Functional Analysis 1
A.A 2011–12

Giuseppe De Marco

0.1. Unordered sums. If Λ is a finite set, $(E,+)$ is a commutative semigroup with neutral element written additively, and $x\mapsto x(\lambda)=x_\lambda$ is a family of elements of E indexed by Λ, it is clear how to define the sum

$$\sum_{\lambda\in\Lambda} x_\lambda=\sum_\Lambda x$$

of all elements of the family: e.g. one can define $\sum_\emptyset x=0$ (the zero of the semigroup E), and then proceed by induction on the cardinality of Λ; the definition is also such that if $\sigma: M\to\Lambda$ is a bijection of M onto Λ then $\sum_{\lambda\in\Lambda}x_\lambda=\sum_{\mu\in M}x_{\sigma(\mu)}$, and uses associativity and commutativity of the addition of E: we regard all that as obvious. If Λ is infinite, there is no way to define an infinite sum $\sum_{\lambda\in\Lambda}x_\lambda$ in the absence of a topology on E. But if such a topology exists, and is compatible with the semigroup structure, then there is a very natural way of defining this infinite sum (which however will not always exist), as the limit of the finite sums, " as the finite sets get larger and larger". We restrict our attention to (the additive group of) normed spaces, particularly Banach spaces. It should however be clear that many results are valid in more general contexts.

0.1.1.

Definition. Let X be a normed space, Λ an infinite set, $x\mapsto x(\lambda)=x_\lambda$ a family of elements of X indexed by Λ. We say that the family x is summable, with $s\in X$ as sum, and we write

$$\sum_{\lambda\in\Lambda}x_\lambda\left(=\sum_\Lambda x\right)=s,$$

if for every neighborhood U of s there exist a finite subset $F(U)\subseteq\Lambda$ such that

$$\sum_{\lambda\in F}x_\lambda\in U,$$

for every finite subset F of Λ containing $F(U)$, $F(U)\subseteq F\subseteq\Lambda$.

Remark. For readers acquainted with the theory of generalized sequences, also called nets: we consider the set $\mathcal{F}(\Lambda)$ of finite subsets of Λ, partially ordered by inclusion; since the union of two finite sets is also finite this is a directed set, so that $S:\mathcal{F}(\Lambda)\to X$ defined by $S(F)=\sum_{\lambda\in F}x_\lambda$ is a net with values in X: if this net has a limit $s\in X$, then this limit is called *sum* of x on Λ.

A simple imitation of the argument given for series with positive terms yields:

Exercise 0.1.1.1. A family $(x_\lambda)_{\lambda\in\Lambda}$ of positive real numbers, $x_\lambda\ge 0$ for every $\lambda\in\Lambda$, is summable in $\mathbb{R}$ if and only if $s=\sup\{\sum_F x:\ F\subseteq\Lambda,\ F\text{ finite}\}<+\infty$, and in this case $\sum_\Lambda x=s$.

Notice that there is commutative convergence, as expected from the definition.

Exercise 0.1.1.2. Let X be a normed space, and let $x=(x_\lambda)_{\lambda\in\Lambda}$ be a family of elements of X. If x is summable, and $\sigma:\Lambda\to\Lambda$ is a self–bijection of Λ, then $x\circ\sigma$ is also summable, and with the same sum as x.

In other words: summable families are commutatively summable.

Solution. Let $s=\sum_\Lambda x$; given a nbhd U of s let $F(U)$ be as in the definition of summability, and let $G(U)=\sigma^{-1}(F(U))$; if $F\supseteq G(U)$ then $\sigma(F)\supseteq\sigma(G(U))=F(U)$ so that $\sum_{\sigma(F)}x\in U$; but $\sum_{\sigma(F)}x=\sum_F x\circ\sigma$, and the proof is obtained. $\square$

0.1.2. *The Cauchy criterion.* If $\sum_\Lambda x$ exists in the normed space X, then the finite sums $\sum_F x$ must "vanish at infinity": given $\varepsilon > 0$ there is a finite subset $F(\varepsilon) \subseteq \Lambda$ such that for all finite $F \subseteq \Lambda \setminus F(\varepsilon)$ we have

$$\left\| \sum_F x \right\| \le \varepsilon.$$

In fact, if the sum is s, given $\varepsilon > 0$ there exists a finite subset $F(\varepsilon)$ of Λ such that $\|\sum_G x - s\| \le \varepsilon/2$ for every finite subset G of Λ containing $F(\varepsilon)$; if $F \subseteq \Lambda \setminus F(\varepsilon)$ we have

$$\left\| \sum_{F(\varepsilon)\cup F} x - s \right\| \le \varepsilon/2, \quad \text{and} \quad \left\| \sum_{F(\varepsilon)} x - s \right\| \le \varepsilon/2$$

but

$$\varepsilon/2 \ge \left\| \sum_{F(\varepsilon)\cup F} x - s \right\| = \left\| \sum_{F(\varepsilon)} x + \sum_F x - s \right\| \ge \left\| \sum_F x \right\| - \left\| \sum_{F(\varepsilon)} x - s \right\|,$$

which implies

$$\left\| \sum_F x \right\| \le \varepsilon/2 + \left\| \sum_{F(\varepsilon)} x - s \right\| \le \varepsilon/2 + \varepsilon/2 = \varepsilon.$$

In a complete normed space this condition is also sufficient to ensure summability:

. The Cauchy summability criterion *Let X be a Banach space, and let $x : \Lambda \to X$ be a family of elements of X. The family x is summable if and only if for every $\varepsilon > 0$ there is a finite subset $F(\varepsilon) \subseteq \Lambda$ such that for all finite $F \subseteq \Lambda \setminus F(\varepsilon)$ we have*

$$\left\| \sum_F x \right\| \le \varepsilon.$$

Moreover in this case the support *of the family, i.e. the set $S = \{\lambda \in \Lambda :\ x_\lambda \ne 0\}$ is (finite or) countable; and if $\sigma : \mathbb{N} \to \Lambda$ is a bijection of $\mathbb{N}$ onto a set containing the support of x, then the series $\sum_{n=0}^\infty x_{\sigma(n)}$ is convergent, and its sum coincides with $\sum_\Lambda x$.*

Proof. Necessity was seen above. For the sufficiency let's first prove countability of the support: given any $\varepsilon > 0$ the set $S(\varepsilon) = \{\lambda \in \Lambda :\ \|x_\lambda\| > \varepsilon\}$ is finite (it is contained in $F(\varepsilon)$, if $F(\varepsilon)$ is as in the statement of the Cauchy's condition, since for any singleton $\{\lambda\} \subseteq \Lambda \setminus F(\varepsilon)$ we have $\left\|\sum_{\{\lambda\}} x\right\| = \|x_\lambda\| \le \varepsilon$), hence the support is countable, being a countable union of finite sets ($S = \bigcup_{n=1}^\infty S(1/n)$). Assume now that $\sigma : \mathbb{N} \to \Lambda$ is injective and that $\sigma(\mathbb{N}) \supseteq S$. The sequence $s_m = \sum_{n=0}^m x_{\sigma(n)}$ of partial sums of the series $\sum_{n=0}^\infty x_{\sigma(n)}$ is a Cauchy sequence: in fact, given $\varepsilon > 0$ and $F(\varepsilon)$ as in the Cauchy condition above, if $m_\varepsilon = \max(\sigma^{\leftarrow}(F(\varepsilon)))$ we have that for every $m > m_\varepsilon$ and every integer $p \ge 1$ the set $\sigma(\{m+1, \dots, m+p\})$ is finite and disjoint from $F(\varepsilon)$, so that for these m and p:

$$\|s_{m+p} - s_m\| = \left\| \sum_{k=m+1}^{m+p} x_{\sigma(k)} \right\| \le \varepsilon$$

proving that in fact the sequence $(s_m)_m$ is a Cauchy sequence. Since X is complete this sequence converges to a limit s. We claim that $s = \sum_\Lambda x$. In fact, given $\varepsilon > 0$ take $n_\varepsilon \in \mathbb{N}$ such that $\|s_m - s\| \le \varepsilon$ for every $m \ge n_\varepsilon$ and let F_1 be as in the statement of the criterion, i.e. such that the sums $\sum_G x$ have norm less than ε if G is disjoint from F_1; we may also assume that $F_1 \supseteq \sigma(\{0, \dots, n_\varepsilon\})$; put $m_\varepsilon = \max(\sigma^{\leftarrow}(F_1))$, so that $m_\varepsilon \ge n_\varepsilon$, and finally set $F(\varepsilon) = F_1 \cup \sigma(\{0, \dots, m_\varepsilon\}$. Then for every finite F containing $F(\varepsilon)$ we have:

$$\left\| \sum_F x - s \right\| = \left\| \sum_{F(\varepsilon)} x + \sum_{F\setminus F(\varepsilon)} x - s \right\| = \left\| \sum_{n=0}^{m_\varepsilon} x_{\sigma(n)} - s + \sum_{F\setminus F(\varepsilon)} x \right\| \le \left\| \sum_{n=0}^{m_\varepsilon} x_{\sigma(n)} - s \right\| + \left\| \sum_{F\setminus F(\varepsilon)} x \right\| \le 2\varepsilon,$$

thus concluding the proof. □

0.1.3. *Summability and series.* The notion of converging series is different from that of a summable family: series are ordered sums, the sum of the series is defined using the order of the natural numbers, whereas the sums above defined clearly ignore any ordering which may be present on Λ. There are of course strict relations: as seen above, a summable family is different from zero on a set which is at most countable, and if this support is ordered by the natural numbers, in an arbitrary way, we get a series which converges to the sum of the family. In particular, if a a sequence $n \mapsto x_n$ is summable, then the series $\sum_{n=0}^{\infty} x_n$ is *commutatively convergent*, that is, every rearrangement $\sum_{n=0}^{\infty} x_{\sigma(n)}$, where σ is a self–bijection of $\mathbb{N}$, converges (and to the same sum; see also exercise 0.1.1.2). In fact, commutative convergence is equivalent to summability:

. *If a series of vectors $\sum_{n=0}^{\infty} x_n$ on a Banach space X is commutatively convergent, then it is summable, and all rearrangements converge to $\sum_{\mathbb{N}} x$.*

The proof is not particularly relevant, and may be skipped; for curious readers is reported in 0.1.7.

0.1.4. *Normal and commutative convergence.* It is well–known that in $\mathbb{R}$ or $\mathbb{C}$ commutative convergence of a series $\sum_{n=0}^{\infty} x_n$ is equivalent to *absolute convergence*, i.e. convergence of the series of absolute values $\sum_{n=0}^{\infty} |x_n|$. The most spontaneous extension to normed spaces of the notion of absolute convergence is that of *normal convergence*, in general of normal summability. We say that a family $x : \Lambda \to X$ in a normed space X is *normally summable* if the family $\lambda \mapsto \|x_\lambda\|$ of the norms is summable in $\mathbb{R}$. As we might expect:

. *In a Banach space X, every normally summable family $(x_\lambda)_{\lambda\in\Lambda}$ is summable, and $\|\sum_\Lambda x\| \le \sum_\Lambda \|x_\lambda\|$.*

Proof. The proof is a trivial application of Cauchy's criterion: since $\|x\|$ is summable in $\mathbb{R}$, given $\varepsilon > 0$ there exists finite $F(\varepsilon) \subseteq \Lambda$ such that $\sum_G \|x\| \le \varepsilon$ for every finite $G \subseteq \Lambda \setminus F(\varepsilon)$, by the necessity part of Cauchy's criterion; then, by triangle inequality:

$$\left\|\sum_G x\right\| \le \sum_G \|x\| \le \varepsilon$$

and by the sufficiency part of the same criterion the family is summable in X. □

It is easy to see that a family $(x_\lambda)_{\lambda\in\Lambda}$ of real numbers is summable if and only if it is normally (equivalently, absolutely) summable. We only have to prove that in $\mathbb{R}$ summability implies normal summability; to this end simply observe that for every finite $F \subseteq \Lambda$ there exists $G \subseteq F$ such that $|\sum_G x| \ge (\sum_F |x|)/2$ (either $G = \{\lambda \in F : x_\lambda \ge 0\}$ or $G = \{\lambda \in F : x_\lambda < 0\}$ will do), so that if $F(\varepsilon) \subseteq \Lambda$ is such that $|\sum_F x| \le \varepsilon$ for F finite and disjoint from $F(\varepsilon)$, we have for such an F:

$$\frac{1}{2}\sum_F |x| \le \left|\sum_G x\right| \le \varepsilon,$$

thus concluding the argument. In finite dimensional normed spaces $\mathbb{K}^n$ summability is clearly equivalent to summability of the components, and the same is true for normal summability, so that we conclude that in finite dimensional spaces summability is equivalent to normal summability and also to commutative convergence. This fails in infinite dimensional spaces (see 1.1.3.4).

0.1.5. *Sums on subsets.* If X is a Banach space and $x : \Lambda \to X$ is a summable family, then for every subset A of Λ the restriction $x_{|A}$ is summable on A: this is immediate, by Cauchy's criterion (given ε and $F(\varepsilon)$ as usual, $F(\varepsilon) \cap A$ is as required). For summable families $x : \Lambda \to X$ we thus get a function "sum of x over A" from the power set $\mathcal{P}(\Lambda)$ of Λ to X, $A \mapsto \sum_A x$. It is easy to see that this function is finitely additive: if $\Lambda_1, \dots, \Lambda_m$ is a disjoint family of subsets of Λ then

$$\sum_{\bigcup_{k=1}^m \Lambda_k} x = \sum_{k=1}^m \sum_{\Lambda_k} x.$$

which may be seen as an expression of the associativity of (finitely many) infinite sums. It is clearly enough to prove it for $m = 2$, that is if A, B are disjoint subsets of Λ then $\sum_{A\cup B} x = \sum_A x + \sum_B x$; by disjointness of A and B we get, for every finite subset $F \subseteq A \cup B$:

$$\sum_F x = \sum_{F\cap A} x + \sum_{F\cap B} x,$$

and the finishing argument is clear (take the limit as F gets large and use the fact that the sum of the limits is the limit of the sum). But more is true, as shown in the next paragraph. We first need an observation (B_X is the closed unit ball of X):

. *Let X be a Banach space, $x : \Lambda \to X$ a summable family, with sum s. Given $\varepsilon > 0$, let $F(\varepsilon)$ be such that $\sum_F x \in s + \varepsilon B_X$ for every finite $F \supseteq F(\varepsilon)$. Then $\sum_A x \in s + \varepsilon B_X$ for every subset A of Λ containing $F(\varepsilon)$.*

Proof. The proof depends on the fact that $s + \varepsilon B_X$ is a closed subset of X, and since it contains $\sum_F x$ for every finite F with $F(\varepsilon) \subseteq F \subseteq A$ it also contains the limit of these sums "as F tends to A running among the finite subsets of A containing $F(\varepsilon)$". □

0.1.6. *Unrestricted associativity (summation by blocs).*

. *Let X be a Banach space, let $X : \Lambda \to X$ be a summable family of elements of X, let $(\Lambda(j))_{j\in J}$ be a disjoint family of subsets of Λ, and let $y_j = \sum_{\Lambda(j)} x$, for every $j \in J$. Then the family $y : j \mapsto y_j$ is summable over J, and we have*

$$\sum_J y = \sum_{\bigcup_{j\in J}\Lambda(j)} x, \quad \textit{in other words} \quad \sum_{j\in J}\left(\sum_{\lambda\in\Lambda(j)} x_\lambda\right) = \sum_{\lambda\in\bigcup_{j\in J}\Lambda(j)} x_\lambda.$$

Proof. Let s be the sum of x on $M = \bigcup_{j\in J}\Lambda(j)$. Given $\varepsilon > 0$, let $F(\varepsilon)$ be such that $\sum_F x \in s + \varepsilon B_X$ for every finite $F \supseteq F(\varepsilon)$. Let $J(\varepsilon) = \{j \in J : F \cap \Lambda(j) \neq \emptyset\}$. Then $J(\varepsilon)$ is a finite subset of J, and $F(\varepsilon) \subseteq \bigcup_{j\in J(\varepsilon)}\Lambda(j)$; by 0.1.5 we have, for every finite subset K of J containing $J(\varepsilon)$:

$$\sum_{\bigcup_{j\in K}\Lambda(j)} x \in S + \varepsilon B_X;$$

by the finite case for associativity the sum on the left is $\sum_{j\in K} y_j$, and the proof ends: given $\varepsilon > 0$ we have found a finite $J(\varepsilon) \subseteq J$ such that $\sum_{j\in K} y_j \in s + \varepsilon B_X$ for every finite $K \supseteq J(\varepsilon)$. □

Trivial examples show that summability of the entire family (at least on the union $M = \bigcup_{j\in J}\Lambda(j)$) is essential for the validity of the result: x can be summable with sum y_j on every $\Lambda(j)$, and the family $j \mapsto y_j$ be summable on J without x being summable on M (partition $\Lambda = \mathbb{N}$ in infinitely many pairs $\Lambda(j) = \{2j, 2j+1\}$, and set $x(n) = (-1)^n$; we have $y_j = 0$ for every j, but clearly x is not summable on $\mathbb{N}$).

0.1.7. *Commutative convergence implies summability.* We prove here what announced in 0.1.3: if a sequence $n \mapsto x_n$ in a Banach space is not summable, then some rearrangement does not converge. In fact the Cauchy criterion 0.1.2 does not hold, so that there is $\varepsilon > 0$ with the following property:

for each $n \in \mathbb{N}$ there is a finite subset G_n of $\mathbb{N}$, with $\min G_n > n$, such that

$$\left\|\sum_{k\in G_n} x_k\right\| > \varepsilon.$$

Now we construct the rearrangement inductively: list first all elements of G_0, in increasing order, then all natural numbers up to of $\max G_0$, also in increasing order. Let $m(1) = \max G(0)$; list then all elements of $G_{m(1)}$ in increasing order, then all natural numbers not already listed up to $\max G_{m(1)}$; let $m(2) = \max G_{m(1)}$ etc.. It is clear that the construction gives a rearrangement $\sum_{k=0}^{\infty} x_{\sigma(k)}$ whose sequence of partial sums is not Cauchy: the numbers $m(n)$ get arbitrarily large, and there is a consecutive block of $\mathrm{Card}(G_{m(n)})$ elements with sum exceeding ε in norm.

1. Banach spaces

1.1. Some important Banach spaces. We list here some of the most important Banach spaces treated in this course. There is a very significant omission, that of Sobolev spaces, not treated here.

1.1.1. *Spaces* $L^p(\mu)$. If $(X,\mathcal{M},\mu)$ is a measure space (meaning that $\mathcal{M}$ is a $\sigma-$algebra of subsets of X, and $\mu:\mathcal{M}\to[0,+\infty]$ is a countably additive set–function with $\mu(\emptyset)=0$), we denote by $L_{\mathcal{M}}(X,\mathbb{K})$ or simpy $L(X)$ the space of all $\mathcal{M}-$measurable $\mathbb{K}-$valued functions. For $p>0$ and $f\in L(X)$ we set

$$\|f\|_p=\left(\int_X |f|^p\,d\mu\right)^{1/p};$$

Denote by $L^p(\mu)$ the set of all $f\in L(X)$ for which $\|f\|_p<\infty$; it is well known that $L^p(\mu)$ is a vector subspace of $L(X)$; if $p\ge 1$ then $f\mapsto\|f\|_p$ is a seminorm on $L^p(\mu)$, which becomes a norm on the quotient space modulo the subspace of a.e zero functions, those for which $\|f\|_p=0$. Also, $L^p(\mu)$ is a Banach space for $p\ge 1$ (see e.g. [Folland]). For $p=\infty$ and $f\in L(X)$ we set:

$$\|f\|_\infty=\operatorname{esssup}|f|=\inf\{t>0:\mu(\{|f|>t\})=0\},$$

(essential supremum norm) and we denote by $L^\infty(\mu)$ the set of all $f\in L(X)$ such that $\|f\|_\infty<\infty$; modulo a.e zero functions this is a norm which makes $L^\infty(\mu)$, space of *essentially bounded* functions, into a Banach space.

1.1.2. *Spaces* $\ell^p(\Lambda)$. An interesting particular case of L^p spaces is that of a *discrete measure space*, $(\Lambda,\mathcal{P}(\Lambda),\varkappa)$, where Λ is a set and $\varkappa$ is the counting measure on the power set $\mathcal{P}(\Lambda)$ of Λ. In this case it is easy to see that $x:\Lambda\to\mathbb{K}$ belongs to $L^1(\varkappa)$ if and only if x is summable on Λ, in the sense of the unordered sums treated in the previous section. And essentially bounded functions are exactly the bounded functions since for every nonempty subset $A\subseteq\Lambda$ we have $\varkappa(A)\ge 1$. It is customary to write $\ell^p(\Lambda)$ in place of $L^p(\varkappa)$.

1.1.3. *The space* c_0. In $\ell^\infty(\Lambda)$ there is an important subspace, the space $c_0(\Lambda)$ of functions which are 0 at infinity: the function $x:\Lambda\to\mathbb{K}$ belongs to $c_0(\Lambda)$ if for every $\varepsilon>0$ there exists a finite subset $F(\varepsilon)\subseteq\Lambda$ such that $|x(\lambda)|\le\varepsilon$ for $\lambda\in\Lambda\smallsetminus F(\varepsilon)$, equivalently: for every $\varepsilon>0$ the set $\{|x|>\varepsilon\}$ is finite.

Exercise 1.1.3.1. Prove that $c_0(\Lambda)$ is a closed subspace of $\ell^\infty(\Lambda)$ and that if $x\in c_0(\Lambda)$ then $\|x\|_\infty=\max\{|x(\lambda)|:\lambda\in\Lambda\}$.

Solution. In fact $x\in\ell^\infty(\Lambda)\smallsetminus c_0(\Lambda)$ if and only if x is bounded, and for some $\varepsilon>0$ the set $\{|x|\ge\varepsilon\}$ is infinite. In the $\varepsilon/2$ ball around x there are only functions which are larger than $\varepsilon/2$ on every $\lambda\in\{|x|\ge\varepsilon\}$, so that none of these is in $c_0=c_0(\Lambda)$, thus proving that $\ell^\infty(\Lambda)\smallsetminus c_0(\Lambda)$ is open in $\ell^\infty(\Lambda)$. And if $x\in c_0$ and $x\ne 0$, pick any $\mu\in\Lambda$ such that $x(\mu)\ne 0$; then the set $\{|x|\ge|x(\mu)|\}$ is finite, and the maximum of $|x|$ in this set is clearly the maximum of $|x|$ on all of Λ. □

Exercise 1.1.3.2. Prove that if $1\le p<\infty$ then $\ell^p(\Lambda)\subseteq c_0(\Lambda)$ and $\|x\|_\infty\le\|x\|_p$; and if $x\in\ell^p$ and $q>p$ then $\ell^p\subseteq\ell^q$ and $\|x\|_q\le\|x\|_p$.

Solution. Assume that $x\in\ell^p(\Lambda)$. Given $\varepsilon>0$, by Cauchy's criterion we have that there is a finite subset $F(\varepsilon)$ of Λ such that $\sum_G|x|^p\le\varepsilon^p$ for every G disjoint from $F(\varepsilon)$; in particular for every $\lambda\in\Lambda\smallsetminus F(\varepsilon)$ we have $\sum_{\{\lambda\}}|x|^p=|x(\lambda)|^p\le\varepsilon^p$, so that $|x(\lambda)|>\varepsilon$ can happen only if $\lambda\in F(\varepsilon)$. Thus $x\in c_0(\Lambda)$. Moreover $|x(\mu)|\le\left(\sum_\Lambda|x|^p\right)^{1/p}=\|x\|_p$ for every $\mu\in\Lambda$, so that $\|x\|_\infty\le\|x\|_p$. Assume now $1\le p<q<\infty$ and $x\in\ell^p$. Then $|x(\lambda)|^q=|x(\lambda)|^{q-p}\,|x(\lambda)|^p\le\|x\|_\infty^{q-p}\,|x(\lambda)|^p$; summing over Λ we get

$$\sum_\Lambda|x|^q\le\|x\|_\infty^{q-p}\sum_\Lambda|x|^p\le\|x\|_p^{q-p}\,\|x\|_p^p=\|x\|_p^q\Longrightarrow\|x\|_q\le\|x\|_p.$$

□

For every $\lambda\in\Lambda$ we define $e_\lambda:\Lambda\to\mathbb{K}$ as the characteristic function of the singleton $\{\lambda\}$, that is, $e_\lambda(\mu)=0$ if $\mu\ne\lambda$, and $e_\lambda(\lambda)=1$. Clearly $\{e_\lambda:\lambda\in\Lambda\}$ is a linearly independent subset of all the spaces ℓ^p, and of c_0 as well. The vector space spanned by this set is sometimes written $c_{00}(\Lambda)$; it clearly consists of the functions of finite support.

Exercise 1.1.3.3. Prove that $c_{00}(\Lambda)$ is dense in all spaces $\ell^p(\Lambda)$ for $1\le p<\infty$, and that its closure in $\ell^\infty(\Lambda)$ is $c_0(\Lambda)$.

Solution. For every function $x : \Lambda \to \mathbb{K}$ and every finite subset of Λ let $x_F = \sum_{\lambda\in F} x(\lambda)\, e_\lambda$: then $x_F \in c_{00}(\Lambda)$ (it is the function which coincides with x on F, and is zero outside F, i.e. $x_F = x\,\chi_F$). Given now $\varepsilon > 0$, and $x \in \ell^p$, let $F(\varepsilon)$ be a finite subset of Λ such that $\sum_F |x|^p \le \varepsilon^p$ if $F \cap F(\varepsilon) = \emptyset$. Then $\|x - x_{F(\varepsilon)}\|_p \le \varepsilon$. In fact since $x - x_{F(\varepsilon)}$ is zero on $F(\varepsilon)$ and coincides with x on $\Lambda \setminus F(\varepsilon)$ we have

$$\|x - x_{F(\varepsilon)}\|_p^p = \sup\left\{\sum_F |x|^p : F \subseteq \Lambda \setminus F(\varepsilon),\ F \text{ finite}\right\},$$

and this supremum is not larger than ε^p, since all sums are dominated by ε^p. In the case of c_0 we simply take $F(\varepsilon) = \{\lambda \in \Lambda : |x(\lambda)| > \varepsilon\}$. □

EXERCISE 1.1.3.4. Let $\Lambda = \mathbb{N}$, and put $x_n = e_n/(n+1)$; the norm of x_n is $1/(n+1)$ in every ℓ^p, $p \ge 1$, and also in c_0. If $x : \mathbb{N} \to \mathbb{K}$ is defined as $x(n) = 1/(n+1)$, then $x \in c_0$ and $x \in \ell^p$ if $p > 1$; prove that in all these spaces the family $(x_n)_{n\in\mathbb{N}}$ is summable, and $\sum_{n\in\mathbb{N}} x_n = x$; but the family of the norms is not summable, $\sum_{n\in\mathbb{N}} 1/(n+1) = \infty$ (thus normal summability is stronger than summability in Banach spaces).

1.1.4. *Spaces of bounded continuous functions.* Other Banach spaces of interest are those of bounded continuous functions. If X is a topological space we set $C_b(X) = C_b(X, \mathbb{K}) = C(X, \mathbb{K}) \cap \ell^\infty(X, \mathbb{K})$; with the sup–norm, this is a closed, hence complete, subspace of $\ell^\infty(X)$. Of course if X is compact then $C(X) = C_b(X)$, by Weierstrass theorem.

1.1.5. *Spaces of linear continuous mappings; topological dual of a normed space.* If X and Y are normed spaces, we know that a linear mapping $T : X \to Y$ is continuous iff it is Lipschitz continuous, and that the smallest Lipschitz constant of T is

$$\|T\| = \sup\{\|T(x)\|_Y : \|x\|_X \le 1\} = \sup\{\|T(x)\|_Y : \|x\|_X = 1\}$$

(see e.g. Analisi Due). Thus, if $L_{\mathbb{K}}(X, Y) = L(X, Y)$ denotes the space of continuous linear mappings from X to Y, $\|T\| = \|T\|_{L(X,Y)}$ is a norm on $L(X, Y)$, the *operator norm.*

Let's prove that

. *If Y is complete then $L(X, Y)$ is complete in the operator norm.*

Proof. Assume that $(T_n)_{n\in\mathbb{N}}$ is a Cauchy sequence in $L(X, Y)$. Then, for every $x \in X$, $T_n x = T_n(x)$ is a Cauchy sequence in Y: in fact (we may assume $x \ne 0$):

$$\|T_m x - T_n x\|_Y = \|(T_m - T_n)(x)\|Y \le \|T_m - T_n\|_{L(X,Y)}\, \|x\|_X;$$

given $\varepsilon > 0$ there is $n(\varepsilon) \in \mathbb{N}$ such that $\|T_m - T_n\|_{L(X,Y)} \le \varepsilon$ for $m, n \ge n(\varepsilon)$ so that:

$$(*) \qquad \|T_m x - T_n x\|_Y \le \varepsilon\, \|x\|_X \qquad \text{for } m, n \ge n(\varepsilon) \text{ and every } x \in X,$$

proving that $T_n(x)$ is indeed a Cauchy sequence in Y; since Y is complete, this sequence has a limit in Y; we call Tx this limit. In this way we define $T : X \to Y$, pointwise limit of the sequence of linear maps T_n. That T is linear is obvious: simply pass to the limit in the relations $T_n(x + y) = T_n x + T_n y$ and $T_n(\alpha\, x) = \alpha\, T_n x$. And T is also continuous: simply pass to the limit as m tends to ∞ in (*), keeping n fixed, to obtain

$$\|Tx - T_n x\|_Y \le \varepsilon\, \|x\|_X \qquad \text{for } n \ge n(\varepsilon) \text{ and every } x \in X,$$

thus proving that $\|T - T_n\|_{L(X,Y)} \le \varepsilon$ for $n \ge n(\varepsilon)$. □

A most important application of this result is to the case $Y = \mathbb{K}$; recall that in this case $L_{\mathbb{K}}(X, \mathbb{K})$ is denoted by X^* and is the *normed topological dual* of X, which is then always complete.

1.1.6. *Two natural dual spaces.* Let Λ be an infinite set. We prove that the dual space of $c_0 = c_0(\Lambda)$ is isometrically isomorphic, in a natural way, to $\ell^1 = \ell^1(\Lambda)$. Let $\varphi \in (c_0)^*$ be given. Define $b : \Lambda \to \mathbb{K}$ by $b(\lambda) = \varphi(e_\lambda)$, where $e_\lambda = \chi_{\{\lambda\}}$. We prove that $b \in \ell^1$ and that $\|b\|_1 = \|\varphi\|_{c_0^*}$, the norm of φ in c_0^*. For every finite subset F of Λ consider $a_F = \sum_{\lambda\in F} \overline{\operatorname{sgn} b_\lambda}\, e_\lambda \in c_{00}(\Lambda)$; we clearly have $\|a_F\|_\infty \le 1$ (this norm is either 0 or 1) so that $|\varphi(a_F)| \le \|\varphi\|_{c_0^*} \|a_F\|_\infty \le \|\varphi\|_{c_0^*}$; but $\varphi(a_F) = \sum_{\lambda\in F} \overline{\operatorname{sgn} b_\lambda}\, \varphi(e_\lambda) = \sum_{\lambda\in F} \overline{\operatorname{sgn} b_\lambda}\, b_\lambda = \sum_{\lambda\in F} |b_\lambda|$; we have proved that

$$\sum_{\lambda\in F} |b_\lambda| \le \|\varphi\|_{c_0^*} \quad \text{for every finite subset } F \text{ of } \Lambda, \text{ so that } b \in \ell^1 \text{ and } \|b\|_1 \le \|\varphi\|_{c_0^*}.$$

We now reverse the question in this way: given $b \in \ell^1$, we *define* the functional $\varphi_b : c_0 \to \mathbb{K}$ by the formula

$$\varphi_b(a) = \sum_\Lambda a\, b = \sum_{\lambda\in\Lambda} a(\lambda)\, b(\lambda) \quad \text{for every } a \in c_0.$$

This is well defined, because $a\, b \in \ell^1$ (a is bounded and $b \in \ell^1$). Moreover

$$|\varphi_b(a)| = \left|\sum_{\lambda\in\Lambda} a(\lambda)\, b(\lambda)\right| \le \sum_{\lambda\in\Lambda} |a(\lambda)|\, |b(\lambda)| \le \sum_{\lambda\in\Lambda} \|a\|_\infty\, |b(\lambda)| = \|a\|_\infty\, \|b\|_1,$$

thus proving that $\varphi_b \in c_0^*$ and that $\|\varphi_b\|_{c_0^*} \le \|b\|_1$. If $\varphi \in c_0^*$ is given and $b(\lambda) = \varphi(e_\lambda)$ as above, we have that φ and φ_b coincide on c_{00}, and $\|b\|_1 \le \|\varphi\|_{c_0^*}$. Since c_{00} is dense in c_0, we have $\varphi = \varphi_b$ on all of c_0; moreover, since $\|\varphi_b\|_{c_0^*} \le \|b\|_1$, we actually have $\|\varphi_b\|_{c_0^*} = \|b\|_1$. We can collect the results obtained in the following statement:

. *Let Λ be a set. For every $b \in \ell^1 = \ell^1(\Lambda)$ we define $\varphi_b : c_0 \to \mathbb{K}$ by the formula $\varphi_b(a) = \sum_{\lambda\in\Lambda} a(\lambda)\, b(\lambda)$. Then $\varphi_b \in c_0^*$, and $\|\varphi_b\|_{c_0^*} = \|b\|_1$; moreover $b \mapsto \varphi_b$ is norm–preserving isomorphism of ℓ^1 onto c_0^*.*

There is an analogous representation of $(\ell^1)^*$ by ℓ^∞. Explicitly, for $b \in \ell^\infty$ we define $\varphi_b : \ell^1 \to \mathbb{K}$ by the formula $\varphi_b(a) = \sum_\Lambda a\, b = \sum_{\lambda\in\Lambda} a(\lambda)\, b(\lambda)$, for every $a \in \ell^1$. We have

$$|\varphi_b(a)| = \left|\sum_{\lambda\in\Lambda} a(\lambda)\, b(\lambda)\right| \le \sum_{\lambda\in\Lambda} |\alpha(\lambda)|\, |b(\lambda)| \le \sum_{\lambda\in\Lambda} |\alpha(\lambda)|\, \|b\|_\infty = \|a\|_1\, \|b\|_\infty,$$

which proves that φ_b belongs to $(\ell^1)^*$, and also $\|\varphi_b\|_{(\ell^1)^*} \le \|b\|_\infty$. And every $\varphi \in (\ell^1)^*$ is obtained in this way: putting $b(\lambda) = \varphi(e_\lambda)$ we get a function $b : \Lambda \to \mathbb{K}$; this function is dominated by $\|\varphi\|_{(\ell^1)^*}$ because, for every $\lambda \in \Lambda$:

$$|b(\lambda)| = \varphi(\overline{\operatorname{sgn} b(\lambda)}\, e_\lambda) \le \|\varphi\|_{(\ell^1)^*}\, \|\overline{\operatorname{sgn} b(\lambda)}\, e_\lambda\|_1 \le \|\varphi\|_{(\ell^1)^*}.$$

Then φ and φ_b coincide on the subspace c_{00}, dense in ℓ^1, and by continuity they coincide on all of ℓ^1; moreover $\|\varphi\|_{(\ell^1)^*} = \|\varphi_b\|_{(\ell^1)^*} = \|b\|_\infty$.

1.1.7. *The dual of ℓ^p is ℓ^q.* Let's prove:

. *If $1 < p < \infty$, and $1/p + 1/q = 1$ then the mapping $b \mapsto \varphi_b$ of $\ell^q(\Lambda)$ into $(\ell^p(\Lambda))^*$ where $\varphi_b : \ell^p \to \mathbb{K}$ is defined by $\varphi_b(a) = \sum_\Lambda a\, b = \sum_{\lambda\in\Lambda} a(\lambda)\, b(\lambda)$ is an isometric isomorphism of ℓ^q onto $(\ell^p(\Lambda))^*$.*

Proof. Hölder inequality immediately implies that $\varphi_b \in (\ell^p)^*$ and that $\|\varphi_b\|_{(\ell^p)^*} \le \|b\|_q$:

$$|\varphi_b(a)| = \left|\sum_{\lambda\in\Lambda} a(\lambda)\, b(\lambda)\right| \le \sum_{\lambda\in\Lambda} |a(\lambda)|\, |b(\lambda)| \le \|a\|_p\, \|b\|_q.$$

Assume next that $\varphi \in (\ell^p)^*$; define $b : \Lambda \to \mathbb{K}$ by $b(\lambda) = \varphi(e_\lambda)$. We prove that $b \in \ell^q$ and that $\|b\|_q \le \|\varphi\|_{(\ell^p)^*}$; this concludes the proof, since then the continuous functionals φ and φ_b agree on the dense subspace c_{00}, and moreover $\|b\|_q \le \|\varphi\| = \|\varphi_b\| \le \|b\|_q$, so that $\|b\|_q = \|\varphi\|$. For every finite subset F of Λ put $a_F = \sum_{\lambda\in F} \overline{\operatorname{sgn} b(\lambda)}\, |b(\lambda)|^{q-1}\, e_\lambda$ and $b_F = \sum_{\lambda\in F} b(\lambda)\, e_\lambda = b\, \chi_F$; then

$$\varphi(a_F) = \sum_{\lambda\in F} \overline{\operatorname{sgn} b(\lambda)}\, |b(\lambda)|^{q-1}\, b(\lambda) = \sum_{\lambda\in F} |b(\lambda)|^q = \|b_F\|_q^q.$$

Let's now estimate $\|a_F\|_p$: we have

$$\|a_F\|_p^p = \sum_{\lambda\in F} |b(\lambda)|^{qp-p} = \sum_{\lambda\in F} |b(\lambda)|^q = \|b_F\|_q^q \quad \text{hence} \quad \|a_F\|_p = \|b_F\|_q^{q/p};$$

since φ is by hypothesis continuous we have

$$\varphi(a_F) = |\varphi(a_F)| \le \|\varphi\|\, \|a_F\|_p = \|\varphi\|\, \|b_F\|_q^{q/p};$$

and since $\varphi(a_F) = \|b_F\|_q^q$ we have obtained

$$\|b_F\|_q^q \le \|\varphi\|\, \|b_F\|_q^{q/p} \iff \|b_F\|_q^{q-q/p} \le \|\varphi\|,$$

and since $q - q/p = 1$

$$\|b_F\|_q = \left(\sum_{\lambda\in F} |b(\lambda)|^q\right)^{1/q} \le \|\varphi\|, \quad \text{for every finite subset } F \text{ of } \Lambda;$$

this clearly is equivalent to $\|b\|_q \le \|\varphi\|$, and the proof is concluded. □

1.1.8. *The dual of $c = c(\Lambda)$.* Given a set Λ the space "small c", $c(\Lambda)$ of functions with finite limit at infinity is simply $c(\Lambda) = c_0(\Lambda) + \mathbb{K}$, that is, the constant functions added to functions which are zero at infinity. In other words, $x : \Lambda \to \mathbb{K}$ is in $c = c(\Lambda)$ if it may be written as $x = u + k$, with $u \in c_0$, k a constant function; $c = c_0 \oplus \mathbb{K}$, a direct sum of c_0 and the subspace $\mathbb{K}$ of constant functions. Clearly c is a subspace of ℓ^∞; prove that it is closed in ℓ^∞. Moreover, for every $b \in \ell^1$ the mapping

$$\varphi_b : c \to \mathbb{K} \quad \text{defined by} \quad \varphi_b(a) = \sum_\Lambda a\, b = \sum_{\lambda\in\Lambda} a(\lambda)\, b(\lambda)$$

is an element of c^*, with $\|\varphi_b\|_{c^*} \le \|b\|_1$, and actually equality holds. However this time the map $b \mapsto \varphi_b$ is not surjective from ℓ^1 to c^*: the functional $\delta_\infty(u + k) = k$ is continuous of norm 1, but cannot be represented as φ_b: we should have $\varphi_b(u) = 0$ for every $u \in c_0$, and this implies $b = 0$, as we have seen above. It is not difficult to see that c^* is isomorphic to a direct product $\ell^1 \times \mathbb{K}\,\delta_\infty$, with norm $\|(b, \lambda\,\delta_\infty)\| = \|b\|_1 + |\lambda|$.

1.2. The dual of $L^p(\mu)$. The preceding results for discrete measure spaces can be generalized to arbitrary measure spaces: the dual of $L^p(\mu)$ is $L^q(\mu)$, where $q = \tilde{p} = p/(p-1)$ for an any measure space $(X, \mathcal{M}, \mu)$, if $1 \le p < \infty$, but the proof is considerably more difficult; we prove it only for $\sigma-$finite measure spaces, for the general case the reader can see [Folland].

1.2.1. *Embedding of $L^q(\mu)$ in $(L^p(\mu))^*$.* We begin with the easy part:

. *Let $(X, \mathcal{M}, \mu)$ be a measure space, let $1 \le p < \infty$, and let $q > 1$ be the exponent conjugate to p. For $g \in L^q(\mu)$ let $\varphi_g : L^p(\mu) \to \mathbb{K}$ be the mapping defined by $\varphi_g(f) = \int_X f\, g\, d\mu$. Then $\varphi_g \in (L^p(\mu))^*$, and $\|\varphi_g\|_{(L^p(\mu))^*} = \|g\|_q$ if $p > 1$; this holds also for $p = 1$ and $q = \infty$ if μ is semifinite.*

Proof. Hölder's inequality proves that $\|\varphi_g\|_{(L^p(\mu))^*} \le \|g\|_q$, in particular then we have $\varphi_g \in (L^p(\mu))^*$. To get equality, assume first $p > 1$ and set $f = \operatorname{sgn} \bar{g}\, |g|^{q-1} \in L^p(\mu)$; then

$$\varphi_g(f) = \int_X f\, g = \int_X |g|^q = \|g\|_q^q \quad \text{and} \quad \|f\|_p = \left(\int_X |g|^{pq-p}\right)^{1/p} = \|g\|_q^{q/p},$$

so that

$$\varphi_g(f/\|f\|_p) = \frac{1}{\|f\|_p}\, \varphi_g(f) = \frac{\|g\|_q^q}{\|g\|_q^{q/p}} = \|g\|_q,$$

which proves that $\|g\|_q$ is a value of φ_g on the unit vector $f/\|f\|_p$, and thus $\|\varphi_g\|_{(L^p(\mu))^*} = \|g\|_q$. If $p = 1$ then $\|\varphi_g\|_{(L^1(\mu))^*}$ is no longer (in general) a value assumed by φ_g on the unit sphere of $L^1(\mu)$, but is $\|g\|_\infty$ all the same if μ is semifinite. In fact, if $\alpha < \|g\|_\infty$, then the set $\{|g| > \alpha\}$ has strictly positive $\mu-$measure; by semifiniteness of μ, it contains a subset E of finite nonzero $\mu-$measure; if $f = \operatorname{sgn} \bar{g}\, \chi_E$, we have $f \in L^1(\mu)$ and

$$\varphi_g(f) = \int_X f\, g = \int_E |g| \ge \alpha\, \mu(E) = \alpha\, \|f\|_1 \quad \text{but} \quad \varphi_g(f) \le \|\varphi_g\|_{(L^1(\mu))^*}\, \|f\|_1,$$

so that

$$\alpha \le \|\varphi_g\|_{(L^1(\mu))^*},$$

and since α is an arbitrary number less than $\|g\|_\infty$ we get $\|g\|_\infty \le \|\varphi_g\|_{(L^1(\mu))^*}$, as required. □

The preceding statement says that

. *In the above hypotheses the mapping $g \mapsto \varphi_g$ is an isometric embedding of $L^q(\mu)$ into $(L^p(\mu))^*$.*

1.2.2. *The converse of Hölder inequality.* The most interesting and non–trivial part of the duality between $L^p(\mu)$ and $L^q(\mu)$ is surjectivity of the map $g \mapsto \varphi_g$ above described. An indispensable ingredient of this proof, of independent interest, is the converse of Hölder inequality, which we shall now prove. In a measure space $(X, \mathcal{M}, \mu)$ we denote by $S(\mu)$ the linear space of all measurable simple functions on X which vanish outside of a set of finite measure; clearly $S(\mu)$ is a subspace of $L^p(\mu)$ for every p, with $p \le \infty$; and $S(\mu)$ is dense in $L^p(\mu)$ for every $p < \infty$.

. Converse of Hölder inequality *Let $1 \le p < \infty$, let $(X, \mathcal{M}, \mu)$ be a semifinite measure space, and let $g : X \to \mathbb{K}$ be a measurable function such that for every $f \in S(\mu)$ we have $f\, g \in L^1(\mu)$, and the functional $\varphi_g : S(\mu) \to \mathbb{K}$ defined by*

$$\varphi_g(f) = \int_X f\, g$$

is a continuous linear functional on the space $(S(\mu), \|\#\|_p)$. Then g belongs to $L^q(\mu)$, with q the conjugate exponent of p, and $\|g\|_q$ coincides with the norm of φ_g (as element of the dual of $S(\mu)$ with the L^p norm).

Proof. Denote by k the norm of φ_g. We first prove that if $f \in L^\infty(\mu)$ is zero outside a set of finite measure then $|\varphi_g(f)| \le k\, \|f\|_p$. For, there exists a sequence f_n of simple functions, zero outside the set $E = \{f \ne 0\}$, such that $f_n \to f$ and $|f_n| \uparrow |f|$ pointwise; then $f_n\, g$ is dominated by $\|f\|_\infty\, |g|\, \chi_E$; by hypothesis $g\, \chi_E$ is in $L^1(\mu)$; then in the inequality

$$\left| \int_X f_n\, g \right| \le k\, \|f_n\|_p,$$

we may apply Lebesgue's dominated convergence theorem to show that

$$\left| \int_X f\, g \right| \le k\, \|f\|_p.$$

Assume now $1 < p < \infty$. Given $\alpha > 0$ we prove that $\{|g| > \alpha\}$ has finite measure, for every $\alpha > 0$. In fact, if $E \subseteq \{|g| > \alpha\}$ has finite measure, the function $f = \operatorname{sgn} \bar{g}\, \chi_E$ is bounded and zero outside of E, so that

$$\int_X f\, g = \int_E |g| \le k\, (\mu(E))^{1/p} \quad \text{and on the other hand} \quad \int_E |g| \ge \alpha\, \mu(E),$$

so that

$$\alpha\, \mu(E)^{1-1/p} \le k \iff (\mu(E))^{1/q} \le k/\alpha.$$

Since the measure is by hypothesis semifinite, this proves that $\mu(\{|g| > \alpha\}) < \infty$. Let now $f_n = \operatorname{sgn} \bar{g}\, |g|^{q-1}\, \chi(E(n)$, where $E(n) = \{1/n < |g| \le n\}$; then f_n is bounded and zero outside a set of finite measure, so that

$$\int_X f_n\, g \le k\, \|f_n\|_p \quad \text{that is} \quad \int_X |g|\, \chi_{E(n)} \le k\, \|f_n\|_p = k \left(\int_X |g|^{pq-p}\, \chi_{E(n)} \right)^{1/p},$$

which implies

$$\left(\int_X |g|^q\, \chi_{E(n)} \right)^{1/q} \le k,$$

and taking limits, by monotone convergence we get

$$\|g\|_q \le k.$$

The case $p = 1$ can be proved by imitating the last part of the argument given in the previous section. □

1.2.3. We now prove:

. If $(X, \mathcal{M}, \mu)$ is a measure space and $1 < p < \infty$, $q = p/(p-1)$ then the mapping $g \mapsto \varphi_g$ described in 1.2.1 is an isometric isomorphism of $L^q(\mu)$ onto $(L^p(\mu))^*$. The same is true for $p = 1$ and $q = \infty$ if μ is $\sigma-$finite.

Proof. (only for $\sigma-$finite μ). By 1.2.1 only surjectivity of $g \mapsto \varphi_g$ is left to prove. Take $\varphi \in (L^p(\mu))^*$.

Assume first that $\mu(X) < \infty$. Define $\nu : \mathcal{M} \to \mathbb{K}$ by $\nu(E) = \varphi(\chi_E)$. Then ν is a $\mathbb{K}-$valued measure on $\mathcal{M}$: if $(E_n)_n$ is a disjoint sequence of elements of $\mathcal{M}$, $E = \bigcup_{n\in\mathbb{N}} E_n$, and f_m is the characteristic function of $\bigcup_{n=0}^m E_n$, then f_m converges to χ_E in $L^p(\mu)$ (because $p < \infty$), so that

$$\varphi(\chi_E) = \nu(E) = \lim_m \varphi(f_m) = \lim_m \left(\sum_{n=0}^m \varphi(\chi_{E_n}) \right) = \lim_m \sum_{n=0}^m \nu(E_n) = \sum_{n=0}^\infty \nu(E_n).$$

Since clearly we have $\nu \ll \mu$ the Radon–Nikodym theorem asserts that there is $g \in L^1(\mu)$ such that

$$\nu(E) = \int_E g\, d\mu \qquad \text{for every} \quad E \in \mathcal{M}.$$

By linearity we have

$$\varphi(f) = \int_X f\, d\nu = \int_X f\, g\, d\mu \quad \text{for every } f \in S(\mu)$$

(remember that $f = \sum_{k=1}^m \alpha_k\, \chi_{E(k)}$ for every $f \in S(\mu)$); continuity of φ shows that

$$|\varphi(f)| \le \|\varphi\|_{(L^p(\mu))^*}\, \|f\|_p \quad \text{for every } f \in S(\mu),$$

and the converse of Hölder inequality allows us to conclude that $g \in L^q(\mu)$ and $\|g\|_q = \|\varphi\|_{(L^p(\mu))^*}$.

By density of $S(\mu)$ in $L^p(\mu)$ we have that $\varphi(f) = \int_X f\, g$ for every $f \in L^p(\mu)$. If μ is $\sigma-$finite, let $(X(n))_{n\in\mathbb{N}}$ be an increasing sequence of sets in $\mathcal{M}$ of finite measure whose union is X. By the preceding argument, for every $n \in \mathbb{N}$ we get a (essentially unique) function $g_n \in L^q(\mu_n)$ (where $\mu_n(E) = \mu(E\cap X(n))$) such that $\varphi(f) = \int_{X(n)} f\, g_n\, d\mu_n$ for every $f \in L^p(\mu)$ which is zero outside of $X(n)$. Moreover the norm of φ, restricted to the subspace $L^p(X(n), \mu)$ of functions in $L^p(\mu)$ which are zero outside of $X(n)$ is $\|g_n\|_q$, and clearly this norm is dominated by $k = \|\varphi\|_{(L^p(\mu))^*}$. Since $g_{n+1|X(n)} = g_n$ q.o. in $X(n)$, up to modifications on a negligible set we have a measurable $g : X \to \mathbb{K}$ such that $g_{|X(n)} = g_n$ for every n. Then $g \in L^q(\mu)$: in fact $g\, \chi_{X(n)} = g_n$ and $\|g_n\|_q \le k$ for every n; if $q = \infty$ we immediately have $\|g\|_\infty \le k$, whereas if $q < \infty$:

$$\|g_n\|_q^q = \int_{X(n)} |g|^q\, d\mu = \|g_n\|_q^q \le k^q \quad \text{so that, by monotone convergence} \quad \int_X |g|^q \le k^q.$$

If $f \in L^p(\mu)$ and $f_n = f\, \chi_{X(n)}$ we have that f_n converges to f in $L^p(\mu)$, so that $\varphi(f_n)$ converges in $\mathbb{K}$ to $\varphi(f)$; but $\varphi(f_n) = \int_{X(n)} f_n\, g\, d\mu$ and since g is in $L^q(\mu)$ we have $|f\, g| \in L^1(\mu)$ by Hölder inequality; then dominated convergence may be applied to $f_n\, g$ to prove that also $\int_{X(n)} f_n\, g$ converges to $\int_X f\, g$, for every $f \in L^p(\mu)$. □

Exercise 1.2.3.1. Let $(X, \mathcal{M}, \mu)$ be a measure space. Let $g \in L(X)$ be a measurable function such that for every $f \in L^\infty(\mu)$ we have $f\, g \in L^1(\mu)$. Prove that then g is in $L^1(\mu)$.

Assume now that μ is also semifinite. Let $g \in L(X)$ be such that $f\, g \in L^1(\mu)$ for every $f \in L^1(\mu)$. Prove that then $g \in L^\infty(\mu)$ (that is, boundedness of the linear map $f \mapsto \int_X f\, g$ can be removed from the hypotheses in 1.2.2, adding the hypothesis that $f\, g \in L^1(\mu)$ for every $f \in L^1(\mu)$ and not only for $f \in S(\mu)$, in the case $p = 1$. We shall prove later that this holds for every $p \ge 1$).

1.3. Completions. There is often the need of enlarging a non–complete metric space to a complete metric space, to add elements to it to give a limit to every Cauchy sequence. The first encountered instance of this problem is the enlarging of the rational field $\mathbb{Q}$ to the real field $\mathbb{R}$. Here we discuss the general procedure for solving the problem, different however from the one followed for getting $\mathbb{R}$ from $\mathbb{Q}$ (we can exploit the availability of the complete field $\mathbb{R}$). Completions are indispensable in various settings, not only in Analysis, but in Algebra, Geometry, etc.

1.3.1. *Extension of uniformly continuous functions.*

Theorem. *Let X be a metric space, Y a complete metric space, D a dense subspace of X, $f : D \to Y$ a function. Then:*

(i) Extension theorem for uniformly continuous functions, *If f is uniformly continuous, then f has a unique uniformly continuous estension $\bar{f} : X \to Y$; and if f is Lipschitz continuous, then the extension is also Lipschitz continuous, with the same Lipschitz constant.*
(ii) Extension theorem for continuous linear operators *Let X, Y, D be normed spaces, with Y Banach, D a dense subspace of X, and $f : D \to Y$ linear continuous, then f has a unique continuous linear extension to X, with the same norm.*

Proof. (i) Recall that a uniformly continuous function transforms Cauchy sequences into Cauchy sequences; if $x \in X$, since D is dense in X there is a sequence $x_n \in D$ which converges in X to x; being convergent, this sequence is Cauchy, so that $f(x_n)$ is a Cauchy sequence in Y by uniform continuity of f, hence converges in Y to a (unique) limit that we call $\bar{f}(x)$. Notice that this is a good definition, in the sense that if ξ_n is another sequence in D converging to x, then $\lim_n f(\xi_n) = \lim_n f(x_n)$: this may be seen because $x_0, \xi_0, x_1, \xi_1, \ldots$ is another sequence of D converging to x, whose image by f has to have a limit in Y, and this image has both sequences $f(x_n)$ and $f(\xi_n)$ as subsequences, or directly because $d_X(x_n, \xi_n) \to 0$ and uniform continuity of f together imply $d_Y(f(x_n), f(\xi_n)) \to 0$. Moreover $\bar{f}(x) = f(x)$ if $x \in D$, because f is continuous on D.

Let's now prove uniform continuity of $\bar{f}$: given $\varepsilon > 0$ find $\delta > 0$ such that $d_X(x,y) < \delta$ implies $d_Y(f(x), f(y)) \le \varepsilon$; if $x, y \in X$, $d_X(x,y) < \delta$, and x_n, y_n are sequences in D converging to x, y respectively, by continuity of d_X we have $d_X(x_n, y_n) < \delta$ for $n \ge n_\varepsilon$, so that $d_Y(f(x_n), f(y_n)) \le \varepsilon$ if $n \ge n_\varepsilon$, and by continuity of d_Y we get, passing to the limit, $d_Y(\bar{f}(x), \bar{f}(y)) \le \varepsilon$. Uniqueness is by the density of D in X. For the Lipschitz version: since f is Lipschitz continuous, it is uniformly continuous, hence it has a unique uniformly continuous extension $\bar{f}$, by (i). If $\lambda > 0$ is a Lipschitz constant for f, the same λ is also a Lipschitz constant for $\bar{f}$: if $x, y \in X$ and x_n, y_n are sequences in D converging to x, y respectively, we have

$$d_Y(f(x_n, f(y_n)) \le \lambda\, d_X(x_n, y_n) \Longrightarrow d_Y(\bar{f}(x), \bar{f}(y)) \le \lambda d_X(x,y),$$

simply passing to the limit as $n \to \infty$.

(ii) Since continuous linear operators between normed spaces are Lipschitz continuous, with the best Lipschitz constant as norm, the only thing that remains to check is linearity of the extension, which is immediate from the continuity of vector operations (e.g.: the mapping $\bar{f}(x+y) - (\bar{f}(x) + \bar{f}(y))$ is zero on $D \times D$, a dense subset of $X \times X$, so it is zero on all $X \times X$, because the mapping is continuous from $X \times X$ to X). □

It may be worth noting that, while continuous linear maps of normed space are Lipschitz continuous, in particular uniformly continuous, the only uniformly continuous bilinear map is the zero map! Assume in fact that $\beta : X \times Y \to Z$, continuous and hence verifying $\|\beta(x,y)\| \le L\|x\|\,\|y\|$ for some $L > 0$ is uniformly continuous, i.e. that given $\varepsilon > 0$ there is $\delta > 0$ such that $\|x - a\| \le \delta$ and $\|y - b\| \le \delta$ imply $\|\beta(x,y) - \beta(a,b)\| \le \delta$. Then, for every $b \in Y$ we have $\|\beta(x,b) - \beta(0,b)\| \le \varepsilon$, provided that $\|x\| \le \delta$. Since $\beta(0,b) = 0$ in fact we have:

$$\|\beta(x,b)\| \le \varepsilon \quad \text{if } \|x\| \le \delta, \text{ for every } b \in Y.$$

This is possible if and only if $\beta(x,b) = 0$ ($\varepsilon \ge \|\beta(x,tb)\| = t\,\|\beta(x,b)\|$ for every $t > 0$ implies $\|\beta(x,b)\| = 0$). And clearly $\beta(x,b) = 0$ for every $b \in Y$ and every $x \in \delta\, B_X$ implies $\beta(x,y) = 0$ for every $(x,y) \in X \times Y$ (write $x = (\|x\|/\delta)\, u$ with $u = \delta x/\|x\| \in \delta\, B_X$). Thus extension of bilinear mappings is not entirely straightforward.

Exercise 1.3.1.1. Extension of bilinear continuous mappings Let X, Y, Z be normed spaces, with Z complete; let E, F be linear subspaces dense in X and Y respectively, and let $\beta : E \times F \to Z$ be bilinear and continuous. Then β has a unique continuous extension $\bar{\beta} : X \times Y \to Z$; this extension is bilinear and with the same norm of β.

Solution. It is understood that the norm on $X \times Y$ is the product norm $\|(x,\, y)\|_{X\times Y} = \|x\|_X \vee \|y\|_Y$. Since β is bilinear continuous there is $L > 0$ such that $\|\beta(x,y)\|_Z \le L\|x\|_X\, \|y\|_Y$ for every $x \in E$ and $y \in F$ (the smallest such constant is the norm of β). This inequality says that for every given $x \in E$ the mapping $\beta_x : y \mapsto \beta(x,\, y)$ is linear continuous from F to Z, of norm not exceeding $L\,\|x\|_X$, so that β_x extends to a linear map of the same norm $\bar{\beta}_x : Y \to Z$; in this way $\bar{\beta}(x,y)$ is defined for all $x \in E$ and all $y \in Y$, is linear in y for every given $x \in E$, and moreover $\|\beta(x,y)\|_Z \le L\|x\|_X\, \|y\|_Y$ for every $x \in E$ and $y \in Y$. But $\bar{\beta}(x,y)$ is linear also in x for every given $y \in Y$: given x_1, $x_2 \in E$ and $\lambda, \mu \in \mathbb{K}$ we have

$$\beta(\lambda\, x_1 + \mu\, x_2, y_n) = \lambda\, \beta(x_1,\, y_n) + \mu\, \beta(x_2,\, y_n),$$

and passing to the limit as $y_n \in F$ tends to $y \in Y$ we get

$$\bar{\beta}(\lambda\, x_1 + \mu\, x_2, y) = \lambda\, \bar{\beta}(x_1,\, y) + \mu\, \bar{\beta}(x_2,\, y),$$

as desired.

We repeat the procedure, keeping now $y \in Y$ fixed, and extend to X the linear map $x \mapsto \beta(x,y)$, Lipschitz continuous on E . We get also

$$\|\bar{\beta}(x,y)\|_Z \le L\|x\|_X\, \|y\|_Y \quad \text{for all } x \in X \text{ and } y \in Y,$$

which proves continuity of the bilinear map $\bar{\beta}$. □

1.3.2. *Completion of a metric space.* Given a metric space (X,d), we call *completion* of X any metric space $(\tilde{X}, \tilde{d})$ which is complete and of which X is a dense metric subspace: that is $\tilde{d}_{|X\times X} = d$, the space $(\tilde{X},\, \tilde{d})$ is complete, and X is dense in $\tilde{X}$. If X is complete and $\tilde{X}$ is completion of X, then X is closed in $\tilde{X}$ (a complete metric space is closed in any metric space in which it is embedded), and by density $\tilde{X} = X$: completing a complete space we do not get a new space. Completions clearly cannot be unique in a strict sense; they are unique in the following sense:

. *If* $(\tilde{X},\tilde{d})$ *and* $(\bar{X},\bar{d})$ *are completions of the same metric space* (X,d)*, then there is an isometry of* $(\tilde{X},\tilde{d})$ *onto* $(\bar{X},\bar{d})$*, which is the identity on* X*.*

The proof is easy to do with the extension theorem for Lipschitz continuous functions.

To construct a completion we may embed isometrically a metric space in some complete metric space, and take the closure of the embedding. We know (see Analisi Due, 2.3.9) that every metric space (X,d) can be isometrically embedded in $\ell^\infty(X)$ (fix a point $a \in X$ and for every $x \in X$ define $\varphi_x : X \to \mathbb{R}$ by $\varphi_x(y) = d(x,y) - d(a,y)$; it is easy to prove that $x \mapsto \varphi_x$ is an isometry of X onto a subset of $\ell^\infty(X)$); since ℓ^∞ is complete, the closure of $\{\varphi_x : x \in X\}$ in ℓ^∞ will give a space isometric to a completion of X).

A normed space X has a completion as a metric space which can be given in a unique way a linear space structure making it into a Banach space with X as a dense subspace; addition is extended by uniform continuity, and multiplication scalars×vectors as a bilinear continuous map. But we can also consider the isometric embedding $J : X \to X^{**}$ of X into its double dual, and take the closure of X in X^{**}: in this way the vector space structure is already present. A scalar product space has a completion which becomes a Hilbert space: scalar multiplication is extended as an $\mathbb{R}-$bilinear map.

It is however important to keep in mind that, although an "abstract" completion of a space always exists and is unique (in the sense described) this completion is almost useless without a lot of further work, especially without some concrete handy way of representing its elements. It is very easy to describe $L^1(\mathbb{R})$ in this way: consider the space $C_c(\mathbb{R})$ of all continuous functions with compact support on $\mathbb{R}$, normed with

$$\|f\|_1 = \int_{\mathbb{R}} |f(x)|\, dx;$$

then $L^1(\mathbb{R})$ is its completion. Unfortunately this easy definition cannot spare us the work of constructing Lebesgue measure!

Exercise 1.3.2.1. The completion of a metric space is compact if and only if the space is totally bounded.

1.3.3. *Some topological facts.* Recall that a topological space is said to be *separable* if it has a countable dense subset. A topological space is said to be *second countable* when it satisfies the second axiom of countability, that is, it has a countable base for its open sets. Recall that if τ is a topology on the set X, a base for τ is a subset $\mathcal{B} \subseteq \tau$ such that every $A \in \tau$ is a union of members of $\mathcal{B}$, that is, $A = \bigcup\{B \in \mathcal{B} : B \subseteq A\}$. The following statement is trivial:

. *If* (X,τ) *is a topological space, a subset* $\mathcal{B}$ *of* τ *is a base for* τ *if and only if given* $p \in X$ *and* $A \in \tau$ *with* $p \in A$ *we have that there is* $B \in \mathcal{B}$ *with* $p \in B$ *and* $B \subseteq A$*.*

Next we observe:

. *Every second countable space is separable.*

Proof. Given a countable base $\mathcal{B}$, pick $x_B \in B$ for every non–empty $B \in \mathcal{B}$; the subset D of X so obtained is countable and dense: given an open non–empty subset A of X, A contains some non–empty $B \in \mathcal{B}$, so $x_B \in A$, and $A \cap D$ is non–empty. □

But separability does not imply second countability: the Sorgenfrey line S, described in Analisi Due, 2.4.15. is an example of a separable not second countable space. In metrizable spaces this cannot happen:

Proposition. *A metrizable space is second countable if and only if it is separable.*

Proof. Let D be a countable dense subset of the metrizable space X, and let d be a topology determining metric on X. Let $\mathcal{B}$ be the set of all open balls with center at points $x \in D$ and rational strictly positive radii,

$$\mathcal{B} = \{B(x,r[: x \in D,\, r \in \mathbb{Q}^>\}.$$

Clearly $\mathcal{B}$ is countable, and we prove that it is a base for τ. If $p \in X$ and $A \in \tau$ with $p \in A$, pick first $r > 0$ such that $B(p,r[\subseteq A$. Next, pick $x \in D$ such that $d(p,x) < r/3$ (possible because D is dense in X) and a rational ρ with $r/3 < \rho < 2r/3$. Then $p \in B(x,\rho[$, and $B(x,\rho[\subseteq B(p,r[\subseteq A$. In fact, if $y \in B(x,\rho[$ we have

$$d(y,p) \le d(y,x) + d(x,p) < \rho + \frac{r}{3} < \frac{2r}{3} + \frac{r}{3} = r.$$

□

Second countability is a hereditary property, which means that every subspace of a second countable space is also second countable (trivial: if $\mathcal{B}$ is a base for τ and $S \subseteq X$ then $\mathcal{B}_S = \{B \cap S : B \in \mathcal{B}\}$ is a base for the induced topology on S). Hence a second countable space is also *hereditarily separable*, every subspace of a second countable space is separable. In particular, in a metrizable separable space every subspace discrete in the induced topology is at most countable (in a discrete space, no proper subset is dense), in particular every uniformly discrete subset is at most countable.

We prove here non separability of some spaces (cfr. 1.9.13.1).

EXAMPLE 1.3.3.1. The unit sphere of $\ell^\infty(\Lambda)$ is not separable if Λ is infinite.

In fact, consider the elements $\{e_S : S \subseteq \Lambda\}$, where e_S is the characteristic function of the subset S of Λ. If $S \neq T$, then $\|e_S - e_T\|_\infty = 1$, so that this set is uniformly discrete, with cardinality at least equal to the continuum.

EXAMPLE 1.3.3.2. None of the spaces $\ell^p(\Lambda)$ and $c_0(\Lambda)$ is separable if Λ is uncountable.

In fact we have, if λ, $\mu \in \Lambda$ with $\lambda \neq \mu$:

$$\|e_\lambda - e_\mu\|_p = 2^{1/p}; \qquad \|e_\lambda - e_\mu\|_\infty = 1.$$

1.4. The Hahn–Banach theorem.

1.4.1. *Functionals.* Real or complex valued functions defined on linear spaces X of interest in Analysis are often called *functionals* instead of functions; the elements of X frequently are functions, and functionals act on these functions. The name functional thus serves to distinguish them from the elements of the space; the term is perhaps a bit old fashioned, but still very much in use.

1.4.2. *Subadditive and sublinear functionals.* On a $\mathbb{K}-$vector space X a (real valued) functional $p : X \to \mathbb{R}$ is said to be *subadditive* if $p(x + y) \leq p(x) + p(y)$ for every x, $y \in X$, and is called *sublinear* if it is subadditive and *positively homogeneous* (of degree 1), meaning that $p(t\,x) = t\,p(x)$ for every real $t \geq 0$. Notice that if $p : X \to \mathbb{R}$ is sublinear, then $p(0) = p(0\,x) = 0\,p(x) = 0$. As usual, the term *linear* is reserved for those functionals $f : X \to \mathbb{K}$ such that $f(x + y) = f(x) + f(y)$ for every x, $y \in X$, and $f(t\,x) = t\,f(x)$ for every $t \in \mathbb{K}$ and $x \in X$: linear functionals are also called *linear forms*; they are the homomorphisms of the given $\mathbb{K}-$vector space into the scalar field $\mathbb{K}$, and clearly form a vector subspace, sometimes denoted $\mathrm{Hom}_{\mathbb{K}}(X, \mathbb{K})$, of the space $\mathbb{K}^X$ of all functions from X to $\mathbb{K}$, with pointwise operations.

EXAMPLE 1.4.2.1. Let $X = \ell^\infty(\mathbb{N}, \mathbb{R})$ be the $\mathbb{R}-$space of bounded real valued sequences. It is well–known that X is a Banach space under the sup–norm $\|x\|_\infty = \sup\{|x(n)| : n \in \mathbb{N}\}$ (1.1.2). Then

$$p(x) := \limsup_{n\to\infty} x(n)$$

is a sublinear functional, which is not positive, and not linear: sublinearity depends on well–known properties of lim sup. Notice that $p(-x) = -\liminf_{n\to\infty} x_n$.

1.4.3. *Seminorms and norms.* A subadditive functional $p : X \to \mathbb{R}$ which is also *absolutely homogeneous* (of degree 1), i.e. such that $p(\alpha\,x) = |\alpha|\,p(x)$ for every $\alpha \in \mathbb{K}$ and every $x \in X$, is called *seminorm* (an $\mathbb{R}-$seminorm if $p(\alpha\,x) = |\alpha|\,p(x)$ holds for $\alpha \in \mathbb{R}$). A seminorm is sublinear and necessarily positive: in fact

$$0 = p(0) = p(x - x) \leq p(x) + p(-x) = 2p(x) \qquad \text{for every } x \in X.$$

Recall that a seminorm p is called *norm* when $p(x) = 0$ implies $x = 0$: a norm is a seminorm which is zero only at the origin.

EXAMPLE 1.4.3.1. For every linear functional $f : X \to \mathbb{K}$ the functional $x \mapsto |f(x)|$ is a seminorm.

EXERCISE 1.4.3.2. If, $p, q : X \to \mathbb{R}$ are sublinear and $\lambda, \mu > 0$ are real numbers, then also $\lambda\,p + \mu\,q$ and $p \vee q$ (defined on $x \in X$ as $\max\{p(x), q(x)\}$) are sublinear, and are seminorms if p and q are seminorms. Also, $\tilde{p}$ (defined as $\tilde{p}(x) = p(-x)$) is sublinear. Prove that if p is sublinear then $p \vee \tilde{p}$ is an $\mathbb{R}-$seminorm.

EXERCISE 1.4.3.3. If $p : X \to [0, +\infty[$ is a seminorm, then the nullspace of p, that is $N_p = \{x \in X : p(x) = 0\}$ is a linear subspace of X. On the quotient space X/N_p the functional $q(x + N_p) = p(x)$ is well defined, and is a norm (cfr with the spaces of summable functions of 1.1.1).

EXERCISE 1.4.3.4. Let $(X, \|\#\|)$ be a normed space, and let V be a subspace. Prove that the function $x \mapsto \mathrm{dist}(x, V)$ is a seminorm on X (recall that $\mathrm{dist}(x, V) := \inf\{\|x - y\| : y \in V\}$), whose nullspace is the closure of V.

Solution. Let $x, y \in X$; for every $u, v \in V$ we have, since $u+v \in V$:

$$\operatorname{dist}(x+y, V) \le \|(x+y)-(u+v)\| = \|(x-u)+(y-v)\| \le \|x-u\| + \|y-v\|,$$

so that

$$\operatorname{dist}(x+y,V) \le \inf\{\|x-u\|+\|y-v\| : u,v \in V\} = \inf\{\|x-u\| : u \in V\} + \inf\{\|y-v\| : v \in V\} = \operatorname{dist}(x,V)+\operatorname{dist}(y,V).$$

If $\alpha \in \mathbb{K} \setminus \{0\}$ every vector in V may be written as u/α, with $u \in V$, so that

$$\operatorname{dist}(\alpha\, x, V) = \inf\{\|\alpha\, x - u\| : u \in V\} = \inf\{|\alpha|\, \|x-u/\alpha\| : u \in V\} = |\alpha| \inf\{\|x-u/\alpha\| : u \in V\} = |\alpha| \operatorname{dist}(x,V).$$

Finally, it is well–known that in every metric space the function "distance from a subset" is zero exactly on the closure of the subset. □

Recall that in a normed space $(X, \|\cdot\|)$ a linear functional $f \in \operatorname{Hom}_{\mathbb{R}}(X,\mathbb{R})$ is continuous if and only if there exists a constant $k \ge 0$ such that $|f(x)| \le k\,\|x\|$, for all $x \in X$. Since $x \mapsto k\,\|x\|$ is also a norm if $k > 0$, continuity of f is equivalent to f being dominated by a norm multiple of the original norm. This motivates what follows.

1.4.4. *The Hahn–Banach theorem: analytic form.*

. *Let X be a real vector space, let $p : X \to \mathbb{R}$ be a sublinear functional. Assume that V is a linear subspace of X, that $f : V \to \mathbb{R}$ is linear, and that $f(x) \le p(x)$ for every $x \in X$. Then f has an extension to a linear functional $g : X \to \mathbb{R}$ such that $g(x) \le p(x)$, for every $x \in X$.*

Proof. We first prove the

Claim: if $V \ne X$, then given any $a \in X \setminus V$, the functional f may be extended to a linear functional on the linear space $V + \mathbb{R}\,a$ generated by V and a, which is still under p. To extend f linearly we have to pick $\alpha \in \mathbb{R}$ as image of a; in other words the extension is defined by $x + t\,a \mapsto f(x) + t\,\alpha$, for every $x \in V$ and every $t \in \mathbb{R}$. We have to prove that $\alpha \in \mathbb{R}$ may be chosen in such a way that

$$f(x) + t\,\alpha \le p(x+t\,a), \quad \text{for every } X \in V \text{ and every } t \in \mathbb{R}.$$

If $t = 0$ any α will do; if $t > 0$ the inequality is equivalent to $f(x/t) + \alpha \le p(x/t + a)$; since x/t is an arbitrary element $v \in V$, $\alpha \in \mathbb{R}$ must satisfy the condition

$$f(v) + \alpha \le p(v+a) (\iff \alpha \le p(v+a) - f(v)), \quad \text{for every } v \in V.$$

If $t = -s < 0$, the inequality is equivalent to $f(x/s) - \alpha \le p(x/s - a)$; again x/s is an arbitrary element $u \in V$, so that α must also satisfy the condition

$$f(u) - \alpha \le p(u-a) (\iff \alpha \ge f(u) - p(u-a)), \quad \text{for every } u \in V.$$

A real number α with the required conditions will then exist if and only if

$$f(u) - p(u-a) \le p(v+a) - f(v), \quad \text{for every } u, v \in V.$$

This condition is equivalent to $f(u) + f(v) \le p(v+a) + p(u-a)$, for every $u, v \in V$; but $f(u) + f(v) = f(u+v)$ by linearity of f on V, and by the hypotesis we also have

$$f(u+v) \le p(u+v) = p(u-a+a+v) \le p(u-a) + p(a+v), \quad \text{by subadditivity.}$$

The claim is proved.

If X is finite dimensional, more generally if the codimension of V in X is finite, we can now easily prove the theorem by induction on the codimension of V. In the general case transfinite induction is needed, and the theorem follows from an easy application of Zorn's lemma (see, e.g. Analisi Uno, Appendix), as we now show.

We consider the partially ordered set $\mathcal{P}$ consisting of all pairs (g, W), where W is an $\mathbb{R}-$subspace of X containing V, i.e. $V \subseteq W \subseteq X$, and $g : W \to \mathbb{R}$ is an $\mathbb{R}$-linear function with domain W which concides with f on V, and is majorized by p on W, i.e. $g(x) \le p(x)$ for all $x \in W$; the partial order $\preceq$ on $\mathcal{P}$ is by extension, that is, $(g_1, W_1) \preceq (g_2, W_2)$ means that $W_1 \subseteq W_2$, and that $g_2|W_1 = g_1$. If $\mathcal{W}$ is a chain(=totally ordered subset) of $\mathcal{P}$ it is easy to find an upper bound for $\mathcal{W}$ in $\mathcal{P}$: let

$$Z = \bigcup_{(g,W)\in\mathcal{W}} W;$$

then Z is a linear subspace (it is the set theoretic union of a chain of subspaces) of X containing V, and clearly if $h : Z \to \mathbb{R}$ is defined by setting $h(x) = g(x)$ for $x \in W$ with $(g, W) \in \mathcal{W}$ then h is well–defined, is linear, extends all g with $(g, W) \in \mathcal{W}$ and verifies $h(x) \le p(x)$ for every $x \in Z$; in short, (h, Z) belongs to $\mathcal{P}$ and clearly $(g, W) \preceq (h, Z)$ for every $(g, W) \in \mathcal{W}$.

Then Zorn's lemma applies, and by the claim proved at the beginning any maximal element of $\mathcal{P}$ is of the form (g, X), i.e. its domain space is all of X. □

1.4.5. The Hahn–Banach theorem is essential for proving that in a normed space there is a rich supply of continuous linear functionals, as we show in the next subsection. We first observe that if $p : X \to \mathbb{R}$ is a (real) seminorm, f is a real linear functional, and $f(x) \le p(x)$ for every $x \in X$, then also $|f(x)| \le p(x)$ for every $x \in X$: this is immediate since $f(-x) \le p(-x)$ by the hypothesis, but $f(-x) = -f(x)$ and $p(-x) = p(x)$ if p is a real seminorm, so that we have $f(x) \le p(x)$ and also $-f(x) \le p(x)$, equivalent to $|f(x)| \le p(x)$. In other word, if a real seminorm is pointwise above a real linear functional, then it also dominates it. To prove a similar result for complex linear functional we need a little work. If X is a $\mathbb{C}-$vector space and $f : X \to \mathbb{C}$ is $\mathbb{C}-$linear, then $\operatorname{Re} f, \operatorname{Im} f$ are both $\mathbb{R}-$linear functionals, of course. But the interesting and perhaps unexpected fact is that (trivially!) f is entirely determined by its real part, and that if f is continuous with respect to some norm on X, then $\|f\|_{\mathrm{L}_{\mathbb{C}}(X,\mathbb{C})} = \|\operatorname{Re} f\|_{\mathrm{L}_{\mathbb{R}}(X,\mathbb{R})}$. In fact, shortening $\operatorname{Re} f(x)$ to $u(x)$ and $\operatorname{Im} f(x)$ to $v(x)$ we have $f(x) = u(x) + i\,v(x)$, and by $\mathbb{C}-$linearity $f(ix) = i\,f(x) = i\,u(x) - v(x)$, but $f(ix) = u(ix) + i\,v(ix)$ so that $v(x) = -u(ix)$ for every $x \in X$, and one can write $f(x) = u(x) - i\,u(ix)$. Moreover, if p is a complex seminorm then $|f(x)| \le p(x)$ holds for every $x \in X$ iff $\operatorname{Re} f(x) \le p(x)$ for every $x \in X$: clearly $\operatorname{Re} f(x) \le |f(x)| \le p(x)$, so that we have to prove that $\operatorname{Re} f(x) \le p(x)$ for every x implies $|f(x)| \le p(x)$ for every x. If $f(x) \ne 0$ let $\alpha = \overline{\operatorname{sgn} f(x)}$, so that $\alpha\, f(x) = |f(x)|$ and $|\alpha| = 1$: we have $|f(x)| = \alpha\, f(x) = f(\alpha\, x) = \operatorname{Re} f(\alpha\, x)$; but by hypothesis $\operatorname{Re} f(\alpha\, x) \le p(\alpha\, x) = |\alpha|\, p(x) = p(x)$, and we conclude. We can collect these results in a statement:

. *Let X be a $\mathbb{K}-$linear space, and let p be a $\mathbb{K}-$seminorm on X. For a linear functional $f : X \to \mathbb{K}$ the following are equivalent: (i) $|f(x)| \le p(x)$ for every $x \in X$; (ii) $\operatorname{Re} f(x) \le p(x)$ for every $x \in X$.*

1.4.6. We leave it to the reader to extract a proof of the following statement from the preceding considerations:

. *Let X be a complex linear space. Then $f \mapsto \operatorname{Re} f$ is an $\mathbb{R}-$linear isomorphism of $\operatorname{Hom}_{\mathbb{C}}(X, \mathbb{C})$ onto $\operatorname{Hom}_{\mathbb{R}}(X, \mathbb{R})$, with inverse $u \mapsto u(\#) - i\,u(i\#)$. If X is normed, the restriction to continuous linear forms is an isometric $\mathbb{R}-$isomorphism of the normed complex dual $\mathrm{L}_{\mathbb{C}}(X, \mathbb{C})$ onto the normed real dual $\mathrm{L}_{\mathbb{R}}(X, \mathbb{R})$.*

1.4.7. *Extension of continuous linear forms.*

. *Let X be a normed $\mathbb{K}-$space. Assume that V is a subspace, and that $f : V \to \mathbb{K}$ is linear and continuous. Then f may be extended to a linear mapping g on all of X, with g and f having the same norm.*

Proof. We have to prove that if $k \ge 0$ is such that $|f(x)| \le k\,\|x\|$ for all $x \in V$, then there is a linear $g : X \to \mathbb{K}$ such that $g(x) = f(x)$ for $x \in V$, and $|g(x)| \le k\,\|x\|$ for all $x \in X$. If $\mathbb{K} = \mathbb{R}$ this is exactly the Hahn–Banach theorem applied with $p(x) = k\,\|x\|$ (taking account of the preceding section for the lack of the absolute value). If $\mathbb{K} = \mathbb{C}$ we apply the theorem to the real part $u = \operatorname{Re} f$ of f, which of course verifies the inequality $\operatorname{Re} f(x) \le |f(x)| \le k\,\|x\|$ for every $x \in V$, obtaining a real linear $U : X \to \mathbb{R}$ with $|U(x)| \le k\,\|x\|$. Now $g(x) = U(x) - i\,U(ix)$ is a complex linear extension with the same norm. □

1.4.8. From now on, given a normed space X, its dual X^*, already mentioned in 1.1.5, will be the space of all continuous linear forms on the space, equipped with the operator norm (the space $\operatorname{Hom}_{\mathbb{K}}(X, \mathbb{K})$ of all linear forms, continuous or not, will be called *algebraic dual* of X, but we will have almost no use for it). The dual of a normed space separates the points of the space, as we now explain.

. *Given a normed space X and a nonzero $a \in X$ there exists a continuous linear functional f in the dual of X such that $\|f\| = 1$ and $f(a) = \|a\| \ne 0$.*

Proof. Define the form $f : \mathbb{K}\,a \to \mathbb{K}$ by $f(\alpha\, a) = \alpha\,\|a\|$; clearly $\|f\| = 1$ and $f(a) = \|a\|$. The linear map f may be extended to a linear map (still called f) on all of X, without enlarging the norm. □

Exercise 1.4.8.1. If X, Y are normed spaces and $T \in L(X, Y)$, we can define the *transpose* $T^t : Y^* \to X^*$ by $T^t(\varphi) = \varphi \circ T$, for every $\varphi \in Y^*$ (some authors call T^t the *adjoint* of T, but we shall reserve the name adjoint for the Hilbert spaces case). Prove that T^t is continuous and that $\|T^t\|_{L(Y^*,X^*)} = \|T\|_{L(X,Y)}$.

Solution. For every $\varphi \in Y^*$:

$$\|T^t(\varphi)\|_{X^*} = \|\varphi \circ T\|_{X^*} \leq \|\varphi\|_{Y^*} \|T\|_{L(X,Y)},$$

which immediately implies $\|T^t\|_{L(Y^*,X^*)} \leq \|T\|_{L(X,Y)}$. To get equality: given $\alpha < \|T\|_{L(X,Y)}$ pick $u \in X$, of unit norm, such that $\|Tu\|_Y > \alpha$, and use the preceding result to find $\varphi \in Y^*$ of unit norm such that $\varphi(Tu) = \|Tu\|_Y$. Then:

$$\|T^t(\varphi)\|_{X^*} = \|\varphi \circ T\|_{X^*} = \sup\{|\varphi(T(v))| : \|v\|_X = 1\} \geq |\varphi(T(u))| = \|Tu\|_Y > \alpha;$$

and since α is an arbitrary real number smaller that $\|T\|_{L(X,Y)}$ we get $\|T^t\|_{L(Y^*,X^*)} \geq \|T\|_{L(X,Y)}$. □

1.4.9. *Separation of points and closed subspaces.*

. *Let X be a normed space, let V be a closed subspace of X, and let $a \in X \setminus V$. There exists a linear functional $f \in X^*$ which is 0 on V and nonzero on a.*

Proof. From exercise 1.4.3.4 the function $x \mapsto \mathrm{dist}(x, V)$ is a seminorm. Define u on $\mathbb{R}\,a$ by $u(a) = \mathrm{dist}(a, V)$, and then extending linearly, $u(\alpha\, a) = \alpha\, \mathrm{dist}(a, V)(= \mathrm{sgn}\,\alpha\, \mathrm{dist}(\alpha\, a, V) \leq \mathrm{dist}(\alpha\, a, V))$, for $\alpha \in \mathbb{R}$. Then $u(x) \leq \mathrm{dist}(x, V)$ for every $x \in \mathbb{R}\,a$. By the Hahn–Banach theorem we can extend u to an $\mathbb{R}-$linear functional, still called u, on all of X, such that $u(x) \leq \mathrm{dist}(x, V)$ for every $x \in X$. Setting $f(x) = u(x)$ if $\mathbb{K} = \mathbb{R}$, and $f(x) = u(x) - i\, u(ix)$ if $\mathbb{K} = \mathbb{C}$, the functional f is dominated by $\mathrm{dist}(\#, V)$ on X, hence it is 0 on V, and since $\mathrm{dist}(x, V) \leq \|x\|$ for every $x \in X$, f is in X^*. □

Exercise 1.4.9.1. Let X be a normed space, and let V be a closed linear subspace of X. Given $a \in X \setminus V$, prove that $V + \mathbb{K}\,a$ is also closed. Deduce from it that if W is a finite dimensional subspace of X, then $V + W$ is closed in X. In particular, every finite dimensional subspace of X is closed in X.

Solution. By the preceding proposition we have a continuous linear map $f : V + \mathbb{K}\,a \to \mathbb{K}$ such that $f(V) = \{0\}$ and $f(a) = 1$. Let $x_k + \lambda_k\, a$ be a sequence in $V + \mathbb{K}\,a$ converging to $x \in X$; we prove that $x \in V + \mathbb{K}a$. In fact, since f is continuous the sequence $f(x_k + \lambda_k\, a) = \lambda_k$ converges to $\lambda = f(x)$; then $x_k = (x_k + \lambda_k\, a) - \lambda_k\, a \in V$ converges to $x - \lambda\, a$; since V is closed we have $x - \lambda\, a \in V$, and hence $x \in V + \mathbb{K}\,a$, as claimed. All remaining statements are then easy, by induction on the dimension of W. □

1.4.10. *Embedding in the double dual.* There is a natural mapping $J : X \to X^{**}$ of a normed space X into its double dual, $J : X \to X^{**}$, defined as follows: $J(x) \in X^{**}$ acts on linear functionals $f \in X^*$ as f acts on x, i.e.:

$$\langle J(x), f\rangle := f(x)(= \langle f, x\rangle).$$

Linearity of J is trivial. Given $x \in X$, let's compute the norm of $J(x)$ as an element of X^{**}. We have

$$\|J(x)\|_{X^{**}} = \sup\{|\langle J(x), f\rangle| : \|f\|_{X^*} = 1\},$$

and since $\langle J(x), f\rangle = f(x)$, we have

$$\|J(x)\|_{X^{**}} = \sup\{|f(x)| : \|f\|_{X^*} = 1\}.$$

clearly $|f(x)| \leq \|f\|_{X^*} \|x\| \leq \|x\|$ for every f of unit norm in X^*; but as proved in 1.4.8 we have in X^* elements f of unit norm such that $f(x) = \|x\|$. Then $\sup\{|f(x)| : \|f\|_{X^*} = 1\} = \max\{|f(x)| : \|f\|_{X^*} = 1\} = \|x\|$. We have proved:

. *For every normed space X, the natural map $J : X \to X^{**}$ is an isometric embedding.*

Not necessarily the natural map J is also surjective, i.e. an isomorphism of X onto X^{**}. When this happens, the space X is said to be *reflexive*. Since X^{**} is necessarily complete, as every dual space is (1.1.5), a reflexive normed space must be complete, i.e a Banach space. It is immediate, from 1.1.7, that if $1 < p < \infty$ then for every Λ the space $\ell^p = \ell^p(\Lambda)$ is reflexive: the dual of ℓ^p is naturally identified to ℓ^q, and in the same way the dual of ℓ^q is ℓ^p, if $1/p + 1/q = 1$.

1.4.11. *Example of a non–reflexive Banach space.* Every finite dimensional normed space is of course reflexive: the normed dual X^* has the same dimension as X, so that $\dim X^{**} = \dim X^* = \dim X$, and J, being an injective linear map between spaces of the same finite dimension is also surjective. The spaces $c_0 = c_0(\Lambda)$ (1.1.3) are not reflexive, as soon as Λ is infinite: in 1.1.6, the dual of c_0 has been identified naturally with ℓ^1, and the dual of ℓ^1 with ℓ^∞. By the identifications made of the dual spaces it is clear that the natural embedding J of c_0 into its double dual $c_0^{**} \approx \ell^\infty$ is simply the inclusion of c_0 in its superspace ℓ^∞, much larger than c_0. The non reflexivity of c_0 is proved. Also the dual of ℓ^∞ is (much) larger than ℓ^1; we differ the proof to the following number.

1.4.12. In 1.4.2.1 we defined the sublinear functional $p : X \to \mathbb{R}$ by $p(x) = \limsup_{n\to\infty} x_n$, for every $x \in X = \ell^\infty(\mathbb{N}, \mathbb{R})$. In the closed subspace $c = c(\mathbb{N}, \mathbb{R})$ consisting of converging subsequences we have the continuous linear functional lim of norm 1, defined by $\lim x = \lim_{n\to\infty} x_n$ (the proof is trivial), and clearly $\lim x \le p(x)$ (actually $\lim x = p(x)$!) for every $x \in c$. By the Hahn–Banach theorem we can extend lim to a continuous functional li of norm 1 on all of X, thus attributing a sort of "limit" (sometimes called a "Banach limit") to every bounded sequence, in such a way that the limit of a sum is still the sum of the limits, and that for converging sequences this is the true limit. Notice also that we have, for every $x \in \ell^\infty$:

$$\liminf_{n\to\infty} x_n \le \operatorname{li} x \le \limsup_{n\to\infty} x_n.$$

Then li is a continuous linear functional on ℓ^∞ which is identically 0 on c_0 and then also on the subspace $\ell^1 \subseteq c_0$: it is an element of the dual of ℓ^∞ which is not in the natural embedded copy of ℓ^1 into its double dual, thus proving non–reflexivity of ℓ^1.

1.5. Compactness in normed spaces. It is well known that in a finite dimensional normed space a subset is compact if and only if it is closed and bounded. This is *never* true in infinite dimensional normed spaces, as we now prove.

. Riesz lemma *Let X be a normed space, and let V be a closed proper linear subspace of X. Then* $\sup\{\operatorname{dist}(u, V) : \|u\| = 1\} = 1$. *In other words, for every $\alpha < 1$ there exists u with $\|u\| = 1$ such that* $\operatorname{dist}(u, V) > \alpha$.

Proof. Recall (1.4.3.4) that $x \mapsto d(x, V)$ is a seminorm of X, and that $\operatorname{dist}(x, V) = \operatorname{dist}(x - v, V)$ for every $x \in X$ and every $v \in V$. Given $x \notin V$ there is $v \in V$ such that $\|x - v\| < \operatorname{dist}(x, V)/\alpha$. Then, since $\operatorname{dist}(x - v, V) = \operatorname{dist}(x, V)$ we have

$$\|x-v\| < \frac{\operatorname{dist}(x, V)}{\alpha} \iff \|x-v\| < \frac{\operatorname{dist}(x - v, V)}{\alpha} \iff 1 < \frac{1}{\alpha}\operatorname{dist}\left(\frac{x - v}{\|x - v\|}, V\right) \iff \alpha < \operatorname{dist}(u, V),$$

having set $u = (x - v)/\|x - v\|$. □

Remark. In 1.4.9 we proved that there is a nonzero continuous linear functional $f : X \to \mathbb{R}$ such that $f(x) \le \operatorname{dist}(x, V)$, for every $x \in X$. Then $g = f/\|f\|$ has norm 1 in X^*, so that for some unit vector $u \in X$ we have $g(u) > \alpha$; and g satisfies the same condition $g(x) \le \operatorname{dist}(x, V)$, for every $x \in X$, so that $\operatorname{dist}(u, V) \ge g(u) > \alpha$, thus giving another proof of Riesz lemma.

Corollary. *In an infinite dimensional normed space the closed unit ball is not sequentially compact*

Proof. Recall that finite dimensional subspaces of a normed space are closed (because they are complete in the induced norm, or else by 1.4.9.1). Fix α with $0 < \alpha < 1$. Given $u_0 \in X$ of norm 1, $\mathbb{K}\, x_0$ is a closed subspace which does not exhaust X, so that by Riesz lemma we can find x_1 of norm 1 such that $\operatorname{dist}(x_1, \mathbb{K}\, x_0) > \alpha$. It is clear how, by induction, we can define a sequence $x_0, x_1, \dots$ of unit vectors such that $\operatorname{dist}(x_m, \langle x_0, x_1, \dots, x_{m-1}\rangle) > \alpha$ for every $m \ge 1$ (the process never stops: $\langle x_0, x_1, \dots, x_{m-1}\rangle$ is a closed subspace of X, for every $m \ge 1$, proper because X is infinite dimensional). It is clear that the sequence $(x_n)_{n\in\mathbb{N}}$ cannot have converging subsequences: we have $\operatorname{dist}(x_n, x_m) > \alpha > 0$ if $n \ne m$. □

Exercise 1.5.0.1. Prove that in an infinitely dimensional normed space every compact subset has empty interior.

1.5.1. *Quotient of a Banach space.* If X is a normed space and V is a closed subspace, we have observed that the quotient space X/V is normed by $\|x + V\|_{X/V} = \operatorname{dist}(x, V)$ (the quotient norm). Riesz lemma (1.5) says that the natural projection $\pi : X \to X/V$ is of operator norm 1 (barring the case $V = X$, which makes X/V the trivial space). Then

. *If X is Banach space and V is a closed vector subspace then X/V is Banach space under the quotient norm.*

Proof. Left to the reader : recall that a normed space is Banach if and only if every normally summable sequence is summable; if $(x_n + V)_n$ is a sequence such that $\sum_{n=0}^\infty \operatorname{dist}(x_n, V) < \infty$, we can pick $y_n \in x_n + V$ such that also $\sum_{n=0}^\infty \|y_n\|_X < \infty \dots$. □

Exercise 1.5.1.1. Let X be a normed space and V be a closed linear subspace of V; put the quotient norm on X/V and let $\pi : X \to X/V$ be the quotient map. Prove that π is an open map and that if Y is another normed space and $g : X/V \to Y$ a map (not necessarily linear) then g is continuous iff $g \circ \pi$ is continuous.

Hint: prove that the image of the open unit ball of X is exactly the open unit ball of X/V; from this openness is trivial, and what remains follows from openness ....

1.5.2. *Totally bounded subsets of a metric space.* It is well–known that a compact metric space is also complete. As seen above, even in normed spaces adding boundedness to completeness is not enough to ensure compactness. The right condition is that of *total boundedness.*

Definition. In a metric space (X, dist) a subset $E \subseteq X$ is said to be *totally bounded* if for every $\varepsilon > 0$ there exists a finite subset $F \subseteq E$ such that $E \subseteq \bigcup_{x \in F} B(x, \varepsilon]$.

The same notion is obtained if we remove the requirement that the centers of the balls belong to E: that is, if for every $\varepsilon > 0$ there exists a finite subset $G \subseteq X$ such that $E \subseteq \bigcup_{x \in G} B(x, \varepsilon]$, then E is totally bounded: we leave the easy proof to the reader, as the equally easy proof of the following statements:

. *In a metric space:*
a subset is totally bounded if and only if for every $\varepsilon > 0$ is has a finite cover by sets of diameter less than ε;
every subset of a totally bounded subset is totally bounded;
The closure of a totally bounded subset is totally bounded.

In the previous number we have proved that the unit ball of an infinite dimensional normed space is never totally bounded: the sequence $\{x_0, \dots\}$ there constructed cannot be covered by a finite set of sets of diameter less than α.

Recall that in a metric space a subset is said to be *bounded* if it has finite diameter; trivially every totally bounded set is also bounded; the converse holds in finite dimensional normed spaces, never in infinite dimensional ones.

Exercise 1.5.2.1. Prove that a uniformly continuous function between metric spaces maps totally bounded sets into totally bounded sets, but that continuity is not enough for this.

Exercise 1.5.2.2. We say that a subset S of a metric space (X, dist) is *uniformly discrete* if there is $\alpha > 0$ such that $\mathrm{dist}(x, y) \geq \alpha$ for every pair $x \neq y$ of distinct points of S. Prove that X is totally bounded iff every uniformly discrete subset of X is finite.

1.5.3. *Converging subsequences: a reminder.* We recall some notions linked to limits of subsequences. Given a sequence $(x_n)_{n \in \mathbb{N}}$ in a topological space X, a point $c \in X$ is said to be a *cluster point* (by some: a limit point) of the sequence if:

$$\text{for every nbhd } U \text{ of } c \text{ in } X \text{ the set} \quad x^{\leftarrow}(U) = \{n \in \mathbb{N} : x_n \in U\} \quad \text{is infinite.}$$

Sometimes this is expressed by saying that the sequence is *frequently* in every nbhd U of c. We assume as known the following fact:

. *Let X be a metrizable space, and let $(x_n)_{n \in \mathbb{N}}$ be a sequence in X. A point $c \in X$ is the limit of a subsequence $(x_{\mu(k)})_{k \in \mathbb{N}}$ of the given sequence $(x_n)_{n \in \mathbb{N}}$ if and only if it is a cluster point of the sequence $(x_n)_{n \in \mathbb{N}}$.*

If, given the sequence $(x_n)_{n \in \mathbb{N}}$, for every $m \in \mathbb{N}$ we set $F_m = \mathrm{cl}(\{x_n : n \geq m\})$ (the closed $m-$tail of the sequence), then it is clear that $\bigcap_{m \in \mathbb{N}} F_m$ is exactly the set of all cluster points of the sequence.

Exercise 1.5.3.1. Let X be a T_1 topological space (meaning that finite subsets are closed sets), and let $(x_n)_{n \in \mathbb{N}}$ be a sequence in X. Prove that $c \in X$ is a cluster point of $(x_n)_{n \in \mathbb{N}}$ if and only if either c is a value of the sequence repeated infinitely many times, i.e. the set $\{n \in \mathbb{N} : x_n = c\}$ is infinite, or c is an accumulation point for the range $\{x_n : n \in \mathbb{N}\}$ of the sequence (of course the two condition are not mutually exclusive).

1.5.4. *Completeness and compactness in metric spaces.* In this number we prove that compactness in metric spaces is equivalent to completeness plus total boundedness.

. *For a metric space (X, dist) the following are equivalent:*

(i) *X is sequentially compact(:=every sequence in X has a subsequence converging to a point of X).*
(ii) *X is complete and totally bounded.*
(iii) *X is compact(:=every open cover of X has a finite subcover).*

Proof. (i) implies (ii): completeness follows immediately from the well–known fact that if a Cauchy sequence has a converging subsequence then the whole sequence converges. Next we prove that sequential compactness implies total boundedness: if the space is not totally bounded there is $\varepsilon > 0$ such that no finite set of balls of radius ε covers X; starting with any $x_0 \in X$ there is then $x_1 \in X \setminus B(x_0, \varepsilon]$, because X is not covered by $B(x_0, \varepsilon]$; and next we may pick $x_2 \in X \setminus (B(x_0, \varepsilon] \cup B(x_1, \varepsilon])$, etc. Inductively we construct a sequence $(x_n)_{n\in\mathbb{N}}$ such that if $m < n$ then $x_n \notin B(x_m, \varepsilon]$ so that $\mathrm{dist}(x_n, x_m) > \varepsilon$. Such a sequence clearly cannot have converging subsequences, thus contradicting sequential compactness of X.

(ii) implies (iii). Assume that X is not compact, so that there exists an open cover $(A_\lambda)_{\lambda\in\Lambda}$ of X such that for no finite subset $M \subseteq \Lambda$ we have $X = \bigcup_{\lambda\in M} A_\lambda$. We get a contradiction with (ii). Since X is totally bounded there exists a finite subset E_0 of X such that $\bigcup_{x\in E_0} B(x, 1/2^0] = X$. Then at least one of these balls is not covered by any finite subfamily of $(A_\lambda)_{\lambda\in\Lambda}$; call x_0 is center. Since $B(x_0, 1]$ is also totally bounded (as every subset of X is), there is a finite subset E_1 of $B(x_0, 1]$ such that $B(x_0, 1] \subseteq \bigcup_{x\in E_1} B(x, 1/2]$. Hence some ball $B(x, 1/2]$, for some $x \in E_1$, is not covered by a finite subfamily of the cover $(A_\lambda)_{\lambda\in\Lambda}$; call x_1 is center. It is now clear how to proceed inductively: assume that $x_0, \ldots, x_{m-1}$ have been defined in such a way that for $k = 1, \ldots, m-1$ we have $x_k \in B(x_{k-1}, 1/2^{k-1}]$, and for $k = 0, \ldots, m-1$ we have that $B(x_k, 1/2^k]$ is not covered by a finite subfamily of the cover $(A_\lambda)_{\lambda\in\Lambda}$. By total boundedness of $B(x_{m-1}, 1/2^{m-1}]$ there is a finite subset E_m of this ball such that $B(x_{m-1}, 1/2^{m-1}] \subseteq_{x\in E_m} B(x, 1/2^m]$; then some ball $B(x, 1/2^m]$ for some $x \in E_m$, is not covered by a finite subfamily of the given open cover, call its center x_m. The sequence $(x_n)_{n\in\mathbb{N}}$ is a Cauchy sequence, being of bounded variation (we have $\mathrm{dist}(x_n, x_{n+1}) \le 1/2^n$); hence it has a limit $c \in X$, since X is complete. Then $c \in A_\mu$ for some $\mu \in \Lambda$; since A_μ is open, there is $r > 0$ such that $B(c, r] \subseteq A_\mu$. For m large enough we have $\mathrm{dist}(c, x_m) \le r/2$ and also $1/2^m \le r/2$. Then $B(x_m, 1/2^m] \subseteq B(c, r]$ (in fact, if $x \in B(x_m, 1/2^m]$ we have $\mathrm{dist}(c, x) \le \mathrm{dist}(c, x_m) + \mathrm{dist}(x_m, x) \le 1/2^m + 1/2^m \le r$); and since $B(c, r] \subseteq A_\mu$ we also have $B(x_m, 1/2^m] \subseteq A_\mu$, contradicting the fact that $B(x_m, 1/2^m]$ cannot be covered by a finite subfamily of $(A_\lambda)_{\lambda\in\Lambda}$.

(iii) implies (i). Let $(x_n)_{n\in\mathbb{N}}$ be a sequence in X, and let $F_n = \mathrm{cl}(\{x_k : k \ge n\})$, for every $n \in \mathbb{N}$. The set $C = \bigcap_{n\in\mathbb{N}} F_n$ is exactly the set of *cluster points* of the sequence $(x_n)_{n\in\mathbb{N}}$, i.e. the set of limits of converging subsequences, as observed in 1.5.3; we have to prove that it is non–empty. Otherwise, $A_n = X \setminus F_n$ is an open cover of X, which has then a finite subcover; since $A_0 \subseteq A_1 \subseteq \ldots$, this means that $A_m = X$ for some $m \in \mathbb{N}$, equivalently that $F_m = \emptyset$ for some $m \in \mathbb{N}$; this is plainly impossible. □

1.5.5. *Compact subspaces of a metric space.*

. *A subspace of a metric space is compact if and only if it is totally bounded and complete in the induced metric. A subspace of a complete metric space is compact if and only if it is closed and totally bounded.*

Proof. Left to the reader: it is an easy corollary of previous results. □

1.5.6. *An example.* Let Λ be a non empty set, and let $p \ge 1$ be a real number. Fix $a \in \ell^p = \ell^p(\Lambda)$, and let

$$K_a = \{x \in \ell^p : |x(\lambda)| \le |a(\lambda)| \text{ for every } \lambda \in \Lambda\}$$

be the set of functions dominated by a. Then K_a is a compact subset of ℓ^p. In fact it is closed (it is pointwise closed, and convergence in ℓ^p implies pointwise convergence) and totally bounded. For this last assertion, given $\varepsilon > 0$, find a finite subset $F(\varepsilon)$ of Λ such that $\sum_{\Lambda\setminus F(\varepsilon)} |a|^p \le \varepsilon^p$; map ℓ^p to $\mathbb{K}^{F(\varepsilon)}$ by restriction, $\rho(x) = x_{|F(\varepsilon)}$; then $\rho(K_a)$ is a bounded subset of the finite dimensional space $\mathbb{K}^{F(\varepsilon)}$. Then $\rho(K_a)$ is also totally bounded and hence there exist $a_1, \ldots, a_m \in \rho(K_a)$ such that for each $x \in K_a$ there is $j \in \{1, \ldots, m\}$ such that $\sum_{F(\varepsilon)} |a_j - \rho(x)|^p \le \varepsilon^p$. We may consider every a_j also as an element of c_{00}, simply by extending it to be identically zero outside of $F(\varepsilon)$. Then for every $x \in K_a$, if j is as before we have

$$\sum_{\Lambda} |a_j - x|^p = \sum_{F(\varepsilon)} |a_j - x|^p + \sum_{\Lambda\setminus F(\varepsilon)} |a_j - x|^p \le \varepsilon^p + \sum_{\Lambda\setminus F(\varepsilon)} |x|^p \le \varepsilon^p + \sum_{\Lambda\setminus F(\varepsilon)} |a|^p \le 2\,\varepsilon^p,$$

so that K_a is totally bounded. The most common version is with $\Lambda = \mathbb{N}$, $a(n) = 1/(n+1)$ and $p = 2$; in this case $K_a = \{x : \mathbb{N} \to \mathbb{K} : |x(n)| \le 1/(n+1)\}$ is called *Hilbert cube.* If $a \in c_0$ the set K_a is still compact. But if $a \in \ell^\infty \setminus c_0$ the set K_a is not a compact subset of ℓ^∞. It is closed in ℓ^∞ but not totally bounded. In fact there is $\varepsilon > 0$ such that the set $S = \{\lambda \in \Lambda : |a(\lambda)| > \varepsilon\}$ is infinite; the set $\{\varepsilon\, e_\lambda : \lambda \in S\}$ is then a subset of K_a, but if $\lambda, \mu \in S$ and $\lambda \ne \mu$ then $\|\varepsilon\, e_\lambda - \varepsilon\, e_\mu\|_\infty = \varepsilon$, and K_a cannot be covered by finitely many balls of radius smaller that $\varepsilon/2$.

EXERCISE 1.5.6.1. Verify by direct computation that K_a, as above defined, has empty interior in any ℓ^p or in c_0. For $a \in \ell^\infty$, prove that K_a has non–empty interior if and only if $\inf\{|a(\lambda)| : \lambda \in \Lambda\} > 0$ (to save a little work observe first that every K_a is convex and symmetric, so that if $\mathrm{int}(K_a)$ is non empty then $0 \in \mathrm{int}(K_a)$, see 2.2.4...).

1.6. Compactness in spaces of continuous functions.

1.6.1. *Equicontinuity.* Let X be a topological space and Y a normed space; let us recall one of the possible definitions of continuity of a function $f : X \to Y$ at a point $x \in X$: for every $\varepsilon > 0$ there exists a neighborhood $U(\varepsilon)$ of x such that if $\xi \in U(\varepsilon)$ then $|f(\xi) - f(x)| \le \varepsilon$. Here $U(\varepsilon)$ depends on ε but also on f, it is really a $U(f, \varepsilon)$ if various functions f are considered; clearly different functions f will have in general different $U(f, \varepsilon)$ for the same ε; equicontinuity is the requirement that $U(f, \varepsilon)$ can be chosen to depend on ε only. That is:

DEFINITIONS. Given a set $F \subseteq C(X, Y)$, with X topological and Y normed and $x \in X$, we say that F is *equicontinuous at* x if for every $\varepsilon > 0$ there is a neighborhood $U(\varepsilon)$ of x in X such that

$$\|f(\xi) - f(x)\|_Y \le \varepsilon \qquad \text{for every } \xi \in U(\varepsilon), \text{ and every } f \in F.$$

A set $F \subseteq C(X, Y)$ is said to be *equicontinuous* in X, or simply equicontinuous, if it is equicontinuous at every $x \in X$.

Of course a finite set F of continuous functions is always equicontinuous; but it is easy to give examples of non-equicontinuous sets; e.g. if $X = [-1, 1]$ and $Y = \mathbb{R}$ the set of functions $f_\alpha(x) = |x|^\alpha$, with $0 < \alpha < 1$, is not equicontinuous at 0: given $0 < \varepsilon$ observe that

$$\{\xi \in X : |f_\alpha(\xi) - f_\alpha(0)| \le \varepsilon\} = \{\xi : |\xi| \le \varepsilon^{1/\alpha}\} = U(f_\alpha, \varepsilon),$$

and since $\lim_{\alpha \to 0^+} \varepsilon^{1/\alpha} = 0$ if $0 < \varepsilon < 1$ we have $\bigcap_{0<\alpha<1} U(f_\alpha, \varepsilon) = \{0\}$ so that there is no common neighborhood of 0 where all functions f_α are dominated by ε. An analogous easy computation (exercise) shows that however $\{f_\alpha : 0 < \alpha < 1\}$ is equicontinuous at every non zero $x \in X$. And $\{f_\alpha : \alpha > 1\}$ is equicontinuous at every $x \in X \setminus \{-1, 1\}$.

EXERCISE 1.6.1.1. Let X be a topological space, Y a normed space, $(f_n)_{n\in\mathbb{N}}$ a sequence in $C(X, Y)$; assume that f_n converges pointwise on X to a function $f : X \to Y$. If $(f_n)_{n\in\mathbb{N}}$ is equicontinuous at some $c \in X$, then the limit function is continuous at c.

Solution. Given $\varepsilon > 0$ find a nbhd $U(c)$ of c in X such that

$$(*) \qquad \|f_n(x) - f_n(c)\| \le \varepsilon \quad \text{for every } x \in U(c),$$

and find $n(\varepsilon) \in \mathbb{N}$ such that $\|f(c) - f_n(c)\| \le \varepsilon$ for $n \ge n(\varepsilon)$. Passing to the limit for $n \to \infty$ in the inequality (*) we find $\|f(x) - f(c)\| \le \varepsilon$ for $x \in U(c)$, proving continuity of f at c. □

EXERCISE 1.6.1.2. ⊙⊙ Let X be a topological space and let $(f_n)_{n\in\mathbb{N}}$ be an equicontinuous sequence of $C(X, \mathbb{K})$; assume that there is a dense subset D of X such that $\lim_{n\to\infty} f_n(x)$ exists for every $x \in D$ (in other words the sequence converges pointwise on D). Then the sequence converges pointwise to a continuous function f on all of X. If X is compact, the convergence is also uniform on X.

Solution. We call $f(c)$ the limit $\lim_n f_n(c)$, for every $c \in D$. Given $x \in X$ and $\varepsilon > 0$ find a neighborhood $U(x)$ of x such that $|f_n(\xi) - f_n(x)| \le \varepsilon$ for every $\xi \in U(x)$ and every $n \in \mathbb{N}$; since D is dense in X there is $c \in D \cap U(x)$, and we find $n(\varepsilon) \in \mathbb{N}$ such that $|f(c) - f_n(c)| \le \varepsilon$ for $n \ge n(\varepsilon)$. Then for $n, m \ge n(\varepsilon)$:

$$(*) \qquad \begin{aligned} |f_m(x) - f_n(x)| &= |f_m(x) - f_m(c) + f_m(c) - f_n(c) + f_n(c) - f_n(x)| \le \\ &\le |f_m(x) - f_m(c)| + |f_m(c) - f_n(c)| + |f_n(c) - f_n(x)| \le 4\varepsilon \end{aligned}$$

(the first and third term are less than ε for equicontinuity, the middle term is less than 2ε because $|f(c)-f_n(c)| \le \varepsilon$ for $n \ge n(\varepsilon)$). Then $f_n(x)$ is a Cauchy sequence for every $x \in X$, so that f_n converges pointwise on all of X to a function $f : X \to \mathbb{K}$, which is continuous at every point by the preceding exercise. Uniform convergence for X compact is easy, but easier using Ascoli's theorem, see exercise 1.6.2.2. □

1.6.2. *Compactness in $C(X, \mathbb{K})$ with X compact.* We now give a very neat characterization of total boundedness in $C(X, \mathbb{K})$ with X compact.

. THEOREM OF ASCOLI–ARZELÀ. *Assume X compact; a subset F of $C(X) = C(X, \mathbb{K})$ is totally bounded (in the uniform norm $\|\cdot\|_u$) if and only if it is bounded and equicontinuous. Consequently a subset S of $C(X, \mathbb{K})$ is compact iff it is closed, bounded, and equicontinuous.*

Proof. *Necessity* As observed, totally bounded sets are also bounded. Equicontinuity: given $\varepsilon > 0$, by total boundedness we find a finite subset $\{f_1, \dots, f_m\} \subseteq F$ such that $F \subseteq \bigcup_{k=1}^m B(f_k, \varepsilon]$. Given $x \in X$, the finite set $\{f_1, \dots, f_m\}$ is equicontinuous at x, so that there is a nbhd $U(\varepsilon)$ of x in X for which we have $|f_j(\xi) - f_k(x)| \leq \varepsilon$ for $k = 1, \dots, m$. Given $f \in F$ there is $k \in \{1, \dots, m\}$ such that $\|f - f_k\|_\infty \leq \varepsilon$; then, for $\xi \in U(\varepsilon)$:

$$\begin{aligned}|f(\xi) - f(x)| =& |f(\xi) - f_k(\xi) + f_k(\xi) - f_k(x) + f_k(x) - f(x)| \leq \\ &\leq |f(\xi) - f_k(\xi)| + |f_k(\xi) - f_k(x)| + |f_k(x) - f(x)| \leq \\ &\leq \|f - f_k\|_u + |f_k(\xi) - f_k(x)| + \|f_k - f\|_u \leq 3\varepsilon,\end{aligned}$$

thus proving equicontinuity of F at x.

Sufficiency Since F is equicontinuous, given $\varepsilon > 0$ and $x \in X$ there is a neighborhood $U(x)$ of x in X such that $|f(\xi) - f(x)| \leq \varepsilon$ for every $\xi \in U(x)$ and every $f \in F$. By compactness of X we find $x_1, \dots, x_p \in X$ such that $X = \bigcup_{k=1}^p U(x_k)$. Given the p−tuple $\{x_1, \dots, x_p\}$ we consider the linear map $\rho : f \mapsto (f(x_1), \dots f(x_p))$ of $C(X)$ into $\mathbb{K}^p$, where $\mathbb{K}^p$ is given the product norm $\|u\|_\infty = \max\{|u_k| : k = 1, \dots, p\}$; clearly ρ has operator norm 1, hence it maps the bounded set F into a bounded set $\rho(F)$. Being bounded in a finite–dimensional space, $\rho(F)$ is also totally bounded, in other words there exist $f_1, \dots, f_m \in F$ such that for every $f \in F$ there exists $k \in \{1, \dots, m\}$ such that $\|\rho(f) - \rho(f_k)\|_{\mathbb{K}^p} \leq \varepsilon$ (equivalently $|f(x_j) - f_k(x_j)| \leq \varepsilon$ for every $j \in \{1, \dots, p\}$). We prove that $F \subseteq \bigcup_{k=1}^m B(f_k, 3\varepsilon]$: given $f \in F$, choose f_k as above; and given $\xi \in X$ there is $j \in \{1, \dots, p\}$ such that $\xi \in U(x_j)$ so that

$$\begin{aligned}|f(\xi) - f_k(\xi)| =& |f(\xi) - f(x_j) + f(x_j) - f_k(x_j) + f_k(x_j) - f_k(\xi)| \leq \\ &\leq |f(\xi) - f(x_j)| + |f(x_j) - f_k(x_j)| + |f_k(x_j) - f_k(\xi)| \leq 3\varepsilon\end{aligned}$$

(the first and the third term are smaller than ε for equicontinuity, the middle term because of the choice of k). Thus F is totally bounded. □

Exercise 1.6.2.1. Let F be a subset of $C = C([0,1], \mathbb{R})$ such that $f(0) = 0$ for every $f \in F$, and such that every f is everywhere differentiable in $[0,1]$, with $\|f'\|_u \leq 1$ for every $f \in F$. Prove that F is totally bounded.

Exercise 1.6.2.2. Let X be a compact topological space, and let $(f_n)_{n\in\mathbb{N}}$ be a sequence in $C(X, \mathbb{K})$. Prove that the following are equivalent:

(i) $(f_n)_{n\in\mathbb{N}}$ is uniformly convergent on X.
(ii) $(f_n)_{n\in\mathbb{N}}$ is equicontinuous and pointwise convergent in X.

(recall that the terms of a converging sequence, together with the limit, form a compact set …)

Exercise 1.6.2.3. (*Continuous convergence*) Let X be a compact space; assume that $f_n \in C(X, \mathbb{K})$ converges uniformly to $f \in C(X, \mathbb{K})$. Then for every sequence $c_n \in X$, converging to $c \in X$ we have that $f_n(c_n)$ converges to $f(c)$ (use equicontinuity of the sequence f_n).

⊙⊙The converse, more difficult, also holds: if X is metrizable compact, and $f_n, f \in C(X, \mathbb{K})$ are such that $f_n(c_n)$ converges to $f(c)$, for every sequence $c_n \in X$ converging to $c \in X$, then f_n converges uniformly to f on X.

Exercise 1.6.2.4. Let $K : [0,1] \times [0,1] \to \mathbb{K}$ be a continuous function.

(i) Prove that the formula

$$Tf(x) = \int_0^1 K(x,y)\, f(y)\, dy$$

defines a (linear) bounded operator T from $L^1 = L^1([0,1])$ into $C = C([0,1])$ (this last with sup–norm).

(ii) Prove that, if B_1 is the closed unit ball of L^1, then $T(B_1)$ is totally bounded in C (hint: use uniform continuity of K on the square $[0,1] \times [0,1]$), so that $\mathrm{cl}_{C([0,1])}(T(B_1))$ is compact.

(iii) Given any p with $1 \leq p, q \leq \infty$, is it true that T is also a bounded operator of L^p into C?

(iv) With p as in (iii), is it true that T can be interpreted also as a bounded operator of L^p into L^q? and is $\mathrm{cl}_{L^q}(T(B_p))$ compact in L^q, if B_p is the closed unit ball of L^p?

1.6.3. *Compactness in $C(X,Y)$ with X compact and Y a Banach space.* The preceding theorem holds with the same statement for $C(X,Y)$, if X is compact and Y a finite dimensional normed space instead of $Y = \mathbb{K}$. If Y is a normed space, not necessarily finite dimensional, then boundedness of F must be replaced by pointwise total boundedness of F, i.e. :

. *Let X be a compact topological space and let Y be a normed space; equip $C(X,Y)$ with the sup–norm. Then a subset F of $C(X,Y)$ is totally bounded if and only if it is equicontinuous, and for every $x \in X$ the set $F(x) = \{f(x) : f \in F\}$ (the values of members of F at x) is totally bounded in Y.*

Proof. We omit the proof, which is anyway easy to do by imitating the given one for $Y = \mathbb{K}$; recall also that a finite product of totally bounded spaces is totally bounded. □

The following exercise generalizes the previous one on differentiable functions in $C([0,1])$:

Exercise 1.6.3.1. Let X be a metrizable compact space, let Y be a finite dimensional space, and let $F \subseteq C(X,Y)$ be a set of functions such that:

- for some $a \in X$ the set $F(a) = \{f(a) : f \in F\}$ is bounded;
- all functions in F are Lipschitz continuous, with a common Lipschitz constant $\lambda > 0$

Prove that F is totally bounded in $C(X,Y)$.

1.7. Baire's theorem. The intersection of two dense subsets of a topological space is not in general dense in the space: it may even be empty, e.g the sets $\mathbb{Q}$ and $\sqrt{2}+\mathbb{Q}$ are both dense in $\mathbb{R}$ but with empty intersection. However, given two dense subsets A and B of a topological space, if one of them is *open* then $A \cap B$ is still dense in X: if U is open and non–empty then $U \cap A$ is open if A is open, and non–empty if A is dense, so that $U \cap A \cap B$ is non–empty if B is dense. By induction it is clear that the intersection of a finite family of dense open subsets of a topological space X is still (open and) dense in X. A countable intersection of open dense subsets may well be no longer dense in the space: e.g. ordering the rationals in a sequence $(q_n)_{n\in\mathbb{N}}$, the sets $A_n = \mathbb{Q} \smallsetminus \{q_0, \dots, q_n\}$ are all open and dense, but the intersection of them all is empty. In a complete metric space this cannot happen: this fact is extremely useful.

1.7.1. *Cantor formulation of completeness.* We first prove a simple characterization of complete metric spaces, due to Georg Cantor, one of the founding fathers of the modern set theory:

. Cantor theorem *A metric space is complete if and only if every decreasing sequence of non–empty closed subsets with arbitrarily small diameters has a non–empty intersection.*

Explicitly: assume $F_0 \supseteq F_1 \supseteq \dots$ is a decreasing sequence of non–empty closed subsets of the metric space (X, dist), with $\inf\{\mathrm{diam}(F_n) : n \in \mathbb{N}\} = 0$ (equivalently, $\lim_n \mathrm{diam}(F_n) = 0$);

then (X, dist) is complete iff for every such sequence $\bigcap_{n\in\mathbb{N}} F_n \neq \emptyset$.

Proof. In fact, pick a point $x_n \in F_n$, for every $n \in \mathbb{N}$. Then the sequence $(x_n)_{n\in\mathbb{N}}$ is Cauchy: given $\varepsilon > 0$, pick $n(\varepsilon) \in \mathbb{N}$ such that $\mathrm{diam}(F_{n(\varepsilon)}) \le \varepsilon$, which is possible by the condition $\inf\{\mathrm{diam}(F_n) : n \in \mathbb{N}\} = 0$: if $m, n \ge n(\varepsilon)$ then $x_m \in F_m \subseteq F_{n(\varepsilon)}$ and $x_n \in F_n \subseteq F_{n(\varepsilon)}$ so that $x_m, x_n \in F_{n(\varepsilon)}$ and then $\mathrm{dist}(x_m, x_n) \le \mathrm{diam}(F_{n(\varepsilon)}) \le \varepsilon$. If the space is complete, the sequence has a limit c, and clearly c is the (unique) point of the intersection: for every $m \in \mathbb{N}$, we have $x_n \in F_n \subseteq F_m$ if $n \ge m$, so that $c \in F_m$, since F_m is closed in X. And this holds for every $m \in \mathbb{N}$, so that $c \in \bigcap_{m\in\mathbb{N}} F_m$.

The easy converse, that this condition implies completeness, is left to the reader with the following hint: given a Cauchy sequence $(x_n)_{n\in\mathbb{N}}$ take $F_n = \mathrm{cl}(\{x_k : k \ge n\})$.

Remark. Of course under the above hypotheses the intersection must be a singleton: if G is this intersection, then $\mathrm{diam}(G) = 0$, and in a metric space this is possible iff G is either empty or consists of a single point.

□

Exercise 1.7.1.1. A set $\mathcal{B}$ of non–empty subsets of a set X is said to be a *filterbase* if given $U, V \in \mathcal{B}$ there is $W \in \mathcal{B}$ such that $W \subseteq U \cap V$. Prove that a metric space (X, dist) is complete if and only if every filterbase of closed sets of X which contains sets of arbitrarily small diameter has non–empty intersection.

1.7.2. *Baire's theorem.* We say that a topological space is *completely metrizable* if its topology is the topology of a metric which makes the space a complete metric space. Every complete metric space is then completely metrizable as a topological space.

. Theorem of Baire *In a completely metrizable space the intersection of a sequence of open dense subsets of the space is still dense in the space.*

Proof. Let X be the space, and let $(A_n)_{n\in\mathbb{N}}$ be a sequence of open dense subsets of X. Given a nonempty open set U we construct a decreasing sequence $B(x_0, r_0] \supseteq B(x_1, r_1] \supseteq \dots$ such that for every n we have $B(x_n, r_n] \subseteq U \cap A_0 \cap A_1 \cap \dots \cap A_n$, and such that $0 < r_n \le 1/(n+1)$. By Cantor theorem the intersection of all these balls is nonempty, and is clearly contained in $U \cap \left(\bigcap_{n\in\mathbb{N}} A_n\right)$, thus proving that this intersection is non empty; by arbitrariness of the nonempty open set U this proves density of the intersection $\bigcap_{n\in\mathbb{N}} A_n$. Since $U \cap A_0$ is nonempty, we can find $x_0 \in U \cap A_0$, and since it is also open we find r_0, with $0 < r_0 \le 1$, such that $B(x_0, r_0] \subseteq U \cap A_0$. Since $B(x_0, r_0[\cap U \cap A_0 \cap A_1$ is nonempty we can find x_1 in it, and since it is also open we find r_1, with $0 < r_1 \le 1/2$ such that $B(x_1, r_1] \subseteq B(x_0, 1[\cap U \cap A_0 \cap A_1$. It is plain that induction never stops and leads to the required conclusion. □

Clearly the intersection $\bigcap_{n\in\mathbb{N}} A_n$ will not be open, in general! the most important fact is however that this intersection is nonempty.

1.7.3. *Meager sets.* There is a dual statement which is also useful: the complement of an open dense set is a closed set with empty interior, and the complement of a dense set is a set with empty interior. Then the theorem of Baire has the following equivalent formulation:

. *In a completely metrizable space a countable union of closed subsets with empty interior has empty interior.*

Call a subset of a topological space *nowhere dense* if its closure has empty interior, and call *meager* any subset which is a countable union of nowhere dense subsets. Then we also have

. *In a completely metrizable space a meager set has empty interior.*

And here, also, in general the most important fact is that the whole space cannot be meager.

1.7.4. *An application: nowhere differentiable functions.* To illustrate the power of Baire's theorem we prove that in the space $C([a, b])$ of continuous functions (with sup–norm) the set of functions which admit a derivative, even at only one point of $[a, b]$, is meager. Given $k > 0$, let

$$E(k) = \{f \in C([a, b]) : \text{ there is } c = c_f \in [a, b] \text{ such that } |f(x) - f(c)| \le k|x - c| \quad \text{for every } x \in [a, b].\}$$

It is clear that if $f \in C([a, b])$ has a derivative, even at only one point $c \in [a, b]$, then $f = \bigcup_{k\in\mathbb{N}} E(k)$. Now $E(k)$ is a closed subset of $C([a, b])$: in fact, if if f_n is a sequence in $E(k)$ which converges uniformly to f, for every $n \in \mathbb{N}$ there is $c_n \in [a, b]$ for which $|f_n(x) - f_n(c_n)| \le k|x - c_n|$; by considering a subsequence we may assume that c_n converges to $c \in [a, b]$, and passing to the limit in the preceding formula we have

$$|f(x) - f(c)| \le k|x - c| \quad \text{for every } x \in [a, b]$$

(we have $f(c) = \lim_n f_n(c_n)$, continuous convergence, see 2.2.2). But $E(k)$ has empty interior. Let in fact $g_{c,\delta}(x) = \delta \sin((x - c)/\delta^2)$ and notice that $\|g_{c,\delta}\|_u \le \delta$, so that every uniform nbhd of f contains functions as $f + g_{c,\delta}$ for δ small enough. If $f \in E(k)$ consider the difference quotient, for $c = c_f$:

$$\frac{f(x) - f(c)}{x - c} + \delta\,\frac{\sin((x - c)/\delta^2)}{x - c};$$

since $f \in E(k)$ the first term is dominated by k, but as $x \to c$ the second term tends to $1/\delta$, as large as we please for δ small.

We have proved: the set of $f \in C([a, b])$ which admit a derivative at some point of $[a, b]$ is contained in $\bigcup_{n\ge1} E(n)$, a countable union of closed sets with empty interior.

Exercise 1.7.4.1. Let X be a completely metrizable space. Prove that if $(F_n)_{n\in\mathbb{N}}$ is a sequence of closed subsets of X such that $\bigcup_{n=0}^{\infty} F_n = X$, then $A = \bigcup_{n=0}^{\infty} \mathrm{int}_X(F_n)$ is an open dense subset of X.

Solution. If $G_n = F_n \setminus \mathrm{int}_X(F_n)$, clearly G_n is nowhere dense (it is a closed set with empty interior) so that $G = \bigcup_{n\in\mathbb{N}} G_n$ has empty interior; since $X = G \cup A$, A is dense in X (if U is nonempty open in X, then $U \subseteq X = G \cup A$, but U cannot be contained in G, hence it meets A). □

Exercise 1.7.4.2. Prove that a countable metric space without isolated points is not complete. Find a compact countably infinite metric space.

Exercise 1.7.4.3. Prove that the open interval $]-1, 1[$, although non–complete in the usual metric, is completely metrizable, and find a complete admissible metric for it. More generally, any open subset A of a Banach space X is completely metrizable (although never complete in the induced metric, barring the extreme cases $A = X$ or $A = \emptyset$). Hint: if $\rho(x) = \mathrm{dist}(x, X \setminus A)$, consider the metric $d : A \times A \to [0, +\infty[$ given by

$$d(x, y) = \|x - y\| + \left| \frac{1}{\rho(x)} - \frac{1}{\rho(y)} \right|.$$

1.7.5. *Equicontinuity and boundedness in spaces of linear operators.* Equicontinuity is equivalent to boundedness in normed spaces of linear maps:

. *Let X and Y be normed spaces, let $L(X, Y) = L_{\mathbb{K}}(X, Y)$ be the space of continuous linear operators from X to Y, with the operator norm, and let $\mathcal{T}$ be a subset of $L(X, Y)$. The following are equivalent:*

(i) *$\mathcal{T}$ is equicontinuous;*
(ii) *$\mathcal{T}$ is equicontinuous at 0;*
(iii) *$\mathcal{T}$ is uniformly bounded on some neighborhood of 0.*
(iv) *$\mathcal{T}$ is bounded in the normed space $L(X, Y)$, i.e. there exists $M > 0$ such that $\|T\| \le M$, for every $T \in \mathcal{T}$.*

Proof. (i) implies (ii): trivial.

(ii) implies (iii): equicontinuity at 0 implies that there is $\delta > 0$ such that $T(\delta B_X) \subseteq B_Y$ for every $T \in \mathcal{T}$, equivalently $\|Tx\|_Y \le 1$ for every $x \in \delta B_X$ and every $T \in \mathcal{T}$.

(iii) implies (iv): if there exist $\delta, M > 0$ such that $\|Tx\|_Y \le M$ for every $x \in \delta B_X$ and every $T \in \mathcal{T}$, then $T(\delta B_X) \subseteq M B_Y$, equivalently $T(B_X) \subseteq (M/\delta) B_Y$, equivalently $\|T\|_{L(X,Y)} \le M/\delta$ for every $T \in \mathcal{T}$.

(iv) implies (i): equicontinuity is immediately implied by the fact that all members of $\mathcal{T}$ have M as Lipschitz constant. □

1.7.6. *Convex neighborhoods of* 0. In 2.2.4 we shall prove the following:

. *A symmetric convex subset with non–empty interior in a normed space X is a neighborhood of 0 in X.*

We use this result for the next proof.

1.7.7. *The uniform boundedness theorem of Banach and Steinhaus.* If the domain space X is a Banach space then every pointwise bounded subset $\mathcal{T} \subseteq L(X, Y)$ is bounded in norm, that is, uniformly bounded on the unit ball of X(equivalently, on all bounded subsets of X). This explains why the following is called "uniform boundedness theorem":

. Theorem of Banach–Steinhaus *Let X be a Banach space, Y a normed space, and let $\mathcal{T}$ be a subset of $L(X, Y)$. Assume that there is a non–meager subset E of X on which $\mathcal{T}$ is pointwise bounded, i.e. that for every $x \in E$ the set*

$$\mathcal{T}(x) = \{Tx : T \in \mathcal{T}\} \quad \textit{is bounded in } Y.$$

Then $\mathcal{T}$ is bounded in $L(X, Y)$, i.e. there is $M > 0$ such that $\|T\| \le M$ for every $T \in \mathcal{T}$ (equivalently, $\mathcal{T}$ is equicontinuous).

Proof. For every $n \in \mathbb{N}$ consider the set

$$K_n = \{x \in X : \|Tx\|_Y \le n, \text{ for all } T \in \mathcal{T}\} = \bigcap_{T \in \mathcal{T}} T^{\leftarrow}(n B_Y).$$

Clearly K_n is closed, convex and symmetric in X (it is the intersection of inverse images by continuous linear maps of the closed, convex and symmetric ball $n B_Y$). Pointwise boundedness of $\mathcal{T}$ on E implies that $E \subseteq \bigcup_{n \in \mathbb{N}} K_n$. Since X is complete, and E is by hypothesis non–meager, for some n the interior of K_n is non–empty; then K_n is neighborhood of the origin, as recalled above , on which all maps $T \in \mathcal{T}$ are bounded by n.

□

1.7.8. *Applications of Banach–Steinhaus.* The first is a very useful corollary:

. *Let X be a Banach space, Y a normed space, and let $(T_n)_{n \in \mathbb{N}}$ be a sequence of $L(X, Y)$ which converges pointwise to a mapping $T : X \to Y$. Then T belongs to $L(X, Y)$, and $\{T_n : n \in \mathbb{N}\}$ is equicontinuous.*

Proof. Since the sequence $T_n x$ is convergent, it is bounded, so that $\{T_n : n \in \mathbb{N}\}$ is equicontinuous, and uniformly bounded, say by $M > 0$, on the unit ball of X. Clearly T is linear; since $\|T_n x\|_Y \le M$ for every $x \in B_X$, we also have, passing to the limit, $\|Tx\|_Y \le M$ for every $x \in B_X$, so that $\|T\| \le M$, and T is continuous. □

Remark. It is *not* true, in general, that in the above hypotheses T_n converges to T in the norm of $L(X, Y)$. For instance, in the space $\ell^p = \ell^p(\mathbb{N})$ with $1 \le p < \infty$, the evaluation functionals $e_n^* : \ell^p \to \mathbb{K}$ (defined by $e_n^*(x) = x(n)$) are continuous, and the sequence e_n^* converges pointwise to the zero functional, since $\ell^p \subseteq c_0$ if $p < \infty$; but being all e_n^* of norm 1 in $(\ell^p)^*$, the sequence does not converge to 0 in the norm of $(\ell^p)^*$.

It can only be proved that $\|T\|_{L(X,Y)} \le \liminf_{n\to\infty} \|T_n\|_{L(X,Y)}$: in fact, for every $x \in X$ and every $n \in \mathbb{N}$ we have

$$\|T_n(x)\| \le \|T_n\| \, \|x\| \quad \text{taking } \liminf \text{ on both sides} \quad \liminf_{n\to\infty} \|T_n x\| \le \liminf_{n\to\infty} \|T_n\| \, \|x\|,$$

but $\liminf_{n\to\infty} \|T_n x\| = \lim_{n\to\infty} \|T_n x\| = \|Tx\|$, so that

$$\|Tx\| \le (\liminf_{n\to\infty} \|T_n\|) \, \|x\|, \quad \text{for every } x \in X,$$

which clearly implies $\|T\|_{L(X,Y)} \le \liminf_{n\to\infty} \|T_n\|_{L(X,Y)}$.

Exercise 1.7.8.1. Let Λ be a set. Assume that a function $b : \Lambda \to \mathbb{K}$ is such that for every sequence $x \in c_0 = c_0(\Lambda)$ the product sequence $x\, b$ is in $\ell^1 = \ell^1(\Lambda)$. Prove that then $b \in \ell^1$.

Solution. Given a finite subset $F \subseteq \Lambda$, we consider $b_F \in c_{00}$ defined as $b_F = a\, \chi_F$. Then

$$T_F : x \mapsto \sum_{\lambda \in F} x(\lambda)\, b_F(\lambda)$$

is in the dual space of c_0. Given $x \in c_0$ we have

$$|T_F(x)| = \left| \sum_{\lambda \in F} x(\lambda)\, b_F(\lambda) \right| \le \sum_{\lambda \in F} |x(\lambda)|\, |a_F(\lambda)| \le \sum_{\lambda \in \Lambda} |x(\lambda)|\, |b(\lambda)|,$$

finite since $x\, b \in \ell^1$ by the hypothesis. By the Banach–Steinhaus theorem, there is $L > 0$ such that $\|T_F\| \le L$ for every finite subset $F \subseteq \Lambda$. Then, arguing as in 1.1.6, if $a_F = \sum_{\lambda \in F} \overline{\operatorname{sgn} b(\lambda)}\, e_\lambda$, we have $a_F \in c_{00}$, $\|a_F\|_\infty \le 1$ so that:

$$T_F(a_F) = \sum_{\lambda \in F} \overline{\operatorname{sgn} b(\lambda)}\, b(\lambda) = \sum_{\lambda \in F} |b(\lambda)| \le \|T_F\| \|a_F\|_\infty \le \|T_F\| \le L, \quad \text{for every finite subset } F \subseteq \Lambda,$$

so that $b \in \ell^1$ and $\|b\|_1 \le L$. □

Imitating this solution it is easy to see that:

Exercise 1.7.8.2. Let $p > 1$ and let q be the conjugate exponent of p. Assume that a function $b : \Lambda \to \mathbb{K}$ is such that for every sequence $x \in \ell^p = \ell^p(\Lambda)$ the product sequence $x\, b$ is in $\ell^1 = \ell^1(\Lambda)$. Prove that then $a \in \ell^q(\Lambda)$.

1.8. The open mapping theorem.

. The open mapping theorem *Let X and Y be Banach spaces. If $T : X \to Y$ is linear, continuous and surjective then it is an open mapping, i.e. $T(A)$ is open in Y for every open subset A of X.*

Proof. It is clearly equivalent to prove that T is open at 0, i.e that $T(U)$ is neighborhood of 0 in Y for every nbhd U of 0 in X; by homogeneity, we are reduced to prove that $T(B_X)$ is neighborhood of 0 in Y, i.e. that for some $s > 0$ we have $T(B_X) \supseteq s\, B_Y$. To simplify notation let $S = T(B_X)$; notice that $n\, S = T(n\, B_X)$ for every $n \in \mathbb{N}$, so that

$$Y = T(X) = T\left(\bigcup_{n \ge 1} n\, B_X \right) = \bigcup_{n \ge 1} T(n\, B_X) = \bigcup_{n \ge 1} n\, S;$$

(in other words, S is absorbing in Y; this is where surjectivity of T enters the proof). We want to prove that S is a nbhd of the origin. If $\bar{S} = \mathrm{cl}_Y(S)$ is the closure of S in Y, since $\bigcup_{n\ge1} n\, \bar{S} \supseteq \bigcup_{n \ge 1} n\, S = Y$ and Y is complete, by Baire's theorem $n\, \bar{S}$ has non empty interior for some $n \ge 1$, equivalently the set $\bar{S}$ itself has nonempty interior; then, since S is convex and symmetric, $0 \in \mathrm{int}_Y\, \bar{S}$ (2.2.4), i.e. for some

$r > 0$ we have $\bar{S} \supseteq r\,B_Y$. We prove that $T(B_X) = S \supseteq (r/2)B_Y$, which concludes the proof. Given $y = y_0 \in (r/2)\,B_Y$ and $\varepsilon > 0$ we can find x_0 in $(1/2)B_X$ such that $\|y_0 - Tx_0\|_Y \le \varepsilon$; in fact

$$(1/2)\bar{S} = \mathrm{cl}\,(T((1/2)B_X) \supseteq (r/2)\,B_Y.$$

We pick $x_0 \in (1/2)B_X$ such that $\|y_0 - Tx_0\|_Y \le r/2^2$. Set $y_1 = y_0 - Tx_0$; then we can find $x_1 \in (1/2^2)B_X$ such that $\|y_1 - Tx_1\|_Y \le r/2^3$. Inductively, assume that $y_0 \dots, y_m$ and $x_0, \dots, x_{m-1}$ have been picked in such a way that:

$$y_k \in (r/2^{k+1})B_Y;\ x_k \in (1/2^{k+1})B_X,\ \|y_k - Tx_k\| \le \frac{r}{2^{k+1}};\ y_{k+1} = y_k - Tx_k \quad \text{for } k = 0, \dots, m-1.$$

Since

$$(1/2^{m+1})\bar{S} = \mathrm{cl}\left(T((1/2^{m+1})B_X\right) \supseteq (r/2^{m+1})\,B_Y$$

we can pick $x_m \in (1/2^{m+1})B_X$ such that $\|y_m - Tx_m\|_Y \le r/(2^{m+2})$, and induction can continue. We have

$$y = y_0 = y_0 - Tx_0 + Tx_0 = y_1 + Tx_0 = y_1 - Tx_1 + Tx_1 + Tx_0 = y_2 + (Tx_0 + Tx_1) = \dots = y_{m+1} + \sum_{k=0}^{m} Tx_k,$$

for every $m \in \mathbb{N}$; now the series $\sum_{k=0}^{\infty} x_k$ is normally convergent in the Banach space X to a vector $x \in B_X$ (we have $\|x_k\|_X \le 1/2^{k+1}$), thus the series $\sum_{k=0}^{\infty} Tx_k$ is also normally convergent in Y to the vector Tx; since

$$\left\| y - \sum_{k=0}^{m} Tx_k \right\|_Y = \|y_{m+1}\|_Y \le \frac{r}{2^{m+2}},$$

we have that y is the sum of the series $\sum_{k=0}^{\infty} Tx_k$; thus $y = Tx \in T(B_X)$. □

Corollary. *A continuous isomorphism of Banach spaces is a homeomorphism (i.e., the inverse is also continuous).*

Proof. Left to the reader: recall that a continuous bjiective map is a homemorphism iff it maps open sets into open sets. □

Exercise 1.8.0.3. (!) *Linear maps with closed range* Let X, Y be Banach spaces. Assume that $T : X \to Y$ is a linear continuous map. Prove that the following are equivalent:

(a) T is injective and the range $T(X)$ of T is closed in Y
(b) T is a homeomorphism of X onto $T(X)$.
(c) There exists $k > 0$ such that $\|Tx\|_Y \ge k\|x\|_X$, for every $x \in X$.

Deduce from the preceding equivalence that T has closed range, i.e. $T(X)$ is closed in Y, if and only if there exists $k > 0$ such that $\|Tx\|_Y \ge k\ \mathrm{dist}(x, \mathrm{Ker}\,T)$, for every $x \in X$.

Solution. (a) implies (b): if $T(X)$ is closed in the complete space Y, then $T(X)$ is also complete ,and by the corollary the map $T^{-1} : T(X) \to X$ is continuous, in particular T is a homeomorphism onto the image $T(X)$.

(b) implies (c) If T^{-1} is continuous there is $L > 0$ such that $\|T^{-1}y\|_X \le L\|y\|_Y$, for every $y \in T(X)$. If $y = Tx$ we get $\|x\|_X \le L\|Tx\|_Y \iff \|Tx\|_Y \ge (1/L)\|x\|_X$.

(c) implies (a): clearly T is injective (if $\|Tx\| = 0$ then $0 \ge k\|x\|_X$ implies $x = 0$) and the "inverse" $T^{-1} : T(X) \to X$ is Lipschitz with constant $1/k$. We only have to prove that $T(X)$ is closed in Y: if $y_n = Tx_n$ converges to $y \in Y$, then y_n is a Cauchy sequence, and since

$$\|y_n - y_m\|_Y = \|Tx_n - Tx_m\|_Y \ge k\,\|x_n - x_m\|,$$

$(x_n)_{n\in\mathbb{N}}$ is also a Cauchy sequence in X; if x is its limit, we have $y = Tx \in X$.

For the last statement: T induces a continuous bijection $S : X/\mathrm{Ker}\,T \to T(X)$ (by well–known algebraic statements and 1.5.1.1); then we have $\|S(x + \mathrm{Ker}\,T)\|_Y \ge k\ \mathrm{dist}(x, \mathrm{Ker}\,T)$ for every $x \in X$, that is $\|Tx\|_Y \ge k\ \mathrm{dist}(x, \mathrm{Ker}\,T)$ for every $x \in X$.

□

1.8.1. *Topological supplement in a Banach space.* If $Y$and Z are normed spaces, the product $Y \times Z$, under the product topology, is normable, and the *product norm* $\|(y,z)\|_{Y\times Z} = \max\{\|y\|_Y,\ \|z\|_Z\}$ gives the product topology. It is well–known and easy to prove that with this norm $Y \times Z$ is complete if and only if Y and Z are both complete. Assume now that we have a Banach space X with two linear subspaces Y and Z such that $Y \cap Z = \{0\}$ and $X = Y + Z$, so that, algebraically, we have $X = Y \oplus Z$; in other words, we have a bijective linear map $Y \times Z \to X$ given by $(y, z) \mapsto y + z$. If we put the product norm on $Y \times Z$, this map is clearly continuous, and is a homeomorphism if and only if Y and Z are both closed in X. In fact, if Y and Z are closed, then they are complete in the induced norm, and so $Y \times Z$ is a Banach space, and as just observed any continuous isomorphism between Banach spaces is a homeomorphism. The converse is immediate: a linear homeomorphism of normed spaces preserves completeness, so that if X is Banach then so is $Y \times Z$, and this is true iff Y and Z are both complete, which happens iff they are closed in X.

1.8.2. *The closed graph theorem.* If $f : X \to Y$, with X and Y topological spaces, is a continuous map, then provided that Y is Hausdorff, the graph of f, $\mathrm{graph}(f) = \{(x, f(x)) : x \in X\}$ is closed subset of $X \times Y$, when this space is given the product topology (for the proof, simply recall that $(x, y) \mapsto y$ and $(x, y) \mapsto f(x)$ are both continuous maps from $X \times Y$ to Y, and that the coincidence set ot two continuous maps arriving from some space to a Hausdorff space is closed). Even with very simple spaces this condition in general does not imply continuity. For linear maps between Banach spaces we however have:

. The closed graph theorem. *Let X, Y be Banach spaces, and let $T : X \to Y$ be linear. Then T is continuous if and only if the graph of T, i.e $G = \{(x, Tx) : x \in X\}$ is closed in $X \times Y$.*

Proof. Necessity is true in general, we only need to prove sufficiency: if G is closed, then T is continuous. Recall that $X\times Y$, in the product norm $\|(x, y)\|_{X\times Y} = \max\{\|x\|_X,\ \|y\|_Y\}$ is a Banach space; then G, as a closed subspace, is also a Banach space. Let S be the restriction of the first projection pr_X : $X\times Y \to X$ to G: then S is a continuous isomorphism of Banach spaces, and as such a homeomorphism, by the previous corollary. Then $T = (\mathrm{pr}_Y |G) \circ S^{-1}$ is continuous from X to Y. □

Exercise 1.8.2.1. Let $X = C([0, 1])$ with sup–norm and $Y = C^1([0, 1])$ with the induced norm. Prove that Y is a dense proper subspace of X and that (maybe ”so that”?) Y is not a Banach space. Prove that the derivation operator $D : Y \to X$ is not continuous, but has closed graph in $Y \times X$ (this is essentially the theorem on ”passing to the limit under the differentiation operator”).

Exercise 1.8.2.2. Let X, Y be Banach spaces and let $T \in \mathrm{Hom}_{\mathbb{K}}(X, Y)$ be such that $\varphi \circ T \in X^*$ for every $\varphi \in Y^*$. Use the closed graph theorem to prove that T is continuous.

Solution. Assume that x_n converges to $x \in X$ and that $y_n = Tx_n$ converges to $y \in Y$. We wish to prove that $y = Tx$; we prove it by proving that $\varphi(y) = \varphi(Tx)$ for every $\varphi \in Y^*$. This is clear because $\varphi \circ T(x_n) = \varphi(y_n)$ converges to $\varphi \circ T(x)$ by continuity of $\varphi \circ T$, and also to $\varphi(y)$ by continuity of φ. □

Exercise 1.8.2.3. *(from Folland)* Let $Y = \ell^1(\mathbb{N}^>)$, where $\mathbb{N}^> = \{1, 2, 3, \dots\}$, let $\omega : \mathbb{N}^> \to \mathbb{K}$ be the inclusion $\omega(n) = n$, and let

$$X = \{x \in Y : \omega\, x \in Y\}.$$

Then:

(i) X is a proper dense subspace of Y, hence it is not complete in the induced norm.
(ii) The mapping $T(x) = \omega\, x$ is a linear isomorphism of X onto Y; it is not continuous, but it has closed graph.
(iii) The inverse $S = T^{-1} : Y \to X$ is continuous but not an open mapping.

Exercise 1.8.2.4. Let X, Y, Z be Banach spaces, and let $\langle\#, \#\rangle : X \times Y \to Z$ be a bilinear map. Prove that if this map is separately continuous in each variable, then it is continuous in the product topology (in other words: if each of the mappings $\langle x, \#\rangle : \eta \mapsto \langle x, \eta\rangle$ and $\langle\#,\ y\rangle : \xi \mapsto \langle\xi,\ y\rangle$ is continuous, for every $x \in X$ and every $y \in Y$, then $(x, y) \mapsto \langle x,\ y\rangle$ is continuous from $X \times Y$, with the product topology, to Z).

Solution. It is enough to prove that the bilinear mapping is continuous at $(0, 0)$. Assume that $x_n \in X$ and $y_n \in Y$ converge to 0 in X, Y respectively. Then, if $T_n = \langle x_n, \#\rangle$ and $T = \langle a, \#\rangle$ we have that $T_n(y) \to \langle 0, y\rangle = 0$ for every given $y \in Y$, by continuity of $\langle\#,\ y\rangle$, so that T_n is equicontinuous by 1.7.8. It follows that there is a constant $L > 0$ such that $\|T_n\|_{L(Y,Z)} \le L$ for every $n \in \mathbb{N}$, so that

$$\|T_n(y)\|_Z = \|\langle x_n,\ y\rangle\|_Z \le L\|y\|_Y \quad \text{for every } n \in \mathbb{N} \text{ and every } y \in Y;$$

then

$$\|\langle x_n,\, y_n\rangle\|_Z \le L\|y_n\|_Y, \quad \text{for every } n \in \mathbb{N},$$

proving continuity at $(0,0)$. □

1.8.3. *An application of the open mapping theorem to Fourier analysis.* Given $\tau > 0$ let $X = L^1_\tau$ be the space of $\tau-$periodic functions $f : \mathbb{R} \to \mathbb{C}$, with norm $\|f\|_1 = \int_{(\tau)} |f(x)|\, dx/\tau$ (integral on an interval of length τ). The mapping $c : X \to c_0 = c_0(\mathbb{Z}, \mathbb{C})$ given by

$$c(f) = (c_n(f))_{n\in\mathbb{Z}} \quad \text{where} \quad c_n(f) = \int_{(\tau)} f(x)\, e^{-2\pi i n x/\tau}\, \frac{dx}{\tau}$$

is linear, continuous of norm 1, and injective (accept this last fact). If it were onto, by the open mapping theorem it would be a homeomorphism. The Dirichlet kernel $D_m(x) = \sin((2m+1)\omega x/2)/\sin(\omega x/2)$ is mapped by c into the characteristic function of $[-m, m]\cap\mathbb{Z}$, of norm 1 in c_0. But if we prove that $\|D_m\|_1$ tends to infinity as $m \to \infty$, then c cannot be a homeomorphism; hence $c(L^1_\tau) \subsetneqq c_0(\mathbb{Z})$. Same for the Fourier transform $\Phi : L^1(\mathbb{R}) \to C_0(\mathbb{R})$; the functions $f_m = m\,\mathrm{sinc}(mx)\,\mathrm{sinc}(x)$ are transformed by Φ in

$$\widehat{f_m} = \mathrm{rect}(\#/m) * \mathrm{rect},$$

and it is easy to show that the right–hand member has sup–norm 1; if $\lim_{m\to\infty} \|f_m\|_1 = \infty$ we get a contradiction with the Fourier transform being a homeomorphism of $L^1(\mathbb{R})$ onto $C_0(\mathbb{R})$, thus proving that $\Phi(L^1(\mathbb{R})) \subsetneqq C_0(\mathbb{R})$.

Notice that, if $0 \le x \le 1/2$

$$|f_m(x)| = \frac{|\sin(\pi m x)|\,|\sin(\pi x)|}{\pi^2 x^2} \ge \frac{2}{\pi}\,\frac{|\sin(\pi m x)|}{\pi x}, \quad (x \in [0, 1/2])$$

so that

$$\|f_m\|_1 \ge \frac{2}{\pi}\int_0^{1/2} \frac{|\sin(\pi m x)|}{\pi x}\, dx = \frac{2}{\pi}\int_0^{\pi/2} \frac{|\sin(mt)|}{t}\, dt,$$

and that

$$\|D_m\|_1 = 2\int_0^{\tau/2} \frac{|\sin((2m+1)\omega x/2)|}{\sin(\omega x/2)}\,\frac{dx}{\tau} = 2\int_0^{\pi/2} \frac{|\sin((2m+1)t)|}{\sin t}\,\frac{dt}{\pi} \ge \frac{2}{\pi}\int_0^{\pi/2} \frac{|\sin((2m+1)t)|}{t}\, dt,$$

since $0 < \sin t < t$ for $0 < t < \pi/2$ Thus we need only to prove that the last integral tends to ∞ as $\lambda = 2m+1$ tends to ∞. In fact, as $\lambda \to +\infty$:

$$\int_0^{\pi/2} \frac{|\sin(\lambda t)|}{t}\, dt = \int_0^{\lambda\pi/2} \frac{|\sin\theta|}{\theta}\, d\theta \to \int_0^{\infty} \frac{|\sin\theta|}{\theta}\, d\theta = +\infty.$$

An explicit function in $C_0(\mathbb{R})$ not in the image of $L^1(\mathbb{R})$ is for instance

$$g(\xi) = \frac{\operatorname{sgn}\xi}{1 + |\log|\xi||},$$

but the proof of this fact will not be presented here.

1.9. Hilbert spaces. The notion of Hilbert space is supposed to be known, and we quickly recall here the fundamentals. A *scalar product* on a $\mathbb{K}-$space H is a positive definite hermitian form $(\# \mid \#) : H \times H \to \mathbb{K}$, that is each mapping $x \mapsto (x \mid y)$ is linear from H to $\mathbb{K}$, we have $\overline{(x \mid y)} = (y \mid x)$, and $(x \mid x) > 0$ if $x \in H \setminus \{0\}$. Setting $|x| = (x \mid x)^{1/2}$, we have the *Cauchy–Schwarz inequality*:

$$|(x \mid y)| \leq |x|\,|y|, \quad \text{for every } x, y \in H, \text{ with equality iff } x \text{ and } y \text{ are linearly dependent;}$$

it is easy to deduce from this the *triangle inequality*, or Minkowski inequality:

$$|x + y| \leq |x| + |y|, \quad \text{for every } x, y \in H,$$

with equality if and only if $y = 0$, or $x = \lambda y$ with $\lambda \geq 0$. Thus $x \mapsto |x|$ is a norm on H; and if H is complete in this norm, then H is said to be a Hilbert space. Among normed spaces, scalar product spaces are characterized by the *parallelogram identity*:

$$|x + y|^2 + |x - y|^2 = 2|x|^2 + 2|y|^2 \quad \text{often in the form} \quad \left|\frac{x+y}{2}\right|^2 + \left|\frac{x-y}{2}\right|^2 = \frac{|x|^2 + |y|^2}{2}.$$

1.9.1. *Dual of a scalar product space.* The Cauchy–Schwarz inequality proves that, given $a \in H$, the mapping $\varphi_a(x) = (x \mid a)$ is a continuous linear functional on H, with norm $\|\varphi_a\|_{H^*} = |a|$. In fact by the mentioned inequality $|\varphi_a(x)| = |(x \mid a)| \leq |x|\,|a|$ so that $\|\varphi_a\| \leq |a|$; and if $a \neq 0$, then $\varphi_a(a) = (a \mid a) = |a|^2$, so that $\|\varphi_a\|_{H^*} = |a|$. Then:

. *If H is a scalar product space the mapping $a \mapsto \varphi_a$, where $\varphi_a(x) = (x \mid a)$ for every $x \in H$, is an isometric conjugate linear embedding of H into its dual space H^*.*

(conjugate linear means that we have $\varphi_{\lambda a} = \bar{\lambda}\,\varphi_a$ for every $\lambda \in \mathbb{K}$ and every $a \in H$; of course the map is $\mathbb{R}-$linear) The Riesz completeness theorem 1.9.8 will assert that this map in surjective if and only if H is complete, i.e. a Hilbert space.

1.9.2. *Orthogonality.* In a scalar product space H a pair of vectors x, $y \in H$ are said to be orthogonal if $(x \mid y) = 0$; the relation of orthogonality is clearly symmetric, and almost antireflexive, in the sense that the only vector orthogonal to itself is the zero vector. A family $(x_\lambda)_{\lambda \in \Lambda}$ of vectors is said to be *orthogonal* if $(x_\lambda \mid x_\mu) = 0$ for $\lambda, \mu \in \Lambda$ and $\lambda \neq \mu$; it is *orthonormal* if it is orthogonal and each vector has norm 1. It is plain that an orthogonal set non containing 0 is linearly independent, so that in particular any orthonormal set is linearly independent.

Definition. If H is scalar product space and S is a subset of H, the *orthogonal* of S in H is the subset of H defined by

$$S^\perp = \{y \in H : (y \mid x) = 0 \text{ for every } x \in S.\}$$

It is immediate to see that

. *For every subset S of H the set $S^\perp$ is a closed vector subspace of H. And if T is the closure of the linear subspace spanned by S, then $T^\perp = S^\perp$.*

Proof. By definition $S^\perp = \bigcap_{a \in S} \operatorname{Ker}(\varphi_a)$, an intersection of closed linear subspaces. Since $S \subseteq T$ we have $S^\perp \supseteq T^\perp$; and if $a \in S^\perp$, then $\varphi_a(x) = 0$ for every $x \in S$, that is $\operatorname{Ker}\varphi_a \supseteq S$; but since $\operatorname{Ker}\varphi_a$ is a closed linear subspace, we also have $\operatorname{Ker}\varphi_a \supseteq T$, i.e. $a \in T^\perp$. □

1.9.3. *The Pythagorean theorem.* It is a simple calculation to check that:

. *If H is a scalar product space and $(u_1, \dots, u_m)$ is an orthogonal family, then*

$$\left|\sum_{k=1}^m u_k\right|^2 = \sum_{k=1}^m |u_k|^2.$$

For an infinite orthogonal family the following holds:

. Unrestricted pythagorean theorem *If H is a Hilbert space, and $(u_\lambda)_{\lambda \in \Lambda}$ is an orthogonal family in H, then this family is summable in H if and only if the family $(|u_\lambda|^2)_{\lambda \in \Lambda}$ of the squares of the norms is summable in $\mathbb{R}$, and in this case we have:*

$$\left|\sum_{\lambda \in \Lambda} u_\lambda\right|^2 = \sum_{\lambda \in \Lambda} |u_\lambda|^2.$$

Proof. Given $\varepsilon > 0$ we want to find a finite subset $F(\varepsilon)$ of Λ such that $\left|\sum_{\lambda\in F} u_\lambda\right|^2 \le \varepsilon^2$ if $F \subseteq \Lambda \setminus F(\varepsilon)$. Since the family is orthogonal the finite pythagorean theorem implies that

$$\left|\sum_{\lambda\in F} u_\lambda\right|^2 = \sum_{\lambda\in F} |u_\lambda|^2,$$

and the Cauchy condition for the summability of the family of vectors and for the square of the norms are the same. Passing to the limit in the preceding equality "as F gets larger and larger" we get the stated equality of the norms squared.

□

1.9.4. *Projection on convex subsets.* A fundamental result for Hilbert spaces is the following:

THEOREM. *Let C be a closed convex non–empty subset of a Hilbert space H. For every $x \in H$ there exists a unique $c = c_x \in C$ which realizes the minimum distance of x from C, i.e. such that*

$$|x - c| = \min\{|x - y| : y \in C\}.$$

The parallelogram identity may in fact be used to show that any minimizing sequence $c_k \in C$, i.e. such that $\lim_k |x - c_k| = \inf\{|x - y|; y \in C\}$ is Cauchy, and hence converges to a $c \in C$. A picture illustrates the following:

PROPOSITION. *Given a convex subset C and a point x in a scalar product spaces, a point $c \in C$ realizes the minimum distance of x from C if and only if, for every $y \in C$:*

$$\mathrm{Re}(x - c \mid y - c) \le 0.$$

Proof. The point $c \in C$ realizes the minimum distance of x from C if and only if for every $y \in C$ we have $|x - c| \le |x - y|$, equivalently

$$|x - c|^2 \le |x - y|^2 = |(x - c) - (y - c)|^2 = |x - c|^2 + |y - c|^2 - 2\,\mathrm{Re}(x - c \mid y - c),$$

or else

$$(*) \qquad 2\,\mathrm{Re}(x - c \mid y - c) \le |y - c|^2 \quad \text{for every } y \in C.$$

If $\mathrm{Re}(x - c \mid y - c) \le 0$ this condition is trivially verified; conversely, note that by convexity of C we have $y_t = c + t(y - c) \in C$ for every $t \in [0, 1]$ and $y \in C$; substituting this y_t in (*) we get

$$2t\ \mathrm{Re}(x - c \mid y - c) \le t^2\,|y - c|^2 \iff 2\ \mathrm{Re}(x - c \mid y - c) \le t\,|y - c|^2, \quad 0 < t \le 1,$$

and letting $t \to 0^+$ we get $\mathrm{Re}(x - c \mid y - c) \le 0$, as required. □

REMARK. The preceding proposition immediately implies uniqueness of the point of minimum distance: if d also realizes the minimum distance from x we have

$$\mathrm{Re}(x - d \mid c - d) \le 0; \quad \mathrm{Re}(x - c \mid d - c) \le 0 \Longrightarrow \mathrm{Re}(x - d \mid c - d) + \mathrm{Re}(x - c \mid d - c) \le 0,$$

that is

$$\mathrm{Re}(x - d - (x - c) \mid c - d) \le 0 \iff \mathrm{Re}(c - d \mid c - d) \le 0 \iff |c - d|^2 \le 0,$$

equivalently, $c = d$.

The unique point $c = c_x$ is sometimes called the *orthogonal projection* of x into C, and may be written $\pi_C(x)$. The function π_C is extensively studied. But here we are interested only in a particular case, that in which C is a closed affine subspace.

1.9.5. *Distance from an affine subspace.* An affine subspace of a linear space H is simply a translate of a linear subspace V, i.e. a lateral class $a + V$ of this subspace, for some $a \in H$; V is then the *directing subspace* of the affine subspace . Clearly $a + V$ is closed in H iff V is closed. Observe that

. *Let H be a scalar product space, and let $a + V$ be an affine subspace of H. Given $x \in H$, a point $c \in a + V$ realizes the smallest distance of x from the affine subspace $a + V$ iff $x - c$ is orthogonal to the directing subspace V.*

Proof. Observe that $a + V = c + V$ for every $c \in a + V$; then Proposition 1.9.4 says that $c \in a + V$ is of minimum distance if and only if

$$\mathrm{Re}(x - c \mid v) \le 0 \quad \text{for every } v \in V.$$

Since V is a linear subspace this is equivalent to $x - c \in V^\perp$: in fact $-v \in V$ for every $v \in V$, so that we also have $\mathrm{Re}(x - c \mid -v) \le 0$ for every $v \in V$, hence $-\mathrm{Re}(x - c \mid v) \le 0$ for every $v \in V$; and finally, since also $i\,v \in V$ for every $v \in V$ we have $0 = \mathrm{Re}(x - c \mid i\,v) = -\mathrm{Re}(i\,(x - c \mid v)) = \mathrm{Im}(x - c \mid v)$. Thus $(x - c \mid v) = 0$ for every $v \in V$. □

1.9.6. *Orthogonal projection onto a closed subspace.*

PROPOSITION. *Let H be a Hilbert space, and let V be a closed non zero subspace. There is a map $\pi_V : H \to V \hookrightarrow H$ which associates to every $x \in V$ the unique vector $\pi_V(x) \in V$ such that $x - \pi_V(x) \in V^\perp$; this map is linear, idempotent, of norm 1, and its kernel is* $\mathrm{Ker}(\pi_V) = V^\perp$.

Proof. Existence of π_V is 1.9.4 coupled with the preceding result on affine subspaces. Linearity is due to the fact that $x+y-(\pi_V(x)+\pi_V(y)) = (x-\pi_V(x))+(y-\pi_V(y)) \in V^\perp$ since $x-\pi_V(x)$, $y-\pi_V(y) \in V^\perp$, and $V^\perp$ is a subspace, similarly for $\alpha\,x - \alpha\,\pi_V(x) = \alpha(x - \pi_V(x)$. Moreover π_V is the identity on V: if $v \in V$, then $\pi_V(v) = v$, so that π_V is idempotent; since $V \neq \{0\}$ this implies that the operator norm of π_V is not smaller than 1. Moreover for every $x \in H$ we have

$$x = \pi_V(x) + (x - \pi_V(x)) \quad \text{with} \quad (\pi_V(x) \mid x - \pi_V(x)) = 0,$$

so that

$$|x|^2 = |\pi_V(x)|^2 + |x - \pi_V(x)|^2 \geq |\pi_V(x)|^2,$$

hence $|\pi_V(x)| \leq |x|$, and the operator norm of π_V is exactly 1. That $\mathrm{Ker}(\pi_V) = V^\perp$ is now trivial. □

Of course if $W = V^\perp$ then $x - \pi_V(x) = \pi_W(x)$. And there are also the trivial extreme cases: the projection onto the subspace $\{0\}$, the zero endomorphism of H, with kernel H, and the identity of H, projection of H onto itself, with zero kernel.

1.9.7. *Orthogonal direct sum.* Notice that we have also proved:

PROPOSITION. *If H is a Hilbert space and V is a closed vector subspace then $H = V \oplus V^\perp$.*

Proof. If π_V is the orthogonal projection on V then $x = \pi_V(x) + (x - \pi_V(x))$ with $x - \pi_V(x) \in V^\perp$, so that $H = V + V^\perp$; and $V \cap V^\perp = \{0\}$, since the only isotropic vector is 0. □

It is clear that $(V^\perp)^\perp = V$ for every closed subspace V. Sometimes, if $H = V \oplus W$, with V, W closed subspaces orthogonal to each other we write also $H = V \boxplus W$, and say that H is the *orthogonal direct sum* of its two Hilbert subspaces V and W.

EXERCISE 1.9.7.1. (!) Let H be a Hilbert space, let S be a subset of H, and let T be the closure of the linear subspace spanned by S. Then the double orthogonal of S coincides with T, $S^{\perp\perp} = T$.

Solution. Denote $S^\perp$ by V; then V is a closed linear subspace, and we have proved in 1.9.2 that $S^\perp = T^\perp$. Since T is a closed linear subspace the theorem above says that for every $x \in H$ we have $x = \pi_T(x) + \pi_V(x)$. Then $T = V^\perp$.

Alternatively: we clearly have $T \subseteq V^\perp$, and we want to prove that if $a \notin T$ then also $a \notin V^\perp$. By 1.4.9 there is $\varphi \in H^*$ such that $\varphi_{|T} = 0$ but $\varphi(a) \neq 0$; by Riesz representation theorem, to be proved next, we have $\varphi = \varphi_b (= (\# \mid b))$ for some $b \in H$; $\varphi_T = 0$ means that $b \in T^\perp = V$; if $a \in V$ we then have $\varphi_b(a) = (a \mid b) = 0$, a contradiction with $\varphi(a) = \varphi_b(a) \neq 0$ □

1.9.8. *The Riesz completeness theorem.* We prove here the fact that in a Hilbert space every element of the dual is of the form stated in 1.9.1,

. THE RIESZ COMPLETENESS THEOREM *Let H be a Hilbert space. Then the mapping $a \mapsto \varphi_a$, where $\varphi_a(x) = (x \mid a)$ for every $x \in H$, is an isometric conjugate linear isomorphism of H onto its dual space H^*.*

Proof. In view of 1.9.1 only surjectivity is left to prove. Assume that $\varphi : H \to \mathbb{K}$ is a non zero linear bounded functional. Then $K = \mathrm{Ker}\,\varphi$ is a closed subspace of H, of codimension 1, so that $K^\perp = \mathbb{K}u$ for some non–zero $u \in H$. We prove that for a convenient scalar $\lambda \in \mathbb{K}$ we have $\varphi = \varphi_{\lambda\,u}$, that is, $\varphi(x) = (x \mid \lambda\,u)$ for every $x \in H$. In fact $\varphi_{\lambda\,u}(u) = (u \mid \lambda\,u) = \bar{\lambda}(u \mid u) = \bar{\lambda}\,|u|^2$, and this coincides with $\varphi(u)$ iff $\lambda = \overline{\varphi(u)}/|u|^2$; if $b = \lambda\,u = (\overline{\varphi(u)}/|u|^2)\,u$ we then have $\varphi = \varphi_b$ (linear functionals with the same nullspace K which coincide on a vector not in K coincide on the entire space). □

1.9.9. *Reflexivity.* We called the preceding proposition "completeness theorem" because in fact it is equivalent to the fact that H is complete in its norm: we know that H^*, as every dual space, is complete, so that H, if isometric to a complete space, will also be complete. For obvious reasons the theorem is also known as "Riesz representation theorem": every continuous functional is representable as the scalar product for some fixed vector. In the case of spaces $L^2(\mu)$ the theorem has also been proved directly. There we obtained a linear isometry, not a conjugate linear one: this is clearly due to the fact that in $L^2(\mu)$ we have a natural conjugation $f \mapsto \bar{f}$ and we can compose the isometry with it. Clearly H^* can be made also into a Hilbert space: simply define

$$(\varphi_a \mid \varphi_b)_{H^*} := (b \mid a)_H;$$

the exchange is to ensure linearity in the first variable: since $\alpha\,\varphi_a = \varphi_{\bar{\alpha}\,a}$ we in fact have

$$(\alpha\,\varphi_a \mid \varphi_b)_{H^*} = (\varphi_{\bar{\alpha}\,a} \mid \varphi_b)_{H^*} := (b \mid \bar{\alpha}\,a)_H = \alpha\,(b \mid a)_H = \alpha\,(\varphi_a \mid \varphi_b)_{H^*}.$$

A very important corollary of the Riesz theorem is:

. *Every Hilbert space is reflexive.*

The proof is left to the reader (if in doubt see below) : the double dual is naturally identified with H, in such a way that the canonical embedding $J : H \to H^{**}$ becomes the identity. Notice that a real Hilbert space is self–dual: the mapping $a \mapsto \varphi_a$ is a natural isometric isomorphism of H onto its dual H^*.

Help with the proof of reflexivity: since H^* is also a Hilbert space, for every $\varphi \in (H^*)^*$ there is a unique $\varphi_a \in H^*$ such that $\varphi(\varphi_x) = (\varphi_x \mid \varphi_a)_{H^*} := (a \mid x)_H$ for every $x \in H$. We denote by $\rho_a : H^* \to \mathbb{K}$ this mapping. The mapping $a \mapsto \rho_a$ is then a linear bijection of H into H^{**}; it is obtained by composing the two Riesz conjugate linear isomorphisms of a Hilbert space onto is dual, so it is linear; and coincides with the canonical mapping J; in fact $J(a)$ acts on $\varphi_b \in H^*$ as an evaluation, $\langle a, \varphi_b\rangle = \varphi_b(a) = (a \mid b)_H$, while ρ_a acts on $\varphi_b \in H^*$ as a scalar product in H^*, that is $\langle \rho_a, \varphi_b\rangle = (\varphi_b \mid \varphi_a)_{H^*} := (a \mid b)_H$. Thus $J(a) = \rho_a$, for every $a \in H$.

1.9.10. *The Lebesgue–Radon–Nikodym theorem.* To show the power of Riesz representation theorem we use it to prove the following theorem of measure theory:

. Theorem of Lebesgue–Radon–Nikodym *Let $(X, \mathcal{M}, \mu)$ and $(X, \mathcal{M}, \nu)$ be finite measure spaces. Then there exist measures $\nu_a, \nu_s : \mathcal{M} \to [0, +\infty[$ such that $\nu = \nu_a + \nu_s$, with $\nu_a \ll \mu$ and $\nu_s \perp \mu$; moreover there exists $\rho \in L^1(\mu)$ such that $\nu_a(E) = \int_E \rho\, d\mu$ for every $E \in \mathcal{M}$; and ν_a, ν_s, ρ are unique.*

Proof. We assume the measure theoretic notions as known, and simply recall that $\nu_a \ll \mu$ means that if $\mu(E) = 0$ then also $\nu_a(E) = 0$, whereas $\nu_s \perp \mu$ means that ν_s and μ live on disjoint sets, that is we may split X into a disjoint union $X = S \cup T$ with $S, T \in \mathcal{M}$, $\mu(S) = 0$ and $\nu_s(T) = 0$. Uniqueness of the decomposition is easy to prove; we prove only existence. We have to distiguish here between the spaces $\mathcal{L}^p(\mu)$, space of all $\mathcal{M}-$measurable functions f such that $|f|^p$ has finite integral, and its quotient $L^p(\mu) = \mathcal{L}^p(\mu)/\mathcal{N}(\mu)$ modulo null functions. It is clear that $\mathcal{L}^1(\mu + \nu) = \mathcal{L}^1(\mu) \cap \mathcal{L}^1(\nu)$ and that for $f \in \mathcal{L}^1(\mu + \nu)$ we have

$$\int_X f\, d(\mu + \nu) = \int_X f\, d\mu + \int_X f\, d\nu$$

(this is true for every characteristic function of a set $E \in \mathcal{M}$, by the very definition of $\mu+\nu$; by linearity it is true for every positive simple function, an by monotone convergence for every $\mathcal{M}-$measurable positive function). We consider the real Hilbert space $L^2(\mu + \nu) = L^2_{\mu+\nu}(X, \mathbb{R})$, and we define a linear functional $\varphi : L^2(\mu + \nu) \to \mathbb{R}$ by

$$\varphi(f) = \int_X f\, d\nu;$$

it is a bounded functional since

$$\left|\int_X f\, d\nu\right| \le \int_X |f|\, d\nu \le \int_X |f|\ \cdot 1\, d(\mu + \nu) \le \|f\|_2\, \sqrt{(\mu + \nu)(X)}.$$

By Riesz theorem then there exists $u \in L^2(\mu + \nu)$ such that

$$\int_X f\, d\nu = \int_X f\, u\, d(\mu + \nu) \quad \text{for every} \quad f \in L^2(\mu + \nu),$$

equivalently

$$(*) \qquad \int_X f\,(1 - u)\, d\nu = \int_X f\, u\, d\mu \quad \text{for every} \quad f \in L^2(\mu + \nu).$$

In particular this formula holds for every measurable simple function. Let us now prove that $0 \le u(x) \le 1$ holds for $\mu+\nu-$almost every $x \in X$. In fact, if $A = \{x \in X : u(x) < 0\}$, formula (*) with χ_A in place of f says that

$$\int_A (1-u)\, d\nu = \int_A u\, d\mu;$$

but the left–hand side is larger than $\nu(A)$ and the right–hand side is negative; so they are both 0; since u has constant sign -1 on A we have $\mu(A) = 0$, so that $\mu(A) + \nu(A) = 0$. Similarly, if $B = \{x \in X : u(x) > 1\}$ we have that $\int_B (1-u)\, d\nu \le 0$ and $\int_B u\, d\mu \ge \mu(B)$, so that equality is possible if and only if $\mu(B) = 0$ and $\int_B (1-u)\, d\nu = 0$, and since $1-u$ has constant sign -1 on B this implies $\nu(B) = 0$. Having proved that u and $1-u$ are both positive, we can use monotone convergence in formula (*) to show that

(**) $$\int_E f\,(1-u)\, d\nu = \int_E f\, u\, d\mu \quad \text{for every } f \text{ positive measurable and every } E \in \mathcal{M}.$$

Let $S = \{x \in X : u(x) = 1\}$, $T = X \setminus S$. From (**), with $E = S$ and $f = 1$ we get $\mu(S) = 0$. Since:

$$\int_E f\,(1-u)\, d\nu = \int_{E\cap S} f\,(1-u)\, d\nu + \int_{E\cap T} f\,(1-u)\, d\nu = \int_{E\cap T} f\,(1-u)\, d\nu,$$

we have

$$\int_{E\cap T} f\,(1-u)\, d\nu = \int_E f\, u\, d\mu \quad \text{for every } f \text{ positive measurable and every } E \in \mathcal{M}.$$

Consider $v : X \to \mathbb{R}$ defined by $v(x) = 1/(1-u(x))$ for $x \in T$, and $v(x) = 0$ for $x \in S$. Then v is positive and measurable, hence

$$\int_{E\cap T} d\nu = \int_E v\, u\, d\mu \quad \text{for every } E \in \mathcal{M};$$

the left hand side is $\nu(E \cap T)$; setting $\rho(x) = u(x)\, v(x)$ we have proved that

$$\nu(E\cap T) = \int_E \rho\, d\mu \quad \text{for every } E \in \mathcal{M},$$

and the proof is concluded: we set

$$\nu_a(E) = \int_E \rho\, d\mu; \qquad \nu_s(E) = \nu(E \cap S) \quad \text{for every } E \in \mathcal{M},$$

and this is the required decomposition. □

Remark. We may observe that the density function ρ is exactly the (pointwise) sum of the series

$$\rho(x) = u(x) + (u(x))^2 + \ldots \quad \text{for } x \in T;\ \rho(x) = 0 \text{ for } x \in S;$$

this density is in $L^1(\mu)$, in general not in $L^2(\mu)$.

The theorem also holds when μ is a $\sigma-$finite measure; this is easy to see, simply partitioning X into a countable family of sets of finite measure and applying the above result to each piece.

1.9.11. *Bessel inequality and Parseval identity.* In a scalar product space, let $(u_\lambda)_{\lambda\in\Lambda}$ be an orthonormal subfamily. For every finite subset F of Λ denote by $V(F)$ the subspace spanned by the subfamily $(u_\lambda)_{\lambda\in F}$; it is easy to see that for every $x \in H$ the orthogonal projection $\pi_F(x)$ of x into $V(F)$ is given by

$$\pi_F(x) = \sum_{\lambda\in F} (x \mid u_\lambda)\, u_\lambda;$$

we simply check that $x - \pi_F(x) \in V(F)^\perp$, easily done by checking that $(x - \pi_F(x) \mid u_\mu) = 0$ for every $\mu \in F$. We then have, by the pythagorean theorem:

(*)

$$|x|^2 = |\pi_F(x)|^2 + |x-\pi_F(x)|^2 = \left|\sum_{\lambda\in F} (x \mid u_\lambda)\, u_\lambda\right|^2 + (\operatorname{dist}(x, V(F)))^2 = \sum_{\lambda\in F} |(x \mid u_\lambda)|^2 + (\operatorname{dist}(x, V(F)))^2.$$

In particular, for every finite subset F of Λ we have $\sum_{\lambda\in F} |(x \mid u_\lambda)|^2 \le |x|^2$ so that

(Bessel inequality) $$\sum_{\lambda\in\Lambda} |(x \mid u_\lambda)|^2 \le |x|^2.$$

The unrestricted pythagorean theorem 1.9.3 then says that the family $\sum_{\lambda\in\Lambda}(x \mid u_\lambda)\, u_\lambda$ is summable in H if H is complete; we can pass to the limit in (*) as F increases, obtaining

$$(**)\qquad |x|^2 = \left|\sum_{\lambda\in\Lambda}(x \mid u_\lambda)\, u_\lambda\right|^2 + (\mathrm{dist}(x, V)^2 = \sum_{\lambda\in F}|(x \mid u_\lambda)|^2 + (\mathrm{dist}(x, V)^2,$$

having denoted by V the closure of the subspace U spanned by the family $(u_\lambda)_{\lambda\in\Lambda}$. We then get that the orthogonal projection of x into V is exactly

$$\pi_V(x) = \sum_{\lambda\in\Lambda}(x \mid u_\lambda)\, u_\lambda,$$

and that Bessel inequality is an equality:

Parseval identity
$$|x|^2 = \sum_{\lambda\in\Lambda}(x \mid u_\lambda)\, u_\lambda \;\left(= |\pi_V(x)|^2\right) \quad (x \in V)$$

exactly if and only if $x \in V$, the closure of the space spanned by the family, $V = \mathrm{cl}_H\langle\{u_\lambda : \lambda \in \Lambda\}\rangle$.

1.9.12. *Orthonormal bases.* An orthonormal family of vectors in a Hilbert space is said to be an *orthonormal basis* for the closure V of the linear space it spans. Observe that the unrestricted pythagorean theorem 1.9.3 may be interpreted as expressing the following fact:

. *If $(u_\lambda)_{\lambda\in\Lambda}$ is an orthonormal basis for the closed subspace V of the Hilbert space H, then the mapping*

$$\varphi : \ell^2(\Lambda) \to V \hookrightarrow H \quad \text{given by} \quad \varphi(\xi) = \sum_{\lambda\in\Lambda}\xi_\lambda\, u_\lambda,$$

is an isometric isomorphism of $\ell^2(\Lambda)$ onto V. In other words, every vector $x \in V$ admits a unique representation of the form $x = \sum_{\lambda\in\Lambda}\xi_\lambda\, u_\lambda$, for a unique $\xi = (\xi_\lambda)_{\lambda\in\Lambda} \in \ell^2(\Lambda)$.

In particular, if $V = H$ every vector $x \in H$ is representable in this way.

The question spontaneously arises: does a Hilbert space always have an orthonormal basis?

The answer is affirmative, and is a consequence of Zorn's lemma (in general). In fact it is easy to prove the following

. Theorem on orthonormal bases *Let H be a Hilbert space, and let S be an orthonormal subset of H. Then the following are equivalent:*

(i) *The set S is an orthonormal basis for H*
(ii) *The set S spans a dense subspace of H.*
(iii) *The orthogonal $S^\perp$ of S is $\{0\}$.*
(iv) *The set S is maximal as an orthonormal subset (that is, not properly contained in a larger orthonormal subset of H).*

Proof. (i) and (ii) are equivalent by definition. And if V is the closure of the subspace spanned by S, and $W = V^\perp = S^\perp$ then $x = \pi_V(x) + \pi_W(x)$, so that $V = H$ if and only if $W = \{0\}$, hence (ii) and (iii) are equivalent. Finally, equivalence of (iii) and (iv) is clear: S can be enlarged as an orthonormal subset only by adding to it a vector of unit norm belonging to $S^\perp$. □

1.9.13. *Every Hilbert space has an orthonormal basis.* A trivial application of Zorn's lemma proves that every scalar product space has maximal orthonormal subsets. Then every Hilbert space has an orthonormal basis, hence every Hilbert space is isometrically isomorphic to a space $\ell^2(\Lambda)$. It can be proved that the cardinality of an orthonormal basis is uniquely determined by the space, that is, all orthonormal bases have the same cardinality. When this cardinality is finite, it is of course the linear dimension of the space; it is countable only for *separable spaces*, i.e. spaces which admit a countable dense subset. The separable Hilbert spaces are then the finite dimensional spaces and the spaces isometrically isomorphic to $\ell^2(\mathbb{N})$.

Exercise 1.9.13.1. If Λ is countable, then all spaces $\ell^p(\Lambda)$ and $c_0(\Lambda)$ are separable (the $\mathbb{Q}-$linear subspace consisting of all rational valued functions with finite support, $c_{00}(\Lambda) \cap \mathbb{Q}^\Lambda$ is dense in all these spaces if $\mathbb{K} = \mathbb{R}$, otherwise take $c_{00}(\Lambda) \cap (\mathbb{Q} + i\,\mathbb{Q})^\Lambda$).

The space $\ell^\infty(\mathbb{N})$ is not separable, and neither are the spaces $\ell^p(\Lambda)$ and $c_0(\Lambda)$ for Λ uncountable: the proof is easy, but requires a bit of knowledge on metrizable separable spaces that we do not have (and do not care to acquire at present).

Caution: if the dimension is infinite, an orthonormal basis for a Hilbert space is *never* an algebraic (or Hamel) basis: the space generated by the orthonormal basis B is the set of *finite* linear combinations of elements of the basis, much smaller than the Hilbert space, which is the set of finite or infinite linear combinations of the elements of the basis, with string of coefficients in $\ell^2(B)$.

1.10. Adjoint of a bounded linear operator. If X and Y are normed spaces, and $T : X \to Y$ is a bounded linear operator, we can define the *transpose* $T^t : Y^* \to X^*$ by setting $T^t(\varphi) = \varphi \circ T$ for every $\varphi \in Y^*$. Since Hilbert spaces are (almost) self–dual, the transpose may, when X and Y are Hilbert spaces, be thought of as a map $T^* : Y \to X$. That is, we have the diagram

$$\begin{array}{ccc} X & \xrightarrow{\ T\ } & Y \\ {\scriptstyle \rho_X}\big\downarrow & & \big\downarrow{\scriptstyle \rho_Y} \\ X^* & \xleftarrow{\ T^t\ } & Y^* \end{array}$$

where ρ_X and ρ_Y are the Riesz semilinear isometries onto the dual spaces, $\rho_X(a) = (\# \mid a)_X$ and analogously for ρ_Y. We define the *Hilbert space adjoint*, or simply the *adjoint* T^* of T as:

$$T^* := \rho_X^{-1} \circ T^t \circ \rho_Y.$$

Then T^* is a linear bounded operator if so is T. The most important fact about T^* is:

. *If X and Y are Hilbert spaces and $T \in L(X, Y)$ is a bounded linear operator then its adjoint is characterized by the fact that the identity:*

(Adjoint identity) $$(Tx \mid y)_Y = (x \mid T^*y)_X$$

holds for every $x \in X$ and every $y \in Y$.

Proof. If T^*y is interpreted as the linear functional $u \mapsto (u \mid T^*y)_X$, then it must be the composition with T of the linear functional $v \mapsto (v \mid y)_Y$ on Y, so it has to act on $x \in X$ as $(Tx \mid y)_X$. □

1.10.1. *First facts on adjoints.* We have defined a map $T \mapsto T^*$ from $L(X, Y)$ to $L(Y, X)$, with X and Y Hilbert spaces. Clearly we have $(S+T)^* = S^* + T^*$ and $(\alpha\, S)^* = \bar{\alpha}\, S^*$, i.e. this map is conjugate linear. If in the adjoint identity we exchange the roles of X and Y and write the identity for an $S \in L(Y, X)$ we get

$$(Sy \mid x)_X = (y \mid S^*x)_Y \quad \text{putting now } T^* \text{ in place of } S: \quad (T^*y \mid x)_X = (y \mid (T^*)^*x)_Y,$$

so that, exchanging the arguments:

$$(x \mid T^*y)_X = ((T^*)^*x \mid y)_Y;$$

we have then $(Tx \mid y)_Y = ((T^*)^*x \mid y)_Y$ for every $x \in X$ and every $y \in Y$; this clearly implies $(T^*)^*x = Tx$ for every $x \in X$, that is $(T^*)^* = T$. The norm of the adjoint of T coincides with the norm of T: this follows from 1.4.8.1, but we can prove more:

$$\|T^* \circ T\| = \|T \circ T^*\| = \|T\|^2 = \|T^*\|^2,$$

as we now see. Taking absolute values in the adjoint identity and using Cauchy–Schwarz inequality we get

$$|(Tx \mid y)_Y| = |(x \mid T^*y)_X| \le |x|\, |T^*y| \quad x \in X,\, y \in Y;$$

taking now $y = Tx$:

$$|Tx|^2 \le |x|\, |T^*(Tx)| \le |x|\, \|T^* \circ T\|\, |x|;$$

assuming $|x| = 1$ we get

$$|Tx|^2 \le \|T^* \circ T\| \quad \text{for } |x| = 1, \text{ then} \quad \|T\|^2 \le \|T^* \circ T\| \le \|T^*\|\, \|T\|,$$

so that

$$\|T\| \le \|T^*\|$$

for every bounded operator T; but then

$$\|T^*\| \le \|(T^*)^*\| = \|T\|,$$

and finally $\|T\| = \|T^*\|$. Moreover we have $\|T\|^2 \le \|T^* \circ T\| \le \|T\|^2$, so that $\|T^* \circ T\| = \|T\|^2$. We have proved:

. *A bounded operator and its adjoint have the same operator norm; and $\|T^* \circ T\| = \|T\|^2 = \|T^*\|^2$.*

Next, if $S \in L(X,Y)$ and $T \in L(Y,Z)$ then $(T\circ S)^* = S^*\circ T^*$: in the formula $(Ty \mid z)_Z = (y \mid T^*z)_Y$ put Sx in place of y obtaining $(T(Sx) \mid z)_Z = (Sx \mid T^*z)_Y$; this last equals $(x \mid S^*(T^*z))_X$, so that we have $(T(Sx) \mid z)_Z = (x \mid S^*(T^*z))_X$ for every $x \in X$ and every $z \in Z$, proving that indeed $(T\circ S)^* = S^*\circ T^*$. Then

. *A bounded linear operator $T : X \to Y$ is a homeomorphism if and only if its adjoint T^* is a homeomorphism, and in this case $(T^{-1})^* = (T^*)^{-1}$ (the adjoint of the inverse is the inverse of the adjoint.*

(simply remember that the adjoint of the identity is the identity).

Directly from the adjoint identity we get:

. *If $T \in L(X,Y)$ then* $\operatorname{Ker} T = (T^*(Y))^\perp$ *and* $(T(X))^\perp = \operatorname{Ker}(T^*)$. *Hence* $\operatorname{Ker}(T\circ T^*) = \operatorname{Ker} T^*$ *and* $\operatorname{Ker}(T^*\circ T) = \operatorname{Ker} T$.

EXERCISE 1.10.1.1. Let $(X,\mathcal{M},\mu)$ and $(Y,\mathcal{N},\nu)$ be $\sigma-$finite measure spaces, and let $K : X\times Y \to \mathbb{K}$ belong to $L^2(\mu\otimes\nu)$.

(i) $\odot\odot$ Prove that the formula

$$Tf(x) = \int_Y K(x,y)\, f(y)\, d\nu(y)$$

defines a bounded linear operator from $L^2(\nu)$ into $L^2(\mu)$, and that $\|T\| \le \|K\|_2$, where $\|K\|_2 = \left(\int_{X\times Y} |K(x,y)|^2\, d(\mu\otimes\nu)(x,y)\right)^{1/2}$ (hint:use Cauchy–Schwarz inequality. . .).

(ii) Prove that $T^* : L^2(\mu) \to L^2(\nu)$ is defined by the formula

$$T^*g(y) = \int_X \overline{K(x,y)}\, g(x)\, d\mu(x).$$

Solution. (i)

$$\int_X |Tf(x)|^2\, d\mu(x) = \int_X \left|\int_Y K(x,y)\, f(y)\, d\nu(y)\right|^2 d\mu(x) \le \int_X \left(\int_Y |K(x,y)|\,|f(y)|\, d\nu(y)\right)^2 d\mu(x);$$

by the Cauchy–Schwarz inequality, applied to the integral in $d\nu(y)$:

$$\left(\int_Y |K(x,y)|\,|f(y)|\, d\nu(y)\right)^2 \le \int_Y |K(x,y)|^2\, d\nu(y) \int_Y |f(y)|^2\, d\nu(y) = \int_Y |K(x,y)|^2\, d\nu(y)\, \|f\|_2^2,$$

so that

$$\int_X |Tf(x)|^2\, d\mu(x) \le \int_X \left(\int_Y |K(x,y)|^2\, d\nu(y)\, \|f\|_2^2\right) d\mu(x) = \int_{X\times Y} |K(x,y)|^2\, d(\mu\otimes\nu)(x,y)\, \|f\|_2^2;$$

taking square roots we have

$$\|Tf\|_2 \le \|K\|_2 \|f\|_2,$$

as required.

(ii) Take $f \in L^2(\nu)$ and $g \in L^2(\mu)$. Then

$$\begin{aligned}(Tf \mid g)_{L^2(\mu)} &= \int_X Tf(x)\,\overline{g(x)}\, d\mu(x) = \int_X \left(\int_Y K(x,y)\, f(y)\, d\nu(y)\right)\overline{g(x)}\, d\mu(x) = \\ &\int_{X\times Y} K(x,y)\,\overline{g(x)}\, f(y)\, d(\mu\otimes\nu)(x,y) = \int_Y f(y)\left(\int_X K(x,y)\,\overline{g(x)}\, d\mu(x)\right) d\nu(y) = \\ &\int_Y f(y)\left(\int_X \overline{K(x,y)}\, g(x)\, d\mu(x)\right)^- d\nu(y) = (f \mid Sg)_{L^2(\mu)}\end{aligned}$$

if $S : L^2(\mu) \to L^2(\nu)$ is defined by

$$Sg(y) = \int_X \overline{K(x,y)}\, g(x)\, d\mu(x).$$

The interchange of integrals is allowed because the measures are by hypothesis $\sigma-$finite, and the function $(x,y) \mapsto K(x,y)\,\overline{g(x)}\, f(y)$ is in $L^1(\mu\otimes\nu)$, as the product of $K \in L^2(\mu\otimes\nu)$ and $(x,y) \mapsto \overline{g(x)}\, f(y)$, a tensor product of functions in $L^2(\mu)$ and $L^2(\nu)$ and so in $L^2(\mu\otimes\nu)$. $\square$

1.10.2. *Positive operators.* We say that a linear self–operator $A : H \to H$ of a scalar product space is *positive* if $(Ax \mid x) \geq 0$ for every $x \in H$; *strictly positive* if $(Ax \mid x) > 0$ for $x \in H \setminus \{0\}$, *strongly positive* if there is $k > 0$ such that $(Ax \mid x) \geq k\,|x|^2$ for every $x \in H$ (caution: this terminology is far from universal; for instance, many call *monotone operators* what we called here positive).

1.10.3. *Self–adjoint operators.* If X is a Hilbert space, and $T \in L(X)$ is bounded operator of X into itself, it may happen that $T^* = T$, i.e that

$$(Tx \mid y) = (x \mid Ty) \quad \text{for every } x, y \in H;$$

in this case T is said to be self–adjoint, sometimes symmetric. If $S \in L(X, Y)$, then $S^* \circ S$ (the circle is often omitted) is self–adjoint and positive. Self–adjointness is immediate from $(S^* \circ S)^* = S^* \circ (S^*)^* = S^* \circ S$, and positivity is also immediate

$$(S^* \circ Sx \mid x) = (Sx \mid Sx) \geq 0 \quad \text{for every } x \in H;$$

and $S^* \circ S$ is strictly positive if and only if S is one–to–one.

Exercise 1.10.3.1. If $T : H \to H$ is bounded and H is a Hilbert space, then $T + T^*$ is self–adjoint; and if H is a $\mathbb{C}$–space, then also $i(T - T^*)$ is self–adjoint.

Exercise 1.10.3.2. A bounded operator $P : X \to X$ of a normed space X into itself is said to be a *projector* if it is idempotent, that is $P^2 = P \circ P = P$.

(i) Prove that if P is a projector then the image $P(X)$ and the kernel $\operatorname{Ker} P$ are closed subspaces of X, and $X = P(X) \oplus \operatorname{Ker} P$ (algebraically, and topologically if X is a Banach space).
(ii) Assuming now that X is a Hilbert space prove that P is self–adjoint if and only if P is an orthogonal projection, that is $P = \pi_{P(X)}$.

Solution. (i) If 1 =identity of X, then also $1 - P$ is a projector: in fact $(1 - P) \circ (1 - P) = 1 - P - P + P^2 = 1 - P$, and $\operatorname{Ker}(1 - P) = P(X)$ (if $x - Px = 0$ then $x = Px$, so that $x \in P(X)$, and $\operatorname{Ker}(1 - P) \subseteq P(X)$; and if $y = Px$ then by the idempotence of P we get $Py = P(Px) = Px = y$, so that $y - Py = 0$ and $y \in \operatorname{Ker}(1 - P)$ so that $P(X) \subseteq \operatorname{Ker}(1 - P)$). Then $P(X)$ is closed, being the null–space of the bounded linear operator $1 - P$. Next $x = Px + (x - Px)$, and $x - Px \in \operatorname{Ker} P$=(image of $1 - P$).

(ii) If $P = P^*$ then $\operatorname{Ker} P = (P^*(X))^\perp = (P(X))^\perp$, so that $P(X) = \operatorname{Ker}(1 - P)$ and $\operatorname{Ker} P$ are mutually orthogonal closed subspaces of H, and clearly then P is the orthogonal projection onto $P(X)$. If V is a closed subspace of H then π_V is self–adjoint: in fact, writing $x = \pi_V(x) + (x - \pi_V(x))$ and $y = \pi_V(y) + (y - \pi_V(y))$ we get

$$(\pi_V(x) \mid y) = (\pi_V(x) \mid \pi_V(y) + (y - \pi_V(y))) = (\pi_V(x) \mid \pi_V(y)) + (\pi_V(x) \mid y - \pi_V(y)) = (\pi_V(x) \mid \pi_V(y))$$

(recall that $y - \pi_V(y)$ and $\pi_V(x)$ are orthogonal) and in exactly the same way we get

$$(x \mid \pi_V(y)) = (\pi_V(x) \mid \pi_V(y));$$

thus

$$(\pi_V(x) \mid y) = (x \mid \pi_V(y)), \quad \text{for every } x, y \in X,$$

proving that $\pi_V^* = \pi_V$.

□

1.10.4. *Unitary operators.*

Exercise 1.10.4.1. Let X, Y be Hilbert spaces, and let $T \in L(X, Y)$. Prove that the following are equivalent:

(i) T is a norm preserving linear operator onto a (necessarily closed) subspace $T(X)$ of Y,
(ii) $(Tx \mid Ty)_Y = (x \mid y)_X$ for every $x, y \in X$.
(iii) T^* is a left inverse for T, i. e. $T^* \circ T$ is the identity of X.
(iv) T is injective and $T \circ T^*$ is an orthogonal projection in Y.

Moreover, if (iii) holds, then $T \circ T^* = \pi_{T(X)}$, orthogonal projection onto the image of T.

Solution. (i) implies (ii) Given $x, y \in X$ we have $|T(x + y)|^2 = |x + y|^2$, hence

$$|Tx|^2 + 2\operatorname{Re}(Tx \mid Ty) + |Ty|^2 = |x|^2 + 2\operatorname{Re}(x \mid y) + |y|^2 \Longrightarrow \operatorname{Re}(Tx \mid Ty) = \operatorname{Re}(x \mid y);$$

we have proved that (i) implies that the real part of the scalar product is preserved; and since $\operatorname{Re}(x \mid iy) = \operatorname{Im}(x \mid y)$ and $\operatorname{Re}(Tx \mid T(iy)) = \operatorname{Im}(Tx \mid Ty)$, the complex scalar product is preserved. (ii) implies (iii): from $(Tx \mid Ty) = (x \mid y)$ we get $(x \mid T^*(Ty)) = (x \mid y)$ for every $x, y \in X$ so that $T^*(Ty) = y$ for every

$y \in X$. (iii) implies (iv): since T has a left–inverse T is clearly injective; and $T \circ T^*$ obviously self–adjoint and is an idempotent if (iii) holds, since $(T \circ T^*) \circ (T \circ T^*) = T \circ (T^* \circ T) \circ T^* = T \circ 1_X \circ T^* = T \circ T^*$. By 1.10.3.2 $T \circ T^*$ is an orthogonal projection in Y. (iv) implies (i): since the kernel of $T \circ T^*$ is the kernel of T^*, the orthogonal of $T(X)$, $T \circ T^*$ has to be the orthogonal projection onto the closure of $T(X)$; but the image is clearly contained in $T(X)$, and so $T(X)$ is closed. Given $x, y \in X$ we have that $T \circ T^*$ is the identity on $T(X)$, so that $T(T^*(Tx)) = Tx$ for every $x \in X$ is equivalent, by injectivity of T, to $T^*(Tx) = x$ for every $x \in X$. And if $T^*(Tx) = x$ for every $x \in X$, we have $|x|^2 = (T^*(Tx) \mid x)$ for every $x \in X$, so that $(Tx \mid Tx) = (x \mid x)$ for every $x \in X$, and (i) holds. □

Exercise 1.10.4.2. In the space $\ell^2 = \ell^2(\mathbb{N}, \mathbb{C})$ the shift operator, or, more precisely, the *right shift* operator, is defined by $Sx = (0, x_0, x_1, \dots)$, i.e. $Sx(0) = 0$ and $Sx(n) = x(n-1)$ if $n \geq 1$. Clearly S is an isometry of ℓ^2 onto the orthogonal of e_0. Prove that the adjoint $T = S^*$ is the *left shift* $Tx(n) = x(n+1)$. Verify that ST is the orthogonal projection $x \mapsto x\,\chi$, where $\chi : \mathbb{N} \to \mathbb{C}$ is the characteristic function of $\mathbb{N}^>$ in $\mathbb{N}$, while TS is the identity of ℓ^2.

Definition. A norm preserving isomorphism of Hilbert spaces is called *unitary transformation.*

By what precedes, $T : X \to Y$ is unitary if and only if it is invertible and $T^{-1} = T^*$, i.e the inverse coincides with the adjoint. Among the most important examples of unitary transformations are the Fourier maps $f \mapsto (c_n(f))_{n\in\mathbb{Z}}$ from L^2_τ to $\ell^2(\mathbb{Z})$ and the Fourier–Plancherel transform $\Phi : L^2(\mathbb{R}) \to L^2(\mathbb{R})$.

1.10.5. *The Lax – Milgram lemma.*

. (Lax–Milgram) *Let X be a Hilbert space, and let $B : X \times X \to \mathbb{K}$ be a continuous sesquilinear mapping (=linear in the first variable, and conjugate linear in the second variable). If B is coercive, i.e. if there exists a constant $k > 0$ such that $|B(x,x)| \geq k\,|x|^2$ for every $x \in X$, then $y \mapsto B(\#, y)$ is a conjugate linear isomorphism of X onto X^*; and there exists a linear homemorphism $T \in \mathrm{Aut}(X)$ of X onto itself such that $B(x,y) = (x \mid T(y))$, for every x, $y \in X$.*

Proof. Given $y \in Y$, the mapping $x \mapsto B(x,y)$ is clearly an element of X^*, hence there is a unique $T(y) \in X$ such that $B(x,y) = (x \mid T(y))$ for every $x \in X$. Clearly the map $y \mapsto T(y)$ is linear (it is the composition of the conjugate linear map $y \mapsto B(\#, y)$ from X to X^* with the inverse of the Riesz isomorphism $y \mapsto (\# \mid y)$, also conjugate–linear). Moreover, by coercivity of B:

$$|(y \mid T(y))| = |B(y,\, y)| \geq k\,|y|^2 \quad \text{so that} \quad |y|\,|T(y)| \geq |(y \mid T(y))| \geq k\,|y|^2,$$

which implies

$$|T(y)| \geq k\,|y| \quad \text{for every} \quad y \in X.$$

From 1.8.0.3, we get that T is a linear homeomorphism of X onto $T(X)$, with $T(X)$ a closed subspace of X. If $z \in T(X)^\perp$ we get $0 = (z \mid T(z)) = |(z \mid T(z))| = |B(z,z)| \geq k\,|z|^2$, which implies $z = 0$. Then $T(X)^\perp = \{0\}$, and since $T(X)$ is closed we have $T(X) = X$. □

1.10.6. *The Cauchy–Schwarz inequality in the semidefinite case.* For future use we prove the Cauchy–Schwarz inequality for a semidefinite hermitian product. Recall that a sesquilinear form is said to be *hermitian* if $B(y,\,x) = \overline{B(x,y))}$ for every x, $y \in X$, and positive if $B(x,x) \geq 0$ for every $x \in X$. Then we have, for x, $y \in X$:

. Cauchy–Schwarz inequality *Let $B : X \times X \to \mathbb{K}$ be a positive hermitian form. Then, for every x, $y \in X$ we have*

$$|B(x,y)| \leq \sqrt{B(x,\,x)}\,\sqrt{B(y,\,y)}.$$

Proof. Given x, $y \in X$ and $t \in \mathbb{R}$ consider $x + t\,y$; then, for every $t \in \mathbb{R}$:

$$0 \leq B(x + t\,y, x + t\,y) = B(x,x) + 2t \operatorname{Re} B(x,y) + t^2\, B(y,y);$$

If $B(y,y) = 0$ positivity for all $t \in \mathbb{R}$ of this expression forces $\operatorname{Re} B(x,y) = 0$; if $B(y,y) > 0$ the expression is positive for all $t \in \mathbb{R}$ iff $(\operatorname{Re} B(x,y))^2 - B(y,y)\,B(x,x) \geq 0$, i.e. iff

$$|\operatorname{Re} B(x,y)| \leq \sqrt{B(x,\,x)}\,\sqrt{B(y,\,y)};$$

If $B(x,y) = 0$ the proof is achieved; if $B(x,y) \neq 0$, let $\alpha = \operatorname{sgn} B(x,y)$; since $B(x, \alpha\,y) = \bar\alpha\,B(x,y) = |B(x,y)|$ the preceding inequality gives

$$|\operatorname{Re} B(x, \alpha\,y)| \leq \sqrt{B(x,\,x)}\,\sqrt{B(\alpha\,y,\,\alpha\,y)} = \sqrt{B(x,\,x)}\,\sqrt{|\alpha|^2\,B(y,\,y)} = \sqrt{B(x,\,x)}\,\sqrt{B(y,\,y)},$$

and the left–hand side is

$$|\operatorname{Re} B(x, \alpha\,y)| = |\bar\alpha\,B(x,y)| = |B(x,y)|;$$

the proof is concluded. □

We call *isotropic* a vector $x \in X$ such that $B(x,x) = 0$; the inequality shows that isotropic vectors are orthogonal to every vector in the space, that is $B(x, y) = 0$ for every $y \in X$ if x is isotropic; moreover isotropic vectors are a linear subspace:

$$B(x+y, x+y) = B(x,x) + B(y,y) = 0$$

if $B(x,x) = B(y,y) = 0$

EXERCISE 1.10.6.1. Prove that the Cauchy–Schwarz inequality is an equality if and only if x and y are linearly dependent modulo the subspace of isotropic vectors.

1.11. Radon measures. How can the dual space of the space $(C(X), \|\#\|_u)$ be described, if X is a compact Hausdorff space? It turns out that this dual is the space of all *Radon measures* on X, which we shall describe. We do not give proofs here, and refer the reader to [Folland] or [Rudin] for this.

1.11.1. *Finite* $\mathbb{K}-$*valued measures.* If $(X, \mathcal{M})$ is a measurable space, that is, $\mathcal{M}$ is a $\sigma-$subalgebra of $\mathcal{P}(X)$, a finite $\mathbb{K}-$valued measure on $\mathcal{M}$ is a countably additive function $\mu : \mathcal{M} \to \mathbb{K}$. We shall drop the adjective finite unless needed for emphasis or to avoid ambiguity. When $\mathbb{K} = \mathbb{R}$ one also speaks of a *finite signed measure, or charge* (signed measures are often allowed to take $-\infty$ or $+\infty$ as values, we exclude this here). If μ is a $\mathbb{K}-$valued measure, there is a minimal positive measure among those which dominate μ: i.e. taking the set of all positive measures $\nu : \mathcal{M} \to [0, +\infty]$ such that $|\mu(E)| \le \nu(E)$ holds for every $E \in \mathcal{M}$, there is a smallest such measure, called *total variation* of the $\mathbb{K}-$valued measure μ, which is also defined, for every $E \in \mathcal{M}$, as

$$|\mu|(E) = \sup\left\{\sum_{k=1}^{m} |\mu(E_k)| : \text{ with } (E_k)_{1\le k\le m} \text{ ranging in the set of finite } \mathcal{M}\text{−partitions of } E\right\}.$$

It can be proved that $|\mu|$ is actually finite, $|\mu|(X) < \infty$. Moreover, if $\lambda : \mathcal{M} \to [0, +\infty]$ is a positive measure and $\rho \in \mathcal{L}^1_\lambda(X, \mathbb{K})$, then $d\mu = \rho\, d\lambda$ is a finite $\mathbb{K}-$valued measure, and $d|\mu| = |\rho|\, d\lambda$. In any case there exists a function $\sigma \in \mathcal{L}^1(|\mu|)$, with $|\sigma(x)| = 1$ for every $x \in X$, such that

$$\frac{d\mu}{d|\mu|} = \sigma, \quad \text{equivalently, for every } E \in \mathcal{M}, \quad \mu(E) = \int_E \sigma\, d|\mu|;$$

all this can be found in [Folland]. For a signed finite measure $\mu : \mathcal{M} \to \mathbb{R}$ we can define the *positive and negative parts* of μ:

$$\mu^+ := \frac{|\mu| + \mu}{2}; \qquad \mu^- := \frac{|\mu| - \mu}{2},$$

so that μ^+ and μ^- are both positive measures and

$$\mu = \mu^+ - \mu^-; \quad |\mu| = \mu^+ + \mu^-.$$

It is clear that any finite linear combination with coefficients in $\mathbb{K}$ of $\mathbb{K}-$valued measures on $\mathcal{M}$ is still a $\mathbb{K}-$valued measure on $\mathcal{M}$, so that these measure are naturally a $\mathbb{K}-$linear space. We have

. *For every measurable space* $(X, \mathcal{M})$ *the set* $M(X) = M(X, \mathcal{M})$ *of all finite* $\mathbb{K}-$*valued measures on* $\mathcal{M}$ *is a Banach space under the norm* $\|\mu\| = |\mu|(X)$.

Proof. If $\mu, \nu \in M(X, \mathcal{M})$ then

$$|(\mu + \nu)(E)| \le |\mu(E)| + |\nu(E)| \le |\mu|(E) + |\nu|(E) = (|\mu| + |\nu|)(E)$$

for every $E \in \mathcal{M}$; by minimality we get $|\mu + \nu| \le |\mu| + |\nu|$, in particular for $E = X$, we have

$$|\mu + \nu(X)| \le (|\mu| + |\nu|)(X) = |\mu|(X) + |\nu|(X),$$

so that $\mu \mapsto |\mu|(X)$ is subadditive and hence a norm (the remaining conditions are trivial). Completeness: assume that $\sum_{n\in\mathbb{N}} \mu_n$ is a normally convergent series of $\mathbb{K}-$valued measures, that is, the series $\sum_{n\in\mathbb{N}} \|\mu_n\| = \sum_{n\in\mathbb{N}} |\mu_n|(X)$ is convergent. Given $E \in \mathcal{M}$ we set

$$\mu(E) := \sum_{n\in\mathbb{N}} \mu_n(E)$$

(note that this series is absolutely convergent, since $|\mu_n(E)| \le |\mu_n|(E) \le |\mu_n|(X) = \|\mu_n\|$, so that the definition makes sense), and we have to prove that μ is countably additive. If $E = \bigcup_{m\in\mathbb{N}} E_m$, with the $E_m \in \mathcal{M}$ pairwise disjoint, we get

$$\mu(E) = \sum_{n\in\mathbb{N}} \mu_n(E) = \sum_{n\in\mathbb{N}} \left(\sum_{m\in\mathbb{N}} \mu_n(E_m)\right) =$$

we interchange the sums, proving later the admissibility

$$\sum_{m\in\mathbb{N}}\left(\sum_{n\in N}\mu_n(E_m)\right)=\sum_{m\in\mathbb{N}}\mu(E_m),$$

as required. The interchange is admissible because the sum of absolute values:

$$\sum_{n\in\mathbb{N}}\left(\sum_{m\in\mathbb{N}}|\mu_n(E_m)|\right)\le\sum_{n\in\mathbb{N}}\left(\sum_{m\in\mathbb{N}}|\mu_n|(E_m)\right)=\sum_{n\in\mathbb{N}}|\mu_n|(E)\le\sum_{n\in\mathbb{N}}|\mu_n|(X)<\infty$$

is finite; we can apply Tonelli's theorem in the space $\ell^1(\mathbb{N}\times\mathbb{N})$, or the more general results of 0.1.6. Thus μ is $\mathbb{K}-$valued measure and since we have

$$\left|\mu(E)-\sum_{n=0}^{m}\mu_n(E)\right|\le\sum_{n=m+1}^{\infty}|\mu_n(E)|\le\sum_{n=m+1}^{\infty}|\mu_n|(E),$$

we have (remember that the sum of a series of positive measures is always a (not necessarily finite) positive measure):

$$\left|\mu-\sum_{n=0}^{m}\mu_n\right|(E)\le\sum_{n=m+1}^{\infty}|\mu_n|(E),$$

in particular, for $E=X$, we get

$$\left\|\mu-\sum_{n=1}^{m}\mu_n\right\|\le\sum_{n=m+1}^{\infty}\|\mu_n\|,$$

and the proof ends. □

Finally, for every $\mathbb{K}-$valued measure on $\mathcal{M}$ we define $L^1(\mu):=L^1(|\mu|)$ and set, if $d\mu=\sigma\,d|\mu|$:

$$\int_X f\,d\mu:=\int_X f\,\sigma\,d|\mu|;$$

it can be proved that if $d\mu=\rho\,d\lambda$, with λ a positive measure and $\rho\in\mathcal{L}^1(\lambda)$, then $f\in\mathcal{L}^1(\mu)$ iff $f\rho\in\mathcal{L}^1(\lambda)$ and moreover

$$\int_X f\,d\mu=\int_X f\,\rho\,d\lambda.$$

Of course we have the fundamental inequality:

$$\left|\int_X f\,d\mu\right|\le\int_X|f|\,d|\mu|.$$

And if $\mu,\nu:\mathcal{M}\to\mathbb{K}$ are $\mathbb{K}-$valued measures, then since $|\mu+\nu|\le|\mu|+|\nu|$ and $\mathcal{L}^1(|\mu|+|\nu|)=\mathcal{L}^1(|\mu|)\cap\mathcal{L}^1(|\nu|)$ we have $\mathcal{L}^1(|\mu|)\cap\mathcal{L}^1(|\nu|)\subseteq\mathcal{L}^1(|\mu+\nu|)$ and

$$\int_X f\,d(\mu+\nu)=\int_X f\,d\mu+\int_X f\,d\nu$$

for every $f\in\mathcal{L}^1(|\mu|)\cap\mathcal{L}^1(|\nu|)$.

1.11.2. *Locally compact spaces.* We call *locally compact* a Hausdorff topological space in which every point has a compact neighborhood. We accept the following fact, whose proof is at the end of the section:

. *In a locally compact space every point has a neighborhood base consisting of compact neighborhoods.*

Every compact Hausdorff space is of course locally compact. If a space is locally compact but not compact, it can be *compactified* by adding just one point, i.e. the construction made for $\mathbb{R}^n$ by adding a point at infinity can be repeated for a locally compact non–compact space X, obtaining a compact Hausdorff space αX of which X is an open dense subspace: αX is the *one–point compactification*, or Alexandroff compactification, of X. Simply take an object $\infty=\infty_X$ not in X, consider the set $\alpha X=X\cup\{\infty\}$, and declare open in this set all the open subsets of X, and all the complements in αX of the compact subsets of X. It is not difficult to prove that αX is compact Hausdorff and that X is open and dense in αX (accept it; the curious reader can see at the end of the section, better still try to get a proof, not at all difficult). By $C_c(X)=C_c(X,\mathbb{K})$ (sometimes also $C_{00}(X)$) we denote the set of continuous $\mathbb{K}-$valued functions f from X to $\mathbb{K}$ whose support

$$\mathrm{Supp}(f):=\mathrm{cl}_X\{x\in X:\ f(x)\ne 0\}$$

is a compact subset of X. Clearly, if X is compact then $C_c(X) = C(X)$. By $C_0(X) = C_0(X, \mathbb{K})$ for X locally compact non compact we denote the set of all continuous $\mathbb{K}-$valued functions on X whose limit at infinity is zero:

$$C_0(X) = \{f \in C(X) : \lim_{x\to\infty} f(x) = 0\};$$

for X compact we set $C_0(X) = C(X)$.

. *All functions in $C_0(X)$ are bounded, and their modulus has a maximum. In fact a continuous $f \in C(X)$ is in $C_0(X)$ if and only if for every $\varepsilon > 0$ the set $\{x \in X : |f(x)| \geq \varepsilon\}$ is compact.*

Proof. The limit at infinity is zero if and only if, given $\varepsilon > 0$ there is a nbhd U_ε of ∞, which we may assume open, such that $|f(x)| < \varepsilon$ for $x \in U_\varepsilon \setminus \{\infty\}$; since $U_\varepsilon \setminus \{\infty\} = X \setminus K_\varepsilon$, with K_ε a compact subset of X, we have that $\{x \in X : |f(x)| \geq \varepsilon\} \subseteq K_\varepsilon$; since $|f|$ is continuous, the set $\{x \in X : |f(x)| \geq \varepsilon\}$ is closed, hence compact if contained in the compact set K_ε. Then, if $f \in C_0(X)$ is not identically zero, simply pick $c \in X$ such that $f(c) \neq 0$ and consider $K = \{x \in X : |f(x)| \geq |f(c)|\}$: this set is compact, and $\max\{|f(x)| : x \in K\}$ is clearly $\|f\|_u = \max\{|f(x)| : x \in X\}$. □

Clearly $C_c(X) \subseteq C_0(X)$: in this context the functions of $C_c(X)$ are those which are identically zero on a punctured nbhd of ∞. Trivially $C_c(X)$ and $C_0(X)$ are linear subspaces and also subalgebras and sublattices of $C_b(X)$, space of all continuous bounded functions on X.

. *$C_0(X)$ is the uniform closure of $C_c(X)$.*

Proof. Assume that $f \in C_b(X) \setminus C_0(X)$; then there exists $\varepsilon > 0$ such that $F = \{x \in X : |f(x)| \geq \varepsilon\}$ is non compact. If $\|g - f\|_u \leq \varepsilon/2$ we have $\{x \in X : |g(x)| \geq \varepsilon/2\} \supseteq F$, so that $\{x \in X : |g(x)| \geq \varepsilon/2\}$ is non–compact, hence $g \notin C_0(X)$. We have proved that $B(f, \varepsilon/2[\subseteq C_b(X) \setminus C_0(X)$; and the set $C_b(X) \setminus C_0(X)$ is then open in $C_b(X)$. Next we prove that $C_c(X)$ is dense in $C_0(X)$. Assume that $f \in C_0(X)$ is real and positive, $f(x) \geq 0$ for every $x \in X$. Given $\varepsilon > 0$ let $f_\varepsilon = (f - \varepsilon) \vee 0 = (f - \varepsilon)^+$, positive part of the function $f - \varepsilon$. Then $f_\varepsilon \in C_c(X)$: in fact $\{x \in X : f_\varepsilon(x) \neq 0\} = \{x \in X : f_\varepsilon(x) > 0\} \subseteq \{x \in X : f(x) \geq \varepsilon\}$; this latter set is compact, hence closed. so that $\mathrm{Supp}(f_\varepsilon) \subseteq \{x \in X : f(x) \geq \varepsilon\}$ is also compact. And $0 \leq f - f_\varepsilon \leq \varepsilon$, trivially (if $f(x) > \varepsilon$ then $f_\varepsilon(x) = f(x) - \varepsilon$ and $f(x) - f_\varepsilon(x) = \varepsilon$, if $f(x) \leq \varepsilon$ then $f_\varepsilon(x) = 0$ and $f(x) - f_\varepsilon(x) = f(x) \leq \varepsilon$) so that $\|f - f_\varepsilon\|_u \leq \varepsilon$. For $f \in C_0(X)$ arbitrary we can write $f = (\mathrm{Re}\, f)^+ - (\mathrm{Re}\, f)^- + i((\mathrm{Im}\, f)^+ - (\mathrm{Im}\, f)^-)$ and approximate each of the four functions $(\mathrm{Re}\, f)^\pm$, $(\mathrm{Im}\, f)^\pm$. □

1.11.3. *Some proofs.*

- We first prove that in a locally compact space at every point there is a nbhd base consisting of compact nbhds, that is, given $p \in X$ and an open nbhd U of p, U contains a compact nbhd V of p. We know that p has a compact nbhd W, and it is not restrictive to assume that $U \subseteq W$. We assume as known the following result:

  . *Let X be a Hausdorff space, let K be a compact subset of X, and p a point of X not in K. Then p and K have disjoint nbhds in X, that is, there are disjoint open subsets A, B of X such that $p \in A$ and $K \subseteq B$.*

  Proof. Analisi Due, 2.15.9. □

  With p, U, W as above let $K = W \setminus U$; it is a closed subset of W, hence compact, and $p \notin K$ since $p \in U$. Then there are disjoint open sets A and B of X with $p \in A$ and $K \subseteq B$. Then $C = A \cap U$ is an open subset of W containing p and disjoint from B. It follows that $V = \mathrm{cl}_X(C) \subseteq (X \setminus B) \cap W$ is a closed nbhd of p contained in W; then V is compact; and V is disjoint from K, hence contained in U. □
- Let's prove that αX is compact and Hausdorff. Compactness: let $(A_i)_{i\in I} \cup (\alpha X \setminus K_j)_{j\in J}$ be an open cover of αX. Then ∞ is covered by some $\alpha X \setminus K_0$ with $0 \in J$; and then

  $$K_0 \subseteq \left(\bigcup_{i\in I} A_i\right) \cup \left(\bigcup_{j\in J} X \setminus K_j\right);$$

  since K_0 is by hypothesis a compact subset of X, the open cover of K_0 in the above formula has a finite subcover, that is, there are finite subsets F of I and G of J such that

  $$K_0 \subseteq \left(\bigcup_{i\in F} A_i\right) \cup \left(\bigcup_{j\in G} X \setminus K_j\right);$$

  and $(A_i)_{i\in F} \cup (\alpha X \setminus K_j)_{j\in G\cup\{0\}}$ is a finite subcover of the original open cover.

αX is a Hausdorff space; the space X is an open subspace, by hypothesis Hausdorff, so pair of points of X have disjoint nbhds in X, which are also nbhds in αX. And given $x \in X$, pick a compact nbhd W of x in X; then W and $\alpha X \setminus W$ are disjoint nbhds of x and ∞ in $\alpha\, X$. □

EXERCISE 1.11.3.1. Prove that if X is locally compact then all open, and all closed subspaces of X are locally compact in the induced topology. Deduce from it that if $X \cup \{*\}$, a space obtained by adding to X a singleton $\{*\}$, is locally compact, then X is also locally compact; thus the only spaces with a Hausdorff one–point– compactification are the locally compact ones. But not all subspaces are locally compact: the subspace $\mathbb{Q}$ of $\mathbb{R}$ consisting of the rational numbers is not locally compact: the only open subset of $\mathbb{Q}$ contained in a compact subset of $\mathbb{Q}$ is the empty set.

1.11.4. *The dual of* $C_0(X)$. Having a finite $\mathbb{K}-$ valued measure μ on a $\sigma-$algebra $\mathcal{M}$ of subsets of X, integration with respect to this measure clearly gives a bounded linear functional I_μ on the Banach space $L^\infty(X) = L(X) \cap \ell^\infty(X)$ of all bounded $\mathcal{M}-$measurable functions:

$$I_\mu(f) = \int_X f\, d\mu \quad \text{bounded because} \quad |I_\mu(f)| \le \int_X |f|\, d|\mu| \le \int_X \|f\|_u\, d|\mu| = \|f\|_u\, |\mu|(X),$$

so that $\|I_\mu\| \le |\mu|(X)$. Having a topological space (X, τ) we can consider on X the $\sigma-$algebra $\mathcal{B}(X)$ of all Borel subsets of X; given any finite $\mathbb{K}-$valued Borel measure $\mu : \mathcal{B}(X) \to \mathbb{K}$, integration with respect to this measure will give a bounded functional $I_\mu : C_b(X) \to \mathbb{K}$, since all continuous functions are Borel measurable. In this way we get a linear map $\mu \mapsto \varphi_\mu$ from the Banach space of finite Borel measures to the dual of $C_b(X)$; this map is norm–reducing, i.e. $\|I_\mu\| \le \|\mu\| := |\mu|(X)$. Of course we immediately ask if this map is an isomorphism, and if it is norm–preserving. In general this map is neither injective nor surjective; the question is quite difficult, and involves a lot of non–constructive processes. We restrict ourselves to locally compact Hausdorff space, and to the subspace $C_0(X)$ of functions which are zero at infinity (which of course contains the important case of X compact Hausdorff). In this case the mapping is surjective, but not injective (in some pathological cases). To get injectivity we must restrict the measure from the Borel case to a better–behaved type of measure, the *Radon measure*. We first define the positive Radon measures, which are allowed to take also the value $+\infty$

DEFINITION. A positive Radon measure $\mu : \mathcal{B}(X) \to [0, +\infty]$ on the locally compact space X is a positive measure on the $\sigma-$algebra of all Borel subsets of X such that

(0) μ is finite on the compact subsets of X.

(i) μ is outer regular at all measurable subsets, meaning that for every $E \in \mathcal{B}(X)$ we have

$$\mu(E) = \inf_{\tilde{\mathbb{R}}}\{\mu(U) : U \text{ open}, U \supseteq E\}.$$

(ii) μ is inner regular at all open subsets of X. That is, if U is open in X then

$$\mu(U) = \sup_{\tilde{\mathbb{R}}}\{\mu(K),\, K \text{ compact},\, K \subseteq U\}.$$

It can be proved that Lebesgue measure on $\mathbb{R}^n$ is a positive Radon measure, in fact also:

. *If a locally compact space is metrizable and separable, then every positive Borel measure on it which is finite on compact subsets is a Radon measure.*

Proof. Omitted for now. □

Next, a finite $\mathbb{K}-$valued measure on a locally compact space X is said to be a *Radon measure* if its total variation is a positive Radon measure.

It can be proved that finite Radon measures are a closed subspace of the Banach space of all finite Borel measures on the locally compact space X; we accept this result. Given the locally compact space X, we denote by $M(X) = M(X, \mathbb{K})$ the space of all finite $\mathbb{K}-$valued Radon measures on X (normed by $\|\mu\| = |\mu|(X)$). Then:

. *If X is a locally compact space, the mapping $\mu \mapsto I_\mu$, where $I_\mu : C_0(X) \to \mathbb{K}$ is integration with respect to μ, i.e.*

$$I_\mu(f) = \int_X f\, d\mu$$

is a norm– preserving isomorphism of the space $M(X)$ of all finite $\mathbb{K}-$valued Radon measures on X onto the dual $(C_0(X))^*$ of the space of continuous functions on X which vanish at infinity.

We omit the proof, long and quite demanding.

1.11.5. *An example.* We consider the space $C(\mathbb{U},\mathbb{C})$ with uniform norm, where $\mathbb{U}=\{z\in\mathbb{C}:\ |z|=1\}$ is the unit circle. We want to prove that for every point $e^{2\pi ix}$ of the circle there is a non–empty subset S_x of $C(\mathbb{U})$, in fact even a non–meager subset S_x of $C(\mathbb{U})$, such that the Fourier series of every $f\in S_x$ does not converge at $e^{2\pi ix}$.We have the linear functionals $S_mf(x):C(\mathbb{U})\to\mathbb{C}$ (as usual we identify functions on the circle with periodic functions on the real line, choosing here period 1) defined by:

$$S_mf(x)=\sum_{n=-m}^{m}c_n(f)\,e^{2\pi inx}=\int_{-1/2}^{1/2}f(x+t)\,D_m(2\pi t)\,dt;$$

it is not restrictive to assume that $x=0$; we then have the sequence of functionals $\varphi_m:C(\mathbb{U})\to\mathbb{C}$ given by

$$\varphi_m(f)=\int_{-1/2}^{1/2}f(t)\,D_m(2\pi t)\,dt,\quad \text{with norm } \|\varphi_m\|=\|D_m\|_1=\int_{-1/2}^{1/2}|D_m(2\pi t)|\,dt.$$

If $\lim_{m\to\infty}\varphi_m(f)$ esists in $\mathbb{C}$ for a non–meager subset of $C(\mathbb{U})$ the uniform boundedness theorem implies that $\|D_m\|_1$ is bounded, but we have proved in 1.8.3 that $\lim_{m\to\infty}\|D_m\|_1=\infty$.

2. Topological vector spaces

2.1. All the linear(=vector) spaces considered have as field of scalars either the field $\mathbb{R}$ of the reals, or the field $\mathbb{C}$ of the complexes; the letter $\mathbb{K}$ denotes either of these fields. The field $\mathbb{K}$ is considered with its usual topology, which makes it a topological field, meaning that the field operations are continuous. We assume the notion of $\mathbb{K}-$vector space as well-known, as well as the notion of linear mapping between vector spaces, of linear dependence, etc. We give here the notion of *topological vector space*, a $\mathbb{K}-$linear space equipped with a topology compatible with the linear space structure. Explicitly:

2.1.1.

Definition. A *topological vector space* (briefly TVS, or LTS, linear topological space) is an ordered pair (X, τ) consisting of a $\mathbb{K}-$vector space X and a topology on X such that:

(i) The addition $\sigma : (x, y) \mapsto x + y$ is a continuous map from $X \times X$, equipped with the product topology, into X.

(ii) The multiplication of scalars and vectors $\mu : (\lambda, x) \mapsto \lambda\, x$ is continuous from $\mathbb{K} \times X$ to X (it is understood that $\mathbb{K} \times X$ is equipped with the product topology of the usual topology of $\mathbb{K}$ and the topology τ given on X).

2.1.2. The topology τ is not necessarily a separated (=Hausdorff) topology. Trivially, the indiscrete topology is a vector topology. At the opposite extreme it is easy to show that the discrete topology is *never* a vector topology, barring the trivial case of the space consisting only of the zero element: in fact multiplication $(\alpha,\, x) \to \alpha\, x$ is not continuous, at no point of $\mathbb{K} \times X$ (prove it).

We are interested only in certain features of the extensive TVS theory, mainly, but not only, in the subclass of normed spaces, assumed already known in some degree to the reader. However we develop some easy generalities. The first elementary properties are

. *In a TVS translations and homotheties are self–homeomorphisms of the space. That is, for every given $a \in X$ the mapping $x \mapsto a + x$ is a self–homeomorphism of X. And for every given nonzero $\alpha \in \mathbb{K}$ the mapping $x \mapsto \alpha\, x$ is also a self–homeomorphism of X.*

Proof. The translation $x \mapsto a + x$ may be written as the composition map of the map $x \mapsto (a, x)$ of X into $X \times X$, clearly continuous because its components are the constant function $x \mapsto a$ and the identity $x \mapsto x$, with the addition map $(x, y) \mapsto x + y$. Hence the translation is continuous; and it is a homeomorphism because its inverse is another translation, the map $x \mapsto -a + x$.

A similar argument for homotheties: the map $x \mapsto (\alpha, x)$ is continuous from X into $\mathbb{K} \times X$, so its composition with multiplication is continuous; hence $x \mapsto \alpha\, x$ is continuous, and a homeomorphism because its inverse is the homothety $x \mapsto \alpha^{-1}\, x$, also continuous. □

2.1.3. *Continuity of linear maps.* The fact that translations are homeomorphisms implies that TVS's are *homogeneous spaces*, they look the same at every point, and this allows the localization at 0 of many local notions. Of course, given $a, c \in X$ and $\alpha \in \mathbb{K} \setminus \{0\}$, maps like $x \mapsto a + \alpha\,(x - c)$ are also self–homeomorphisms of X; and if U is a neighborhood of 0 in X then $a + \alpha\, U$ is a neighborhood of $a \in X$, for any $\alpha \in \mathbb{K} \setminus \{0\}$. Another illustration of this phenomenon is:

. *Let X, Y be topological vector spaces, and let $T : X \to Y$ be a linear map. Then T is continuous if and only if it is continuous at* 0.

Proof. If T is continuous at 0, then for every nbhd V of 0 in Y there exists a nbhd U of 0 in X such that $T(U) \subseteq V$ (recall that $T(0) = 0$, since T is linear!). Given $a \in X$, we prove that T is continuous at a: every nbhd of $T(a)$ in Y is of the form $T(a) + V$, with V a nbhd of 0 in Y; taking the nbhd U of 0 in X as above, such that $T(U) \subseteq V$, we have that $a + U$ is nbhd of a in X and $T(a+U) = T(a) + T(U) \subseteq T(a) + V$, thus concluding the proof. □

Moreover the antipodal map $x \mapsto -x$ is a homeomorphism, so that if U is a neighborhood of 0 then $-U$ is also a neighborhood of 0, and hence $U \cap (-U)$ is a neighborhood of 0, contained in U. Symmetric neighborhoods of 0 are then a basis for the neighborhood system. But much more is true: call D the open unit ball of $\mathbb{K}$, that is $D = \{\alpha \in \mathbb{K} : |\alpha| < 1\}$, and let B be its closure, that is $B = \{\alpha \in \mathbb{K} : |\alpha| \le 1\}$, so that $D/\ B$ are respectively the open/closed interval $]-1, 1[/[-1, 1]$ if $\mathbb{K} = \mathbb{R}$, the open/closed unit disc in the case $\mathbb{K} = \mathbb{C}$. Call a subset U of X *balanced* if $B\,U \subseteq U$ (we set $B\,U := \{\alpha\, x : \ \alpha \in B,\, x \in U\}$). Verify that in a real linear space U is balanced if and only if $x \in U$ implies that $[x, -x] \subseteq U$, while if $\mathbb{K} = \mathbb{C}$ the subset U is balanced if $x \in U$ implies that U contains also the whole "disc" $B\,x$. For every subset S of X the set $B\,S$ is balanced: this is due to the fact that $\beta\,B \subseteq B$ for every $\beta \in B$ (that is, to the fact that B is a *subsemigroup* of the multiplicative semigroup of $\mathbb{K}$); and clearly $B\,S$ is the smallest balanced subset of X containing S. Then:

2.1.4.

. *In a TVS every neighborhood of* 0 *contains a balanced neighborhood of* 0.

Proof. This is a simple consequence of the continuity of $\mu(\alpha,\,x) = \alpha\,x$ at $(0,0)$: given U we pick V, nbhd of 0 in X, and $\delta > 0$ such that $\mu((\delta\,B)\times V) = \delta B\,V \subseteq U(= B(\delta\,V))$. Then $W = B(\delta\,V)$ is contained in U, is a nbhd of 0 (it contains $\delta\,V$, which is a nbhd of 0) and is balanced, as observed above. □

Exercise 2.1.4.1. Observe that the image of a balanced set by a linear map is also balanced. Prove that if U is balanced then $\alpha\,U = U$ for every α with $|\alpha| = 1$.

2.1.5. *Balanced subsets of* $\mathbb{K}$. It is easy to determine all balanced (non–empty) subsets of $\mathbb{K}$: they are discs centered at the origin, either open or closed, or the entire field $\mathbb{K}$. As observed above, if $E \subseteq \mathbb{K}$ is balanced, and $\beta \in E$, then E contains the whole closed disc (interval in the case of $\mathbb{R}$) of radius $|\beta|$ (explicitly: if $|\alpha| \le |\beta|$, then $\alpha = (\alpha/\beta)\beta$, with $|\alpha|/|\beta| \le 1$). Let $R = \sup_{\tilde{\mathbb{R}}}\{|\beta| : \beta \in E\}$. If $R = +\infty$, equivalently, if E is unbounded, then $E = \mathbb{K}$; if R is finite then $E = R\,D$ (in case $R \notin E$) or $E = R\,B$ (in case $R \in E$). An easy corollary is:

Exercise 2.1.5.1. There are exactly two vector topologies on $\mathbb{K}$: the indiscrete and the usual.

2.1.6. *Continuous linear functionals.* We are now in a position to prove the following characterization of continuous linear functionals on a TVS:

Exercise 2.1.6.1.

. *Let X be a topological vector space, and let $f : X \to \mathbb{K}$ be linear; then the following are equivalent:*

(i) *f is continuous.*
(ii) *f is continuous at* 0
(iii) *The nullspace* $\mathrm{Ker}(f) = f^{\leftarrow}(0)$ *is closed in X.*
(iv) *Some non empty open subset of X is mapped by f into a proper subset of* $\mathbb{K}$.

Solution. That (i) and (ii) are equivalent is 2.1.3. (i) implies (iii): $\{0\}$ is closed in $\mathbb{K}$ so its inverse image by the continuous f is closed in X. (iii) implies (iv): this is trivially true for $f = 0$, and if $f \ne 0$ then $X \smallsetminus \mathrm{Ker}(f)$ is a nonempty open set mapped by f onto $\mathbb{K} \smallsetminus \{0\}$. (iv) implies (i). Let A be open nonempty and $f(A) \ne \mathbb{K}$. Since A is open nonempty it contains a subset of the form $a + U$, where U is a balanced nbhd of the origin. Then $f(a+U) = f(a) + f(U) \ne \mathbb{K}$, so that also $f(U) \ne \mathbb{K}$, and $f(U)$ is balanced. By the previous result, we have that $f(U) \subseteq R\,B$ for some $R > 0$. Given now $\varepsilon > 0$ we have $f((\varepsilon/R)U) = (\varepsilon/R)\,f(U) \subseteq (\varepsilon/R)\,R\,B = \varepsilon\,B$, proving that f is continuous at the origin. □

2.1.7. Continuity of vector operations implies also that:

. *For any given $a \in X$ the mapping $\xi \mapsto \xi\,a$ is continuous, from $\mathbb{K}$ to X.*

Proof. It may be written as the composition of $\xi \mapsto (\xi,\,a)$, continuous from $\mathbb{K}$ to $\mathbb{K}\times X$, with the multiplication scalars×vectors $\mu : \mathbb{K}\times X \to X$.

□

Notice that continuity implies: for every $a \in X$ and every neighborhood U of 0 in X there is $\delta = \delta(a) > 0$ such that $\alpha\,a \in U$ if $|\alpha| \le \delta$.

If X is Hausdorff, then $\{0\}$ is a closed subspace of $\mathbb{K}\,a$ so that the linear functional $\xi\,a \mapsto \xi$ is continuous, by the preceding exercise, so that these maps are homeomorphisms.

2.1.8. *The closure of a vector subspace is a vector subspace.*

. *In a TVS, the closure of a linear subspace is also a linear subspace.*

Proof. Recall that a subset V of a linear space X is a subspace if and only if it is non empty, and $V + V \subseteq V$, $\mathbb{K}\,V \subseteq V$. The closure of $V \times V$ in $X \times X$ is $\bar{V}\times\bar{V}$, the cartesian product of the closures, and similarly the closure of $\mathbb{K}\times V$ in $\mathbb{K}\times X$ is $\mathbb{K}\times\bar{V}$. Since V is a subspace, the addition map σ maps $V\times V$ into V, i.e $\sigma(V) = V + V \subseteq V$; recall now that for continuous maps the image of the closure of any subset of the domain is contained in the closure of the image. Hence $\bar{V} + \bar{V} := \sigma(\bar{V}\times\bar{V}) \subseteq \bar{V}$, meaning that $\bar{V}$ is closed under addition. In the same way one proves that $\bar{V}$ is closed under multiplication by scalars. □

In particular the closure N of the trivial subspace $\{0\}$ is also a linear subspace. The topology induced on N is discrete: if an open subset U of X contains 0 then it contains all of N: in fact, pick a symmetric nbhd V of 0 contained in U; if $x \in N$, then $0 \in x + V$ (if x in the closure of $\{0\}$, every neighborhood of x must contain 0); then $x \in -V = V \subseteq U$, so that $N \subseteq U$.

EXERCISE 2.1.8.1. Let X be a TVS. Then the following are equivalent

(i) X is Hausdorff (in other words, X is a T_2 topological space).
(ii) Singletons of X are closed sets (in other words, X is a T_1 topological space).
(iii) $\{0\}$ is closed in X.
(iv) The intersection of all nbhds of 0 is $\{0\}$.

Solution. (i) implies (ii): well known and true for every topological space. (ii) implies (iii): trivial. (iii) implies (ii) because translations are homeomorphisms. Then, given $x \neq 0$, $x \in X$, $X \setminus \mathrm{cl}(\{x\})$ is nbdhd of 0 which misses x; we have proved that (iii) implies (iv). It remains to prove that (iv) implies (i). If $x, y \in X$ and $x \neq y$ by (iv) there is a nbhd U of 0 such that $x - y \notin U$. If V is a symmetric nbhd of 0 such that $V + V \subseteq U$ (V exists because addition is continuous at $(0,0)$) then $(x+V) \cap (y+V) = \emptyset$. □

EXERCISE 2.1.8.2. Prove that in a TVS the closure of balanced subset is also balanced.

By contrast, the *interior* of a linear subspace of a TVS is always empty, unless the subspace is the entire space (see Exercise 2.1.11.1).

2.1.9. *Segments, rays, lines.* Every $\mathbb{K}-$vector space X is an $\mathbb{R}-$vector space, by restriction of the scalars to $\mathbb{R}$ in the case $\mathbb{K} = \mathbb{C}$. Many concepts are related to the $\mathbb{R}-$linear structure only. Given $a, b \in X$, the *closed segment* with extremes a, b is the set $[a, b] = \{a + t(b-a) : t \in [0,1]\}$, image of the unit interval $[0,1]$ under the map $\sigma(t) = a + t(b-a)$; if the segment is written in this way it is understood that a is the origin and b the endpoint of the segment, which is reduced to the singleton $\{a\}$ if $a = b$ and called *degenerate* in this case; if $a \neq b$ then clearly the map σ is injective. If X carries a vector topology then σ is clearly continuous from $[0,1]$ to X, so that $[a, b]$ is a compact connected subset of X; if X is Hausdorff, and $a \neq b$, then $[a, b]$ is homeomorphic to $[0,1]$ (why?).

It is clear that we can speak of open segments $]a, b[= \{a+t(b-a) : t \in]0,1[\}$ or of half–open segments, e.g $[a, b[= \{a + t(b-a) : t \in [0,1[\}$, etc. But beware that unless X is one or zero–dimensional an open segment will never be an open subset of X. If $a, b \in X$ are distinct, the line through a and b is of course the set $\{a + t(b-a) : t \in \mathbb{R}\}$, a real affine subspace of real dimension one, while the *ray*, or *half–line of origin a in the direction of* $b - a$ is the set $\{a + t(b-a) : t \geq 0\}$ (this is the closed ray; the open ray is $\{a + t(b-a) : t > 0\}$).

2.1.10. Observe that the presence of segments says that

. *Every topological vector space is path connected.*

for, given a, b in the space X, the map $t \mapsto a + t(b-a)$ is a path in X from a to b.

EXERCISE 2.1.10.1. Prove that in a topological vector space every balanced subset is path–connected. Conclude that every TVS is locally path–connected. Hence, the connected components of open subsets of a TVS are open and path–connected.

2.1.11. *Absorbing sets.* In a linear space X, a subset U *absorbs* the subset A if there is $t_A > 0$ such that $tU \supseteq A$ for every $t \geq t_A$. A subset U is *absorbing* if it absorbs every singleton of X. It is very easy to prove that:

. *$U \subseteq X$ is absorbing if and only if for every non zero vector u there is $\delta = \delta(u) > 0$ such that the segment $[0, \delta\, u]$ is contained in U (briefly: a set is absorbing iff it contains a segment from the origin in every direction).*

Proof. Necessity: given $u \in X \setminus \{0\}$, let t_u be such that $u \in t\,U$ for $t \geq t_u$. Then $[0, u/t_u] \subseteq U$: in fact $u \in t\,U \iff u/t \in U$, so that $]0, u/t_u] \subseteq U$; but since U absorbs 0, we also have $0 \in U$, so that $[0, u/t_u] \subseteq U$.

Sufficiency: if for ever nonzero u there is $\delta = \delta(u) > 0$ such that $[0, \delta\, u] \subseteq U$ then $u \in t\,U$ for $t \geq 1/\delta$. □

It is important to observe that in a linear topological space, *all neighborhoods of* 0 *are absorbing*: this is an immediate consequence of the continuity at 0 of the map $\xi \mapsto \xi\, x$, from $\mathbb{R}$ to X, for every $x \in X$ (see 2.1.7).

Notice also that if U is absorbing then $X = \bigcup_{n=0}^{\infty} nU$.

Exercise 2.1.11.1. If a linear subspace Y of X has non empty interior then $Y = X$.

Exercise 2.1.11.2. Let X be a TVS and $f : X \to \mathbb{K}$ be linear not identically 0. Then for every open $A \subseteq X$ the set $f(A)$ is an open subset of $\mathbb{K}$.

(hint: the discontinuous case is trivial, by 2.1.6.1. For the continuous case, prove openness at 0; recall that in $\mathbb{K}$ a non trivial balanced subset is a nbhd of 0. . . ; you need also the preceding exercise).

Exercise 2.1.11.3. Let X be a $\mathbb{K}-$linear space, and assume that U is balanced subset of X. Prove that U absorbs a subset A of X if and only if there is $\alpha \in \mathbb{K}$ such that $A \subseteq \alpha U$. Then U is is absorbing if and only if for every $x \in X$ there is $\alpha_x \in \mathbb{K}$ such that $x \in \alpha_x U$.

(Hint: if $t \geq |\alpha|$ and U is balanced then $\alpha U \subseteq tU$)
2.1.12. *Finite dimensional spaces.* It can be proved that:

Proposition. *Let X be a Hausdorff TVS:*
(i) *If X has finite dimension m, then every isomorphism of $\mathbb{K}^m$ onto X is a homeomorphism.*
(ii) *Every finite dimensional subspace of X is closed in X.*

We shall accept this result; the proof is given below for those who are interested; it is elementary, but not immediate. This implies that a finite dimensional space has a unique separated vector space topology.

Proof. We first prove that (i) implies (ii). It is not restrictive to assume that the finite dimensional subspace Y is dense in X, since we may restrict X to the closure $\mathrm{cl}_X(Y)$ of Y in X. If Y is finite dimensional, then it is homeomorphic to $\mathbb{K}^m$, hence it is locally compact, in particular 0 has an open nbhd V in the relative topology, such that $\mathrm{cl}_Y(V)$ is compact. There exists an open nbhd U of 0 In X such that $U \cap Y = V$; now we have, since U is open in X and Y is dense in X:

$$\mathrm{cl}_X(U) = \mathrm{cl}_X(U \cap Y) = \mathrm{cl}_X(V);$$

but $V \subseteq \mathrm{cl}_Y(V)$ which is compact and hence closed in the Hausdorff space X, so that actually $\mathrm{cl}_X(V) = \mathrm{cl}_Y(V) \subseteq Y$. But then $U \subseteq \mathrm{cl}_X(U) \subseteq Y$; this shows that Y is open in X, and hence that $Y = X$ (see Exercise 2.1.11.1).

(i) We argue by induction on m. For $m = 1$ assertion (i) is simply 2.1.5.1. Assuming that (i) holds for $m-1$, if X is $m-$dimensional and $\omega : \mathbb{K}^m \to X$ is an isomorphism, let $\eta_j = \mathrm{pr}_j \circ \omega^{-1}$; then η_j is a linear functional whose kernel is an $(m-1)-$dimensional subspace K_j of X; by the inductive hypothesis and (ii) the subspace K_j is closed in X, so that by 2.1.6.1 η_j is continuous. But then ω^{-1} is continuous, and ω is a homeomorphism, as desired. □

2.2. Convexity in topological vector spaces. By far the most important subclass of the class of all linear topological spaces is that of *locally convex spaces.* These are linear topological spaces which have a neighborhood basis at 0 consisting of convex sets, and they also are exactly the linear spaces whose vector topology is generated by a family of seminorms, in a sense that will be made explicit. We begin with some reminders on the notion of convexity.

2.2.1. *Convex sets.* Recall that in a real vector space X, a subset $C \subseteq X$ is said to be *convex* if for every pair a, b of points of C the segment $[a, b]$ is contained in C. Clearly every linear subspace is convex; every segment and every line or half–line are also convex. Convex subsets of $\mathbb{R}$ have been called *intervals* of $\mathbb{R}$. From the definition it is clear that the intersection of any family of convex sets is also convex. Then for every subset $S \subseteq X$ there is a smallest convex subset of X containing it, the *convex hull* of S, denoted by $[S]$: it is simply

$$[S] = \bigcap \{K : S \subseteq K \subseteq X, K \text{ convex}\}.$$

If Y is another $\mathbb{R}-$linear space and $T : X \to Y$ is $\mathbb{R}-$linear, it is immediate to see that if $C \subseteq X$ is a convex set then $T(C)$ is convex, and if $K \subseteq Y$ is convex then also $T^{\leftarrow}(K)$ is convex. We leave as an exercise:

. *Images and inverse images of convex sets by linear maps are convex. Translates of convex sets are convex.*

Then, if $C \subseteq X$ is convex, also $a + C$, $-C$, $a + \alpha C$ are convex, for any scalar α and any $a \in X$. Also:

Exercise 2.2.1.1. If A, B are convex subsets in the $\mathbb{R}-$linear space X, then $A + B = \{a + b : a \in A, b \in B\}$ is convex.

This may be proved directly, or by proving first that $A \times B$ is convex in the product space $X \times X$, then applying the addition map $(x, y) \mapsto x + y$, linear from $X \times X$ to X.

2.2.2. *Convex combinations.* Given a, b in the real linear space X, the points of the segment $[a, b]$ may be written in various ways

$$a + t\,(b - a) = (1 - t)\,a + t\,b \quad t \in [0, 1] \quad \text{or also } \alpha\,a + \beta\,b \quad \alpha + \beta = 1,\ \alpha, \beta \geq 0,$$

having set $\alpha = 1 - t$, $\beta = t$. The last is said to be a (linear) *convex combination* of the pair of points a, b (with coefficients α, β). Thus a set is convex iff it contains all convex combinations of all of its pairs. This may also be expressed as:

. *If X is a real linear space, a subset C of X is convex if and only if, for every pair $\alpha, \beta \geq 0$ with $\alpha + \beta = 1$ we have $\alpha\,C + \beta\,C \subseteq C$.*

REMARK. The notion can be generalized to arbitrary finite families of points: given an $n-$tuple $a_1, \dots, a_n$ of points of the real linear space X and positive scalars $\alpha_1, \dots, \alpha_n \geq 0$ with sum $\sum_{j=1}^n \alpha_j = 1$, the linear combination

$$\sum_{j=1}^{n} \alpha_j\,a_j$$

is called convex combination of the n points $a_1, \dots, a_n$. It is not difficult to prove, by induction on n, that

. *A convex set contains all convex combinations of its $n-$tuples of points.*

2.2.3. *Convexity and closure.* It is very easy to see that

. *In a linear topological space X the closure of a convex set C is also convex.*

Proof. Given $\alpha, \beta \geq 0$ with $\alpha + \beta = 1$ the map $X \times X \to X$ given by $(x, y) \mapsto \alpha\,x + \beta\,y$ is clearly continuous, and maps $C \times C$ into C (see 2.2.2). Then it also maps $\bar{C} \times \bar{C}$ into $\bar{C}$. □

2.2.4. *Convexity and interior.* It is less easy to see that in a TVS the interior of a convex set is also convex.

LEMMA. *Let X be a topological linear space, and let C be a convex set. If a is in the interior of C, and $b \in C$, then the half-open segment $[a, b[= \{a + t(b - a) : t \in [0, 1[\}$ is contained in the interior of C.*

Proof. Let U be open, contained in C, with $a \in U$. The sets $U_t = U + t(b - U) = (1 - t)\,U + t\,b$ are all open if $t \neq 1$ (U_t is the image of U by the map $x \mapsto (1 - t)x + t\,b$, which is a self–homeomorphism of X) and are contained in C if $0 \leq t < 1$. Their union is then contained in the interior of C, and clearly contains $[a, b[$. □

. *In a topological linear space X, in particular in a normed space, the interior of a convex set is convex.*
If a convex subset C has non–empty interior, then $\mathrm{int}(C) \subseteq C \subseteq \mathrm{cl}(\mathrm{int}(C))$.
If a symmetric convex set has non empty interior, then the origin is in its interior.

Proof. The lemma implies immediately that if $a, b \in \mathrm{int}(C)$, where C is convex, then the entire segment $[a, b]$ is contained in $\mathrm{int}(C)$, which is then convex.
Given $b \in C \setminus \mathrm{int}(C)$ and $a \in \mathrm{int}(C)$ we have $[a, b[\subseteq \mathrm{int}(C)$, and $b \in \mathrm{cl}([a, b[) \subseteq \mathrm{cl}(\mathrm{int}(C))$, so that $C \subseteq \mathrm{cl}(\mathrm{int}(C))$; taking closures in the inclusion $\mathrm{int}(C) \subseteq C \subseteq \mathrm{cl}(\mathrm{int}(C))$ we get $\mathrm{cl}(\mathrm{int}(C)) \subseteq \mathrm{cl}(C) \subseteq \mathrm{cl}(\mathrm{int}(C))$, so that $\mathrm{cl}(C) = \mathrm{cl}(\mathrm{int}(C))$.
And if $a \in \mathrm{int}(C)$ with $C = -C$, then either $a = 0$, or $[a, -a[\subseteq \mathrm{int}(C)$, in both cases $0 \in \mathrm{int}(C)$ □

EXERCISE 2.2.4.1. In a linear topological space X a *barrel* is a subset B which is closed convex balanced and absorbing. Prove that if X is topologized by a complete metric, then every barrel is a neighborhood of the origin in X.

EXERCISE 2.2.4.2. In a linear space X:

(i) A convex subset C is balanced if and only if for every $\alpha \in \mathbb{K}$ with $|\alpha| = 1$ we have $\alpha\,C \subseteq C$.
(ii) The convex hull of a balanced subset is balanced.
(iii) A locally convex space has a base of neighborhoods at 0 consisting of convex balanced subsets.

Solution. (i) First observe that actually $\alpha\,C = C$ for every α of absolute value 1: in fact $C = (\alpha^{-1}\alpha)\,C \subseteq \alpha\,C \subseteq C$ (that is, a set stable under the action of the sign group is actually invariant under this action; sometimes these sets are called *circled*). Notice that $C = -C$ so that $0 \in C$, by convexity. If $0 < |\alpha| < 1$ we write $\alpha = \mathrm{sgn}\,\alpha\,|\alpha|$, with $\mathrm{sgn}\,\alpha = 1$; for $x \in C$, $\alpha\,x = |\alpha|\,(\mathrm{sgn}\,\alpha\,x)$ is on the segment $[0, \mathrm{sgn}\,\alpha x]$, contained in C by convexity.

(ii) Let S be balanced, so that $\alpha S \subseteq S$ for every $\alpha \in B$, and $\alpha S = S$ if $|\alpha| = 1$. If C is a convex set containing S, and $|\alpha| = 1$, then $S = \alpha S \subseteq \alpha C$ and αC is convex, so that $[S] \subseteq \alpha C$; it follows that a convex subset C of X contains S iff αC also contains S, for every α of absolute value 1; then $\alpha[S] = [S]$, and by (i) the convex hull $[S]$ of S is balanced.

(iii) Given a neighborhood U of 0, by local convexity there is a convex neighborhood C of 0 contained in U; inside C we can find a balanced neighborhood W of 0. Its convex hull $V = [W]$ is then a convex balanced nbhd of 0 contained in $C \subseteq U$. □

2.2.5. *Sublinear functionals and convexity.* If X is a real linear space, and $p : X \to \mathbb{R}$ is a sublinear functional, then for every $t \in \mathbb{R}$ the sublevel sets

$$\{x \in X : p(x) < t\} \quad \text{and} \quad \{x \in X : p(x) \le t\},$$

are both convex (the proof is immediate: if $\alpha x + \beta y$ is a convex combination of elements x, y in, say, the first set, then $p(\alpha x + \beta y) \le p(\alpha x) + p(\beta y) = \alpha p(x) + \beta p(y) < \alpha t + \beta t = t$). If X has a linear topology and p is continuous at 0, then for every $t > 0 (= p(0))$ the set $\{p < t\}$ is a convex nbhd of the origin.

2.2.6. *Minkowski functional of a convex neighborhood of the origin.* In a topological vector space, a convex set which contains 0 in its interior is the "unit ball" of a positive sublinear functional, its *Minkowski functional.* Any neighborhood U of 0 is absorbing, as seen in 2.1.11. If C is the convex set and $0 \in \text{int}(C)$, for every $x \in X$ the set $I(x) = \{t > 0 : x/t \in C\}$ is then non empty, hence it has a finite least upper bound. The Minkowski functional of C, $p_C = p : X \to [0, +\infty[$ is defined by

$$p(x) = \inf I(x), \quad \text{for every } x \in X.$$

Notice also that by convexity of C the set $I(x)$ is a half–line in $\mathbb{R}$: if $s \in I(x)$ and $t > s$, then $t \in I(x)$ (in fact x/t belongs to the segment $[0, x/s]$). We now prove

. *Let C be a convex absorbing subset of the linear space X. The functional $p = p_C$ above defined is sublinear, and we have*

$$\{x \in X : p(x) < 1\} \subseteq C \subseteq \{x \in X : p(x) \le 1\}.$$

Moreover if C is symmetric ($C = -C$) then p is a real seminorm; and if C is balanced, ($\alpha C \subseteq C$ for every $\alpha \in \mathbb{C}$ with $|\alpha| \le 1$) then p is a complex seminorm. If X has a linear topology and $0 \in \text{int}(C)$, then we also have

$$\{x \in X : p(x) < 1\} = \text{int}(C); \qquad \{x \in X : p(x) \le 1\} = \text{cl}\, C.$$

Proof. We prove that for $x, y \in X$ we have $I(x+y) \supseteq I(x) + I(y)$; this clearly implies $\inf I(x+y) \le \inf(I(x) + I(y)) = \inf I(x) + \inf I(y)$, i.e. subadditivity of p. In fact, if $s \in I(x)$ and $t \in I(y)$ we have

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

which is in C, being a convex combination of $x/s \in C$ and $y/t \in C$. If $\alpha > 0$ we clearly have $I(\alpha x) = \alpha I(x)$ so that p is positively homogeneous. If $p(x) < 1$ then $1 \in I(x)$, so that $x = x/1 \in C$; and if $p(x) > 1$ then $1 \notin I(x)$ equivalently $x/1 = x \notin C$,

Symmetry of C implies $I(-x) = I(x)$ for every $x \in C$; and if C is balanced then for every α of absolute value 1 we have $\alpha C = C$, equivalently $I(\alpha x) = I(x)$ for every $x \in C$ if $|\alpha| = 1$.

If $p(x) < 1$ then $t \in I(x)$ for some $t < 1$; then x is in the half–open segment $[0, x/t[$, all contained in $\text{int}(C)$ by 2.2.4; if $x \in \text{int}(C)$ then the set $I(x)$ contains 1 in its interior, by continuity of the map $t \mapsto x/t$ from $]0, +\infty[$ to X so that $p(x) = \inf I(x)$ must be strictly smaller than 1. And if $p(x) = 1$ then $x/t \in C$ for all $t \ge 1$; but since also $x/t \in [0, x[$ for all $t > 1$, and $x \in \text{cl}([0, x[)$, we get $x \in \text{cl}\, C$. □

Exercise 2.2.6.1. Let X be a TVS, and assume that $p : X \to \mathbb{R}$ is sublinear. Then the following are equivalent:

(i) p is continuous on X.
(ii) p is continuous at 0.
(iii) $C = \{x \in X : p(x) \le 1\}$ is a neighborhood of 0.

Solution. Recall that if p is sublinear then $p(0) = 0$ (in fact $p(0) = p(0\,0) = 0\,p(0) = 0$); (i) implies (ii) and (ii) implies (iii) are trivial.

(iii) implies (i) If $\tilde{p}$ is defined by $\tilde{p}(x) = p(-x)$ then also $\tilde{p}$ is sublinear, and $-C = \{x \in X; \tilde{p}(x) \le 1\}$, so that also $-C$ is a nbhd of 0 in X. Set $q(x) = p(x) \vee \tilde{p}(x) = \max\{p(x), p(-x)\}$; it is easy to see that q is a real seminorm, and is exactly the Minkowski functional of $E = C \cap (-C)$, which is a nbhd of 0, being the intersection of two nbhds of 0. This immediately implies continuity of q at 0: given $\varepsilon > 0$ we have $\{x \in X : q(x) \le \varepsilon\} = \varepsilon E$, by positive homogeneity of q. And an $\mathbb{R}$−seminorm q continuous at 0 is continuous at every $x \in X$: simply recall that subadditivity and $q(u) = q(-u)$ imply $|q(y) - q(x)| \le q(y - x)$, so that $x + \varepsilon E \subseteq \{y \in X : |q(y) - q(x)| \le \varepsilon\}$. Finally we clearly have also $|p(y) - p(x)| \le q(y - x)$ for every $x, y \in X$, so that also p is continuous at every $x \in X$. □

2.2.7. *Separation of convex subsets.* Let X be a real linear space, and let A, B be subsets of X. We say that a real linear functional $f : X \to \mathbb{R}$ *separates* A and B if $f(A) \le f(B)$, meaning that $f(x) \le f(y)$ for every $x \in A$ and every $y \in B$. If $\xi \in \mathbb{R}$ is a *separator* of $f(A)$, $f(B)$, that is if $\sup f(A) \le \xi \le \inf f(B)$, then $A \subseteq f^{\leftarrow}(]-\infty, \xi])$ and $B \subseteq f^{\leftarrow}([\xi, +\infty[)$ and $f^{\leftarrow}(]-\infty, \xi])$, $f^{\leftarrow}([\xi, +\infty[)$ are convex subsets, opposite *half–spaces* of X if f is nonzero. If X is a TVS and f is continuous, these half–spaces are closed and have a non empty interior (of course $\mathrm{int}(f^{\leftarrow}(]-\infty, \xi])) = f^{\leftarrow}(]-\infty, \xi[)$, *open half–space*, and analogously for the other half–space). Both half–spaces have the closed *hyperplane* $\{f = \xi\} = f^{\leftarrow}(\xi)$ as their common boundary: we also say that this hyperplane separates A and B, meaning that A and B are contained in closed half–spaces opposite relatively to this hyperplane. These considerations prepare for the following geometric form of the Hahn–Banach theorem.

2.2.8. *Separation of a convex set and a point.*

. *Let X be a TVS, let C be a convex set with non–empty interior, and let a be a point not in the interior of C. Then there exists a closed hyperplane which separates C and a; equivalently, there is a nonzero continuous real valued linear functional f on X such that $f(x) \le f(a)$, for every $x \in C$.*

Proof. First assume that 0 is in the interior of C. Let p be the Minkowski functional of C; then $p(a) \ge 1$. Define $f : \mathbb{R}\, a \to \mathbb{R}$ by $f(\alpha\, a) = \alpha\, p(a)$, for every $\alpha \in \mathbb{R}$. Then f is majorized by p on $\mathbb{R}\, a$, and by the analytic form of the Hahn–Banach theorem f extends to a linear real functional majorized by p on all of X. Then $f(x) \le p(x) \le 1$ for all $x \in C$, while $f(a) = p(a) \ge 1$. Since $f(C)$ is not all of $\mathbb{R}$ (it is contained in the half–line $]-\infty, p(a)]$) the functional f is continuous. If $0 \notin \mathrm{int}(C)$, pick $c \in \mathrm{int}(C)$; then 0 is in the interior of $C - c$, and there exists a real linear $f : X \to \mathbb{R}$ such that $f(x - c) \le f(a - c)$ for all $x \in C$, but by linearity of f this is the same as $f(x) \le f(a)$ for every $x \in C$. □

The closed hyperplane $\{f = f(a)\}$ is a *supporting hyperplane* for the convex set C through f. If $a \notin \mathrm{cl}\, C$, then f also *strongly separates* C and a, in the sense that $\sup f(C) < f(a)$, as appears in the above proof (if $a \notin \mathrm{cl}\, C$ then $p(a) > 1$).

Corollary. *In a locally convex linear topological space X if C is a closed convex subset of X, and $a \in X \setminus C$, there exists a linear continuous functional f which strongly separates C and a.*

Proof. Since X is locally convex, and $X \setminus C$ is an open set containing a, there is an open convex nbhd $a + U$ of a such that $(a + U) \subseteq X \setminus C$. Pick now an open convex symmetric nbhd. V of 0 such that $V + V \subseteq U$. Then $C + V$ is an open convex set, and $a \notin \mathrm{cl}(C + V)$, since $(a + V) \cap (C + V) = \emptyset$, as is easy to verify. By the preceding proposition there is a continuous functional which strongly separates $C + V$ and a. □

2.2.9. *Separation of two convex sets.*

. *Let X be a TVS, let A and B be convex subsets of X. Assume that the interior of A is non–empty and that B is disjoint from the interior of A. Then there exists a closed hyperplane that separates A and B; equivalently, there exists a continuous nonzero real valued linear functional such that $f(x) \le f(y)$, for every $x \in A$ and every $y \in B$.*

Proof. Let $C = \mathrm{int}\, A$; then $C - B$ is an open convex set (see 2.2.1.1) and $0 \notin C - B$, so that there is a linear continuous $f : X \to \mathbb{R}$ such that $f(C-B) \le f(0) = 0$, in other words $f(x-y) \le 0 \iff f(x) \le f(y)$ for every $y \in B$ and every $x \in C = \mathrm{int}(A)$. By continuity, since $A \subseteq \mathrm{cl}\, C$ (2.2.4) we have $f(x) \le f(y)$ also for every $x \in A$. □

Not necessarily two disjoint convex closed subsets with non empty interiors are strongly separated: in $\mathbb{R}^2$ take the set $A = \{(x, y) \in \mathbb{R}^2 : xy \ge 1,\, x > 0\}$ (points above the right branch of the hyperbola $xy = 1$) and $B = \{(x, y) \in \mathbb{R}^2 : y \le 0\}$ (lower half–plane). The functional $(x, y) \mapsto y$ is (essentially) the only functional separating the two sets, in the sense that every separating functional is a multiple of it by some non zero scalar (prove it!). If however one of the two is compact, and not both have empty interior, then trivially we have strong separation.

EXERCISE 2.2.9.1. (*Existence of nonzero continuous linear functionals*) In a topological vector space X, given $a \in X \setminus \{0\}$ there exists a continuous linear functional which is nonzero at a if and only if there is an open convex subset of X which contains 0 but not a.

Solution. Necessity: if $f : X \to \mathbb{K}$ is linear continuous and $f(a) \neq 0$, simply take $C = f^{\leftarrow}(|f(a)|\, D)$, inverse image of the open disk in $\mathbb{K}$ of center 0 and radius $|f(a)|$. And if C is convex open containing 0 but not a there is by 2.2.8 a real linear functional u such that $u(x) < 1 \leq u(a)$ for every $x \in C$; if $\mathbb{K} = \mathbb{R}$ we are done, otherwise take $f(x) = u(x) - i\, u(i\, x)$. □

2.2.10. *A non locally convex space. There is no need to memorize the following definitions, which anyway are far from universal; we only want to give an example, and do not intend to develop any further theory.*

DEFINITION. Let X be a $\mathbb{K}-$linear space. A *pseudoseminorm* on X is a function $[\#] : X \to [0, +\infty[$ which is subadditive ($[x + y] \leq [x] + [y]$ for every $x,\, y \in X$), $[\alpha\, x] \leq [x]$ for every $\alpha \in \mathbb{K}$ with $|\alpha| \leq 1$, and finally $\lim_{n\to\infty}[x/n] = 0$ for every $x \in X$. If $[x] = 0$ implies $x = 0$ then $[\#]$ is a pseudonorm.

First of all observe that $[\alpha\, x] = [x]$ for every α with $|\alpha| = 1$, as is immediate from

$$[x] = [\alpha^{-1}(\alpha\, x)] \leq [\alpha\, x] \leq [x],$$

and also notice that, although we do not have absolute homogeneity, we still have

$$[m\, x] = [x + \cdots + x] \leq [x] + \cdots + [x] = m\, [x] \quad \text{for every} \quad m \in \mathbb{N}.$$

The formula $d(x, y) = [x - y]$ then defines a semimetric on X (a metric in which two distinct points can have 0 distance) which in turn gives a topology on X. We prove that with this topology X is a TVS. Notice that given $x \in X$ and $r > 0$ the ball of center x and radius r is $x + B_r$, if $B_r = \{y \in X : [y] \leq r\}$; notice also that B_r is balanced. Continuity of addition is easy: given $(x,\, y) \in X \times X$ and $\varepsilon > 0$ consider $x + B_{\varepsilon/2}$ and $y + B_{\varepsilon/2}$; then addition maps $(x + B_{\varepsilon/2}) \times (y + B_{\varepsilon/2})$ into $x + y + B_\varepsilon$ as subadditivity immediately implies. Continuity of multiplication scalars×vectors is more involved; given $\alpha \in \mathbb{K}$, $a \in X$ and $\varepsilon > 0$ we have to find $\delta > 0$ such that

$$(\alpha + \delta\, B)(a + B_\delta) \subseteq \alpha\, a + B_\varepsilon, \quad \text{equivalently} \quad \alpha\, u + \beta\, a + \beta\, u \in B_\varepsilon,$$

if $\beta \in \mathbb{K}$ with $|\beta| \leq \delta$ and $u \in B_\delta$. If $\delta < 1$ we have $[\beta\, u] \leq [u] \leq \delta$. Next, given $a \in X$, and $n \in \mathbb{N}$ such that $[a/n] \leq \varepsilon/3$, assume that $\delta \leq 1/n$, so that $\beta = \gamma/n$ with $|\gamma| \leq 1$; then

$$[\beta\, a] = [(\gamma/n)a] = [\gamma\, (a/n)] \leq [a/n] \leq \frac{\varepsilon}{3},$$

for every β with $|\beta| \leq \delta \leq 1/n$. Finally, take an integer $m \geq 1$ such that $|\alpha|/m \leq 1$. Then we have, for $[u] \leq \delta$:

$$[\alpha\, u] = [(\alpha/m)\,(mu)] \leq [mu] \leq m[u] \leq m\delta;$$

so we simply take $\delta \leq \varepsilon/(3m)$, $\delta \leq 1/n$ and we conclude.

The *topology of convergence in measure* on a finite measure space $(X, \mathcal{M}, \mu)$ is given by the pseudoseminorm on the space $L(X) = L_{\mathcal{M}}(X)$ of all measurable functions

$$[f] = \int_X \frac{|f|}{1 + |f|}\, d\mu.$$

It is defined because $1 \in L^1(\mu)$, since $\mu(X) < \infty$ by hypothesis; it is a pseudoseminorm because $t \mapsto t/(1 + t)$ is increasing, subadditive on $[0, +\infty[$ and 0 for $t = 0$. Moreover, given $f \in L(X)$ the sequence $|f/n|/(1+|f/n|)$ converges pointwise to 0, and is dominated by $1 \in L^1(\mu)$, so that $[f/n] \to 0$. So $L(X)$ is TVS with the topology of this seminorm. This topology is called *topology of the convergence in measure*: a sequence $f_n \in L(X)$ is said to converge in measure to $f \in L(X)$ if for every $\alpha > 0$ we have

$$\lim_{n\to\infty} \mu(\{|f - f_n| > \alpha\}) = 0.$$

Let us prove that if X has finite measure then $f_n \to 0$ in measure iff $[f_n] \to 0$ (assuming of course that $f_n \in L(X)$).

Assume first that $f_n \to 0$ in measure, and fix $\varepsilon > 0$. We compute

$$[f_n] = \int_X \frac{|f_n|}{1 + |f_n|}\, d\mu = \int_{\{|f_n| \leq \varepsilon\}} \frac{|f_n|}{1 + |f_n|}\, d\mu + \int_{\{|f_n| > \varepsilon\}} \frac{|f_n|}{1 + |f_n|}\, d\mu \leq \int_{\{|f_n| \leq \varepsilon\}} |f_n|\, d\mu + \int_{\{|f_n| > \varepsilon\}} d\mu \leq$$
$$\leq \varepsilon\, \mu(X) + \mu(\{|f_n| > \varepsilon\});$$

Pick then $n_\varepsilon \in \mathbb{N}$ such that $\mu(\{|f_n| > \varepsilon\}) \le \varepsilon$ for $n \ge n_\varepsilon$; the conclusion is immediate.

Assume now that $[f_n] \to 0$, and fix $\alpha > 0$. Observe that (always because $t \mapsto t/(1+t)$ is increasing and continuous) we have $\{|f| > \alpha\} = \{|f|/(1+|f|) > \alpha/(1+\alpha)\}$ so that:

$$[f_n] = \int_X \frac{|f_n|}{1+|f_n|}\,d\mu \ge \int_{\{|f_n|>\alpha\}} \frac{|f_n|}{1+|f_n|}\,d\mu \ge \int_{\{|f_n|>\alpha\}} \frac{\alpha}{1+\alpha}\,d\mu = \frac{\alpha}{1+\alpha}\,\mu(\{|f_n| > \alpha\}),$$

so that

$$\mu(\{|f_n| > \alpha\}) \le (1/\alpha + 1)\,[f_n],$$

and we conclude by letting n tend to infinity.

Exercise 2.2.10.1. In the space $(L([0,1]), [\#])$ with Lebesgue measure prove that the only non–empty open convex set is the entire space (prove that for every $\varepsilon > 0$ the convex hull of B_ε is the entire space; partition $[0,1]$ into n subintervals of length $1/n < \varepsilon$, and given $f \in L(X)$ write $nf = \sum_{k=1}^n (nf)\chi_{[(k-1)/n, k/n[}\ \dots$). Deduce that the only continuous linear functional on $L(X)$ is identically zero.

Exercise 2.2.10.2. Prove that if $(X, \mathcal{M}, \mu)$ is a finite measure space, then setting

$$q(f) = \int_X |f| \wedge 1\,d\mu \quad f \in L(X),$$

we get a pseudo seminorm topologically equivalent to $[f]$ previously introduced.

Exercise 2.2.10.3. Let $0 < p < 1$; $L^p(\mu) = \{f \in L(X) : |f|^p \in L^1(\mu)\}$. Setting

$$\nu(f) = \int_X |f|^p\,d\mu \quad f \in L^p(\mu)$$

we get a pseudonorm on $L^p(\mu)$. ⊙⊙Again, if $X = [0,1]$ with Lebesgue measure, there are no proper open convex subsets of $L^p(\mu)$ (more difficult than the previous one).

2.3. Topologies generated by families of seminorms. If X is a vector space, and $(p_\alpha)_{\alpha\in A}$ is a family of seminorms on X, we can consider on X the topologies τ_α of the various seminorms p_α; the topology generated by $(p_\alpha)_{\alpha\in A}$ is, by definition, the weakest topology among those which are stronger than every τ_α: in other words, τ is the intersection of all topologies on X which contain $\bigcup_{\alpha\in A}\tau_\alpha$. This topology can be so described:

. *A subset A of X is open in the topology τ if and only if it is the union of sets of the form $\bigcap_{\alpha\in F} A_\alpha$, where each $A_\alpha \in \tau_\alpha$, and F is a finite subset of A.*

Proof. Clearly all these sets are open in the topology τ; and is easy to see that the sets of this form are a topology, which is then τ. □

We shorten $B_{p_\alpha}(x, r[$ to $B_\alpha(x, r[$, that is

$$B_\alpha(x, r[:= \{y \in X : p_\alpha(x-y) < r\} = x + B_\alpha(0, r[.$$

A neighborhood basis for τ at the point $x \in X$ is then

$$x + \bigcap_{\alpha\in F} B_\alpha(0, r[\qquad r > 0,$$

and a nbhd. basis at 0 is then

$$B_q(0, r[= \{y \in X : q(y) < r\} \qquad (r > 0),\ q = q_F := \bigvee_{\alpha\in F} p_\alpha.$$

Remember in fact that the supremum of a finite set of seminorms is a seminorm; clearly

$$B_{\bigvee p_\alpha}(0, r[= \bigcap_{\alpha\in F} B_\alpha(0, r[.$$

The topology τ is a vector space topology: we know that every seminorm generates a vector space topology, so that continuity of the vector space operations is ensured from their continuity in the topologies of the seminorms $q_F := \bigvee_{\alpha\in F} p_\alpha$. Moreover, the topology τ is Hausdorff if and only if for every nonzero $x \in X$ there exists an $\alpha \in A$ such that $p_\alpha(x) > 0$: simply remember that a vector topology is Hausdorff iff the intersection of all nbhds of 0 is $\{0\}$. It is perhaps in order to note that instead of the seminorms $q_F := \bigvee_{\alpha\in F} p_\alpha$, with F varying in the set of finite subsets of A, we could have picked the seminorms $s_F := \sum_{\alpha\in F} p_\alpha$, or $x \mapsto \|(p_\alpha(x))_{\alpha\in F}\|_p$, where $\|\#\|_p$ is the ℓ^p–norm on $\mathbb{R}^F$; we chose the maximum because the ball of a given radius is then exactly the intersection of the balls of same radius of the seminorms p_α, $\alpha \in F$.

2.3.1. *Convergence of sequences in spaces with seminorms.* Assume that in the linear space X the topology τ is given by a family $(p_\alpha)_{\alpha\in A}$ of seminorms. Then

. *A sequence $(x_n)_{n\in\mathbb{N}}$ of X converges to $x \in X$ in the topology τ if and only if the sequence converges to x in every seminorm p_α, i.e. $\lim_{n\to\infty} p_\alpha(x - x_n) = 0$, for every $\alpha \in A$.*

Proof. The easy proof is left as an exercise. □

2.3.2. *Continuity of linear maps in spaces with seminorms.* The following condition for continuity is very important.

Proposition. *Let X and Y be linear spaces, $T : X \to Y$ a linear map, and let $(p_\alpha)_{\alpha\in A}$ and $(q_\beta)_{\beta\in B}$ be families of seminorms on X and Y respectively, with respective topologies τ on X and σ on Y. Then T is continuous if and only if for every $\beta \in B$ there exist a finite subset $F = F_\beta$ of A and $L = L_\beta \geq 0$ such that*

$$(*)\qquad q_\beta(Tx) \leq L \bigvee_{\alpha\in F} p_\alpha(x),$$

for every $x \in X$.

Proof. By linearity, continuity of T needs to be checked only at the origin. For that, we need only to prove that for every $\beta \in B$ and every $\varepsilon > 0$ the inverse image of the ball $B_{q_\beta}(0, \varepsilon[$ is a neighborhood of the origin in X: every nbhd. of 0 in Y contains in fact an intersection of finitely many such balls. By homogeneity we may assume $\varepsilon = 1$. Assuming T continuous, we get that $T^{\leftarrow}(B_{q_\beta}(0, 1[)$ contains $B_p(0, \delta]$ (closed ball) for some $\delta > 0$, where for simplicity we have put $p = \bigvee_{\alpha\in F} p_\alpha$. Then (*) is satisfied, with $L = 1/\delta$. In fact, if $x \in X$ then either $p(x) = 0$, or $p(x) > 0$; in the first case $p(\lambda x) = \lambda p(x) = 0$ for every $\lambda > 0$, so that $\lambda x \in B_p(0, \delta]$, so that $T(\lambda x) \in B_{q_\beta}(0, 1[$ for every $\lambda > 0$, i.e $q_\beta(T(\lambda x)) < 1$ for every $\lambda > 0$; but since $q_\beta(T(\lambda x)) = \lambda q_\beta(Tx)$ we must have $q_\beta(Tx) = 0$.

If $p(x) > 0$ we have $\delta x/p(x) \in B_p(0, 1]$ so that

$$q_\beta(\delta x/(p(x)) \leq 1 \iff q_\beta(x) \leq \frac{1}{\delta} p(x),$$

and necessity of the condition is proved. Sufficiency is trivial: under (*) we have

$$T^{\leftarrow}(B_{q_\beta}(0, \varepsilon[) \supseteq B_p(0, \varepsilon/L[\quad \text{with } p = \bigvee_{\alpha\in F} p_\alpha.$$

□

Exercise 2.3.2.1. Let X be a linear space. Prove that a linear functional $f : X \to \mathbb{K}$ is continuous in the topology of the seminorm $p_f(x) = |f(x)|$, and that f is continuous in a linear topology if and only if this topology is stronger than that of the seminorm p_f.

2.3.3. *Weak topologies.* If $(Y_\alpha)_{\alpha\in A}$ is a family of normed spaces, X is a linear space, and for every $\alpha \in A$ we have a linear map $T_\alpha : X \to Y_\alpha$, there exists a coarsest linear topology τ which makes all functions T_α continuous; it is easy to check that this topology is that of the family of seminorms $(p_\alpha)_{\alpha\in A}$ given by

$$p_\alpha(x) = \|T_\alpha(x)\|_\alpha \quad \text{where } \|\#\|_\alpha \text{ is the norm on } Y_\alpha.$$

In particular we may have $Y_\alpha = \mathbb{K}$ for every $\alpha \in A$. and $T_\alpha = f_\alpha$ is, for each $\alpha \in A$, a functional from X to $\mathbb{K}$. The topology of the seminorms $p_\alpha(x) = |f_\alpha(x)|$ is called *weak topology* of the family $(f_\alpha)_{\alpha\in A}$. Let us prove:

. *Given a family $(f_\alpha)_{\alpha\in A}$ of linear functionals on X, a linear functional $f : X \to \mathbb{K}$ is continuous in the weak topology of the family if and only if f belongs to the vector space spanned by the family.*

Proof. If f is in this vector space, then $f = \sum_{\alpha\in F} \lambda_\alpha f_\alpha$ is a finite linear combination of the functionals in the family, which are all continuous, so that f is also continuous. And if $f : X \to \mathbb{K}$ is continuous there exists a finite family $(f_\alpha)_{\alpha\in F}$ and $L > 0$ such that

$$|f(x)| \leq L \bigvee_{\alpha\in F} p_\alpha(x) \qquad (p_\alpha(x) := |f_\alpha(x)|);$$

in particular we have $f(x) = 0$ if $p_\alpha(x) = 0$ for all $\alpha \in F$, equivalently $\mathrm{Ker}(f) \supseteq \bigcap_{\alpha\in F} \mathrm{Ker}(f_\alpha)$. A well–known theorem of linear algebra says that this is equivalent to the assertion that f is in the linear space spanned by $(f_\alpha)_{\alpha\in F}$, i.e. there exist $\{\lambda_\alpha \in \mathbb{K} : \alpha \in F\}$, such that $f = \sum_{\alpha\in F} \lambda_\alpha f_\alpha$. □

Notation: if X is a linear space, and $S \subseteq \mathrm{Hom}_{\mathbb{K}}(X, \mathbb{K})$ is a set of linear forms on X, the weak topology of this set on X will be denoted by $\sigma(X, S)$; we clearly have $\sigma(X, S) = \sigma(X, \langle S\rangle)$ if $\langle S\rangle$ is the subspace of $\mathrm{Hom}_{\mathbb{K}}(X, \mathbb{K})$ generated by S. Proposition 2.3.3 then shows that the topological dual of $(X, \sigma(X, S))$ is exactly the space $\langle S\rangle$ generated by S.

Exercise 2.3.3.1. (!) Let X be a linear space. Assume that S and T are sets of linear functionals on X, i.e. $S, T \subseteq \mathrm{Hom}_{\mathbb{K}}(X, \mathbb{K})$. Prove that $\sigma(X, S) \subseteq \sigma(X, T)$ if and only if $\langle S\rangle \subseteq \langle T\rangle$, with equality iff $\langle S\rangle = \langle T\rangle$.

Example 2.3.3.2. If E is a topological space, the set $X = C(E, \mathbb{K})$ of all continuous $\mathbb{K}-$valued functions is a linear space under pointwise operations. The *topology of pointwise convergence*, or simply *pointwise topology*, on X, is the topology $\sigma(X, \{\delta_x : x \in E\})$ of all evaluation functionals δ_x, where $\delta_x(f) = f(x)$, for every $x \in E$ and every $f \in X$.

Exercise 2.3.3.3. (!) Let X be a linear space and let S be a subset of $\mathrm{Hom}_{\mathbb{K}}(X, \mathbb{K})$, Prove that the following are equivalent:

(i) $\sigma(X, S)$ is a Hausdorff topology
(ii) S separates the points of X, that is, if $x, y \in X$ and $x \neq y$ then $f(x) \neq f(y)$ for some $f \in S$
(iii) For every nonzero $x \in X$ there exists $f \in S$ such that $f(x) \neq 0$.

2.3.4. *The weak topology of a normed space.* If X is normed, the weak topology $\sigma(X, X^*)$ of all norm–continuous linear functionals is simply called *weak topology of* X.

Exercise 2.3.4.1. (!!) Prove that the weak topology is Hausdorff (this is the main application of the Hahn–Banach theorem).

Of course the norm topology has many more open sets than the weak topology:

. *If X is not finite dimensional, then every non–empty weakly open subset of X is unbounded with respect to the norm; in fact it contains an affine subspace of finite codimension.*

Proof. By translation, we may assume that the open set A contains 0; then the open set contains a subset of the form

$$\bigcap_{f\in F} \{x \in X : |f(x)| < \varepsilon\},$$

for some finite subset F of X^* and some $\varepsilon > 0$. Then A contains the subspace $\bigcap_{f\in F} \mathrm{Ker}\, f$, of finite codimension an hence nontrivial; and no nontrivial subspace is norm–bounded. □

The norm topology has of course also many more closed sets than the weak topology. The following result is then remarkable.

. *A convex norm–closed subset of a normed space X is closed also in the weak topology.*

Proof. Given a norm–closed convex subset C and $a \notin C$, by corollary 2.2.8 there is a continuous linear real functional f which strongly separates C and a, i.e. such that $\sup f(C) < f(a)$; then $f^{\leftarrow}(]\sup f(C), +\infty[$ is a weakly open subset of X containing a and not meeting C. □

In particular the closed unit ball B_X of the norm is also weakly closed. But:

. *If X is an infinite dimensional normed space the weak closure of the sphere*

$$S = S_X = \{u \in X : \|u\|_X = 1\}$$

of vectors of unit norm is the closed unit ball $B = B_X = \{x \in X : \|x\|_X \leq 1\}$.

Proof. Since the closed unit ball is weakly closed, as we have just seen, the weak closure of S is contained in B. And if $x \in X$ with $\|x\| < 1$, since every weak neighborhood U of x contains a non–trivial affine subspace $x + V$, we have $U \cap S \neq \emptyset$: if $v \in V \setminus \{0\}$, for a convenient $t \in \mathbb{R}$ a vector $x + t\,v$ is of norm 1. □

Exercise 2.3.4.2. Say that a subset E of the normed space X is *weakly bounded* if $f(E)$ is a bounded subset of $\mathbb{K}$ for every $f \in X^*$. Prove that a weakly bounded subset is also norm bounded. Deduce that all weakly compact subsets of X are norm–bounded.

Solution. Apply the uniform boundedness theorem to $\{\delta_x : x \in E\}$ considered as evaluation functionals on X^*. □

EXERCISE 2.3.4.3. Assume that a sequence $(x_n)_n$ in the normed space X is weakly convergent to $x \in X$ (sometimes this is written $x_n \rightharpoonup x$). Then $\{x_n : n \in \mathbb{N}\}$ is norm–bounded in X. If moreover X is a scalar product space, and $\|x_n\| \to \|x\|$, then $x_n \to x$ in norm.

Solution. For the last part: simply compute ($\|x\| = |x| = (x \mid x)^{1/2}$)

$$|x - x_n|^2 = |x|^2 - 2\,\mathrm{Re}(x_n \mid x) + |x_n|^2 \dots$$

□

2.3.5. *The weak* topology of the normed dual.* If X is a normed space, the weak topology $\sigma(X^*, X)$ given by the evaluation functionals $\delta_x : X^* \to \mathbb{K}$, $\delta_x(f) = f(x)$, is called weak* topology on X^*; trivially it is a Hausdorff topology. In this topology the only continuous linear functionals are the δ_x themselves (2.3.3); if the space X is reflexive, these are all the continuous linear functionals on X^*, so that this topology is exactly the weak topology $\sigma(X^*, X^{**})$ of X^* as a normed space; but if X is not reflexive then no functional in $X^{**} \setminus J(X)$ is continuous in the weak* topology of X^*, which then is strictly coarser than the weak topology $\sigma(X^*, X^{**})$ of X^*.

EXAMPLE 2.3.5.1. (Riemann–Lebesgue lemma). Let $p \in L^\infty(\mathbb{R})$ be τ−periodic. For $\lambda > 0$, $p_\lambda(\#) = p(\lambda\#)$ is a family of periodic functions, all with the same L^∞ norm. Let $\mu = \int_0^\tau p(t)\,(dt/\tau)$ the average of p on a period. Then $\lim_{\lambda\to+\infty} p_\lambda = \mu$, in the weak* topology of $L^\infty(\mathbb{R})$, considered as the dual space of $L^1(\mathbb{R})$: in other words

$$\lim_{\lambda\to+\infty} \int_{\mathbb{R}} f(t)\, p(\lambda\, t)\, dt = \int_{\mathbb{R}} f(t)\, \mu\, dt, \quad \text{for every } f \in L^1(\mathbb{R}).$$

Recall the proof: first one proves that the formula is true for a dense subspace of $L^1(\mathbb{R})$, e.g. that of integrable step–functions; by equicontinuity of the p_λ (when considered as linear functionals on $L^1(\mathbb{R})$, we have $\|p_\lambda\|_\infty = \|p\|_\infty$ for every λ) the formula is then true for all $f \in L^1(\mathbb{R})$.

EXERCISE 2.3.5.2. Let X be the real Hilbert space $L^2([0,1], \mathbb{R})$. Let $(f_n)_n$ be a bounded sequence in X, let $f \in X$, and set, for $x \in [0,1]$:

$$F_n(x) = \int_0^x f_n(t)\, dt; \quad F(x) = \int_0^x f(t)\, dt.$$

Prove that:

(i) Every F_n is hölderian with exponent $1/2$.
(ii) If $f_n \rightharpoonup f$ weakly in X, then $F_n \to F$ uniformly on $[0,1]$.
(iii) If $F_n \to F$ pointwise on $[0,1]$, then $f_n \rightharpoonup f$ weakly in X.

Proof. (i) Assuming $0 \le x_1 < x_2 \le 1$, and using C.S.:

$$|F_n(x_2) - F_n(x_1)| = \left|\int_{x_1}^{x_2} f_n(t)\, dt\right| \le \int_{x_1}^{x_2} |f_n(t)|\, dt = \left(\int_{x_1}^{x_2} (f_n(t))^2\, dt\right)^{1/2} |x_1 - x_2|^{1/2} \le$$
$$\le \left(\int_0^1 (f_n(t))^2\, dt\right)^{1/2} |x_1 - x_2|^{1/2} = \|f_n\|_2\, |x_1 - x_2|^{1/2} \le M\, |x_1 - x_2|^{1/2},$$

if M is a constant which dominates all $\|f_n\|_2$.

Notice that the functions are equihölderian, with a common Hölder constant, so that they are equicontinuous; and if $M \ge \|f\|_2$ then M is an Hölder constant also for F.

(ii) Weak convergence implies immediately pointwise convergence of F_n to F: if, for every $x \in [0,1]$ we denote by φ_x the characteristic function of $[0,x]$ we have

$$F_n(x) = \int_0^x f_n(t)\, dt = \int_0^1 f_n(t)\, \varphi_x(t)\, dt = (f_n \mid \varphi_x) \to (f \mid \varphi_x) = \int_0^x f(t)\, dt = F(x).$$

Equicontinuity then implies uniform convergence (see Exercise 1.5.1.2). We can repeat the argument: given $\varepsilon > 0$, pick $\delta > 0$ which will have to be small with respect to ε, and partition the interval $[0,1]$ into m subintervals shorter than δ. By pointwise convergence, if $x_1 = 0, x_2, \dots x_m = 1$ are the points of subdivision there is n_ε such that if $n \geq n_\varepsilon$ then $|F(x_k) - F_n(x_k)| \leq \varepsilon$, for $k = 1, \dots, m$. By equihölderianity we get $|F_n(x) - F_n(x_k)| \leq M\,\delta^{1/2}$ if $x \in [x_{k-1}, x_k]$, so that, if $x \in [x_{k-1}, x_k]$ and $n \geq n_\varepsilon$:

$$\begin{aligned}|F(x) - F_n(x)| =& |F(x) - F(x_k) + F(x_k) - F_n(x_k) + F_n(x_k) - F_n(x)| \leq \\ & |F(x) - F(x_k)| + |F(x_k) - F_n(x_k)| + |F_n(x_k) - F_n(x)| \leq \\ & M\,\delta^{1/2} + \varepsilon + M\,\delta^{1/2} \leq 3\varepsilon \quad \text{if} \quad \delta \leq (\varepsilon/M)^2.\end{aligned}$$

Alternatively we can observe that all functions F_n are 0 in 0, and by equihölderianity we get $|F_n(x)| \leq M\sqrt{x} \leq M$ for every n and every x, that is, the set $\{F_n : n \in \mathbb{N}\}$ is uniformly bounded, so that it has compact closure by Ascoli's theorem; every subsequence has then a uniformly converging subsequence, but pointwise convergence to F implies that all these sequences converge uniformly to F, which is therefore the uniform limit of the entire sequence.

(iii) As remarked above, pointwise convergence of F_n to F is equivalent to say that for every $x \in [0, 1]$ we have

$$(f_n \mid \varphi_x) \to (f \mid \varphi_x) \qquad \text{where} \quad \varphi_x = \chi_{[0,x]}.$$

By linearity this implies that we have $(f_n \mid g) \to (f \mid g)$ for every $g \in V$, where V is the linear space generated by the set of functions $\{\varphi_x : x \in [0, 1]\}$; since $\varphi_b - \varphi_a = \chi_{]a,b]}$, for $0 \leq a < b \leq 1$, the space V is exactly the space of all step functions on $[0, 1]$, a space which is dense in $X = L^2([0, 1])$. Then we prove that $(f_n \mid h) \to (f \mid h)$ for every $h \in X$; this is essentially due to the fact that f_n is bounded in X, so that the f_n, considered as elements of the dual $X^* = X$ are equicontinuous, and if they converge to the linear form $(\# \mid f)$ on a dense subspace V of X they converge on all the space X: explicitly, given $\varphi \in X$ and $\varepsilon > 0$, pick $g \in V$ such that $\|\varphi - g\|_2 \leq \varepsilon$; then

$$\begin{aligned}& |(f_n \mid \varphi) - (f \mid \varphi)| = |(f_n - f \mid \varphi)| = |(f_n - f \mid \varphi - g + g)| = \\ & |(f_n - f \mid \varphi - g) + (f_n - f \mid g)| \leq |(f_n - f \mid \varphi - g)| + |(f_n - f \mid g)| \\ & \leq \|f_n - f\|_2\,\|\varphi - g\|_2 + |(f_n - f \mid g)| \leq (\|f_n\|_2 + \|f\|_2)\,\|\varphi - g\|_2 + |(f_n - f \mid g)| \leq \\ & 2M\,\varepsilon + |(f_n - f \mid g)|.\end{aligned}$$

If $n \to \infty$ then $|(f_n - f \mid g)| \to 0$, and we conclude. □

In the two previous exercises we have used twice the same technique. There is a useful result which is better to state explicitly.

PROPOSITION. *Let X and Y be normed spaces, with Y complete. Let $T_n \in L(X, Y)$ be an equicontinuous sequence of linear operators (i.e. $\sup_n\{\|T_n\|_{L(X,Y)}\} < \infty$). Let*

$$V = \{x \in X : \lim_{n\to\infty} T_n(x) \text{ exists in } Y\};$$

then V is a closed linear subspace of X, and the formula $T(x) = \lim_{n\to\infty} T_n(x)$ defines a bounded linear operator on V, with norm $\|T\| \leq \liminf_{n\to\infty} \|T_n\|$.

Proof. That V is a subspace is immediate: if $T_n x \to Tx$ and $T_n y \to Ty$ then $T_n(x+y) = T_n x + T_n y \to Tx + Ty$, and $T_n(\alpha\,x) = \alpha\,T_n x \to \alpha\,Tx$, so that V is a subspace and T is linear on V. Moreover, if $L = \sup_n\{\|T_n\|\}$ we have $\|Tx\|_Y = \lim_{n\to\infty} \|T_n x\|_Y \leq L\|x\|$, for every $x \in V$, so that T is continuous and $\|T\| \leq L$; by uniform continuity T extends to a continuous linear operator on $\mathrm{cl}_X(V)$ (here completeness of Y is needed), which we still call T. We have to prove that Tx equals $\lim_{n\to\infty} T_n x$ for $x \in \mathrm{cl}_X(V)$ (thus proving that $V = \mathrm{cl}_X(V)$). Given $\varepsilon > 0$ pick $y \in V$ such that $\|x - y\| \leq \varepsilon$ and estimate:

$$\begin{aligned}\|Tx - T_n x\| =& \|Tx - Ty + Ty - T_n y + T_n y - T_n x\| \leq \\ & \|Tx - Ty\| + \|Ty - T_n y\| + \|T_n y - T_n x\| \leq \\ & \leq L\,\|x - y\| + \|Ty - T_n y\| + L\|x - y\| \leq 2L\,\varepsilon + \|Ty - T_n y\|,\end{aligned}$$

and if $n \geq n_\varepsilon$ the last term is dominated by ε. The last inequality, $\|T\| \leq \liminf_{n\to\infty} \|T_n\|$, is proved in the remark after 1.7.8. □

EXERCISE 2.3.5.3. Prove that the sequences

$$f_n(x) = \sin(n^2 x^2); \quad g_n(x) = \sin(\sqrt{n\,x})$$

converge weakly to 0 in $L^2([0, 1])$.

Solution. Since $|f_n(x)| \le 1$ and $|g_n(x)| \le 1$, both sequences are bounded in $L^\infty([0,1])$, so also in $L^2([0,1])$. By the previous proposition we then only need to prove that $\int_0^1 f_n g \to 0$ and $\int_0^1 g_n\, g \to 0$ for all g in some dense subset of $L^2([0,1])$.

We prove that this is true for every function g of the form $\chi_{[0,b]}$, for every $b \in [0,1]$; the linear space generated by this set of functions is the space of step-functions, known to be dense in $L^p([0,1])$ for every p with $1 \le p < \infty$ (incidentally, we then prove weak convergence to 0 of the sequence f_n in each of these spaces). In fact

$$\int_0^1 \sin(n^2x^2)\,\chi_{[0,b]}(x)\,dx = \int_0^b \sin(n^2x^2)\,dx = \int_0^{n^2b^2} \sin t\,\frac{dt}{2n\sqrt{t}} = \frac{1}{2n}\int_0^{n^2b^2} \frac{\sin t}{\sqrt{t}}\,dt;$$

If $F(x) = \int_0^x (\sin t/\sqrt{t})\,dt$, by the Abel–Dirichlet criterion the generalized integral

$$F(+\infty) = \int_0^{\uparrow\infty} \frac{\sin t}{\sqrt{t}}\,dt$$

is finite (Fresnel's integral); thus

$$\lim_{n\to\infty} \int_0^b \sin(n^2x^2)\,dx = \lim_{n\to\infty}\left(\frac{1}{2n}\int_0^{n^2b^2} \frac{\sin t}{\sqrt{t}}\,dt\right) = 0,$$

as desired.

Similarly, and more easily:

$$\int_0^b \sin(\sqrt{nx})\,dx = \int_0^{\sqrt{nb}} \sin t\,\frac{2t}{n}\,dt = \frac{2}{n}\int_0^{\sqrt{nb}} t\,\sin t\,dt;$$

A primitive of $t\sin t$ is $-t\cos t + \sin t$, so that

$$\int_0^b \sin(\sqrt{nx})\,dx = \frac{2}{n}\left[\sin t - t\cos t\right]_{t=0}^{t=\sqrt{nb}} = 2\,\frac{\sin(\sqrt{nb}) - \sqrt{nb}\,\cos(\sqrt{nb})}{n} \to 0.$$

□

EXERCISE 2.3.5.4. Let $(f_n)_{n\in\mathbb{N}}$ be a sequence in $L^p = L^p([0,1])$. Prove that if $p > 1$ then the sequence is weakly convergent in L^p if and only if the sequence is bounded in L^p, and for every $c \in [0,1]$ the limit

$$\lim_{n\to\infty} \int_0^c f_n(x)\,dx \quad \text{exists in } \mathbb{K}.$$

⊙⊙ Prove that if $p = 1$ the conditions are not sufficient.

We may also consider L^1 as a subspace of the dual of $C([0,1])$, the space $M([0,1])$ of finite Radon measures on $[0,1]$, and consider on L^1 the vague topology. Are the preceding conditions sufficient for convergence in the vague topology?

Solution. Necessity: It is well–known that a weakly converging sequence is bounded (uniform boundedness, consider the elements of f_n as functionals acting on $(L^p)^* = L^q$), and if f is the weak limit then

$$\lim_{n\to\infty} \int_0^1 f_n\,g = \int_0^1 f\,g \quad \text{for every } g \in L^q, \text{ in particular for } g = \chi_{[0,c]}.$$

Sufficiency: Since the sequence is by hypothesis bounded, the functionals f_n acting on the dual space L^q are an equicontinuous family and the set of all $g \in L^q$ for which $\lim_{n\to\infty}\langle f_n, g\rangle$ exists is a closed linear subspace of L^q, which in our case then has to contain the linear subspace spanned by the set of functions $\{\chi_{[0,c]} : c \in [0,1]\}$; it is well–known and easy to see that this space is the space of step–functions. The space of step–functions is known to be dense in L^q if $q < \infty$; thus if $p > 1$ the conclusion is reached. If $p = 1$ we shall construct an example which shows that the conditions are not sufficient. First we address the last question: the uniform closure of the space of step–functions is the space of *regulated functions*, those functions which have a finite left and right limit at every point of $[0,1]$ (this is found e.g. in Analisi Uno, 15.9.5, even if there the question is not stated in terms of uniform approximation); but it is anyway immediate to prove that this closure contains the space of continuous functions on $[0,1]$ (a simple argument of uniform continuity), so that the answer to the last question is yes.

Example Consider in $[0,1]$ a strictly decreasing sequence $a_0 = 1 > a_1 > \dots$ with limit 0. For every n let $c_n = (a_n + a_{n+1})/2$ be the center of the interval $[a_{n+1}, a_n]$, let $\varepsilon_n = a_n - a_{n+1}$ be its length, and put

$$f_n = \frac{1}{\varepsilon_n}\,(\chi_{[c_n,a_n]} - \chi_{[a_{n+1},c_n]});$$

Then $f_n \in L^1$ and $\|f_n\|_1 = 1$ for every n; if $c \in]0,1]$ we have, as soon as $a_n < c$:

$$\int_0^c f_n(x)\,dx = \int_0^1 \frac{1}{\varepsilon_n}\,(\chi_{[c_n,a_n]}(x) - \chi_{[a_{n+1},c_n]}(x))\,dx = \frac{1}{\varepsilon_n}((a_n - c_n) - (c_n - a_{n+1})) = 0;$$

but if $g = \sum_{k=0}^{\infty} \chi_{[c_k,a_k]}$ we get

$$\int_0^1 f_n(x)\,g(x)\,dx = \int_0^1 \frac{1}{\varepsilon_n}\chi_{[c_n,a_n]}(x)\,dx = \frac{1}{2}.$$

□

Exercise 2.3.5.5. Let X be a Banach space, and let $\phi : [0,1] \to X$ be continuous with respect to the usual topology on $[0,1]$ and the weak topology on X. Let $f : [0,1] \to \mathbb{R}$ be defined by $f(t) = \|\phi(t)\|$. Prove that:

(i) f is bounded.
(ii) f has a minimum on X.

Solution. The image $K = \phi([0,1])$ is weakly compact being the continuous image of the compact space $[0,1]$. Weakly compact subsets are weakly bounded (every $u \in X^*$ is weakly continuous, so that $u(K)$ is bounded in $\mathbb{K}$). But we have seen that weakly bounded sets are also norm–bounded (exercise 2.3.4.2), hence f is bounded. And we know that the norm is sequentially lower semicontinuous in the weak topology, that is, if $x_k \rightharpoonup x$ in the normed space X, then $\|x\| \le \liminf_k \|x_k\|$; this implies that if $t_k \to t$ in $[0,1]$ then $f(t) \le \liminf_k f(t_k)$; this and the compactness of $[0,1]$ easily imply that f has a minimum (take a minimizing sequence $f(t_k) \to \inf f([0,1])$, with $t_k \to t \in [0,1] \dots$). □

Exercise 2.3.5.6. Let $X = C([0,1],\mathbb{R})$; we shall consider various norms on X, the $\|\#\|_p$ norms and the uniform norm $\|\#\|_u = \|\#\|_\infty$. Consider the functional $\varphi : X \to \mathbb{R}$ given by

$$\varphi(f) = \int_0^1 \sin(\pi t)\,f(t)\,dt.$$

(i) Put on X the norm $\|\#\|_1$: verify that φ is continuous and compute its norm.
(ii) If we put on X the L^p norm, how does the norm of φ change?
(iii) Prove that φ is not continuous with respect to the topology of pointwise convergence on X (the topology $\sigma(X,F)$ where $F = \{\delta_x : x \in [0,1]\}$).

Solution. (i) We write for simplicity $a(t)$ in place of $\sin(\pi t)$; we have

$$|\varphi(f)| = \left|\int_0^1 a(t)\,f(t)\,dt\right| \le \int_0^1 |a(t)|\,|f(t)|\,dt \le \int_0^1 \|a\|_\infty\,|f(t)|\,dt = \|a\|_\infty\,\|f\|_1,$$

so that φ is bounded with norm $\|\varphi\| \le \|a\|_\infty$. To show that $\|\varphi\| = \|a\|_\infty$; in this case a direct argument is not difficult, but we can resort to general theorems in this way: in the L^1 norm, X is dense in $L^1([0,1])$, its completion, and extends to an element of the dual of L^1, with the same norm; the dual of L^1 is L^∞, and so $\|\varphi\| = \|a\|_\infty$, in this case $\|a\|_\infty = 1$.

(ii) If we consider the L^p norm on X, with $1 \le p < \infty$, we know that φ is still bounded, since $\|f\|_1 \le \|f\|_p$ if $p > 1$; still the completion of X in the L^p norm is $L^p([0,1])$ and φ extends by uniform continuity to a continuous linear functional with the same norm; since the dual of L^p is L^q with $q = p/(p-1)$ we get

$$\|\varphi\| = \|a\|_q = \left(\int_0^1 |a(t)|^q\,dt\right)^{1/q}$$

(in our case the norm could be explicitly computed with the beta and gamma functions). Finally, with the sup–norm X is a Banach space, with dual the space of Radon measures on $[0,1]$, and norm the total variation measure of $[0,1]$, that is

$$\|\varphi\| = \int_0^1 |a(t)|\,dt \quad \text{in our case} \quad \|\varphi\| = \frac{2}{\pi}$$

(In our case the direct computation of the norm is trivial, without recourse to general theorems, simply use $f(t) = 1$, because $a(t) = \sin(\pi t)$ has constant sign in $[0,1]$; it is less immediate when a changes sign).

(iii) If φ is continuous there is a finite subset $\{x_1, \dots, x_m\} \subseteq [0,1]$ and a constant $L > 0$ such that

$$|\varphi(f)| \le L \max\{|f(x_k)| : 1 \le k \le m\}, \quad \text{for every} \quad f \in X.$$

In particular, if f is zero on the finite set $\{x_1, \dots, x_m\}$ we have $\varphi(f) = 0$. This is plainly impossible: if a is continuous not identically 0, it has constant sign on some open interval I, and we may also assume, restricting I if necessary, that this interval does not contain any of the x_k; if f is the function $f(x) = \operatorname{dist}(x, [0,1] \setminus I)$ we have

$$\varphi(f) = \int_0^1 a(t)\, f(t)\, dt = \int_I a(t)\, f(t)\, dt \neq 0,$$

the last integral non–zero because $a\, f$ is a continuous function with constant sign on the open interval I □

EXERCISE 2.3.5.7. Let $(X, \mathcal{M}, \mu)$ be a finite measure space. Prove that if a sequence $f_n \in L(X)$ of measurable functions converges in measure to $f \in L(X)$, then some subsequence converges a.e to f.

Next, let $f_n,\ f \in L^2([0,1])$, and assume that $f_n \rightharpoonup f$ weakly in L^2. Let $g_n \in L^\infty([0,1])$ be a sequence bounded in L^∞ which converges in measure to a measurable function g.

(i) Prove that $g \in L^\infty$.

(ii) ⊙⊙ Prove that the sequence $f_n\, g_n$ converges to $f\, g$ weakly in L^2.

Solution. As observed in a previous exercise, f_n converges to f in measure if and only if the sequence $|f - f_n|/(1 + |f - f_n|)$ converges to 0 in $L^1(\mu)$. Then some subsequence of it converges to 0 a.e.; and since $1 + |f - f_n| \ge 1$ this clearly implies that the subsequence of the f_n's converges to f a.e..

(i) From what just proved we have $|g(x)| \le M$ for a.e. $x \in [0,1]$.

(ii) We have to prove that $(f_n\, g_n - f\, g \mid h) \to 0$ for every $h \in L^2$; write

$$(f_n\, g_n - f\, g \mid h) = (f_n\, g_n - f_n\, g + f_n\, g - f\, g \mid h) = (f_n(g_n - g) \mid h) + ((f_n - f)\, g \mid h) =$$
$$(f_n(g_n - g) \mid h) + (f_n - f \mid gh);$$

since $g\, h \in L^2$ and $f_n - f \rightharpoonup 0$ in L^2, the second term tends to 0; we have to show that the same is true for the first; fix $\varepsilon > 0$:

$$\left|\int_0^1 (f_n(g_n - g)\, h\right| \le \int_0^1 |g_n - g|\, |f_n|\, |h| = \int_{\{|g_n - g| > \varepsilon\}} |g_n - g|\, |f_n|\, |h| + \int_{\{|g_n - g| \le \varepsilon\}} |g_n - g|\, |f_n|\, |h|$$
$$\le \int_{\{|g_n - g| > \varepsilon\}} 2M\, |f_n|\, |h| + \int_{\{|g_n - g| \le \varepsilon\}} \varepsilon\, |f_n|\, |h|$$

Recall that f_n is bounded in L^2, being weakly convergent, so that $\|f_n\|_2 \le N$ for some $N > 0$. Then, writing $B(n)$ in place of $\{|g_n - g| \le \varepsilon\}$:

$$\int_{B(n)} |f_n|\, |h| \le \int_0^1 |f_n|\, |h| \le \|f_n\|_2\, \|h\|_2 \le N \|h\|_2;$$

and writing $A(n) = \{|g_n - g| > \varepsilon\}$ we get

$$\int_{A(n)} |f_n|\, |h| \le \left(\int_{A(n)} |f_n|^2\right)^{1/2} \left(\int_{A(n)} |h|^2\right)^{1/2} \le N \left(\int_{A(n)} |h|^2\right)^{1/2};$$

by absolute continuity, since $|h|^2 \in L^1$, given $\varepsilon > 0$ there is $\delta > 0$ such that if $\operatorname{mis}(A(n)) \le \delta$ then $\int_{A(n)} |h|^2 \le \varepsilon^2$; for $n \ge n_\varepsilon$ we have $\operatorname{mis}(A(n)) \le \delta$; thus, if $n \ge n_\varepsilon$:

$$\left|\int_0^1 (f_n(g_n - g)\, h\right| \le 2M\, N\, \varepsilon + \varepsilon\, N\, \|h\|_2.$$

□

2.3.6. *Examples.*

EXERCISE 2.3.6.1. Let X be a set, and let $\mathcal{K}$ be a set of subsets of X closed under union (if $K, L \in \mathcal{K}$ then $K \cup L \in \mathcal{K}$), and such that $\bigcup_{K\in\mathcal{K}} K = X$ (in most of the cases X is a topological space and $\mathcal{K}$ is the set of all compact subsets of X). For every $K \in \mathcal{K}$ and every $f \in \mathbb{K}^X$ we set:

$$p_K(f) = \|f\|_K := \sup_{x\in X} |f(x)| \quad \text{(finite or } +\infty).$$

($\mathbb{K}^X$ denotes the $\mathbb{K}-$linear space of all functions from X into $\mathbb{K}$).

(i) Prove that $p_K(f + g) \le p_K(f) + p_K(g)$ and $p_K(\lambda f) = |\lambda|\, p_K(f)$ (here $0 \cdot \infty = 0$ is used). Deduce that the set $\mathcal{B}(K) = \{f \in \mathbb{K}^X : p_K(f) < \infty\}$ is a linear subspace of $\mathbb{K}^X$, and that p_K is a seminorm on it.
(ii) Notice that $p_K \vee p_L = p_{K\cup L}$, for every $K, L \in \mathcal{K}$.
(iii) For $f \in X$, $K \in \mathcal{K}$, and $\varepsilon > 0$ we let:

$$B_K(f, \varepsilon[= \{g \in \mathbb{K}^X : p_K(g - f) < \varepsilon\}.$$

Show that $\{B_K(f, \varepsilon[: f \in X, K \in \mathcal{K}, \varepsilon > 0\}$ is a base for a Hausdorff topology on $\mathbb{K}^X$, called *topology of uniform convergence on members of* $\mathcal{K}$. Show also that in this topology addition and subtraction are continuous as mappings from $\mathbb{K}^X \times \mathbb{K}^X$ in $\mathbb{K}^X$ i.e., the additive group $(\mathbb{K}^X, +)$ is a *topological group*).
(iv) Assuming that $f \in \mathbb{K}^X$ is such that for some $K \in \mathcal{K}$ we have $p_K(f) = \infty$, show that the function from $\mathbb{K}$ into $\mathbb{K}^X$ given by $\lambda \mapsto \lambda f$ is never continuous, at no point of $\mathbb{K}$, and that the topology induced on the one–dimensional subspace $\mathbb{K} f$ is the discrete topology.
(v) The linear subspace $\mathcal{B}_\mathcal{K} \subseteq \mathbb{K}^X$ consisting of functions which are bounded on every member of $\mathcal{K}$ is a linear topological space under the induced topology.

If X is a topological space and $\mathcal{K}$ is the set of all compact subspaces of X, the above is the *topology of uniform convergence on compacta*, which in general is not a vector space topology on $\mathbb{K}^X$, but it is so on $\mathcal{C}(X, \mathbb{K})$, subspace of continuous $\mathbb{K}-$valued functions, because continuous functions are bounded on compacta.

Another interesting topology is obtained when $\mathcal{K}$ is the set of all finite subsets of X; it is a vector space topology on $\mathbb{K}^X$, and it is called *topology of pointwise convergence* on X. At the other extreme there is the case $\mathcal{K} = \{X\}$; the topology is that of uniform convergence on X, the strongest possible among topologies of this kind (not a vector space topology, only an additive group topology if X is infinite, but a Banach vector space topology on the subspace $\ell^\infty(X) = \ell^\infty(X, \mathbb{K})$ of bounded $\mathbb{K}-$valued functions, normed by the sup–norm $\|f\|_\infty = \sup_{x\in X} |f(x)|$).

2.3.7. *Metrizability; Fréchet spaces.*

PROPOSITION. *Let X be a linear space, and let $\{p_\alpha\}_{\alpha\in A}$ be a family of seminorms on X; let $\mathcal{T}$ be the topology of this family, assumed to be Hausdorff. The following are equivalent:*

(i) *The topology $\mathcal{T}$ is metrizable.*
(ii) *The topology $\mathcal{T}$ is first countable: that is, there is a countable base of neighborhoods at 0 (hence also at every point).*
(iii) *There exists a sequence $p_n \in \{p_\alpha\}_{\alpha\in A}$ whose topology coincides with $\mathcal{T}$.*

Proof. (i) implies (ii) is trivial. (ii) implies (iii): Let $\{U_0, U_1, \dots\}$ be a neighborhood base at 0 for $\mathcal{T}$. For every n there exists a finite subset $F(n)$ of A, and $r_n > 0$ such that if $q_= \bigvee_{k\in F(n)} p_k$ then $B_{q_n}(r_n)(:= \{y \in X : q_n(y) < r_n\}) \subseteq U_n$. The set $N = \bigcup_{n\in\mathbb{N}} F(n)$, a countable union of finite sets, is a countable subset of A; the topology $\mathcal{Q}$ of $\{p_k\}_{k\in N}$ is then coarser than $\mathcal{T}$, but since every neighborhood U_n in the $\mathcal{T}$ topology contains a neighborhood in the $\mathcal{Q}-$topology, the $\mathcal{Q}-$ topology is finer that the $\mathcal{T}$ topology, hence the two topologies coincide. The proof will be concluded from the following fact:

. *Let $\{p_n\}_{n\in\mathbb{N}}$ be a sequence of seminorms on X which generates a Hausdorff topology on X. Then the formula:*

$$\rho(x, y) = \max\{p_n(x - y) \wedge 2^{-n} : n \in \mathbb{N}\}, \qquad x, y \in X,$$

defines a metric on X whose topology is that of $\{p_n\}_{n\in\mathbb{N}}$.

Observe that ρ is translation invariant, in the sense that we have $\rho(x + h, y + h) = \rho(x, y)$ for every $x, y, h \in X$. Thus $\rho(x, y) = \rho(x - y, 0)$; setting

$$[x] = \max_{n\in\mathbb{N}} p_n(x) \wedge 2^{-n},$$

we have $\rho(x,y) = [x-y]$; to prove the triangle inequality for ρ we simply prove subadditivity of $[\cdot]$, which is easy given subadditivity of all p_n ($[\#]$ is a pseudonorm, see 2.2.10). Next observe that:

$$(*)\qquad B_\rho(0, 2^{-m}[= \{x \in X : [x] < 2^{-m}\} = \bigcap_{0\le n\le m} B_{p_n}(0, 2^{-m}[\quad \text{for every } m \in \mathbb{N}$$

which immediately implies that the topology of the metric ρ coincides with that of the family $\{p_n\}_{n\in\mathbb{N}}$. To prove (*): $[x] < 2^{-m}$ means $\max_{n\in\mathbb{N}}\{p_n(x), 2^{-n}\} < 2^{-m}$, equivalently $\min\{p_n(x),\, 2^{-n}\} < 2^{-m}$ for all $n \in \mathbb{N}$; if $n > m$ then $\min\{p_n(x), 2^{-n}\} \le 2^{-n} < 2^{-m}$, for all $x \in X$, so that (*) is equivalent to $\min\{p_n(x),\, 2^{-n}\} < 2^{-m}$ for every $n \le m$, and since $2^{-n} \ge 2^{-m}$ if $n \le m$, this is in turn equivalent to $p_n(x) < 2^{-m}$ for $n \le m$, proving the assertion. We may also observe that we have $\rho(\lambda\, x, \lambda\, y) \le \rho(x,y)$ for $|\lambda| \le 1$, since $[\lambda\, x] \le [x]$ if $|\lambda| \le 1$. □

2.3.8. *Fréchet spaces.* Locally convex metrizable and complete spaces are called *Fréchet spaces.* Banach spaces are of course also Fréchet spaces, but many Fréchet spaces are not normable. Observe that completeness in a metrizable TVS is expressible with the topology of the space:

. *If X is a metrizable TVS, and $\rho : X \times X \to [0, +\infty[$ is a translation–invariant metric which gives the topology of X, then a sequence x_n of X is Cauchy for the metric ρ if and only if given a neighborhood U of 0 there is $n_U \in \mathbb{N}$ such that $x_n - x_m \in U$ for $n, m \ge n_U$.*

Proof. Assume x_n Cauchy for ρ, and assume that U is nbhd of 0 in X. Then $B_\rho(0, \varepsilon[\subseteq U$; if $n_\varepsilon \in \mathbb{N}$ is such that $\rho(x_n, x_m) < \varepsilon$ for $n, m \ge n_\varepsilon$, by translation invariance we have $\rho(x_n - x_m, 0) < \varepsilon$, hence $x_n - x_m \in U$ for $n, m \ge n_\varepsilon$. The converse is trivial. □

2.3.9. *Completeness and total boundedness.* First a lemma

Lemma. *Let p, q be semimetrics on the set X. A subset E of X is totally bounded for $p \vee q$ if and only if it is totally bounded for p and for q. Consequently, if $p_1, \dots, p_m$ is a finite family of semimetrics on X, a subset E is $p_1 \vee \cdots \vee p_m$–totally bounded iff it is totally bounded for every semimetric p_k, $1 \le k \le m$.*

Proof. Assume that E is totally bounded for p and for q. Given $\varepsilon > 0$, there is a finite subset F of E such that $E \subseteq B_p(x, \varepsilon/2[$. For every $x \in F$, put $E(x) = E \cap B_p(x, \varepsilon[$. Each $E(x)$ is totally bounded for the semimetric q, being a subset of the totally bounded set E; then there exists a finite subset $G(x)$ of $E(x)$ such that $E(x) \subseteq \bigcup_{y\in E(x)} B_q(y, \varepsilon[$. If $G = \bigcup_{x\in F} G(x)$, then G is a finite subset of E, and we claim that $E \subseteq \bigcup_{y\in G} B_{p\vee q}(x, \varepsilon[$: given $z \in E$, if $z \in E(x)$ then $q(z, y) < \varepsilon$ for some $y \in G(x)$; since $z, y \in B_p(x, \varepsilon/2[$ we have also $p(z, y) < \varepsilon$, so that $p \vee q(z, y) < \varepsilon$. By induction this is true for every finite family of semimetrics. The reciprocal proposition is trivial. □

. *If the sequence p_n of seminorms determines the (Hausdorff) topology of X, and ρ is the associated metric, then a sequence $(x_k)_k$ of X converges to $x \in X$ if and only if $\lim_k p_n(x - x_k) = 0$ for every n; the sequence is Cauchy if and only if the sequence is Cauchy for every seminorm p_n. And a subset E of X is totally bounded for the metric ρ if and only if it is totally bounded for every seminorm p_n.*

Proof. Easy from the lemma. It is also true that X is complete iff it is complete in every seminorm, but we skip it (it requires some non difficult, but a little delicate acting with semimetric non metric spaces). □

2.3.10. *The Banach–Alaoglu theorem.* An immediate consequence of the Tychonoff product theorem, that a product of (infinitely many) compact spaces is compact, is the following:

. Theorem of Banach–Alaoglu *If X is a normed space the closed unit ball B_{X^*} of the dual is compact in the weak* topology.*

We omit the proof. We remark immediately the important

Corollary. *If X is a reflexive Banach space, then every closed convex bounded subset of X is compact in the weak topology.*

Proof. If X is reflexive, then the Banach–Alaoglu theorem says that the closed unit ball B_X of X is weakly compact. Every bounded weakly closed subset of X is then weakly compact, being a closed subset of the weakly compact space $r\, B_X$, for some $r > 0$. And we know that norm–closed convex sets are weakly closed (2.2.8). □

2.4. Separability and metrizability. The weak topology of an infinite dimensional space is never metrizable (see 2.4.2, at the end). However, when X^* is separable, then on bounded subsets of X the weak topology induces a metrizable topology, as we prove in 2.4.2. We discuss separability first.

2.4.1. *Separability.* Observe that:

. *A topological vector space is separable (as a topological space) if and only if it has a countably generated dense linear subspace.*

Proof. Trivially a linear subspace generated by a countable dense subset of X is dense in X. Assume now that $A = \{a_n : n \in \mathbb{N}\}$ generates a dense subspace V of X. Consider the vector subspace D generated by A over the field $\mathbb{F}$, where $\mathbb{F} = \mathbb{Q}$ in the case $\mathbb{K} = \mathbb{R}$, and $\mathbb{F} = \mathbb{Q} + i\,\mathbb{Q}$ if $\mathbb{K} = \mathbb{C}$. Clearly D is countable and its closure contains V, hence its closure is X. □

. *Let X be a normed space. If the dual X^* is separable, then X is separable.*

Proof. If X^* is separable, then so is every subspace, in particular the unit sphere $S^* = \{f \in X^* : \|f\|_{X^*} = 1\}$ is separable. Let $\{f_n\}_n$ be a countable dense subset of S^*. For every n pick $u_n \in X$, with $\|u_n\|_X = 1$, such that $|f_n(u_n)| > 1/2$. Let V be the subspace generated by $\{u_n\}_n$; we claim that V is a dense subspace of X; by the preceding argument X is then separable. If $\mathrm{cl}_X(V) \subsetneq X$, we have a continuous linear functional f of norm 1, i.e. with $f \in S^*$, which is identically zero on V. By density of $\{f_n\}_n$ in S^* there exist (infinitely many) $n \in \mathbb{N}$ such that $\|f - f_n\|_{X^*} < 1/2$. But $\|f - f_n\|_{X^*} \geq |(f - f_n)(u_n)| = |f(u_n) - f_n(u_n)| = |f_n(u_n)| > 1/2$, a contradiction.

□

The converse is not true: $\ell^1 = \ell^1(\mathbb{N})$ is separable, having the countably generated subspace $c_{00}(\mathbb{N})$ as a dense subspace, but its dual $\ell^\infty = \ell^\infty(\mathbb{N})$ is not separable, having a uniformly discrete subset of continuum cardinality (see 1.3.3.1).

2.4.2. *Metrizability.* For every countable subset D of X^* which separates points of X, the topology $\sigma(X, D)$ is a metrizable topology (see 2.3.7), certainly strictly weaker than $\sigma(X, X^*)$ if D generates a proper subspace of X^*. If X^* is separable, and $D \subseteq X^*$ is a countable dense subset of X^*, the topology $\sigma(X, X^*)$ and $\sigma(X, D)$ however agree on bounded subsets of X. Let us see that they agree on the closed unit ball B of X; this concludes the proof, since every bounded subset is contained in some $n\,B$, and the mapping $x \mapsto n\,x$ is a self–homeomorphism of X for every linear topology on X. We have to prove that, given $x \in B$ and a weak nbhd U of x in B, this contains a $\sigma(X, D)$ nbhd of x in B; we may assume that U is of the form $U(f, \varepsilon) \cap B = \{y \in B : |f(y) - f(x)| < \varepsilon\}$, in fact every weak nbhd of x in B contains a finite intersection of such nbhds. Given $\delta > 0$ we can pick $g \in D$ such that $\|f - g\|_{X^*} < \delta$; we want to prove that we can find δ so small that $\{y \in B : |g(y) - g(x)| < \delta\} \subseteq U(f, \varepsilon)$. In fact:

$$\begin{aligned}|f(y) - f(x)| =&|f(y) - g(y) + g(y) - g(x) + g(x) - f(x)| \leq \\ &\leq |f(y) - g(y)| + |g(y) - g(x)| + |g(x) - f(x)| = \\ &|(f - g)(y)| + |g(y) - g(x)| + |(g - f)(x)| \leq \|f - g\|_{X^*}\,\|y\| + \delta + \|f - g\|_{X^*}\,\|x\| \\ &< \delta + 2\|f - g\|_{X^*} < 3\delta,\end{aligned}$$

and we just take $\delta = \varepsilon/3$. An entirely analogous proof says that if X is separable then the weak* topology induces a metrizable topology on all norm bounded subsets of X^*. In particular, if X is separable then the closed unit ball of X^* is a compact metrizable space. We have proved:

. *Let X be a Banach space. If the dual space X^* is separable, then the weak topology is a metrizable topology on all norm–bounded subsets of X.The weak* topology is metrizable on all norm bounded subsets of X^*, and weakly* closed bounded subsets of X^* are metrizable compact spaces in the weak* topology. In particular, on a separable Hilbert spaces all weakly closed bounded subsets, in particular all closed bounded convex subsets are weakly compact.*

But the weak topology is *never metrizable on all of* X. A topology like $\sigma(X, V)$, where V is a point–separating subspace of $\mathrm{Hom}_{\mathbb{K}}(X, \mathbb{K})$ is metrizable if and only if V is countably generated; and a Banach space is never countably generated (unless it is finite dimensional): by Baire's theorem it cannot be a countable union of finite dimensional subspaces, which are closed nowhere dense subsets of it.

2.4.3. *Bounded subsets of a linear topological space.* In topological vector spaces there is a natural notion of boundedness. Recall that we defined: in a linear space X, a subset U *absorbs* the subset A if there is $t_A > 0$ such that $tU \supseteq A$ for every $t \geq t_A$ (2.1.11). We say that a subset A of a linear topological space is *bounded* if it is absorbed by every nbhd of 0. Since every nbhd of 0 contains a balanced nbhd of

0, A is bounded if and only if for every nbhd U of 0 there is $t > 0$ such that $tU \supseteq A$. Clearly subsets of bounded sets are also bounded, and finite unions of bounded sets are bounded.

EXERCISE 2.4.3.1. Let X be a vector space and $\{p_\alpha\}_{\alpha\in A}$ be a family of seminorms on X. Prove that $A \subseteq X$ is bounded in the topology generated by the family if and only if A is bounded for every seminorm p_α, for every $\alpha \in A$ (meaning that there is $M_\alpha > 0$ such that $p_\alpha(x) \le M_\alpha$, for every $x \in A$).

In particular, the notion of weakly bounded set given in 2.3.4.2 coincides with boundedness in the weak topology.

EXERCISE 2.4.3.2. No nbhd of 0 is bounded in the weak topology, if X is not finite dimensional. ⊙⊙ More generally, a locally convex space with a bounded nbhd of 0 is seminormable.

Solution. Every nbhd of 0 contains a linear subspace V of finite codimension; if $v \in V$ is non–zero, and $f \in X^*$ is such that $f(v) = 1$, then $U = \{x \in X : |f(x)| < 1\}$ is a (convex balanced) nbhd of 0, and tU does not contain V, for no $t > 0$: in fact $2t\,v \notin tU$. For the last part we only give a hint: we may suppose this nbhd to be convex and balanced; consider its Minkowski functional .... □

There is also a notion of *total boundedness*: a subset E is totally bounded in the linear topological space X if for every nbhd U of 0 there is a finite subset $F \subseteq E$ such that $E \subseteq F + U$.

EXERCISE 2.4.3.3. Prove that every compact subset of X is totally bounded, and that every totally bounded subset of X is bounded.

Solution. Assume that $E \subseteq X$ is compact. Given a nbhd U of 0, for every $x \in E$ the set $x + U$ is nbhd of x; by compactness there is a finite subset F of E such that $E \subseteq \bigcup_{x\in F}(x + U) = F + U$. Next, assume that E is totally bounded; we have to prove that for every nbhd V of 0 we have a $t > 0$ such that $E \subseteq tV$. Let U be a balanced nbhd of 0 such that $U + U \subseteq V$, and let F be a finite subset of E such that $E \subseteq F + U$. Since F is finite and U is absorbing there is $t > 1$ such that $F \subseteq tU$. Then $E \subseteq F + U \subseteq tU + U = t(U + t^{-1}U) \subseteq t(U + U) \subseteq tV$ (since U is balanced and $t > 1$ we have $t^{-1}U \subseteq U$). □

2.5. The Stone Weierstrass theorem. Towards the end of 1800 Weierstrass proved the following:

. THEOREM OF WEIERSTRASS. *Let $X \subseteq \mathbb{R}^n$ be compact. Then the $\mathbb{R}-$algebra $\mathbb{R}[x_1, \dots, x_n]$ of (the restrictions to X of) real polynomial functions is uniformly dense in $C(X,\mathbb{R})$.*

2.5.1. *Subalgebras of $C(X,\mathbb{K})$.* Weierstrass theorem was substantially generalized by the american mathematician Marshall Stone in the first half of 1900. We first recall what an algebra is: we only need to speak here of subalgebras of the $\mathbb{K}-$algebra $C(X,\mathbb{K})$ of continuous $\mathbb{K}-$valued functions on the topological space X: we say that $A \subseteq C(X,\mathbb{K})$ is a $\mathbb{K}-$subalgebra of $C(X,\mathbb{K})$ if it is a linear subspace closed under pointwise multiplication. We consider on $C(X,\mathbb{K})$ the *topology of uniform convergence on compacta*, called also *compact–open topology*, the locally convex topology generated by the seminorms

$$p_K(f) = \|f\|_K = \max\{|f(x)| : x \in K\}.$$

If X is compact, this topology is a of course the topology of uniform convergence on X. We know that this topology makes $C(X,\mathbb{K})$ into a linear topological space; and also into a topological algebra:

. *Pointwise multiplication $\mu(f, g) = f\,g$ is a continuous map $\mu : C(X,\mathbb{K}) \times C(X,\mathbb{K}) \to C(X,\mathbb{K})$.*

as easily follows from the inequality $\|f\,g\|_K \le \|f\|_K\,\|g\|_K$. Thus the closure of a $\mathbb{K}-$algebra is still a $\mathbb{K}-$algebra.

The Stone–Weierstrass theorem is essentially a "real" theorem, for $\mathbb{R}-$algebras. To adapt it to the complex case we need to consider $\mathbb{C}-$algebras A which are *closed under conjugation*, i.e. such that $f \in A$ implies $\bar{f} \in A$; equivalently, $f \in A$ implies $\mathrm{Re}\,f(= (f + \bar{f})/2) \in A$, so that $A = A \cap C(X,\mathbb{R}) + i\,(A \cap C(X,\mathbb{R}))$. It is easy to see that if A is closed under conjugation then so is its closure in the topology of uniform convergence on compact subsets of X (the mapping $f \mapsto \bar{f}$ preserves the seminorms and is then continuous).

2.5.2. *Statement of the theorem.* A $\mathbb{K}-$subalgebra A of $C(X,\mathbb{K})$ is said to *separate the points of X* if for any pair x, y, $x \ne y$ of *distinct* points of X there exists $f \in A$ such that $f(x) \ne f(y)$. Then we have (this version is not the most general, we shall prove more, but it is the one easy to remember):

. THE STONE–WEIERSTRASS THEOREM *Let X be a topological space. An $\mathbb{R}-$subalgebra of $C(X,\mathbb{R})$ that separates the point of X and contains the constant functions is dense in $C(X,\mathbb{R})$ in the topology of uniform convergence on compacta of X. A $\mathbb{C}-$subalgebra of $C(X,\mathbb{C})$ that separates the points of X, contains the constant functions, and is closed under conjugation is dense in $C(X,\mathbb{C})$ in the topology of uniform convergence on compacta of X.*

We now proceed to the proof.

2.5.3. *Sublattices of $C(X,\mathbb{R})$.* A subset $L\subseteq C(X,\mathbb{R})$ is a sublattice of $C(X,\mathbb{R})$ if it is closed under the lattice operations $\vee$ and $\wedge$:

$$f\vee g(x)=f(x)\vee g(x)=:\max\{f(x),\,g(x)\}\qquad f\wedge g(x)=f(x)\wedge g(x):=\min\{f(x),\,g(x)\}.$$

A linear subspace V of $C(X,\mathbb{R})$ is a sublattice of $C(X,\mathbb{R})$ if and only if $f\in V$ implies $|f|\in V$: in fact if V is a sublattice and $f\in V$ then $|f|=f\vee(-f)\in V$; and $f\vee g=(f+g)/2+|f-g|/2$, while $f\wedge g=(f+g)/2-|f-g|/2$, for $f,\,g\in C(X,\mathbb{R})$. We want to prove the following:

PROPOSITION. *The closure of a subalgebra A of $C(X,\mathbb{R})$ in the compact–open topology is sublattice of $C(X,\mathbb{R})$.*

The proof is at the end of the next number.

2.5.4. *Approximation of the square root.*

LEMMA. *There is a sequence of polynomials u_n with zero constant term which converges to the function $\sqrt{x}$ uniformly on the interval $[0,1]$.*

Proof. The following sequence arises out of Newton's method for constructing solutions of equations, applied to the equation $x-y^2=0$ in the unknown y for finding $\sqrt{x}$; it is defined inductively by $u_0(x)=0$ and

$$u_{n+1}(x)=u_n(x)+\frac{1}{2}(x-(u_n(x))^2).$$

Note that $u_n(0)=0$ for every $n\in\mathbb{N}$. Let us prove by induction that $0<u_n(x)<u_{n+1}(x)$ and that $(u_n(x))^2<x$ for $0<x\le 1$; the first assertion is a consequence of the second; squaring $u_{n+1}(x)>u_n(x)>0$ we get

$$\left(u_n(x)+\frac{1}{2}(x-(u_n(x))^2)\right)^2=(u_n(x))^2+u_n(x)\,(x-(u_n(x))^2)+\frac{(x-(u_n(x))^2)^2}{4}<x\iff$$
$$(x-(u_n(x))^2)\left(u_n(x)+\frac{x-(u_n(x))^2}{4}\right)<x-(u_n(x))^2;$$

since $x-(u_n(x))^2>0$ by the inductive hypothesis, the preceding relation is equivalent to

$$u_n(x)+\frac{x-(u_n(x))^2}{4}<1\iff(u_n(x))^2-4u_n(x)+4>x\iff(2-u_n(x))^2>x,$$

certainly true because $2-u_n(x)>2-x>2-1\ge x$. So u_n is an increasing sequence of polynomial functions, all 0 at 0, and all dominated by x in $[0,1]$, and hence pointwise convergent to $\ell(x)=\ell(x)+(x-(\ell(x))^2)/2\iff\ell(x)=\sqrt{x}$. Since the limit function is continuous, the monotone convergence is also uniform (Dini's theorem: given $\varepsilon>0$, $A_n(\varepsilon)=\{x\in[0,1]:\sqrt{x}-u_n(x)<\varepsilon\}$ is an increasing sequence of open subsets of $[0,1]$ whose union is the compact space $[0,1]\ldots$). □

Proof. (of 2.5.3) Given $f\in A$, a compact subset K of X, and $\varepsilon>0$, we find a polynomial p with zero constant term such that $\|p(f)-|f|\|_K\le\varepsilon$, which proves that $|f|\in\operatorname{cl}A$, and hence that $\operatorname{cl}A$ is a sublattice of $C(X,\mathbb{R})$. If $\|f\|_K=0$, then the zero polynomial will do. Otherwise, pick $n\in\mathbb{N}$ so that $\sqrt{t}-u_n(t)<\varepsilon/\|f\|_K$ for $t\in[0,1]$, where u_n is as in the previous lemma; for $x\in K$ we have $0\le\sqrt{(f/\|f\|_K)^2}(x)-u_n((f/\|f\|_K)^2)(x)<\varepsilon/\|f\|_K$, in other words

$$0\le|f(x)|-\|f\|_K\,u_n((f/\|f\|_K)^2)(x)<\varepsilon,\quad\text{for every }x\in K,$$

as desired.

□

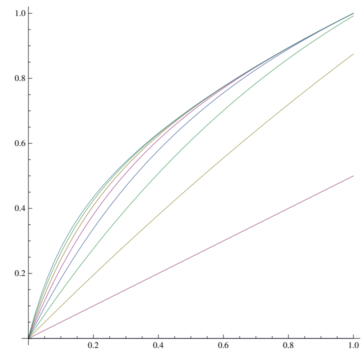

FIGURE 1. Some polynomials u_n.

2.5.5. *Closure of a sublattice.* In the next proof we use this terminology: given a subset L of $C(X) = C(X, \mathbb{R})$, and a function $f \in C(X)$, we say that f is *approximable at every pair of points of* X *by functions in* L if, given $\varepsilon > 0$ and $x, y \in X$ there exists $f_{xy} \in L$ such that

$$|f(x) - f_{xy}(x)| < \varepsilon, \qquad |f(y) - f_{xy}(y)| < \varepsilon.$$

then:

PROPOSITION. *Let L be a sublattice of $C(X, \mathbb{R})$. The closure of L in the compact–open topology is the set of all functions in $C(X, \mathbb{R})$ which are approximable at every pair of points by functions in L.*

Proof. Since one and two–elements subsets are compact sets, the condition is clearly necessary. Assume now that $f \in C(X)$ is approximable at every pair of points by functions in L; we want to prove that $f \in \operatorname{cl} L$, i.e that given $\varepsilon > 0$ and a compact subset K of X there is $g \in L$ such that $|f(x) - g(x)| < \varepsilon$ for every $x \in K$. We first prove that, given $a \in K$, there is a function $g_a \in L$ such that $|f(a) - g_a(a)| < \varepsilon$ and $g_a(x) > f(x) - \varepsilon$ for every $x \in K$. In fact we have, for every $x \in K$, a function $f_x \in L$ such that $|f(a) - f_x(a)| < \varepsilon$ and $|f(x) - f_x(x)| < \varepsilon$; let $U_x = \{y \in X : |f(y) - f_x(y)| < \varepsilon\}$. The set U_x is open in X and contains x. Then there exist $x(1), \dots, x(p) \in K$ such that $K \subseteq \bigcup_{j=1}^{p} U_{x(j)}$. Set $g_a = f_{x(1)} \vee \cdots \vee f_{x(p)}$: then $g_a(a) \in]f(a) - \varepsilon, f(a) + \varepsilon[$ and if $y \in K$ then $y \in U_{x(j)}$ for some $j \in \{1, \dots, p\}$ so that $g_a(y) \geq f_{x(j)}(y) > f(y) - \varepsilon$.

Thus for every $a \in K$ we get $g_a \in L$ as above: let $V_a = \{y \in X : |f(y) - g_a(y)| < \varepsilon\}$; since V_a is neighborhood of a we get $a(1), \dots, a(q) \in K$ such that $K \subseteq \bigcup_{k=1}^{q} V_{a(k)}$; if $g = g_{a(1)} \wedge \cdots \wedge g_{a(q)}$ we have $g \in L$ and $|f(y) - g(y)| < \varepsilon$ for every $y \in K$ (clearly $g(y) > f(y) - \varepsilon$ for every $y \in K$; and every $y \in K$ belongs to some $V_{a(k)}$, so that $g(y) \leq g_{a(k)}(y) < f(y) + \varepsilon$).

□

2.5.6. *A little algebra.* We need a simple algebraic:

LEMMA. *The only subalgebras of $\mathbb{R}^2$, under coordinatewise addition and multiplication, are $\{(0,0)\}$, $\mathbb{R} \times \{0\}$, $\{0\} \times \mathbb{R}$, the diagonal $\{(t,t) : t \in \mathbb{R}\}$, and $\mathbb{R} \times \mathbb{R}$.*

Proof. If A is a subalgebra of $\mathbb{R}^2$ and $(a, b) \in A$ with $a \neq b$ and a, b both non–zero, then (a, b) and (a^2, b^2) are linearly independent and thus generate all of $\mathbb{R}^2$ as a linear space. The remaining cases give the other four subalgebras. □

2.5.7. *Proof of the Stone–Weierstrass theorem.* Given a $\mathbb{K}-$subalgebra A of $C(X, \mathbb{K})$, its zero–set is the set

$$Z_A = \{x \in X : f(x) = 0 \text{ for every } f \in A\} = \bigcap_{f \in A} Z(f).$$

If A is point–separating its zero set is empty, or contains at most one point a; if A contains the constant functions clearly Z_A is empty. Then:

. *The closure in the compact–open topology of a point separating real subalgebra is the set of all continuous real functions which are 0 on the zero set of A. And if A is a conjugation closed point separating $\mathbb{C}-$algebra, then its closure is the set of all continuous complex valued functions which are 0 on the zero–set of A,* so this closure is either the maximal ideal $M_a = \{f \in C(X) : f(a) = 0\}$, or all of $C(X)$ if Z_A is empty. The complex case, in the hypothesis of conjugation–closure, comes immediately from the real case, so we assume that A is an $\mathbb{R}-$algebra of real functions.

Pick $x, y \in X$, $x \neq y$, and consider the restriction map $f \mapsto f_{|\{x,y\}}$; then $A_{|\{x,y\}}$ is a subalgebra of $\mathbb{R}^{\{x,y\}}$; if neither x nor y are in Z_A, point separation implies that $A_{|\{x,y\}} = \mathbb{R}^{\{x,y\}}$, by the algebraic lemma; by 2.5.5, if Z_A is empty this proves that $\operatorname{cl} A = C(X, \mathbb{R})$; if $Z_A = \{a\}$ and $x = a$, $y \neq a$ we get $A_{|\{x,y\}} = \{0\} \times \mathbb{R}$, and again by 2.5.5 we conclude that $\operatorname{cl} A = M_a$.

2.5.8. *Importance of conjugation.* Of course Weierstrass theorem is a particular case of the theorem just proved: the projection maps $(x_1, \dots, x_n) \mapsto x_j$ separate the points, and the presence of the constants implies that there is no common zero, so that the polynomial algebra is uniformly dense in $C(X, \mathbb{R})$ for every compact subset X of $\mathbb{R}^n$. Of course, if $x_1, \dots, x_n$ are *real* variables then $\mathbb{C}[x_1, \dots, x_n]$ is a conjugation closed point separating, constant containing $\mathbb{C}-$algebra, and is then uniformly dense in $C(X, \mathbb{C})$, if X is a compact subset of $\mathbb{R}^n$. If X is a compact subset of $\mathbb{C}^n$, then $\mathbb{C}[z_1, \dots, z_n]$ is *not* conjugation closed, and in general its uniform closure is strictly smaller than $C(X, \mathbb{C})$. For instance, take $n = 1$ and $X = \bar{\Delta}$, the closed unit disk of $\mathbb{C}$. The uniform closure of $\mathbb{C}[z]$ in $C(\bar{\Delta}, \mathbb{C})$ consists of functions which are continuous on $\bar{\Delta}$, and holomorphic on the open unit disc $\Delta = \{z \in \mathbb{C} : |z| < 1\}$ (recall that a function continuous on the simply connected open set Δ is holomorphic on Δ iff its integral on every circuit inside Δ is 0; if f is a uniform limit of functions holomorphic on Δ this is certainly true). So the conjugate function $z \mapsto \bar{z}$ is not in the uniform closure of the polynomial algebra on $\bar{\Delta}$.

2.5.9. *Trigonometric polynomials.* On the unit circle $\mathbb{U}$ we consider the subalgebra A of $C(\mathbb{U}) = C(\mathbb{U}, \mathbb{C})$ generated by $\{1, z, \bar{z}\}$. The elements of this algebra may be written as

$$\sum_{n=0}^{m} c_n z^n + \sum_{n=1}^{m} c_{-n} \bar{z}^n;$$

(they can also be written as $\sum_{n=-m}^{m} c_n z^n$, but while in the first form they represent a function continuous on the whole disc, in the second form there may be a pole at the origin). This algebra is clearly point separating and conjugation closed, and contains the constants, so it is uniformly dense in $C(\mathbb{U})$. Clearly then A is uniformly dense in $L^p(\mathbb{U})$ for every p, $1 \le p < \infty$; the measure on $\mathbb{U}$ is the normalized Lebesgue measure $dz/(2\pi i z)$ (the arc length divided by 2π). Recall that $\{z^n : n \ge 0\} \cup \{\bar{z}^n : n \ge 1\}$ are an orthonormal set for $L^2(\mathbb{U})$ (orthonormality of characters); for this we take advantage of $\bar{z} = 1/z$ on $\mathbb{U}$ and we get, for $m, n \in \mathbb{Z}$

$$\int_{\mathbb{U}} z^m \, \bar{z}^n \, \frac{dz}{2\pi i z} = \frac{1}{2\pi i} \int_{\mathbb{U}} z^{m-n-1} \, dz = \begin{cases} 0 & \text{if } m \neq n \\ 1 & \text{if } m = n \end{cases}.$$

Functions on $\mathbb{U}$ are often interpreted as periodic functions on $\mathbb{R}$; fix $\tau > 0$ and put $\omega = 2\pi/\tau$; then $f : \mathbb{U} \to \mathbb{C}$ becomes the $\tau-$periodic function $g_f(x) = f(e^{i\omega x})$ (by abuse of notation we often write $f(x)$ for $f(e^{i\omega x})$); to preserve integrals we use integration on a period long interval, divided by τ. Trigonometric polynomials become the functions

$$\sum_{n=-m}^{m} c_n \, e^{in\omega x};$$

the algebra of trigonometric polynomials is uniformly dense in the algebra af all continuous $\tau-$periodic functions on $\mathbb{R}$ (notice that on periodic continuous functions the uniform topology on $C_b(\mathbb{R}, \mathbb{C})$ and the topology of uniform convergence on compacta coincide: the sup–norm on $\mathbb{R}$ of a periodic function g is $\|g\|_{[0,\tau]}$. Trigonometric polynomials are then dense in all spaces L^p_τ; and since these spaces are, apart from a rescaling of the measure, identifiable with any $L^p([a, b])$ if $b - a = \tau$, we have:

. *The algebra of all trigonometric polynomials of period $b - a$, i.e. the algebra generated by the functions $\{e^{(2\pi/(b-a))ni\,x} : n \in \mathbb{Z}\}$ is dense in $L^p([a, b])$, for every p with $1 \le p < \infty$.*

Caution: not only trigonometric polynomials are not dense in $L^\infty([a, b])$, they are not even all of $C([a, b])$; it is in fact immediate to see that:

EXERCISE 2.5.9.1. The uniform closure in $C([a, b])$ of the algebra of trigonometric polynomials of period $b - a$ is the space of all $f \in C([a, b])$ such that $f(a) = f(b)$. Prove directly that this subspace of $C([a, b])$ is dense in $L^p([a, b])$ for all $p < \infty$.

Exercise 2.5.9.2. Let $f_n \in L^\infty([0,1])$ be a bounded sequence (i.e. there is M such that $\|f_n\|_\infty \le M$ for every n). Assume that for every $k \in \mathbb{Z}$ we have

$$\lim_{n\to\infty} c_k(f_n) = 0, \quad \text{where} \quad c_k(f_n) = \int_0^1 f_n(x)\, e^{-2\pi i k x}\, dx$$

is the $k-$th Fourier coefficient of f_n.

(i) Prove that then f_n converges to 0 weakly in $L^p([0,1])$, for every p, $1 \le p < \infty$.
(ii) Give an example of a sequence which satisfies all the requirements and does not converge strongly to 0 in $L^1([0,1])$.

Solution. (i) We have to prove that for every $q \ge 1$ and every $g \in L^q$ we have

$$\lim_{n\to\infty} \int_0^1 f_n\, g = 0.$$

Since $L^\infty([0,1]) \subseteq L^p([0,1])$ for all p, we have that f_n is in the dual space of every L^q, so that we have a sequence of continuous operators $T_n : L^q \to \mathbb{C}$

$$T_n g = \int_0^1 f_n\, g \quad \text{with norm} \quad \|T_n\|_{(L^q)^*} = \|f_n\|_p \le \|f_n\|_\infty \le M,$$

for every $n \in \mathbb{N}$. Assume first $q < \infty$, i.e $p > 1$. The statement on Fourier coefficients says that $T_n(e_{-k}) \to 0$ as $n \to \infty$ for every $e_k(x) = e^{2\pi i k x}$, hence for every trigonometric polynomial, and then for every $g \in L^q$, by 2.3.5, and by density of trigonometric polynomials in L^q. For $p = 1$ we can consider the operators $S_n : L^1 \to \mathbb{C}$ given by $S_n g = \int_0^1 f_n\, g$, and the same argument says that they converge to 0 on all of L^1, in other words $f_n \to 0$ in the weak* topology of L^∞. And since $L^\infty \subseteq L^1$, we have also $\int_0^1 f_n\, g \to 0$ for every $g \in L^\infty$, and hence weak convergence in L^1, as required.

(ii) The very sequence $e_n(x) = e^{2\pi i n x}$ satisfies all the requirements ($c_k(e_n) = 0$ if $k \ne n$) and is of constant norm 1 in L^1. □

Exercise 2.5.9.3. Consider the linear space $\mathbb{R}[x]$ of all real polynomials in one variable. Prove that $\|p\| = \max\{|p(x)| : x \in [0,1]\}$ is a norm on $\mathbb{R}[x]$. Is the functional $\varphi(p) = p(2)$ bounded on $\mathbb{R}[x]$? (you may use Weierstrass theorem, but there is a completely elementary solution …).

2.5.10. *Bernstein polynomials.* There is a direct proof, due to the russian mathematician Bernstein, that every continuous function on $[0,1]$ is a uniform limit of polynomials; the proof is constructive, and can be interpreted in a probabilistic setting (see [Folland], Thm 10.7). Given $f \in C([0,1])$ its $n-$th Bernstein polynomial is

$$B_n f(x) = \sum_{k=0}^{n} f\left(\frac{k}{n}\right) \binom{n}{k} x^k\, (1-x)^{n-k}.$$

then

. Theorem of Bernstein. *If $n \to \infty$ then $B_n f$ converges uniformly to f on $[0,1]$.*

2.5.11. *A general Stone–Weierstrass theorem.* Point separation is a condition easy to remember and in general easy to check, and it is customary to state the Stone–Weierstrass theorem using it. However, the same proof we gave in 2.5.7 yields easily the general result we are going to state. Given a real subalgebra A, or a conjugation closed complex subalgebra of $C(X, \mathbb{K})$, we define an equivalence relation $\overset{A}{\sim}$ on X by declaring that $x \overset{A}{\sim} y$ means $f(x) = f(y)$ for every $f \in A$ (that is x and y are $A-$equivalent if they are *not* separated by any $f \in A$). These equivalence classes are closed sets, and Z_A, if non–empty, is one of them; they are also called constancy sets for A, since clearly every $f \in A$ is constant on each such class. Clearly A and its closure have the same constancy sets. Then

. *Let X be a topological space and let A be a real subalgebra, or a conjugation closed complex subalgebra of $C(X, \mathbb{K})$. Then the closure of A in the compact–open topology is the algebra of all continuous functions on X which are zero on Z_A, and constant on each constancy set of A.*

The proof is identical to the one given in 2.5.7.

2.6. Compact linear operators.

Definition. Let X, Y be normed spaces. A linear operator $T : X \to Y$ is said to be compact if it maps bounded subsets of X into subsets with compact closure in Y.

2.6.1. *First properties.* Equivalently, the linear map $T : X \to Y$ is compact iff $\mathrm{cl}_Y(T(B_X))$ is compact. Since compact sets are bounded, compact operators are also bounded. And if Y is a Banach space, the closure of a subset of Y is compact iff that set is totally bounded, so that in this case compact operators could be called *totally bounded operators*, those that map bounded subsets of X into totally bounded subsets of Y. Of course:

. *If X, Y are normed spaces and $T : X \to Y$ is linear, T is compact if and only if for every bounded sequence $x_n \in X$ the sequence Tx_n has a converging subsequence in Y.*

Proof. Immediate. □

2.6.2. *The space of compact operators.* Given two normed spaces X, Y we denote $K(X, Y)$ the subset of $L(X, Y)$ consisting of compact operators. Then

. *$K(X, Y)$ is a linear subspace of $L(X, Y)$; and if Y is complete then $K(X, Y)$ is closed in $L(X, Y)$.*

Proof. If T is compact then αT is trivially compact for every $\alpha \in \mathbb{K}$. If $S, T : X \to Y$ are compact then $\mathrm{cl}_Y(S(B))$ and $\mathrm{cl}_Y(T(B))$ are compact subsets of Y ($B = B_X$ is the unit ball of X; then $\mathrm{cl}_Y(S(B))+$ $\mathrm{cl}_Y(T(B)$ is also compact and contains $(S + T)(B)$. FInally, assume that T is in the closure of $K(X, Y)$ in $L(X, Y)$. Given $\varepsilon > 0$. pick $S \in K(X, Y)$ such that $\|T - S\| \le \varepsilon/2$. Since $S(B_X)$ is totally bounded there are $x_1, \dots, x_p \in B_X$ such that

$$S(B_X) \subseteq \bigcup_{j=1}^{p} (S(x_j) + \varepsilon\, B_Y).$$

But then we have also

$$T(B_X) \subseteq \bigcup_{j=1}^{p} (T(x_j) + 3\varepsilon\, B_Y);$$

in fact, if $x \in B_X$,and $S(x) \in S(x_j) + \varepsilon\, B_Y$ we have

$$\|Tx - Tx_j\|_Y = \|Tx - Sx + Sx - Sx_j + Sx_j - Tx_j\|_Y \le \|(T - S)x\|_Y + \|Sx - Sx_j\|_Y + \|(S - T)x_j\|_Y \le 3\varepsilon.$$

□

We have encountered many compact operators in the exercises, e.g. the one described in exercise 1.5.2.4. Multiplication by a function $a \in \ell^p$, with $p < \infty$, gives a compact map of ℓ^∞ into ℓ^p. The mapping $Tf(x) = \int_0^x f(t)\, dt$ is a compact mapping of $C([0, 1])$ (with uniform norm) into itself.

2.6.3. *Compact operators under composition.* In a composition of bounded linear operators it is enough that one of them is compact for the composition to be compact:

. *Let X, Y, Z be normed spaces, and assume that $T : X \to Y$ is a compact linear operator. If $S : Y \to Z$ is bounded, then $S \circ T : X \to Z$ is compact. And if $S : Z \to X$ is bounded then $T \circ S : Z \to Y$ is compact.*

Proof. If $\mathrm{cl}_Y(T(B_X))$ is compact, then $S(\mathrm{cl}_Y(T(B_X)))$ is compact, being a continuous image of a compact set, and contains $S \circ T(B_X)$, so that $S \circ T$ is compact. And $S(B_Z)$ is bounded in X, so $T(S(B_Z))$ has compact closure in Y. □

Corollary. *If $T : X \to Y$ is compact and V is a linear subspace of X, then $T_{|V}$ is compact.*

Exercise 2.6.3.1. Let X be a normed space, Y a Banach space, $T : X \to Y$ a bounded linear operator. Assume that V is a dense linear subspace of X; prove that T is compact if and only if $T_{|V}$ is compact.

2.6.4. *Finite rank operators.* An important class of compact operators is that of *bounded operators of finite rank*, those bounded operators $T : X \to Y$ whose image $T(X)$ is finite–dimensional (by definition, $\mathrm{rank}(T) = \dim T(X)$): they are all compact operators, prove it. By the preceding theorem any limit in $L(X, Y)$ of operators of finite rank is a compact operator, if the range space Y is complete. When Y is a Hilbert space, this condition is also necessary:

. *If X is normed space, Y a Hilbert space, and $T : X \to Y$ a compact operator, then T belongs to the closure in $L(X, Y)$ of the subspace of operators of finite rank.*

Proof. Since $T(B)$ is totally bounded, given $\varepsilon > 0$ there are $x_1, \dots, x_p \in B = B_X$ such that $T(B) \subseteq \bigcup_{j=1}^p (T(x_j) + \varepsilon\, B_Y)$. Let V be the space generated by $\{T(x_j) : j = 1, \dots, p\}$, let $\pi_V : Y \to V \hookrightarrow Y$ be the orthogonal projection onto V, and set $S = \pi_V \circ T$. Then $\operatorname{rank}(S) \le p$, and if $x \in B$ then $T(x) \in T(x_j) + \varepsilon\, B_Y$ for some $j \in \{1, \dots, p\}$ and

$$\|Tx - Sx\|_Y = \|Tx - Tx_j + Tx_j - Sx\|_Y \le \|Tx - Tx_j\|_Y + \|\pi_V(Tx_j) - \pi_V(Tx)\|_Y \le \varepsilon + \|Tx_j - Tx\|_Y \le 2\varepsilon.$$

(recall that π_V has norm 1). So $\|Tx - Sx\|_Y \le 2\varepsilon$ for every $x \in B$, i.e. $\|T - S\|_{L(X,Y)} \le 2\varepsilon$. □

EXAMPLE 2.6.4.1. In Exercise 1.9.1.1 we considered two σ−finite measure spaces $(X, \mathcal{M}, \mu)$ and $(Y, \mathcal{N}, \nu)$ and $K : X \times Y \to \mathbb{K}$ belonging to $L^2(\mu \otimes \nu)$, and proved that the formula

$$Tf(x) = \int_Y K(x, y)\, f(y)\, d\nu(y)$$

defines a bounded linear operator from $L^2(\nu)$ into $L^2(\mu)$, with norm not larger than $\|K\|_2$. Let us prove that T is a compact operator, limit of operators of finite rank. By the definition of the product measure there is a sequence $K_n \in L^2(\mu) \otimes L^2(\nu)$ that converges in $L^2(\mu \otimes \nu)$ to K. Thus, it T_n is the operator associated to the kernel K_n, T_n converges to T in $L(L^2(\mu), L^2(\nu))$ (in fact $\|T - T_n\| \le \|K - K_n\|_2$). Now every K_n is of finite rank: we have $K_n(x, y) = \sum_{k=1}^{m(n)} u_{nk}(x)\, v_{nk}(y)$, with $u_{nk} \in L^2(\mu)$ and $v_{nk} \in L^2(\nu)$, so that:

$$T_n f(x) = \int_Y \sum_{k=1}^{m(n)} u_{nk}(x)\, v_{nk}(y)\, f(y)\, d\nu(y) = \sum_{k=1}^{m(n)} u_{nk}(x) \int_Y v_{nk}(y)\, f(y)\, d\nu(y) = \sum_{k=1}^{m(n)} c_{nk}(f)\, u_{nk}(x),$$

with

$$c_{nk}(f) = \int_Y v_{nk}(y)\, f(y)\, d\nu(y),$$

so that the image of T_n is contained in the subspace of $L^2(\mu)$ spanned by u_{nk}, $k = 1, \dots, m(n)$.

EXERCISE 2.6.4.2. Let X, Y be Banach spaces, and let $T : X \to Y$ be bounded with closed image, i.e. $T(X)$ is a closed subspace of Y. Prove that T is compact if and only if T has finite rank (hint: open mapping theorem).

Compact operators transform weakly convergent sequences into strongly convergent sequences:

EXERCISE 2.6.4.3. Let X and Y be normed spaces, and let $T : X \to Y$ be a compact linear operator. Then for every sequence x_n weakly convergent to $x \in X$ the sequence Tx_n strongly converges to Tx in Y. Prove next the converse: if X is reflexive, X^* is separable and $T : X \to Y$ is a linear operator which transforms weakly convergent sequences of X into strongly convergent sequences of Y, then T is compact.

Solution. If x_n is weakly convergent then it is norm–bounded and hence Tx_n is contained in a compact subset of Y. Let's argue by contradiction: if Tx_n does not converge strongly to $y = Tx$ there is $\varepsilon > 0$ and a subsequence $x_{\mu(k)}$ of x_n such that $y_{\mu(k)} \notin y + \varepsilon\, B_Y$ for all k. Since the sequence $y_{\mu(k)}$ is contained in a compact subset of Y, it has a subsequence strongly converging to some $z \in Y$, and we may as well assume that $y_{\mu(k)} \to z$ as $k \to \infty$; of course $\|z - y\| \ge \varepsilon$. Let $\phi \in Y^*$ be such that $\phi(y) = 0$ and $\phi(z) = 1$. Then $\lim_{n\to\infty} \phi \circ T(x_n) = \phi \circ T(x) = \phi(y) = 0$; but $\lim_{k\to\infty} \phi \circ T(x_{\mu(k)}) = \lim_{k\to\infty} \phi(y_{\mu(k)}) = \phi(z) = 1$, a contradiction because this is a subsequence of the previous one.

If X is reflexive and X^* is separable the closed unit ball is weakly metrizable and weakly compact, and the condition then says that the image $T(B_X)$ is compact in the strong topology. □

2.6.5. *Transpose and adjoint of a compact operator.* It is not overly difficult to prove the following:

. *Let X and Y be Banach spaces, and let $T : X \to Y$ be a bounded linear operator. Then T is compact if and only if its transpose $T^t : Y^* \to X^*$ is compact.*

However we omit the proof (but see next exercise), and prove instead the easier version for Hilbert spaces:

. COMPACTNESS OF THE ADJOINT *Let $T : X \to Y$ be a bounded linear operator, where X and Y are Hilbert spaces. Then T is compact if and only if its adjoint $T^* : Y \to X$ is compact.*

Proof. If T is compact, by 2.6.4 it is the limit in $L(X, Y)$ of a sequence T_n of finite rank operators. The adjoint S^* of a finite rank operator S is of course also of finite rank (in fact the same rank!) since its image is contained in the orthogonal of $\operatorname{Ker} S$, a space of finite codimension. Since $T \mapsto T^*$ is an isometry of $L(X, Y)$ onto $L(Y, X)$ we have $\|T^* - T_n^*\|_{L(Y,X)} = \|T - T_n\|_{L(X,Y)}$ so T_n^* converges to T^*. □

Exercise 2.6.5.1. Prove that the transpose of the compact operator $T : X \to Y$ of normed spaces is compact.

Solution. Let $K = \mathrm{cl}_Y(T(B_X))$; by hypothesis K is compact. Consider the set $F = \{\phi_K : \phi \in B^*_Y\}$ of the restrictions to K of the functionals in the unit ball of Y^*; it is a subset of $C(K, \mathbb{K})$ equicontinuous (all functions have Lipschitz constant ≤ 1) and uniformly bounded (by $\max\{\|y\|_Y : y \in K\}$. Then F is totally bounded in $C(K, \mathbb{K})$, considered with the uniform norm, by Ascoli–Arzelà theorem, and given $\varepsilon > 0$ we find $\phi_1, \dots, \phi_m \in F$ such that for every $\phi \in F$ there is $k \in \{1, \dots, m\}$ such that $\|\phi - \phi_k\|_K \le \varepsilon$. This immediately implies that $T^t(B^*_Y)$ is totally bounded: given $\phi \in B^*_Y$, with the same $k \in \{1, \dots, m\}$ we have

$$\begin{aligned}\|T^t(\phi) - T^t(\phi_k)\|_{X^*} &= \sup\{|\phi \circ T(x) - \phi_k \circ T(x)| : x \in B_X\} = \sup\{|\phi(y) - \phi_k(y)| : y \in T(B_x)\}\\ &\le \sup\{|\phi(y) - \phi_k(y)| : y \in K\} \le \varepsilon.\end{aligned}$$

□

2.7. Spectrum of a continuous operator. We consider a Banach space, and the Banach algebra $L(X)$ of continuous operators of X into itself. Remember that $\|ST\| \le \|S\| \|T\|$ for every couple of operators $S, T \in L(X)$; we omit the $\circ$ in the composition. We call $G(X)$ the group of invertible elements of $L(X)$. Recall that if $\|T\| < 1$ then $1 - T \in G(X)$, and

$$(1 - T)^{-1} = \sum_{n=0}^{\infty} T^n;$$

in fact the series is normally convergent since $\|T^n\| \le \|T\|^n$ and $\|T\| < 1$ by hypothesis; moreover

$$(1 - T) \sum_{n=0}^{m} T^n = \left(\sum_{n=0}^{m} T^n\right)(1 - T) = 1 - T^{m+1} \to 1 \quad \text{as} \quad m \to \infty.$$

It follows that $G(X)$ is an open subset of $L(X)$; given $A \in G(X)$ consider $A - T = A(1 - A^{-1}T)$; if $\|A^{-1}T\| < 1$ then $A - T$ is invertible, and if $\|T\| < \|A^{-1}\|^{-1}$ this happens. Given an operator $T \in L(X)$ we define the *resolvent set* of T in $\mathbb{K}$ as $\rho(T) = \{\zeta \in \mathbb{C} : \zeta - T \in G(X)\}$, and the spectrum $\sigma(T) = \mathbb{K} \setminus \rho(T)$. Clearly $\rho(T)$ is open in $\mathbb{K}$, since $\rho(T)$ is the inverse image of the open set $G(X)$ under the map $\zeta \mapsto \zeta - T$, continuous from $\mathbb{K}$ to $L(X)$. Then $\sigma(T)$ is closed; and it is also bounded, since:

. *If $|\zeta| > \|T\|$, then $\zeta \in \rho(T)$.*

In fact if $|\zeta| > \|T\|$ we have $\|T\|/|\zeta| < 1$ so that:

$$(\zeta - T)^{-1} = \zeta^{-1}(1 - \zeta^{-1}T)^{-1} = \frac{1}{\zeta} \sum_{n=0}^{\infty} \frac{T^n}{\zeta^n} = \sum_{n=1}^{\infty} \frac{T^{n-1}}{\zeta^n}.$$

Thus $\sigma(T)$ is compact, and contained in $\{\zeta \in \mathbb{K} : |\zeta| \le \|T\|\}$. The resolvent function of T is the function $R_T : \rho(T) \to G(X)$ defined by $R_T(\zeta) = (\zeta - T)^{-1}$. It is trivial to verify that R_T is analytic in $\rho(T)$; if $c \in \rho(T)$ write

$$(\zeta - T)^{-1} = ((\zeta - c) - (T - c))^{-1} = -(T - c)^{-1}(1 - (\zeta - c)(T - c)^{-1})^{-1} = -(T - c)^{-1} \sum_{n=0}^{\infty} (\zeta - c)^n (T - c)^{-n},$$

provided that $|\zeta - c| < \|(T - c)^{-1}\|^{-1}$; that is, in the disc centered at c with this radius we have

$$R_T(\zeta) = \sum_{n=0}^{\infty} (\zeta - c)^n(-(T - c)^{-(n+1)}),$$

with normal convergence in $L(X)$ of the series. Then we have:

. *In a complex Banach space the spectrum of a bounded linear operator is a compact non–empty subset of $\mathbb{C}$.*

Proof. In fact an empty spectrum says that R_T is analytic on $\mathbb{C}$. Given a continuous linear $\phi \in (L(X))^*$ we have that $\phi \circ R_T : \mathbb{C} \to \mathbb{C}$ is an everywhere analytic, hence holomorphic entire , function. For $|\zeta| > \|T\|$ we have

$$\phi \circ R_T(\zeta) = \sum_{n=1}^{\infty} \frac{c_n}{\zeta^n} \quad \text{with} \quad c_n = \phi(T^{n-1}).$$

Then $\lim_{\zeta\to\infty} \phi \circ R_T(\zeta) = 0$, so that if the function is everywhere continuous it is bounded, hence constant by Liouville's theorem, hence actually everywhere zero. Since this is true for every continuous linear functional we have that $R_T(\zeta) = 0$, plainly impossible. □

Elements of the spectrum are of two types: the point spectrum, consisting of *eigenvalues*, those $\lambda \in \mathbb{K}$ such that $\operatorname{Ker}(\lambda - T) \neq \{0\}$ (in this case $\operatorname{Ker}(\lambda - T)$ is the *eigenspace* associated to λ), and the others, for which $\lambda - T$ is injective but not surjective (a finer distinction is made in Exercise 2.8.7.7). When T has finite rank then its spectrum has only eigenvalues (see Exercise 2.8.1.1) But otherwise an operator can have no eigenvalues at all.

Example 2.7.0.2. Consider $X = \ell^2(\mathbb{N})$. The *shift operator* $S : X \to X$ is given by the formula $Sx = (0, x_0, x_1, \dots)$. It clearly is isometric, hence injective, but not surjective, since $e_0 = (1, 0, 0, \dots)$ is not in the image of S; thus $0 \in \sigma(S)$, and it is easy to see that S has no eigenvalues.

Exercise 2.7.0.3. Work out the spectrum of $Tf(x) = \int_0^x f(t)\,dt$ of $C([0,1])$.

Solution. Clearly $0 \in \sigma(T)$, since the operator is not surjective; but it is obviously injective, so that 0 is not an eigenvalue. Assume now $\lambda \neq 0$. We prove that $\lambda - T$ is bijective, so that $\lambda \in \rho(T)$. We have to prove that given $g \in X$ there exists exactly one $f \in X$ such that $\lambda f - Tf = g$. If $F(x) = Tf(x)$ we have $F(0) = 0$ and F must verify the differential equation

$$\lambda F'(x) - F(x) = g(x) \qquad F(0) = 0.$$

this is a Cauchy problem for an affine first order differential equation, and so will have exactly one solution F, whose derivative is the required f. Thus $\rho(T) = \mathbb{C} \setminus \{0\}$. The exercise is concluded.

As a further exercise we compute a formula for the inverse $(\lambda - T)^{-1}$:

$$F'(x) - \frac{1}{\lambda} F(x) = \frac{1}{\lambda}\, g(x) \iff e^{-x/\lambda}\, F'(x) - \frac{e^{-x/\lambda}}{\lambda}\, F(x) = \frac{e^{-x/\lambda}}{\lambda}\, g(x);$$

the left hand side is the derivative of $e^{-x/\lambda}\, F(x)$; integrating both sides from 0 to x we get, remembering that $F(0) = 0$:

$$e^{-x/\lambda}\, F(x) = \int_0^x \frac{e^{-t/\lambda}}{\lambda}\, g(t)\,dt \iff F(x) = \frac{1}{\lambda} \int_0^x e^{(x-t)/\lambda}\, g(t)\,dt.$$

so that, differentiating:

$$f(x) = \frac{1}{\lambda^2} \int_0^x e^{(x-t)/\lambda}\, g(t)\,dt + \frac{1}{\lambda}\, g(x) (= ((\lambda - T)^{-1} g)(x)).$$

□

Exercise 2.7.0.4. Let E be a compact metrizable space and $X = C(E, \mathbb{K})$ with uniform norm. Given $g \in X$, the operator of multiplication by g is $T : X \to X$ defined by $T(f) = g\,f$. Prove that T is bounded of norm $\|g\|_u$, and that $\sigma(T) = g(E)$. When is a $\lambda \in g(E)$ an eigenvalue?

Solution. We have

$$\|Tf\|_u = \|g\,f\|_u \le \|g\|_u\, \|f\|_u, \quad \text{thus} \quad \|T\| \le \|g\|_u,$$

and if $f = 1$ we get equality, so that $\|T\| = \|g\|_u$. Clearly $\lambda - T$ is multiplication by $\lambda - g$, invertible if and only if $\lambda - g$ is never 0, i.e. $\lambda \notin g(E)$. For $\lambda \in g(E)$ to be an eigenvalue we need that $(\lambda - g)\,h = 0$ for some nonzero $h \in C(E)$, which clearly happens if and only if the interior of $\{x \in E : g(x) = \lambda\}$ is non–empty. □

Exercise 2.7.0.5. Let X be a Banach space.

(i) If $S, T \in L(X)$ are permutable then ST is invertible if and only if S and T are both invertible. Show by example that TS may be invertible with S, T non invertible if $ST \neq TS$.
(ii) Let now $\mathbb{K} = \mathbb{C}$, let $T \in L(X)$ be a bounded operator and $p \in \mathbb{C}[\zeta]$ a polynomial, with zero–set $Z(p)$. Prove that $p(T)$ is invertible if and only if $Z(p) \cap \sigma(T) = \emptyset$, and deduce that $\sigma(p(T)) = p(\sigma(T))$.
(iii) Extend (ii) to the case $p \in \mathbb{C}(\zeta)$, p a rational function with no poles on $\sigma(T)$.

Solution. (i) A product of invertible elements is invertible even if they do not commute. If R is the inverse of $ST = TS$ then $R(ST) = (ST)R = 1$ implies $(RS)T = T(SR) = 1$ so that T has both a left and a right inverse, hence an inverse; same for S. If S and T are the left and right shift in $\ell^2(\mathbb{N})$ then $TS = 1$, but neither is invertible, and ST is the orthogonal projection onto the subspace spanned by all the e_n's but the first.

(ii) Let's factor the polynomial in $\mathbb{C}[\zeta]$:

$$p(\zeta) = a\prod_{k=1}^{m}(\zeta - \lambda_k),$$

where $\lambda_1, \dots, \lambda_m$ are the zeros of p, repeated according to multiplicity, in particular $Z(p) = \{\lambda_1, \dots, \lambda_m\}$. From (i) follows that $p(T) = a\prod_{k=1}^{m}(T - \lambda_k)$ is invertible if and only if $\lambda_k \in \rho(T)$ for every $k = 1, \dots, m$, equivalently $Z(p) \cap \sigma(T) = \emptyset$. If $p \in \mathbb{C}[\zeta]$ and $\lambda \in \mathbb{C}$ the polynomial $p(\zeta) - \lambda$ has zeroes in $\sigma(T)$ if and only if $\lambda = p(\alpha)$ for some $\alpha \in \sigma(T)$, i.e. iff $\lambda \in p(\sigma(T))$; then $p(T) - \lambda$ is non–invertible, equivalently $\lambda \in \sigma(p(T))$, iff $\lambda \in p(\sigma(T))$.

(iii) This function may be written $p(\zeta) = a\prod_{j=1}^{m}(\zeta - \alpha_j)^{m_j}\prod_{k=1}^{n}(\zeta - \beta_k)^{-n_k}$, where $m_j,\ n_k \geq 1$, and $\beta_k \notin \sigma(T)$ for $k = 1, \dots, n$; an argument entirely analogous to the preceding one allows us to conclude. □

2.8. Spectrum of a compact operator.

2.8.1. *Preliminaries.*

. *Let X be Banach space, and let $T : X \to X$ be compact. Then:*

(i) $\mathrm{Ker}(1 - T)$ *is finite–dimensional.*
(ii) *The image* $(1 - T)(X)$ *of* X *by* $1 - T$ *is closed in* X*.*
(iii) *If* $\mathrm{Ker}(1 - T) = \{0\}$ *then* $1 - T$ *is a linear automorphism of* X*.*

Proof. (i) On $\mathrm{Ker}(1 - T)$ the operator T induces the identity; since this must be a compact map, $\mathrm{Ker}\,T$ is finite dimensional.

(ii) We use the criterion in 1.8.0.3: setting $K = \mathrm{Ker}(1 - T)$ we prove that there is $k > 0$ such that $\|x - Tx\| \geq k\ \mathrm{dist}(x, K)$ for every $x \in X$; if such a k does not exist, then there is a sequence $x_n \in X \setminus K$ such that

$$\lim_{n\to\infty}\frac{\|x_n - Tx_n\|}{\mathrm{dist}(x_n, K)} = 0.$$

Since K is finite dimensional, its closed bounded subsets are compact, hence for every n there is $z_n \in K$ such that $\|x_n - z_n\| = \mathrm{dist}(x_n, K)$; since $z_n \in K$, we have $x_n - Tx_n = (x_n - z_n) - T(x_n - z_n)$ so that

$$\frac{\|x_n - Tx_n\|}{\mathrm{dist}(x_n, K)} = \frac{\|(x_n - z_n) - T(x_n - z_n)\|}{\|x_n - z_n\|} = \left\|\frac{x_n - z_n}{\|x_n - z_n\|} - T\left(\frac{x_n - z_n}{\|x_n - z_n\|}\right)\right\| = \|u_n - Tu_n\|,$$

where we have set $u_n = (x_n - z_n)/\|x_n - z_n\|$; notice that $\|u_n\| = 1$ and $\mathrm{dist}(u_n, K) = 1$. We get a contradiction with $u_n - Tu_n \to 0$; since u_n is bounded, Tu_n has a converging subsequence (and we still call it u_n); if v is the limit, since $\lim_n(u_n - Tu_n) = 0$ we get $\lim_{n\to\infty} u_n = v$; then $v - Tv = 0$, hence $v \in K$, and hence also $\lim_{n\to\infty}\mathrm{dist}(u_n, K) = 0$, contradicting the fact that $\mathrm{dist}(u_n, K) = 1$ for every $n \in \mathbb{N}$.

(iii) Set $S = 1 - T$, and consider the iterates S^n of S; they all are of the form $1 - L_n$ with L_n a compact operator of X, as is easy to recognize by induction ($S^{n+1} = S\,S^n = (1 - T)(1 - L_n) = 1 - L_n - T + TL_n$; recall that a sum of compact operators is compact). The images $E_n = S^n(X)$ form then a decreasing sequence of closed subspaces of X, $X = E_0 \supseteq E_1 \supseteq E_2 \supseteq \dots$. If for some n we have $E_n = E_{n+1}$, then $E_m = E_n$ for all $m \geq n$ (simply apply S to $S^n(X) = S^{n+1}(X)$ to get $S^{n+1}(X) = S^{n+2}(X)$, then $E_n = E_{n+1} = E_{n+2}$, etc.)

The sequence must stop. Otherwise, Riesz lemma 1.5 allows us to pick a vector $u_n \in E_n \setminus E_{n+1}$, with $\|u_n\| = 1$, such that $\mathrm{dist}(u_n, E_{n+1}) \geq 1/2$; then, if $m > n$ we have

$$\|Tu_n - Tu_m\| = \|Tu_n - u_n + u_n - (Tu_m - u_m + u_m)\| = \|u_n - (u_m - Su_m + Su_n)\| \geq \mathrm{dist}(u_n, E_{n+1}) \geq 1/2,$$

(simply observe that $u_m - Su_m + Su_n \in E_{n+1}$) and this contradicts compactness of T.

If S is injective and $E_n = E_{n+1}$ then S is surjective: given $y \in X$, there is $x \in X$ such that $S^n y = S^{n+1}x$, that is $S^n y = S^n(Sx)$; since S^n is injective then $y = Sx$, so that $E_1 = S(X) = X$. □

Remark. If $K_n = \mathrm{Ker}(S^n)$, then $K_0 = \{0\} \subseteq K_1 \subseteq K_2 \subseteq K_3 \subseteq \dots$; these spaces are all finite dimensional. It can be proved that also this sequence becomes stationary, exactly at the same "instant" n as the sequence E_n does, and X is the topological direct sum of K_n and E_n.

Corollary. *If $\lambda \neq 0$ is in the spectrum of the compact operator T, then λ is an eigenvalue of T; and the eigenspace is finite–dimensional.*

Proof. Since $\lambda - T = \lambda(1 - \lambda^{-1}T)$ and $\lambda^{-1}T$ is also compact we are reduced to the case $\lambda = 1$. Then (i) of the preceding theorem says that $\mathrm{Ker}(1-T)$ is finite–dimensional; and if $\mathrm{Ker}(1-T) = \{0\}$ then (iii) of the preceding theorem says that $1-T$ is also onto X, so that $1 - T \in G(X)$ by the open mapping theorem, and $1 \notin \sigma(T)$. □

Exercise 2.8.1.1. Let $T : X \to X$ be of finite rank. Then T has a finite spectrum consisting only of eigenvalues.

Solution. This is trivially true if X is finite dimensional. Assume then that $\dim X$ is infinite; let $E = T(X)$ and $K = \mathrm{Ker}\, T$. Clearly $K \neq \{0\}$, since K has finite codimension, so $0 \in \sigma(T)$ is an eigenvalue; and T induces an operator from E to E, whose eigenvalues are also eigenvalues for T. We claim that 0 and these other eigenvalues make the spectrum of T. In fact T is compact, being of finite rank, so non zero elements of the spectrum are eigenvalues; and a nonzero eigenvalue has an eigenspace necessarily contained in $T(X)$. □

2.8.2. *Structure of the spectrum of a compact operator.* Unless X is finite–dimensional the number $\{0\}$ is always in the spectrum of a compact operator: a compact operator in an infinite dimensional space is never surjective (see exercise 2.6.4.2); it can be an eigenvalue or not (think about multiplication by some $a \in c_0$ as a compact mapping $T : \ell^\infty \to \ell^\infty$, $T(x) = a\,x$: then $\{0\}$ is an eigenvalue if and only if $a(n) = 0$ for some n. The spectrum of a compact operator can be finite, and even contain only 0. If infinite, being compact it must have accumulation points.

We now prove that for a compact operator $\{0\}$ is the only possible accumulation point of the spectrum.

. Let $T \in L(X)$ be a compact operator in the Banach space X. For every $\alpha > 0$ the set of eigenvalues of T with absolute value larger than α is finite.

Proof. Argue by contradiction and assume that $(\lambda_n)_{n\ge1}$ is a sequence of distinct eigenvalues with $|\lambda_n| \ge \alpha$ for every $n \ge 1$. For every $n \ge 1$ pick a non–zero $e_n \in \mathrm{Ker}(\lambda_n - T)$. We know that eigenvectors associated to distinct eigenvalues are linearly independent, so that if $E_n = \langle e_1, \dots, e_n\rangle$ is the vector space spanned by the first n eigenvectors, the sequence E_n is strictly increasing, $E_0 = \{0\} \subsetneq E_1 \subsetneq E_2 \subsetneq \dots$. Thus Riesz lemma allows us to pick a sequence $u_n \in E_n \setminus E_{n-1}$, for $n \ge 1$, such that $\|u_n\| = 1$ and $\mathrm{dist}(u_n, E_{n-1}) \ge 1/2$. If $1 \le m < n$ we get:

$$\left\|\frac{Tu_n}{\lambda_n} - \frac{Tu_m}{\lambda_m}\right\| = \left\|\frac{Tu_n - \lambda_n u_n + \lambda_n u_n}{\lambda_n} - \frac{Tu_m}{\lambda_m}\right\| =$$
$$\left\|\frac{Tu_n - \lambda_n u_n}{\lambda_n} - \frac{Tu_m}{\lambda_m} + u_n\right\| = \|u_n - v\| \ge \mathrm{dist}(u_n, E_{n-1}) \ge \frac{1}{2}$$

(simply observe that $v = \lambda_n^{-1}(\lambda_n u_n - Tu_n) + \lambda_m^{-1} Tu_m \in E_{n-1}$; the key fact is that $(\lambda_n - T)(E_n) \subseteq E_{n-1}$, and all spaces E_m are stable under the action of T and $\lambda - T$, since $T(e_k) = \lambda_k e_k$). Thus we have

$$\left\|T\left(\frac{u_n}{\lambda_n}\right) - T\left(\frac{u_m}{\lambda_m}\right)\right\| \ge \frac{1}{2} \quad \text{for } 1 \le m < n.$$

If $|\lambda_k| \ge \alpha$ for every k the sequence u_n/λ_n is bounded (by $1/\alpha$), and the preceding inequality contradicts compactness of T. □

Corollary. *Let $T \in L(X)$ be a compact operator with infinite spectrum. Then $\sigma(T)$ is countably infinite, and has 0 as its unique accumulation point: for every bijective indexing $(\lambda_n)_{n\ge1}$ of $\sigma(T) \setminus \{0\}$, we have $\lim_{n\to\infty} \lambda_n = 0$.*

Proof. For $n \ge 1$ let $S_n = \{\lambda \in \sigma(T) : |\lambda| \ge 1/n\}$; we have just proved that S_n is finite. Then $\sigma(T) \setminus \{0\} = \bigcup_{n\ge1} S_n$, a countable union of finite set, is countable. And if $n \mapsto \lambda_n$ is a bijective indexing of $\sigma(T) \setminus \{0\}$, for every $\varepsilon > 0$ the set $\{n \in \mathbb{N}^> : |\lambda_n| \ge \varepsilon\}$ is finite; if n_ε is the maximum of this set, for $n > n_\varepsilon$ we have $|\lambda_n| < \varepsilon$. □

2.8.3. *Spectrum of a self–adjoint operator in a Hilbert space.*

. *Let X be a Hilbert space, and assume that $T \in L(X)$ is self–adjoint.Then the eigenvalues of T are real; and eigenspaces associated to distinct eigenvalues are orthogonal.*

Proof. Let λ be an eigenvalue of T, and $a \neq 0$ an associated eigenvector; from $(Ta \mid a) = (a \mid Ta)$ we get $(\lambda\, a \mid a) = (a \mid \lambda\, a)$, that is $\lambda\, |a|^2 = \bar\lambda\, |a|^2$, i.e. $\lambda = \bar\lambda$. And if λ, μ are distinct eigenvalues with associated eigenvectors a, b, from $(Ta \mid b) = (a \mid Tb)$ we get $\lambda\,(a \mid b) = \mu\,(a \mid b)$, which implies $(a \mid b) = 0$ since $\lambda \neq \mu$. □

One can also prove that if T is self–adjoint then $\sigma(T) \subseteq \mathbb{R}$ (also for non–eigenvalues):

EXERCISE 2.8.3.1. Let X be a complex Hilbert space and let $T \in L(X)$ be a bounded operator.

(i) Prove that $\lambda \in \sigma(T)$ if and only if $\bar\lambda \in \sigma(T^*)$.

From now on $T = T^*$ is assumed.

(ii) Prove that if λ is an eigenvalue of T then $\lambda \in \mathbb{R}$ (proved above, repeated here for emphasis).
(iii) Let $\lambda = a + i\,b$ with $b \neq 0$. Prove that $|(\lambda - T)(x)| \geq |b|\,|x|$ for every $x \in X$; deduce that $\lambda - T$ is a homeomorphism onto a closed subspace Y of X (compute

$$|(\lambda - T)(x)|^2 = ((a - T)(x) + ib\,x \mid (a - T)(x) + ib\,x) = |(a - T)(x)|^2 + \dots$$

and prove that it is larger than $b^2\,|x|^2$).
(iv) Prove that $Y = X$ (the orthogonal of Y is an eigenspace for ...).

2.8.4. *Self–operators of a Hilbert space, sesquilinear forms and quadratic forms.* If $T : X \to X$ is a (continuous) linear operator, the formula

$$B_T(x, y) = B(x, y) := (Tx \mid y) \quad \text{defines a (continuous) sesquilinear form:} \quad B : X \times X \to \mathbb{K}.$$

It is easy to see that B is hermitian (meaning that $\overline{B(x, y)} = B(y,\, x)$) if and only if T is self–adjoint: in fact

$$B(y,\, x) = (Ty \mid x) = (y \mid T^*x) = \overline{(T^*x \mid y)} = \overline{B_{T^*}(x,\, y)},$$

for this reason, self–adjoint operators are often also called hermitian.

To a (continuous) *hermitian* form $B : X \times X \to \mathbb{K}$ it is associated a (continuous) *real quadratic form* $Q = Q_B : X \times X \to \mathbb{R}$, given by $Q(x) = B(x,\, x)$, the restriction to the diagonal of the given hermitian form. Quadratic forms are absolutely homogeneous of degree 2: for $\alpha \in \mathbb{K}$ and $x \in X$ we have $Q(\alpha\, x) = |\alpha|^2\, Q(x)$, as is easy to check. For this reason they are completely determined by their values on the unit sphere of X, $S = \{u \in X : |u| = 1\}$. The hermitian form can be reconstructed from the quadratic form. In fact

$$Q(x \pm y) = B(x \pm y, x \pm y) = B(x, x) \pm B(x, y) \pm B(y, x) + B(y, y);$$

subtracting we get

$$Q(x + y) - Q(x - y) = 2(B(x, y) + B(y, x)) = 4\ \mathrm{Re}\, B(x, y),$$

that is

$$(*) \qquad \mathrm{Re}\, B(x, y) = \frac{1}{4}(Q(x + y) - Q(x - y)).$$

(of course we also have $\mathrm{Im}\, B(x, y) = \mathrm{Re}\, B(x, iy) = (Q(x + iy) - Q(x - iy))/4$, so that we have the *polar identity*:

$$B(x, y) = \frac{1}{4}((Q(x + y) - Q(x - y)) + i(Q(x + iy) - Q(x - iy))),$$

but we have no use for this fact). What we are after is the following:

. *Let X be a Hilbert space, let $T : X \to X$ be bounded and self–adjoint, and let $Q(x) = Q_T(x) := (Tx \mid x)$ be the quadratic form associated to T. Then*

$$\|T\| = \sup\{|Q(u)| : |u| = 1\}.$$

Proof. We have, for $|u| = 1$:

$$|Q(u)| = |(Tu \mid u)| \le |Tu|\,|u| \le \|T\|,$$

so that the sup-norm b of Q is not larger than $\|T\|$. And from the polar identity (*) we get, setting $u = \operatorname{sgn}(x+y) = (x+y)/|x+y|$ and $v = \operatorname{sgn}(x-y) = (x-y)/|x-y|$ (the sign is zero if either vector is 0):

$$\begin{aligned}\operatorname{Re}(Tx \mid y) &\le \frac{1}{4}(|Q(x+y)| + |Q(x-y)|) = \frac{1}{4}(|Q(u)|\,|x+y|^2 + |Q(v)|\,|x-y|^2)\\ &\le \frac{b}{4}(|x+y|^2 + |x-y|^2) = \frac{b}{2}(|x|^2 + |y|^2),\end{aligned}$$

by the parallelogram identity. If $|x| = 1$ and $y = Tx/|Tx|$ we get

$$|Tx| \le \frac{b}{2}(1+1) = b,$$

for every x in the unit sphere, so that $\|T\| \le b$. □

2.8.5. *Spectrum of a self–adjoint operator and associated quadratic form.*

Proposition. *Let X be a Hilbert space and let $T : X \to X$ be linear, bounded and self–adjoint. Let $a = \inf\{(Tu \mid u) : |u| = 1\}$, $b = \sup\{(Tu \mid u) : |u| = 1\}$. Then a and b are finite, $\|T\| = \max\{|a|, |b|\}$, $\sigma(T) \subseteq [a, b]$ and $a, b \in \sigma(T)$.*

Proof. That a and b are finite and $\|T\| = \max\{|a|, |b|\}$ is what just proved in the preceding number. Let us prove that if $\lambda > b$ then $\lambda \in \rho(T)$. Consider the sesquilinear form $B : X \times X \to \mathbb{K}$ given by $B(x, y) = ((\lambda - T)(x) \mid y)$; it is clearly continuous, and moreover

$$B(x, x) = (\lambda x - Tx \mid x) = \lambda\,|x|^2 - (Tx \mid x) \ge \lambda\,|x|^2 - b\,|x|^2 = (\lambda - b)\,|x|^2,$$

so that B is coercive. By the Lax–Milgram lemma (see 1.10.5) we have that $B(x, y) = (x \mid Sy) = (S^*x \mid y)$ for some linear self–homeomorphism of X; but then $((\lambda - T)(x) \mid y) = (S^*x \mid y)$ for $x, y \in X$ if and only if $S^* = \lambda - T$, equivalently $S = \lambda - T$. Then $\lambda - T$ is bijective, equivalently $\lambda \in \rho(T)$. Consider now $B(x, y) = ((b - T)(x) \mid y)$. It is a hermitian form, because $b - T$ is self–adjoint, and it is positive:

$$B(x, x) = ((b - T)(x) \mid x) = b\,|x|^2 - (Tx \mid x) \ge b\,|x|^2 - b\,|x|^2 = 0.$$

Then Cauchy–Schwarz inequality implies that

$$|B(x, y)| \le \sqrt{B(x, x)}\,\sqrt{B(y, y)} = \sqrt{b\,|x|^2 - (Tx \mid x)}\,\sqrt{b\,|y|^2 - (Ty \mid y)};$$

If $|y| = 1$ we have $b\,|y|^2 - (Ty \mid y) \le |b| + \max\{|a|, |b|\} = C^2$, so that, for every y in the unit sphere:

$$|B(x, y)| = |(b\,x - Tx|y)| \le C\,\sqrt{b\,|x|^2 - (Tx \mid x)},$$

which implies, assuming also $|x| = 1$:

$$|b\,x - Tx| \le C\,\sqrt{b - (Tx \mid x)}.$$

The preceding inequality implies that if u_n is a sequence in the unit sphere such that $(Tu_n \mid u_n) \to b$, then $b\,u_n - Tu_n \to 0$ in X; and if $b \in \rho(T)$, then $u_n = (b - T)^{-1}(b\,u_n - Tu_n)$ must tend to 0, impossible because u_n has constant norm 1. Then $b \in \sigma(T)$. For a we argue with $-T$ in place of T. □

It is now clear the:

Corollary. *If T is self–adjoint we have $\|T\| = \max\{|\lambda| : \lambda \in \sigma(T)\}$. Consequently, a self–adjoint operator is zero if and only if its spectrum is $\{0\}$.*

2.8.6. *Spectral decomposition of a compact self–adjoint operator.* Let X be a Hilbert space and let $T : X \to X$ be compact and self–adjoint. Let $\{\lambda_k : k \in N\}$ be a bijective indexing of the non–zero eigenvalues of T, where $N = \{1, \dots, m\}$ if N is finite, otherwise $N = \mathbb{N}^{>}$, and let $\lambda_0 = 0$. For every $k \in N \cup \{0\}$ let P_k be the orthogonal projection onto the eigenspace $E_k = \operatorname{Ker}(\lambda_k - T)$. Recall that if $k \ne 0$ then E_k is finite dimensional, and that the E_k's are pairwise orthogonal. Then:

THEOREM. (HILBERT–SCHMIDT) *In the preceding terminology and notations, the space*

$$\bigoplus_{k\in N\cup\{0\}} E_k$$

direct sum of all eigenspaces of T, is dense in X, and for every $x \in X$ we have

$$x = P_0(x) + \sum_{k\in N} P_k(x); \qquad Tx = \sum_{k\in N} \lambda_k\, P_k(x).$$

Proof. The only thing left to prove is that the sum $E = \bigoplus_{k\in N\cup\{0\}} E_k$ is dense in X. If Y is the orthogonal in X of this sum we have that T induces a compact self–adjoint operator S of Y into itself (if $(x \mid v) = 0$ for every $v \in E$ then, since $Tv \in E$ if $v \in E$ we have $(Tx \mid v) = (x \mid Tv) = 0$, so that $Tx \in Y$). If λ is a non zero element of the spectrum of S, then it is an eigenvalue, but then it is also an eigenvalue of T, i.e $\lambda = \lambda_k$ for some $k \in N$, and so every vector in the eigenspace is orthogonal to itself, a contradiction. By corollary 2.8.5, $S = 0$, and so $Y \subseteq \operatorname{Ker} T = E_0$, hence $Y = \{0\}$. □

Another way of stating the theorem is:

. *If $T : X \to X$ is a compact self–adjoint operator in the Hilbert space X, then X has an orthonormal basis consisting of eigenvectors of T; and eigenspaces of nonzero eigenvalues are finite dimensional.*

EXAMPLE 2.8.6.1. Let $X = L^2([0,1])$, and let $K : [0,1]\times[0,1] \to \mathbb{C}$ be defined as $K(x,y) = (1-x)y$ for $0 \le y \le x$, $K(x,y) = (1-y)x$ for $x \le y \le 1$; define $T : X \to X$ by $Tf(x) = \int_0^1 K(x,y)\, f(y)\, dy$. Then T is compact (because $K \in L^2([0,1]\times[0,1])$) and self–adjoint (because $K(x,y) = \overline{K(y,x)}$). We want to find eigenvalues and eigenspaces. Write

$$Tf(x) = (1-x)\int_0^x y\, f(y)\, dy + x\int_x^1 (1-y)\, f(y)\, dy.$$

Since T is compact and X infinite dimensional, T cannot be surjective, hence $0 \in \sigma(T)$; on the other hand it is easy to observe that the image of T contains only continuous functions (but we do not yet know if 0 is an eigenvalue or not). Note that $Tf(0) = 0$ and that $Tf(1) = 0$, for every $f \in X$. Assume that $\lambda \neq 0$ is an eigenvalue, so that

$$\lambda\, f(x) = (1-x)\int_0^x y\, f(y)\, dy + x\int_x^1 (1-y)\, f(y)\, dy$$

for some nonzero $f \in X$. Since, as observed, the right–hand side is continuous, f must be continuous and hence also C^1; remember that we also have $f(0) = f(1) = 0$. Taking derivatives:

$$\lambda\, f'(x) = -\int_0^x y\, f(y)\, dy + (1-x)\, x\, f(x) + \int_x^1 (1-y)\, f(y)\, dy - x(1-x)\, f(x),$$

that is

$$\lambda f'(x) = \int_x^1 f(y)\, dy - \int_0^1 y\, f(y)\, dy \qquad f(0) = f(1) = 0;$$

this formula implies that $f' \in C^1([0,1])$; a further derivation brings to

$$\lambda\, f''(x) = -f(x) \qquad f(0) = f(1) = 0.$$

So we have to solve the second order differential equation $f''(x) = (-1/\lambda)\, f(x)$, with the boundary conditions $f(0) = f(1) = 0$. If $\omega^2 = 1/\lambda$ the solutions are

$$f(x) = c_1\cos(\omega x) + c_2\sin(\omega x);$$

the boundary conditions give

$$c_1 = 0, \quad c_1\, \cos\omega + c_2 \sin\omega = 0$$

Non–trivial solutions exist iff $\sin\omega = 0$, that is, iff $\omega = n\pi$ with $n \in \mathbb{Z}$; since $\lambda = 1/\omega^2$, this means $\lambda = 1/(n\pi)^2$; the corresponding eigenspace is $\mathbb{K}\sin(n\pi x)$. So: nonzero eigenvalues are $\lambda_n = 1/(n\pi)^2$, with corresponding eigenspace $E_n = \operatorname{Ker}(\lambda_n - T) = \mathbb{K}\,\sin(n\pi\#)$, $n \ge 1$.

It remains to see if 0 is an eigenvalue or not. This can be solved observing that the space generated by the orthogonal family $\{\sin(n\pi\#) : n \ge 1\}$ of eigenfunctions is actually all of $L^2([0,1])$: every $f \in L^2([0,1])$ can be expanded in a sine series, simply consider the odd extension to $[-1,1]$. Thus $K = \operatorname{Ker}(T) = \{0\}$, because K is contained in the orthogonal of the sum $\bigoplus_{n\ge 1} E_n$; 0 is not an eigenvalue.

But it is perhaps easier to argue as before: by Lebesgue differentiation theorem we have that Tf is differentiable and that, a.e. in $[0,1]$:

$$(Tf)'(x) = \int_x^1 f(y)\,dy - \int_0^1 y\,f(y)\,dy;$$

if $Tf = 0$, then $(Tf)' = 0$ and differentiating again we get $0 = -f(x)$, a.e. in $[0,1]$.

Exercise 2.8.6.2. Recall that if $(X, \mathcal{M}, \mu)$ is a semifinite measure space, and $g : X \to \mathbb{K}$ a measurable function, then multiplication by g defines a bounded operator $T_g(f) = g\,f$ of $L^p(\mu)$ into itself $(1 \le p \le \infty)$ if and only if $g \in L^\infty(\mu)$, and in this case $\|T_g\| = \|g\|_\infty$. We fix $p = 2$ and a $g \in L^\infty(\mu)$; so $T = T_g$ is the map of $L^2(\mu)$ into itself defined by $T(f) = g\,f$.

(i) Describe the adjoint T^* of T.

Given a measurable function $h \in L(X)$ its *essential range* $h_{\mathrm{ess}}(X)$ is the smallest closed subset R of $\mathbb{K}$ such that $h^{\leftarrow}(\mathbb{K} \setminus R)$ has measure 0. The essential range exists: consider a countable basis B_n for the open sets of $\mathbb{K}$, and take the union of all B_n whose inverse image has zero measure; its complement is the essential range of h: it is clear that $\lambda \in h_{\mathrm{ess}}(X)$ if and only if for every nbhd V of $\lambda \in \mathbb{K}$ we have $\mu(h^{\leftarrow}(V)) > 0$. For a $g \in L^\infty(\mu)$, the essential range is compact, and $\|g\|_\infty = \max\{|y|;\, y \in h_{\mathrm{ess}}(X)$.

(ii) Prove that the spectrum $\sigma(T)$ of $T = T_g$ is exactly $g_{\mathrm{ess}}(X)$, and that $\lambda \in g_{\mathrm{ess}}(X)$ is an eigenvalue if and only if $\mu(\{g = \lambda\}) > 0$.

(iii) Give an example of a self–adjoint operator $T = T_g$ on $L^2([0,1])$ for which the norm is an eigenvalue, and one for which the norm is not an eigenvalue.

Exercise 2.8.6.3. Let X be a normed space and let X^* be its dual.

(i) Prove that if f_n converges to f strongly in X^* and x_n converges to x weakly in X, then $f_n(x_n) \to f(x)$ in $\mathbb{K}$.

(ii) Let $T : X \to Y$ be compact (Y is another normed space). Prove that for every sequence x_n in X weakly converging to $x \in X$, the sequence Tx_n converges to Tx strongly in Y.

From now on X is a separable Hilbert space, and $T : X \to X$ is compact and self–adjoint.

(iii) Prove that $Q : X \to \mathbb{K}$ defined by $Q(x) = (Tx \mid x)$ is continuous if restricted to the unit ball B_X of X, in the weak topology.

(iv) Let $b = \sup Q(S)$, where S is the unit sphere of X. Assume that $b > 0$, and prove directly that $b = \max Q(S)$.

Solution. (i):

$$\begin{aligned}|f_n(x_n) - f(x)| &= |f_n(x_n) - f(x_n) + f(x_n) - f(x)| \le |f_n(x_n) - f(x_n)| + |f(x_n) - f(x)| \\ &\le \|f - f_n\|_{X^*}\,\|x\|_n + |f(x_n) - f(x)|;\end{aligned}$$

recall that a weakly convergent sequence is bounded, so that $\|x_n\| \le M$ for some $M > 0$; thus

$$|f_n(x_n) - f(x)| \le \|f - f_n\|_{X^*}\,M + |f(x_n) - f(x)|,$$

and as $n \to \infty$ both terms in the right hand–side tend to 0, by the assumptions made.

(ii) This is exercise 2.6.4.3

(iii) The first question is now clear: the unit ball B_X is metrizable and compact in the weak topology, and Q restricted to B_X is continuous in this weak topology: if $x_n \rightharpoonup x$, then $Tx_n \to Tx$ strongly, and by (i) we have that $Q(x_n) = (Tx_n \mid x_n) \to (Tx \mid x)$.

(iv) Pick a sequence $u_n \in S$ such that $\lim_{n\to\infty} Q(u_n) = b$; then u_n has a subsequence, still called u_n, which weakly converges to $v \in B_X$; then $Q(v) = b$, and this implies that actually $v \in S$: in fact $Q(v/|v|) = 1/|v|^2\,Q(v) \ge b$, with equality if and only if $|v| = 1$. Notice that in this case u_n converges to v also strongly, since $|u_n| = 1$ converges to $|v| = 1$ (2.3.4.3).

□

Exercise 2.8.6.4. Let X be a Hilbert space and let $T : X \to X$ be self–adjoint and bounded; call Q the quadratic map Q associated to T, $Q(x) = (Tx \mid x)$, and let $S = \{x \in X : |x| = 1\}$ be the unit sphere of X. In 2.8.5 we proved that $\inf Q(S)$ and $\sup Q(S)$ are both in the spectrum of T. Here we prove that if a or b are values assumed by Q, then they are eigenvalues.

. *Assume that a point $u \in S$ is of local maximum or minimum for Q. Prove that then u is an eigenvalue of T.*

(given $v \in u^\perp \cap S$, consider the map $g : [-\pi, \pi] \to \mathbb{R}$ given by $g(t) = Q(\cos t\, u + \sin t\, v)$; then the derivative $g'(t)$ vanishes for $t = 0 \dots$).

Solution. The map g is obtained by composition of the vector valued map $h(t) = \cos t\, u + \sin t\, v$ with the quadratic map Q. Clearly the Fréchet $\mathbb{R}-$differential of Q at an $x \in X$ (with Q considered as a map from X to $\mathbb{K}$), is the $\mathbb{R}-$linear map from $X \to \mathbb{K}$ given by:

$$Q(x)[\delta x] = (T(\delta x) \mid x) + (Tx \mid \delta x) \quad \text{whence} \quad g'(t) = (Th'(t) \mid h(t)) + (Th(t) \mid h'(t)),$$

that is:

$$g'(t) = (-\sin t\, Tu + \cos t\, Tv \mid \cos t\, u + \sin t\, v) + (\cos t\, Tu + \sin t\, Tv \mid -\sin t\, u + \cos t\, v) \quad \text{hence}$$
$$g'(0) = (Tv \mid u) + (Tu \mid v) = 2(Tu \mid v),$$

so that $(Tu \mid v) = 0$ for every $v \in u^\perp \cap S$, which implies $Tu \in \mathbb{K}\, u$, that is $Tu = \lambda\, u$ for some $\lambda \in \mathbb{K}$; but $Q(u) = (Tu \mid u) = (\lambda\, u \mid u) = \lambda\, |u|^2 = \lambda$: the "Lagrange multiplier" λ is exactly the value of Q at u. $\square$

2.8.7. *The Fredholm alternative.* The results proved in 2.8.1 and 2.8.2 on the spectrum of a compact operator are classically formulated, most often in the context of integral equations, in (more or less) the following way, called *Fredholm alternative*.

. Fredholm alternative *Let X be a Banach space, and let $T : X \to X$ be a linear compact operator. Given $y \in X$ and $\lambda \in \mathbb{K}$, we consider the (affine) equation (Fredholm equation of the second kind):*

$$y = x - \lambda\, Tx \quad \textit{in the unknown } x \in X.$$

Then, given $\lambda \neq 0$ in $\mathbb{K}$, either the equation has a unique solution, or else the associated homogeneous equation

$$0 = x - \lambda\, Tx$$

has a non–trivial solution. In this second case:

the solutions of the homogeneous equation are a finite dimensional subspace of X;

λ is called characteristic value *for T, and the characteristic values are a closed discrete subset of $\mathbb{K}$.*

Proof. Since $\lambda \neq 0$, we rewrite the equation as

$$\lambda^{-1}\, y = \lambda^{-1}\, x - Tx;$$

if $\lambda^{-1} \notin \sigma(T)$ then $\lambda^{-1} - T$ is invertible and

$$x = (\lambda^{-1} - T)^{-1}(y/\lambda) = (1 - \lambda\, T)^{-1} y$$

is the unique solution. Otherwise, $\lambda^{-1} \in \sigma(T)$, and being non–zero λ^{-1} is an eigenvalue, with a finite dimensional eigenspace $\mathrm{Ker}(\lambda^{-1} - T) = \mathrm{Ker}(1 - \lambda\, T)$; and since eigenvalues, if infinitely many, form a sequence converging to 0, characteristic values, their reciprocals, form a sequence diverging to infinity. $\square$

Exercise 2.8.7.1. Let $X = C([0, 1])$ with uniform norm. Prove that the operator $T : X \to X$ given by $Tf(x) = \int_0^1 x \wedge y\, f(y)\, dy$ is compact, and find eigenvalues and eigenfunctions.

Next, determine the set of all $\lambda \in \mathbb{R}$ for which there exists at most one $\varphi \in C([0, 1])$ such that, for $x \in [0, 1]$:

$$\varphi(x) = \lambda - \lambda^2 \int_0^1 y\, \varphi(y)\, dy - \lambda^2 \int_x^1 (x - y)\, \varphi(y)\, dy.$$

Solution. Clearly $K(x, y) = \min\{x, y\}$ is continuous, so that T is compact (1.6.2.4). Write:

$$Tf(x) = \int_0^x y\, f(y)\, dy + x \int_x^1 f(y)\, dy,$$

and note that $Tf(0) = 0$; moreover $T(X) \subseteq C^1([0, 1])$, so that eigenfunctions associated to nonzero eigenvalues are necessarily C^1:

$$\lambda\, f(x) = \int_0^x y\, f(y)\, dy + x \int_x^1 f(y)\, dy;$$

differentiating we get

$$\lambda\, f'(x) = x\, f(x) + \int_x^1 f(y)\, dy - x\, f(x) = \int_x^1 f(y)\, dy;$$

note that $f'(1) = 0$. Differentiating again:

$$\lambda f''(x) = -f(x) \iff f''(x) + (1/\lambda) f(x) = 0;$$

if $1/\lambda = \omega^2$ the solutions are

$$f(x) = c_1 \cos(\omega x) + c_2 \sin(\omega x);$$

the boundary condition $f(0) = 0$ gives $c_1 = 0$, so we are reduced to $f(x) = c \sin(\omega x)$; the other condition $f'(1) = 0$ gives $c\omega \cos(\omega) = 0$ so that $\omega = \omega_n = (\pi/2 + n\pi) = \pi(n + 1/2)$, with $n \in \mathbb{Z}$; this implies $\lambda_n = 1/(\omega_n)^2 = \pi^{-2}(n + 1/2)^{-2}$. Thus: eigenvalues $\lambda_n = \pi^{-2}(n + 1/2)^{-2}$ with eigenspace $\mathbb{K} \sin(\pi(n + 1/2)\#)$.

For the kernel we argue as before: if $Tf = 0$ we get

$$0 = (Tf)'(x) = \int_x^1 f(y)\, dy,$$

which implies $f = 0$. So: 0 is not an eigenvalue for T (but of course $0 \in \sigma(T)$).

Second question: notice first that if $\lambda = 0$ there is exactly one solution, the 0 function. Assume then $\lambda \neq 0$. It is easy to check that the right hand side is

$$\lambda - \lambda^2 \int_0^1 x \wedge y\, \varphi(y)\, dy = \lambda - \lambda^2\, T\varphi(x),$$

so that an equivalent form of the equation is

$$\lambda = \varphi(x) + \lambda^2\, T\varphi(x) \quad \text{or else} \quad \lambda = \varphi(x) - (-\lambda^2)\, T\varphi(x),$$

By the Fredholm alternative, if $-\lambda^2$ is not a characteristic value of T, equivalently if $-1/\lambda^2$ is not an eigenvalue of T there is exactly one solution; but eigenvalues of T are all positive. Consequently, for every real λ there is exactly one solution.

□

EXERCISE 2.8.7.2. Let E be a locally compact Hausdorff space, and let $X = C_0(E)$ with uniform norm. Prove that a sequence $f_n \in X$ is weakly convergent to $f \in X$ if and only if f_n is uniformly bounded in X, and f_n converges to f pointwise on E, i.e. we have $\lim_{n\to\infty} f_n(x) = f(x)$ for every $x \in E$.

Solution. Any weakly convergent sequence is norm bounded. And for every $c \in E$ the evaluation functional $\delta_c : X \to \mathbb{K}$ given by $\delta_c(f) = f(c)$ is in X^* (δ_c, the Dirac's measure, unit mass concentrated at c, is plainly a Radon measure), so also pointwise convergence, $f_n(c) \to f(c)$ for every $c \in E$, is implied by weak convergence. For the converse: assume that $\|f_n\|_u \le M$ for every n, and pick a finite Radon measure μ on E, $d\mu = \sigma\, d|\mu|$, with σ measurable and $|\sigma(x)| = 1$ for every $x \in E$. Then the sequence $f_n\,\sigma$ converges pointwise to $f\,\sigma$ and is dominated by the constant M, which is in $L^1(|\mu|)$ because $|\mu|(E) < \infty$. By Lebesgue dominated convergence theorem we have that

$$\lim_{n\to\infty} \int_E f_n\, \sigma\, d|\mu| = \int_E f\, \sigma\, d|\mu|.$$

□

EXERCISE 2.8.7.3. We denote by $M(\mathbb{R})$ the space of all finite Radon measures on $\mathbb{R}$. Let $\mu_n \in M(\mathbb{R})$ be the measure

$$d\mu_n = \frac{1}{\pi} \frac{n}{1 + (nx)^2}\, dx \quad n = 1, 2, 3, \ldots.$$

(i) Compute $\mu_n(\mathbb{R})$. Prove that for every $\delta > 0$ and every bounded Borel measurable function $f : \mathbb{R} \to \mathbb{R}$ we have

$$\lim_{n\to\infty} \int_{\mathbb{R}\smallsetminus[-\delta,\delta]} f\, d\mu_n = 0.$$

(ii) Prove that for every $f \in C_0(\mathbb{R})$ we have

$$\lim_{n\to\infty} \int_{\mathbb{R}} f\, d\mu_n = f(0)$$

(split the integral in an integral over $[-\delta, \delta]$ and the complement $\mathbb{R} \smallsetminus [-\delta, \delta]$ and use continuity at 0 of f ...).

The weak* topology on $M(\mathbb{R})$ is often called *vague* topology.

(iii) To which measure μ does μ_n converge in the vague topology? Compute the limits:

$$\lim_{n\to\infty} \mu_n([-1,1]) \qquad \lim_{n\to\infty} \mu_n([0,1]).$$

Do these limits coincide with $\mu([-1,1])$ and $\mu([0,1])$?

(iv) The space $L^1(\mathbb{R})$ is isometrically embedded in $M(\mathbb{R})$: any $f \in L^1(\mathbb{R})$ is considered as the measure $f\,dx$. On it we may consider the topology induced by the vague topology of $M(\mathbb{R})$ and the weak topology of $L^1(\mathbb{R})$. Which one is stronger?

(the computations are all easy. Of course the weak topology on $L^1(\mathbb{R})$ is stronger than the vague topology: the first is the topology $\sigma(L^1(\mathbb{R}), L^\infty(\mathbb{R}))$, the second is $\sigma(L^1(\mathbb{R}), C_0(\mathbb{R}))$ and $L^\infty(\mathbb{R}) \supsetneq C_0(\mathbb{R})$)

Exercise 2.8.7.4. Prove or disprove: in a locally convex space X if C is a closed convex subset and $a \in X \smallsetminus C$ there is a continuous real functional φ such that $\sup\varphi(C) < \varphi(a)$.

Given a measure space $(X, \mathcal{M}, \mu)$ and an integer $n \geq 0$, we define $L^1_\mu(X, \mathbb{K}^n)$ as the set of measurable functions $f : X \to \mathbb{K}^n$ whose components $\pi_k \circ f$ are in $L^1(\mu)$ for every $k = 1, \dots, n$, and the integral is of course

$$\int_X f\,d\mu = \left(\int_X f_1\,d\mu, \dots, \int_X f_n\,d\mu\right).$$

With this definition, it is easy to prove (accept it) that for every linear map $T : \mathbb{K}^n \to \mathbb{K}^m$ we have that $T \circ f \in L^1_\mu(X, \mathbb{K}^m)$ if $f \in L^1_\mu(X, \mathbb{K}^n)$, and moreover:

$$\int_X T \circ f\,d\mu = T\left(\int_X f\,d\mu\right).$$

Suppose now that $E \subseteq X$ has finite non–zero measure, that $C \subseteq \mathbb{K}^n$ is closed convex, and that $f \in L^1_\mu(X, \mathbb{K}^n)$ is such that $f(E) \subseteq C$. Prove that then the average of f over E,

$$A_E(f) = \frac{1}{\mu(E)} \int_E f\,d\mu$$

belongs to C. Recall that if C is a convex subset of a real linear space Y, a function $\varphi : C \to \mathbb{R}$ (sometimes the range is assumed to be $]-\infty, +\infty]$) is said to be *convex* when its epigraphic

$$\operatorname{epi}(\varphi) := \{(x,t) \in Y \times \mathbb{R} : t \geq \varphi(x)\}$$

is a convex subset of $Y \times \mathbb{R}$. Prove *Jensen's inequality*:

. *If C is a compact convex subset of $\mathbb{K}^n$, $\varphi : C \to \mathbb{R}$ is convex and continuous, $(X, \mathcal{M}, \mu)$ is a probability space, and $f : X \to C$ belongs to $L^1_\mu(X, \mathbb{K}^n)$, then $\varphi \circ f$ is in $L^1_\mu(X, \mathbb{R})$ and*

$$\varphi\left(\int_X f\,d\mu\right) \leq \int_X \varphi \circ f\,d\mu.$$

(consider the map $g : x \mapsto (f(x), \varphi(f(x)))$ from X to $\operatorname{epi}(\varphi) \subseteq \mathbb{K}^n \times \mathbb{K}$; continuity of φ says that $\operatorname{epi}(\varphi)$ is closed in $\mathbb{K}^n \times \mathbb{R} \dots$; compactness of C is not really needed, but it speeds up the proof).

Solution. The answer is yes (2.2.8). Arguing by contradiction, assume next that $a = A_E(f)$ is not in C. Then we find a functional $\varphi : \mathbb{K}^n \to \mathbb{R}$ such that $\varphi(y) \leq \alpha < \varphi(a)$ for every $y \in C$; thus

$$\varphi(f(x)) \leq \alpha < \varphi(a) \quad \text{for every } x \in E\text{; hence, integrating over } E$$

$$\int_E \varphi(f(x))\,d\mu(x) \leq \alpha\,\mu(E) < \varphi(a)\,\mu(E),$$

so that

$$\frac{1}{\mu(E)} \int_E \varphi(f(x))\,d\mu(x) \leq \alpha < \varphi(a);$$

but by linearity of φ we have

$$\frac{1}{\mu(E)} \int_E \varphi(f(x))\,d\mu(x) = \varphi\left(\frac{1}{\mu(E)} \int_E f\,d\mu\right) = \varphi(a),$$

so $\varphi(a) < \varphi(a)$, a contradiction.

For the last question: clearly g is in $L^1(X, \mathbb{K}^n \times \mathbb{K})$, since it is bounded on a set of finite measure; as said, continuity of φ says that $\operatorname{epi}(\varphi)$ is closed in $\mathbb{K}^n \times \mathbb{R}$, which is closed in $\mathbb{K}^n \times \mathbb{K}$; and by what just proved we have, since $\mu(X) = 1$:

$$\int_X g\,d\mu \in \operatorname{epi}(\varphi), \quad \text{equivalently} \quad \left(\int_X f\,d\mu, \int_X \varphi \circ f\ d\mu\right) \in \operatorname{epi}(\varphi),$$

in turn equivalent to

$$\varphi\left(\int_X f\,d\mu\right) \le \int_X \varphi\circ f\,d\mu.$$

□

EXERCISE 2.8.7.5. Consider the real Hilbert space $L^2 = L^2([-1,1])$. Let :

(i) $V = \{f \in L^2 : \int_{-1}^1 f(x)\,x^n\,dx = 0 \quad \text{for} \quad n = 0,1,2,3,\dots\}$.
(ii) $W = \{f \in L^2 : \int_{-1}^1 f(x)\,x^{2n}\,dx = 0 \quad \text{for} \quad n = 0,1,2,3,\dots\}$.

Can you say wether V, $W = \{0\}$ or not? if not, can you describe V, W?

Solution. Clearly V is the subspace of $L^2 = L^2([-1,1])$ orthogonal to the space of all polynomials, and W is the subspace of L^2 orthogonal to all polynomials in x^2. By the Stone–Weierstrass theorem polynomials are uniformly dense in $C([-1,1])$, hence their L^2 closure also contains all continuous functions (uniform convergence in $[-1,1]$ implies L^2 convergence), and since continuous functions are dense in L^2 this orthogonal is 0, $V = \{0\}$.

It is quite clear also that only *even* continuous functions on $[-1,1]$ can be uniformly approximated by polynomials in x^2, and all even continuous functions are so approximable: the constancy sets of the algebra $\mathbb{R}[x^2]$ are exactly the pairs $\{-t,t\}$, with $t \in]0,1]$, and the singleton $\{0\}$. The space D of *odd* functions in $L^2([-1,1]$ and the space P of *even* functions are clearly a pair of mutually orthogonal subspaces of L^2: if $u \in P$ and $v \in D$ then $u\,v$ is odd, hence $\int_{-1}^1 u\,v\,dx = 0$; and given $f \in L^2$ we set $u = (f+\tilde f)/2$, $v = (f-\tilde f)/2$, where $\tilde f(x) = f(-x)$, so that $L^2 = P + D$. It is clear that even continuous functions are dense in the space P of even functions of L^2 (if $u \in L^2$ is even, and $\varepsilon > 0$, pick $a \in C([0,1])$ such that $\int_0^1 |u-a|^2 < \varepsilon$; if b is the even extension of a to $[0,1]$ then $\int_{-1}^1 |u-b|^2 < 2\varepsilon$). So W is the space D of odd functions in L^2. □

EXERCISE 2.8.7.6. Let $L^1 = L^1([0,1])$ and $Y = C([0,1])$, the second with uniform norm, and consider the operator $T : L^1 \to Y$ given by $Tf(x) = \int_0^x f(t)\,dt$.

(i) Prove that T is bounded but not compact.
(ii) Let $\alpha \in L^1$ be positive, let $E_\alpha = \{f \in L^1 : |f(t)| \le \alpha(t) \text{ for every } t \in [0,1]\}$, and $K_\alpha = T(E_\alpha)$. Prove that K_α is totally bounded in Y (hint: remember absolute continuity ...).

Solution. (i) Trivially T is bounded of norm 1. But T is not compact: if $u_n = n\,\chi_{[0,1/n]}$ then $\|u_n\|_1 = 1$, but $Tu_n(x) = \min\{n\,x, 1\}$ has no converging subsequence: it is easy to see that Tu_n converges pointwise to 0, but has uniform norm constantly 1, so that no subsequence is uniformly convergent (see figure)

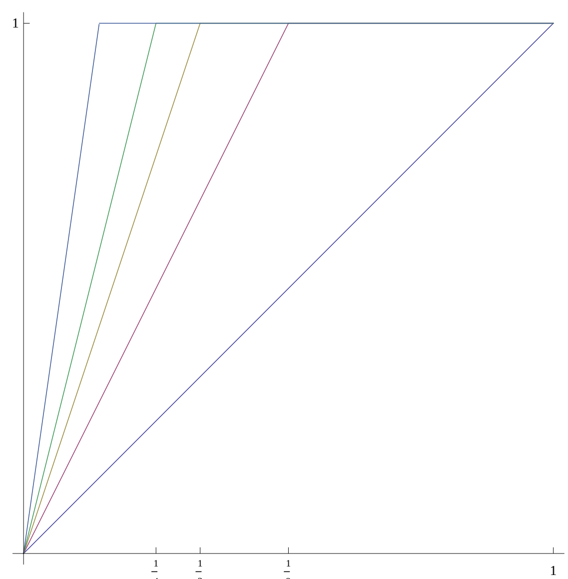

FIGURE 2. Some functions Tu_n.

(ii) Since $\alpha \in L^1([0,1])$ the indefinite integral $A \mapsto \int_A \alpha(x)\,dx$ is absolutely continuous with respect to Lebesgue measure; this means that for every $\varepsilon > 0$ there is $\delta = \delta_\varepsilon > 0$ such that $\left|\int_A \alpha(x)\,dx\right| \le \varepsilon$ if $m(A) \le \delta$, where m is Lebesgue measure. But then K_α is equicontinuous: if $u = Tf$, with $|u| \le \alpha$, and $x, y \in [0,1]$, with $x < y$ we get

$$|u(y) - u(x)| = \left|\int_x^y \alpha(t)\,dt\right| = \left|\int_{[x,y]} \alpha(t)\,dt\right| \le \varepsilon \quad \text{if} \quad |x-y| \le \delta,$$

exactly equiuniformcontinuity for K_α. Uniform boundedness of K_α is trivial:

$$|u| \le \alpha \quad \text{implies} \quad \left|\int_0^x u(t)\,dt\right| \le \left|\int_0^x |u(t)|\,dt\right| \le \left|\int_0^x \alpha(t)\,dt\right| \le \int_0^1 \alpha(t)\,dt = \|\alpha\|_1,$$

so that $\|u\|_u \le \|\alpha\|_1$, for every $u \in K_\alpha$. Thus K_α is totally bounded in $C([0,1])$. □

Exercise 2.8.7.7. There is a finer classification of the spectrum of a bounded operator $T \in L(X)$ of a Banach space X, divided into three disjoint pieces:

Point spectrum $\sigma_p(T)$: the eigenvalues of T, those $\lambda \in \mathbb{K}$ for which $\mathrm{Ker}(\lambda - T) \ne \{0\}$, i.e. $\lambda - T$ not injective.
Continuous spectrum $\sigma_c(T)$: those $\lambda \in \mathbb{K}$ for which $\lambda - T$ is injective, and the image $(\lambda - T)(X)$ is a proper dense subspace of X.
Residual spectrum: all other $\lambda \in \sigma(T)$, i.e. $\lambda - T$ is injective, and the image $(\lambda - T)(X)$ is not dense in X.

The right and left shift S, T in $\ell^2 = \ell^2(\mathbb{N}, \mathbb{C})$ have been described in 1.10.4.2, where one was asked to prove that $S^* = T$. One can prove that:

$\sigma_p(T) = \sigma_r(S) = \Delta := \{\zeta \in \mathbb{C} : |\zeta| < 1\}$.
$\sigma_c(T) = \sigma_c(S) = \mathbb{U} := \{\zeta \in \mathbb{C} : |\zeta| = 1\}$.
$\sigma_r(T) = \sigma_p(S) = \emptyset$.

2.8.8. *Assorted exercises.* Recall that the function $u : \mathbb{R} \to \mathbb{R}$ defined by $u(t) = 0$ if $t \le 0$, $u(t) = e^{-1/t}$ for $t > 0$ belongs to $C^\infty(\mathbb{R}, \mathbb{R})$ (see e,g, Analisi Uno, 16.9.4). Then the function $\rho : \mathbb{R}^n \to \mathbb{R}$ defined by $\rho(x) = u(1 - |x|^2)$ is in $C^\infty(\mathbb{R}^n)$, being the composition of u with the polynomial function $x \mapsto 1 - \sum_{k=1}^n x_k^2$; notice that $\rho(x) > 0$ if $|x| < 1$, and $\rho(x) = 0$ if $|x| \ge 1$, so that $\mathrm{Supp}(\rho) = B$, the closed unit ball of $\mathbb{R}^n$. Given $c \in \mathbb{R}^n$ and $r > 0$ set $\rho_{cr}(x) = \rho((x-c)/r)$; then $\rho_{cr}(x) > 0$ if $x \in B(c, r[$, and $\mathrm{Supp}(\rho_{cr}) = B(c, r]$. It is a common usage to call *smooth* the functions in C^∞.

Exercise 2.8.8.1. The set of functions $\{\rho_{cr} : c \in \mathbb{Q}^n, r \in \mathbb{Q}^>\}$ is a countable subset of $C_c^\infty(\mathbb{R}^n, \mathbb{K})$, which generates a subalgebra $A_c^\infty(\mathbb{R}^n, \mathbb{K})$. Prove that this subalgebra is dense in $C(\mathbb{R}^n, \mathbb{K})$ in the compact open topology and in $C_0(\mathbb{R}^n, \mathbb{K})$ in the uniform topology.

Given $a \in C_c(\mathbb{R}^n)$ the set A of all $\varphi \in C_c^\infty(\mathbb{R}^n, \mathbb{K})$ such that $\mathrm{Supp}(\varphi) \subseteq \mathrm{Supp}(a)$ is a subalgebra. Prove that given a pair of points x, $y \in \mathbb{R}^n$ there is $\varphi \in A$ such that $\varphi(x) = a(x)$ and $\varphi(y) = a(y)$, and conclude that there is a sequence φ_n of smooth functions converging uniformly to a and whose supports are all contained in the support of a.

Solution. Clearly all subalgebras are conjugation closed, and it is then not restrictive to assume $\mathbb{K} = \mathbb{R}$. Let us prove that A_c^∞ separates points in $\mathbb{R}^n$: given distinct points x, $y \in \mathbb{R}^n$ we find a ball $B(c, r[$ which contains one point, say x, but not the other y, so that $\rho_{cr}(x) > 0$ but $\rho_{cr}(y) = 0$. Moreover the functions ρ_{cr} have no common zero in $\mathbb{R}^n$. Then A_c^∞ is dense in $C(\mathbb{R}^n)$ in the compact–open topology. For the uniform closure it is better to consider all functions as continuous functions on the one–point compactification $\alpha\mathbb{R}^n = \mathbb{R}^n \cup \{\infty_n\}$ of $\mathbb{R}^n$: still A_c^∞ separates points, but now the extended functions have the point at infinity as their common zero; so the closure is the set of all functions on $\alpha\mathbb{R}^n$ which are 0 at infinity, exactly $C_0(\mathbb{R}^n)$ when restricted to $\mathbb{R}^n$.

That A is subalgebra is clear. Given then $x, y \in \mathbb{R}^n$, with $x \ne y$, if $a(x) = a(y) = 0$ the zero function will do. If $a(x) \ne 0$ but $a(y) = 0$ pick $r > 0$ such that $B(x, r] \subseteq \mathrm{Coz}(a)$; then $\varphi(\#) = a(x)(/\rho_{xr}(x))\,\rho_{xr}(\#)$ belongs to a and $\varphi(x) = a(x)$. And if $x \ne y$ and $a(x)$, $a(y)$ are both nonzero simply pick $r > 0$ so small that $B(x, r] \cap B(y, r] = \emptyset$ and $B(x, r] \cup B(y, r] \subseteq \mathrm{Coz}(a)$; then

$$\varphi(\#) = \frac{a(x)}{\rho_{xr}(x)}\,\rho_{xr}(\#) + \frac{a(y)}{\rho_{yr}(y)}\,\rho_{yr}(\#),$$

is as required. By the lattice theorem, the closure of A in the compact–open topology, being a sublattice of $C(X,\mathbb{R})$, contains A. But all functions involved are zero outside the support of a, so that uniform convergence on this support means uniform convergence on all of X. □

Exercise 2.8.8.2. Let X be a locally compact space and let $M(X)$ be the space of all Radon measures on X, dual of the space $C_0(X)$ of continuous functions which are zero at infinity. We want to prove that, excluding trivial cases, $C_0(X)$ is never reflexive.

(i) For every $\mu \in M(X)$ define $\Phi(\mu) = \sum_{x\in X} \mu(\{x\})$, where the sum is in the sense of summable families (0.1.1); prove that it is well-defined, and that $|\Phi(\mu)| \leq \|\mu\|$, so that $\Phi \in (M(X))^*$, and $\|\Phi\| = 1$

(ii) Let $\mu \in M(X)$ be positive. Prove that for every bounded Borel measurable function $f : X \to \mathbb{K}$ the measure ν defined by $d\nu = f\,d\mu$ is also a Radon measure on X.

(iii) Consider now $X = \mathbb{R}$: then Φ is not in the image of the canonical embedding $J : C_0(\mathbb{R}) \to (M(\mathbb{R}))^*$, hence $C_0(\mathbb{R})$ is not reflexive (argue by contradiction, using a measure like $d\mu = dx/(1+x^2)$).

Solution. (i) Given $\mu \in M(X)$, we prove that $x \mapsto \mu(\{x\})$ is a function in $\ell^1(X)$. In fact, if F is any finite subset of X:

$$\left|\sum_{x\in F} \mu(\{x\})\right| \leq \sum_{x\in F} |\mu(x)| = \sum_{x\in F} |\mu|(\{x\}) \leq |\mu|(X) = \|\mu\|,$$

which proves that $\Phi(\mu) = \sum_{x\in x} \mu(\{x\})$ exists, and moreover

$$|\Phi(\mu)| = \left|\sum_{x\in x} \mu(\{x\})\right| \leq \|\mu\|;$$

and since for every Dirac's measure we have $\Phi(\delta_x) = 1$ we actually have $\|\Phi\|_{(M(X))^*} = 1$.

(ii) We have to prove that the total variation $d|\nu| = |f|\,d\mu$ is a Radon measure. Given $E \subseteq X$ Borel measurable and $\varepsilon > 0$ we find an open U containing E and a compact K contained in E such that $\mu(U \setminus K) \leq \varepsilon$. Then

$$|\nu|(U \setminus K) = \int_{U\setminus K} |f|\,d\mu \leq \int_{U\setminus K} \|f\|_\infty\,d\mu = \|f\|_\infty \mu(U \setminus K) \leq \|f\|_\infty\,\varepsilon.$$

(iii) If $\Phi = J(\varphi)$, with $\varphi \in C_0(\mathbb{R})$ we have $\Phi(\nu) = \int_{\mathbb{R}} \varphi\,d\nu$, for every $\nu \in M(\mathbb{R})$. If $d\mu = dx/(1+x^2)$, then μ is a non-zero measure on X for which every singleton has zero measure; and $d\nu = \bar\varphi\,d\mu$ is also a Radon measure on $\mathbb{R}$, for which every singleton has measure zero, so that $\Phi(\nu) = 0$; but if $\Phi = J(\varphi)$ we have

$$\Phi(\nu) = \int_{\mathbb{R}} \varphi\,d\nu = \int_X \varphi\,\bar\varphi\,d\mu = \int_{\mathbb{R}} \frac{|\varphi(x)|^2}{1+x^2}\,dx = 0,$$

which implies $\varphi(x) = 0$ for every $x \in \mathbb{R}$. But then $\Phi(\delta) = 1 = \varphi(0) = 0$, a contradiction. □

Contents