A)) = 0$ and so μ is a real measure.
Assume that l_μ is non-negative. With the same choice of A, K, U, f as in the previous paragraph, we get

$$0 \leq \int_\Omega f \, d\mu = \mu(K) + \int_{U \setminus K} f \, d\mu \leq \mu(A) + 2|\mu|(U \setminus K) \leq \mu(A) + 2\epsilon.$$

Again, since $\epsilon > 0$ is arbitrary, we find $\mu(A) \geq 0$ and so μ is a non-negative measure.
We consider the function $T : \mathcal{A}_r(\Omega, \mathcal{B}(\Omega)) \to C(\Omega)'$ defined for every $\mu \in \mathcal{A}_r(\Omega, \mathcal{B}(\Omega))$ by

$$T(\mu) = l_\mu.$$

Then T is linear and we have already seen that $\|T(\mu)\| = \|l_\mu\| = \|\mu\|$ for every $\mu \in \mathcal{A}_r(\Omega, \mathcal{B}(\Omega))$.
So it remains to prove that T is onto $C(\Omega)'$, i.e. to prove that for every $l \in C(\Omega)'$ there is a complex (if $F = \mathbb{C}$) or real (if $F = \mathbb{R}$) Borel measure μ so that $l(f) = \int_\Omega f\, d\mu$ for every $f \in C(\Omega)$.
At first we assume that $l \in C(\Omega)'$ is non-negative.
For every open $U \subseteq \Omega$ and every $f \in C(\Omega)$ we write

$$f \prec O,$$

if $f : \Omega \to [0, 1]$ and $\text{supp}(f) \subseteq U$.
Now, for every open $U \subseteq \Omega$ we define

$$\mu(U) = \sup\{l(f) \,|\, f \prec U\}$$

and, then, for every $E \subseteq \Omega$ we define

$$\mu^*(E) = \inf\{\mu(U) \,|\, U \text{ open } \supseteq E\}.$$

If $U, U' \subseteq \Omega$ are open and $U \subseteq U'$, then $f \prec U$ implies $f \prec U'$, and hence $\mu(U) \leq \mu(U')$. Therefore,

$$\mu^*(U) = \mu(U)$$

for every open U.
If $f \prec U$, then $l(f) \leq \|l\| \|f\|_u \leq \|l\|$. So $\mu(U) \leq \|l\|$ and hence $\mu^*(E) \leq \|l\|$ for every $E \subseteq \Omega$.
It is clear that $\mu^*(\emptyset) = \mu(\emptyset) = 0$, and that $\mu^*(E) \leq \mu^*(E')$ for every $E, E' \subseteq \Omega$ with $E \subseteq E'$.
Now, let $E = \bigcup_{j=1}^{+\infty} E_j$. For each j we find an open $U_j \supseteq E_j$ so that $\mu(U_j) < \mu^*(E_j) + \frac{\epsilon}{2^j}$ and we consider the open $U = \bigcup_{j=1}^{+\infty} U_j$. Let $f \prec U$, and let $K = \text{supp}(f) \subseteq U$. Then there is n so that $K \subseteq \bigcup_{j=1}^{n} U_j$ and we consider a partition of unity $\{f_1, \ldots, f_n\}$ for K with respect to $\{U_1, \ldots, U_n\}$. Then $f = ff_1 + \cdots + ff_n$ and $\text{supp}(ff_j) \prec U_j$ for every j, and so

$$l(f) = \sum_{j=1}^{n} l(ff_j) \leq \sum_{j=1}^{n} \mu(U_j) \leq \sum_{j=1}^{+\infty} \mu(U_j) \leq \sum_{j=1}^{+\infty} \mu^*(E_j) + \epsilon.$$

Taking the supremum of $l(f)$ over all $f \prec U$, we get $\mu(U) \leq \sum_{j=1}^{+\infty} \mu^*(E_j) + \epsilon$. Since $E \subseteq U$, we get $\mu^*(E) \leq \sum_{j=1}^{+\infty} \mu^*(E_j) + \epsilon$. Finally, since $\epsilon > 0$ is arbitrary, we find

$$\mu^*(E) \leq \sum_{j=1}^{+\infty} \mu^*(E_j).$$

Thus μ^* is an outer measure on Ω.
Now the process of Caratheodory defines the σ-algebra of μ^*-measurable subsets of Ω, and then μ^* restricted on this σ-algebra is a measure.
We take any open $U \subseteq \Omega$ and any $E \subseteq \Omega$. We take any $\epsilon > 0$, and then there is an open $U' \supseteq E$ with $\mu(U') < \mu^*(E) + \epsilon$, and a $f \prec U' \cap U$ with $l(f) > \mu(U' \cap U) - \epsilon$. Then $U' \setminus \text{supp}(f)$ is open, and there is a $g \prec U' \setminus \text{supp}(f)$ so that $l(g) > \mu(U' \setminus \text{supp}(f)) - \epsilon$. We observe that $f + g \prec U'$, and so

$$\begin{aligned}\mu^*(E) + \epsilon > \mu(U') \geq l(f+g) = l(f) + l(g) > \mu(U' \cap U) + \mu(U' \setminus \text{supp}(f)) - 2\epsilon \\ \geq \mu^*(E \cap U) + \mu^*(E \setminus U) - 2\epsilon.\end{aligned}$$

Since $\epsilon > 0$ is arbitrary, we find

$$\mu^*(E) \geq \mu^*(E \cap U) + \mu^*(E \setminus U)$$

and so U is μ^*-measurable. Therefore, the σ-algebra of μ^*-measurable sets contains all open sets, and so it contains $\mathcal{B}(\Omega)$. We define μ to be the restriction of μ^* on $\mathcal{B}(\Omega)$, and so μ is a non-negative Borel measure on Ω. Then μ is identical with the already defined μ on all open sets, since we have proved that $\mu^*(U) = \mu(U)$ for every open U.
Now we shall prove that

$$\mu(K) = \inf\{l(f) \,|\, f \in C(\Omega), \chi_K \leq f \text{ in } \Omega\} \tag{3.6}$$

for every compact $K \subseteq \Omega$.
We take any $f \in C(\Omega)$ so that $f \geq \chi_K$, i.e. $f \geq 0$ in Ω and $f \geq 1$ in K. We take any t with $0 < t < 1$, and we consider the open set $U = \{x \in \Omega \,|\, f(x) > t\} \supseteq K$. If $g \prec U$, then $tg \leq f$ in Ω, and then $tl(g) \leq l(f)$, since l is non-negative. From this, taking the supremum of $l(g)$ over all $g \prec U$, we find $t\mu(U) \leq l(f)$, and hence $t\mu(K) \leq l(f)$. Then we take the supremum over $t < 1$, and we get $\mu(K) \leq l(f)$. Thus,

$$\mu(K) \leq \inf\{l(f) \,|\, f \in C(\Omega), \chi_K \leq f \text{ in } \Omega\}.$$

Now we take any $\epsilon > 0$, and then there is an open $U \supseteq K$ with $\mu(U) < \mu(K) + \epsilon$, and a continuous $f : \Omega \to [0, 1]$ with $f = 1$ in K and $\text{supp}(f) \subseteq U$. Then $f \geq \chi_K$ and $f \prec U$, and so $l(f) \leq \mu(U) < \mu(K) + \epsilon$. Since $\epsilon > 0$ is arbitrary,

$$\inf\{l(f) \,|\, f \in C(\Omega), \chi_K \leq f \text{ in } \Omega\} \leq \mu(K),$$

and the proof of (3.6) is finished.
We shall now prove the regularity of μ.
For any Borel set E we have

$$\mu(E) = \mu^*(E) = \inf\{\mu(U) \,|\, U \text{ open } \supseteq E\}$$

by the definition of $\mu^*(E)$, and this is the first regularity condition.
Now we take any Borel set E and any $\epsilon > 0$. Then there is an open $U \supseteq E$ so that $\mu(U) < \mu(E)+\epsilon$, and a $g \prec U$ so that $l(g) > \mu(U) - \epsilon$. We consider the compact $K = \text{supp}(g) \subseteq U$. For every $f \in C(\Omega)$ with $f \geq \chi_K$, we have $f \geq g$, and hence $l(f) \geq l(g)$. From (3.6) we get $\mu(K) \geq l(g)$ and hence $\mu(K) > \mu(U) - \epsilon$. Since $\mu(U \setminus E) = \mu(U) - \mu(E) < \epsilon$, there is an open $U' \supseteq U \setminus E$ so that $\mu(U') < 2\epsilon$. Now we set $L = K \setminus U'$ and we observe that L is a compact subset of E and that $E \setminus L \subseteq (U \setminus K) \cup U'$. Hence $\mu(E) - \mu(L) \leq \mu(U \setminus K) + \mu(U') < 3\epsilon$ and so

$$\mu(E) = \sup\{\mu(L) \,|\, L \text{ compact } \subseteq E\}.$$

This is the second regularity condition.
Finally, we shall prove that $l(f) = \int_\Omega f\, d\mu$ for every $f \in C(\Omega)$. Because of the linearity of l, it is enough to prove this for real f. (Of course, if $F = \mathbb{R}$, then all our functions are real.) If f is real we consider $f^+ = \frac{1}{2}(|f| + f) \geq 0$ and $f^- = \frac{1}{2}(|f| - f) \geq 0$, and then $f = f^+ - f^-$. Therefore, in proving $l(f) = \int_\Omega f\, d\mu$ it is enough to consider $f \geq 0$ and, multiplying with an appropriate constant, we may assume that $0 \leq f \leq 1$ in Ω.
We take any $n \in \mathbb{N}$ and we consider $K_k = \{x \in \Omega \,|\, f(x) \geq \frac{k}{n}\}$ for $0 \leq k \leq n$. Then K_k is compact, and $K_0 = \Omega$. Also, for every $j = 0, \ldots, n-1$ we consider the function

$$f_j = \min\Big\{ \max\Big\{f, \frac{j}{n}\Big\}, \frac{j+1}{n}\Big\} - \frac{j}{n}.$$

Then every f_j is continuous in Ω and

$$\frac{1}{n}\chi_{K_{j+1}} \le f_j \le \frac{1}{n}\chi_{K_j}$$

for every $j = 0, \dots, n-1$ and also

$$f = \sum_{j=0}^{n-1} f_j.$$

Adding the last inequalities and integrating we find

$$\frac{1}{n}\sum_{j=1}^{n}\mu(K_j) \le \int_\Omega f\,d\mu \le \frac{1}{n}\sum_{j=0}^{n-1}\mu(K_j).$$

From $\chi_{K_{j+1}} \le nf_j$ and (3.6) we get $\mu(K_{j+1}) \le l(nf_j) = nl(f_j)$. Now, we take any open $U \supseteq K_j$. From $nf_j \le \chi_{K_j}$ we get $nf_j \prec U$ and hence $nl(f_j) = l(nf_j) \le \mu(U)$. So from the definition of $\mu(K_j) = \mu^*(K_j)$ we get $nl(f_j) \le \mu(K_j)$. Therefore,

$$\frac{1}{n}\mu(K_{j+1}) \le l(f_j) \le \frac{1}{n}\mu(K_j),$$

and, adding, we find

$$\frac{1}{n}\sum_{j=1}^{n}\mu(K_j)) \le l(f) \le \frac{1}{n}\sum_{j=0}^{n-1}\mu(K_j).$$

Therefore,

$$\Big|\int_\Omega f\,d\mu - l(f)\Big| \le \frac{1}{n}\sum_{j=0}^{n-1}\mu(K_j) - \frac{1}{n}\sum_{j=1}^{n}\mu(K_j) = \frac{1}{n}\mu(K_0 \setminus K_n) \le \frac{1}{n}\mu(\Omega).$$

Since n is arbitrary, we get

$$l(f) = \int_\Omega f\,d\mu.$$

We finished the proof in the case of a non-negative $l \in C(\Omega)'$: we proved that there is a non-negative $\mu \in \mathcal{A}_r(\Omega, \mathcal{B}(\Omega))$ so that $l(f) = \int_\Omega f\,d\mu$ for every $f \in C(\Omega)$.
Now let $l \in C(\Omega)'$ be real. For every non-negative $f \in C(\Omega)$ we define

$$l^+(f) = \sup\{l(g)\,|\,g \in C(\Omega), 0 \le g \le f \text{ in } \Omega\}.$$

Obviously, $l^+(f) \ge l(0) = 0$ and $l^+(f) \ge l(f)$.
If $0 \le g \le f$, then $l(g) \le |l(g)| \le \|l\|\|g\|_u \le \|l\|\|f\|_u$, and so

$$0 \le l^+(f) \le \|l\|\|f\|_u < +\infty.$$

For every $\lambda > 0$ and every non-negative $f \in C(\Omega)$ we have

$$\begin{aligned} l^+(\lambda f) &= \sup\{l(g)\,|\,g \in C(\Omega), 0 \le g \le \lambda f \text{ in } \Omega\} \\ &= \sup\{l(\lambda h)\,|\,h \in C(\Omega), 0 \le h \le f \text{ in } \Omega\} \\ &= \lambda \sup\{l(h)\,|\,h \in C(\Omega), 0 \le h \le f \text{ in } \Omega\} = \lambda l^+(f). \end{aligned}$$

If $f_1, f_2 \in C(\Omega)$ are non-negative, and $0 \le g_1 \le f_1$ and $0 \le g_2 \le f_2$, then $l(g_1) + l(g_2) = l(g_1 + g_2)$ and, since $0 \le g_1 + g_2 \le f_1 + f_2$, we get $l(g_1) + l(g_2) \le l^+(f_1 + f_2)$. Taking the supremum over g_1 and g_2, we find $l^+(f_1) + l^+(f_2) \le l^+(f_1 + f_2)$.
Now let $0 \le g \le f_1 + f_2$. We set $g_1 = \min\{f_1, g\}$, and then $0 \le g_1 \le f_1$ and $g_1 \le g$. If

we set $g_2 = g - g_1$, then it is easy to see that $0 \leq g_2 \leq f_2$ and, of course, $g = g_1 + g_2$. Thus $l(g) = l(g_1) + l(g_2) \leq l^+(f_1) + l^+(f_2)$, and, taking the supremum over g, we get $l^+(f_1 + f_2) \leq l^+(f_1) + l^+(f_2)$.
We conclude that

$$l^+(f_1 + f_2) = l^+(f_1) + l^+(f_2).$$

Until now, $l^+(f)$ is defined only for non-negative $f \in C(\Omega)$. Now, for any real $f \in C(\Omega)$ we consider $f^+ = \frac{1}{2}(|f| + f) \geq 0$ and $f^- = \frac{1}{2}(|f| - f) \geq 0$, so that $f = f^+ - f^-$. Then for every real $f \in C(\Omega)$ we define

$$l^+(f) = l^+(f^+) - l^+(f^-).$$

We observe that, if $f = g - h$ for any non-negative $g, h \in C(\Omega)$, then $f^+ + h = f^- + g$, and so

$$l^+(f^+) + l^+(h) = l^+(f^+ + h) = l^+(f^- + g) = l^+(f^-) + l^+(g).$$

Thus,

$$l^+(f) = l^+(f^+) - l^+(f^-) = l^+(g) - l^+(h).$$

If $f_1, f_2 \in C(\Omega)$ are real, then $f_1 + f_2 = (f_1^+ + f_2^+) - (f_1^- + f_2^-)$, and from the last identity we have

$$l(f_1 + f_2) = l(f_1^+ + f_2^+) - l(f_1^- + f_2^-) = l(f_1^+) + l(f_2^+) - l(f_1^-) - l(f_2^-) = l(f_1) + l(f_2).$$

If $f \in C(\Omega)$ is real and $\lambda \geq 0$, then

$$l^+(\lambda f) = l^+(\lambda f^+) - l^+(\lambda f^-) = \lambda l^+(f^+) - \lambda l^+(f^-) = \lambda l^+(f),$$

while if $\lambda < 0$, then

$$l^+(\lambda f) = l^+(|\lambda| f^-) - l^+(|\lambda| f^+) = |\lambda| l^+(f^-) - |\lambda| l^+(f^+) = \lambda l^+(f).$$

If $F = \mathbb{R}$, we have proved that $l^+ : C(\Omega) \to \mathbb{R}$ is linear.
If $F = \mathbb{C}$, then for every $f \in C(\Omega)$ we define

$$l^+(f) = l^+(\mathrm{Re}(f)) + i l^+(\mathrm{Im}(f))$$

and it is easy to see that $l^+ : C(\Omega) \to \mathbb{C}$ is linear. If $f \in C(\Omega)$ is real then

$$\begin{aligned} |l^+(f)| &= |l^+(f^+) - l^+(f^-)| \leq \max\{l^+(f^+), l^+(f^-)\} \leq \max\{\|l\| \|f^+\|_u, \|l\| \|f^-\|_u\} \\ &= \|l\| \|f\|_u. \end{aligned}$$

If $f \in C(\Omega)$ is complex, then there is $\lambda \in \mathbb{C}$ with $|\lambda| = 1$ so that $\lambda l^+(f) = |l^+(f)|$, and then we have

$$|l^+(f)| = \lambda l^+(f) = l^+(\lambda f) = \mathrm{Re}(l^+(\lambda f)) = l^+(\mathrm{Re}(\lambda f)) \leq \|l\| \|\mathrm{Re}(\lambda f)\|_u \leq \|l\| \|f\|_u.$$

So l^+ is a non-negative linear functional on $C(\Omega)$ with $\|l^+\| \leq \|l\|$.
We also define $l^- = l^+ - l : C(\Omega) \to F$. This is a bounded linear functional on $C(\Omega)$ and it is non-negative, since for every non-negative $f \in C(\Omega)$ we have $l^-(f) = l^+(f) - l(f) \geq 0$. So there are non-negative $\mu_1, \mu_2 \in \mathcal{A}_r(\Omega, \mathcal{B}(\Omega))$ so that $l^+(f) = \int_\Omega f \, d\mu_1$ and $l^-(f) = \int_\Omega f \, d\mu_2$ for every $f \in C(\Omega)$. Now we consider $\mu = \mu_1 - \mu_2$ and then $\mu \in \mathcal{A}_r(\Omega, \mathcal{B}(\Omega))$ is real and

$$l(f) = l^+(f) - l^-(f) = \int_\Omega f \, d\mu_1 - \int_\Omega f \, d\mu_2 = \int_\Omega f \, d\mu$$

for every $f \in C(\Omega)$.
At this point the proof is complete if $F = \mathbb{C}$ and l is real, or if $F = \mathbb{R}$ (and so l is automatically

real).
If $F = \mathbb{C}$ and l is complex, then $\mathrm{Re}(l)$ and $\mathrm{Im}(l)$ are real continous $\mathbb{R}$-linear functionals on $C(\Omega)$ and hence they are continuous $\mathbb{R}$-linear functionals on $C_{\mathbb{R}}(\Omega)$, the $\mathbb{R}$-linear space of the real continuous functions on Ω. So there are real $\mu_1, \mu_2 \in \mathcal{A}_r(\Omega, \mathcal{B}(\Omega))$ so that $\mathrm{Re}(l)(f) = \int_\Omega f\, d\mu_1$ and $\mathrm{Im}(l)(f) = \int_\Omega f\, d\mu_2$ for every real $f \in C(\Omega)$. So if we set $\mu = \mu_1 + i\mu_2$, then $\mu \in \mathcal{A}_r(\Omega, \mathcal{B}(\Omega))$ and for every real $f \in C(\Omega)$ we get

$$l(f) = \mathrm{Re}(l)(f) + i\,\mathrm{Im}(l)(f) = \int_\Omega f\, d\mu_1 + i \int_\Omega f\, d\mu_2 = \int_\Omega f\, d\mu.$$

So for every $f \in C(\Omega)$ we get

$$l(f) = l(\mathrm{Re}(f)) + il(\mathrm{Im}(f)) = \int_\Omega \mathrm{Re}(f)\, d\mu + i \int_\Omega \mathrm{Im}(f)\, d\mu = \int_\Omega f\, d\mu.$$

□

The main result of the theorem of F.Riesz-Radon-Banach-Kakutani is its "onto" part:
Let Ω be a compact, Hausdorff topological space. Then for every $l \in C(\Omega)'$ there is a unique $\mu \in \mathcal{A}_r(\Omega, \mathcal{B}(\Omega))$ so that

$$\|l\| = \|\mu\|, \qquad l(f) = \int_\Omega f\, d\mu \quad \text{for every } f \in C(\Omega).$$

If l is non-negative, i.e. $l(f) \geq 0$ for every non-negative $f \in C(\Omega)$, then μ is non-negative.
If l is real, i.e. $l(f) \in \mathbb{R}$ for every real $f \in C(\Omega)$, then μ is real.

3.6 The theorem of Hahn-Banach.

3.6.1 The analytic form.

Definition. *Let X be a linear space over F. Then $p : X \to \mathbb{R}$ is called* **positive-homogenuous and subadditive functional** *on X, if*
(i) $p(tx) = tp(x)$ for every $x \in X$ and every $t \geq 0$,
(ii) $p(x + y) \leq p(x) + p(y)$ for every $x, y \in X$.

Definition. *Let X be a linear space over F. Then $p : X \to \mathbb{R}$ is called* **seminorm** *on X, if*
(i) $p(\lambda x) = |\lambda| p(x)$ for every $x \in X$ and every $\lambda \in F$,
(ii) $p(x + y) \leq p(x) + p(y)$ for every $x, y \in X$.

Every seminorm is a positive-homogenuous and subadditive functional.

Lemma 3.4. *(i) If p is a positive-homogenuous and subadditive functional on X, then $p(0) = 0$, $-p(-x) \leq p(x)$ for every $x \in X$, and $p(x) - p(y) \leq p(x - y) \leq p(x) + p(-y)$ for every $x, y \in X$.*
(ii) If p is a seminorm on X, then $p(0) = 0$, $p(-x) = p(x)$ for every $x \in X$, and $|p(x) - p(y)| \leq p(x - y)$ for every $x, y \in X$. In particular, $p(x) \geq 0$ for every $x \in X$.

Proof. Exercise. □

Assume that a seminorm $p : X \to \mathbb{R}$ has the additional property: $p(x) = 0$ implies $x = 0$. Then, clearly, p is a norm on X.

The theorem of Hahn-Banach. *We consider $F = \mathbb{R}$. Let X be a linear space and Y be a linear subspace of X, let p be a positive-homogenuous and subadditive functional on X, and let l be a linear functional on Y. We assume that $l(y) \leq p(y)$ for every $y \in Y$. Then there is a linear functional L on X so that*
(i) $L(y) = l(y)$ for every $y \in Y$, i.e. L is an extension of l,
(ii) $L(x) \leq p(x)$ for every $x \in X$.

Proof. We consider the set $\mathcal{K}$ of all k with the properties:
(i) $k : D(k) \to \mathbb{R}$ is a linear functional on a linear subspace $D(k)$ of X,
(ii) k is an extension of l, i.e. $Y = D(l) \subseteq D(k)$ and $l(y) = k(y)$ for every $y \in Y$,
(iii) $k(z) \leq p(z)$ for every $z \in D(k)$.
Then l is an element of $\mathcal{K}$, and so $\mathcal{K}$ is non-empty. We define an order relation $\prec$ on $\mathcal{K}$ in the following way: $k_1 \prec k_2$ means that k_2 is an extension of k_1, i.e. $D(k_1) \subseteq D(k_2)$ and $k_1(z) = k_2(z)$ for every $z \in D(k_1)$. It is very easy to see that $\prec$ has the properties of an order relation.
Now let $\mathcal{M}$ be any totally ordered subset of $\mathcal{K}$. We consider the set

$$Z_0 = \bigcup_{k \in \mathcal{M}} D(k).$$

Since $D(l) \subseteq D(k)$ for every $k \in \mathcal{M}$, we see that $Y = D(l) \subseteq Z_0 \subseteq X$.
If $z_1, z_2 \in Z_0$, then there are $k_1, k_2 \in \mathcal{M}$ so that $z_1 \in D(k_1)$ and $z_2 \in D(k_2)$. Since one of k_1, k_2 is an extension of the other, we have that either $D(k_1) \subseteq D(k_2)$ or $D(k_2) \subseteq D(k_1)$. If $D(k_1) \subseteq D(k_2)$ then $z_1, z_2 \in D(k_2)$, and, since $D(k_2)$ is a linear subspace of X, we have that $z_1 + z_2 \in D(k_2)$, and so $z_1 + z_2 \in Z_0$. Obviously, the same is true if $D(k_2) \subseteq D(k_1)$.
If $z \in Z_0$ and $\lambda \in \mathbb{R}$, then there is $k \in \mathcal{M}$ so that $z \in D(k)$. Since $D(k)$ is a linear subspace of X, we have $\lambda z \in D(k)$, and so $\lambda z \in Z_0$.
Therefore, Z_0 is a linear subspace of X.
Now we take any $z \in Z_0$. Then $z \in D(k)$ for some $k \in \mathcal{M}$. If $z \in D(k')$ for any other $k' \in \mathcal{M}$, then, since one of k, k' is an extension of the other, we get that $k(z) = k'(z)$. So we may consider the function

$$k_0 : Z_0 \to \mathbb{R}$$

defined for every $z \in Z_0$ by

$$k_0(z) = k(z) \quad \text{for any } k \in \mathcal{M} \text{ with } z \in D(k).$$

We saw that, if $z_1, z_2 \in Z_0$, then there is $k \in \mathcal{M}$ so that $z_1, z_2 \in D(k)$, and hence

$$k_0(z_1 + z_2) = k(z_1 + z_2) = k(z_1) + k(z_2) = k_0(z_1) + k_0(z_2).$$

Similarly, if $z \in Z_0$ and $\lambda \in \mathbb{R}$, then there is $k \in \mathcal{M}$ so that $z \in D(k)$, and hence

$$k_0(\lambda z) = k(\lambda z) = \lambda k(z) = \lambda k_0(z).$$

Thus, k_0 is a linear functional on Z_0.
It is obvious that k_0 is an extension of l and that $k_0(z) \leq p(z)$ for every $z \in Z_0$.
Thus, $k_0 \in \mathcal{K}$. It is also clear that k_0 is, by its definition, an extension of every $k \in \mathcal{M}$, and so k_0 is an upper bound of $\mathcal{M}$.
Now Zorn's lemma implies that $\mathcal{K}$ has at least one maximal element. In other words, there is L with the properties (i), (ii) and (iii), and there is no k with the same properties which is a proper extension of L.
We shall prove that $D(L) = X$ and this will finish the proof.
Towards a contradiction, we assume that $D(L) \neq X$, and we take any $x_0 \in X \setminus D(L)$. We consider the linear subspace

$$Z = \{a + \lambda x_0 \,|\, a \in D(L), \lambda \in \mathbb{R}\}$$

of X. Then $D(L)$ as a proper linear subspace of Z. We shall define a linear functional $k : Z \to \mathbb{R}$ so that $k(a) = L(a)$ for every $a \in D(L)$, and $k(z) \leq p(z)$ for every $z \in Z$. This means that k is a proper extension of L with the properties (i), (ii) and (iii), and we shall arrive at a contradiction.
Now we take any $\lambda_0 \in \mathbb{R}$ and we consider $k : Z \to \mathbb{R}$ defined for every $a + \lambda x_0 \in Z$ (i.e. for every $a \in D(L)$ and every $\lambda \in \mathbb{R}$) by

$$k(a + \lambda x_0) = L(a) + \lambda \lambda_0.$$

Then it is very easy to see that k is a linear functional on Z and that $k(a) = L(a)$ for every $a \in D(L)$. So we only have to choose λ_0 so that $k(a + \lambda x_0) \le p(a + \lambda x_0)$ for every $a \in D(L)$ and every $\lambda \in \mathbb{R}$. This is equivalent to

$$L(a) + \lambda\lambda_0 \le p(a + \lambda x_0) \quad \text{for every } a \in D(L), \lambda \in \mathbb{R}.$$

So for $\lambda = 0$ we must have $L(a) \le p(a)$ for every $a \in D(L)$, which is true.
Then we must have

$$\lambda_0 \le \frac{1}{\lambda}\, p(a + \lambda x_0) - \frac{1}{\lambda}\, L(a) = p\Big(\frac{a}{\lambda} + x_0\Big) - L\Big(\frac{a}{\lambda}\Big) \quad \text{for every } a \in D(L), \lambda > 0$$

or, equivalently,

$$\lambda_0 \le p(a + x_0) - L(a) \quad \text{for every } a \in D(L).$$

Finally, we must have

$$\lambda_0 \ge \frac{1}{\lambda}\, p(a + \lambda x_0) - \frac{1}{\lambda}\, L(a) = -p\Big(\frac{a}{|\lambda|} - x_0\Big) + L\Big(\frac{a}{|\lambda|}\Big) \quad \text{for every } a \in D(L), \lambda < 0$$

or, equivalently,

$$\lambda_0 \ge -p(a - x_0) + L(a) \quad \text{for every } a \in D(L).$$

In other words, we must choose λ_0 so that

$$-p(a - x_0) + L(a) \le \lambda_0 \le p(a + x_0) - L(a) \quad \text{for every } a \in D(L).$$

The existence of such λ_0 is clearly equivalent to the inequality

$$\sup\{-p(a - x_0) + L(a) \,|\, a \in D(L)\} \le \inf\{p(a + x_0) - L(a) \,|\, a \in D(L)\},$$

and this is equivalent to

$$-p(a_1 - x_0) + L(a_1) \le p(a_2 + x_0) - L(a_2) \quad \text{for every } a_1, a_2 \in D(L).$$

But this last inequality is true, since

$$L(a_1) + L(a_2) = L(a_1 + a_2) \le p(a_1 + a_2) = p(a_1 - x_0 + a_2 + x_0) \le p(a_1 - x_0) + p(a_2 + x_0)$$

for every $a_1, a_2 \in D(L)$. $\square$

Definition. *Let X be a linear space over $\mathbb{C}$. Then $l : X \to \mathbb{C}$ is called* $\mathbb{R}$**-linear functional** *on X, if*

$$l(x_1 + x_2) = l(x_1) + l(x_2), \quad l(\lambda x) = \lambda l(x)$$

for every $x, x_1, x_2 \in X$ and every $\lambda \in \mathbb{R}$.

So the difference between a linear functional and a $\mathbb{R}$-linear functional is that the first satisfies $l(\lambda x) = \lambda l(x)$ for every $\lambda \in \mathbb{C}$ and the second satisfies the same equality for every $\lambda \in \mathbb{R}$. So, obviously, every linear functional is also a $\mathbb{R}$-linear functional.

Lemma 3.5. *Let X be a linear space over $\mathbb{C}$.*
(i) If $l : X \to \mathbb{C}$ is a linear functional, then its real part $\mathrm{Re}(l) : X \to \mathbb{R}$ *is a $\mathbb{R}$-linear functional, and*

$$l(x) = \mathrm{Re}(l)(x) - i\,\mathrm{Re}(l)(ix)$$

for every $x \in X$.
(ii) For every $\mathbb{R}$-linear functional $l_0 : X \to \mathbb{R}$, there is a unique linear functional $l : X \to \mathbb{C}$ so that $\mathrm{Re}(l) = l_0$ *in X.*

Proof. (i) Equating real parts in $l(x_1 + x_2) = l(x_1) + l(x_2)$ and $l(\lambda x) = \lambda l(x)$ with $\lambda \in \mathbb{R}$, we see that $\mathrm{Re}(l)$ is a $\mathbb{R}$-linear functional.
From $l(x) = \mathrm{Re}(l)(x) + i\,\mathrm{Im}(l)(x)$ we get

$$i\,\mathrm{Re}(l)(x) - \mathrm{Im}(l)(x) = il(x) = l(ix) = \mathrm{Re}(l)(ix) + i\,\mathrm{Im}(l)(ix),$$

and hence $\mathrm{Im}(l)(x) = -\mathrm{Re}(l)(ix)$. Thus $l(x) = \mathrm{Re}(l)(x) - i\,\mathrm{Re}(l)(ix)$ for every $x \in X$.
(ii) We consider $l : X \to \mathbb{C}$ defined for every $x \in X$ by

$$l(x) = l_0(x) - il_0(ix).$$

For every $x_1, x_2 \in X$ we get

$$l(x_1+x_2) = l_0(x_1+x_2) - il_0(ix_1+ix_2) = l_0(x_1)+l_0(x_2)-il_0(ix_1)-il_0(ix_2) = l(x_1)+l(x_2).$$

Also, if $\lambda = \mu + i\nu \in \mathbb{C}$, then

$$\begin{aligned} l(\lambda x) &= l(\mu x + i\nu x) = l(\mu x) + l(i\nu x) = l_0(\mu x) - il_0(i\mu x) + l_0(i\nu x) - il_0(-\nu x) \\ &= \mu l_0(x) - i\mu l_0(ix) + \nu l_0(ix) + i\nu l_0(x) = \mu l(x) + i\nu l(x) = \lambda l(x). \end{aligned}$$

Hence l is a linear functional and, clearly, $\mathrm{Re}(l) = l_0$.
If $\mathrm{Re}(l_1) = \mathrm{Re}(l_2)$ for two linear functionals l_1, l_2, then

$$l_1(x) = \mathrm{Re}(l_1)(x) - i\,\mathrm{Re}(l_1)(ix) = \mathrm{Re}(l_2)(x) - i\,\mathrm{Re}(l_2)(ix) = l_2(x)$$

for every $x \in X$, and so $l_1 = l_2$. □

The next result is the "complex" version of the theorem of Hahn-Banach.

The theorem of Bohnenblust-Sobczyk. *We consider $F = \mathbb{C}$. Let X be a linear space and Y be a linear subspace of X, let p be a seminorm on X, and let l be a linear functional on Y. We assume that $|l(y)| \le p(y)$ for every $y \in Y$. Then there is a linear functional L on X so that*
(i) $L(y) = l(y)$ for every $y \in Y$, i.e. L is an extension of l,
(ii) $|L(x)| \le p(x)$ for every $x \in X$.

Proof. We can obviously consider X (and hence also Y) as a linear space over $\mathbb{R}$. Lemma 3.5 implies that $\mathrm{Re}(l) : Y \to \mathbb{R}$ is a $\mathbb{R}$-linear functional on Y and that

$$l(y) = \mathrm{Re}(l)(y) - i\,\mathrm{Re}(l)(iy)$$

for every $y \in Y$.
We also have $\mathrm{Re}(l)(y) \le |l(y)| \le p(y)$ for every $y \in Y$.
The theorem of Hahn-Banach implies that there is a $\mathbb{R}$-linear functional $L_0 : X \to \mathbb{R}$ so that $L_0(y) = \mathrm{Re}(l)(y)$ for every $y \in Y$, and $L_0(x) \le p(x)$ for every $x \in X$.
Lemma 3.5 implies that there is a linear functional $L : X \to \mathbb{C}$ so that $\mathrm{Re}(L) = L_0$ in X. Then for every $y \in Y$ we have

$$L(y) = \mathrm{Re}(L)(y) - i\,\mathrm{Re}(L)(iy) = L_0(y) - iL_0(iy) = \mathrm{Re}(l)(y) - i\,\mathrm{Re}(l)(iy) = l(y),$$

and so L is an extension of l.
Finally, for every $x \in X$ there is $\lambda \in \mathbb{C}$ so that $|\lambda| = 1$ and $|L(x)| = \lambda L(x)$. Then

$$|L(x)| = \lambda L(x) = L(\lambda x) = \mathrm{Re}(L)(\lambda x) = L_0(\lambda x) \le p(\lambda x) = |\lambda| p(x) = p(x)$$

for every $x \in X$. □

3.6.2 The geometric form.

Definition. *Le X be a linear space over F, and let $A \subseteq X$.*
(i) We say that A **absorbs** *X, if for every $x \in X$ there is $r_0 > 0$ so that $[0, x] \subseteq r_0 A$.*
(ii) We say that A is **balanced**, *if $\lambda a \in A$ for every $a \in A$ and every $\lambda \in F$, $|\lambda| \leq 1$.*

It is obvious that 0 belongs to every A which absorbs X, and also to every balanced A.

Assume that A absorbs X and take any $x \in X$. Then there is $r_0 > 0$ so that $[0, x] \subseteq r_0 A$, i.e. $sx \in r_0 A$ for all s, $0 \leq s \leq 1$. This implies that $x \in rA$ for all $r \geq r_0$. Therefore, the set $\{r > 0 \,|\, x \in rA\}$ is a halfline in $(0, +\infty)$.

Definition. *Let X be a linear space, and assume that $A \subseteq X$ is convex and absorbs X. We consider the function $p_A : X \to [0, +\infty)$ defined for every $x \in X$ by*

$$p_A(x) = \inf\{r > 0 \,|\, x \in rA\}.$$

The function p_A is called **Minkowski functional** *of A.*

From the remarks before the definition, it is clear that $x \in rA$ for every $r > p_A(x)$. It is also clear that, if $0 < r < p_A(x)$, then $x \notin rA$.

Proposition 3.5. *Let X be a linear space, and assume that $A \subseteq X$ is convex and absorbs X. If p_A is the Minkowski functional of A, then*
(i) p_A is positive-homogenuous and subadditive on X.
(ii) if A is also balanced, then p_A is a seminorm on X.
(iii) $\{x \in X \,|\, p_A(x) < 1\} \subseteq A \subseteq \{x \in X \,|\, p_A(x) \leq 1\}$.

Proof. (i) If $t > 0$, then

$$\begin{aligned} p_A(tx) &= \inf\{r > 0 \,|\, tx \in rA\} = \inf\Big\{r > 0 \,\Big|\, x \in \frac{r}{t}\,A\Big\} = \inf\{ts > 0 \,|\, x \in sA\} \\ &= t \inf\{s > 0 \,|\, x \in sA\} = t p_A(x). \end{aligned}$$

Also, $0 \in rA$ for every $r > 0$, and so $p_A(0) = 0$. Thus, $p_A(tx) = t p_A(x)$ holds also for $t = 0$.
Now, take any $r > p_A(x)$ and any $s > p_A(y)$. Then $x \in rA$ and $y \in sA$, and so $\frac{1}{r}\,x \in A$ and $\frac{1}{s}\,y \in A$. Then the convexity of A implies that

$$\frac{1}{r+s}\,(x+y) = \frac{r}{r+s}\,\frac{1}{r}\,x + \frac{s}{r+s}\,\frac{1}{s}\,y \in A.$$

Therefore, $x + y \in (r+s)A$ and so $p_A(x+y) \leq r + s$. Since this holds for every $r > p_A(x)$ and every $s > p_A(y)$, we get $p_A(x+y) \leq p_A(x) + p_A(y)$.
(ii) If $\lambda \neq 0$, then

$$\begin{aligned} p_A(\lambda x) &= \inf\{r > 0 \,|\, \lambda x \in rA\} = \inf\Big\{r > 0 \,\Big|\, \frac{\lambda}{|\lambda|}\,\frac{|\lambda|}{r}\,x \in A\Big\} = \inf\Big\{r > 0 \,\Big|\, \frac{|\lambda|}{r}\,x \in A\Big\} \\ &= \inf\Big\{r > 0 \,\Big|\, x \in \frac{r}{|\lambda|}\,A\Big\} = |\lambda| \inf\{s > 0 \,|\, x \in sA\} = |\lambda| p_A(x). \end{aligned}$$

We saw in the proof of (i) that $p_A(\lambda x) = |\lambda| p_A(x)$ holds also for $\lambda = 0$.
(iii) If $p_A(x) < 1$, then $x \in 1A = A$. If $x \in A = 1A$, then $p_A(x) \leq 1$. □

Proposition 3.6. *Let X be a linear space and let $p : X \to \mathbb{R}$ be a positive-homogenuous and subadditive functional on X.*
(i) $B = \{x \in X \,|\, p(x) < 1\}$ and $C = \{x \in X \,|\, p(x) \leq 1\}$ are convex and they absorb X.
(ii) If p is a seminorm, then B, C are also balanced.
(iii) If A is convex and $B \subseteq A \subseteq C$, then $p_A = \max\{p, 0\}$. If p is a seminorm, then $p_A = p$.

Proof. (i) If $x, y \in B$ and $0 \leq t \leq 1$, then

$$p(tx + (1-t)y) \leq p(tx) + p((1-t)y) = tp(x) + (1-t)p(y) < t + (1-t) = 1.$$

Thus, $tx + (1-t)y \in B$ and so B is convex. The same argument shows that C is convex.
Let $x \in X$, and take any $r_0 > \max\{p(x), 0\}$. If $0 \leq s \leq 1$, then $p(\frac{s}{r_0}\, x) = \frac{s}{r_0}\, p(x) < 1$, and hence $\frac{s}{r_0}\, x \in B$. Thus $[0, x] \subseteq r_0 B$ and so B absorbs X. Now, C absorbs X, since $B \subseteq C$.
(ii) Let $x \in B$ and $|\lambda| \leq 1$. Then $p(\lambda x) = |\lambda| p(x) < 1$ and hence $\lambda x \in B$. So B is balanced, and the same argument shows that C is balanced.
(iii) A absorbs X, since B absorbs X. From proposition 3.5 we have that

$$\{x \in X \,|\, p_A(x) < 1\} \subseteq A \subseteq \{x \in X \,|\, p_A(x) \leq 1\}.$$

Thus, $p_A(x) < 1$ implies $p(x) \leq 1$. Also, $p(x) < 1$ implies $p_A(x) \leq 1$.
If $\lambda > \max\{p(x), 0\}$, then $p(\frac{1}{\lambda}\, x) < 1$, then $p_A(\frac{1}{\lambda}\, x) \leq 1$, and so $p_A(x) \leq \lambda$. Therefore, $p_A(x) \leq \max\{p(x), 0\}$.
If $\lambda > p_A(x)(\geq 0)$, then $p_A(\frac{1}{\lambda}\, x) < 1$, then $p(\frac{1}{\lambda}\, x) \leq 1$, and so $p(x) \leq \lambda$. Therefore, $p(x) \leq p_A(x)$. Since $0 \leq p_A(x)$, we get $\max\{p(x), 0\} \leq p_A(x)$.
Hence, $p_A(x) = \max\{p(x), 0\}$ for every $x \in X$.
If p is a seminorm, then $p(x) \geq 0$, and hence $p_A(x) = p(x)$ for every $x \in X$. $\square$

Definition. *Let X be a linear space, $A \subseteq X$ and $a \in X$. We say that A* **absorbs** *X with* **center** *a, if $A - a$ absorbs X.*

Clearly, if A absorbs X with center a, then $0 \in A - a$, and so $a \in A$.

It is easy to see that, if A absorbs X with center a, then, for every $b \in X$ and every $\lambda \in F$, we have that $A + b$ absorbs X with center $a + b$, and that λA absorbs X with center λa.

We know from Linear Algebra that, if $l \neq 0$ is a linear functional on X, then its null space (or kernel) $\mathrm{N}(l) = \{x \in X \,|\, l(x) = 0\}$ is a linear subspace of X of codimension equal to 1. Conversely, if Y is a linear subspace of X of codimension equal to 1 then there is a linear functional $l \neq 0$ on X so that $Y = \mathrm{N}(l)$. Moreover, any set of the form $Y + a$, where Y is a linear subspace of X of codimension equal to 1 and $a \in X$, is called **hyperplane** of X. Then it is easy to see that a subset of X is a hyperplane if and only if it is of the form $\{x \in X \,|\, l(x) = \lambda\}$, where $l \neq 0$ is a linear functional on X and $\lambda \in F$. Then we say that $\{x \in X \,|\, l(x) < \lambda\}$ and $\{x \in X \,|\, l(x) > \lambda\}$ are the **open halfspaces**, and $\{x \in X \,|\, l(x) \leq \lambda\}$ and $\{x \in X \,|\, l(x) \geq \lambda\}$ are the **closed halfspaces** determined by the hyperplane.

Theorem 3.5. *We consider $F = \mathbb{R}$. Let X be a linear space, and let $A \subseteq X$ be convex and absorb X with every $a \in A$ as center. If $b \notin A$, then there is a hyperplane of X which contains b and so that A is contained in one of the two open halfspaces determined by this hyperplane. Therefore, A is equal to the intersection of all open halfspaces which contain A.*

Proof. At first we assume that $0 \in A$. Then A is convex and absorbs X, and we consider the Minkowski functional p_A of A. Proposition 3.5 implies that $p_A(a) \leq 1$ for every $a \in A$, and also that $p_A(b) \geq 1$ for every $b \notin A$.
If $a \in A$, then $A - a$ absorbs X. Then there is $r > 0$ so that $a \in r(A - a)$, and hence $\frac{1+r}{r}\, a \in A$. Then $p_A(\frac{1+r}{r}\, a) \leq 1$ and hence $p_A(a) \leq \frac{r}{1+r} < 1$. Therefore, $p_A(a) < 1$ for every $a \in A$.
Now we take any $b \notin A$ (and so $b \neq 0$) and we consider the linear subspace $Y = \{\lambda b \,|\, \lambda \in \mathbb{R}\}$ of X, of dimension equal to 1, and the linear functional $l : Y \to \mathbb{R}$ defined for every $\lambda b \in Y$ by

$$l(\lambda b) = \lambda.$$

If $\lambda \leq 0$, then

$$l(\lambda b) = \lambda \leq 0 \leq p_A(\lambda b).$$

If $\lambda > 0$, then
$$l(\lambda b) = \lambda \leq \lambda p_A(b) = p_A(\lambda b).$$
Therefore, $l(y) \leq p_A(y)$ for every $y \in Y$.
The theorem of Hahn-Banach implies that there is a linear functional $L : X \to \mathbb{R}$ which is an extension of l and so that $L(x) \leq p_A(x)$ for every $x \in X$. Then $L(b) = l(b) = 1$ and $L(a) \leq p_A(a) < 1$ for every $a \in A$. So the hyperplane $\{x \in X \mid L(x) = 1\}$ contains b, and A is contained in the open halfspace $\{x \in X \mid L(x) < 1\}$.
Now, assume that $0 \notin A$. We take any $a_0 \in A$ and we consider the set $A_0 = A - a_0$. Then $0 \in A_0$, and A_0 is convex and absorbs X with every $a \in A_0$ as center.
Now we take any $b \notin A$, and we consider $b_0 = b - a_0 \notin A_0$. We have proved that there is a hyperplane L_0 which contains b_0 and so that A_0 is contained in one of the two open halfspaces determined by L_0. Then the hyperplane $L_0 + a_0$ contains b, and A is contained in one of the two open halfspaces determined by $L_0 + a_0$. □

Theorem 3.6. *We consider $F = \mathbb{R}$. Let X be a linear space, and let $A \subseteq X$ be convex and absorb X with some $a \in A$ as center. If $b \notin A$, then there is a hyperplane of X which contains b and so that A is contained in one of the two closed halfspaces determined by this hyperplane.*

Proof. We just repeat the proof of theorem 3.5, ommiting the part which proves that $p_A(a) < 1$ for every $a \in A$. It is enough that $p_A(a) \leq 1$ holds for every $a \in A$. □

Theorem 3.7. *We consider $F = \mathbb{R}$. Let X be a linear space, let $A \subseteq X$ be convex and absorb X with some (or every) $a \in A$ as center, let $B \subseteq X$ be convex, and $A \cap B = \emptyset$. Then there is a hyperplane of X so that A is contained in one of the two closed (or open) halfspaces determined by this hyperplane, and B is contained in the complementary closed halfspace.*

Proof. We consider the set $C = A - B$. Then C is convex and $0 \notin C$. Also, it is easy to show that C absorbs X with some $c \in C$ as center. Indeed, assume that A absorbs X with center $a_0 \in A$ and take any $b_0 \in B$. Then $A - b_0$ absorbs X with center $a_0 - b_0$. Since $A - b_0 \subseteq A - B$, we have that $A - B$ absorbs X with center $a_0 - b_0$.
Then theorems 3.5 and 3.6 imply that there is a hyperplane which contains 0 (i.e. a linear subspace of X of codimension equal to 1) so that C is contained in one of the two closed halfspaces determined by this hyperplane. In other words, there is a linear functional $l : X \to \mathbb{R}$, $l \neq 0$, so that $l(c) \leq 0$ for every $c \in C$. This implies that $l(a - b) \leq 0$, i.e. $l(a) \leq l(b)$ for every $a \in A$ and every $b \in B$.
Therefore,
$$\sup\{l(a) \mid a \in A\} \leq \inf\{l(b) \mid b \in B\}.$$
Now if we consider any $\lambda \in \mathbb{R}$ between these supremum and infimum, then
$$A \subseteq \{x \in X \mid l(x) \leq \lambda\}, \quad B \subseteq \{x \in X \mid l(x) \geq \lambda\}.$$
Now, assume that A absorbs X with center a. We take any $x \in X$ so that $l(x) > \lambda$. Then there is $t_0 > 0$ so that $x - a \in t_0(A - a)$ and so there is $a_0 \in A$ so that $x - a = t_0(a_0 - a)$, i.e. $x = t_0 a_0 + (1 - t_0)a$. Then
$$\lambda < l(x) = t_0 l(a_0) + (1 - t_0)l(a) \leq t_0 \lambda + (1 - t_0)l(a).$$
This excludes the case $t_0 \leq 1$. So $t_0 > 1$ and then we get $l(a) < \lambda$.
Therefore, if A absorbs X with every $a \in A$ as center, then
$$A \subseteq \{x \in X \mid l(x) < \lambda\}, \quad B \subseteq \{x \in X \mid l(x) \geq \lambda\}.$$
Now, the hyperplane we need is the $\{x \in X \mid l(x) = \lambda\}$. □

3.7 Implications of the theorem of Hahn-Banach.

The following is one of the fundamental results in the theory of normed spaces. It is very often called *theorem of Hahn-Banach*, but it is actually a corollary of the theorem of Hahn-Banach and its "complex" version, the theorem of Bohnenblust-Sobczyk.

Let X be a normed space, and Y be a subspace of X. Assume that $y' \in Y'$, that $x' \in X'$, and that x' is an extension of y', i.e. that $x'(y) = y'(y)$ for every $y \in Y$. Then it is very easy to show that $\|y'\| \leq \|x'\|$. Indeed,

$$\|y'\| = \sup_{y\in Y, \|y\|\leq 1} |y'(y)| = \sup_{y\in Y, \|y\|\leq 1} |x'(y)| \leq \sup_{x\in X, \|x\|\leq 1} |x'(x)| = \|x'\|.$$

We may say that *extensions have larger norms*.

Theorem 3.8. *Let X be a normed space, and Y be a subspace of X. Then for every $y' \in Y'$ there is $x' \in X'$ so that $x'(y) = y'(y)$ for every $y \in Y$, and $\|x'\| = \|y'\|$.*

Proof. We consider the seminorm $p : X \to \mathbb{R}$ defined for every $x \in X$ by

$$p(x) = \|y'\|\|x\|.$$

The linear functional y' on Y satisfies $|y'(y)| \leq p(y)$ for every $y \in Y$.
Let $F = \mathbb{C}$. Then the theorem of Bohnenblust-Sobczyk implies that there is a linear functional x' on X so that $x'(y) = y'(y)$ for every $y \in Y$ and

$$|x'(x)| \leq p(x) = \|y'\|\|x\|$$

for every $x \in X$. Therefore, $x' \in X'$ and $\|x'\| \leq \|y'\|$. Since $\|y'\| \leq \|x'|$ is trivially satisfied, we get that $\|x'\| = \|y'\|$.
Let $F = \mathbb{R}$. Since $y'(y) \leq p(y)$ for every $y \in Y$, the theorem of Hahn-Banach implies that there is a linear functional x' on X so that $x'(y) = y'(y)$ for every $y \in Y$ and

$$x'(x) \leq p(x) = \|y'\|\|x\|$$

for every $x \in X$. If we replace x with $-x$ in this inequality, we get $-\|y'\|\|x\| \leq x'(x)$ and so

$$|x'(x)| \leq \|y'\|\|x\|$$

for every $x \in X$. Thus, $\|x'\| \leq \|y'\|$, and hence $\|x'\| = \|y'\|$. $\square$

Definition. *Let X be a normed space. If $A \subseteq X$, we define*

$$A^{\perp} = \{x' \in X' \mid x'(x) = 0 \textit{ for every } x \in A\}.$$

If X is a space with inner product $\langle\cdot,\cdot\rangle$ and $A \subseteq X$, then $A^{\perp}$ has been defined in two different forms. The old form is

$$A^{\perp,\text{old}} = \{z \in X \mid \langle x, z\rangle = 0 \text{ for every } x \in A\} \subseteq X,$$

and the new form is

$$A^{\perp,\text{new}} = \{x' \in X' \mid x'(x) = 0 \text{ for every } x \in A\} \subseteq X'.$$

According to the theorem of F.Riesz, there is a conjugate-linear isometry T of X into X', given by:

$$T(z)(x) = \langle x, z\rangle \quad \text{for every } x \in X \text{ and every } z \in X.$$

Therefore, if we write $A^{\perp,\text{old}} = \{z \in X \mid \langle x, z\rangle = 0 \text{ for every } x \in A\}$, then we see that

$$\begin{aligned} T(A^{\perp,\text{old}}) &= \{T(z) \in X' \mid \langle x, z\rangle = 0 \text{ for every } x \in A\} \\ &= \{T(z) \in X' \mid T(z)(x) = 0 \text{ for every } x \in A\} \\ &= T(X) \cap A^{\perp,\text{new}} \subseteq A^{\perp,\text{new}}. \end{aligned}$$

If X is a Hilbert space, then T is onto X' and so

$$T(A^{\perp,\text{old}}) = A^{\perp,\text{new}}.$$

Proposition 3.7. *Let X be a normed space. If $A \subseteq X$, then $A^\perp$ is a closed subspace of X'.*

Proof. Exercise. □

Theorem 3.9. *Let X be a normed space, $x \in X$ and Y be a subspace of X. Then*

$$\max_{x' \in Y^\perp, \|x'\| \leq 1} |x'(x)| = \inf_{y \in Y} \|x - y\|.$$

Proof. For every $x' \in Y^\perp$ with $\|x'\| \leq 1$ and for every $y \in Y$ we have

$$|x'(x)| = |x'(x) - x'(y)| = |x'(x - y)| \leq \|x'\|\|x - y\| \leq \|x - y\|.$$

Hence

$$\sup_{x' \in Y^\perp, \|x'\| \leq 1} |x'(x)| \leq \inf_{y \in Y} \|x - y\|.$$

So it is enough to prove that there is $x' \in Y^\perp$ so that $\|x'\| \leq 1$ and $|x'(x)| = \inf_{y \in Y} \|x - y\|$.
If $x \in Y$, then $\inf_{y \in Y} \|x - y\| = 0$ and $|x'(x)| = 0$ for every $x' \in Y^\perp$. So in this case the proof is complete.
If $x \notin Y$, we consider the linear subspace Y_1 of X which is spanned by $Y \cup \{x\}$:

$$Y_1 = \{y + \lambda x \mid y \in Y, \lambda \in F\}.$$

We consider $y' : Y_1 \to F$ defined by

$$y'(y + \lambda x) = \lambda d$$

for every $y \in Y$ and every $\lambda \in F$, where $d = \inf_{y \in Y} \|x - y\|$. It is clear that y' is a linear functional on Y_1.
If $\lambda = 0$, then

$$|y'(y + \lambda x)| = |\lambda| d = 0 \leq \|y + \lambda x\|.$$

If $\lambda \neq 0$, then

$$|y'(y + \lambda x)| = |\lambda| d \leq |\lambda| \Big\| x - \Big(-\frac{1}{\lambda} y \Big) \Big\| = \|y + \lambda x\|.$$

Hence $y' \in Y_1'$ and $\|y'\| \leq 1$.
Now theorem 3.8 implies that there is $x' \in X'$ so that $x'(y + \lambda x) = \lambda d$ for every $y \in Y$ and every $\lambda \in F$, and $\|x'\| = \|y'\| \leq 1$. Now, $x'(y) = x'(y + 0x) = 0d = 0$ for every $y \in Y$, and so $x' \in Y^\perp$, and $|x'(x)| = |x'(0 + 1x)| = 1d = d$. □

Theorem 3.10. *Let X be a normed space, and $x \in X$. Then*

$$\|x\| = \max_{x' \in X', \|x'\| \leq 1} |x'(x)|.$$

Proof. This is a corollary of theorem 3.9. We consider the linear subspace $Y = \{0\}$, and then we have $\{0\}^\perp = X'$ and $\inf_{y \in \{0\}} \|x - y\| = \|x - 0\| = \|x\|$. □

Theorem 3.11. *Let X be a normed space, $A \subseteq X$ and $x \in X$. Then $x \in \operatorname{clspan}(A)$ if and only if $x'(x) = 0$ for every $x' \in A^{\perp}$.*

Proof. This is a corollary of theorem 3.9. We take $Y = \operatorname{clspan}(A)$, and then $A^{\perp} = Y^{\perp}$ and

$$\max_{x' \in A^{\perp}, \|x'\| \leq 1} |x'(x)| = \inf_{y \in Y} \|x - y\|.$$

Since Y is closed, we have that $x \in Y$ if and only if $\inf_{y \in Y} \|x - y\| = 0$ if and only if $x'(x) = 0$ for every $x' \in A^{\perp}$ with $\|x'\| \leq 1$ if and only if $x'(x) = 0$ for every $x' \in A^{\perp}$. $\square$

Proposition 3.8. *Let X be a normed space. If X' is separable, then X is separable.*

Proof. Let $\{x'_n \,|\, n \in \mathbb{N}\}$ be a countable dense subset of X'.
For each n we consider $x_n \in X$ so that $\|x_n\| = 1$ and $|x'_n(x_n)| \geq \frac{1}{2} \|x'_n\|$ and we define the set $A = \{x_n \,|\, n \in \mathbb{N}\}$.
Assume that there is $x \in X$ so that $x \notin \operatorname{clspan}(A)$. Then theorem 3.11 implies that there is $x' \in A^{\perp}$ so that $x'(x) \neq 0$. Hence, $x'(x_n) = 0$ for every n and $x' \neq 0$ and so $\|x'\| > 0$.
Since $\{x'_n \,|\, n \in \mathbb{N}\}$ is dense, there is n so that $\|x' - x'_n\| < \frac{1}{3} \|x'\|$. Then

$$\|x'_n\| \geq \|x'\| - \|x' - x'_n\| > 2\|x' - x'_n\|,$$

and so

$$\frac{1}{2} \|x'_n\| \leq |x'_n(x_n)| = |x'_n(x_n) - x'(x_n)| \leq \|x'_n - x'\| < \frac{1}{2} \|x'_n\|$$

and we have a contradiction.
Therefore, for every $x \in X$ we have that $x \in \operatorname{clspan}(A)$. Thus, for every $\epsilon > 0$ there are $n \in \mathbb{N}$ and $\lambda_1, \ldots, \lambda_n \in F$ so that

$$\|x - (\lambda_1 x_1 + \cdots + \lambda_n x_n)\| < \epsilon.$$

Now we take rational $\kappa_1, \ldots, \kappa_n \in F$ so that $|\lambda_j - \kappa_j| < \frac{\epsilon}{n\|x_j\|}$ for every j and we easily see that

$$\|x - (\kappa_1 x_1 + \cdots + \kappa_n x_n)\| < 2\epsilon.$$

So the countable set, whose elements are all linear combinations of elements of A with rational coefficients, is dense in X. $\square$

Corollary 3.1. *l^1 is not linearly isometric with $(l^{\infty})'$. In fact l^1 is not even topologically homeomorphic with $(l^{\infty})'$.*

Proof. l^1 is separable, so, if the two spaces are topologically homeomorphic, then $(l^{\infty})'$ is separable. But then proposition 3.8 implies that l^{∞} is also separable and this is not true. $\square$

Theorem 3.12. *Let X be a normed space, let Y be a subspace of X, and $x' \in X'$. Then*

$$\sup_{y \in Y, \|y\| \leq 1} |x'(y)| = \min_{z' \in Y^{\perp}} \|x' - z'\|.$$

Proof. For every $y \in Y$ with $\|y\| \leq 1$, and every $z' \in Y^{\perp}$ we have

$$|x'(y)| = |x'(y) - z'(y)| \leq \|x' - z'\| \|y\| \leq \|x' - z'\|.$$

Hence,

$$\sup_{y \in Y, \|y\| \leq 1} |x'(y)| \leq \inf_{z' \in Y^{\perp}} \|x' - z'\|.$$

So it is enough to prove that there is $z' \in Y^\perp$ so that $\sup_{y \in Y, \|y\| \le 1} |x'(y)| = \|x' - z'\|$.
We consider y' to be the restriction of x' on Y, and then

$$\|y'\| = \sup_{y \in Y, \|y\| \le 1} |y'(y)| = \sup_{y \in Y, \|y\| \le 1} |x'(y)|.$$

Theorem 3.8 implies that there is $x_1' \in X'$ so that $x_1'(y) = y'(y) = x'(y)$ for every $y \in Y$ and $\|x_1'\| = \|y'\|$. Now we take $z' = x' - x_1' \in X'$, and then

$$\|x' - z'\| = \|y'\| = \sup_{y \in Y, \|y\| \le 1} |x'(y)|.$$

Also $z'(y) = x'(y) - x_1'(y) = 0$ for every $y \in Y$, and hence $z' \in Y^\perp$. □

3.8 The second dual.

The **second dual** $X'' = (X')'$ of a normed space X is a Banach space with norm

$$\|x''\| = \sup_{x' \in X', \|x'\| \le 1} |x''(x')|$$

for every $x'' \in X''$.

Definition. *Let X be a normed space. For every $x \in X$ we consider the function $l_x : X' \to F$ defined for every $x' \in X'$ by*

$$l_x(x') = x'(x).$$

Theorem 3.13. *Let X be a normed space. For every $x \in X$ the function l_x is an element of X'' and the function*

$$J : X \to X'',$$

defined for every $x \in X$ by $J(x) = l_x$, is a linear isometry.

Proof. We have

$$l_x(x_1' + x_2') = (x_1' + x_2')(x) = x_1'(x) + x_2'(x) = l_x(x_1') + l_x(x_2'),$$

$$l_x(\lambda x') = (\lambda x')(x) = \lambda x'(x) = \lambda l_x(x')$$

for every $x', x_1', x_2' \in X'$ and every $\lambda \in F$. Thus, l_x is a linear functional on X'.
Theorem 3.10 implies that

$$\sup_{x' \in X', \|x'\| \le 1} |l_x(x')| = \sup_{x' \in X', \|x'\| \le 1} |x'(x)| = \|x\|.$$

This means that $l_x \in X''$ and $\|l_x\| = \|x\|$.
Now,

$$l_{x_1+x_2}(x') = x'(x_1 + x_2) = x'(x_1) + x'(x_2) = l_{x_1}(x') + l_{x_2}(x')$$

for every $x' \in X'$ and hence $l_{x_1+x_2} = l_{x_1} + l_{x_2}$. Also,

$$l_{\lambda x}(x') = x'(\lambda x) = \lambda x'(x) = \lambda l_x(x')$$

for every $x' \in X'$ and hence $l_{\lambda x} = \lambda l_x$. Thus, J is linear:

$$J(x_1 + x_2) = l_{x_1+x_2} = l_{x_1} + l_{x_2} = J(x_1) + J(x_2), \quad J(\lambda x) = l_{\lambda x} = \lambda l_x = \lambda J(x).$$

We saw that $\|J(x)\| = \|l_x\| = \|x\|$ for every $x \in X$, and so J is a linear isometry. □

Definition. *Let X be a normed space. The linear isometry $J : X \to X''$ defined in theorem 3.13 is called* **natural embedding** *of X into X''.*
If J is onto X'', then we say that X is **reflexive**.

Thus, if X is reflexive, then X is linearly isometric with X''. The converse is not true in general: there are normed spaces X which are linearly isometric with their second dual X'' but their natural embeddings J are not onto X''.

We observe that a necessary condition for a normed space X to be reflexive is its completeness. Indeed, $X'' = (X')'$ is a dual space, and so it is complete. Hence, if X is linearly isometric with X'', then X is also complete.

A second observation is the following. X_1 is complete, since it is a closed subspace of the Banach space X''. Also, $J(X)$ is dense in X_1, since X_1 is the closure of $J(X)$. Now, $J(X) \subseteq X''$ is linearly isometric with X, and we conclude that X_1 is a completion of X.

Proposition 3.9. *Every Hilbert space is reflexive.*

Proof. We consider the conjugate-linear isometry of the theorem of F. Riesz, $T : X \to X'$, given by

$$T(z)(x) = \langle x, z\rangle$$

for every $z \in X$ and $x \in X$.
We shall prove that the natural embedding $J : X \to X''$ is onto X''.
We take any $x'' \in X''$ and we consider the function $\overline{x'' \circ T} : X \to F$. It is easy to prove that this is linear:

$$\begin{aligned}\overline{x'' \circ T}(z_1 + z_2) &= \overline{x''(T(z_1 + z_2))} = \overline{x''(T(z_1) + T(z_2))} = \overline{x''(T(z_1)) + x''(T(z_2))} \\ &= \overline{x''(T(z_1))} + \overline{x''(T(z_2))} = \overline{x'' \circ T}(z_1) + \overline{x'' \circ T}(z_2),\end{aligned}$$

$$\overline{x'' \circ T}(\lambda z) = \overline{x''(T(\lambda z))} = \overline{x''(\overline{\lambda}\, T(z))} = \overline{\overline{\lambda}\, x''(T(z))} = \lambda\, \overline{x''(T(z))} = \lambda\, \overline{x'' \circ T}(z).$$

Also

$$|\overline{x'' \circ T}(z)| = |\overline{x''(T(z))}| = |x''(T(z))| \le \|x''\|\|T(z)\| = \|x''\|\|z\|,$$

and hence $\overline{x'' \circ T} \in X'$.
Since T is onto X', there is $x \in X$ so that $\overline{x'' \circ T} = T(x)$. Then,

$$x''(T(z)) = \overline{T(x)(z)} = \overline{\langle z, x\rangle} = \langle x, z\rangle = T(z)(x) = J(x)(T(z))$$

for every $z \in X$. Since T is onto X', the last equality implies that

$$x''(x') = J(x)(x')$$

for every $x' \in X'$. Therefore, $x'' = J(x)$, and so J is onto X''. □

Proposition 3.10. *If $1 < p < +\infty$ and (Ω, Σ, μ) is a measure space, then l^p and $L^p(\Omega, \Sigma, \mu)$ are reflexive.*

Proof. We consider q given by $\frac{1}{p} + \frac{1}{q} = 1$ and the linear isometries $T^{(p)} : l^q \to (l^p)'$ and $T^{(q)} : l^p \to (l^q)'$ given by

$$T^{(p)}(x)(y) = \sum_{k=1}^{+\infty} \lambda_k \mu_k = T^{(q)}(y)(x)$$

for every $x = (\lambda_k) \in l^q$ and $y = (\mu_k) \in l^p$.
We also consider the natural embedding $J : l^p \to (l^p)''$.

We take any $y'' \in (l^p)''$, and we consider the function $y'' \circ T^{(p)} : l^q \to F$. Then $y'' \circ T^{(p)}$ is a composition of linear functions and, hence, it is a linear functional on l^q. Also,

$$|y'' \circ T^{(p)}(x)| = |y''(T^{(p)}(x))| \leq \|y''\| \|T^{(p)}(x)\| = \|y''\| \|x\|$$

for every $x \in l^q$. Hence $y'' \circ T^{(p)} \in (l^q)'$.
Since $T^{(q)}$ is onto $(l^q)'$, there is $y \in l^p$ so that $y'' \circ T^{(p)} = T^{(q)}(y)$. Thus,

$$J(y)(T^{(p)}(x)) = T^{(p)}(x)(y) = T^{(q)}(y)(x) = y''(T^{(p)}(x))$$

for every $x \in l^q$. Since $T^{(p)}$ is onto $(l^p)'$, the last equality implies

$$J(y)(y') = y''(y')$$

for every $y' \in (l^p)'$. Thus $Jy = y''$ and so J is onto $(l^p)''$.
The proof of the reflexivity of $L^p(\Omega, \Sigma, \mu)$ is similar. □

Theorem 3.14. *Let X be a normed space, and Y be a closed subspace of X. If X is reflexive, then Y is reflexive.*

Proof. Take any $y'' \in Y''$. We consider $x'' : X' \to F$ defined for every $x' \in X'$ by

$$x''(x') = y''(x'|_Y),$$

where $x'|_Y \in Y'$ is the restriction of x' on Y. It is easy to see that x'' is linear. Also

$$|x''(x')| = |y''(x'|_Y)| \leq \|y''\| \|x'|_Y\| \leq \|y''\| \|x'\|$$

for every $x' \in X'$. Therefore, $x'' \in X''$ and, since X is reflexive, there is $x \in X$ so that $J(x) = x''$, where J is the natural embedding of X onto X''. This implies

$$y''(x'|_Y) = x''(x') = J(x)(x') = x'(x)$$

for every $x' \in X'$.
Now, we take any $x' \in Y^{\perp}$. Then $x'|_Y(y) = x'(y) = 0$ for every $y \in Y$, and so $x'|_Y = 0$. The last equality above implies that $x'(x) = y''(0) = 0$. So, theorem 3.9 implies $\inf_{y \in Y} \|x - y\| = 0$ and, since Y is closed, we get that $x \in Y$. Therefore,

$$y''(x'|_Y) = x'(x) = x'|_Y(x)$$

for every $x' \in X'$.
Theorem 3.8 says, in particular, that for every $y' \in Y'$ there is $x' \in X'$ so that $x'|_Y = y'$. So for every $y' \in Y'$ we have, by the last equality,

$$y''(y') = y'(x) = J'(x)(y'),$$

where J' is the natural embedding of Y into Y''. Thus $y'' = J'(x)$ with $x \in Y$, and we conclude that J' is onto Y''. □

3.9 The uniform boundedness principle.

Lemma 3.6. *Let X be a complete metric space with metric d. If C_n is a non-empty closed subset of X for every $n \in \mathbb{N}$, so that $C_{n+1} \subseteq C_n$ for every n and $\operatorname{diam}(C_n) \to 0$, then $\bigcap_{n=1}^{+\infty} C_n$ contains exactly one element.*

Proof. We take any $x_n \in C_n$. If $n \leq m$, then $x_n, x_m \in C_n$, and so $d(x_n, x_m) \leq \text{diam}(C_n) \to 0$ when $n, m \to +\infty$. Thus (x_n) is a Cauchy sequence, and, since X is complete, there is $x \in X$ so that $x_n \to x$. Since the sequence (x_n) is contained, after the index m, in the closed set C_m, we get that $x \in C_m$. Thus, $x \in \bigcap_{n=1}^{+\infty} C_n$.
If also $y \in \bigcap_{n=1}^{+\infty} C_n$, then $x, y \in C_n$ for every n. Therefore, $d(x, y) \leq \text{diam}(C_n)$ for every n, and so $d(x, y) = 0$. Thus $x = y$. □

The theorem of Baire. *Let X be a complete metric space. If U_n is an open and dense subset of X for every $n \in \mathbb{N}$, then $\bigcap_{n=1}^{+\infty} U_n$ is dense in X.*

Proof. We consider the set $U = \bigcap_{n=1}^{+\infty} U_n$, and we take any $r > 0$.
Since U_1 is dense, there is $x_1 \in B(x; r) \cap U_1$. Since $B(x; r) \cap U_1$ is open, there is $r_1 > 0$ so that $r_1 \leq \frac{1}{2} r$ and

$$\text{cl}(B(x_1; r_1)) \subseteq \overline{B}(x_1; r_1) \subseteq B(x; r) \cap U_1.$$

Since U_2 is dense, there is $x_2 \in B(x_1; r_1) \cap U_2$. Since $B(x_1; r_1) \cap U_2$ is open, there is $r_2 > 0$ so that $r_2 \leq \frac{1}{2} r_1 \leq \frac{1}{2^2} r$ and

$$\text{cl}(B(x_2; r_2)) \subseteq \overline{B}(x_2; r_2) \subseteq B(x_1; r_1) \cap U_2.$$

We continue inductively, and we see that for every $n \in \mathbb{N}$ there is a ball $B(x_n; r_n)$ so that $r_n \leq \frac{1}{2^n} r$, and so that

$$\text{cl}(B(x_{n+1}; r_{n+1})) \subseteq B(x_n; r_n) \cap U_{n+1} \subseteq B(x_n; r_n) \subseteq \text{cl}(B(x_n; r_n))$$

for every n. We aplly lemma 3.6 to the non-empty closed sets $\text{cl}(B(x_n; r_n))$ and we get that that there is some

$$y \in \bigcap_{n=1}^{+\infty} \text{cl}(B(x_n; r_n)).$$

Now, this implies that $y \in \text{cl}(B(x_1; r_1))$ and hence $y \in B(x; r)$. It also implies that $y \in \text{cl}(B(x_n; r_n))$ and hence $y \in U_n$ for every n. Therefore, $y \in B(x; r) \cap U$ and we conclude that U is dense in X. □

If A is a subset of a metric space Y with metric d, then A is **bounded** if the distances of the elements of A from any fixed element y_0 of Y are bounded, i.e.

$$\sup_{a \in A} d(a, y_0) < +\infty.$$

The uniform boundedness principle. *Let X be a complete metric space, let Y be a metric space with metric d, let $y_0 \in Y$, and let $\mathcal{F}$ be a collection of continuous functions $f : X \to Y$. We assume that*

$$\sup_{f \in \mathcal{F}} d(f(x), y_0) < +\infty \quad \textit{for every } x \in X.$$

Then there is a non-empty open $U \subseteq X$ and a $M \geq 0$ so that $d(f(x), y_0) \leq M$ for every $x \in U$ and every $f \in \mathcal{F}$, i.e.

$$\sup_{x \in U, f \in \mathcal{F}} d(f(x), y_0) < +\infty.$$

Proof. For each $n \in \mathbb{N}$ we define

$$P_n = \{x \in X \,|\, d(f(x), y_0) \leq n \text{ for every } f \in \mathcal{F}\} = \bigcap_{f \in \mathcal{F}} \{x \in X \,|\, d(f(x), y_0) \leq n\}.$$

It is easy to see that the continuity of the metric d and the continuity of each function f_n imply that $\{x \in X \,|\, d(f(x), y_0) \leq n\}$ is a closed set. Since P_n is the intersection of closed sets, it is closed.

Also, the assumption that $\sup_{f\in\mathcal{F}} d(f(x), y_0) < +\infty$ for every $x \in X$, implies that for every $x \in X$ there is $n \in \mathbb{N}$ so that $\sup_{f\in\mathcal{F}} d(f(x), y_0) \leq n$, and hence $x \in P_n$. Therefore $X = \bigcup_{n=1}^{+\infty} P_n$. If we define $U_n = X \setminus P_n$, then U_n is open and $\bigcap_{n=1}^{+\infty} U_n = \emptyset$.
Now, the theorem of Baire implies that there is $M \in \mathbb{N}$ so that U_M is *not* dense in X, i.e.

$$\mathrm{cl}(U_M) \neq X.$$

We consider the set $U = X \setminus \mathrm{cl}(U_M)$, Then U is open and non-empty, and $U \cap \mathrm{cl}(U_M) = \emptyset$. So $U \cap U_M = \emptyset$ and hence $U \subseteq P_M$. Of course, this implies that $d(f(x), y_0) \leq M$ for every $x \in U$ and every $f \in \mathcal{F}$. ☐

Regarding the uniform boundedness principle, $\sup_{f\in\mathcal{F}} d(f(x), y_0) < +\infty$ is equivalent with $\{f(x) \,|\, f \in \mathcal{F}\}$ being a bounded subset of the metric space Y. So we may say that the assumption that $\sup_{f\in\mathcal{F}} d(f(x), y_0) < +\infty$ for every $x \in X$ means that the collection of functions $\mathcal{F}$ is *pointwise bounded* in X. The result of the uniform boundedness principle, is that, under the assumption of its pointwise boundedness, the collection $\mathcal{F}$ is *uniformly bounded* in some open subset U of X. Of course, another central assumption is the completeness of X.

The next two theorems are just a few, among many, applications of the uniform boundedness principle. For both theorems we consider the metric space $Y = F$ with its usual euclidean metric. The role of $y_0 \in Y$ is played by $0 \in F$. So $d(f(x), y_0)$ is simply $|f(x)|$ for functions $f : X \to F$. In other words, we have the following special case of the uniform boundedness principle.

The uniform boundedness principle *Let X be a complete metric space, and let $\mathcal{F}$ be a collection of continuous functions $f : X \to F$. We assume that*

$$\sup_{f\in\mathcal{F}} |f(x)| < +\infty \quad \textit{for every } x \in X.$$

Then there is a non-empty open $U \subseteq X$ and a $M \geq 0$ so that $|f(x)| \leq M$ for every $x \in U$ and every $f \in \mathcal{F}$, i.e.

$$\sup_{x\in U, f\in\mathcal{F}} |f(x)| < +\infty.$$

Theorem 3.15. *Let X be a Banach space and let $\mathcal{F} \subseteq X'$ satisfy $\sup_{x'\in\mathcal{F}} |x'(x)| < +\infty$ for every $x \in X$. Then $\sup_{x'\in\mathcal{F}} \|x'\| < +\infty$.*

Proof. We apply the uniform boundedness principle to the collection X' of functions $x' : X \to F$, and we get that there is a non-empty open $U \subseteq X$ and a $M \geq 0$ so that $|x'(x)| \leq M$ for every $x' \in \mathcal{F}$ and every $x \in U$.
Now we take any $x_0 \in U$ and then there is $R > 0$ so that $B(x_0; R) \subseteq U$. Therefore, we have that $|x'(x)| \leq M$ for every $x' \in \mathcal{F}$ and every $x \in B(x_0; R)$.
Take any $x' \in \mathcal{F}$, any $x \neq 0$ and any $t > 1$. Then $x_0 \in B(x_0; R)$ and $x_0 + \frac{R}{t\|x\|}\, x \in B(x_0; R)$. Hence

$$|x'(x)| = \frac{t\|x\|}{R} \Big| x'\Big(\frac{R}{t\|x\|}\, x\Big)\Big| = \frac{t\|x\|}{R} \Big| x'\Big(x_0 + \frac{R}{t\|x\|}\, x\Big) - x'(x_0)\Big| \leq \frac{t\|x\|}{R}\, 2M.$$

Since $t > 1$ is arbitrary, we get

$$|x'(x)| \leq \frac{2M}{R}\, \|x\|.$$

This is true also for $x = 0$, and hence $\|x'\| \leq \frac{2M}{R}$ for every $x' \in \mathcal{F}$. ☐

Since $\|x'\| = \sup_{x\in X, \|x\|\leq 1} |x'(x)| = \sup_{x\in\overline{B}(0;1)} |x'(x)|$, theorem 3.15 says that *if X is a Banach space and the collection $\mathcal{F} \subseteq X'$ is pointwise bounded in X, then $\mathcal{F}$ is uniformly bounded in the closed unit ball $\overline{B}(0;1)$ of X.*

Theorem 3.16. *Let X be a normed space and let $\mathcal{F} \subseteq X$ satisfy $\sup_{x\in\mathcal{F}} |x'(x)| < +\infty$ for every $x' \in X'$. Then $\sup_{x\in\mathcal{F}} \|x\| < +\infty$.*

Proof. We consider the natural embedding $J : X \to X''$ and the collection $J(\mathcal{F}) \subseteq X''$ of the functions $J(x) : X' \to F$ for every $x \in \mathcal{F}$. We apply the previous theorem for the Banach space X' and for the collection $J(\mathcal{F}) \subseteq (X')'$, since

$$\sup_{J(x)\in J(\mathcal{F})} |J(x)(x')| = \sup_{x\in\mathcal{F}} |x'(x)| < +\infty$$

for every $x' \in X'$.
We conclude that $\sup_{x\in\mathcal{F}} \|x\| = \sup_{J(x)\in J(\mathcal{F})} \|J(x)\| < +\infty$. □

3.10 Weak convergence and weak-star convergence.

Definition. *Let X be a normed space.*
(i) We say that the sequence (x_n) in X **converges weakly** *to $x \in X$, if $x'(x_n) \to x'(x)$ for every $x' \in X'$. Then we write*

$$x_n \overset{w}{\to} x.$$

(ii) We say that the sequence (x'_n) in X' **converges weakly$*$** *to $x' \in X'$, if $x'_n(x) \to x'(x)$ for every $x \in X$. Then we write*

$$x'_n \overset{w*}{\to} x'.$$

Of course, when we write $x_n \to x$ or $x'_n \to x'$ we mean $\|x_n - x\| \to 0$ or $\|x'_n - x'\| \to 0$, respectively. To stress the difference between the various notions of convergence, we may say that (x_n) **converges strongly** to x, if $x_n \to x$, and we may say that (x'_n) **converges strongly** to x', if $x'_n \to x'$. This terminology is justified by the:

Proposition 3.11. *Let X be a normed space.*
(i) In X: if $x_n \to x$, then $x_n \overset{w}{\to} x$.
(ii) In X': if $x'_n \to x'$, then $x'_n \overset{w*}{\to} x'$.

Proof. (i) If $x_n \to x$, then for every $x' \in X'$ we have $|x'(x_n) - x'(x)| \leq \|x'\|\|x_n - x\| \to 0$. Hence $x_n \overset{w}{\to} x$.
(ii) If $x'_n \to x'$, then for every $x \in X$ we have $|x'_n(x) - x'(x)| \leq \|x'_n - x'\|\|x\| \to 0$. Hence $x'_n \overset{w*}{\to} x'$. □

Proposition 3.12. *Let X be a normed space.*
(i) In X: if $x_n \overset{w}{\to} x$, $y_n \overset{w}{\to} y$ and $\lambda_n \to \lambda$, then $x_n + y_n \overset{w}{\to} x + y$ and $\lambda_n x_n \overset{w}{\to} \lambda x$.
(ii) In X': if $x'_n \overset{w*}{\to} x'$, $y'_n \overset{w*}{\to} y'$ and $\lambda_n \to \lambda$, then $x'_n + y'_n \overset{w*}{\to} x' + y'$ and $\lambda_n x'_n \overset{w*}{\to} \lambda x'$.
(iii) In X: if $x_n \overset{w}{\to} y$ and $x_n \overset{w}{\to} z$, then $y = z$.
(iv) In X': if $x'_n \overset{w*}{\to} y'$ and $x'_n \overset{w*}{\to} z'$, then $y' = z'$.

Proof. (i) For every $x' \in X'$ we have

$$x'(x_n + y_n) = x'(x_n) + x'(y_n) \to x'(x) + x'(y) = x'(x + y),$$

$$x'(\lambda_n x_n) = \lambda_n x'(x_n) \to \lambda x'(x) = x'(\lambda x).$$

Hence $x_n + y_n \overset{w}{\to} x + y$ and $\lambda_n x_n \overset{w}{\to} \lambda x$.
(ii) Similar to the proof of (i).
(iii) For every $x' \in X'$ we have $x'(x_n) \to x'(y)$ and $x'(x_n) = x'(z)$, and hence $x'(y - z) = x'(y) - x'(z) = 0$. Theorem 3.10 implies that $y - z = 0$ and so $y = z$.
(iv) For every $x \in X$ we have $x'_n(x) \to y'(x)$ and $x'_n(x) = z'(x)$, and hence $y'(x) = z'(x)$. Therefore, $y' = z'$. □

We must stress the difference between the natures of (iii) and (iv) of the last proposition, i.e. the uniqueness of the weak limit and the weak$*$ limit, which is reflected in the difference between the difficulties of their proofs.

Example 3.10.1. If $1 < p \leq +\infty$, then $e_n \overset{\mathrm{w}}{\to} 0$ in l^p and also in c, c_0. But (e_n) does not have a weak limit in l^1.
In all cases the norms of the e_n are equal to 1, and (e_n) does not converge since the norms of the differences $e_n - e_m$ are constant and $\neq 0$.

Example 3.10.2. If $\{a_n \,|\, n \in \mathbb{N}\}$ is an orthonormal set in an inner product space X, then $a_n \overset{\mathrm{w}}{\to} 0$ in X.

Theorem 3.17. *Let X be a normed space, and $x_n \overset{\mathrm{w}}{\to} x$ in X. Then $\sup_n \|x_n\| < +\infty$ and $\|x\| \leq \liminf_{n\to+\infty} \|x_n\|$.*

Proof. For every $x' \in X'$ the sequence $(x'(x_n))$ converges to $x'(x)$ in F, and so it is bounded. Theorem 3.16 implies that $\sup_n \|x_n\| < +\infty$.
Let $q = \liminf_{n\to+\infty} \|x_n\|$. Then there is a subsequence (x_{n_k}) so that $\|x_{n_k}\| \to q$. For every $x' \in X'$ with $\|x'\| \leq 1$ we have $|x'(x_{n_k})| \leq \|x_{n_k}\|$. Since $x'(x_{n_k}) \to x'(x)$, we find $|x'(x)| \leq q$. Now theorem 3.10 implies that $\|x\| = \max_{x' \in X', \|x'\| \leq 1} |x'(x)| \leq q$. □

Theorem 3.18. *Let X be a Banach space and $x'_n \overset{\mathrm{w}*}{\to} x'$ in X'. Then $\sup_n \|x'_n\| < +\infty$ and $\|x'\| \leq \liminf_{n\to+\infty} \|x'_n\|$.*

Proof. For every $x \in X$ the sequence $(x'_n(x))$ converges to $x'(x)$ in F, and so it is bounded. Theorem 3.15 implies that $\sup_n \|x'_n\| < +\infty$.
Let $q = \liminf_{n\to+\infty} \|x'_n\|$. Then there is a subsequence (x'_{n_k}) so that $\|x'_{n_k}\| \to q$. For every $x \in X$ with $\|x\| \leq 1$ we have $|x'_{n_k}(x)| \leq \|x'_{n_k}\|$. Since $x'_{n_k}(x) \to x'(x)$, we find $|x'(x)| \leq q$. Therefore, $\|x'\| = \sup_{x \in X, \|x\| \leq 1} |x'(x)| \leq q$. □

Chapter 4

Weak topologies 1

4.1 Generalities about topological spaces.

4.1.1 Open sets and closed sets.

Definition. *Let A be a non-empty set, and $\mathcal{T}$ be a collection of subsets of A, with the properties:*
(i) $\emptyset \in \mathcal{T}$, $A \in \mathcal{T}$.
(ii) The union of any elements of $\mathcal{T}$ is an element of $\mathcal{T}$. In other words, if $U_i \in \mathcal{T}$ for every $i \in I$, then $\bigcup_{i\in I} U_i \in \mathcal{T}$.
(iii) The intersection of any finitely many elements of $\mathcal{T}$ is an element of $\mathcal{T}$. In other words, if $U_1, \dots, U_n \in \mathcal{T}$, then $\bigcap_{i=1}^n U_i \in \mathcal{T}$.
Then $\mathcal{T}$ is called **topology** *of A, and the elements of $\mathcal{T}$ are called* **open** *(with respect to $\mathcal{T}$) subsets of A. Finally, A equipped with a topology is called* **topological space**.

If A a topological space, then (ii) says that the union of any open subsets of A is an open subset of A, and (iii) says that the intersection of any finitely many open subsets of A is an open subset of A.

Example 4.1.1. Let A be a non-empty set. Then $\{\emptyset, A\}$ is a topology of A.

Example 4.1.2. Let A be a non-empty set. Then $\mathcal{P}(A)$, the collection of *all* subsets of A, is a topology of A.

Example 4.1.3. Let A be a metric space with metric d. Then

$$\mathcal{T} = \{U \subseteq A \,|\, U \text{ is open with respect to } d\}$$

is a topology in A. In this case we say that the topology $\mathcal{T}$ is **induced** by d.
To be more precise, $U \subseteq A$ is open with respect to d if for every $x \in U$ there is a radius $r > 0$ so that the ball $B(x, r) = \{y \in A \,|\, d(y, x) < r\}$ is included in U.
It is easy to see that $\emptyset$ and A are open with respect to d.
Now assume that U_i is open with respect to d for every $i \in I$, and take any $x \in \bigcup_{i\in I} U_i$. Then $x \in U_{i_0}$ for some $i_0 \in I$, and then there is $r > 0$ so that $B(x; r) \subseteq U_{i_0}$. Hence $B(x; r) \subseteq \bigcup_{i\in I} U_i$ and we have that $\bigcup_{i\in I} U_i$ is open with respect to d.
Finally, assume that $U_1, \dots, U_n$ are open with respect to d, and take any $x \in \bigcap_{i=1}^n U_i$. Then for every $i = 1, \dots, n$ there is $r_i > 0$ so that $B(x; r_i) \subseteq U_i$. If we take $r = \min\{r_1, \dots, r_n\} > 0$, then $B(x; r) \subseteq B(x; r_i) \subseteq U_i$ for every $i = 1, \dots, n$, and hence $B(x; r) \subseteq \bigcap_{i=1}^n U_i$. So $\bigcap_{i=1}^n U_i$ is open with respect to d.
Every ball $B(x; r)$ is open with respect to d. Indeed, take any $y \in B(x; r)$. Then $d(y, x) < r$, and we consider $s = r - d(y, x) > 0$. Now, if $z \in B(y; s)$, then

$$d(z, x) \le d(z, y) + d(y, x) < s + d(y, x) = r$$

and hence $z \in B(x; r)$. Thus $B(y; s) \subseteq B(x; r)$.

Definition. *Let A be a topological space, and $F \subseteq A$. We say that F is* **closed**, *if $A \setminus F$ is open.*

Proposition 4.1. *Let A be a topological space. Then*
(i) $\emptyset$ and A are closed.
(ii) The intersection of any closed subsets of A is a closed subset of A.
(iii) The union of any finitely many closed subsets of A is a closed subset of A.

Proof. The proof is a trivial corollary of the definition of closed set, of the properties of open sets, and of the laws of de Morgan for the complements of unions and intersections. □

Definition. *Let A be a topological space, and $x \in A$. Every open set containing x is called* **open neighborhood** *of x.*

Definition. *Let A be a topological space, and $M \subseteq A$. Then the set $\bigcap\{F \mid F \supseteq M \text{ is closed}\}$ is called* **closure** *of M and it is denoted* $\mathrm{cl}(M)$.

Proposition 4.2. *Let A be a topological space, and $M \subseteq A$.*
(i) $\mathrm{cl}(M)$ is the smallest closed subset of A which includes M.
(ii) $x \in \mathrm{cl}(M)$ if and only if $U \cap M \neq \emptyset$ for every open neighborhood U of x.

Proof. (i) $\mathrm{cl}(M)$ is the intersection of closed sets which include M, and so it closed and includes M. Also, if F is closed and includes M, then $\mathrm{cl}(M) \subseteq F$. So $\mathrm{cl}(M)$ is the smallest closed subset of A which includes M.
(ii) Let $x \in \mathrm{cl}(M)$, and take any open neighborhod U of x. Then $A \setminus U$ is closed and, since $x \notin A \setminus U$, we have that $\mathrm{cl}(M)$ is not included in $A \setminus U$. According to (i), M is not included in $A \setminus U$, and hence $U \cap M \neq \emptyset$.
Conversely, assume that $U \cap M \neq \emptyset$ for every open neighborhood U of x. We take any closed $F \supseteq M$, and then $A \setminus F$ is open and $(A \setminus F) \cap M = \emptyset$. Therefore, $x \notin A \setminus F$ and so $x \in F$. We conclude that $x \in \mathrm{cl}(M)$. □

Definition. *Let A be a topological space with topology $\mathcal{T}$, and let (x_n) be a sequence in A. We say that (x_n)* **converges (with respect to $\mathcal{T}$)** *to $x \in A$, if for every open neighborhood U of x there is n_0 so that $x_n \in U$ for every $n \geq n_0$.*
Then we say that x is a **limit** *of (x_n), and we write $x_n \to x$.*

4.1.2 Continuous functions.

Definition. *Let A, B be two topological spaces, $M \subseteq A$, and $f : M \to B$.*
(i) We say that f is **continuous** *at $x \in M$ if for every open $V \subseteq B$ such that $f(x) \in V$ there is an open $U \subseteq A$ so that $x \in U$ and $f(U \cap M) \subseteq V$, i.e. so that $x \in U$, and $f(y) \in V$ for every $y \in U \cap M$.*
(ii) We say that f is **continuous** *in M if it is continuous at every $x \in M$.*

Proposition 4.3. *Let A, B be two topological spaces, $M \subseteq A$, and $f : M \to B$. Then f is continuous in M if and only if for every open $V \subseteq B$ there is an open $U \subseteq A$ so that $f^{-1}(V) = U \cap M$.*

Proof. Let f be continuous in M, and let $V \subseteq B$ be open. Then for every $x \in f^{-1}(V)$ we have $f(x) \in V$, and so there is an open $U_x \subseteq A$ such that $x \in U_x$ and $f(U_x \cap M) \subseteq V$. Then $U = \bigcup_{x \in f^{-1}(V)} U_x \subseteq A$ is open, and it is easy to see that $f^{-1}(V) = U \cap M$. Indeed, if $y \in f^{-1}(V)$, then $y \in U_y \cap M$ and hence $y \in U \cap M$. Also, if $y \in U \cap M$, then $y \in U_x \cap M$ for some $x \in f^{-1}(V)$. Then $f(y) \in V$ and hence $y \in f^{-1}(V)$.
Conversely, take any $x \in M$ and any open $V \subseteq B$ so that $f(x) \in V$. Then there is an open $U \subseteq A$ so that $f^{-1}(V) = U \cap M$. Then $x \in U$ and $f(U \cap M) \subseteq V$, and so f is continuous at x. □

Proposition 4.4. *(i) Let A, B, C be topological spaces, $M \subseteq A$, $N \subseteq B$, $f : M \to N$ and $g : N \to C$. If f is continuous at $x \in M$ and g is continuous at $f(x) \in N$, then $g \circ f$ is continuous at x.*
(ii) Let A be a topological space, $M \subseteq A$, $f, g : M \to \mathbb{R}^n$ and $\lambda \in \mathbb{R}$. If $\mathbb{R}^n$ has the topology induced by the euclidean metric and f, g are continuous at $x \in M$, then $f + g, \lambda f$ are continuous at x.

Proof. (i) Let $W \subseteq C$ be open and $(g \circ f)(x) = g(f(x)) \in W$. Then there is an open $V \subseteq B$ so that $f(x) \in V$ and $g(V \cap N) \subseteq W$. Then there is an open $U \subseteq A$ so that $x \in U$ and $f(U \cap M) \subseteq V$. Then $f(U \cap M) \subseteq V \cap N$ and hence

$$(g \circ f)(U \cap M) = g(f(U \cap M)) \subseteq g(V \cap N) \subseteq W.$$

Thus $g \circ f$ is continuous at x.
(ii) Let $V \subseteq \mathbb{R}^n$ be open and $f(x) + g(x) \in V$. Then there is $r > 0$ so that $z \in V$ for every $z \in \mathbb{R}^n$ with $\|z - (f(x) + g(x))\| < r$. Now, there is an open $U_1 \subseteq A$ so that $x \in U_1$ and $\|f(y) - f(x)\| < \frac{r}{2}$ for every $y \in U_1 \cap M$. Also, there is an open $U_2 \subseteq A$ so that $x \in U_2$ and $\|g(y) - g(x)\| < \frac{r}{2}$ for every $y \in U_2 \cap M$. Then $U = U_1 \cap U_2 \subseteq A$ is open, and $x \in U$, and for every $y \in U \cap M$ we have

$$\|(f(y) + g(y)) - (f(x) + g(x))\| \leq \|f(y) - f(x)\| + \|g(y) - g(x)\| < \frac{r}{2} + \frac{r}{2} = r,$$

and hence $f(y) + g(y) \in V$. So $f + g$ is continuous at x.
The proof for λf is similar. $\square$

Definition. *Let A, B be topological spaces, and $f : A \to B$. We say that f is a* **homeomorphism** *of A onto B, if f is one-to-one in A and onto B, and f is continuous in A and f^{-1} is continuous in B. In this case we say that A, B are* **homeomorphic**.

If A, B are homeomorphic topological spaces, and $f : A \to B$ is a homeomorphism of A onto B, then we may *identify* the two spaces: we identify every $a \in A$ with the corresponding $b = f(a) \in B$ and, conversely, we identify every $b \in B$ with the corresponding $a = f^{-1}(b) \in A$. Then every open $U \subseteq A$ is identified with the open $V = f(U) \subseteq B$ and, conversely, every open $V \subseteq B$ is identified with the open $U = f^{-1}(V) \subseteq A$.

4.1.3 Compact sets.

Definition. *Let A be a topological space. We say that A is a* **Hausdorff** *topological space, if for every $x_1, x_2 \in A$, $x_1 \neq x_2$, there are disjoint open $U_1, U_2 \subseteq A$ so that $x_1 \in U_1$ and $x_2 \in U_2$.*

Proposition 4.5. *Every metric space is Hausdorff.*

Proof. If d is the metric of A and $x_1, x_2 \in A$, $x_1 \neq x_2$, we take $r = \frac{1}{2} d(x_1, x_2) > 0$, and then $B(x_1; r) \cap B(x_2; r) = \emptyset$. The balls $B(x_1; r), B(x_2; r)$ are open with respect to d. $\square$

Proposition 4.6. *Let A be a Hausdorff topological space. If a sequence in A has a limit, then this limit is unique.*

Proof. Let $x_n \to y$ and $x_n \to z$. If $y \neq z$, then there are disjoint open $U, V \subseteq A$ so that $y \in U$ and $z \in V$. But then there is n_0 so that $x_n \in U$ and $x_n \in V$ for every $n \geq n_0$, and obviously this is impossible. $\square$

Definition. *Let A be a topological space, and $K \subseteq A$. We say that a collection $\{U_i \,|\, i \in I\}$ of open subsets of A is an* **open cover** *of K, if $K \subseteq \bigcup_{i \in I} U_i$.*

Definition. *Let A be a topological space, and $K \subseteq A$. We say that K is* **compact**, *if for every open cover of K there is a finite subcover of K. More precisely, K is compact, if for every open cover $\{U_i \,|\, i \in I\}$ of K there are $i_1, \ldots, i_n \in I$ so that $\{U_{i_k} \,|\, 1 \leq k \leq n\}$ is also an open cover of K.*

Proposition 4.7. *Let A be a topological space.*
(i) If $K \subseteq A$ is compact and A is Hausdorff, then K is closed.
(ii) If $K \subseteq A$ is compact and $K' \subseteq K$ is closed, then K' is compact.

Proof. (i) Take any $x \in A \setminus K$. For every $z \in K$ there are disjoint open $U_z, V_z \subseteq A$ so that $z \in U_z$ and $x \in V_z$. Then $\{U_z \,|\, z \in K\}$ is an open cover of K. Since K is compact, there are $z_1, \ldots, z_n \in K$ so that $K \subseteq U_{z_1} \cup \cdots \cup U_{z_n}$. Then $V_{z_1} \cap \cdots \cap V_{z_n}$ is open, it is included in $A \setminus K$, and contains x. Therefore, $A \setminus K$ is open, and so K is closed.
(ii) Let $\{U_i \,|\, i \in I\}$ be any open cover of K'. Then $\{U_i \,|\, i \in I\} \cup \{A \setminus K'\}$ is an open cover of K. Since K is compact, there are $i_1, \ldots, i_n \in I$ so that $K \subseteq \left(\bigcup_{k=1}^n U_{i_k}\right) \cup (A \setminus K')$. Then $K' \subseteq \bigcup_{k=1}^n U_{i_k}$, and so K' is compact. $\square$

Proposition 4.8. *Let A, B be topological spaces, $M \subseteq A$, and let $f : M \to B$ be continuous. If $K \subseteq M$ is compact, then $f(K)$ is compact.*

Proof. Let $\{V_i \,|\, i \in I\}$ be any open cover of $f(K)$, i.e. $f(K) \subseteq \bigcup_{i \in I} V_i$. Since each $V_i \subseteq B$ is open and f is continuous, proposition 4.3 implies that there is a corresponding open $U_i \subseteq A$ so that $f^{-1}(V_i) = U_i \cap M$. Then,

$$K \subseteq f^{-1}\Big(\bigcup_{i \in I} V_i\Big) = \bigcup_{i \in I} f^{-1}(V_i) = \bigcup_{i \in I} (U_i \cap M) \subseteq \bigcup_{i \in I} U_i.$$

Since K is compact, there are $i_1, \ldots, i_n \in I$ so that $K \subseteq \bigcup_{k=1}^n U_{i_k}$. Then

$$f(K) = f(K \cap M) \subseteq f\Big(\bigcup_{k=1}^{n} (U_{i_k} \cap M)\Big) = \bigcup_{k=1}^{n} f(U_{i_k} \cap M) \subseteq \bigcup_{k=1}^{n} V_{i_k}.$$

So $f(K)$ is compact. $\square$

Proposition 4.9. *Let A be a topological space, $M \subseteq A$, and let $f : M \to \mathbb{R}$ be continuous. If $K \subseteq M$ is compact, then f has a maximum value and a minimum value in K.*

Proof. According to proposition 4.8, $f(K)$ is a compact subset of $\mathbb{R}$, and hence it is closed and bounded. Since $f(K)$ is bounded, $u = \sup(f(K))$ is a real number. Then for every $\epsilon > 0$ there is $a \in f(K)$ so that $a \in (u - \epsilon]$ and hence $u \in \mathrm{cl}(f(K))$. Since $f(K)$ is closed, we conclude that $u \in f(K)$ and so u is the maximum value of f in K. The case of the minimum value is similar. $\square$

Definition. *Let A be a non-empty set, and $\mathcal{C}$ be a non-empty collection of subsets of A. We say that $\mathcal{C}$ has the* **finite intersection property**, *if $\bigcap_{k=1}^n C_k \neq \emptyset$ for every $C_1, \ldots, C_n \in \mathcal{C}$.*

Proposition 4.10. *Let A be a topological space, and $K \subseteq A$. Then K is compact if and only if for every collection $\mathcal{F}$ of subsets of K with the finite intersection property we have that $K \cap \bigcap_{F \in \mathcal{F}} \mathrm{cl}(F) \neq \emptyset$.*

Proof. Assume that K is compact, and consider any collection $\mathcal{F}$ of subsets of K with the finite intersection property. Then $\mathcal{G} = \{A \setminus \mathrm{cl}(F) \,|\, F \in \mathcal{F}\}$ is a collection of open subsets of A. For every $F_1, \ldots, F_n \in \mathcal{F}$ we get $K \cap \bigcap_{k=1}^n F_k = \bigcap_{k=1}^n F_k \neq \emptyset$ and so $K \cap \bigcap_{k=1}^n \mathrm{cl}(F_k) \neq \emptyset$ which implies $\bigcup_{k=1}^n (A \setminus \mathrm{cl}(F_k)) \neq K$. So there is no finite subcollection of $\mathcal{G}$ which is a cover of K. Since K is compact, $\mathcal{G}$ is not a cover of K. Thus, $\bigcup_{F \in \mathcal{F}} (A \setminus \mathrm{cl}(F)) \neq K$ and this implies $K \cap \bigcap_{F \in \mathcal{F}} \mathrm{cl}(F) \neq \emptyset$.
Now, assume that for every collection $\mathcal{F}$ of subsets of K with the finite intersection property we

have that $K \cap \bigcap_{F \in \mathcal{F}} \operatorname{cl}(F) \neq \emptyset$. Take any open cover $\mathcal{G}$ of K. If $x \in K$, then $x \in G_0$ for some $G_0 \in \mathcal{G}$, and, since $(K \setminus G_0) \cap G_0 = \emptyset$, we get $x \notin \operatorname{cl}(K \setminus G_0)$. Therefore, $K \cap \bigcap_{G \in \mathcal{G}} \operatorname{cl}(K \setminus G) = \emptyset$. Now $\mathcal{F} = \{K \setminus G \,|\, G \in \mathcal{G}\}$ is a collection of subsets of K which, according to our assumption, does not have the finite intersection property. So there are $G_1, \ldots, G_n \in \mathcal{G}$ so that $\bigcap_{k=1}^{n} (K \setminus G_k) = \emptyset$, i.e. $K \subseteq \bigcup_{k=1}^{n} G_k$. Therefore, K is compact. □

4.1.4 Subspace topology.

Proposition 4.11. *Let A be a topological space with topology $\mathcal{T}$, and $B \subseteq A$. Then the collection $\mathcal{S} = \{U \cap B \,|\, U \in \mathcal{T}\}$ is a topology of B.*

Proof. $\emptyset = \emptyset \cap B$ and $B = A \cap B$, so $\emptyset$ and B belong to $\mathcal{S}$.
Let $V_i \in \mathcal{S}$ for every $i \in I$. Then there are $U_i \in \mathcal{T}$ so that $V_i = U_i \cap B$ for every $i \in I$. Since $\mathcal{T}$ is a topology, we have $\bigcup_{i \in I} U_i \in \mathcal{T}$. Hence, $\bigcup_{i \in I} V_i = \big(\bigcup_{i \in I} U_i\big) \cap B \in \mathcal{S}$.
Let $V_1, \ldots, V_n \in \mathcal{S}$. Then there are $U_i \in \mathcal{T}$ so that $V_i = U_i \cap B$ for every $i = 1, \ldots, n$. Since $\mathcal{T}$ is a topology, we have $\bigcap_{i=1}^{n} U_i \in \mathcal{T}$. Hence, $\bigcap_{i=1}^{n} V_i = \big(\bigcap_{i=1}^{n} U_i\big) \cap B \in \mathcal{S}$. □

Definition. *Let A be a topological space, and $B \subseteq A$. The topology of B, which is described in proposition 4.11, is called* **subspace topology** *or* **relative topology** *of B (with respect to A).*

In other words, if $B \subseteq A$ has its subspace topology, then the open subsets of B are the intersections with B of the open subsets of A.

Let $V \subseteq B \subseteq A$. Then we say that V is **open in** A, if V is open as a subset of A, i.e. it belongs to the topology of A, and we say that V is **open in** B, if V is open as a subset of B, i.e. it belongs to the subspace topology of B. In the second case, by definition, $V = U \cap B$ for some $U \subseteq A$ which is open in A.

Proposition 4.12. *Let A be a topological space, and let $B \subseteq A$ have its subspace topology. Then $G \subseteq B$ is closed in B if and only if there is $F \subseteq A$ closed in A so that $G = F \cap B$.*

Proof. Let $G \subseteq B$ be closed in B. Then $B \setminus G$ is open in B, and so $B \setminus G = U \cap B$ for some U open in A. We take $F = A \setminus U$, and then $G = F \cap B$ and F is closed in A.
The proof of the converse is similar. □

4.2 Weak topology.

Proposition 4.13. *Let A be a non-empty set, and let $\mathcal{F}$ be a non-empty collection of functions $f : A \to B_f$, where every B_f is a topological space with topology $\mathcal{S}_f$. For every $x \in A$ we consider the collection $\mathcal{N}_x$ of all sets*

$$N_x = \{y \in A \,|\, f_k(y) \in V_k \text{ for every } k = 1, \ldots, n\},$$

with arbitrary $n \in \mathbb{N}$, arbitrary $f_1, \ldots, f_n \in \mathcal{F}$, and arbitrary $V_1 \in \mathcal{S}_{f_1}, \ldots, V_n \in \mathcal{S}_{f_n}$ such that $f_k(x) \in V_k$ for every $k = 1, \ldots, n$.
Observe that $x \in N_x$.
Finally, we consider the collection

$$\sigma(A, \mathcal{F}) = \{U \subseteq A \,|\, \text{for every } x \in U \text{ there is } N_x \in \mathcal{N}_x \text{ so that } N_x \subseteq U\}.$$

Then $\sigma(A, \mathcal{F})$ is a topology of A.
Moreover, for every $x \in X$, every $N_x \in \mathcal{N}_x$ belongs to $\sigma(A, \mathcal{F})$.

Proof. It is easy to see that $\emptyset \in \sigma(A, \mathcal{F})$, and that $A \in \sigma(A, \mathcal{F})$.
Let $U_i \in \sigma(A, \mathcal{F})$ for every $i \in I$, and take any $x \in \bigcup_{i \in I} U_i$. Then $x \in U_{i_0}$ for some $i_0 \in I$, and so there is $N_x \in \mathcal{N}_x$ so that $N_x \subseteq U_{i_0} \subseteq \bigcup_{i \in I} U_i$. Therefore $\bigcup_{i \in I} U_i \in \sigma(A, \mathcal{F})$.

Let $U_1, \dots, U_n \in \sigma(A, \mathcal{F})$, and take any $x \in \bigcap_{k=1}^n U_k$. Then $x \in U_k$ for every $k = 1, \dots, n$, and so there are $N_{x1}, \dots, N_{xn} \in \mathcal{N}_x$ so that $N_{xk} \subseteq U_k$ for every $k = 1, \dots, n$. Now, it is easy to see that $\bigcap_{k=1}^n N_{xk} \in \mathcal{N}_x$, and $\bigcap_{k=1}^n N_{xk} \subseteq \bigcap_{k=1}^n U_k$. Therefore, $\bigcap_{k=1}^n U_k \in \sigma(A, \mathcal{F})$.
Finally, take any $N_x \in \mathcal{N}_x$, i.e.

$$N_x = \{y \in A \,|\, f_k(y) \in V_k \text{ for every } k = 1, \dots, n\}$$

for some $n \in \mathbb{N}$, some $f_1, \dots, f_n \in \mathcal{F}$, and some $V_1 \in \mathcal{S}_{f_1}, \dots, V_n \in \mathcal{S}_{f_n}$ such that $f_k(x) \in V_k$ for every $k = 1, \dots, n$.
We take any $z \in N_x$, and then $f_k(z) \in V_k$ for every $k = 1, \dots, n$. Now, if we define $N_z = N_x$, then clearly we have that $N_z \in \mathcal{N}_z$ and obviously $N_z \subseteq N_x$. This implies that $N_x \in \sigma(A, \mathcal{F})$. $\square$

Definition. *Let A be a non-empty set, and let $\mathcal{F}$ be a non-empty collection of functions $f : A \to B_f$, where every B_f is a topological space with topology $\mathcal{S}_f$. Then the topology $\sigma(A, \mathcal{F})$ of A, which is described in proposition 4.13, is called* **weak topology** *of A induced by the collection of functions $\mathcal{F}$. The elements of $\sigma(A, \mathcal{F})$ are called* **weakly open** *subsets of A with respect to the collection of functions $\mathcal{F}$.*

Definition. *Let A be non-empty set, and let $\mathcal{T}_1, \mathcal{T}_2$ be two topologies of A. We say that $\mathcal{T}_1$ is* **weaker** *than $\mathcal{T}_2$ and that $\mathcal{T}_2$ is* **stronger** *than $\mathcal{T}_1$, if $\mathcal{T}_1 \subseteq \mathcal{T}_2$.*

In other words, $\mathcal{T}_1$ is weaker than $\mathcal{T}_2$ if and only if every $U \subseteq A$ which is open with respect to $\mathcal{T}_1$ is also open with respect to $\mathcal{T}_2$. It is clear that $\mathcal{T}_1$ is weaker than $\mathcal{T}_2$ if and only if every $F \subseteq A$ which is closed with respect to $\mathcal{T}_1$ is also closed with respect to $\mathcal{T}_2$.

Proposition 4.14. *Let A be a non-empty set, and let $\mathcal{F}$ be a non-empty collection of functions $f : A \to B_f$, where every B_f is a topological space with topology $\mathcal{S}_f$. Then $\sigma(A, \mathcal{F})$ is the weakest topology of A with respect to which every $f \in \mathcal{F}$ is continuous.*

Proof. We take any $f \in \mathcal{F}$, and any $x \in A$. We consider any $V \in \mathcal{S}_f$ such that $f(x) \in V$. Then the set $N_x = \{y \in A \,|\, f(y) \in V\}$ clearly belongs to $\mathcal{N}_x$. Now $N_x \in \sigma(A, \mathcal{F})$, $x \in N_x$ and obviously $f(N_x) \subseteq V$. Hence f is continuous at every $x \in A$.
Let $\mathcal{T}$ be any topology of A such that every $f \in \mathcal{F}$ is continuous. We take any $x \in A$ and any $N_x \in \mathcal{N}_x$, i.e.

$$N_x = \{y \in A \,|\, f_k(y) \in V_k \text{ for every } k = 1, \dots, n\}$$

for some $n \in \mathbb{N}$, some $f_1, \dots, f_n \in \mathcal{F}$, and some $V_1 \in \mathcal{S}_{f_1}, \dots, V_n \in \mathcal{S}_{f_n}$ such that $f_k(x) \in V_k$ for every $k = 1, \dots, n$. We observe that

$$N_x = \bigcap_{k=1}^n f_k^{-1}(V_k).$$

Since each f_k is continuous, we have that $f_k^{-1}(V_k) \in \mathcal{T}$ for every $k = 1, \dots, n$, and hence $N_x \in \mathcal{T}$.
Now we consider any $U \in \sigma(A, \mathcal{F})$. Then for every $x \in U$ there is $N_x \in \mathcal{N}_x$ so that $x \in N_x \subseteq U$.
This implies that $U = \bigcup_{x \in U} N_x$, and since $N_x \in \mathcal{T}$ for every $x \in U$, we conclude that $U \in \mathcal{T}$.
Therefore, $\sigma(A, \mathcal{F}) \subseteq \mathcal{S}$. $\square$

Proposition 4.15. *Let A be a non-empty set, and let $\mathcal{F}$ be a non-empty collection of functions $f : A \to B_f$, where every B_f is a topological space with topology $\mathcal{S}_f$, and let A have the weak topology $\sigma(A, \mathcal{F})$. Consider also a topological space D and $g : D \to A$. Then g is continuous if and only if $f \circ g : D \to B_f$ is continuous for every $f \in \mathcal{F}$.*

Proof. If g is continuous, then, obviously, $f \circ g : D \to B_f$ is continuous for every $f \in \mathcal{F}$.
Conversely, let $f \circ g : D \to B_f$ be continuous for every $f \in \mathcal{F}$. We take any $p \in D$ and any

$U \in \sigma(A, \mathcal{F})$ such that $g(p) \in U$. Then there are $f_1, \ldots, f_n \in \mathcal{F}$ and $V_1 \in \mathcal{S}_{f_1}, \ldots, V_n \in \mathcal{S}_{f_n}$, so that $f_k(g(p)) \in V_k$ for every $k = 1, \ldots, n$ and so that

$$N_{g(p)} = \{y \in A \,|\, f_k(y) \in V_k \text{ for every } k = 1, \ldots, n\} \subseteq U.$$

Since each $f_k \circ g$ is continuous, there is $P_k \in \mathcal{R}$, where $\mathcal{R}$ is the topology of D, so that $p \in P_k$ and

$$f_k(g(q)) = (f_k \circ g)(q) \in V_k \quad \text{for every } q \in P_k.$$

Now, if $P = \bigcap_{k=1}^n P_k$, then $P \in \mathcal{R}$, $p \in P$, and

$$g(q) \in N_{g(p)} \subseteq U \quad \text{for every } q \in P.$$

Therefore g is continuous at p. □

Definition. *Let A be a non-empty set, and $\mathcal{F}$ be a non-empty collection of functions $f : A \to B_f$. We say that $\mathcal{F}$ is* **separating**, *if for every $x_1, x_2 \in A$, $x_1 \neq x_2$ there is $f \in \mathcal{F}$ so that $f(x_1) \neq f(x_2)$.*

Proposition 4.16. *Let A be a non-empty set, and let $\mathcal{F}$ be a non-empty collection of functions $f : A \to B_f$, where every B_f is a topological space with topology $\mathcal{S}_f$, and let A have the weak topology $\sigma(A, \mathcal{F})$. If $\mathcal{F}$ is separating, and if every topology $\mathcal{S}_f$ is Hausdorff, then $\sigma(A, \mathcal{F})$ is Hausdorff.*

Proof. Let $x_1, x_2 \in A$, $x_1 \neq x_2$. Since $\mathcal{F}$ is separating, there is $f \in \mathcal{F}$ so that $f(x_1) \neq f(x_2)$. Now, since $f(x_1), f(x_2) \in B_f$ and $\mathcal{S}_f$ is Hausdorff, there are $V_1, V_2 \in \mathcal{S}_f$ so that $f(x_1) \in V_1$, $f(x_2) \in V_2$ and $V_1 \cap V_2 = \emptyset$.
Now we consider $N_{x_1} = \{y \in A \,|\, f(y) \in V_1\}$ and $N_{x_2} = \{y \in A \,|\, f(y) \in V_2\}$. Then $x_1 \in N_{x_1}$, $x_2 \in N_{x_2}$, and $N_{x_1}, N_{x_2} \in \sigma(A, \mathcal{F})$, and $N_{x_1} \cap N_{x_2} = \emptyset$. □

4.3 Product topology.

Definition. *We consider a non-empty set I of indices, and a collection of sets $\{A_i \,|\, i \in I\}$. We define the set*

$$\prod_{i \in I} A_i = \Big\{x \,\Big|\, x : I \to \bigcup_{i \in I} A_i \text{ so that } x(i) \in A_i \text{ for every } i \in I\}.$$

This set is called **cartesian product** *of $\{A_i \,|\, i \in I\}$.*

Axiom of choice. *If I is non-empty and A_i is non-empty for every $i \in I$, then the cartesian product of $\{A_i \,|\, i \in I\}$ is non-empty.*

Proof. We consider the set $\mathcal{X}$ whose elements are all the functions $x : D(x) \to \bigcup_{i \in I} A_i$, where $D(x)$ is any non-empty subset of I and $x(i) \in A_i$ for every $i \in D(x)$.
If we choose any $i_0 \in I$ and any $a_0 \in A_{i_0}$ we may define the function $x_0 : \{i_0\} \to \bigcup_{i \in I} A_i$ by $x_0(i_0) = a_0$. Clearly, $x_0 \in \mathcal{X}$.
We define an order relation in $\mathcal{X}$ as follows. If $x_1, x_2 \in \mathcal{X}$, we write $x_1 \prec x_2$ if x_2 is an extension of x_1, i.e. if $D(x_1) \subseteq D(x_2)$ and $x_1(i) = x_2(i)$ for every $i \in D(x_1)$. It is clear that $\prec$ is an order relation in $\mathcal{X}$.
Let $\mathcal{X}_0$ be any totally ordered subset of $\mathcal{X}$. We consider the set

$$J_0 = \bigcup_{x \in \mathcal{X}_0} D(x) \subseteq I.$$

If $i \in J_0$, there is $x \in \mathcal{X}_0$ so that $i \in D(x)$. If $i \in D(x')$ for any other $x' \in \mathcal{X}_0$, then, since one of the x, x' is an extension of the other, we get that $x(i) = x'(i)$. So we can consider the function

$$x_0 : J_0 \to \bigcup_{i \in I} A_i$$

defined for every $i \in J_0$ by

$$x_0(i) = x(i) \quad \text{for any } x \in \mathcal{X}_0 \text{ with } i \in D(x).$$

Then $D(x_0) = J_0 \subseteq I$, and it is clear that x_0 is an element of $\mathcal{X}$, and that it is an extension of every $x \in \mathcal{X}_0$. Thus, x_0 is an upper bound of $\mathcal{X}_0$ in $\mathcal{X}$.
Therefore, Zorn's lemma implies that there is at least one maximal element x in $\mathcal{X}$. This means that $x : D(x) \to \bigcup_{i \in I} A_i$, where $D(x) \subseteq I$, and $x(i) \in A_i$ for every $i \in I$, and also that there is no $x' \in \mathcal{X}$ which is a proper extension of x.
If $D(x) = I$, then x is an element of the cartesian product $\prod_{i \in I} A_i$.
Assume that $D(x) \neq I$. Then we take any $i_0 \in I \setminus D(x)$ and any $a_0 \in A_{i_0}$, and we consider the function $x' : D(x) \cup \{i_0\} \to \bigcup_{i \in I} A_i$ defined so that: $x'(i) = x(i)$ for every $i \in D(x)$, and $x'(i_0) = a_0$. Obviously, $x' \in \mathcal{X}$ and x' is a proper extension of x. This is a contradiction. $\square$

We proved the axiom of choice using Zorn's lemma. It is possible to prove Zorn's lemma using the axiom of choice, and so the axiom of choice and Zorn's lemma are equivalent.

Exactly as in the case of sequences, a convenient way to denote elements x of the cartesian product $\prod_{i \in I} A_i$ is

$$x = (x_i)_{i \in I},$$

where we denote x_i the value $x(i) \in A_i$ and we call it i-**th coordinate** or i-**th term** of x. If the index set is $I = \{1, 2, \ldots, n\}$, then the cartesian product is denoted $\prod_{i=1}^{n} A_i$ or $A_1 \times \cdots \times A_n$, and its elements are denoted $x = (x_i)_{i=1}^{n}$ or $x = (x_1, \ldots, x_n)$. Similarly, if the index set is $\mathbb{N} = \{1, 2, \ldots\}$, then the cartesian product is denoted $\prod_{i=1}^{+\infty} A_i$ or $A_1 \times A_2 \times \cdots$, and its elements are denoted $x = (x_i)_{i=1}^{+\infty}$ or $x = (x_1, x_2, \ldots)$.

Definition. *For each $j \in I$ we consider the function $\pi_j : \prod_{i \in I} A_i \to A_j$ defined for every $x = (x_i)_{i \in I}$ by*

$$\pi_j(x) = x_j.$$

*This function is called j-**th projection**.*

Definition. *Let I be a non-empty set of indices, and for each $i \in I$ let A_i be a topological space with topology $\mathcal{S}_i$. We also consider the collection $\mathcal{P} = \{\pi_i \,|\, i \in I\}$ of projections $\pi_j : \prod_{i \in I} A_i \to A_j$. The weak topology $\sigma\big(\prod_{i \in I} A_i, \mathcal{P}\big)$ of $\prod_{i \in I} A_i$, is called **product topology** of $\prod_{i \in I} A_i$.*

We recall proposition 4.13 in order to describe the product topology of $\prod_{i \in I} A_i$, i.e. the weak topology $\sigma\big(\prod_{i \in I} A_i, \mathcal{P}\big)$. For every $x \in \prod_{i \in I} A_i$ we consider the collection $\mathcal{N}_x$ of all sets

$$N_x = \Big\{ y \in \prod_{i \in I} A_i \,\Big|\, y_{i_k} = \pi_{i_k}(y) \in V_{i_k} \text{for every } k = 1, \ldots, n \Big\}$$

with arbitrary $n \in \mathbb{N}$, arbitrary $i_1, \ldots, i_n \in I$, and arbitrary $V_{i_1} \in \mathcal{S}_{i_1}, \ldots, V_{i_n} \in \mathcal{S}_{i_n}$ such that $x_{i_k} = \pi_{i_k}(x) \in V_{i_k}$ for every $k = 1, \ldots, n$. Then

$$\sigma\Big(\prod_{i \in I} A_i, \mathcal{P}\Big) = \Big\{ U \subseteq \prod_{i \in I} A_i \,\Big|\, \text{for every } x \in U \text{ there is } N_x \in \mathcal{N}_x \text{ so that } N_x \subseteq U \Big\}.$$

We also have that every set N_x belongs to the product topology $\sigma\big(\prod_{i \in I} A_i, \mathcal{P}\big)$.
The next proposition collects all the basic results about the product topology.

Proposition 4.17. *Let I be a non-empty set of indices, and for each $i \in I$ let A_i be a topological space with topology $\mathcal{S}_i$.*
(i) The product topology is the weakest topology of $\prod_{i\in I} A_i$ with respect to which every projection $\pi_j : \prod_{i\in I} A_i \to A_j$ is continuous.
(ii) Let D be a topological space and $g : D \to \prod_{i\in I} A_i$, and let $\prod_{i\in I} A_i$ have its product topology. Then g is continuous if and only if $\pi_j \circ g : D \to A_j$ is continuous for every $j \in I$.
(iii) If every topology $\mathcal{S}_i$ is Hausdorff, then the product topology is Hausdorff.

Proof. (i) Direct implications of proposition 4.14.
(ii) Direct implication of proposition 4.15.
(iii) The collection $\mathcal{P} = \{\pi_i \,|\, i \in I\}$ is separating. Indeed, take any $x, y \in \prod_{i\in I} A_i$ so that $x \neq y$. Then there is $i \in I$ so that $x_i \neq y_i$, i.e. $\pi_i(x) \neq \pi_i(y)$. Now the result is a direct implication of proposition 4.16. □

The theorem of Tychonov. *Let I be a non-empty set of indices, and for each $i \in I$ let A_i be a topological space with topology $\mathcal{S}_i$. Let $\prod_{i\in I} A_i$ have its product topology. If every A_i is compact, then $\prod_{i\in I} A_i$ is compact.*

Proof. We shall use proposition 4.10.
We take any collection $\mathcal{F}$ of subsets of $\prod_{i\in I} A_i$ with the finite intersection property, and we shall prove that $\bigcap_{F\in\mathcal{F}} \mathrm{cl}(F) \neq \emptyset$.
We consider the collection

$$\mathbf{P} = \Big\{\mathcal{G} \Big| \mathcal{G} \supseteq \mathcal{F} \text{ is a collection of subsets of } \prod_{i\in I} A_i \text{ with the finite intersection property}\Big\}.$$

We also consider the order relation of set inclusion in $\mathbf{P}$.
Now we take any totally ordered $\mathbf{P_0} \subseteq \mathbf{P}$, and we define

$$\mathcal{F}_0 = \bigcup_{\mathcal{G}\in\mathbf{P_0}} \mathcal{G}.$$

This is a collection of subsets of $\prod_{i\in I} A_i$ with the finite intersection property. Indeed, if we take any $C_1, \ldots, C_n \in \mathcal{F}_0$, then $C_1 \in \mathcal{G}_1, \ldots, C_n \in \mathcal{G}_n$ for some $\mathcal{G}_1, \ldots, \mathcal{G}_n \in \mathbf{P_0}$. Since $\mathbf{P_0}$ is totally ordered, there is one of $\mathcal{G}_1, \ldots, \mathcal{G}_n$ which includes all the others. Thus, $C_1, \ldots, C_n$ belong to one $\mathcal{G} \in \mathbf{P_0}$, and so $\bigcap_{k=1}^n C_k \neq \emptyset$. It is also clear that $\mathcal{F} \subseteq \mathcal{F}_0$. Therefore, $\mathcal{F}_0 \in \mathbf{P}$. Since $\mathcal{G} \subseteq \mathcal{F}_0$ for every $\mathcal{G} \in \mathbf{P}_0$, we conclude that $\mathcal{F}_0$ is an upper bound of $\mathbf{P}_0$ in $\mathbf{P}$.
According to the lemma of Zorn, $\mathbf{P}$ has a maximal element, i.e. there is a collection $\mathcal{G} \supseteq \mathcal{F}$ of subsets of $\prod_{i\in I} A_i$ with the finite intersection property, and so that there is no strictly larger collection with the same properties.
This implies, in particular, that every intersection of finitely many elements of $\mathcal{G}$ belongs to $\mathcal{G}$. Indeed, if G is the intersection of finitely many elements of $\mathcal{G}$ so that $G \notin \mathcal{G}$, then $\mathcal{G}' = \mathcal{G} \cup \{G\} \supseteq \mathcal{F}$ is a collection of subsets of $\prod_{i\in I} A_i$ *with the finite intersection property*, and it is strictly larger than $\mathcal{G}$.
Now, since $\mathcal{F} \subseteq \mathcal{G}$, it is enough to prove that $\bigcap_{G\in\mathcal{G}} \mathrm{cl}(G) \neq \emptyset$.
For each $j \in I$ we consider the collection $\mathcal{G}_j = \{\pi_j(G) \,|\, G \in \mathcal{G}\}$ of subsets of A_j. It is easy to see that $\mathcal{G}_j$ has the finite intersection property. Indeed, if we take any $G_1, \ldots, G_n \in \mathcal{G}$, then there is $x \in \bigcap_{k=1}^n G_k$, and so $x_j = \pi_j(x) \in \bigcap_{k=1}^n \pi_k(G_k)$. Therefore, the compactness of A_j implies that $\bigcap_{G\in\mathcal{G}} \mathrm{cl}(\pi_j(G)) \neq \emptyset$. For every $j \in I$ we take any

$$x_j \in \bigcap_{G\in\mathcal{G}} \mathrm{cl}(\pi_j(G)),$$

and we consider the

$$x = (x_i)_{i\in I} \in \prod_{i\in I} A_i.$$

We shall prove that $x \in \bigcap_{G \in \mathcal{G}} \mathrm{cl}(G)$.
Now, let V_j be any open neighborhood of x_j in A_j. Since $x_j \in \bigcap_{G \in \mathcal{G}} \mathrm{cl}(\pi_j(G))$, we have that V_j has non-empty intersection with $\pi_j(G)$ for every $G \in \mathcal{G}$. So if we take any $G \in \mathcal{G}$, then there is $a_j \in V_j \cap \pi_j(G)$, and so there is $y \in G$ so that $\pi_j(y) = a_j \in V_j$, i.e. $y \in \pi_j^{-1}(V_j) \cap G$. Thus, $\pi_j^{-1}(V_j)$ has non-empty intersection with every $G \in \mathcal{G}$. This implies that $\mathcal{G} \cup \{\pi_j^{-1}(V_j)\} \supseteq \mathcal{F}$ is a collection of subsets of $\prod_{i \in I} A_i$ with the finite intersection property. Since $\mathcal{G}$ is a maximal collection with these properties, we get that $\pi_j^{-1}(V_j) \in \mathcal{G}$.
Now we take any $i_i, \ldots, i_n \in I$ and any open neighborhoods $V_{i_1}, \ldots, V_{i_n}$ of $x_{i_1}, \ldots, x_{i_n}$ in $A_{i_1}, \ldots, A_{i_n}$. Then $\bigcap_{k=1}^n \pi_{i_k}^{-1}(V_{i_k}) \in \mathcal{G}$ and hence $\bigcap_{k=1}^n \pi_{i_k}^{-1}(V_{i_k})$ has non-empty intersection with every $G \in \mathcal{G}$.
Now, let U be any open neighborhood of x in $\prod_{i \in I} A_i$. Then there are $i_i, \ldots, i_n \in I$ and open neighborhoods $V_{i_1}, \ldots, V_{i_n}$ of $x_{i_1}, \ldots, x_{i_n}$ in $A_{i_1}, \ldots, A_{i_n}$ so that

$$\bigcap_{k=1}^{n} \pi_{i_k}^{-1}(V_{i_k}) = C_x = \left\{ y \in \prod_{i \in I} A_i \,\middle|\, y_{i_k} = \pi_{i_k}(y) \in V_{i_k} \text{for every } k = 1, \ldots, n \right\} \subseteq U.$$

So every open neighborhood U of x has non-empty intersection with every $G \in \mathcal{G}$. Thus $x \in \mathrm{cl}(G)$ for every $G \in \mathcal{G}$, and we conclude that $x \in \bigcap_{G \in \mathcal{G}} \mathrm{cl}(G)$. □

4.4 Weak topologies of linear spaces.

Lemma 4.1. *Let X be a linear space, and let $\mathcal{L}$ be a non-empty collection of linear functionals $l : X \to F$. We consider F with its usual euclidean topology, and X with the weak topology $\sigma(X, \mathcal{L})$. Then $U \subseteq X$ is weakly open if and only if for every $x \in U$ there are $l_1, \ldots, l_n \in \mathcal{L}$ and $\epsilon_1, \ldots, \epsilon_n > 0$ so that*

$$C_x = \{y \in X \,|\, |l_k(y) - l_k(x)| < \epsilon_k \textit{ for every } k = 1, \ldots, n\} \subseteq U. \tag{4.1}$$

Moreover, every set of the form $C_x = \{y \in X \,|\, |l_k(y) - l_k(x)| < \epsilon_k$ for every $k = 1, \ldots, n\}$ is weakly open.

Proof. Let $U \in \sigma(X, \mathcal{L})$ and $x \in U$. According to the definition of $\sigma(X, \mathcal{L})$, as this appears in proposition 4.13, there are $l_1, \ldots, l_n \in \mathcal{L}$, and open sets $V_1, \ldots, V_n \subseteq F$ so that $l_k(x) \in V_k$ for every $k = 1, \ldots, n$ and so that

$$N_x = \{y \in X \,|\, l_k(y) \in V_k \text{ for every } k = 1, \ldots, n\} \subseteq U.$$

Then for every $k = 1, \ldots, n$ there is $\epsilon_k > 0$ such that $\{\lambda \in F \,|\, |\lambda - l_k(x)| < \epsilon_k\} \subseteq V_k$. Thus,

$$C_x = \{y \in X \,|\, |l_k(y) - l_k(x)| < \epsilon_k \text{ for every } k = 1, \ldots, n\} \subseteq N_x \subseteq U.$$

Conversely, assume that for every $x \in U$ there are $l_1, \ldots, l_n \in \mathcal{L}$ and $\epsilon_1, \ldots, \epsilon_n > 0$ so that (4.1) is true. Then each $V_k = \{\lambda \in F \,|\, |\lambda - l_k(x)| < \epsilon_k\}$ is an open subset of F containing $l_k(x)$, and we clearly have

$$N_x = \{y \in X \,|\, l_k(y) \in V_k \text{ for every } k = 1, \ldots, n\} = C_x \subseteq U.$$

So $U \in \sigma(X, \mathcal{L})$.
Finally, we already noticed that

$$\begin{aligned} C_x = \{y \in X \,|\, |l_k(y) - l_k(x)| < \epsilon_k \text{ for every } k = 1, \ldots, n\} \\ = \{y \in X \,|\, l_k(y) \in V_k \text{ for every } k = 1, \ldots, n\} = N_x \end{aligned}$$

where each $V_k = \{\lambda \in F \,|\, |\lambda - l_k(x)| < \epsilon_k\}$ is an open subset of F, and so, according to proposition 4.13, this set is weakly open. □

Proposition 4.18. *Let X be a linear space, and let $\mathcal{L}$ be a non-empty collection of linear functionals $l : X \to F$. We consider F with its usual euclidean topology.*
(i) $\sigma(X, \mathcal{L})$ is the weakest topology of X with respect to which every $l \in \mathcal{L}$ is continuous.
(ii) Let D be a topological space and $g : D \to X$, and let X have the weak topology $\sigma(X, \mathcal{L})$. Then g is continuous if and only if $l \circ g : D \to F$ is continuous for every $l \in \mathcal{L}$.
(iii) If $\mathcal{L}$ is separating, then $\sigma(X, \mathcal{L})$ is Hausdorff.

Proof. (i) Direct implication of proposition 4.14.
(ii) Direct implication of proposition 4.15.
(iii) Direct implication of proposition 4.16, since F with the euclidean topology is Hausdorff. □

Proposition 4.19. *Let X be a linear space, and let $\mathcal{L}$ be a non-empty collection of linear functionals $l : X \to F$. We consider F with its usual euclidean topology, and X with the weak topology $\sigma(X, \mathcal{L})$. Then the linear space operations of addition and multiplication are continuous.*

Proof. We consider addition: $+ : X \times X \to X$.
Let $x_1, x_2 \in X$ and let $x_1 + x_2 \in U$, where $U \in \sigma(X, \mathcal{L})$. Then there are $l_1, \dots, l_n \in \mathcal{L}$ and $\epsilon_1, \dots, \epsilon_n > 0$ so that

$$C_{x_1+x_2} = \{y \in X \,|\, |l_k(y) - l_k(x_1 + x_2)| < \epsilon_k \text{ for every } k = 1, \dots, n\} \subseteq U.$$

We consider the sets

$$C_{x_1} = \left\{y \in X \,\middle|\, |l_k(y) - l_k(x_1)| < \frac{\epsilon_k}{2} \text{ for every } k = 1, \dots, n\right\},$$

$$C_{x_2} = \left\{y \in X \,\middle|\, |l_k(y) - l_k(x_2)| < \frac{\epsilon_k}{2} \text{ for every } k = 1, \dots, n\right\}.$$

Then $C_{x_1}, C_{x_2} \in \sigma(X, \mathcal{L})$ and $x_1 \in C_{x_1}$, $x_2 \in C_{x_2}$. Now, if $y_1 \in C_{x_1}$, $y_2 \in C_{x_2}$, then for every $k = 1, \dots, n$ we get

$$|l_k(y_1 + y_2) - l_k(x_1 + x_2)| \le |l_k(y_1) - l_k(x_1)| + |l_k(y_2) - l_k(x_2)| < \frac{\epsilon_k}{2} + \frac{\epsilon_k}{2} = \epsilon_k,$$

and hence $y_1 + y_2 \in C_{x_1+x_2} \subseteq U$. Therefore, addition is continuous.
The proof that multiplication $\cdot : F \times X \to X$ is continuous is similar and we leave it as an exercise. □

Definition. *Let X be a linear space equipped with a topology $\mathcal{T}$. If the linear space operations of addition and multiplication on X are continuous with respect to $\mathcal{T}$, then we say that X is a* **topological linear space**.

Example 4.4.1. If X is a linear space equipped with the weak topology which is induced by a non-empty collection of linear functionals in X, then X is a topological linear space.

Example 4.4.2. Every normed space X is a topological linear space.

Lemma 4.2. *Let X be a linear space, and $l, l_1, \dots, l_n : X \to F$ be linear functionals in X. If $l(x) = 0$ for every $x \in X$ such that $l_1(x) = \dots = l_n(x) = 0$, then there are $\kappa_1, \dots, \kappa_n \in F$ so that $l = \kappa_1 l_1 + \dots + \kappa_n l_n$.*

Proof. We consider the linear function $L : X \to F^n$ defined for every $x \in X$ by

$$L(x) = (l_1(x), \dots, l_n(x)).$$

Then we consider the function $M : \mathrm{R}(L) \to F$ defined for every $y \in \mathrm{R}(L)$ by

$$M(y) = l(x) \quad \text{where } y = L(x).$$

This function is well defined, since, if $y = L(x_1)$ and $y = L(x_2)$, then $l(x_1) = l(x_2)$. It is also easy to see that M is linear on the linear subspace $\mathrm{R}(L)$ of F^n.
Now, we extend M to F^n, i.e. we consider any linear functional $\overline{M} : F^n \to F$ so that $\overline{M}(y) = M(y)$ for every $y \in \mathrm{R}(L)$. Then there are $\kappa_1, \ldots, \kappa_n \in F$ so that for every $y = (\lambda_1, \ldots, \lambda_n) \in F^n$ we have

$$\overline{M}(y) = \kappa_1\lambda_1 + \cdots + \kappa_n\lambda_n.$$

This implies

$$l(x) = M(L(x)) = \overline{M}(L(x)) = \overline{M}(l_1(x), \ldots, l_n(x)) = \kappa_1 l_1(x) + \cdots + \kappa_n l_n(x)$$

for every $x \in X$. □

Proposition 4.20. *Let X be a linear space, let $\mathcal{L}$ be a non-empty collection of linear functionals $l : X \to F$, and let X have the weak topology $\sigma(X, \mathcal{L})$. Then a linear functional $l : X \to F$ is continuous in X if and only if $l \in \mathrm{span}(\mathcal{L})$.*

Proof. If $l \in \mathrm{span}(\mathcal{L})$, i.e. if $l = \kappa_1 l_1 + \cdots + \kappa_n l_n$ for some $\kappa_1, \ldots, \kappa_n \in F$ and some $l_1, \ldots, l_n \in \mathcal{L}$, then it is obvious that l is continuous in X.
Conversely, let l be continuous in X. Then l is continuous at $0 \in X$ and so there are $l_1, \ldots, l_n \in \mathcal{L}$ and $\epsilon_1, \ldots, \epsilon_n > 0$ so that $|l(x)| < 1$ for every $x \in C_0$, where

$$C_0 = \{x \in X \mid |l_k(x)| < \epsilon_k \quad \text{for every } k = 1, \ldots, n\}.$$

Now, take any $x \in X$ such that $l_1(x) = \ldots = l_n(x) = 0$. Then for every $t > 0$ we have $l_1(tx) = \ldots = l_n(tx) = 0$ and hence $tx \in C_0$. Thus,

$$t|l(x)| = |l(tx)| < 1,$$

and letting $t \to +\infty$ we get $l(x) = 0$. Now, lemma 4.2 finishes the proof. □

Proposition 4.21. *Let X be a linear space, and let $\mathcal{L}$ be a separating collection of linear functionals $l : X \to F$. We consider the function $\phi : X \to \prod_{l \in \mathcal{L}} F$ defined for every $x \in X$ by $\phi(x) = (l(x))_{l \in \mathcal{L}}$. Then ϕ is one-to-one in X.*
If X has the weak topology $\sigma(X, \mathcal{L})$, and $\prod_{l \in \mathcal{L}} F$ has the product topology (where each F has the euclidean topology), and $\phi(X)$ has the subspace topology, then $\phi : X \to \phi(X)$ is a homeomorphism of X onto $\phi(X)$.

Proof. Take any $x_1, x_2 \in X$ with $\phi(x_1) = \phi(x_2)$. Then $\phi(x_1)_l = \phi(x_2)_l$ and hence $l(x_1) = l(x_2)$ for every $l \in \mathcal{L}$. Since $\mathcal{L}$ is separating, we get $x_1 = x_2$. Thus, ϕ is one-to-one.
It remains to prove that $\phi : X \to \phi(X)$ and $\phi^{-1} : \phi(X) \to X$ are continuous.
If $x \in X$, then the parameters $n \in \mathbb{N}$, $l_1, \ldots, l_n \in \mathcal{L}$ and $\epsilon_1, \ldots, \epsilon_n > 0$ define the open neighborhood

$$C_x = \{y \in X \mid |l_k(y) - l_k(x)| < \epsilon_k \text{ for every } k = 1, \ldots, n\}$$

of x with respect to $\sigma(X, \mathcal{L})$.
If $z \in \prod_{l \in \mathcal{L}} F$, then the same parameters define the open neighborhood

$$N_z = \Big\{ w \in \prod_{l \in \mathcal{L}} F \,\Big|\, |w_{l_k} - z_{l_k}| < \epsilon_k \text{ for every } k = 1, \ldots, n \Big\}$$

of z with respect to the product topology. Now, if we restrict z, w in $\phi(X)$ and set $z = \phi(x), w = \phi(y)$ with $x, y \in X$, we get an open neighborhood $\widetilde{N}_z$ of $z \in \phi(X)$ with respect to the subspace topology of $\phi(X)$. Writing

$$z_{l_k} = \phi(x)_{l_k} = l_k(x), \quad w_{l_k} = \phi(y)_{l_k} = l_k(y)$$

for every $k = 1, \dots, n$, we get

$$\widetilde{N}_z = \{\phi(y) \in \phi(X) \,|\, |l_k(y) - l_k(x)| < \epsilon_k \text{ for every } k = 1, \dots, n\}.$$

So $\widetilde{N}_z = \phi(C_x)$.
Take $x \in X$ and V open in $\phi(X)$ with respect to its subspace topology so that $z = \phi(x) \in V$. Then there is some $\widetilde{N}_z \subseteq V$. We consider the corresponding C_x, which is an open neighborhood of x in X with respect to $\sigma(X, \mathcal{L})$, and then $\phi(C_x) = \widetilde{N}_z \subseteq V$. Therefore, ϕ is continuous at x.
Take $z = \phi(x) \in \phi(X)$ and U open in X with respect to $\sigma(X, \mathcal{L})$ so that $x = \phi^{-1}(z) \in U$. Then there is some $C_x \subseteq U$. We consider the corresponding $\widetilde{N}_z$, which is an open neighborhood of z in $\phi(X)$ with respect to its subspace topology, and then $\phi^{-1}(\widetilde{N}_z) = C_x \subseteq U$. Therefore, ϕ^{-1} is continuous at z. $\square$

4.5 Weak topologies of normed spaces.

If X is a normed space, then theorem 3.10 implies that the collection $\mathcal{L} = X'$ of bounded linear functionals in X is separating. Indeed, let $x_1, x_2 \in X$, $x_1 \neq x_2$. Then

$$0 < \|x_1 - x_2\| = \max_{x' \in X', \|x'\| \leq 1} |x'(x_1 - x_2)|,$$

and so there is $x' \in X'$ such that $x'(x_1) - x'(x_2) = x'(x_1 - x_2) \neq 0$.

Definition. *Let X be a normed space. The topology $\sigma(X, X')$ is called* **weak topology** *of X. A subset of X which is open or closed or compact with respect to $\sigma(X, X')$ is called* **weakly open** *or* **weakly closed** *or* **weakly compact**, *respectively.*

According to lemma 4.1, a basic open neighborhood of $x \in X$ with respect to $\sigma(X, X')$ is

$$C_x = \{y \in X \,|\, |x'_k(y) - x'_k(x)| < \epsilon_k \text{ for every } k = 1, \dots, n\},$$

where $n \in \mathbb{N}$, $x'_1, \dots, x'_n \in X'$ and $\epsilon_1, \dots, \epsilon_n > 0$ are arbitrary.

We know the following about $\sigma(X, X')$. All are consequences of propositions 4.18, 4.19 and 4.20.
(i) If X has the weak topology $\sigma(X, X')$, then the linear space operations of addition and multiplication on X are continuous.
(ii) $\sigma(X, X')$ is the weakest topology of X with respect to which every $x' \in X'$ is continuous.
(iii) Let D be a topological space and $g : D \to X$, and let X have its weak topology $\sigma(X, X')$. Then g is continuous if and only if $x' \circ g : D \to F$ is continuous for every $x' \in X'$.
(iv) $\sigma(X, X')$ is Hausdorff.
(v) If X has its weak topology $\sigma(X, X')$, then a linear functional $l : X \to F$ is continuous in X if and only if $l \in X'$.

We have exactly the same situation for X' and its dual X''. The weak topology on X' is $\sigma(X', X'')$. On the other hand, there is another interesting topology on X'.

Definition. *Let X be a normed space, and consider the natural embedding $J : X \to X''$. Then $J(X) \subseteq X''$ is a collection of linear functionals in X'. The topology $\sigma(X', J(X))$ is called* **weak$*$ topology** *on X'. Because of the identification of X with $J(X)$, the topology $\sigma(X', J(X))$ is traditionally denoted $\sigma(X', X)$. A subset of X' which is open or closed or compact with respect to $\sigma(X', X)$ is called* **weakly$*$ open** *or* **weakly$*$ closed** *or* **weakly$*$ compact**, *respectively.*

A basic open neighborhood of $x' \in X'$ with respect to $\sigma(X', X'')$ is

$$C_{x'} = \{y' \in X' \,|\, |x''_k(y') - x''_k(x')| < \epsilon_k \text{ for every } k = 1, \dots, n\},$$

for arbitrary $n \in \mathbb{N}$, $x''_1, \dots, x''_n \in X''$ and $\epsilon_1, \dots, \epsilon_n > 0$.

Also, a basic open neighborhood of $x' \in X'$ with respect to $\sigma(X', X) = \sigma(X', J(X))$ is

$$C_{x'} = \{y' \in X' \,|\, |J(x_k)(y') - J(x_k)(x')| < \epsilon_k \text{ for every } k = 1, \dots, n\}$$
$$= \{y' \in X' \,|\, |y'(x_k) - x'(x_k)| < \epsilon_k \text{ for every } k = 1, \dots, n\},$$

for arbitrary $n \in \mathbb{N}$, $x_1, \dots, x_k \in X$ and $\epsilon_1, \dots, \epsilon_n > 0$.

We have the following for $\sigma(X', X)$.

(i) If X' has the weak$*$ topology $\sigma(X', X)$, then the linear space operations of addition and multiplication on X' are continuous.

(ii) $\sigma(X', X)$ is the weakest topology of X' with respect to which every $x'' \in J(X)$ is continuous.

(iii) Let D be a topological space and $g : D \to X'$, and let X' have its weak$*$ topology $\sigma(X', X)$. Then g is continuous if and only if $J(x) \circ g : D \to F$ is continuous for every $x \in X$.

(iv) $\sigma(X', X)$ is Hausdorff.

(v) If X' has its weak$*$ topology $\sigma(X', X)$, then a linear functional $l : X' \to F$ is continuous in X' if and only if $l \in J(X)$.

The fact that $\sigma(X', X) = \sigma(X', J(X))$ is Hausdorff follows from proposition 4.18, since $J(X)$ is separating. Indeed, if $x'_1, x'_2 \in X'$, $x'_1 \neq x'_2$, then there is $x \in X$ so that $x'_1(x) \neq x'_2(x)$ and hence $J(x)(x'_1) \neq J(x)(x'_2)$.

In a normed space X we have two topologies: the weak topology and the topology which is induced by the norm of X, which is also called **strong topology** on X.

In X' we have three topologies: the strong topology, the weak topology, and the weak$*$ topology. Clearly, *if X is reflexive, then the weak topology and the weak$*$ topology of X' are the same.*

In X'' we have two topologies: the strong topology and the weak$*$ topology.

The next proposition expresses the relation between the weak topology on X and the weak$*$ topology on X'' through the natural embedding of X into X''. What happens is that, after the identification of X with $J(X)$, the weak topology of X and the weak$*$ topology of $J(X)$ (more precisely, the restriction of the weak$*$ topology of X'' on $J(X)$) are the same.

Proposition 4.22. *Let X be a normed space, and let $J : X \to X''$ be the natural embedding. If X has the weak topology, X'' has the weak$*$ topology, and $J(X) \subseteq X''$ has the subspace topology, then $J : X \to J(X)$ is a homeomorphism.*

Proof. Take any $x \in X$. The parameters $n \in \mathbb{N}$, $x'_1, \dots, x'_n \in X'$ and $\epsilon_1, \dots, \epsilon_n > 0$ give us the open neighborhood

$$C_x = \{y \in X \,|\, |x'_k(y) - x'_k(x)| < \epsilon_k \text{ for every } k = 1, \dots, n\}$$

of x with respect to the weak topology of X. The same parameters give us the open neighborhood

$$C_{J(x)} = \{y'' \in X'' \,|\, |y''(x'_k) - J(x)(x'_k)| < \epsilon_k \text{ for every } k = 1, \dots, n\}$$

of $J(x)$ with respect to the weak$*$ topology of X''. Restricting y'' in $J(X)$, i.e. taking $y'' = J(y)$, and writing $J(x)(x'_k) = x'_k(x)$ and $J(y)(x'_k) = x'_k(y)$, we get the open neighborhood

$$\widetilde{C}_{J(x)} = \{J(y) \in J(X) \,|\, |x'_k(y) - x'_k(x)| < \epsilon_k \text{ for every } k = 1, \dots, n\}$$

of $J(x)$ with respect to the subspace topology of $J(X)$.

It is clear that $\widetilde{C}_{J(x)} = J(C_x)$.

Take any $x \in X$, and any V open with respect to the subspace topology of $J(X)$ so that $J(x) \in V$. Then there is $\widetilde{C}_{J(x)} \subseteq V$, and for the corresponding C^x we get $J(C_x) = \widetilde{C}_{J(x)} \subseteq V$. Therefore, J is continuous at x.

Now take any $z = J(x) \in J(X)$, and any U weakly open in X so that $x \in U$. Then there is $C_x \subseteq U$, and for the corresponding $\widetilde{C}_{J(x)}$ we get $J^{-1}(\widetilde{C}_{J(x)}) = C_x \subseteq U$. Therefore, J^{-1} is continuous at $z = J(x)$. $\square$

Proposition 4.23. *Let X be a normed space.*
(i) $x_n \overset{w}{\to} x$ in X if and only if (x_n) converges to x with respect to the weak topology of X.
(ii) $x'_n \overset{w*}{\to} x'$ in X' if and only if (x'_n) converges to x' with respect to the weak$*$ topology of X'.

Proof. (i) Let $x_n \overset{w}{\to} x$ in X. We take any $U \in \sigma(X, X')$ such that $x \in U$. Then there are $x'_1, \dots, x'_m \in X'$ and $\epsilon_1, \dots, \epsilon_m > 0$ so that

$$C_x = \{y \in X \,|\, |x'_k(y) - x'_k(x)| < \epsilon_k \text{ for every } k = 1, \dots, m\} \subseteq U.$$

Since $x'(x_n) \to x'(x)$ for every $x' \in X'$, there is n_0 so that $|x'_k(x_n) - x'_k(x)| < \epsilon_k$ for every $n \geq n_0$ and every $k = 1, \dots, m$. This means that $x_n \in C_x \subseteq U$ for every $n \geq n_0$. So (x_n) converges to x with respect to the weak topology of X.
Conversely, let (x_n) converge to x with respect to the weak topology of X. We take any $x' \in X'$ and the weakly open neighborhood $C_x = \{y \in X \,|\, |x'(y) - x'(x)| < \epsilon\}$ of x. Then there is n_0 so that $x_n \in C_x$ for every $n \geq n_0$, i.e. $|x'(x_n) - x'(x)| < \epsilon$ for every $n \geq n_0$. Thus, $x'(x_n) \to x'(x)$ for every $x' \in X'$, and so $x_n \overset{w}{\to} x$.
(ii) Similarly. □

Proposition 4.24. *Let X be a normed space.*
(i) The weak topology of X is weaker than the strong topology of X.
(ii) The weak$*$ topology of X' is weaker than the weak topology of X' and this is weaker than the strong topology of X'.

Proof. (i) If X has its strong topology, then every $x' \in X'$ is continuous. Since $\sigma(X, X')$ is the weakest topology of X with respect to which every $x' \in X'$ is continuous, we get that $\sigma(X, X')$ is weaker than the strong topology of X.
(ii) That $\sigma(X', X'')$ is weaker than the strong topology of X' is an immediate consequence of (i). Now, every $x'' \in X''$ is continuous with respect to $\sigma(X', X'')$. In particular, every $x'' \in J(X) \subseteq X''$ is continuous with respect to $\sigma(X', X'')$. Since $\sigma(X', X) = \sigma(X', J(X))$ is the weakest topology of X' with respect to which every $x'' \in J(X)$ is continuous, we get that $\sigma(X', X)$ is weaker than $\sigma(X', X'')$. □

Proposition 4.25. *Let X be a normed space.*
(i) If $K \subseteq X$ is weakly compact, then it is weakly closed and bounded.
(ii) If $K \subseteq X'$ is weakly$*$ compact, then it is weakly$*$ closed and, if X is a Banach space, bounded.

Proof. (i) Since X with the topology $\sigma(X, X')$ is Hausdorff, the weakly compact $K \subseteq X$ is weakly closed.
Every $x' \in X'$ is continuous in X, and hence in K, with respect to $\sigma(X, X')$. Since K is weakly compact, we get $\sup_{x \in K} |x'(x)| < +\infty$ for every $x' \in X'$. According to theorem 3.16, we have that $\sup_{x \in K} \|x\| < +\infty$, and so K is bounded.
(ii) Similarly. □

The theorem of Alaoglou. *Let X be a normed space.*
(i) The closed unit ball of X' with center 0 is weakly$*$ compact.
(ii) If $K \subseteq X'$ is weakly$*$ closed and bounded, then it is weakly$*$ compact.

Proof. (i) We apply proposition 4.21 to the space X' with $\mathcal{L} = J(X) \subseteq X''$ and the induced weak$*$ topology of X'. To do this we consider the homeomorphism

$$\phi : X' \to \phi(X') \subseteq \prod_{J(x) \in J(X)} F = \prod_{x \in X} F$$

defined for every $x' \in X'$ by

$$\phi(x') = (J(x)(x'))_{J(x) \in J(X)} = (x'(x))_{x \in X}.$$

Let $\overline{B}' = \{x' \in X' \mid \|x'\| \leq 1\}$ be the closed unit ball of X' with center 0. Then for every $x \in X$ and every $x' \in \overline{B}'$ we have

$$|\phi(x')_x| = |x'(x)| \leq \|x\|.$$

Therefore, if $x' \in \overline{B}'$, then $\phi(x')_x \in \{\lambda \mid |\lambda| \leq \|x\|\}$ for every $x \in X$. Hence,

$$\phi(\overline{B}') \subseteq \prod_{x \in X} \{\lambda \mid |\lambda| \leq \|x\|\} \subseteq \prod_{x \in X} F.$$

Now, it is enough to prove that $\phi(\overline{B}')$ is a closed subset of $\prod_{x \in X} F$ with respect to the product topology. Indeed, this will imply that $\phi(\overline{B}')$ is a closed subset of $\prod_{x \in X}\{\lambda \mid |\lambda| \leq \|x\|\}$, which is compact by the theorem of Tychonov, and this will imply that $\phi(\overline{B}')$ is compact, and hence that $\overline{B}'$, a continuous image of $\phi(\overline{B}')$, is weakly$*$ compact.
Let $\kappa = (\kappa_x)_{x \in X} \in \prod_{x \in X} F$ belong to $\mathrm{cl}(\phi(\overline{B}'))$. We take any $x_1, x_2 \in X$, any $\lambda \in F$ and any $\epsilon > 0$, and the open neighborhood of κ:

$$\Big\{\mu = (\mu_x)_{x \in X} \in \prod_{x \in X} F \,\Big|\, |\mu_{x_1} - \kappa_{x_1}| < \epsilon, |\mu_{x_2} - \kappa_{x_2}| < \epsilon, \\ |\mu_{x_1 + x_2} - \kappa_{x_1 + x_2}| < \epsilon, |\mu_{\lambda x_1} - \kappa_{\lambda x_1}| < \epsilon\Big\}.$$

Then there is $x' \in \overline{B}'$ so that $\phi(x')$ belongs to this neighborhood. This means that

$$|x'(x_1) - \kappa_{x_1}| < \epsilon, \quad |x'(x_2) - \kappa_{x_2}| < \epsilon, \quad |x'(x_1 + x_2) - \kappa_{x_1 + x_2}| < \epsilon, \quad |x'(\lambda x_1) - \kappa_{\lambda x_1}| < \epsilon.$$

Since x' is linear and $\|x'\| \leq 1$, we easily prove that

$$|\kappa_{x_1 + x_2} - \kappa_{x_1} - \kappa_{x_2}| < 3\epsilon, \quad |\kappa_{\lambda x_1} - \lambda \kappa_{x_1}| < (1 + |\lambda|)\epsilon, \quad |\kappa_{x_1}| \leq \|x_1\| + \epsilon.$$

Finally, since ϵ is arbitrary, we get

$$\kappa_{x_1 + x_2} = \kappa_{x_1} + \kappa_{x_2}, \quad \kappa_{\lambda x_1} = \lambda \kappa_{x_1}, \quad |\kappa_{x_1}| \leq \|x_1\|$$

for every $x_1, x_2 \in X$ and every $\lambda \in F$.
Now we consider $x' : X \to F$ defined for every $x \in X$ by $x'(x) = \kappa_x$. Then

$$x'(x_1 + x_2) = x'(x_1) + x'(x_2), \quad x'(\lambda x_1) = \lambda x'(x_1), \quad |x'(x_1)| \leq \|x_1\|$$

for every $x_1, x_2 \in X$ and every $\lambda \in F$. This means that $x' \in \overline{B}'$ and $\phi(x') = \kappa$, and so $\kappa \in \phi(\overline{B}')$.
We just proved that $x \in \phi(\overline{B}')$ for every $x \in \mathrm{cl}(\phi(\overline{B}'))$, and hence that $\phi(\overline{B}')$ is closed.
(ii) Let $K \subseteq X'$ be weakly$*$ closed and bounded. Then there is $M > 0$ so that $K \subseteq \overline{B}'(0; M)$, where $\overline{B}'(0; M)$ is the closed ball of X' with center 0 and radius M.
Now, $\overline{B}'(0; M)$ is the image of $\overline{B}' = \overline{B}'(0; 1)$ under multiplication by M. Since multiplication is a continuous function with respect to the weak$*$ topology of X' and since $\overline{B}'$ is weakly$*$ compact, we get that $\overline{B}'(0; M)$ is also weakly$*$ compact. Then K is a weakly$*$ closed subset of $\overline{B}'(0; M)$ and so it is weakly$*$ compact. □

The theorem of Mazur. *Let X be a normed space, let $A, B \subseteq X$ be convex and disjoint, and let 0 be an interior point of A, i.e. $B(0; R) \subseteq A$ for some $R > 0$. Then there is $x' \in X'$ so that $\|x'\| \leq \frac{1}{R}$,* $\mathrm{Re}(x'(a)) \leq 1$ *for every $a \in A$, and* $\mathrm{Re}(x'(b)) \geq 1$ *for every $b \in B$. If, moreover, A is open, then we may also have that* $\mathrm{Re}(x'(a)) < 1$ *for every $a \in A$.*

Proof. It is obvious that the ball $B(0;R)$ absorbs X, and so A absorbs X. If, moreover, A is open, then A absorbs X with every $a \in A$ as center.
If $F = \mathbb{R}$, then theorem 3.7 implies that there is a linear functional $l : X \to \mathbb{R}$, $l \neq 0$, and $\lambda \in \mathbb{R}$ so that $l(a) \leq \lambda$ for every $a \in A$, and $l(b) \geq \lambda$ for every $b \in B$. If, moreover, A is open, then we may also have that $l(a) < \lambda$ for every $a \in A$.
Since $l \neq 0$, there is $x_0 \in X$ so that $l(x_0) \neq 0$. Then both points $\pm\frac{R}{2\|x_0\|}\, x_0$ belong to $B(0;R) \subseteq A$, and l has opposite values at these points. This implies $\lambda > 0$.
Now, take any $x \neq 0$ and any $t > 1$. Then $\pm\frac{R}{t\|x\|}\, x \in B(0;R) \subseteq A$, and hence

$$\pm\frac{R}{t\|x\|}\, l(x) = l\Big(\pm\frac{R}{t\|x\|}\, x\Big) \leq \lambda$$

and hence $|l(x)| \leq \frac{t\lambda}{R}\, \|x\|$. Since $t > 1$ is arbitrary, we get

$$|l(x)| \leq \frac{\lambda}{R}\, \|x\|.$$

This is also true for $x = 0$, and so $l \in X'$ with $\|l\| \leq \frac{\lambda}{R}$. Now we consider $x' = \frac{1}{\lambda}\, l \in X'$ and we have that $\|x'\| \leq \frac{1}{R}$, $x'(a) \leq 1$ for every $a \in A$, and $x'(b) \geq 1$ for every $b \in B$. If, moreover, A is open, then we may also have $x'(a) < 1$ for every $a \in A$.
If $F = \mathbb{C}$, we consider at first X as a $\mathbb{R}$-linear space. From the first part we know that there is a $\mathbb{R}$-linear functional $x_0' : X \to \mathbb{R}$ such that $\|x_0'\| \leq \frac{1}{R}$, $x_0'(a) \leq 1$ for every $a \in A$ και $x_0'(b) \geq 1$ for every $b \in B$. Moreover, if A is open, we may also have $x_0'(a) < 1$ for every $a \in A$.
Now, lemma 3.5 implies that there is a linear functional $x' : X \to \mathbb{C}$ so that $\mathrm{Re}(x') = x_0'$. Then, obviously, $\mathrm{Re}(x'(a)) \leq 1$ for every $a \in A$, and $\mathrm{Re}(x'(b)) \geq 1$ for every $b \in B$. If, moreover, A is open, then $\mathrm{Re}(x'(a)) < 1$ for every $a \in A$.
Also, for every $x \in X$ there is $\lambda \in \mathbb{C}$ so that $|\lambda| = 1$ and $|x'(x)| = \lambda x'(x)$, and hence

$$|x'(x)| = \lambda x'(x) = x'(\lambda x) = \mathrm{Re}(x')(\lambda x) = x_0'(\lambda x) \leq \|x_0'\|\|\lambda x\| \leq \frac{1}{R}\, \|x\|.$$

Therefore, $x' \in X'$ with $\|x'\| \leq \frac{1}{R}$. □

Since the weak topology of a normed space X is weaker than its strong topology, every $K \subseteq X$ which is weakly closed is also closed. The next theorem says that the converse is true for *convex* sets K.

Theorem 4.1. *Let X be a normed space, and $K \subseteq X$ be convex. If K closed, then it is weakly closed.*

Proof. Let K be convex and closed. We consider any $x \notin K$, and we shall prove that there is a weakly open neighborhood of x which is disjoint from K.
Since translations are continuous with respect to both the strong and the weak topology of X, we may assume that $x = 0$. Then there is $R > 0$ so that $B(0;R) \cap K = \emptyset$. The theorem of Mazur implies that there is $x' \in X$, so that $\|x'\| \leq \frac{1}{R}$, and $\mathrm{Re}(x'(x)) < 1$ for every $x \in B(0;R)$, and $\mathrm{Re}(x'(x)) \geq 1$ for every $x \in K$. Then $\{x \in X \mid |x'(x)| < 1\}$ is a weakly open neighborhood of 0 which is disjoint from K. □

Chapter 5

Weak topologies 2

5.1 Generalities about topological spaces.

5.1.1 Open sets and closed sets.

Definition. *Let A be a non-empty set, and $\mathcal{T}$ be a collection of subsets of A, with the properties:*
(i) $\emptyset \in \mathcal{T}$, $A \in \mathcal{T}$.
(ii) The union of any elements of $\mathcal{T}$ is an element of $\mathcal{T}$. In other words, if $U_i \in \mathcal{T}$ for every $i \in I$, then $\bigcup_{i\in I} U_i \in \mathcal{T}$.
(iii) The intersection of any finitely many elements of $\mathcal{T}$ is an element of $\mathcal{T}$. In other words, if $U_1, \dots, U_n \in \mathcal{T}$, then $\bigcap_{i=1}^n U_i \in \mathcal{T}$.
Then $\mathcal{T}$ is called **topology** *of A, and the elements of $\mathcal{T}$ are called* **open** *(with respect to $\mathcal{T}$) subsets of A. Finally, A equipped with a topology is called* **topological space**.

If A a topological space, then (ii) says that the union of any open subsets of A is an open subset of A, and (iii) says that the intersection of any finitely many open subsets of A is an open subset of A.

Example 5.1.1. Let A be a non-empty set. Then $\{\emptyset, A\}$ is a topology of A.

Example 5.1.2. Let A be a non-empty set. Then $\mathcal{P}(A)$, the collection of *all* subsets of A, is a topology of A.

Example 5.1.3. Let A be a metric space with metric d. Then

$$\mathcal{T} = \{U \subseteq A \,|\, U \text{ is open with respect to } d\}$$

is a topology in A. In this case we say that the topology $\mathcal{T}$ is **induced** by d.
To be more precise, $U \subseteq A$ is open with respect to d if for every $x \in U$ there is a radius $r > 0$ so that the ball $B(x, r) = \{y \in A \,|\, d(y, x) < r\}$ is included in U.
It is easy to see that $\emptyset$ and A are open with respect to d.
Now assume that U_i is open with respect to d for every $i \in I$, and take any $x \in \bigcup_{i\in I} U_i$. Then $x \in U_{i_0}$ for some $i_0 \in I$, and then there is $r > 0$ so that $B(x; r) \subseteq U_{i_0}$. Hence $B(x; r) \subseteq \bigcup_{i\in I} U_i$ and we have that $\bigcup_{i\in I} U_i$ is open with respect to d.
Finally, assume that $U_1, \dots, U_n$ are open with respect to d, and take any $x \in \bigcap_{i=1}^n U_i$. Then for every $i = 1, \dots, n$ there is $r_i > 0$ so that $B(x; r_i) \subseteq U_i$. If we take $r = \min\{r_1, \dots, r_n\} > 0$, then $B(x; r) \subseteq B(x; r_i) \subseteq U_i$ for every $i = 1, \dots, n$, and hence $B(x; r) \subseteq \bigcap_{i=1}^n U_i$. So $\bigcap_{i=1}^n U_i$ is open with respect to d.
Every ball $B(x; r)$ is open with respect to d. Indeed, take any $y \in B(x; r)$. Then $d(y, x) < r$, and we consider $s = r - d(y, x) > 0$. Now, if $z \in B(y; s)$, then

$$d(z, x) \le d(z, y) + d(y, x) < s + d(y, x) = r$$

and hence $z \in B(x; r)$. Thus $B(y; s) \subseteq B(x; r)$.

Definition. *Let A be a topological space, and $F \subseteq A$. We say that F is* **closed***, if $A \setminus F$ is open.*

Proposition 5.1. *Let A be a topological space. Then*
(i) $\emptyset$ and A are closed.
(ii) The intersection of any closed subsets of A is a closed subset of A.
(iii) The union of any finitely many closed subsets of A is a closed subset of A.

Proof. The proof is a trivial corollary of the definition of closed set, of the properties of open sets, and of the laws of de Morgan for the complements of unions and intersections. □

Definition. *Let A be a topological space, and $x \in A$. Every open set containing x is called* **open neighborhood** *of x.*

Definition. *Let A be a topological space, and $M \subseteq A$. Then the set $\bigcap\{F \mid F \supseteq M \text{ is closed}\}$ is called* **closure** *of M and it is denoted* $\mathrm{cl}(M)$.

Proposition 5.2. *Let A be a topological space, and $M \subseteq A$.*
(i) $\mathrm{cl}(M)$ is the smallest closed subset of A which includes M.
(ii) $x \in \mathrm{cl}(M)$ if and only if $U \cap M \neq \emptyset$ for every open neighborhood U of x.

Proof. (i) $\mathrm{cl}(M)$ is the intersection of closed sets which include M, and so it closed and includes M. Also, if F is closed and includes M, then $\mathrm{cl}(M) \subseteq F$. So $\mathrm{cl}(M)$ is the smallest closed subset of A which includes M.
(ii) Let $x \in \mathrm{cl}(M)$, and take any open neighborhod U of x. Then $A \setminus U$ is closed and, since $x \notin A \setminus U$, we have that $\mathrm{cl}(M)$ is not included in $A \setminus U$. According to (i), M is not included in $A \setminus U$, and hence $U \cap M \neq \emptyset$.
Conversely, assume that $U \cap M \neq \emptyset$ for every open neighborhood U of x. We take any closed $F \supseteq M$, and then $A \setminus F$ is open and $(A \setminus F) \cap M = \emptyset$. Therefore, $x \notin A \setminus F$ and so $x \in F$. We conclude that $x \in \mathrm{cl}(M)$. □

Definition. *Let A be a topological space with topology $\mathcal{T}$, and let (x_n) be a sequence in A. We say that (x_n)* **converges (with respect to $\mathcal{T}$)** *to $x \in A$, if for every open neighborhood U of x there is n_0 so that $x_n \in U$ for every $n \geq n_0$.*
Then we say that x is a **limit** *of (x_n), and we write $x_n \to x$.*

5.1.2 Continuous functions.

Definition. *Let A, B be two topological spaces, $M \subseteq A$, and $f : M \to B$.*
(i) We say that f is **continuous** *at $x \in M$ if for every open $V \subseteq B$ such that $f(x) \in V$ there is an open $U \subseteq A$ so that $x \in U$ and $f(U \cap M) \subseteq V$, i.e. so that $x \in U$, and $f(y) \in V$ for every $y \in U \cap M$.*
(ii) We say that f is **continuous** *in M if it is continuous at every $x \in M$.*

Proposition 5.3. *Let A, B be two topological spaces, $M \subseteq A$, and $f : M \to B$. Then f is continuous in M if and only if for every open $V \subseteq B$ there is an open $U \subseteq A$ so that $f^{-1}(V) = U \cap M$.*

Proof. Let f be continuous in M, and let $V \subseteq B$ be open. Then for every $x \in f^{-1}(V)$ we have $f(x) \in V$, and so there is an open $U_x \subseteq A$ such that $x \in U_x$ and $f(U_x \cap M) \subseteq V$. Then $U = \bigcup_{x \in f^{-1}(V)} U_x \subseteq A$ is open, and it is easy to see that $f^{-1}(V) = U \cap M$. Indeed, if $y \in f^{-1}(V)$, then $y \in U_y \cap M$ and hence $y \in U \cap M$. Also, if $y \in U \cap M$, then $y \in U_x \cap M$ for some $x \in f^{-1}(V)$. Then $f(y) \in V$ and hence $y \in f^{-1}(V)$.
Conversely, take any $x \in M$ and any open $V \subseteq B$ so that $f(x) \in V$. Then there is an open $U \subseteq A$ so that $f^{-1}(V) = U \cap M$. Then $x \in U$ and $f(U \cap M) \subseteq V$, and so f is continuous at x. □

Proposition 5.4. *(i) Let A, B, C be topological spaces, $M \subseteq A$, $N \subseteq B$, $f : M \to N$ and $g : N \to C$. If f is continuous at $x \in M$ and g is continuous at $f(x) \in N$, then $g \circ f$ is continuous at x.*
(ii) Let A be a topological space, $M \subseteq A$, $f, g : M \to \mathbb{R}^n$ and $\lambda \in \mathbb{R}$. If $\mathbb{R}^n$ has the topology induced by the euclidean metric and f, g are continuous at $x \in M$, then $f + g, \lambda f$ are continuous at x.

Proof. (i) Let $W \subseteq C$ be open and $(g \circ f)(x) = g(f(x)) \in W$. Then there is an open $V \subseteq B$ so that $f(x) \in V$ and $g(V \cap N) \subseteq W$. Then there is an open $U \subseteq A$ so that $x \in U$ and $f(U \cap M) \subseteq V$. Then $f(U \cap M) \subseteq V \cap N$ and hence

$$(g \circ f)(U \cap M) = g(f(U \cap M)) \subseteq g(V \cap N) \subseteq W.$$

Thus $g \circ f$ is continuous at x.
(ii) Let $V \subseteq \mathbb{R}^n$ be open and $f(x) + g(x) \in V$. Then there is $r > 0$ so that $z \in V$ for every $z \in \mathbb{R}^n$ with $\|z - (f(x) + g(x))\| < r$. Now, there is an open $U_1 \subseteq A$ so that $x \in U_1$ and $\|f(y) - f(x)\| < \frac{r}{2}$ for every $y \in U_1 \cap M$. Also, there is an open $U_2 \subseteq A$ so that $x \in U_2$ and $\|g(y) - g(x)\| < \frac{r}{2}$ for every $y \in U_2 \cap M$. Then $U = U_1 \cap U_2 \subseteq A$ is open, and $x \in U$, and for every $y \in U \cap M$ we have

$$\|(f(y) + g(y)) - (f(x) + g(x))\| \leq \|f(y) - f(x)\| + \|g(y) - g(x)\| < \frac{r}{2} + \frac{r}{2} = r,$$

and hence $f(y) + g(y) \in V$. So $f + g$ is continuous at x.
The proof for λf is similar. $\square$

Definition. *Let A, B be topological spaces, and $f : A \to B$. We say that f is a* **homeomorphism** *of A onto B, if f is one-to-one in A and onto B, and f is continuous in A and f^{-1} is continuous in B. In this case we say that A, B are* **homeomorphic**.

If A, B are homeomorphic topological spaces, and $f : A \to B$ is a homeomorphism of A onto B, then we may *identify* the two spaces: we identify every $a \in A$ with the corresponding $b = f(a) \in B$ and, conversely, we identify every $b \in B$ with the corresponding $a = f^{-1}(b) \in A$. Then every open $U \subseteq A$ is identified with the open $V = f(U) \subseteq B$ and, conversely, every open $V \subseteq B$ is identified with the open $U = f^{-1}(V) \subseteq A$.

5.1.3 Compact sets.

Definition. *Let A be a topological space. We say that A is a* **Hausdorff** *topological space, if for every $x_1, x_2 \in A$, $x_1 \neq x_2$, there are disjoint open $U_1, U_2 \subseteq A$ so that $x_1 \in U_1$ and $x_2 \in U_2$.*

Proposition 5.5. *Every metric space is Hausdorff.*

Proof. If d is the metric of A and $x_1, x_2 \in A$, $x_1 \neq x_2$, we take $r = \frac{1}{2}\, d(x_1, x_2) > 0$, and then $B(x_1; r) \cap B(x_2; r) = \emptyset$. The balls $B(x_1; r), B(x_2; r)$ are open with respect to d. $\square$

Proposition 5.6. *Let A be a Hausdorff topological space. If a sequence in A has a limit, then this limit is unique.*

Proof. Let $x_n \to y$ and $x_n \to z$. If $y \neq z$, then there are disjoint open $U, V \subseteq A$ so that $y \in U$ and $z \in V$. But then there is n_0 so that $x_n \in U$ and $x_n \in V$ for every $n \geq n_0$, and obviously this is impossible. $\square$

Definition. *Let A be a topological space, and $K \subseteq A$. We say that a collection $\{U_i \mid i \in I\}$ of open subsets of A is an* **open cover** *of K, if $K \subseteq \bigcup_{i \in I} U_i$.*

Definition. *Let A be a topological space, and $K \subseteq A$. We say that K is* **compact**, *if for every open cover of K there is a finite subcover of K. More precisely, K is compact, if for every open cover $\{U_i \mid i \in I\}$ of K there are $i_1, \ldots, i_n \in I$ so that $\{U_{i_k} \mid 1 \leq k \leq n\}$ is also an open cover of K.*

Proposition 5.7. *Let A be a topological space.*
(i) If $K \subseteq A$ is compact and A is Hausdorff, then K is closed.
(ii) If $K \subseteq A$ is compact and $K' \subseteq K$ is closed, then K' is compact.

Proof. (i) Take any $x \in A \setminus K$. For every $z \in K$ there are disjoint open $U_z, V_z \subseteq A$ so that $z \in U_z$ and $x \in V_z$. Then $\{U_z \mid z \in K\}$ is an open cover of K. Since K is compact, there are $z_1, \ldots, z_n \in K$ so that $K \subseteq U_{z_1} \cup \cdots \cup U_{z_n}$. Then $V_{z_1} \cap \cdots \cap V_{z_n}$ is open, it is included in $A \setminus K$, and contains x. Therefore, $A \setminus K$ is open, and so K is closed.
(ii) Let $\{U_i \mid i \in I\}$ be any open cover of K'. Then $\{U_i \mid i \in I\} \cup \{A \setminus K'\}$ is an open cover of K. Since K is compact, there are $i_1, \ldots, i_n \in I$ so that $K \subseteq \left(\bigcup_{k=1}^n U_{i_k}\right) \cup (A \setminus K')$. Then $K' \subseteq \bigcup_{k=1}^n U_{i_k}$, and so K' is compact. □

Proposition 5.8. *Let A, B be topological spaces, $M \subseteq A$, and let $f : M \to B$ be continuous. If $K \subseteq M$ is compact, then $f(K)$ is compact.*

Proof. Let $\{V_i \mid i \in I\}$ be any open cover of $f(K)$, i.e. $f(K) \subseteq \bigcup_{i \in I} V_i$. Since each $V_i \subseteq B$ is open and f is continuous, proposition 5.3 implies that there is a corresponding open $U_i \subseteq A$ so that $f^{-1}(V_i) = U_i \cap M$. Then,

$$K \subseteq f^{-1}\Big(\bigcup_{i \in I} V_i\Big) = \bigcup_{i \in I} f^{-1}(V_i) = \bigcup_{i \in I} (U_i \cap M) \subseteq \bigcup_{i \in I} U_i.$$

Since K is compact, there are $i_1, \ldots, i_n \in I$ so that $K \subseteq \bigcup_{k=1}^n U_{i_k}$. Then

$$f(K) = f(K \cap M) \subseteq f\Big(\bigcup_{k=1}^{n} (U_{i_k} \cap M)\Big) = \bigcup_{k=1}^{n} f(U_{i_k} \cap M) \subseteq \bigcup_{k=1}^{n} V_{i_k}.$$

So $f(K)$ is compact. □

Proposition 5.9. *Let A be a topological space, $M \subseteq A$, and let $f : M \to \mathbb{R}$ be continuous. If $K \subseteq M$ is compact, then f has a maximum value and a minimum value in K.*

Proof. According to proposition 5.8, $f(K)$ is a compact subset of $\mathbb{R}$, and hence it is closed and bounded. Since $f(K)$ is bounded, $u = \sup(f(K))$ is a real number. Then for every $\epsilon > 0$ there is $a \in f(K)$ so that $a \in (u - \epsilon]$ and hence $u \in \mathrm{cl}(f(K))$. Since $f(K)$ is closed, we conclude that $u \in f(K)$ and so u is the maximum value of f in K. The case of the minimum value is similar. □

Definition. *Let A be a non-empty set, and $\mathcal{C}$ be a non-empty collection of subsets of A. We say that $\mathcal{C}$ has the* **finite intersection property**, *if $\bigcap_{k=1}^n C_k \neq \emptyset$ for every $C_1, \ldots, C_n \in \mathcal{C}$.*

Proposition 5.10. *Let A be a topological space, and $K \subseteq A$. Then K is compact if and only if for every collection $\mathcal{F}$ of subsets of K with the finite intersection property we have that $K \cap \bigcap_{F \in \mathcal{F}} \mathrm{cl}(F) \neq \emptyset$.*

Proof. Assume that K is compact, and consider any collection $\mathcal{F}$ of subsets of K with the finite intersection property. Then $\mathcal{G} = \{A \setminus \mathrm{cl}(F) \mid F \in \mathcal{F}\}$ is a collection of open subsets of A. For every $F_1, \ldots, F_n \in \mathcal{F}$ we get $K \cap \bigcap_{k=1}^n F_k = \bigcap_{k=1}^n F_k \neq \emptyset$ and so $K \cap \bigcap_{k=1}^n \mathrm{cl}(F_k) \neq \emptyset$ which implies $\bigcup_{k=1}^n (A \setminus \mathrm{cl}(F_k)) \neq K$. So there is no finite subcollection of $\mathcal{G}$ which is a cover of K. Since K is compact, $\mathcal{G}$ is not a cover of K. Thus, $\bigcup_{F \in \mathcal{F}} (A \setminus \mathrm{cl}(F)) \neq K$ and this implies $K \cap \bigcap_{F \in \mathcal{F}} \mathrm{cl}(F) \neq \emptyset$.
Now, assume that for every collection $\mathcal{F}$ of subsets of K with the finite intersection property we

have that $K \cap \bigcap_{F \in \mathcal{F}} \text{cl}(F) \neq \emptyset$. Take any open cover $\mathcal{G}$ of K. If $x \in K$, then $x \in G_0$ for some $G_0 \in \mathcal{G}$, and, since $(K \setminus G_0) \cap G_0 = \emptyset$, we get $x \notin \text{cl}(K \setminus G_0)$. Therefore, $K \cap \bigcap_{G \in \mathcal{G}} \text{cl}(K \setminus G) = \emptyset$. Now $\mathcal{F} = \{K \setminus G \mid G \in \mathcal{G}\}$ is a collection of subsets of K which, according to our assumption, does not have the finite intersection property. So there are $G_1, \ldots, G_n \in \mathcal{G}$ so that $\bigcap_{k=1}^{n} (K \setminus G_k) = \emptyset$, i.e. $K \subseteq \bigcup_{k=1}^{n} G_k$. Therefore, K is compact. □

5.2 Weak topologies of linear spaces.

Proposition 5.11. *Let X be a linear space, and let $\mathcal{L}$ be a non-empty collection of linear functionals $l : X \to F$. For every $x \in X$ we consider the collection $\mathcal{C}_x$ of all sets*

$$C_x = \{y \in X \mid |l_k(y) - l_k(x)| < \epsilon_k \text{ for every } k = 1, \ldots, n\},$$

with arbitrary $n \in \mathbb{N}$, arbitrary $l_1, \ldots, l_n \in \mathcal{L}$, and arbitrary $\epsilon_1, \ldots, \epsilon_n > 0$.
Observe that $x \in C_x$.
Finally, we consider the collection

$$\sigma(X, \mathcal{L}) = \{U \subseteq X \mid \text{for every } x \in U \text{ there is } C_x \in \mathcal{C}_x \text{ so that } C_x \subseteq U\}.$$

Then $\sigma(X, \mathcal{L})$ is a topology of X.
Moreover, for every $x \in X$, every $C_x \in \mathcal{C}_x$ belongs to $\sigma(X, \mathcal{L})$.

Proof. It is easy to see that $\emptyset \in \sigma(X, \mathcal{L})$, and that $X \in \sigma(X, \mathcal{L})$.
Let $U_i \in \sigma(X, \mathcal{L})$ for every $i \in I$, and take any $x \in \bigcup_{i \in I} U_i$. Then $x \in U_{i_0}$ for some $i_0 \in I$, and so there is $C_x \in \mathcal{C}_x$ so that $C_x \subseteq U_{i_0} \subseteq \bigcup_{i \in I} U_i$. Therefore $\bigcup_{i \in I} U_i \in \sigma(X, \mathcal{L})$.
Let $U_1, \ldots, U_n \in \sigma(X, \mathcal{L})$, and take any $x \in \bigcap_{k=1}^{n} U_k$. Then $x \in U_k$ for every $k = 1, \ldots, n$, and so there are $C_{x1}, \ldots, C_{xn} \in \mathcal{C}_x$ so that $C_{xk} \subseteq U_k$ for every $k = 1, \ldots, n$. Now, it is easy to see that $\bigcap_{k=1}^{n} C_{xk} \in \mathcal{C}_x$, and $\bigcap_{k=1}^{n} C_{xk} \subseteq \bigcap_{k=1}^{n} U_k$. Therefore, $\bigcap_{k=1}^{n} U_k \in \sigma(X, \mathcal{L})$.
Finally, take any $C_x \in \mathcal{C}_x$, i.e.

$$C_x = \{y \in X \mid |l_k(y) - l_k(x)| < \epsilon_k \text{ for every } k = 1, \ldots, n\}$$

for some $n \in \mathbb{N}$, $l_1, \ldots, l_n \in \mathcal{L}$, and $\epsilon_1, \ldots, \epsilon_n > 0$.
We take any $z \in C_x$, and then

$$|l_k(z) - l_k(x)| < \epsilon_k \text{ for every } k = 1, \ldots, n.$$

For each $k = 1, \ldots, n$ we consider $\delta_k = \epsilon_k - |l_k(z) - l_k(x)| > 0$. We also consider

$$C_z = \{y \in X \mid |l_k(y) - l_k(z)| < \delta_k \text{ for every } k = 1, \ldots, n\}.$$

Then for every $y \in C_z$ we get

$$|l_k(y) - l_k(x)| \leq |l_k(y) - l_k(z)| + |l_k(z) - l_k(x)| < \delta_k + |l_k(z) - l_k(x)| = \epsilon_k \text{ for every } k = 1, \ldots, n$$

and hence $y \in C_x$. Therefore, $C_z \subseteq C_x$, and so $C_x \in \sigma(X, \mathcal{L})$. □

Definition. *Let X be a linear space, and let $\mathcal{L}$ be a non-empty collection of linear functionals $l : X \to F$. Then the topology $\sigma(X, \mathcal{L})$ of X, which is described in proposition 5.11, is called* **weak topology** *of X induced by the collection of linear functionals $\mathcal{L}$. The elements of $\sigma(X, \mathcal{L})$ are called* **weakly open** *subsets of X with respect to the collection of linear functionals $\mathcal{L}$.*

Proposition 5.12. *Let X be a linear space, and let $\mathcal{L}$ be a non-empty collection of linear functionals $l : X \to F$. We consider X with the weak topology $\sigma(X, \mathcal{L})$. Then the linear space operations of addition and multiplication are continuous.*

Proof. We consider addition: $+ : X \times X \to X$.
Let $x_1, x_2 \in X$ and let $x_1 + x_2 \in U$, where $U \in \sigma(X, \mathcal{L})$. Then there are $l_1, \ldots, l_n \in \mathcal{L}$ and $\epsilon_1, \ldots, \epsilon_n > 0$ so that

$$C_{x_1+x_2} = \{y \in X \mid |l_k(y) - l_k(x_1 + x_2)| < \epsilon_k \text{ for every } k = 1, \ldots, n\} \subseteq U.$$

We consider the sets

$$C_{x_1} = \left\{y \in X \,\middle|\, |l_k(y) - l_k(x_1)| < \frac{\epsilon_k}{2} \text{ for every } k = 1, \ldots, n\right\},$$

$$C_{x_2} = \left\{y \in X \,\middle|\, |l_k(y) - l_k(x_2)| < \frac{\epsilon_k}{2} \text{ for every } k = 1, \ldots, n\right\}.$$

Then $C_{x_1}, C_{x_2} \in \sigma(X, \mathcal{L})$ and $x_1 \in C_{x_1}$, $x_2 \in C_{x_2}$. Now, if $y_1 \in C_{x_1}$, $y_2 \in C_{x_2}$, then for every $k = 1, \ldots, n$ we get

$$|l_k(y_1 + y_2) - l_k(x_1 + x_2)| \leq |l_k(y_1) - l_k(x_1)| + |l_k(y_2) - l_k(x_2)| < \frac{\epsilon_k}{2} + \frac{\epsilon_k}{2} = \epsilon_k,$$

and hence $y_1 + y_2 \in C_{x_1+x_2} \subseteq U$. Therefore, addition is continuous.
The proof that multiplication $\cdot : F \times X \to X$ is continuous is similar and we leave it as an exercise. □

Definition. *Let X be a linear space equipped with a topology $\mathcal{T}$. If the linear space operations of addition and multiplication on X are continuous with respect to $\mathcal{T}$, then we say that X is a* **topological linear space**.

Example 5.2.1. If X is a linear space equipped with the weak topology which is induced by a non-empty collection of linear functionals in X, then X is a topological linear space.

Example 5.2.2. Every normed space X is a topological linear space.

Definition. *Let A be non-empty set, and let $\mathcal{T}_1, \mathcal{T}_2$ be two topologies of A. We say that $\mathcal{T}_1$ is* **weaker** *than $\mathcal{T}_2$ and that $\mathcal{T}_2$ is* **stronger** *than $\mathcal{T}_1$, if $\mathcal{T}_1 \subseteq \mathcal{T}_2$.*

In other words, $\mathcal{T}_1$ is weaker than $\mathcal{T}_2$ if and only if every $U \subseteq A$ which is open with respect to $\mathcal{T}_1$ is also open with respect to $\mathcal{T}_2$. It is clear that $\mathcal{T}_1$ is weaker than $\mathcal{T}_2$ if and only if every $F \subseteq A$ which is closed with respect to $\mathcal{T}_1$ is also closed with respect to $\mathcal{T}_2$.

Proposition 5.13. *Let X be a linear space, and let $\mathcal{L}$ be a non-empty collection of linear functionals $l : X \to F$. Then $\sigma(X, \mathcal{L})$ is the weakest topology of X with respect to which every $l \in \mathcal{L}$ is continuous.*

Proof. We take any $l \in \mathcal{L}$, $x \in X$, and $\epsilon > 0$. Then the set

$$C_x = \{y \in X \mid |l(y) - l(x)| < \epsilon\}$$

belongs to $\sigma(X, \mathcal{L})$, $x \in C_x$, and we obviously have $|l(y) - l(x)| < \epsilon$ for every $y \in C_x$. Therefore, l is continuous at x.
Now, let $\mathcal{T}$ be any topology of X such that every $l \in \mathcal{L}$ is continuous. We take any $x \in X$ and any $C_x \in \mathcal{C}_x$, i.e.

$$C_x = \{y \in X \mid |l_k(y) - l_k(x)| < \epsilon_k \text{ for every } k = 1, \ldots, n\}$$

for some $l_1, \ldots, l_n \in \mathcal{L}$ and $\epsilon_1, \ldots, \epsilon_n > 0$. We observe that

$$C_x = \bigcap_{k=1}^{n} l_k^{-1}(V_k),$$

where each $V_k = \{\lambda \,|\, |\lambda - l_k(x)| < \epsilon_k\}$ is open in F. Since each l_k is continuous, we have that $l_k^{-1}(V_k) \in \mathcal{T}$ for every $k = 1, \ldots, n$, and hence $C_x \in \mathcal{T}$.
Now we consider any $U \in \sigma(X, \mathcal{L})$. Then for every $x \in X$ there is $C_x \in \mathcal{C}_x$ so that $x \in C_x \subseteq U$. This implies that $U = \bigcup_{x \in U} C_x$, and, since $C_x \in \mathcal{T}$ for every $x \in U$, we conclude that $U \in \mathcal{T}$. In other words $\sigma(X, \mathcal{L}) \subseteq \mathcal{T}$. $\square$

Proposition 5.14. *Let X be a linear space, let $\mathcal{L}$ be a non-empty collection of linear functionals $l : X \to F$, and let X have the weak topology $\sigma(X, \mathcal{L})$. Consider also a topological space D and $g : D \to X$. Then g is continuous if and only if $l \circ g : D \to F$ is continuous for every $l \in \mathcal{L}$.*

Proof. If g is continuous, then, obviously, $l \circ g : D \to F$ is continuous for every $l \in \mathcal{L}$.
Conversely, let $l \circ g : D \to F$ be continuous for every $l \in \mathcal{L}$. We take any $p \in D$ and any $U \in \sigma(X, \mathcal{L})$ such that $g(p) \in U$. Then there are $l_1, \ldots, l_n \in \mathcal{L}$ and $\epsilon_1, \ldots, \epsilon_n > 0$, so that

$$C_{g(p)} = \{y \in X \,|\, |l_k(y) - l_k(g(p))| < \epsilon_k \text{ for every } k = 1, \ldots, n\} \subseteq U.$$

Since each $l_k \circ g$ is continuous, there is $P_k \in \mathcal{R}$, where $\mathcal{R}$ is the topology of D, so that $p \in P_k$ and

$$|l_k(g(q)) - l_k(g(p))| = |(l_k \circ g)(q) - (l_k \circ g)(p)| < \epsilon_k \quad \text{for every } q \in P_k.$$

Now, if $P = \bigcap_{k=1}^n P_k$, then $P \in \mathcal{R}$, $p \in P$, and

$$g(q) \in C_{g(p)} \subseteq U \quad \text{for every } q \in P.$$

Therefore g is continuous at p. $\square$

Definition. *Let X be a linear space, and let $\mathcal{L}$ be a non-empty collection of linear functionals $l : X \to F$. We say that $\mathcal{L}$ is* **separating**, *if for every $x_1, x_2 \in X$, $x_1 \neq x_2$ there is $l \in \mathcal{L}$ so that $l(x_1) \neq l(x_2)$.*

Proposition 5.15. *Let X be a linear space, let $\mathcal{L}$ be a non-empty collection of linear functionals $l : X \to F$, and let X have the weak topology $\sigma(X, \mathcal{L})$. If $\mathcal{L}$ is separating, then $\sigma(X, \mathcal{L})$ is Hausdorff.*

Proof. Let $x_1, x_2 \in X$, $x_1 \neq x_2$. Since $\mathcal{L}$ is separating, there is $l \in \mathcal{L}$ so that $l(x_1) \neq l(x_2)$. Now, we take $\epsilon = \frac{|l(x_1) - l(x_2)|}{2} > 0$ and we consider the sets

$$C_{x_1} = \{y \in X \,|\, |l(y) - l(x_1)| < \epsilon\}, \quad C_{x_2} = \{y \in X \,|\, |l(y) - l(x_2)| < \epsilon\}.$$

Then $C_{x_1}, C_{x_2} \in \sigma(X, \mathcal{L})$ and $x_1 \in C_{x_1}, x_2 \in C_{x_2}$ and it is easy to see that $C_{x_1} \cap C_{x_2} = \emptyset$. $\square$

Lemma 5.1. *Let X be a linear space, and $l, l_1, \ldots, l_n : X \to F$ be linear functionals in X. If $l(x) = 0$ for every $x \in X$ such that $l_1(x) = \ldots = l_n(x) = 0$, then there are $\kappa_1, \ldots, \kappa_n \in F$ so that $l = \kappa_1 l_1 + \cdots + \kappa_n l_n$.*

Proof. We consider the linear function $L : X \to F^n$ defined for every $x \in X$ by

$$L(x) = (l_1(x), \ldots, l_n(x)).$$

Then we consider the function $M : \mathrm{R}(L) \to F$ defined for every $y \in \mathrm{R}(L)$ by

$$M(y) = l(x) \quad \text{where } y = L(x).$$

This function is well defined, since, if $y = L(x_1)$ and $y = L(x_2)$, then $l(x_1) = l(x_2)$. It is also easy to see that M is linear on the linear subspace $\mathrm{R}(L)$ of F^n.
Now, we extend M to F^n, i.e. we consider any linear functional $\overline{M} : F^n \to F$ so that $\overline{M}(y) =$

$M(y)$ for every $y \in \mathrm{R}(L)$. Then there are $\kappa_1, \ldots, \kappa_n \in F$ so that for every $y = (\lambda_1, \ldots, \lambda_n) \in F^n$ we have

$$\overline{M}(y) = \kappa_1 \lambda_1 + \cdots + \kappa_n \lambda_n.$$

This implies

$$l(x) = M(L(x)) = \overline{M}(L(x)) = \overline{M}(l_1(x), \ldots, l_n(x)) = \kappa_1 l_1(x) + \cdots + \kappa_n l_n(x)$$

for every $x \in X$. $\square$

Proposition 5.16. *Let X be a linear space, let $\mathcal{L}$ be a non-empty collection of linear functionals $l : X \to F$, and let X have the weak topology $\sigma(X, \mathcal{L})$. Then a linear functional $l : X \to F$ is continuous in X if and only if $l \in$ span$(\mathcal{L})$.*

Proof. If $l \in \mathrm{span}(\mathcal{L})$, i.e. if $l = \kappa_1 l_1 + \cdots + \kappa_n l_n$ for some $\kappa_1, \ldots, \kappa_n \in F$ and some $l_1, \ldots, l_n \in \mathcal{L}$, then it is obvious that l is continuous in X.
Conversely, let l be continuous in X. Then l is continuous at $0 \in X$ and so there are $l_1, \ldots, l_n \in \mathcal{L}$ and $\epsilon_1, \ldots, \epsilon_n > 0$ so that $|l(x)| < 1$ for every $x \in C_0$, where

$$C_0 = \{x \in X \,|\, |l_k(x)| < \epsilon_k \quad \text{for every } k = 1, \ldots, n\}.$$

Now, take any $x \in X$ such that $l_1(x) = \ldots = l_n(x) = 0$. Then for every $t > 0$ we have $l_1(tx) = \ldots = l_n(tx) = 0$ and hence $tx \in C_0$. Thus,

$$t|l(x)| = |l(tx)| < 1,$$

and letting $t \to +\infty$ we get $l(x) = 0$. Now, lemma 5.1 finishes the proof. $\square$

5.3 Weak topologies of normed spaces.

If X is a normed space, then theorem 3.10 implies that the collection $\mathcal{L} = X'$ of bounded linear functionals in X is separating. Indeed, let $x_1, x_2 \in X$, $x_1 \neq x_2$. Then

$$0 < \|x_1 - x_2\| = \max_{x' \in X', \|x'\| \leq 1} |x'(x_1 - x_2)|,$$

and so there is $x' \in X'$ such that $x'(x_1) - x'(x_2) = x'(x_1 - x_2) \neq 0$.

Definition. *Let X be a normed space. The topology $\sigma(X, X')$ is called* **weak topology** *of X. A subset of X which is open or closed or compact with respect to $\sigma(X, X')$ is called* **weakly open** *or* **weakly closed** *or* **weakly compact**, *respectively.*

According to proposition 5.11, a basic open neighborhood of $x \in X$ with respect to $\sigma(X, X')$ is

$$C_x = \{y \in X \,|\, |x'_k(y) - x'_k(x)| < \epsilon_k \text{ for every } k = 1, \ldots, n\},$$

where $n \in \mathbb{N}$, $x'_1, \ldots, x'_n \in X'$ and $\epsilon_1, \ldots, \epsilon_n > 0$ are arbitrary.

We know the following about $\sigma(X, X')$. All are consequences of propositions 5.12, 5.13, 5.14, 5.15 and 5.16.

(i) If X has its weak topology $\sigma(X, X')$, then the linear space operations of addition and multiplication on X are continuous.
(ii) $\sigma(X, X')$ is the weakest topology of X with respect to which every $x' \in X'$ is continuous.
(iii) Let D be a topological space and $g : D \to X$, and let X have its weak topology $\sigma(X, X')$. Then g is continuous if and only if $x' \circ g : D \to F$ is continuous for every $x' \in X'$.
(iv) $\sigma(X, X')$ is Hausdorff.
(v) If X has its weak topology $\sigma(X, X')$, then a linear functional $l : X \to F$ is continuous in X if and only if $l \in X'$.

We have exactly the same situation for X' and its dual X''. The weak topology on X' is $\sigma(X', X'')$. On the other hand, there is another interesting topology on X'.

Definition. *Let X be a normed space, and consider the natural embedding $J : X \to X''$. Then $J(X) \subseteq X''$ is a collection of linear functionals in X'. The topology $\sigma(X', J(X))$ is called* **weak$*$ topology** *on X'. Because of the identification of X with $J(X)$, the topology $\sigma(X', J(X))$ is traditionally denoted $\sigma(X', X)$. A subset of X' which is open or closed or compact with respect to $\sigma(X', X)$ is called* **weakly$*$ open** *or* **weakly$*$ closed** *or* **weakly$*$ compact***, respectively.*

A basic open neighborhood of $x' \in X'$ with respect to $\sigma(X', X'')$ is

$$C_{x'} = \{y' \in X' \,|\, |x''_k(y') - x''_k(x')| < \epsilon_k \text{ for every } k = 1, \ldots, n\},$$

for arbitrary $n \in \mathbb{N}$, $x''_1, \ldots, x''_k \in X''$ and $\epsilon_1, \ldots, \epsilon_n > 0$.

Also, a basic open neighborhood of $x' \in X'$ with respect to $\sigma(X', X) = \sigma(X', J(X))$ is

$$\begin{aligned} C_{x'} &= \{y' \in X' \,|\, |J(x_k)(y') - J(x_k)(x')| < \epsilon_k \text{ for every } k = 1, \ldots, n\} \\ &= \{y' \in X' \,|\, |y'(x_k) - x'(x_k)| < \epsilon_k \text{ for every } k = 1, \ldots, n\}, \end{aligned}$$

for arbitrary $n \in \mathbb{N}$, $x_1, \ldots, x_n \in X$ and $\epsilon_1, \ldots, \epsilon_n > 0$.

We have the following for $\sigma(X', X)$.

(i) If X' has its weak$*$ topology $\sigma(X', X)$, then the linear space operations of addition and multiplication on X' are continuous.

(ii) $\sigma(X', X)$ is the weakest topology of X' with respect to which every $x'' \in J(X)$ is continuous.

(iii) Let D be a topological space and $g : D \to X'$, and let X' have its weak$*$ topology $\sigma(X', X)$. Then g is continuous if and only if $J(x) \circ g : D \to F$ is continuous for every $x \in X$.

(iv) $\sigma(X', X)$ is Hausdorff.

(v) If X' has its weak$*$ topology $\sigma(X', X)$, then a linear functional $l : X' \to F$ is continuous in X if and only if $l \in J(X)$.

The fact that $\sigma(X', X) = \sigma(X', J(X))$ is Hausdorff follows from proposition 5.15, since $J(X)$ is separating. Indeed, if $x'_1, x'_2 \in X'$, $x'_1 \neq x'_2$, then there is $x \in X$ so that $x'_1(x) \neq x'_2(x)$ and hence $J(x)(x'_1) \neq J(x)(x'_2)$.

In a normed space X we have two topologies: the weak topology and the topology which is induced by the norm of X, which is also called **strong topology** on X.

In X' we have three topologies: the strong topology, the weak topology, and the weak$*$ topology. Clearly, *if X is reflexive, then the weak topology and the weak$*$ topology of X' are the same.*

In X'' we have two topologies: the strong topology and the weak$*$ topology.

Proposition 5.17. *Let X be a normed space.*

(i) $x_n \xrightarrow{\text{w}} x$ in X if and only if (x_n) converges to x with respect to the weak topology of X.

(ii) $x'_n \xrightarrow{\text{w}*} x'$ in X' if and only if (x'_n) converges to x' with respect to the weak$*$ topology of X'.

Proof. (i) Let $x_n \xrightarrow{\text{w}} x$ in X. We take any $U \in \sigma(X, X')$ such that $x \in U$. Then there are $x'_1, \ldots, x'_m \in X'$ and $\epsilon_1, \ldots, \epsilon_m > 0$ so that

$$C_x = \{y \in X \,|\, |x'_k(y) - x'_k(x)| < \epsilon_k \text{ for every } k = 1, \ldots, m\} \subseteq U.$$

Since $x'(x_n) \to x'(x)$ for every $x' \in X'$, there is n_0 so that $|x'_k(x_n) - x'_k(x)| < \epsilon_k$ for every $n \geq n_0$ and every $k = 1, \ldots, m$. This means that $x_n \in C_x \subseteq U$ for every $n \geq n_0$. Thus, (x_n) converges to x with respect to the weak topology of X.

Conversely, let (x_n) converge to x with respect to the weak topology of X. We take any $x' \in X'$ and the weakly open neighborhood $C_x = \{y \in X \,|\, |x'(y) - x'(x)| < \epsilon\}$ of x. Then there is n_0 so that $x_n \in C_x$ for every $n \geq n_0$, i.e. $|x'(x_n) - x'(x)| < \epsilon$ for every $n \geq n_0$. Thus, $x'(x_n) \to x'(x)$ for every $x' \in X'$, and so $x_n \xrightarrow{\text{w}} x$.

(ii) Similarly. $\square$

Proposition 5.18. *Let X be a normed space.*
(i) The weak topology of X is weaker than the strong topology of X.
(ii) The weak$*$ topology of X' is weaker than the weak topology of X' and this is weaker than the strong topology of X'.

Proof. (i) If X has its strong topology, then every $x' \in X'$ is continuous. Since $\sigma(X, X')$ is the weakest topology of X with respect to which every $x' \in X'$ is continuous, we get that $\sigma(X, X')$ is weaker than the strong topology of X.
(ii) That $\sigma(X', X'')$ is weaker than the strong topology of X' is an immediate consequence of (i). Now, every $x'' \in X''$ is continuous with respect to $\sigma(X', X'')$. In particular, every $x'' \in J(X) \subseteq X''$ is continuous with respect to $\sigma(X', X'')$. Since $\sigma(X', X) = \sigma(X', J(X))$ is the weakest topology of X' with respect to which every $x'' \in J(X)$ is continuous, we get that $\sigma(X', X)$ is weaker than $\sigma(X', X'')$. $\square$

Proposition 5.19. *Let X be a normed space.*
(i) If $K \subseteq X$ is weakly compact, then it is weakly closed and bounded.
(ii) If $K \subseteq X'$ is weakly$*$ compact, then it is weakly$*$ closed and, if X is a Banach space, bounded.

Proof. (i) Since X with the topology $\sigma(X, X')$ is Hausdorff, the weakly compact $K \subseteq X$ is weakly closed.
Every $x' \in X'$ is continuous in X, and hence in K, with respect to $\sigma(X, X')$. Since K is weakly compact, we get $\sup_{x\in K} |x'(x)| < +\infty$ for every $x' \in X'$. According to theorem 3.16, we have that $\sup_{x\in K} \|x\| < +\infty$, and so K is bounded.
(ii) Similarly. $\square$

The theorem of Alaoglou. *Let X be a normed space.*
(i) The closed unit ball of X' with center 0 is weakly$*$ compact.
(ii) If $K \subseteq X'$ is weakly$*$ closed and bounded, then it is weakly$*$ compact.

Proof. (i) Let $\overline{B}' = \{x' \in X' \,|\, \|x'\| \leq 1\}$ be the closed unit ball of X' with center 0. To prove that $\overline{B}'$ is weakly$*$ compact, we shall use proposition 5.10.
We consider any collection $\mathcal{F}$ of subsets of $\overline{B}'$ with the finite intersection property, and we shall prove that $\overline{B}' \cap \bigcap_{F\in\mathcal{F}} \mathrm{cl}(F) \neq \emptyset$. (The symbol $\mathrm{cl}(F)$ means the weakly$*$ closure of F.)
We consider

$$\mathbf{P} = \{\mathcal{G} \,|\, \mathcal{G} \supseteq \mathcal{F} \text{ is a collection of subsets of } \overline{B}' \text{ with the finite intersection property}\}.$$

We also consider the order relation of set inclusion in $\mathbf{P}$.
Now we take any totally ordered $\mathbf{P_0} \subseteq \mathbf{P}$, and we define

$$\mathcal{F}_0 = \bigcup_{\mathcal{G}\in\mathbf{P_0}} \mathcal{G}.$$

This is a collection of subsets of $\overline{B}'$ with the finite intersection property. Indeed, if we take any $G_1, \ldots, G_n \in \mathcal{F}_0$, then $G_1 \in \mathcal{G}_1, \ldots, G_n \in \mathcal{G}_n$ for some $\mathcal{G}_1, \ldots, \mathcal{G}_n \in \mathbf{P_0}$. Since $\mathbf{P_0}$ is totally ordered, there is one of $\mathcal{G}_1, \ldots, \mathcal{G}_n$ which includes all the others. Thus, $G_1, \ldots, G_n$ belong to one $\mathcal{G} \in \mathbf{P_0}$, and so $\bigcap_{k=1}^n G_k \neq \emptyset$. It is also clear that $\mathcal{F} \subseteq \mathcal{F}_0$. Therefore, $\mathcal{F}_0 \in \mathbf{P}$. Since $\mathcal{G} \subseteq \mathcal{F}_0$ for every $\mathcal{G} \in \mathbf{P}_0$, we conclude that $\mathcal{F}_0$ is an upper bound of $\mathbf{P}_0$ in $\mathbf{P}$.
According to the lemma of Zorn, $\mathbf{P}$ has a maximal element, i.e. there is a collection $\mathcal{G} \supseteq \mathcal{F}$ of subsets of $\overline{B}'$ with the finite intersection property, and so that there is no strictly larger collection with the same properties.
This implies that every intersection of finitely many elements of $\mathcal{G}$ belongs to $\mathcal{G}$. Indeed, if G is the intersection of finitely many elements of $\mathcal{G}$ so that $G \notin \mathcal{G}$, then $\mathcal{G}' = \mathcal{G} \cup \{G\} \supseteq \mathcal{F}$ is a collection of subsets of $\overline{B}'$ *with the finite intersection property*, and it is strictly larger than $\mathcal{G}$.

Now, since $\mathcal{F} \subseteq \mathcal{G}$, it is enough to prove that $\overline{B}' \cap \bigcap_{G \in \mathcal{G}} \mathrm{cl}(G) \neq \emptyset$.
For each $x \in X$ we consider the collection

$$\mathcal{G}_x = \{J(x)(G) \,|\, G \in \mathcal{G}\}$$

of subsets of F. For each $G \in \mathcal{G}$, we have $|J(x)(x')| = |x'(x)| \leq \|x\|$ for every $x' \in G$ (since $G \subseteq \overline{B}'$), and hence $J(x)(G) \subseteq \{\lambda \,|\, |\lambda| \leq \|x\|\} \subseteq F$. Thus, $\mathcal{G}_x$ is a collection of subsets of $\{\lambda \,|\, |\lambda| \leq \|x\|\}$. Now we take any $G_1, \ldots, G_n \in \mathcal{G}$. Then there exists $x' \in \bigcap_{k=1}^{n} G_k$, and so $J(x)(x') \in \bigcap_{k=1}^{n} J(x)(G_k)$. We conclude that $\mathcal{G}_x$ has the finite intersection property, and now the compactness of $\{\lambda \,|\, |\lambda| \leq \|x\|\}$ implies that $\{\lambda \,|\, |\lambda| \leq \|x\|\} \cap \bigcap_{G \in \mathcal{G}} \mathrm{cl}(J(x)(G)) \neq \emptyset$.
For every $x \in X$ we take any number

$$\mu_x \in \{\lambda \,|\, |\lambda| \leq \|x\|\} \cap \bigcap_{G \in \mathcal{G}} \mathrm{cl}(J(x)(G)).$$

Now, let D_x be any open neighborhood of μ_x in F. Since $\mu_x \in \bigcap_{G \in \mathcal{G}} \mathrm{cl}(J(x)(G))$, we have that D_x has non-empty intersection with $J(x)(G)$ for every $G \in \mathcal{G}$. So if we take any $G \in \mathcal{G}$, then there is $\kappa_x \in D_x \cap J(x)(G)$, and so there is $y' \in G$ so that $J(x)(y') = \kappa_x \in D_x$, i.e. $y' \in J(x)^{-1}(D_x) \cap G$. Thus, $J(x)^{-1}(D_x)$ has non-empty intersection with every $G \in \mathcal{G}$, and, since $G \subseteq \overline{B}'$, we have that $J(x)^{-1}(D_x) \cap \overline{B}'$ has non-empty intersection with every $G \in \mathcal{G}$. This implies that $\mathcal{G} \cup \{J(x)^{-1}(D_x) \cap \overline{B}'\} \supseteq \mathcal{F}$ is a collection of subsets of $\overline{B}'$ with the finite intersection property. Since $\mathcal{G}$ is a maximal collection with these properties, we get that

$$J(x)^{-1}(D_x) \cap \overline{B}' \in \mathcal{G}.$$

Now we take any $x_1, \ldots, x_n \in X$ and any open neighborhoods $D_{x_1}, \ldots, D_{x_n}$ of $\mu_{x_1}, \ldots, \mu_{x_n}$ in F. Then by the finite intersection property of $\mathcal{G}$ we have

$$\Big(\bigcap_{k=1}^{n} J(x_k)^{-1}(D_{x_k})\Big) \cap \overline{B}' = \bigcap_{k=1}^{n} \big(J(x_k)^{-1}(D_{x_k}) \cap \overline{B}'\big) \neq \emptyset, \tag{5.1}$$

and also

$$\Big(\bigcap_{k=1}^{n} J(x_k)^{-1}(D_{x_k})\Big) \cap G = \Big(\bigcap_{k=1}^{n} \big(J(x_k)^{-1}(D_{x_k}) \cap \overline{B}'\big)\Big) \cap G \neq \emptyset \quad \text{for every } G \in \mathcal{G}. \tag{5.2}$$

Now we take any $x_1, x_2 \in X$ and any $\lambda \in F$ and we consider the following open neighbrhoods of $\mu_{x_1}, \mu_{x_2}, \mu_{x_1+x_2}, \mu_{\lambda x_1}$ in F:

$$D_{x_1} = \{\kappa \,|\, |\kappa - \mu_{x_1}| < \epsilon\}, \quad D_{x_2} = \{\kappa \,|\, |\kappa - \mu_{x_2}| < \epsilon\},$$

$$D_{x_1+x_2} = \{\kappa \,|\, |\kappa - \mu_{x_1+x_2}| < \epsilon\}, \quad D_{\lambda x_1} = \{\kappa \,|\, |\kappa - \mu_{\lambda x_1}| < \epsilon\}.$$

Then (5.1), applied to $x_1, x_2, x_1+x_2, \lambda x_1 \in X$ and to the corresponding $D_{x_1}, D_{x_2}, D_{x_1+x_2}, D_{\lambda x_1}$, implies that there is $y' \in \overline{B}'$ so that

$$|y'(x_1) - \mu_{x_1}| < \epsilon, \quad |y'(x_2) - \mu_{x_2}| < \epsilon, \quad |y'(x_1+x_2) - \mu_{x_1+x_2}| < \epsilon, \quad |y'(\lambda x_1) - \mu_{\lambda x_1}| < \epsilon.$$

Since y' is linear and $\|y'\| \leq 1$, we easily prove that

$$|\mu_{x_1+x_2} - \mu_{x_1} - \mu_{x_2}| < 3\epsilon, \quad |\mu_{\lambda x_1} - \lambda\mu_{x_1}| < (1+|\lambda|)\epsilon, \quad |\mu_{x_1}| \leq \|x_1\| + \epsilon.$$

Finally, since ϵ is arbitrary, we get

$$\mu_{x_1+x_2} = \mu_{x_1} + \mu_{x_2}, \quad \mu_{\lambda x_1} = \lambda\mu_{x_1}, \quad |\mu_{x_1}| \leq \|x_1\|$$

for every $x_1, x_2 \in X$ and every $\lambda \in F$.
Now we consider $x' : X \to F$ defined for every $x \in X$ by $x'(x) = \mu_x$. Then

$$x'(x_1 + x_2) = x'(x_1) + x'(x_2), \quad x'(\lambda x_1) = \lambda x'(x_1), \quad |x'(x_1)| \leq \|x_1\|$$

for every $x_1, x_2 \in X$ and every $\lambda \in F$. This means that $x' \in \overline{B}'$.
Now, consider any weakly$*$ open neighborhood U of x'. Then there are $x_1, \ldots, x_n \in X$ and $\epsilon_1, \ldots, \epsilon_n > 0$ so that

$$C_{x'} = \{y' \in X' \,|\, |y'(x_k) - x'(x_k)| < \epsilon_k \text{ for every } k = 1, \ldots, n\} \subseteq U.$$

We apply (5.2) with $D_{x_k} = \{\kappa \,|\, |\kappa - x'(x_k)| < \epsilon_k\} = \{\kappa \,|\, |\kappa - \mu_{x_k}| < \epsilon_k\}$ for $k = 1, \ldots, n$, and we get

$$C_{x'} \cap G \neq \emptyset \quad \text{for every } G \in \mathcal{G}.$$

Therefore, $U \cap G \neq \emptyset$ for every $G \in \mathcal{G}$. Since this is true for every weakly$*$ open neighborhood U of x', we conclude that $x' \in \mathrm{cl}(G)$ for every $G \in \mathcal{G}$.
Hence $x' \in \overline{B}' \cap \bigcap_{G \in \mathcal{G}} \mathrm{cl}(G)$, and so $\overline{B}' \cap \bigcap_{G \in \mathcal{G}} \mathrm{cl}(G) \neq \emptyset$.
(ii) Let $K \subseteq X'$ be weakly$*$ closed and bounded. Then there is $M > 0$ so that $K \subseteq \overline{B}'(0; M)$, where $\overline{B}'(0; M)$ is the closed ball of X' with center 0 and radius M.
Now, $\overline{B}'(0; M)$ is the image of $\overline{B}' = \overline{B}'(0; 1)$ under multiplication by M. Since multiplication is a continuous function with respect to the weak$*$ topology of X' and since $\overline{B}'$ is weakly$*$ compact, we get that $\overline{B}'(0; M)$ is also weakly$*$ compact. Then K is a weakly$*$ closed subset of $\overline{B}'(0; M)$ and so it is weakly$*$ compact. $\square$

Chapter 6

Bounded linear operators

6.1 Bounded linear operators.

Definition. *Let X, Y be two normed spaces (over the same F), and let $T : X \to Y$ be a linear operator. Then T is called* **bounded**, *if there is $C \geq 0$ so that*

$$\|Tx\| \leq C\|x\|$$

for every $x \in X$.

It is more precise to write $\|Tx\|_Y \leq C\|x\|_X$, or something similar, in order to distinguish between the norms of the different spaces X, Y, but most of the time we shall adopt the simpler notation.

Proposition 6.1. *Let X, Y be normed spaces, and let $T : X \to Y$ be a linear operator. The following are equivalent:*
(i) T is continuous in X.
(ii) T is continuous at $0 \in X$.
(iii) T is bounded.

Proof. Let T be continuous at $0 \in X$. Then there is $\delta > 0$ so that $\|T(x)\| < 1$ for every $x \in X$ with $\|x\| < \delta$. Now, if $x \neq 0$, then $y = \frac{\delta}{2\|x\|}\, x$ satisfies $\|y\| < \delta$ and hence

$$\|T(x)\| = \frac{2\|x\|}{\delta}\|T(y)\| < \frac{2}{\delta}\,\|x\|.$$

The inequality $\|T(x)\| \leq \frac{2}{\delta}\|x\|$ is obviously true also for $x = 0$, and so T is bounded.
If T is bounded, then there is $C \geq 0$ so that $\|T(x)\| \leq C\|x\|$ for every $x \in X$. So, if $x_n \to x$ in X, then

$$\|T(x_n) - T(x)\| = \|T(x_n - x)\| \leq C\|x_n - x\| \to 0$$

and so $T(x_n) \to T(x)$. Hence T is continuous in X. □

Proposition 6.2. *Let X, Y be normed spaces, and let $T : X \to Y$ be a linear operator. If T is continuous, then $\mathrm{N}(T)$ is closed in X.*

Proof. $\mathrm{N}(T) = T^{-1}(\{0\})$ is the inverse image of a closed set, and so, if T is continuous in X, then $\mathrm{N}(T)$ is closed in X. □

Definition. *Let X, Y be normed spaces. The set of all continuous or, equivalently, bounded linear operators $T : X \to Y$ is denoted $L(X, Y)$.*
If $Y = X$, then we denote $L(X)$ instead of $L(X, X)$.

If $Y = F$, then, obviously, $L(X, F) = X'$.

Proposition 6.3. *Let X, Y be normed spaces. Then $L(X, Y)$ as a function space, with the usual addition of functions and the usual multiplication of numbers and functions, is a linear space.*

Proof. If $T, T_1, T_2 : X \to Y$ and $\lambda \in F$, we consider the functions $T_1 + T_2 : X \to Y$ and $\lambda T : X \to Y$ defined for every $x \in X$ by

$$(T_1 + T_2)(x) = T_1(x) + T_2(x), \quad (\lambda T)(x) = \lambda T(x).$$

It is known from Linear Algebra that, if T, T_1, T_2 are linear operators, then $T_1 + T_2$ and λT are also linear operators. It is also clear that, if T, T_1, T_2 are continuous, then $T_1 + T_2$ and λT are also continuous. $\square$

Definition. *Let X, Y be normed spaces. For every $T \in L(X, Y)$ we define*

$$\|T\| = \sup_{x \in X, \|x\| \leq 1} \|T(x)\|.$$

Proposition 6.4. *Let X, Y be normed spaces and let $T \in L(X, Y)$. Then $\|T\|$ is the smallest constant C which satifies the inequality $\|T(x)\| \leq C\|x\|$ for every $x \in X$.*

Proof. For every $x \in X$, $x \neq 0$, we have $\left\|\frac{x}{\|x\|}\right\| = 1$, and then, by the definition of $\|T\|$, we get

$$\|T(x)\| = \left\|T\left(\frac{x}{\|x\|}\right)\right\|\|x\| \leq \|T\|\|x\|.$$

The inequality $\|T(x)\| \leq \|T\|\|x\|$ is obviously satistied also if $x = 0$, and so $C = \|T\|$ satisfies the inequality $\|T(x)\| \leq C\|x\|$ for every $x \in X$.
Conversely, let C satisfy the inequality $\|T(x)\| \leq C\|x\|$ for every $x \in X$. Then we have $\|T(x)\| \leq C$ for every $x \in X$ with $\|x\| \leq 1$, and so $\|T\| \leq C$. $\square$

So, if $T \in L(X, Y)$, then

$$\|T(x)\| \leq \|T\|\|x\| \quad \text{for every } x \in X.$$

Also,

$$\|T(x)\| \leq C\|x\| \text{ for every } x \in X \quad \Rightarrow \quad \|T\| \leq C.$$

Proposition 6.5. *Let X, Y be normed spaces. The function $\|\cdot\| : L(X, Y) \to \mathbb{R}$ defined above is a norm on $L(X, Y)$. If Y is a Banach space, then $L(X, Y)$ with this norm is a Banach space.*

Proof. Obviously, $\|T\| \geq 0$ for every $T \in L(X, Y)$. It is also clear that $\|T\| = 0$ if $T = 0$.
If $T \in L(X, Y)$ and $\|T\| = 0$, then $T(x) = 0$ for every $x \in X$, and so $T = 0$.
For every $x \in X$ and every $T_1, T_2 \in L(X, Y)$ we have

$$\|(T_1 + T_2)(x)\| \leq \|T_1(x)\| + \|T_2(x)\| \leq \|T_1\|\|x\| + \|T_2\|\|x\| = (\|T_1\| + \|T_2\|)\|x\|.$$

Hence $\|T_1 + T_2\| \leq \|T_1\| + \|T_2\|$.
For every $T \in L(X, Y)$ and every $\lambda \in F$ we have

$$\|\lambda T\| = \sup_{x \in X, \|x\| \leq 1} \|(\lambda T)(x)\| = \sup_{x \in X, \|x\| \leq 1} |\lambda|\|T(x)\| = |\lambda| \sup_{x \in X, \|x\| \leq 1} \|T(x)\| = |\lambda|\|T\|.$$

Therefore, $\|\cdot\| : L(X, Y) \to \mathbb{R}$ is a norm on $L(X, Y)$.
Now we assume that Y is a Banach space, and we take any Cauchy sequence (T_n) in $L(X, Y)$. For every $x \in X$ we have

$$\|T_n(x) - T_m(x)\| = \|(T_n - T_m)(x)\| \leq \|T_n - T_m\|\|x\| \to 0$$

when $n, m \to +\infty$, and so $(T_n(x))$ is a Cauchy sequence in Y. Since Y is complete, the sequence $(T_n(x))$ has a limit in Y. Now we consider the function $T : X \to Y$ defined for every $x \in X$ by

$$T(x) = \lim_{n \to +\infty} T_n(x) \in Y.$$

Since each T_n is a linear operator, we have for every $x, z \in X$ and $\lambda \in F$ that

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

$$T(\lambda x) = \lim_{n \to +\infty} T_n(\lambda x) = \lambda \lim_{n \to +\infty} T_n(x) = \lambda T(x).$$

So T is a linear operator.
Now, there is n_0 so that $\|T_n - T_m\| \leq 1$ for every $n, m \geq n_0$. Hence

$$\|T_n(x)\| \leq \|T_n(x) - T_{n_0}(x)\| + \|T_{n_0}(x)\| \leq \|T_n - T_{n_0}\|\|x\| + \|T_{n_0}\|\|x\| \leq (1 + \|T_{n_0}\|)\|x\|$$

for every $n \geq n_0$ and every $x \in X$. Taking the limit when $n \to +\infty$, we find

$$\|T(x)\| \leq (1 + \|T_{n_0}\|)\|x\|$$

for every $x \in X$. So T is bounded, i.e. $T \in L(X, Y)$.
Finally, we take any $\epsilon > 0$ and then there is n_0 so that $\|T_n - T_m\| \leq \epsilon$ for every $n, m \geq n_0$. Then

$$\|T_n(x) - T_m(x)\| \leq \|T_n - T_m\|\|x\| \leq \epsilon\|x\|$$

for every $n, m \geq n_0$ and every $x \in X$. Taking the limit when $m \to +\infty$, we find

$$\|T_n(x) - T(x)\| \leq \epsilon\|x\|$$

for every $n \geq n_0$ and every $x \in X$. Therefore, $\|T_n - T\| \leq \epsilon$ for every $n \geq n_0$, and so $T_n \to T$ in $L(X, Y)$. □

Example 6.1.1. If $T : X \to Y$ is a linear isometry from X into Y, then $\|T\| = 1$. Indeed,

$$\|T\| = \sup_{x \in X, \|x\| \leq 1} \|T(x)\| = \sup_{x \in X, \|x\| \leq 1} \|x\| = 1.$$

Example 6.1.2. Let X be an inner product space, let Y be a subspace of X with an orthogonal complement in X, and let $P_Y : X \to X$ be the orthogonal projection of X onto Y. If $Y \neq \{0\}$, then $\|P_Y\| = 1$.
Indeed, for every $x \in X$ we have $\|P_Y(x)\| \leq \|x\|$ and this shows that $\|P_Y\| \leq 1$. Also, for every $y \in Y$, $y \neq 0$, we have

$$\|y\| = \|P_Y(y)\| \leq \|P_Y\|\|y\|,$$

and so $\|P_Y\| \geq 1$. Therefore, $\|P_Y\| = 1$.

Proposition 6.6. *Let X, Y, Z be normed spaces, $T \in L(X, Y)$ and $S \in L(Y, Z)$. Then $S \circ T \in L(X, Z)$, and $\|S \circ T\| \leq \|S\|\|T\|$.*

Proof. For every $x \in X$ we have

$$\|(S \circ T)x\| = \|S(T(x))\| \leq \|S\|\|T(x)\| \leq \|S\|\|T\|\|x\|.$$

So $S \circ T$ is bounded, and $\|S \circ T\| \leq \|S\|\|T\|$. □

We shall use the notation

$$ST \quad \text{instead of} \quad S \circ T.$$

The next proposition is useful when it is convenient to work with a one-to-one operator.

Proposition 6.7. *Let X, Y be normed spaces, and $T \in L(X,Y)$. We consider $\widetilde{T} : X/N(T) \to Y$ defined for every $\xi \in X/N(T)$ by $\widetilde{T}(\xi) = T(x)$, where $x \in X$ is such that $[x] = \xi$. Then $\widetilde{T}$ is a bounded linear operator. Also, $\widetilde{T}$ is one-to-one,* $\mathrm{R}(\widetilde{T}) = \mathrm{R}(T)$, *and* $\|\widetilde{T}\| = \|T\|$.

Proof. If $x_1, x_2 \in X$ are such that $[x_1] = [x_2] = \xi$, then $x_1 - x_2 \in N(T)$, and so $T(x_1) - T(x_2) = T(x_1 - x_2) = 0$. Thus, $\widetilde{T}$ is well defined.
If $[x_1] = \xi_1$, $[x_2] = \xi_2$, then $[x_1 + x_2] = [x_1] + [x_2] = \xi_1 + \xi_2$. Also, if $[x] = \xi$ and $\lambda \in F$, then $[\lambda x] = \lambda[x] = \lambda\xi$. Then

$$\widetilde{T}(\xi_1 + \xi_2) = T(x_1 + x_2) = T(x_1) + T(x_2) = \widetilde{T}(\xi_1) + \widetilde{T}(\xi_2),$$

$$\widetilde{T}(\lambda\xi) = T(\lambda x) = \lambda T(x) = \lambda\widetilde{T}(\xi).$$

Thus, $\widetilde{T}$ is linear. It is also clear that $\mathrm{R}(\widetilde{T}) = \mathrm{R}(T)$.
If $[x] = \xi$, we have

$$\|\widetilde{T}(\xi)\| = \|T(x)\| \le \|T\|\|x\|.$$

Taking the infimum over all x such that $[x] = \xi$, we get $\|\widetilde{T}(\xi)\| \le \|T\|\|\xi\|$. So $\widetilde{T}$ is bounded, and $\|\widetilde{T}\| \le \|T\|$.
Also, for every $x \in X$ we take $\xi = [x]$, and then

$$\|T(x)\| = \|\widetilde{T}(\xi)\| \le \|\widetilde{T}\|\|\xi\| \le \|\widetilde{T}\|\|x\|.$$

Hence $\|T\| \le \|\widetilde{T}\|$, and we conclude that $\|\widetilde{T}\| = \|T\|$. □

Proposition 6.8. *Let X, Y be normed spaces, and $T \in L(X,Y)$. Assume that $\overline{X}, \overline{Y}$ are completions of X, Y, i.e. there are linear isometries $S_X : X \to \overline{X}$ and $S_Y : Y \to \overline{Y}$ so that $S_X(X)$ is a dense subspace of $\overline{X}$ and $S_Y(Y)$ is a dense subspace of $\overline{Y}$. Then there is a unique $\overline{T} \in L(\overline{X}, \overline{Y})$ such that $\overline{T}S_X = S_Y T$. Also, $\|\overline{T}\| = \|T\|$.*

Proof. Take any $\xi \in \overline{X}$. Then there is a sequence $(S_X(x_n))$ in $S_X(X)$ such that $S_X(x_n) \to \xi$ in $\overline{X}$. Then $(S_X(x_n))$ is a Cauchy sequence and, since

$$\|x_n - x_m\| = \|S_X(x_n - x_m)\| = \|S_X(x_n) - S_X(x_m)\| \to 0,$$

we have that (x_n) is a Cauchy sequence in X. Now,

$$\begin{aligned}\|(S_Y T)(x_n) - (S_Y T)(x_m)\| &= \|S_Y(T(x_n)) - S_Y(T(x_m))\| = \|S_Y(T(x_n) - T(x_m))\| \\ &= \|T(x_n) - T(x_m)\| \le \|T\|\|x_n - x_m\| \to 0.\end{aligned}$$

Thus, $((S_Y T)(x_n))$ is a Cauchy sequence in $\overline{Y}$, and so it converges to some element of $\overline{Y}$.
Now we consider the function $\overline{T} : \overline{X} \to \overline{Y}$ defined for every $\xi \in \overline{X}$ by

$$\overline{T}(\xi) = \lim_{n\to+\infty} (S_Y T)(x_n).$$

It is easy to see that $\overline{T}(\xi)$ is well defined, i.e. that it depends on ξ and not on the sequence (x_n). Moreover, using the linearity of T, S_X, S_Y, it is very easy to show that $\overline{T}$ is linear.
Also, for every $\xi \in \overline{X}$,

$$\begin{aligned}\|\overline{T}(\xi)\| &= \lim_{n\to+\infty} \|(S_Y T)(x_n)\| = \lim_{n\to+\infty} \|T(x_n)\| \le \lim_{n\to+\infty} \|T\|\|x_n\| = \lim_{n\to+\infty} \|T\|\|S_X(x_n)\| \\ &= \|T\|\|\xi\|.\end{aligned}$$

This says that $\overline{T} \in L(\overline{X}, \overline{Y})$ and $\|\overline{T}\| \le \|T\|$.
If $\xi \in S_X(X)$, then $\xi = S_X(x)$ for some $x \in X$. Then we may take the constant sequence (x) to define $\overline{T}(\xi)$, and then

$$\overline{T}(S_X(x)) = \overline{T}(\xi) = \lim_{n\to+\infty} (S_Y T)(x) = (S_Y T)(x).$$

Therefore $\overline{T}S_X = S_Y T$.
Now, take any $\overline{\overline{T}} \in L(\overline{X}, \overline{Y})$ such that $\overline{\overline{T}}S_X = S_Y T$. Then for each $\xi \in \overline{X}$ we take, as above, a sequence $(S_X(x_n))$ in $S_X(X)$ such that $S_X(x_n) \to \xi$ in $\overline{X}$, and we get

$$\overline{\overline{T}}(\xi) = \lim_{n\to+\infty} \overline{\overline{T}}(S_X(x_n)) = \lim_{n\to+\infty} (S_Y T)(x_n) = \overline{T}(\xi).$$

Thus, $\overline{\overline{T}} = \overline{T}$.
Finally, since $S_X(X) \subseteq \overline{X}$, we get

$$\begin{aligned}\|\overline{T}\| &= \sup_{\xi\in\overline{X},\|\xi\|\le 1} \|\overline{T}(\xi)\| \geq \sup_{x\in X,\|S_X(x)\|\le 1} \|\overline{T}(S_X(x))\| = \sup_{x\in X,\|x\|\le 1} \|S_Y(Tx)\| \\ &= \sup_{x\in X,\|x\|\le 1} \|Tx\| = \|T\|.\end{aligned}$$

Thus, $\|\overline{T}\| \geq \|T\|$, and we conclude that $\|\overline{T}\| = \|T\|$. □

The relation $\overline{T}S_X = S_Y T$ means, of course, that

$$\overline{T}(S_X(x)) = S_Y(T(x))$$

for every $x \in X$. Now, we may "identify" X with the subspace $S_X(X)$ of $\overline{X}$, and Y with the subspace $S_Y(Y)$ of $\overline{Y}$, by "identifying" every $x \in X$ with the corresponding $S_X(x) \in \overline{X}$, and every $y \in Y$ with the corresponding $S_Y(y) \in \overline{Y}$. Then the above relation becomes

$$\overline{T}(x) = T(x)$$

for every $x \in X$. In other words, it appears as if the operator $\overline{T} \in L(\overline{X}, \overline{Y})$ extends the operator $T \in L(X, Y)$.

6.2 The dual operator.

Proposition 6.9. *Let X, Y be normed spaces, and $T \in L(X,Y)$. We consider $T' : Y' \to X'$ defined for every $y' \in Y'$ by*

$$T'(y') = y' \circ T.$$

Then $T' \in L(Y', X')$, and $\|T'\| = \|T\|$.

Proof. For every $y', y_1', y_2' \in Y'$ and every $\lambda \in F$ we have

$$T'(y_1') + T'(y_2') = y_1' \circ T + y_2' \circ T = (y_1' + y_2') \circ T = T'(y_1' + y_2'),$$

$$T'(\lambda y') = (\lambda y') \circ T = \lambda(y' \circ T) = \lambda T'(y'),$$

and so $T' : Y' \to X'$ is a linear operator.
Take any $y' \in Y'$. Then

$$\|T'(y')\| = \|y' \circ T\| \leq \|y'\|\|T\|.$$

Therefore, $T' \in L(Y', X')$ and $\|T'\| \leq \|T\|$.
Now we take any $x \in X$. According to theorem 3.10, there is $y' \in Y'$ so that $\|y'\| \leq 1$ and $\|T(x)\| = |y'(T(x))|$. Then

$$\|T(x)\| = |(y' \circ T)(x)| = \|T'(y')(x)\| \leq \|T'(y')\|\|x\| \leq \|T'\|\|y'\|\|x\| \leq \|T'\|\|x\|.$$

This implies $\|T\| \leq \|T'\|$, and hence $\|T'\| = \|T\|$. □

Definition. *Let X, Y be normed spaces, and $T \in L(X,Y)$. The operator $T' \in L(Y', X')$ defined in proposition 6.9 is called* **dual** *of T.*

The defining relation $T'(y') = y' \circ T$ means that

$$T'(y')(x) = y'(T(x)) \quad \text{for every } x \in X,\ y' \in Y'.$$

Proposition 6.10. *Let X, Y, Z be normed spaces.*
(i) If $T, T_1, T_2 \in L(X,Y)$ and $\lambda \in F$, then $(T_1 + T_2)' = T_1' + T_2'$ and $(\lambda T)' = \lambda T'$.
(ii) If $T \in L(X,Y)$ and $S \in L(Y,Z)$, then $(ST)' = T'S'$.
(iii) $I' = I$ and $0' = 0$, where I is the identity operator and 0 is the zero operator.
(iv) If $T \in L(X,Y)$ and $T^{-1} \in L(Y,X)$, then $(T')^{-1} = (T^{-1})'$.

Proof. Exercise. □

We recall the definition of $A^\perp$. If X is a normed space, and $A \subseteq X$, we define

$$A^\perp = \{x' \in X' \,|\, x'(a) = 0 \text{ for every } a \in A\}.$$

Here is a similar notion.

Definition. *Let X be a normed space. If $A \subseteq X'$, we define*

$${}^\perp A = \{x \in X \,|\, a(x) = 0 \textit{ for every } a \in A\}.$$

Proposition 6.11. *Let X be a normed space. If $A \subseteq X'$, then ${}^\perp A$ is a closed subspace of X.*

Proof. Exercise. □

If $A \subseteq X$, then $A^\perp \subseteq X'$. If $A \subseteq X'$, then ${}^\perp A \subseteq X$.

Proposition 6.12. *Let X be a normed space.*
(i) If $A \subseteq X$, then $\text{clspan}(A) = {}^\perp(A^\perp)$.
(ii) If $A \subseteq X'$, then $\text{clspan}(A) \subseteq ({}^\perp A)^\perp$.

Proof. (i) This is the content of theorem 3.11. This theorem says: $x \in \text{clspan}(A)$ if and only if $x'(x) = 0$ for every $x' \in A^\perp$. Equivalently: $x \in \text{clspan}(A)$ if and only if $x \in {}^\perp(A^\perp)$.
(ii) Take any $x' \in A$. Then for every $x \in {}^\perp A$ we have $x'(x) = 0$, and so $x' \in ({}^\perp A)^\perp$. Hence $A \subseteq ({}^\perp A)^\perp$. Since $({}^\perp A)^\perp$ is a closed subspace of X', we get $\text{clspan}(A) \subseteq ({}^\perp A)^\perp$. □

Proposition 6.13. *Let X, Y be normed spaces, and $T \in L(X,Y)$. Then:*
(i) $\text{N}(T') = \text{R}(T)^\perp$.
(ii) $\text{N}(T) = {}^\perp\text{R}(T')$.
(iii) $\text{cl}(\text{R}(T')) \subseteq \text{N}(T)^\perp$.
(iv) $\text{cl}(\text{R}(T)) = {}^\perp\text{N}(T')$.
Therefore:
(v) T' is one-to-one if and only if $\text{R}(T)$ is dense in Y.
(vi) T is one-to-one if $\text{R}(T')$ is dense in X'.

Proof. (i) $y' \in \text{N}(T')$ if and only if $T'(y') = 0$ if and only if $y' \circ T = 0$ if and only if $y'(T(x)) = 0$ for every $x \in X$ if and only if $y' \in \text{R}(T)^\perp$.
(ii) $x \in \text{N}(T)$ if and only if $T(x) = 0$ if (theorem 3.10) and only if $y'(T(x)) = 0$ for every $y' \in Y'$ if and only if $T'(y')(x) = 0$ for every $y' \in Y'$ if and only if $x \in {}^\perp\text{R}(T')$.
(iii) Direct implication of (ii) and of (ii) of proposition 6.12.
(iv) Direct implication of (i) and of (i) of proposition 6.12.
(v) Use (i) and (iv).
(vi) Use (ii) and (iii). □

6.3 Finite dimensional spaces.

Let X, Y be two finite dimensional linear spaces, and take any basis $B = \{b_1, \dots, b_n\}$ of X and any basis $C = \{c_1, \dots, c_m\}$ of Y. We know from Linear Algebra that to every linear operator $T : X \to Y$ corresponds the $m \times n$ matrix

$$[T]_{BC} = [a_{ij}],$$

where the $a_{ij} \in F$ are determined by the relations

$$T(b_j) = \sum_{i=1}^{m} a_{ij} c_i, \quad j = 1, \dots, n.$$

Conversely, every $m \times n$ matrix $[a_{ij}]$ determines a linear operator $T : X \to Y$ such that $[T]_{BC} = [a_{ij}]$. Therefore, the linear space of all linear operators $T : X \to Y$ is in a one-to-one correspondence with the linear space of all $m \times n$ matrices $[a_{ij}]$ through the mapping $T \mapsto [T]_{BC}$.

If to every $x \in X$ we assign the $n \times 1$ matrix $[x]_B = [\lambda_j]$, where the λ_j are determined by $x = \sum_{j=1}^{n} \lambda_j b_j$, and to every $y \in Y$ we assign the $m \times 1$ matrix $[y]_C = [\kappa_i]$, where the κ_i are determined by $y = \sum_{i=1}^{m} \kappa_i c_i$, then

$$y = T(x) \quad \Leftrightarrow \quad [y]_C = [T]_{BC}[x]_B.$$

We also know that for every linear operators $T, S : X \to Y$ and every $\lambda \in F$ we have

$$[\lambda T]_{BC} = \lambda\, [T]_{BC}, \quad [T + S]_{BC} = [T]_{BC} + [S]_{BC}.$$

Thus, the mapping $T \mapsto [T]_{BC}$ is a linear space isomorphism between the linear space of all linear operatots $T : X \to Y$ and the linear space of all $m \times n$ matrices $[a_{ij}]$.

If Z is another finite dimensional linear space, with a basis $D = \{d_1, \dots, d_l\}$, then for every linear operators $T : X \to Y$ and $S : Y \to Z$ we have

$$[ST]_{BD} = [S]_{CD}[T]_{BC}.$$

Now let $B' = \{b'_1, \dots, b'_n\}$ be the basis of X' which is dual to the basis B of X. Also let $C' = \{c'_1, \dots, c'_m\}$ be the basis of Y' which is dual to the basis C of Y. Then the relation between the matrices of the linear operator $T : X \to Y$ and of the dual linear operator $T' : Y' \to X'$ is

$$[T']_{C'B'} = ([T]_{BC})',$$

where $[a_{ij}]' = [a_{ji}]$ is the transpose matrix of $[a_{ij}]$.

It is obvious that, if $I : X \to X$ is the identity operator, then $[I]_{BB} = [\delta_{ij}]$ is the unit matrix, where $\delta_{ij} = 1$, if $i = j$, and $\delta_{ij} = 0$, if $i \neq j$. Also, if $0 : X \to Y$ is the zero linear operator, then $[0]_{BC} = [0]$ is the zero matrix.

Finally, in the case $m = n$, the linear operator $T : X \to Y$ is invertible if and only if $[T]_{BC}$ is an invertible matrix, and then

$$([T]_{BC})^{-1} = [T^{-1}]_{CB}.$$

In fact, $[T]_{BC}$ is an invertible matrix if and only if $\det([T]_{BC}) \neq 0$.

Everything we have said up to this point is known from Linear Algebra. Now we shall see that every linear operator $T : X \to Y$ is bounded. We assume that X, Y have arbitrary norms, and then for every $x \in X$ with $x = \sum_{j=1}^{n} \lambda_j b_j$ we get

$$\begin{aligned}\|T(x)\| &= \Big\| \sum_{j=1}^{n} \lambda_j T(b_j) \Big\| \leq \sum_{j=1}^{n} |\lambda_j| \|T(b_j)\| \leq \max_{1 \leq j \leq n} \|T(b_j)\| \sum_{j=1}^{n} |\lambda_j| = \max_{1 \leq j \leq n} \|T(b_j)\| \|x\|_1 \\ &\leq c \max_{1 \leq j \leq n} \|T(b_j)\| \|x\| = C\, \|x\|,\end{aligned}$$

where $c > 0$ is a constant such that $\|x\|_1 \leq c\|x\|$ for every $x \in X$, and $C = c\ \max_{1\leq j\leq n} \|T(b_j)\|$. The existence of such a constant c is implied by the equivalence of the arbitrary norm $\|\cdot\|$ with the 1-norm $\|\cdot\|_1$.

If we consider the p-norm of X and the q-norm of Y, where $1 \leq p, q \leq +\infty$, then we usually denote $\|T\|_{pq}$ the norm of a linear operator $T : X \to Y$, i.e.

$$\|T\|_{pq} = \sup_{x\in X, \|x\|_p \leq 1} \|T(x)\|_q.$$

Since all norms of X are pairwise equivalent and all norms of Y are also pairwise equivalent, it is easy to show that all norms of $L(X, Y)$ are equivalent. This can also be proven in another way. We have seen that there is a linear space isomorphism between $L(X, Y)$ and the linear space M_{mn} of all $m \times n$ matrices. This implies that $L(X, Y)$ is finite dimensional, with dimension equal to mn. Therefore every two norms of $L(X, Y)$ are equivalent.

To get an idea of this kind of calculations, we shall find the exact value of the norm

$$\|T\|_{\infty\infty} = \sup_{x\in X, \|x\|_\infty \leq 1} \|T(x)\|_\infty$$

of a linear operator $T : X \to Y$.

If $[T]_{BC} = [a_{ij}]$, then for every $x = \sum_{j=1}^n \lambda_j b_j \in X$ we have

$$T(x) = \sum_{j=1}^n \lambda_j T(b_j) = \sum_{j=1}^n \lambda_j \Big(\sum_{i=1}^m a_{ij} c_i\Big) = \sum_{i=1}^m \Big(\sum_{j=1}^n a_{ij}\lambda_j\Big) c_i,$$

and so

$$\|T(x)\|_\infty = \max_{1\leq i\leq m} \Big|\sum_{j=1}^n a_{ij}\lambda_j\Big| \leq \max_{1\leq i\leq m} \sum_{j=1}^n |a_{ij}| \max_{1\leq j\leq n} |\lambda_j| = \max_{1\leq i\leq m} \sum_{j=1}^n |a_{ij}|\, \|x\|_\infty.$$

Therefore, $\|T\|_{\infty\infty} \leq \max_{1\leq i\leq m} \sum_{j=1}^n |a_{ij}|$.

Now, there is i_0 so that

$$\sum_{j=1}^n |a_{i_0 j}| = \max_{1\leq i\leq m} \sum_{j=1}^n |a_{ij}|$$

and then we choose $\lambda_1, \ldots, \lambda_n$ so that $|\lambda_j| = 1$ and $a_{i_0 j}\lambda_j = |a_{i_0 j}|$ for $j = 1, \ldots, n$. Then for the particular $x = \sum_{j=1}^n \lambda_j b_j$ we have

$$\|x\|_\infty = \max_{1\leq j\leq n} |\lambda_j| = 1$$

and so

$$\|T\|_{\infty\infty} \geq \|T(x)\|_\infty = \max_{1\leq i\leq m} \Big|\sum_{j=1}^n a_{ij}\lambda_j\Big| \geq \Big|\sum_{j=1}^n a_{i_0 j}\lambda_j\Big| = \sum_{j=1}^n |a_{i_0 j}| = \max_{1\leq i\leq m} \sum_{j=1}^n |a_{ij}|.$$

We conclude that

$$\|T\|_{\infty\infty} = \max_{1\leq i\leq m} \sum_{j=1}^n |a_{ij}|.$$

6.4 Hilbert spaces.

In a Hilbert space, besides the notion of dual operator, we also have the notion of *adjoint operator*.

Proposition 6.14. *Let X, Y be Hilbert spaces, and $T \in L(X,Y)$. Then there is $T^* \in L(Y,X)$ such that*

$$\langle x, T^*(y)\rangle = \langle T(x), y\rangle \tag{6.1}$$

for every $x \in X$, $y \in Y$. Moreover, $\|T^*\| = \|T\|$.

Proof. We take any $y \in Y$ and we consider the function $l_y : X \to F$ defined for every $x \in X$ by

$$l_y(x) = \langle T(x), y\rangle.$$

It is clear that l_y is a linear functional in X. Also

$$|l_y(x)| = |\langle T(x), y\rangle| \le \|T(x)\|\|y\| \le \|T\|\|y\|\|x\|$$

for every $x \in X$, and so $l_y \in X'$.
According to the theorem of F. Riesz, there is an element of X, which we denote $T^*(y)$, such that $\langle x, T^*(y)\rangle = l_y(x)$, i.e.

$$\langle x, T^*(y)\rangle = \langle T(x), y\rangle$$

for every $x \in X$.
Now, for every $y_1, y_2 \in Y$ we get

$$\begin{aligned}\langle x, T^*(y_1 + y_2)\rangle &= \langle T(x), y_1 + y_2\rangle = \langle T(x), y_1\rangle + \langle T(x), y_2\rangle = \langle x, T^*(y_1)\rangle + \langle x, T^*(y_2)\rangle \\ &= \langle x, T^*(y_1) + T^*(y_2)\rangle\end{aligned}$$

for every $x \in X$. This implies $T^*(y_1 + y_2) = T^*(y_1) + T^*(y_2)$. In a similar manner we may show that $T^*(\lambda y) = \lambda T^*(y)$ for every $y \in Y$ and every $\lambda \in F$. Therefore, $T^* : Y \to X$ is linear. Also,

$$\|T^*(y)\|^2 = \langle T^*(y), T^*(y)\rangle = \langle T(T^*(y)), y\rangle \le \|T(T^*(y))\|\|y\| \le \|T\|\|T^*(y)\|\|y\|$$

and hence $\|T^*(y)\| \le \|T\|\|y\|$ for every $y \in Y$. Therefore, T^* is bounded, with $\|T^*\| \le \|T\|$.
Symmetrically,

$$\|T(x)\|^2 = \langle T(x), T(x)\rangle = \langle x, T^*(T(x))\rangle \le \|x\|\|T^*(T(x))\| \le \|x\|\|T^*\|\|T(x)\|$$

and hence $\|T(x)\| \le \|T^*\|\|x\|$ for every $x \in X$. Thus, $\|T\| \le \|T^*\|$, and we conclude that $\|T^*\| = \|T\|$. □

Definition. *Let X, Y be Hilbert spaces, and $T \in L(X,Y)$. The operator $T^* \in L(Y,X)$ defined in proposition 6.14 is called* **adjoint** *of T.*

Proposition 6.15. *Let X, Y, Z be Hilbert spaces.*
(i) If $T, T_1, T_2 \in L(X,Y)$ and $\lambda \in F$, then $(T_1 + T_2)^* = T_1^* + T_2^*$ and $(\lambda T)^* = \overline{\lambda} T^*$.
(ii) If $T \in L(X,Y)$, then $(T^*)^* = T$.
(iii) If $T \in L(X,Y)$ and $S \in L(Y,Z)$, then $(ST)^* = T^* S^*$.
(iv) If $T \in L(X,Y)$ and $T^{-1} \in L(Y,X)$, then $(T^*)^{-1} = (T^{-1})^*$.

Proof. Exercise. □

Proposition 6.16. *Let X, Y be Hilbert spaces, and $T \in L(X,Y)$. Then $\|T^*T\| = \|TT^*\| = \|T\|^2$.*

Proof. We have $\|T^*T\| \le \|T^*\|\|T\| = \|T\|^2$.
Also, for every $x \in X$,

$$\begin{aligned}\|T(x)\|^2 &= \langle T(x), T(x)\rangle = \langle x, T^*(T(x))\rangle \le \|x\|\|(T^*T)(x)\| \le \|x\|\|T^*T\|\|x\| \\ &= \|T^*T\|\|x\|^2,\end{aligned}$$

and so $\|T\|^2 \le \|T^*T\|$. Therefore, $\|T^*T\| = \|T\|^2$.
The equality $\|TT^*\| = \|T\|^2$ can be proved either in the same manner or by using T^* in the place of T in $\|T^*T\| = \|T\|^2$. □

To find the relation between the notions of the dual operator $T' \in L(Y', X')$ and the adjoint operator $T^* \in L(Y, X)$, we consider the conjugate-linear isometries

$$S_X : X \to X', \quad S_Y : Y \to Y'$$

which are defined through the theorem of F. Riesz. We recall that S_X is onto X' and S_Y is onto Y'. The defining formulas of these isometries are

$$S_X(z)(x) = \langle x, z\rangle \quad S_Y(w)(y) = \langle y, w\rangle$$

for every $x, z \in X$ and every $y, w \in Y$.

Proposition 6.17. *Let X, Y be Hilbert spaces. Then $T^* = S_X^{-1} T' S_Y$.*

Proof. For every $x \in X$, $y \in Y$ we have

$$S_X(T^*(y))(x) = \langle x, T^*(y)\rangle = \langle T(x), y\rangle = S_Y(y)(T(x)) = T'(S_Y(y))(x).$$

Thus, $S_X(T^*(y)) = T'(S_Y(y))$ for every $y \in Y$ and so $S_X T^* = T' S_Y$. □

Definition. *Let X be a Hilbert space. We say that $T \in L(X)$ is* **self-adjoint**, *if $T^* = T$.*

In the case of a self-adjoint operator, (6.1) states

$$\langle x_1, T(x_2)\rangle = \langle T(x_1), x_2\rangle$$

for every $x_1, x_2 \in X$.

Example 6.4.1. If Y is a closed subspace of a Hilbert space X, then we have the orthogonal projection $P_Y \in L(X)$.
Now, proposition 2.15 implies that P_Y is self-adjoint. In fact, proposition 2.17 says that orthogonal projections are exactly those operators $P \in L(X)$ which are self-adjoint and satisfy $P^2 = P$ (where P^2 means $PP = P \circ P$).

In an inner product space X we have two notions of "orthogonal" set $A^\perp$ of a set $A \subseteq X$. One of the $A^\perp$ is a subspace of X', exactly as in the case of a general normed space X. The other $A^\perp$ is a subspace of X itself:

$$A^\perp = \{x \in X \,|\, x \perp A\}.$$

The relation between these two notions of $A^\perp$ was determined just before proposition 3.7.

Now we shall see the analogue of proposition 6.13 for the adjoint operator.

Proposition 6.18. *Let X, Y be Hilbert spaces, and $T \in L(X, Y)$. Then:*
(i) $\mathrm{N}(T^*) = \mathrm{R}(T)^\perp$.
(ii) $\mathrm{N}(T) = \mathrm{R}(T^*)^\perp$.
(iii) $\mathrm{cl}(\mathrm{R}(T^*)) = \mathrm{N}(T)^\perp$.
(iv) $\mathrm{cl}(\mathrm{R}(T)) = \mathrm{N}(T^*)^\perp$.
Therefore:
(v) T^* is one-to-one if and only if $\mathrm{R}(T)$ is dense in Y.
(vi) T is one-to-one if and only if $\mathrm{R}(T^*)$ is dense in X.

Proof. (i) $y \in \mathrm{N}(T^*)$ if and only if $T^*(y) = 0$ if and only if $\langle x, T^*(y)\rangle = 0$ for every $x \in X$ if and only if $\langle T(x), y\rangle = 0$ for every $x \in X$ if and only if $y \in \mathrm{R}(T)^\perp$.
(ii) $x \in \mathrm{N}(T)$ if and only if $T(x) = 0$ if and only if $\langle T(x), y\rangle = 0$ for every $y \in Y$ if and only if $\langle x, T^*(y)\rangle = 0$ for every $y \in Y$ if and only if $x \in \mathrm{R}(T^*)^\perp$.
(iii) Direct implication of (ii) and of (ii) of proposition 2.12.
(iv) Direct implication of (i) and of (i) of proposition 2.12.
(v) Use (i) and (iv).
(vi) Use (ii) and (iii). □

6.5 Normed algebras. The normed algebra L(X).

Definition. *The linear space X over F is called* **algebra** *over F, if, besides the (internal) operation of addition of elements of X and the (external) operation of multiplication of numbers in F with elements of X, there is also an (internal) operation of multiplication of elements of X, which to every $(x, y) \in X \times X$ assigns the* **product** *$xy \in X$, so that*
(i) $(xy)z = x(yz)$ for every $x, y, z \in X$,
(ii) $x(y + z) = xy + xz$ and $(x + y)z = xz + yz$ for every $x, y, z \in X$,
(iii) $(\lambda x)y = x(\lambda y) = \lambda(xy)$ for every $\lambda \in F$ and every $x, y \in X$.
If there is some $e \in X \setminus \{0\}$ so that
(iv) $ex = xe = x$ for every $x \in X$,
then e is called **unit** *of the algebra X, and X is called* **algebra with unit**.
Also, if
(v) $xy = yx$ for every $x, y \in X$,
then X is called **commutative algebra**.
If the algebra X has a unit, and if for any $x \in X$, $x \neq 0$, there is some $x^{-1} \in X$ so that $xx^{-1} = x^{-1}x = e$, then x is called **invertible**, *and x^{-1} is called* **inverse** *of x.*

It is very easy to prove that $x0 = 0x = 0$ for every $x \in X$, where 0 is the zero element of X. Also, if the algebra X has a unit, then this is unique. Moreover, if some element of X has an inverse, then this is unique. Finally, if $x, y \in X$ are invertible, then xy is also invertible, and $(xy)^{-1} = y^{-1}x^{-1}$.

Definition. *Let X be an algebra. If $\|\cdot\|$ is a norm on the linear space X such that $\|xy\| \leq \|x\|\|y\|$ for every $x, y \in X$, then X is called* **normed algebra**. *If, moreover, X is complete, then X is called* **Banach algebra**.
If the normed algebra X has a unit e, and $\|e\| = 1$, then X is called **normed algebra with unit**.

Example 6.5.1. In the space l^∞ we consider the operation of multiplication defined for every $x = (\lambda_k), y = (\kappa_k) \in l^\infty$ by

$$xy = (\lambda_k \kappa_k).$$

It is easy to see that l^∞ with this multiplication is a commutative algebra with unit. The unit e is the constant sequence (1). Moreover, l^∞ is a normed algebra, since

$$\|xy\|_\infty = \sup_{k \in \mathbb{N}} |\lambda_k \kappa_k| \leq \sup_{k \in \mathbb{N}} |\lambda_k| \sup_{k \in \mathbb{N}} |\kappa_k| = \|x\|_\infty \|y\|_\infty$$

and $\|e\|_\infty = \sup_{k \in \mathbb{N}} |1| = 1$.

Example 6.5.2. In $L^\infty = L^\infty(\Omega) = L^\infty(\Omega, \Sigma, \mu)$ we consider the standard operation of multiplication of functions: the product of essentially bounded functions is an essentially bounded function. Then L^∞ is a commutative algebra with unit. Its unit is the constant function 1. Moreover, L^∞ is a normed algebra, since

$$\|fg\|_\infty = \text{ess-sup}\, |fg| \leq \text{ess-sup}\, |f|\, \text{ess-sup}\, |g| = \|f\|_\infty \|g\|_\infty$$

and $\|1\|_\infty = \text{ess-sup}\, |1| = 1$.

Example 6.5.3. The spaces $B(A)$ and $BC(A)$ are normed algebras with unit. As in the previous example, multiplication in these spaces is the standard multiplication of functions, and we have

$$\|fg\|_u = \sup_{x \in A} |f(x)g(x)| \leq \sup_{x \in A} |f(x)| \sup_{x \in A} |g(x)| = \|f\|_u \|g\|_u$$

and $\|1\|_u = \sup_{x \in A} |1| = 1$.

Example 6.5.4. A more interesting example is the normed space $l^1(\mathbb{Z})$. This is a variant of the usual space $l^1 = l^1(\mathbb{N})$, and it is the set of all double-sided sequences $x = (\lambda_k) = (\lambda_k)_{k\in\mathbb{Z}}$ with the 1-norm, which is defined by

$$\|x\|_1 = \sum_{k\in\mathbb{Z}} |\lambda_k| = \sum_{-\infty}^{+\infty} |\lambda_k|.$$

Addition and multiplication by numbers are defined in $l^1(\mathbb{Z})$ exactly as in $l^1 = l^1(\mathbb{N})$. With its 1-norm, $l^1(\mathbb{Z})$ is a Banach space.
Now, we define an operation in $l^1(\mathbb{Z})$ as follows. For any $x = (\lambda_k), y = (\kappa_k) \in l^1(\mathbb{Z})$ and any $k \in \mathbb{Z}$ we define

$$\mu_k = \sum_{m\in\mathbb{Z}} \lambda_{k-m}\kappa_m. \tag{6.2}$$

Then

$$\begin{aligned}\sum_{k\in\mathbb{Z}} |\mu_k| &\le \sum_{k\in\mathbb{Z}} \Big(\sum_{m\in\mathbb{Z}} |\lambda_{k-m}||\kappa_m|\Big) = \sum_{m\in\mathbb{Z}} \Big(\sum_{k\in\mathbb{Z}} |\lambda_{k-m}|\Big)|\kappa_m| = \sum_{m\in\mathbb{Z}} \Big(\sum_{k\in\mathbb{Z}} |\lambda_k|\Big)|\kappa_m| \\ &= \sum_{k\in\mathbb{Z}} |\lambda_k| \sum_{m\in\mathbb{Z}} |\kappa_m| = \|x\|_1\|y\|_1 < +\infty.\end{aligned} \tag{6.3}$$

Hence the sequence (μ_k) is in $l^1(\mathbb{Z})$.

Definition. *We denote the sequence (μ_k), defined by (6.2), by the symbol $x * y$ and we call it* **convolution** *of x, y:*

$$x * y = (\mu_k).$$

So we have defined the operation of convolution in $l^1(\mathbb{Z})$, and it is easy to show the properties:

$$(x*y)*z = x*(y*z), \quad x*(y+z) = x*y + x*z, \quad (x+y)*z = x*z + y*z,$$
$$(\lambda x)*y = x*(\lambda y) = \lambda(x*y), \quad x*y = y*x$$

for every $\lambda \in F$ and every $x, y, z \in l^1(\mathbb{Z})$. This means that $l^1(\mathbb{Z})$ is a commutative algebra, with convolution as the operation of multilication. This algebra has a unit: the sequence $e = (\delta_k)$, which is defined by $\delta_k = 1$, if $k = 0$, and $\delta_k = 0$, if $k \neq 0$, satisfies

$$e * x = x * e = x$$

for every $x \in l^1(\mathbb{Z})$.
Now, (6.3) says that the norm of $l^1(\mathbb{Z})$ satisfies

$$\|x * y\|_1 \le \|x\|_1\|y\|_1$$

for every $x, y \in l^1(\mathbb{Z})$. Also

$$\|e\|_1 = \sum_{k\in\mathbb{Z}} |\delta_k| = 1.$$

Therefore, $l^1(\mathbb{Z})$ is a commutative Banach algebra with unit.

Example 6.5.5. Another interesting example is the normed space $L^1 = L^1(\Omega) = L^1(\Omega, \Sigma, \mu)$. We define an operation in L^1 as follows. For any $f, g \in L^1$ one can show that the function $f(x-y)g(y)$ is measurable with respect to the product σ-algebra $\Sigma \times \Sigma$ in $\Omega \times \Omega$. Tonelli's theorem implies

$$\begin{aligned}\iint_{\Omega\times\Omega} |f(x-y)g(y)|\, d(\mu\times\mu)(x,y) &= \int_\Omega \Big(\int_\Omega |f(x-y)||g(y)|\, d\mu(x)\Big)\, d\mu(y) \\ &= \int_\Omega \Big(\int_\Omega |f(x-y)|\, d\mu(x)\Big)|g(y)|\, d\mu(y) = \int_\Omega \Big(\int_\Omega |f(x)|\, d\mu(x)\Big)|g(y)|\, d\mu(y) \\ &= \int_\Omega |f(x)|\, d\mu(x) \int_\Omega |g(y)|\, d\mu(y) = \|f\|_1\|g\|_1 < +\infty.\end{aligned} \tag{6.4}$$

Now, Fubini's theorem implies that the function $f(x-y)g(y)$ is integrable with respect to the product σ-algebra $\Sigma \times \Sigma$ in $\Omega \times \Omega$, that for μ-a.e. $x \in \Omega$ the function $f(x-y)g(y)$ (as a function of y) is in $L^1(\Omega)$, that the function

$$\int_\Omega f(x-y)g(y)\, d\mu(y),$$

as a function of x, is in $L^1(\Omega)$, and that

$$\begin{aligned}\int_\Omega \Big| \int_\Omega f(x-y)g(y)\, d\mu(y) \Big|\, d\mu(x) &\leq \int_\Omega \Big(\int_\Omega |f(x-y)||g(y)|\, d\mu(y) \Big)\, d\mu(x) \\ &= \int_\Omega \Big(\int_\Omega |f(x-y)||g(y)|\, d\mu(x) \Big)\, d\mu(y) = \|f\|_1 \|g\|_1,\end{aligned} \tag{6.5}$$

where the last equality comes from (6.4).

Definition. *For every $f, g \in L^1$ we define the function*

$$(f * g)(x) = \int_\Omega f(x-y)g(y)\, d\mu(y) \quad \text{for } \mu\text{-a.e. } x \in \Omega.$$

The function $f * g$ is called **convolution** *of f, g.*

We saw that $f * g \in L^1$ and (6.5) says that

$$\|f * g\|_1 \leq \|f\|_1 \|g\|_1.$$

It is relatively easy to show the properties:

$$(f * g) * h = f * (g * h), \quad f * (g + h) = f * g + f * h, \quad (f + g) * h = f * h + g * h,$$

$$(\lambda f) * g = f * (\lambda g) = \lambda (f * g), \quad f * g = g * f$$

for every $\lambda \in F$ and every $f, g, h \in L^1$. This means that L^1 is a commutative algebra, with convolution as the operation of multilication. It can be proved that, in general, the algebra L^1 *does not* have a unit.
We conclude that L^1 is a commutative Banach algebra.

The last example is the most important for us in this course.

Example 6.5.6. Let X be a normed space. We know that $L(X) = L(X, X)$ is a normed space, and that, if X is a Banach space, then $L(X)$ is also a Banach space. We have seen in proposition 6.6 that, if we denote ST the composition $S \circ T$ of $S, T \in L(X)$, then $ST \in L(X)$ and

$$\|ST\| \leq \|S\| \|T\|.$$

One can easily prove the properties

$$(TS)R = T(SR), \quad T(S + R) = TS + TR, \quad (T + S)R = TR + SR,$$

$$(\lambda T)S = T(\lambda S) = \lambda (TS).$$

Therefore, $L(X)$ is a normed algebra with composition as the operation of multiplication. The unit of multiplication in $L(X)$ is the identity operator $I : X \to X$, and this satisfies $\|I\| = 1$. The algebra $L(X)$, in general, is not commutative.
Regarding the notion of invertibility, we must be careful. By definition, $T \in L(X)$ is invertible, if there is $T^{-1} \in L(X)$ so that $TT^{-1} = T^{-1}T = I$. The equality $TT^{-1} = T^{-1}T = I$, by itself, is equivalent to the function T being one-to-one in X and onto X, and then T^{-1} is the mapping

which is inverse to T. We also know from Linear Algebra that the linearity of T automatically implies the linearity of T^{-1}. But when we write $T \in L(X)$ and $T^{-1} \in L(X)$ we also mean that T and T^{-1} are *bounded*. Now, the boundedness of T *does not* imply the boundedness of T^{-1}. In other words, for $T \in L(X)$, *the invertibility of T as a mapping is not equivalent to its invertibility as an element of $L(X)$*. In the context of Functional Analysis, *when we say that $T \in L(X)$ is invertible we mean that T is invertible as an element of $L(X)$, i.e. that T is one-to-one in X and onto X, and the inverse linear operator T^{-1} is bounded*.

A little later, in the open mapping theorem, we shall prove that, if X is a Banach space, then for every $T \in L(X)$ which is invertible as a mapping, i.e. which is one-to-one in X and onto Y, T^{-1} is *automatically* bounded, and hence T is an invertible element of $L(X)$.

In the algebra $L(X)$ we use the notations

$$T^0 = I, \qquad T^k = \underbrace{T \circ \cdots \circ T}_{k} \ \text{ when } k \in \mathbb{N}.$$

Also, if T is invertible and $T^{-1} \in L(X)$, we write

$$T^{-k} = (T^{-1})^k \ \text{ when } k \in \mathbb{N}.$$

6.6 The uniform boundedness principle.

Theorem 6.1. *Let X be a Banach space, Y be a normed space, and let $\mathcal{F} \subseteq L(X,Y)$. If* $\sup_{T\in\mathcal{F}} \|T(x)\| < +\infty$ *for every $x \in X$, then* $\sup_{T\in\mathcal{F}} \|T\| < +\infty$.

Proof. According to the uniform boundedness principle, there is a non-empty open $U \subseteq X$ and a $M \geq 0$ so that $\|T(x)\| \leq M$ for every $T \in \mathcal{F}$ and every $x \in U$. We take any $x_0 \in U$ and then there is $R > 0$ so that $B(x_0; R) \subseteq U$. So we have that $\|T(x)\| \leq M$ for every $T \in \mathcal{F}$ and every $x \in B(x_0; R)$.
Now we take any $T \in \mathcal{F}$, any $x \neq 0$ and any $t > 1$. Then $x_0 \in B(x_0; R)$ and $x_0 + \frac{R}{t\|x\|} x \in B(x_0; R)$. Hence

$$\|T(x)\| = \frac{t\|x\|}{R} \Big\| T\Big(\frac{R}{t\|x\|}\, x\Big) \Big\| = \frac{t\|x\|}{R} \Big\| T\Big(x_0 + \frac{R}{t\|x\|}\, x\Big) - T(x_0) \Big\| \leq \frac{t\|x\|}{R}\, 2M.$$

Since $t > 1$ is arbitrary, we get

$$\|T(x)\| \leq \frac{2M}{R}\, \|x\|.$$

This is true also for $x = 0$, and hence $\|T\| \leq \frac{2M}{R}$ for every $T \in \mathcal{F}$. □

Theorem 6.2. *Let X be a Banach space, Y be a normed space, and let $\mathcal{F} \subseteq L(X,Y)$. If* $\sup_{T\in\mathcal{F}} |y'(T(x))| < +\infty$ *for every $x \in X$ and every $y' \in Y'$, then* $\sup_{T\in\mathcal{F}} \|T\| < +\infty$.

Proof. Theorem 3.16 implies that $\sup_{T\in\mathcal{F}} \|Tx\| < +\infty$ for every $x \in X$, and then theorem 6.1 finishes the proof. □

6.7 The open mapping theorem.

Lemma 6.1. *Let X be a Banach space, Y be a normed space, and let $T \in L(X,Y)$ and $K > 0$. If $\{y \in Y \,|\, \|y\| < 1\} \subseteq \mathrm{cl}(\{T(x) \,|\, \|x\| < K\})$, then $\{y \in Y \,|\, \|y\| < 1\} \subseteq \{T(x) \,|\, \|x\| < 2K\}$, and T is onto Y.*

Proof. Using $\{y \in Y \mid \|y\| < 1\} \subseteq \text{cl}(\{T(x) \mid \|x\| < K\})$, we easily get

$$\{y \in Y \mid \|y\| < r\} \subseteq \text{cl}(\{T(x) \mid \|x\| < rK\}) \tag{6.6}$$

for every $r > 0$. Indeed, take any $y \in Y$ with $\|y\| < r$. Then $\|\frac{1}{r}\, y\| < 1$ and so there is a sequence (x_n) in X so that $\|x_n\| < K$ for every n and $T(x_n) \to \frac{1}{r}\, y$. Then the sequence (rx_n) satisfies $\|rx_n\| < rK$ for every n and $T(rx_n) = rT(x_n) \to y$.
Now, take any $y \in Y$ with $\|y\| < 1$. Then (6.6) with $r = 1$ implies that there is $x_1 \in X$ so that

$$\|x_1\| < K, \quad \|y - T(x_1)\| < \frac{1}{2}.$$

Then (6.6) with $r = \frac{1}{2}$ implies that there is $x_2 \in X$ so that

$$\|x_2\| < \frac{K}{2}, \quad \|y - T(x_1) - T(x_2)\| < \frac{1}{2^2}.$$

Then, similarly, there is $x_3 \in X$ so that

$$\|x_3\| < \frac{K}{2^2}, \quad \|y - T(x_1) - T(x_2) - T(x_3)\| < \frac{1}{2^3}.$$

Continuing inductively, we see that for every k there is $x_k \in X$ so that

$$\|x_k\| < \frac{K}{2^{k-1}}, \quad \|y - T(x_1) - \cdots - T(x_k)\| < \frac{1}{2^k}.$$

Since $\sum_{k=1}^{+\infty} \|x_k\| < +\infty$, the series $\sum_{k=1}^{+\infty} x_k$ converges in X, and we consider

$$x = \sum_{k=1}^{+\infty} x_k.$$

Then

$$\|x\| \leq \sum_{k=1}^{+\infty} \|x_k\| < \sum_{k=1}^{+\infty} \frac{K}{2^{k-1}} = 2K.$$

Moreover, by the continuity of T, we have

$$y = \lim_{k\to+\infty} \sum_{j=1}^{k} T(x_j) = \lim_{k\to+\infty} T\Big(\sum_{j=1}^{k} x_j\Big) = T\Big(\lim_{k\to+\infty} \sum_{j=1}^{k} x_k\Big) = T(x).$$

□

The open mapping theorem. *Let X, Y be Banach spaces, and let $T \in L(X, Y)$ be onto Y. Then*
(i) there is $M > 0$ so that $\{y \in Y \mid \|y\| < 1\} \subseteq \{T(x) \mid \|x\| < M\}$.
(ii) $T(U)$ is open in Y for every U open in X.
(ii) if T is one-to-one in X, then $T^{-1} \in L(X, Y)$.

Proof. (i) Since T is onto Y, we have $Y = \bigcup_{m=1}^{+\infty} \{T(x) \mid \|x\| < m\}$, and hence

$$Y = \bigcup_{m=1}^{+\infty} \text{cl}(\{T(x) \mid \|x\| < m\}).$$

The theorem of Baire implies that there is m_0 so that $\text{cl}(\{T(x) \mid \|x\| < m_0\})$ has non-empty interior in Y. So there are $y_0 \in Y$ and $R > 0$ such that

$$\{y \in Y \mid \|y - y_0\| < R\} \subseteq \text{cl}(\{T(x) \mid \|x\| < m_0\}).$$

Now take any $y \in Y$ with $\|y\| < 1$. Then $Ry + y_0$ is in $\{y \in Y \,|\, \|y - y_0\| < R\}$ and so there is a sequence (x_n) so that $\|x_n\| < m_0$ for every n and $T(x_n) \to Ry + y_0$. Also, y_0 is in $\{y \in Y \,|\, \|y - y_0\| < R\}$, and so there is a sequence (x_{0n}) so that $\|x_{0n}\| < m_0$ for every n and $T(x_{0n}) \to y_0$. Then

$$T\Big(\frac{1}{R}\Big(x_n - x_{0n}\Big)\Big) = \frac{1}{R}T(x_n - x_{0n}) = \frac{1}{R}\big(T(x_n) - T(x_{0n})\big) \to y$$

and $\|\frac{1}{R}\,(x_n - x_{0n})\| < \frac{2m_0}{R}$ for every n. Therefore

$$\{y \in Y \,|\, \|y\| < 1\} \subseteq \text{cl}(\{T(x) \,|\, \|x\| < K\}),$$

where $K = \frac{2m_0}{R}$.
Now, lemma 6.1 implies that

$$\{y \in Y \,|\, \|y\| < 1\} \subseteq \{T(x) \,|\, \|x\| < 2K\},$$

and this shows (i) with $M = 2K$.
(ii) Take any open $U \subseteq X$, and any $y_0 = T(x_0) \in T(U)$ with $x_0 \in U$. Then there is $r > 0$ so that

$$\{x \in X \,|\, \|x - x_0\| < r\} \subseteq U.$$

Let $\|y - y_0\| < \frac{r}{M}$. Then $\frac{M}{r}\,(y - y_0)$ is in $\{y \in Y \,|\, \|y\| < 1\}$ and so (i) implies that there is $x \in X$ so that $\|x\| < M$ and $T(x) = \frac{M}{r}\,(y - y_0)$. Then

$$y = T\Big(\frac{r}{M}\,x\Big) + y_0 = T\Big(\frac{r}{M}\,x + x_0\Big)$$

and $\frac{r}{M}\,x + x_0$ is in $\{x \,|\, \|x - x_0\| < r\}$. Therefore,

$$\Big\{y \in Y \,\Big|\, \|y - y_0\| < \frac{r}{M}\Big\} \subseteq \{T(x) \,|\, \|x - x_0\| < r\} \subseteq T(U)$$

and so $T(U)$ is open.
(iii) Let T be one-to-one in X. Then $T^{-1} : Y \to X$ is defined and it is a linear operator.
Now, for any $y \in Y$ with $y \neq 0$ and any $t > 1$ we have that $\frac{1}{t\|y\|}\,y$ is in $\{y \in Y \,|\, \|y\| < 1\}$, and (i) implies that there is $x \in X$ so that $\|x\| < M$ and $\frac{1}{t\|y\|}\,y = T(x)$. Thus

$$\|T^{-1}(y)\| = t\|y\|\|x\| < Mt\|y\|.$$

Since $t > 1$ is arbitrary, we get $\|T^{-1}(y)\| \leq M\|y\|$. This is also true for $y = 0$, and we conclude that $\|T^{-1}(y)\| \leq M\|y\|$ for every $y \in Y$. $\square$

6.8 The closed graph theorem.

Definition. *Let X, Y be normed spaces, and let $T : X \to Y$ be a linear operator. We say that T is* **closed** *if for every sequence (x_n) in X such that $x_n \to x$ in X and $T(x_n) \to y$ in Y it follows that $T(x) = y$.*

If X, Y are linear spaces, then we know from Linear Algebra that their direct sum

$$X \oplus Y = \{(x, y) \,|\, x \in X, y \in Y\}$$

is their cartesian product, equipped with the linear space operations

$$(x_1, y_1) + (x_2, y_2) = (x_1 + x_2, y_1 + y_2), \quad \lambda\,(x, y) = (\lambda x, \lambda y).$$

Definition. *Let X, Y be normed spaces. We consider $\|\cdot\| : X \oplus Y \to \mathbb{R}$ defined for every $(x, y) \in X \oplus Y$ by*

$$\|(x, y)\| = \|x\| + \|y\|.$$

It would be more precise to write $\|(x, y)\| = \|x\|_X + \|y\|_Y$, or something similar, since the spaces X, Y may not have the same norm. But we ignore this, keeping the simpler notation.

Proposition 6.19. *Let X, Y be normed spaces. Then the function $\|\cdot\|$ defined on $X \oplus Y$ is a norm. Moreover, if X, Y are Banach spaces, then $X \oplus Y$ is a Banach space.*

Proof. Exercise. $\square$

Definition. *If $f : A \to B$, then the set $G(f) = \{(a, f(a)) \,|\, a \in A\} \subseteq A \times B$ is called* **graph** *of f.*

It is trivial to show that, if X, Y are linear spaces and $T : X \to Y$ is a linear operator, then $G(T)$ is a linear subspace of $X \oplus Y$.

Lemma 6.2. *Let X, Y be normed spaces, and $T : X \to Y$ be a linear operator. Then T is closed if and only if $G(T)$ is a closed subspace of $X \oplus Y$.*

Proof. Exercise. $\square$

Proposition 6.20. *Let X, Y be normed spaces, and $T : X \to Y$ be a linear operator. If T is bounded, then T is closed.*

Proof. Exercise. $\square$

The closed graph theorem. *Let X, Y be Banach spaces, and $T : X \to Y$ be a linear operator. If T is closed, then T is bounded.*

Proof. Let $T : X \to Y$ be closed. Proposition 6.19 and lemma 6.2 imply that $G(T)$ is a closed subspace of the Banach space $X \oplus Y$ and hence it is a Banach space.
We consider $S : G(T) \to X$ defined for every $x \in X$ by

$$S(x, T(x)) = x.$$

It is clear that S is a linear operator which is one-to-one in $G(T)$ and onto X. Moreover, S is bounded since

$$\|S(x, T(x))\| = \|x\| \leq \|x\| + \|T(x)\| = \|(x, T(x))\| \quad \text{for every } x \in X.$$

The open mapping theorem implies that $S^{-1} : X \to G(T)$ is bounded and so there is $C \geq 0$ so that

$$\|x\| + \|T(x)\| = \|(x, T(x))\| = \|S^{-1}(x)\| \leq C\|x\| \quad \text{για κάθε } x \in X.$$

Therefore, $C \geq 1$, and also $\|T(x)\| \leq (C - 1)\|x\|$ for every $x \in X$. $\square$