V. Averbuch

FUNCTIONAL ANALYSIS

The 2nd version

1995

Introduction

Functional analysis is simply infinite-dimensional analysis.

The main objects of study in analysis are objects, equipped by compatible algebraical and topological structures. So the main objects of study in functional analysis are infinite-dimensional objects, equipped by compatible algebraical and topological structures: topological vector spaces, topological groups, topological algebras etc.

The name "functional analysis" originated from the word "functional". At first this term was used for scalar functions, which had as its argument not scalars or vectors but *functions*. Now one means by a functional *any* scalar function, that is a real- (or complex-) valued function defined on any (in general case, infinite-dimensional) vector space. Spaces of functions (or *function spaces*) are basic and typical examples of infinite-dimensional spaces.

Functional analysis (FA) marks a fundamental change in the point of view in mathematics: one goes from study of individual functions and individual relations, connecting them, to study of *sets* of such objects, viz. to study of function spaces and function transformations. So one consider, say, differentiation and integration not as operationes, applied to individual functions, but as operators, applied to a whole class of functions.

This change is comparable with one that occured, when it came into mathematics the notion of a variable. At that time one went from points to functions, now (dialectics!) from functions back to points: it is *functions* that are now considered as points! One goes from algebra through analysis to geometry.

* * *

The creation of FA was prepared by developping of "concrete" disciplines, viz. of the calculus of variations, where it originated the notions of *functional* and of *variational derivative* (Vito Volterra, 1887), and the theory of integral equations (Eric Ivar Fredholm, 1903), which served as a base for working out of the "operator approach".

From another side, developping of set-theoretical disciplines (topology, abstract geometry) prepared the "abstract" frame for constructing FA.

As a self-depending branch of mathematics FA appeared in 1904–1910, when David Hilbert, Frigues Riesz and Erhard Schmidt developped the theory of operators in infinite-dimensional Hilbert spaces. These authors demonstrated an analogy of the Fredholm theory of integral equations on the one hand and the corresponding algebraic equations on the orther hand. It was Schmidt (1908) who first introduced *geometrical language* into this subject.

Further there appeared *normed spaces*. A set of axioms, close to the axioms of a normed space, was at first introduced by A. Bennett (1916). The axioms of a *complete* normed space was given by F. Riesz in 1918, and independently by Stefan Banach, Hans Hahn and Norbert Wiener in 1922. It was constructed a rich theory of normed spaces (S. Banach, H. Hahn, T. H. Steinhaus, J. Schauder). The crown of this theory and the first *book* on FA was "Operations linéaires" by S. Banach (Warszawa, 1932).

Since it was Banach, who gave the greatest contributions to the theory of such spaces, complete normed spaces became the name "Banach spaces".

Then analysis in normed spaces was developped (based on the notion of Fréchet derivative, (Maurice René Fréchet), 1925). The interest to FA grew when one found applications of FA to theoretical physics, in particular to quantum mechanics (the theory of Hilbert spaces).

Since the frames of normed spaces appeared to be too constraint in some questions, it was introduced a more general notion of *topological vector space* (TVS) (A. N. Kolmogorov, 1934; J. von Neumann, 1935). The most important subclass of TVS's are so-called *locally convex spaces* (LCS's). The complete metrizable TVS's (resp. LCS's) obtained the name *F-spaces* (resp. *Fréchet spaces*). Banach spaces are a special case of Fréchet spaces, and Hilbert spaces are a special case of Banach ones.

[Notice, that afterwards yet more general spaces were introduced, viz. pseudotopological (or convergence, or limit) vector spaces.]

* * *

Since 1932 FA became an *universal language* of analysis. It is hardly possible now to indicate a sharp boundary between the "usual" analysis and FA. The latter embraces such branches of mathematics as measure theory, convex analysis, semi-group theory, spectral theory of linear operators, distribution theory (the theory of generalized functions), the theory of differential (and pseudo-differential) operators, the ergodic theory, the theory of Banach algebras, fixed-point theorems, numerical approximation methods, differential calculus in TVS's, the theory of extremal points, the theory of extremal problems ...

* * *

In this course we shall subsequently consider more and more "concrete" spaces: at first general TVS's, then LCS's, F-spaces, Banach spaces and at last Hilbert spaces, the theory becoming more and more rich.

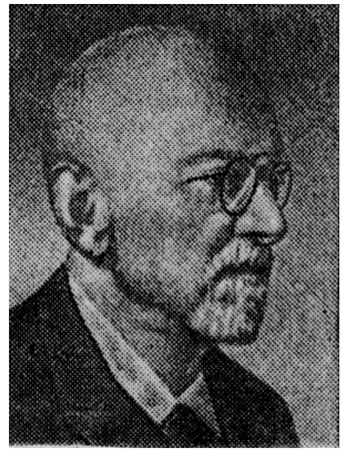
Contents

Short chronological table

1887 V. Volterra: variational derivative
1903 I. Fredholm: the theory of integral equations
1904 D. Hilbert: Hilbert spaces
1908 E. Schmidt: geometrical language for Hilbert spaces
1913 R. Gâteaux: Gâteaux derivative
1916 A. Bennett: axioms of a normed space
1918 F. Riesz: complete normed spaces
1922 H. Hahn: boundedness principle for functionals
1922 S. Banach, 1923 T. H. Hildebrandt: boundedness principle for operators
1925 M. Fréchet: Fréchet derivative
1927 H. Hahn: Hahn-Banach theorem, dual space
1929 S. Banach: Hahn-Banach theorem, dual operator, openness principle
1932 S. Banach: "Operations linéaires", the first and classical book on functional analysis
1934 A. N. Kolmogorov, J. von Neumann: topological vector spaces (TVS)
1935 A. N. Kolmogorov, J. von Neumann: bounded sets in TVS
1950 B. J. Pettis: boundedness and openness principles for TVS

V. Volterra

I. Fredholm



D. Hilbert

H. Minkovski

F. Riesz

S. Banach

M. Fréchet

A. N. Kolmogorov

J. von Neumann

1 Topological vector spaces

In this first chapter we consider rather "poor" objects, on which we have only two structures, algebraic one and topological one. At first we give the definition and some examples of TVS and indicate some basic properties of such spaces. Then we prove three fundamental principles ("three whales") of linear functional analysis: Hahn-Banach theorem, openness principle and boundedness principle. In this connection we study the most important subclass of TVS, viz. locally convex spaces.

1.1 Definition, examples and basic properties

Here we give the definition and some examples of TVS, discuss elementary properties of TVS and prove some results on properties of neighbourhoods in TVS. For that end we introduce notions of balanced set and absorbing one. At last we prove some elementary results on linear mapping of TVS.

1.1.1 Definition, examples and elementary properties

A topological vector space is a set, on which we have simultaneously two structures, one of vector space and one of topological space, these structures being compatible one with another in a natural sence. For definiteness we consider throughout vector spaces over $\mathbb{R}$.

Definition. A *topological vector space* (TVS) is a vector space (v. s.) X, supplied with a topology τ, which is *linear* (or *compatible* with the linear structure of X) in the sence, that arithmetic operations

$$\begin{aligned} + \quad &: X \times X \longrightarrow X,\ (x_1, x_2) \longmapsto x_1 + x_2, \\ \cdot \quad &: \mathbb{R} \times X \longrightarrow X,\ (t, x) \longmapsto tx \end{aligned}$$

are continuous.

Instead of (X, τ) one writes usually simply X. Continuity of adding ("+") means (below $\mathcal{N}(x)$ (or Nb_x) denotes the set of all neighbourhoods of x in τ) that

$$\forall \hat{x}_1, \hat{x}_2 \in X\, \forall V \in \mathcal{N}(\hat{x}_1 + \hat{x}_2)\, \exists U_1 \in \mathcal{N}(\hat{x}_1)\, \exists U_2 \in \mathcal{N}(\hat{x}_2) \,:\, x_1 \in U_1\, x_2 \in U_2 \Longrightarrow x_1 + x_2 \subset V.$$

The last implication may be written in the form

$$U_1 + U_2 \subset V.$$

Continuity of multiplying by scalar ("·") means that

$$\forall t_0 \in \mathbb{R}\, \forall x_0 \in X\, \forall V \in \mathcal{N}(tx)\, \exists \delta > 0\, \exists U \in \mathcal{N}(x_0) \,:\, |t - t_0| \leq \delta,\ x \in U \Longrightarrow tx \in V.$$

The last implication may be writen in the form

$$(t_0 + I_\delta)U \subset V,$$

where we use the notation

$$I_\delta := \{t \in \mathbb{R} \mid |t| \leq \delta\} \quad (\delta \geq 0).$$

It follows immediately from the definition that arithmetical operations define the following natural *homeomorphisms* of TVS:

Lemma on homeomorphisms of a TVS. *Let X be a TVS. Then*

a) translation by any fixed vector a

$$x \longmapsto x + a,\ X \longrightarrow X$$

is a homeomorphism; in particular, for every x a set U is a neighbourhood of 0 iff $x + U$ is a neighbourhood of x;

b) homothetic transformation (with the center at 0*) with any nonzero coefficient* t

$$x \longmapsto tx, \ X \longrightarrow X$$

is a homeomorphism; in particular, ihe image of every neighbourhood of 0 *by a nontrivial homothetic transformation is a neighbourhood of* 0*.*

$\lhd$ a) There exists inverse mapping $x \mapsto x - a$, and it is continuous (as addition with a *fixed* vector $-a$).

b) There exists inverse mapping $x \mapsto \frac{1}{t}x$, and it is continuous (a multiplication by a *fixed* scalar $1/t$). $\rhd$

Exercise. Let X be a TVS. Prove that $\forall n \in \mathbb{N}\ \forall t_1, \ldots, t_n \in \mathbb{R}$, the mapping

$$(t_1, \ldots, t_n, x_1, \ldots, x_n) \longmapsto \sum_{i=1}^{n} t_i x_i, \quad \mathbb{R}^n \times X^n \longrightarrow X$$

is continuous (that is taking linear combination with fixed number of terms is a continuous operation).

Thus, transformations of open sets are open sets and nontrivial homothetic images of open sets are open sets (and the same is true, surely, for closed sets).

Corollary. Let X be a TVS. Each neighbourhood of $x \in X$ has the form $x + U$, where U is a neighbourhood of 0. Besides, if U runs over a base of neighbourhoods of 0, then $x + U$ runs over a base of neighbourhoods of x.

Thus, we may confine ourselves by consideration of neighbourhoods of 0.

Exercise. Prove that a TVS is *Hausdorff* (that is $\forall x_1, x_2,\ x_1 \neq x_2,\ \exists U_1 \in \mathcal{N}(x_1)\ \exists U_2 \in \mathcal{N}(x_2) : U_1 \cap U_2 = 0$), iff

$$\forall x \neq 0\ \exists U \in \mathcal{N}(0) : x \notin U.$$

[*Hint*: Use the fact that $\forall U \in \mathrm{Nb}_0\ \exists V \in \mathrm{Nb}_0 : V - V \subset U$. This latter fact follows from continuity of *substraction* in TVS's (see Exercise on p. 2).]

Examples of TVS.

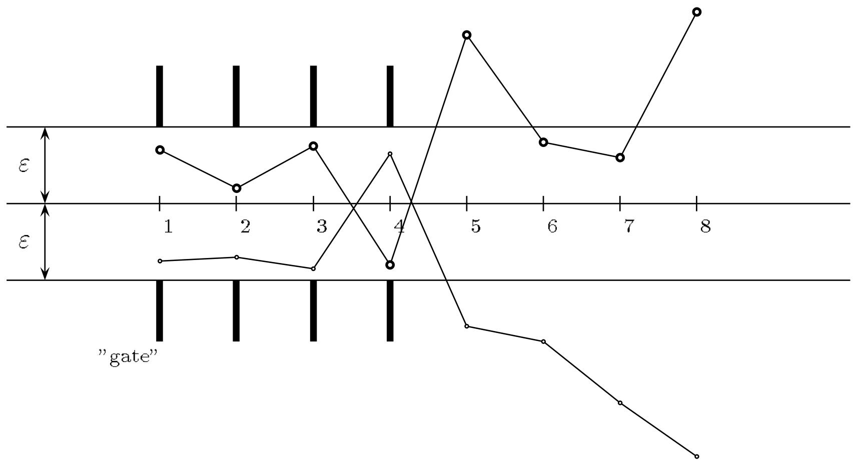
1. $\mathbb{R}^n$ with the usual topology. [Remark. This usual topology is the *unigue* Hausdorff linear topology in $\mathbb{R}^n$.]

2. More generally, all *normed spaces* (see chapter 3), for example $C([a, b])$, $C^1([a, b])$, l_2.

3. $\mathbb{R}^\infty$, the space of all sequences of real numbers $x = (x_1, x_2, \ldots)$, with the following base of neighbourhoods of 0:

$$U_{n,\varepsilon} := \{x \mid |x_i| \leq \varepsilon,\ i = 1, \ldots, n\} \quad n \in \mathbb{N}, \varepsilon > 0$$

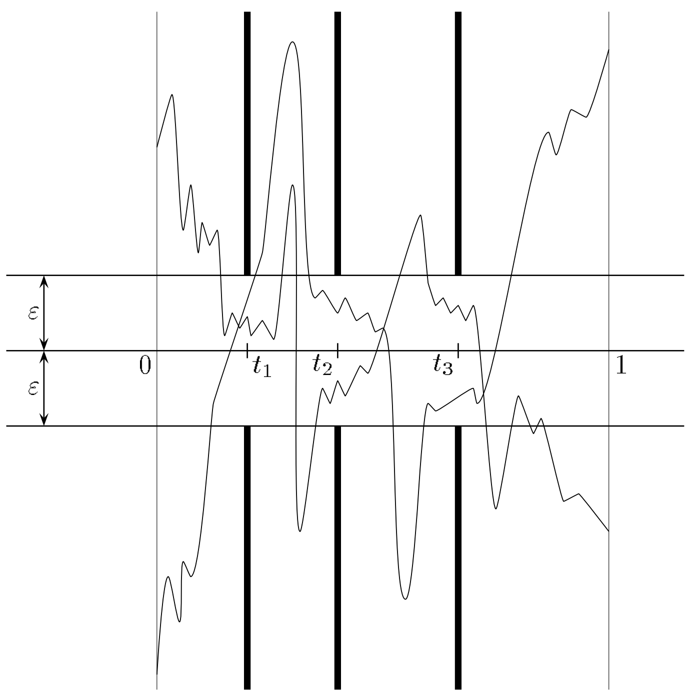
(a finite number of "gates").



In orther words, it is $\mathbb{R} \times \mathbb{R} \times \ldots$ with the product topology.

4. $\mathbb{R}^{[0,1]} = \mathcal{F}([0,1])$, the space of *all* real-valued functions on $[0,1]$, with the *topology of simple* (or *pointwise*) *convergence,* for which a base of neighbourhoods of 0 is formed by the sets

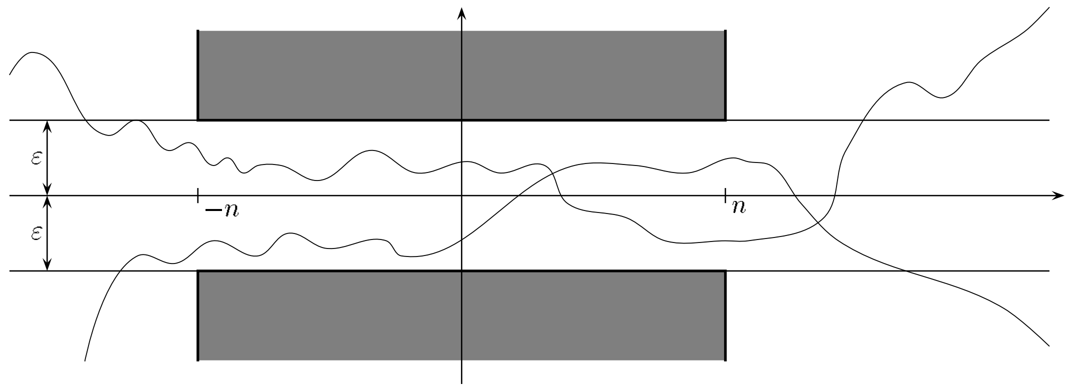
$$U_{t_1,\ldots,t_n,\varepsilon} := \left\{x \middle| \, |x(t_i)| \leq \varepsilon, i = 1,\ldots,n\right\}, \quad t_1,\ldots,t_n \in [0,1], n \in \mathbb{N}, \varepsilon > 0.$$



[Arbitrarily big (but finite!) number of arbitrarily small (but with nonzero width!) "gates".] Again it is the product topology.

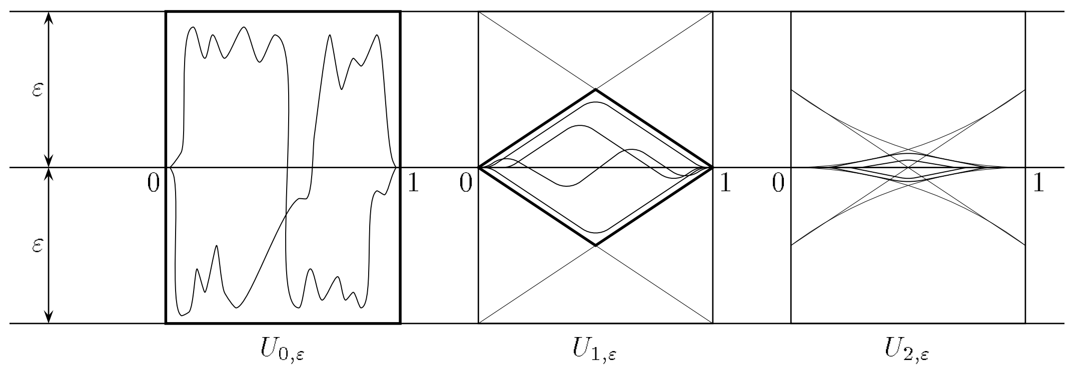
5. $C(\mathbb{R})$, the space of all *continuous* function on $\mathbb{R}$, with the following base of neighbourhoods of 0:

$$U_{n,\varepsilon} := \left\{x \middle| \, |x(t)| \leq \varepsilon \, \forall t \in [-n,n]\right\}, \quad n \in \mathbb{N}, \varepsilon > 0.$$



6. $\mathcal{D}([0,1])$, the space of all infinite-differenciable real-valued functions on $\mathbb{R}$ (sic!), which are equal to zero outside of $[0,1]$, with the base of neighbourhoods of 0

$$U_{n,\varepsilon} := \left\{ x \middle| |x(t)| \le \varepsilon, |x'(t)| \le \varepsilon, \dots, \left| x^{(n)}(t) \right| \le \varepsilon \; \forall t \right\}, \quad n = 0, 1, 2, \dots, \; \varepsilon > 0.$$

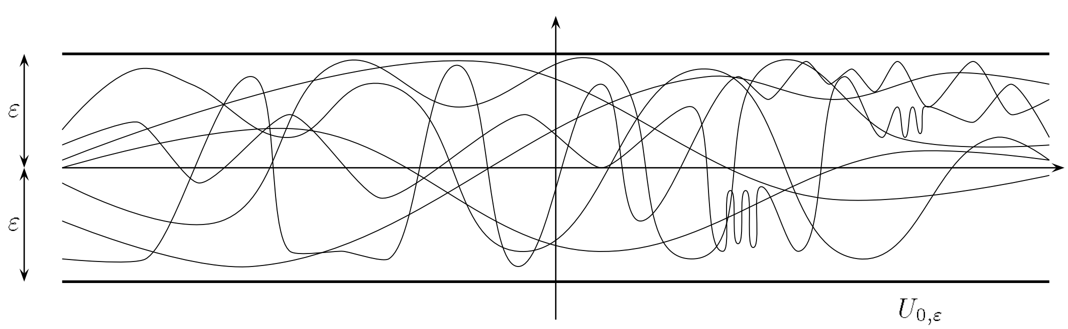


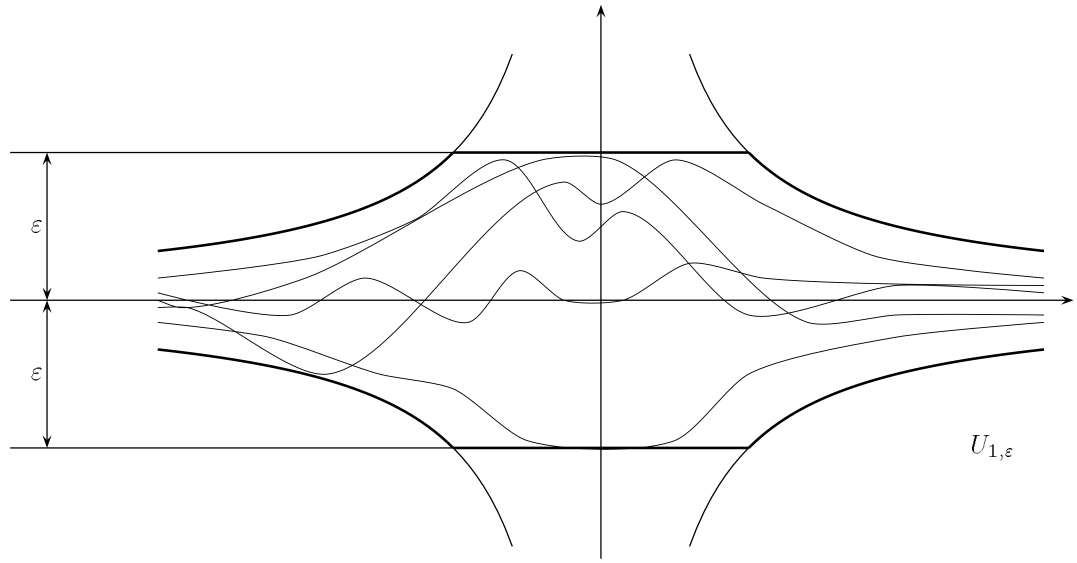
7. $\mathcal{S}(\mathbb{R})$, the space of all infinite-differentiable real-valued functions x on $\mathbb{R}$, that *quickly decrease* on infinity with all their derivatives in the sense, that for every $n = 0, 1, 2, \dots$ we have

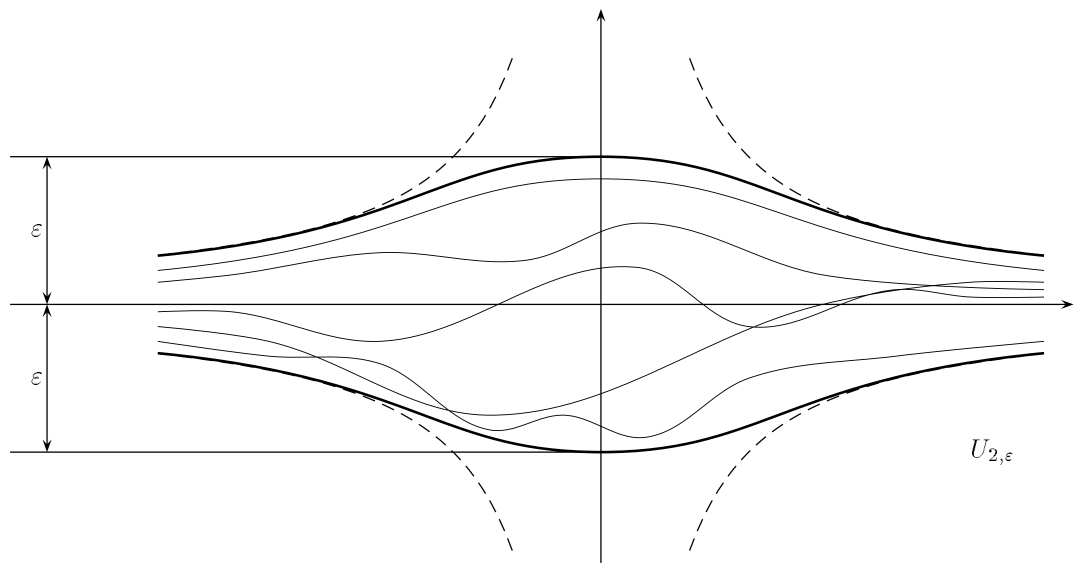
$$\|x\|_n := \mathbb{R} \sup_{\substack{k,q \le n \\ t \in \mathbb{R}}} \left| t^k x^{(q)}(t) \right| < \infty,$$

with the base of neighbourhoods of 0

$$U_{n,\varepsilon} := \left\{ x \middle| \|x\|_n \le \varepsilon \right\}.$$





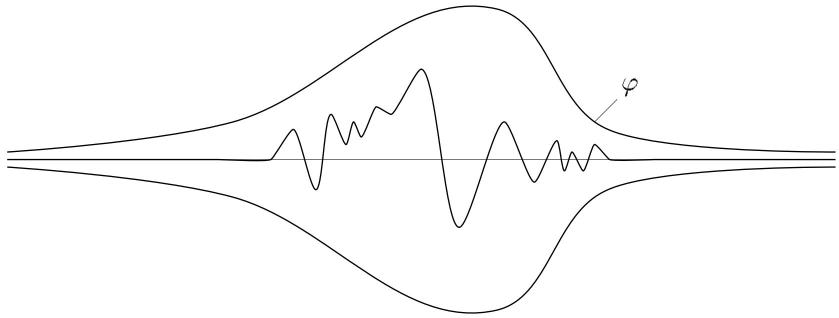


8. $\mathbb{C}_0(\mathbb{R})$, the space of all *finitary* (that is, equal to zero outside some compact interval) continuous real-valued functions on $\mathbb{R}$ (emphasize that this compact interval may depend on the function), with the base of neighbourhoods of 0

$$\{B_\varphi\},$$

where φ is any *everywhere positive* continuous function on $\mathbb{R}$, and

$$B_\varphi := \big\{x \in \mathbb{C}_0(\mathbb{R}) \,\big|\, |x(t)| \le \varphi(t),\, \forall t \in \mathbb{R}\big\}.$$



Exercise. Prove that a *sequence* in $\mathbb{C}_0(\mathbb{R})$ converges to 0 iff all the functions from this sequence vanish outside some *common* compact interval, and on this anterval the sequence converges to 0 uniformly.

Exercise. Verify that in all above example we obtain really TVS.

1.1.2 Balanced sets and absorbing sets

For describing of properties of neighbourhoods of 0 we need two notions.

Definition. A set A in a vector space X is caled *balanced* if

$$I_1 A = A,$$

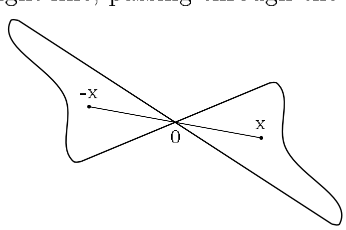
and *absorbing*, if

$$\mathbb{R}A = X.$$

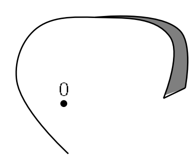
The set $I_1 A$ is called the *balanced hull* of A.

[Recall that, for any $T \subset \mathbb{R}$, $TA := \{tx \,|\, t \in T,\ x \in A\}$, $I_1 := [-1, 1]$.]

In other words, A is balanced if $\forall t, |t| \le 1$, we have $tA \subset A$; and A is absorbing if for every $x \in X$ we have $x \in tA$ for some $t \in \mathbb{R}$. Geometrically balanceness means that A contains with each its point x the segment $[-x, x]$, so in particular every balanced set is *symmetric* with respect to 0 (that is $-A = A$); and absorbingness means that A intersets at least at one *nonzero* point with every straight line, passing through the origin ($\forall x \ne 0\, \exists t \ne 0 : tx \in A.$).



balanced set



absorbing set

Exercise. Prove that $I_1 A$ is balanced and coincides with the intersection of all balanced sets containing A.

In what follows we shall often use the equations

$$I_\delta A = \delta I_1 A,$$
$$I_{\delta_1} I_{\delta_2} A = I_{\delta_1 \delta_2} A.$$

They are immediate corollary of the formula

$$I_\delta A = \{tx \,|\, |t| \le \delta, x \in A\}.$$

1.1.3 Conservation of algebraical properties by topological operations and of topological ones by algebraical ones

Continuity of arithmetic operations in TVS implies that all the properties, which can be expressed in terms of these operations, are conserved by *closure*. More preciously, it holds
Lemma on conservation of algebraical properties by closure. *Let X be a TVS and let $A \subset X$. Then*

$$A \in Lin(resp.\ Aff, Conv, Cone, Bal) \Longrightarrow \bar{A} \in Lin(resp.\ abAff, Conv, Cone, Bal).$$

Here

$$\begin{aligned} A \in \mathrm{Lin} &:\Longleftrightarrow A \textit{is a linear subspace of } X, \\ A \in \mathrm{Aff} &:\Longleftrightarrow A \textit{is an affine subspace of } X, \\ A \in \mathrm{Conv} &:\Longleftrightarrow A \textit{is a convex subset of } X, \\ A \in \mathrm{Cone} &:\Longleftrightarrow A \textit{is a cone in } X \text{ (with vertex at 0)}, \\ A \in \mathrm{Bal} &:\Longleftrightarrow A \textit{is a balanced subset at } X. \end{aligned}$$

$\lhd$ Let $\alpha, \beta \in \mathbb{R}$. We have

$$\begin{array}{lll} A \in \mathrm{Lin} & \Longleftrightarrow \forall \alpha, \beta \;\vdots & \alpha A + \beta A \subset A, \\ A \in \mathrm{Aff} & \Longleftrightarrow \forall \alpha, \beta,\ \alpha + \beta = 1 \;\vdots & \alpha A + \beta A \subset A, \\ A \in \mathrm{Conv} & \Longleftrightarrow \forall \alpha, \beta \in \triangle \;\vdots & \alpha A + \beta A \subset A, \\ A \in \mathrm{Cone} & \Longleftrightarrow \forall \alpha > 0 \;\vdots & \alpha A \subset A, \\ A \in \mathrm{Bal} & \Longleftrightarrow \forall \alpha,\ |\alpha| \le 1 \;\vdots & \alpha A \subset A. \end{array}$$

[Here we use the following notation: $\alpha, \ldots, \beta \in \triangle \;:\Longleftrightarrow \alpha \ge 0, \ldots, \beta \ge 0, \alpha + \ldots + \beta = 1$.]
Prove that $A \in \mathrm{Lin} \Rightarrow \bar{A} \in \mathrm{Lin}$. Let $A \in \mathrm{Lin}$, and let α, β be fixed real numbers. We have to show that $\alpha\bar{A} + \beta\bar{A} \subset \bar{A}$, i.e., that

$$\bar{A} \times \bar{A} \subset f^{-1}(\bar{A}), \tag{1}$$

where f is the mapping

$$f : (x, y) \longmapsto \alpha x + \beta y, \quad X \times X \longrightarrow X.$$

By continuity of arithmetical operations in TVS's this mapping is continuous. By the condition, we have

$$A \times A \subset f^{-1}(A). \tag{2}$$

So

$$\bar{A} \times \bar{A} \overset{\substack{\text{General} \\ \text{topology}}}{\subset} \overline{A \times A} \overset{(2)}{\subset} \overline{f^{-1}(A)} \overset{\text{Obv.}}{\subset} \overline{f^{-1}(\bar{A})} \overset{f \in \mathrm{Cont}}{=} f^{-1}(\bar{A}),$$

whence it follows (1).

The proofs of all the rest assertions are quite analogical. $\rhd$

On the other hand, the property to be *open* is conserved by algebraical operations:
Lemma on conservation of openness by algebraical operations. *Let X be a TVS and A, B be open sets in X. Then*

a) the sum $A + B$ is open;

b) the balanced hull $I_1 A$ of A is open, if A contains 0;

c) the convex hull coA of A is open.

$\lhd$ The proof is based on two properties of TVS (see the lemma on homeomorphisms of TVS):

1) for every nontrivial homothopy transformation the image of an open set is an open set;

2) every translation of an open set is an open set.

1° $A + B = \bigcup\limits_{b\in B}(A + b)$, so a) follows from 0°2).

2° If $0 \in A$ then $I_1 A = \bigcup\limits_{|t|\leq 1} tA \overset{0\in A}{=} \bigcup\limits_{\substack{|t|\leq 1\\ t\neq 0}} tA$, so b) follows from 0°1).

3° $coA = \left\{ \sum\limits_{i=1}^{n} \alpha_i x_i \,\middle|\, x_i \in A,\ \alpha_1, \ldots, \alpha_n \in \triangle,\ n \in \mathbb{N} \right\} = \bigcup\limits_{\substack{\alpha_1,\ldots,\alpha_n\in\triangle\\ n\in\mathbb{N}}} \underbrace{\sum_{i=1}^{n} \underbrace{\alpha_i A}_{i)},}_{ii)}$

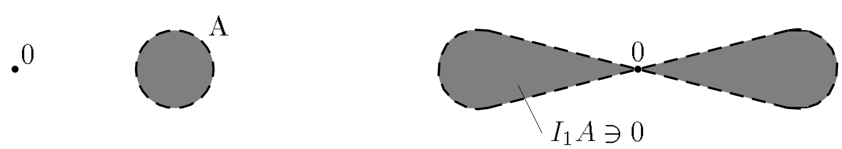
i) is open if $\alpha_i \neq 0$, by 0°1), and is $\{0\}$ if $\alpha_i = 0$

ii) is open (since not all α_i are 0), by c)

whence it follows from c). ▷

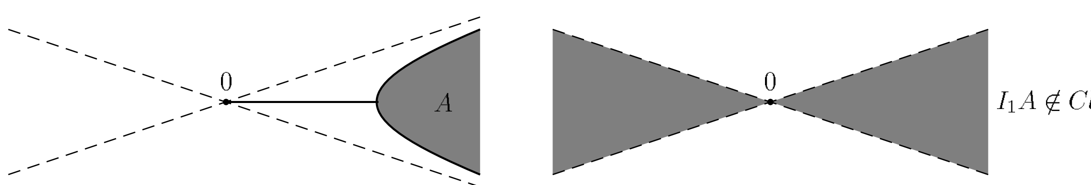
Remarks.

1. Balanced hull of an open set, *not containing* 0, may be non-open:

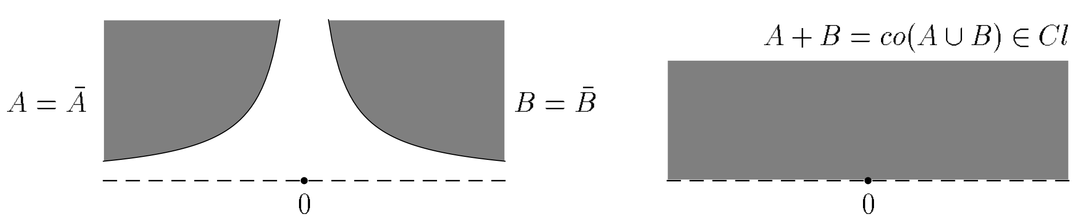


2. The property to be *closed* is not preserved by taking the sum and by passing to the balanced hull (even for sets, containing zero), nor to the convex hull (see the examples below).

Example 1.



Example 2.



1.1.4 Properties of neighbourhoods of 0 in TVS

In TVS we may without loss of generality assume that an arbitrary neighbourhood of 0 is absorbing, balanced and, as one wish, open, resp., closed:

Lemma on properties of neighbourhoods of 0. *Each neighbourhood U of 0 in TVS X*

a) is absorbing and, what is more, satisfies to following conditions:

$$\mathbb{N}U = X; \tag{1}$$

$$\forall x \in X \; \exists \delta > 0 \; : \; I_\delta x \subset U; \tag{2}$$

b) contains an open balanced neighbourhood of 0*;*

c) contains a closed balanced neighbourhood of 0*.*

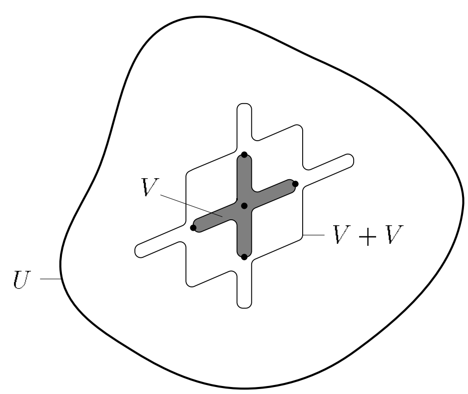
◁ 0° The proof is based on the following facts: [1]

1) the definition and elementary properties of TVS;
2) the lemma on conservation of algebraic properties by closure;
3) the lemma on conservation of openness by algebraic operations.

1° Let U be a neighbourhood of 0 in X. Let us prove a). Let $x \in X$. Since $0x = 0$, and multiplication by scalar is continuous (by 0°1)), there exist $\delta > 0$ and $V \in \mathcal{N}(x)$, such that $I_\delta V \subset U$. Hence $I_\delta x \subset U$, and hence $\frac{1}{n}x \in U$ for some sufficiently big $n \in \mathbb{N}$, so that $x \in nU \subset \mathbb{N}U$.

2° Let us prove b). Since $0 \cdot 0 = 0$, and multiplication by a scalar is continuous (by 0°1)) there exist $\delta > 0$ and $V \in \mathcal{N}(0)$, such that $I_\delta V \subset U$. Without loss of generality ⟨w.l.g.⟩ we may assume that V is open. Then the set $I_\delta V$ is an open balanced neighbourhood of 0. Inheed, $I_\delta V = (I_1\delta)V = I_1(\delta V)$ and δV is open as an image of open set by nontrivial homothetic transformation. So $I_\delta V$ is balanced (as every balanced hull is) and is open (by 0°3)) as the balanced hull of an open set.

3° Let us prove c). Since $0+0 = 0$ and addition in X is continuous (by 0°1), there exists $V \in \mathcal{N}(0)$ such that $V + V \subset U$. By b) we may assume that V is *balanced*. Then by 0°2) $\bar{V}$ is also balanced. It remains to verify that $\bar{V} \subset U$. Let $\hat{x} \in \bar{V}$. Then $\hat{x} + V$ is a neighbourhood of $\hat{x}$ (by 0°1)) and hence contains some point $x \in V$. For this x



we have

$$x \in \hat{x} + V \implies x - \hat{x} \in V \implies \hat{x} - x \in -V = V \implies \hat{x} \in x + V \subset V + V \subset U,$$

that is, $\bar{V} \subset U$. ▷

[1] Later on we shall usually omit this phrase and write simply: "0°1)..., 2)..." or "0° Lemma on ...".

1.1.5 Theorem on base of neighbourhoods of 0

The following result allows to construct TVS by claining a certain system of subsets a base of neighbourhoods of 0:

Theorem on base of neighbourhoods of 0. *For each TVS X there exists a base $\mathcal{B}$ of neighbourhoods of* 0*, such that:*

a) every $U \in \mathcal{B}$ is absorbing and balanced;

b) $U \in \mathcal{B}, t \neq 0 \Longrightarrow tU \in \mathcal{B}$ (homothety stability);

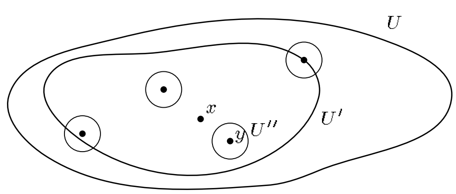
c) $\forall U \in \mathcal{B}\ \exists V \in \mathcal{B} : V + V \subset U$ (continuity of addition at $(0,0)$);

d) $\forall U_1, U_2 \in \mathcal{B}\ \exists U \in \mathcal{B} : U \subset U_1 \cap U_2$ (filter base property).

And Vice versa $\langle v.v. \rangle$, if in a vector space X a system $\mathcal{B}$ of its subsets is given, that satisfies a)–d), then there exists (the unique) linear topology in X, for which $\mathcal{B}$ is a base of neighbourhoods of 0*.*

$\lhd$ 0° Lemma on properties of neighbourhoods of 0;
1° Let X be a TVS. By 0°, we may take as $\mathcal{B}$ the system of all open (resp., closed) balanced neighbourhoods of 0. Indeed, then a) is fulfilled by 0°; b) is true, since for any nontrivial homothetic transformation the image of each open (resp., closed) neighbourhood of 0 is an open (resp., closed) neighbourhood of 0 and the image of each set is a balanced set; c) is true by continuity of addition in TVS (in the point (0,0)) and by the fact that each neighbourhood of 0 contains an open (resp., closed) balanced neighbourhood of 0 (by 0°); d) follows from the same *later* fact.
2° Let us prove the converse assertion. Let $\mathcal{B}$ be a system of subset of v. s. X, satisfying a)–d). Define a topology on X by assuming that for any $x \in X$ the sets $x + U$ where $U \in \mathcal{B}$ from a base of neighbourhoods of x. In order to verify that this definition is correct, we have, as is wellknown from topology, to show that our "neighbourhoods" $x + U$ satisfy three conditions:

1. each "neighbourhood" of x contains x;
2. the intersection of any two "neighbourhoods" of x contains a third "neighbourhood" of x;
3. for every "neighbourhood" U of x there exists a "neighbourhood" $U' \subset U$ of the same point x, such that every point $y \in U'$ is contained in U together with some its "neighbourhood":



$$\forall U \in "Nb_x"\ \exists U' \in "Nb_x"$$
$$\forall y \in U'\ \exists U'' \in "NB_y" :$$
$$U'' \subset U.$$

$\lhd\lhd$ 1. $x \in x + U$, since U contains 0 (as every *balanced* set).
2. This follows from condition d).
3. Consider the typical case of $x = 0$. Let $U \in \mathcal{B}$. We may take as U' the "neighbourhood" V, existence of which is asserted by c). Indeed, if $y \in V$, then $y + V$ is a "neighbourhood" of y and $y + V \subset V + V \subset U$. $\rhd\rhd$

Thus the topology is defined correctly.
3° Let us prove that this topology is linear. Continuity of addition follows easily from c). As to multiplication by scalars, the proof of its continuity requires some efforts.

Let $\hat{x} \in X$ and $\hat{t} \in \mathbb{R}$ be fixed. Let us show that the mapping $(t, x) \mapsto tx$ is continuous at $(\hat{t}, \hat{x})$. We have to show that $\forall V \in \mathcal{B}\, \exists \delta > 0\, \exists U \in \mathcal{B}$:

$$(\hat{t} + I_\delta)(\hat{x} + U) \subset \hat{t}\hat{x} + V,$$

or

$$I_\delta \hat{x} + I_\delta U + \hat{t} U \subset V.$$

By c), $\exists V' \in \mathcal{B}$:

$$V' + V' + V' \subset V. \tag{1}$$

[Indeed, by c), $\exists V'' \in \mathcal{B} : V'' + V'' \subset V$. Again by c), $\exists V' \in \mathcal{B} : V' + V' \subset V''$. Then $V' + V' + V' \subset \underbrace{V' + V'} + \underbrace{V' + V'} \subset V'' + V'' \subset V$.]

By b), we may look for U in the form

$$U = \alpha V' \quad (\alpha > 0).$$

So we need choose α and δ so that

$$I_\delta \hat{x} + I_\delta \alpha V' + \hat{t} \alpha V' \subset V,$$

or, by the fact, that V' is balanced,

$$I_\delta \hat{x} + \delta \alpha V' + \hat{t} \alpha V' \subset V. \tag{2}$$

At first we choose δ so that

$$I_\delta \hat{x} \subset V'$$

($V' \in$ Abs Bal!), then we choose α so that

$$\delta \alpha \leq 1 \quad \text{and} \quad \hat{t} \alpha \leq 1.$$

Then, by (1), it holds (2). $\triangleright$

Remark. The converse assertion remains true if one *omits* the condition b) of stability of $\mathcal{B}$ relative to non-trivial homothetic transformations. (The proof becomes some more complex and uses the following generalisation of (1): $\forall V \in \mathcal{B}\ \exists n \in \mathbb{N}\ \exists V' \in \mathcal{B} : \underbrace{V' + V' + \ldots + V'}_{n \text{ times}} \subset V$.) But in practical situations the condition b) is ever fulfilled.

This theorem allows us to construct TVS's by prescribing a base of neighbourhoods of 0. In particular one may verify with the aid of the theorem that all our above examples of TVS are really examples of TVS's.

Further, we may construct new TVS from given ones with the aid of the following lemma:

Lemma on product and subspaces. *The product $X \times Y$ of two TVS's X and Y (with the product topology) is a TVS. Each linear subspace Y of a TVS X (the notation: $Y \Subset X$) equipped by the induced topology, is a TVS.*

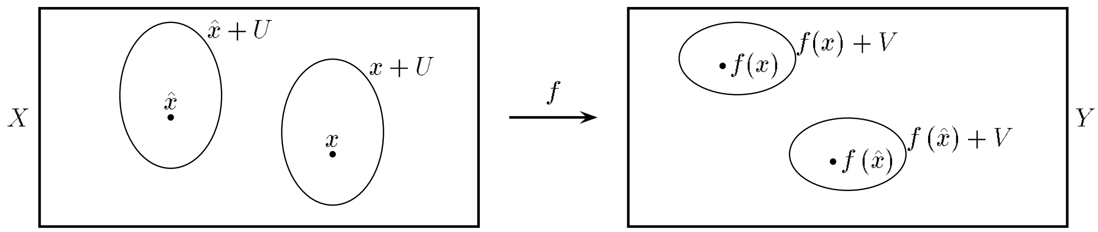
Exercise. This follows immediately from the definition.

1.1.6 Linear mappings of TVS

For TVS the key role is played by *continuous linear mappings* (c.l.m'.s). The set of all c.l.m'.s from a TVS X into a TVS Y is denoted by $\mathcal{L}(X, Y)$ (the set of all linear mapping from X into Y is denoted by $L(x, y)$). For a linear mapping of TVS, in order to establish its continuity, it is sufficient to prove continuity in any *single* point:

Lemma on continuity at one point. *Let $f : X \to Y$ be a linear mapping of TVS's. If f is continuous at any one point, it is continuous (everywhere).*

$\triangleleft$ Let f be continuous at $\hat{x}$. Let us prove, that f is continuous. Let $x \in X$ and let $f(x) + V$ (where $V \in \mathrm{Nb}_0$) be a given neighbourhood of $f(x)$.



Then $f(\hat{x})+V$ will be a neighbourhood of $f(\hat{x})$. Since f is continuous at $\hat{x}$ there exists a neighbourhood $\hat{x}+U$ of $\hat{x}$ such that $f(\hat{x}+U) \subset f(\hat{x})+V$. Then $x+U$ will be a neighbourhood of x, and we have by linearity of f

$$f(x+U) = f(x+\hat{x}-\hat{x}+U) = f(x) - f(\hat{x}) + f(\hat{x}+U) \subset f(x) - \cancel{f(\hat{x})} + \cancel{f(\hat{x})} + V,$$

that is f is continuous at x. $\rhd$

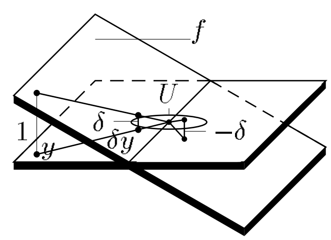
* * *

For linear *functionals* we have the following results.

Lemma on openness. *Let X be a TVS, $f : X \longrightarrow \mathbb{R}$ be any nonzero linear functional and A be any open set in X. Then $f(A)$ is an open set in $\mathbb{R}$. (In another words, any nonzero linear functional is open.)*

(Geometrically it is quite obvious!)

Emphasize that continuity of f is not supposed here.



$\lhd$ 0° In TVS every neighbourhood of 0 is absorbing and contains some balanced neighbourhood of 0.

1° Let $x \in A$ and U be a neighbourhood of 0 such that $x+U \subset A$. By $0^\circ 2)$ we may assume that U is balanced.

2° Since $f \neq 0$, $\exists y \in X$ such that $f(y) \neq 0$. Without loss of generality we may assume that $f(y) = 1$.

3° By $0^\circ 1)$ $\exists \delta > 0$ such that $\delta y \in U$. Since U is balanced, we have $I_\delta y = I_1 \delta y \subset U$.

4° By linearity of f we have $f(I_\delta y) = I_\delta f(y) = I_\delta$, hence

$$f(A) \supset f(x+U) = f(x) + f(U) \supset f(x) + f(I_\delta y) = f(x) + I_\delta,$$

which just means that $f(A)$ is open. $\rhd$

In particular, it follows from this lemma that a nonzero linear functional takes on any neighbourhood of 0 *strictly* positive and *strictly* negative values. It appears that if the values of a linear functional on a TVS are bounded from above or from below *on some neighbourhood of* 0 then the functional is continuous:

Lemma on continuity and boundedness from above. *Let X be a TVS and $f : X \to \mathbb{R}$ be a linear functional. If exists a neighbourhood U of 0 and a number $c \in \mathbb{R}$ such that*

$$f(x) \leq c \ (or\ f(x) \geq c) \quad \forall x \in U,$$

then f is continuous.

$\lhd$ Let, say, $f(x) \leq c \ \forall x \in U$ $(c \geq 0)$. By properties of neighbourhoods of 0 in TVS exists a *balanced* neighbourhood of 0

$$V \subset U.$$

For every $x \in V$ we have $-x \in V$ and hence $f(-x) \leq c$, that is $f(x) \geq -c$. Thus we have

$$|f(x)| \leq c \quad \forall x \in V,$$

that is

$$f(V) \subset I_c.$$

It follows that for every $\varepsilon > 0$

$$f(\varepsilon V) = \varepsilon f(V) \subset \varepsilon I_c = I_{\varepsilon c}.$$

Since εV is a neighbourhood of 0 together with V we may conclude that f is continuous. $\triangleright$

The set of all *continuous linear functionals* on a TVS X is denoted by X^*:

$$X^* := \mathcal{L}(X, \mathbb{R}).$$

Remarks.

1. On $\mathbb{R}^n$, every linear functional is continuous.

2. On the TVS k (see p. 41) the linear functional $x \mapsto x_1 + 2x_2 + 3x_3 + \ldots$ $(x = (x_1, x_2, x_3, \ldots))$ is *not* continuous.

The set of *all linear functionals* on a vector space X is denoted by X':

$$X' := L(X, \mathbb{R}).$$

For linear functionals it is often convenient to use the *symmetric* bracket notation:

$$\langle x', x \rangle := x'(x) \quad (x' \in X', x \in X).$$

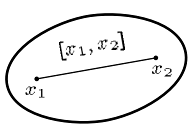
1.2 Hahn-Banach theorem

In this section we consider questions connected with the notation of convexity.

1.2.1 Basic notations of convex analysis

Let X be a vector space. A set $A \subset X$ is called *convex* if

$$x_1, x_2 \in A \Longrightarrow [x_1, x_2] \subset A,$$



where $[x_1, x_2]$ denotes the straight line segment, that joins the points x_1 and x_2:

$$[x_1, x_2] := \{x_1 + t(x_2 - x_1) \,|\, 0 \le t \le 1\}.$$

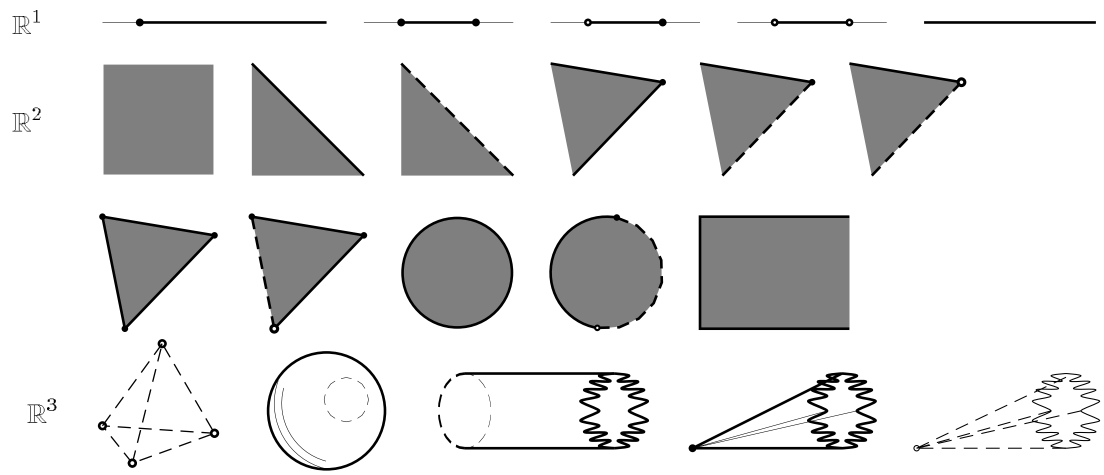
It is convinient to write this equation in the following symmetric form:

$$[x_1, x_2] = \{\alpha_1 x_1 + \alpha_2 x_2 \,|\, \alpha_1 \ge 0, \alpha_2 \ge 0, \alpha_1 + \alpha_2 = 1\}.$$

The expression to the right of the vertical bar will be often used later, and we introduce a special symbol for it:

$$(\alpha_1, \alpha_2) \in \triangle \quad :\Longleftrightarrow \alpha_1 \ge 0, \alpha_2 \ge 0, \alpha_1 + \alpha_2 = 1.$$

Examples of convex sets:



(• endpoint encluded; ∘ endpoint excluded;

—— corresponding points of the boundary encluded;

– – – corresponding points of the boundary excluded

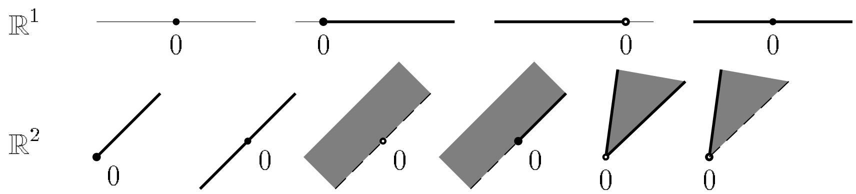
in the case of $\mathbb{R}^3$ the interpretation is left to the reader.)

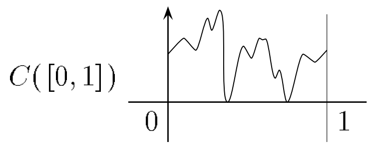
An important role will be played by convex *cones*: A set K in a vector space is called a *cone* with the vertex at 0 if

$$x \in K,\ \alpha > 0 \Longrightarrow \alpha x \in K.$$

As a rule we shall omit the words "with the vertex at 0", since we shall not deal with another cones.

Examples of convex cones:



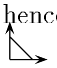


$\{x \,|\, x(t) \geq 0\ \forall t \in [0,1]\}$ (the "non-negative" cone).

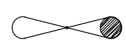
Elementary properties of convex sets. Convexness is conserved by any intersections and by *finite* summation:

$$A_\iota \in \text{Conv}\ \forall \iota \in I \quad \Longrightarrow \quad \bigcap_{\iota \in I} A_\iota \in \text{Conv}, \tag{1}$$

$$A_1, A_2 \in \text{Conv} \quad \Longrightarrow \quad A_1 + A_2 \in \text{Conv}. \tag{2}$$

$\lhd$ (1) Let $x, y \in \bigcap A_\iota$. Then $x, y \in A_\iota\ \forall \iota$, and hence $[x,y] \subset A_\iota\ \forall \iota$, that is $[x,y] \subset \bigcap A_\iota$.
(2) Let $x, y \in A_1 + A_2$, and let $(\alpha + \beta) \in$  . Then $x = x_1 + x_2$, $y = y_1 + y_2$, for some $x_1, y_1 \in A_1$, $x_2, y_2 \in A_2$, and hence

$$\alpha x + \beta y = \alpha(x_1 + x_2) + \beta(y_1 + y_2) = (\alpha x_1 + \beta y_1) + (\alpha x_2 + \beta y_2) \in A_1 + A_2. \rhd$$

Remark. The balanced hull of a convex set may be nonconvex: .

Convex combinations and convex hull. To the notions of linear combination and linear hull there correspond in convex analysis the notions of convex combination and convex hull.

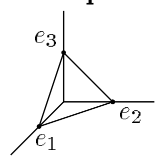
Recall that a *linear combination* of elements $x_1, \ldots, x_n$ in a linear space X is a sum $\sum_{i=1}^{n} \alpha_i x_i$, where $\alpha_i \in \mathbb{R}$. If one imposes on α_i the condition $\sum \alpha_i = 1$, then we obtain an *affine combination.*
In the convex analysis one adds the condition of *nonnegativity* of α_i :

$$\alpha_i \geq 0, \quad i = 1, \ldots, n,$$

and obtaines respectively *convex-conic combinations* and *convex combinations*:

	$\sum_{i=1}^{n} \alpha_i x_i$	**conditions on the signs**: no conditions	$\alpha_i \geq 0$
conditions on the sum	no conditions	linear combination	convex-cone combination
	$\sum \alpha_i = 1$	affine combination	convex combination

Example. The linear combinations of e_1, e_2, e_3 cover all $\mathbb{R}^3$; the affine combinations from the plane, passing through e_1, e_2, e_3; the convex-conic ones form the positive octant; and the convex ones form the triangle with the vertices e_1, e_2, e_3.



The set of all linear (resp. affine, convex-conic, convex) combinations of (arbitrary finite number of) points of a given set A in a vector space is called the *linear* (resp. *affine*, *convex-conic*, *convex*) *hull* of A and is denoted by

$$\mathrm{lin}A\,(\text{resp. } \mathrm{aff}A, \mathrm{cocon}A, \mathrm{co}A).$$

Exercise. Prove that for every set $\hat{A}$ in a vector space X

$$\mathrm{lin}\hat{A} = \bigcap_{\substack{A\in\mathrm{Lin}(X)\\ A\supset\hat{A}}} A, \quad \mathrm{aff}\hat{A} = \bigcap_{\substack{A\in\mathrm{Aff}(X)\\ A\supset\hat{A}}} A, \quad \mathrm{cocon}\hat{A} = \bigcap_{\substack{A\in\mathrm{Cocon}(X)\\ A\supset\hat{A}}} A, \quad \mathrm{co}\hat{A} = \bigcap_{\substack{A\in\mathrm{Conv}(X)\\ A\supset\hat{A}}} A, \tag{3}$$

where $\mathrm{Lin}(X)$ (resp. $\mathrm{Aff}(X)$, $\mathrm{Cocon}(X)$, $\mathrm{Conv}(X)$) denotes the set of all vector subspaces (resp. of all affine subspaces, of all convex cones, containing 0, and of all convex sets) in X.

Convex functions. In the convex analysis it is convinient to consider the so called *extended real line*

$$\overline{\mathbb{R}} := \mathbb{R}\cup\{+\infty\}\cup\{-\infty\}.$$

For a function $f : X \longrightarrow \overline{\mathbb{R}}$ (where X is a vector space) its *domain* is defined as

$$\mathrm{dom}f := \big\{x\in X\big|\, f(x)\neq+\infty\big\}, \tag{4}$$

and its *epigraph* is defined by the formula

$$\mathrm{epi}f := \big\{(x,\alpha)\in X\times\mathbb{R}\big|\,\alpha\geq f(x)\big\}. \tag{5}$$

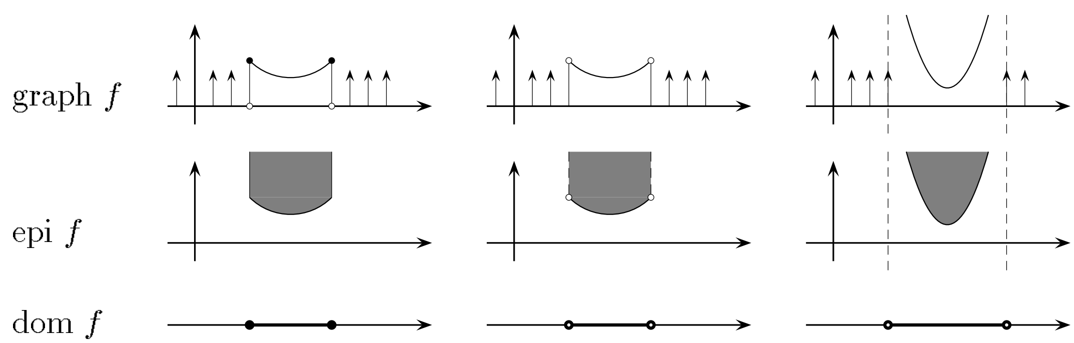
A function $f : X \longrightarrow \overline{\mathbb{R}}$ is called *convex* if its epigraph is a convex set:

$$f\in\mathrm{Conv} :\Longleftrightarrow \mathrm{epi}f\in\mathrm{Conv}.$$

Remark. $\mathrm{dom}f\subset X$ is the projection of $\mathrm{epi}f\subset X\times\mathbb{R}$ onto X.

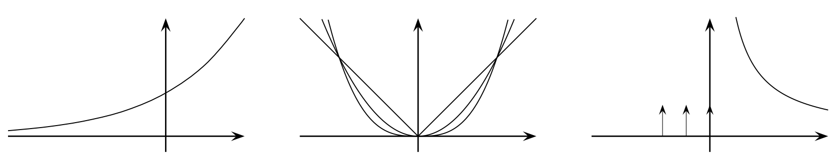
Examples of convex functions:

1. The functions $f:\mathbb{R}\longrightarrow\overline{\mathbb{R}}$ with the following graphs:



(The arrow "$\uparrow$" (resp. "$\downarrow$") means the value $+\infty$ (resp. $-\infty$).)

2. The following analytically defined functions $\mathbb{R}\longrightarrow\overline{\mathbb{R}}$:



e^{ax} $(a\in\mathbb{R})$ $\qquad |x|^p$ $(p\geq 1)$ $\qquad \begin{cases} x^\beta & \text{if } x>0\ (\beta<0)\\ +\infty & \text{if } x\leq 0\end{cases}$

3. On any vector space X functions $f \equiv +\infty$ and $f \equiv -\infty$ (with dom$f = 0$, resp. X, and epi$f = 0$, resp. $X \times \mathbb{R}$).

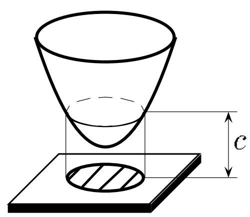
4. Every affine function (that is a linear function + constant).

5. For every *convex* set A its *indicator function* δA:

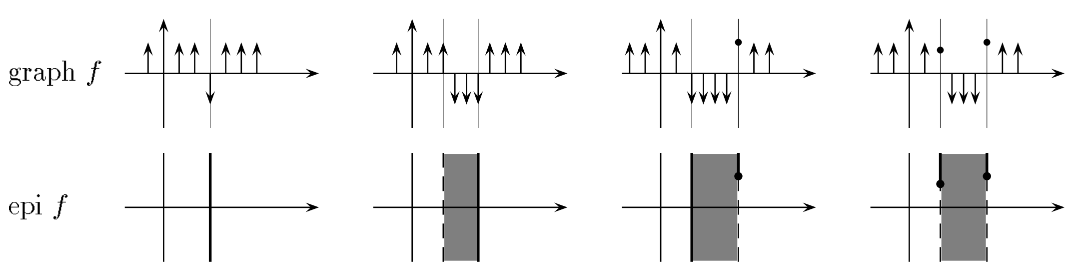
$$\delta A(x) := \begin{cases} 0 & \text{if } x \in A, \\ +\infty & \text{if } x \notin A. \end{cases}$$

Remark. In the "usual" analysis one means by indicator function of a set A the function that is equal to 1 on A and 0 at all other points.

Exercises. 1. Prove that $f \in \text{Conv} \Longrightarrow \text{dom}f \in \text{Conv}$.
2. Verify example 5 above.
3. Prove that $f \in \text{Conv} \Longrightarrow \forall c \in \mathbb{R} \,\vdots\, \left\{x \middle| f(x) \leq c\right\} \in \text{Conv}$.



The function $\equiv +\infty$ with the empty domain is of course "degenerate". Convex functions, that have (if only at one point) the value $-\infty$, are also degenerate: they may have finite values only on the "*boundary*" of its domain. For example, typical such functions f on $\mathbb{R}$ are

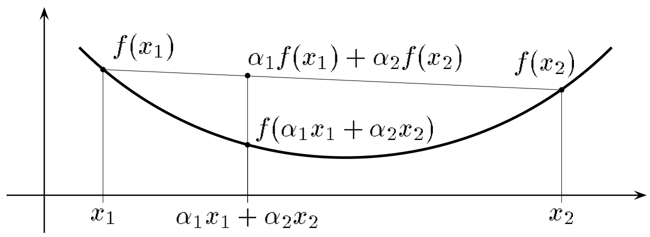


In view of this, a function f is said to be *proper* if $f(x)$ is not $+\infty$ at least at one point and is $-\infty$ at no point x.

Lemma. *A function f is convex iff (= if and only iff) it satisfies Jensen's inequality*

$$f(\alpha_1 x_1 + \alpha_2 x_2) \leq \alpha_1 f(x_1) + \alpha_2 f(x_2) \quad \forall x_1, x_2 \in dom\ X\ \forall(\alpha_1, \alpha_2) \in \Delta \tag{6}$$

◁ **Exercise.** Hint:



Corollary. A function $f : X \longrightarrow \mathbb{R} \cup \{+\infty\}$ is convex iff it satisfies Jensen inequality for *all* x_1, x_2.

▷

Elementary properties of convex functions. Convexity of functions conserves by taking *any* supremum and by summation:

$$f_\iota \in \mathrm{Conv} \quad \forall \iota \in I \quad \Longrightarrow \quad \bigvee_{\iota \in I} f_\iota \in \mathrm{Conv}, \tag{7}$$

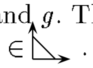
$$f, g \in \mathrm{Conv} \quad \Longrightarrow \quad f + g \in \mathrm{Conv}. \tag{8}$$

We use standard notation

$$\begin{aligned} (f \vee g)(x) &:= max\{f(x), g(x)\}, \\ (f \wedge g)(x) &:= min\{f(x), g(x)\}, \\ \left(\bigvee_{\iota \in I} f_\iota\right)(x) &:= \sup_{\iota \in I} f_\iota(x), \\ \left(\bigwedge_{\iota \in I} f_\iota\right)(x) &:= \inf_{\iota \in I} f_\iota(x). \end{aligned}$$

◁ (7) follows from the evident fact that

$$\mathrm{epi} \bigvee_{\iota \in I} f_\iota = \bigcap_{\iota \in I} \mathrm{epi} f_\iota . \tag{9}$$

As to (8), consider for simplicity the case of finite f and g. Thus we need to verify that for $f + g$ Jensen's inequality holds. Let $x_1, x_2 \in X$ and $(\alpha_1, \alpha_2) \in$  . Then

$$\begin{aligned} (f + g)(\alpha_1 x_1 + \alpha_2 x_2) &= f(\alpha_1 x_1 + \alpha_2 x_2) + g(\alpha_1 x_1 + \alpha_2 x_2) \\ &\leq \alpha_1 f(x_1) + \alpha_2 f(x_2) + \alpha_1 g(x_1) + \alpha_2 g(x_2) \\ &= \alpha_1 (f + g)(x_1) + \alpha_2 (f + g)(x_2). \ \triangleright \end{aligned}$$

Sublinear functions. An important subclass of convex functions is the class of so called "sublinear" functions. The name is not good. It would be better to call them, say, "conic" functions. But such is the tradition.

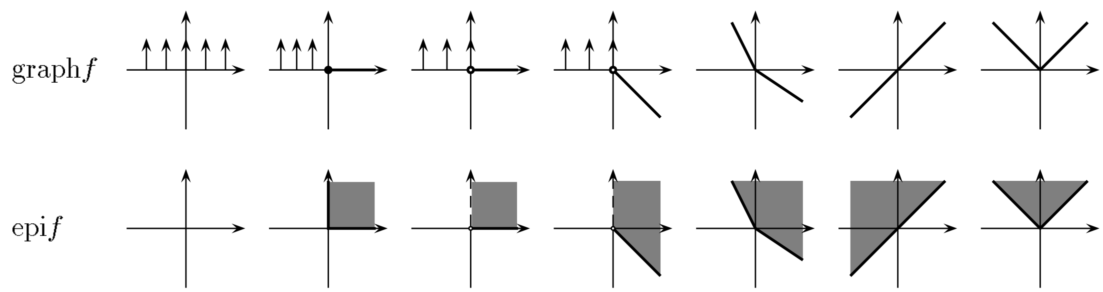
By study of sublinear functions it is convinient to introduce once more notation:

$$R^\bullet := \mathbb{R} \cup \{+\infty\}.$$

A function $p : X \longrightarrow R^\bullet$ (where X is a vector space) is called *sublinear* if epi$p$ is a convex cone:

$$p \in \mathrm{Sublin}(X) :\Longleftrightarrow \mathrm{epi} f \in \mathrm{ConvCone}(X \times \mathbb{R}).$$

Examples of sublinear functions. 1) The functions on $\mathbb{R}$ with the following graphs:



2) On every vector space the function $\equiv +\infty$ (the unique improper sublinear function).
3) Every linear functional.

Lemma. *A function $p : X \longrightarrow R^\bullet$ is sublinear iff it satisfies the following two conditions:*

a) $p(\alpha x) = \alpha p(x) \quad \forall \alpha > 0\ \forall x \in X$ *(positive homogeneity);*

b) $p(x_1 + x_2) \leq p(x_1) + p(x_2) \quad \forall x_1, x_2 \in X$ *(subadditivity).*

◁ **Exercise.** Notice that for positive homogeneous functions condition b) follows from convexity:

$$p(x_1 + x_2) = 2p\Big(\frac{x_1}{2} + \frac{x_2}{2}\Big) \leq 2\Big(\frac{1}{2}p(x_1) + \frac{1}{2}p(x_2)\Big) = p(x_1) + p(x_2). \triangleright$$

Thus the sublinear functions are just the positively homogeneous convex functions (having at no point the value $-\infty$).

Elementary properties of sublinear functions. Sublinearity is conserved by taking any supremum and by (finite) summation:

$$p_\iota \in \text{Sublin}\,\forall \iota \in I \implies \bigvee_{\iota \in I} p_\iota \in \text{Sublin}, \tag{10}$$

$$p, q \in \text{Sublin} \implies p + q \in \text{Sublin}. \tag{11}$$

◁ The conservation of convexity we have already proved, and the conservation of positive homogeneity is obvious. ▷

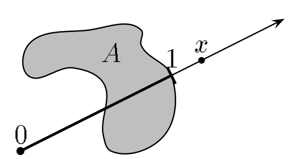
An important example of sublinear functions is Minkowski functions of convex sets. For any set A in a vector space X its *Minkowski function* μA is a function on X, defined by the formula

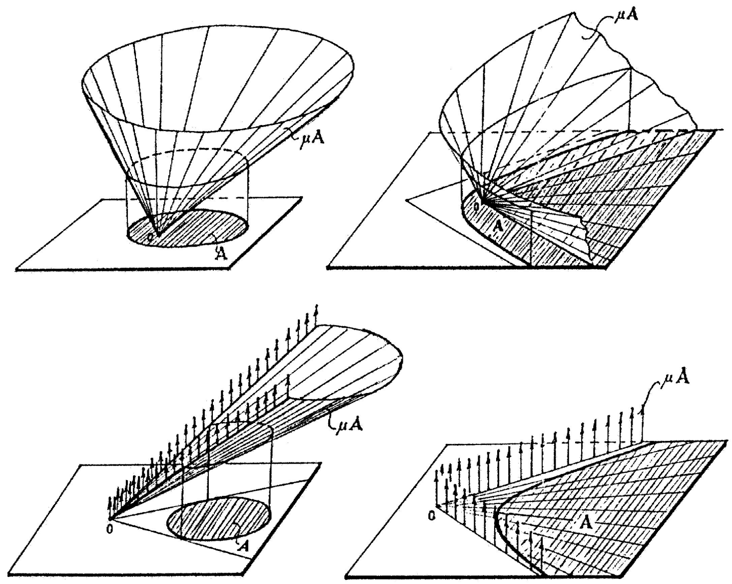
$$\begin{aligned} \mu A(x) &:= \inf\{\alpha > 0 \,|\, x \in \alpha A\} \\ &= \inf\{\alpha > 0 \,|\, \alpha^{-1} x \in A\} \end{aligned}$$

$$(\inf 0 := +\infty)$$

(As to the definition of inf0, see Remark 2 on p. 50)

Roughly speaking, $\mu A(x)$ measures the distance of x from 0, measured by means of the "unit" that is the "maximal radius" of A in the direction of x (see the picture to the left). Some typical cases are represented below on the pictures (for $X = \mathbb{R}^2$).





Exercise. Show that

$$\mu A = \bigwedge_{x \in A} e_x,$$

where

$$e_0 := \delta\{0\},$$

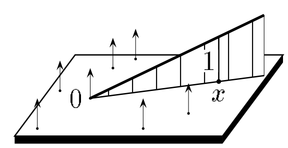
and for $x \neq 0$ the function e_x is defined as the unique sublinear function such that

$$\mathrm{dom} e_x = \mathbb{R}^+ x \setminus \{0\}$$

and

$$e_x(x) = 1.$$

In particular,



$$\mu\{x\} = e_x.$$

[Hint: $\alpha^{-1}x \in A \Leftrightarrow e_{\alpha^{-1}x}(x) = \alpha$.]

So μA is called also the *gauge function* of A.

Elementary properties of Minkowski function.

1) For every *convex* A the function μA is sublinear.

2) $A \subset B \Rightarrow \mu A \geq \mu B$.

$\lhd$ 1) We need to verify positive homogeneity and subadditivity of μA. That $\mu A(tx) = t\mu A(x)$ ($t > 0$), is true for *any* set A and follows at once from the definition of μA. Let us prove that $\mu A(x+y) \leq \mu A(x) + \mu A(y)$. Let $\alpha^{-1}x \in A$ and $\beta^{-1}y \in A$ for some $\alpha > 0$, $\beta > 0$. Then

$$(\alpha+\beta)^{-1}(x+y) = \frac{\alpha}{\alpha+\beta}\frac{x}{\alpha} + \frac{\beta}{\alpha+\beta}\frac{y}{\beta} \in A$$

by convexity of A. Hence $\mu A(x+y) \leq \alpha + \beta$, whence it follows that

$$\mu A(x+y) \leq \inf\{\alpha > 0 \,|\, \alpha^{-1}x \in A\} + \inf\{\beta > 0 \,|\, \beta^{-1}y \in A\} = \mu A(x) + \mu A(y).$$

2) It is obvious. $\rhd$

Exercises.

1. Prove that for every A we have

$$\mu(\,(0,1]A\,) = \mu A. \tag{12}$$

2. Let for $x \in \mathbb{R}^n$
$$\|x\| := \sqrt{x_1^2 + \ldots + x_n^2},$$
and let B be the unit ball
$$B = \left\{x \,\middle|\, \|x\| \le 1\right\}.$$
Verify that
$$\mu B = \|\cdot\|. \tag{13}$$
3. Prove that for every *cone* K
$$\mu K = \delta K, \tag{14}$$
where δK is the indicator function of K (see p. 17)

1.2.2 Hahn-Banach theorem

This is the first from the mentioned "three whales".

Hahn-Banach theorem. *Let X be an arbitrary vector space, let X_0 be a vector subspace in X ($X_0 \Subset X$,) let p be a sublinear functional on X, and let x_0' be a linear functional on X_0 (the notation: $x_0' \in X_0'$). Assume that*
$$x_0' \le p\big|_{X_0}.$$
Then there exists a linear functional x' on X ($x' \in X'$) such that
$$x'\big|_{X_0} = x_0'$$
and
$$x' \le p.$$

[Recall that, say, $p\big|_{X_0}$ denotes the restriction of p onto X_0 and that, say, $x' \le p$ means that $x'(x) \le p(x)$ for every $x \in X$.]

One express the content of the Hahn-Banach theorem by the words: every linear functional on a linear subspace of a vector space, which is majorized by some sublinear functional, defined on the whole space, may be *extended* (or *continued*) onto the whole space with conserving of this property.

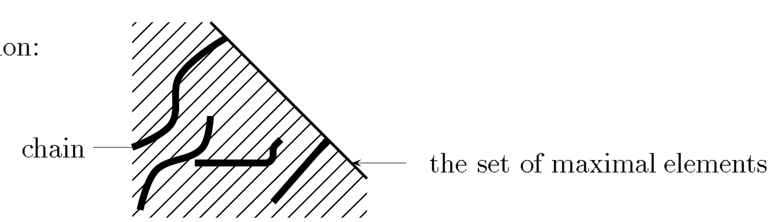
$\lhd$ 0⊳ The proof is based on the *Zorn lemma.* It is one of the axioms of the set theory, which is equivalent to the *axiom of choice.* The latter axiom says that for every set X there exists a function $\varphi : \mathcal{P}(X) \to X$ (where $\mathcal{P}(X)$ denotes the set of all subsets of X), such $\varphi(A) \in A \ \forall A \in \mathcal{P}(X)$.

In order to formulate the Zorn lemma, we need some notions. Let $(M, \prec)$ be a (partially) ordered set. Any its subset, such that each two elements a, b of this subset are comparable (that is either $a \prec b$ or $b \prec a$), is called a *chain.* An element $a \in M$ is called an *upper bound* of a subset $M' \subset M$, if $a' \prec a \ \ \forall a' \in M'$. An element $\hat{a} \in M$ is called *maximal* if
$$\hat{a} \prec a, \ a \in M \Longrightarrow a = \hat{a}.$$

Zorn lemma. *If every chain in an ordered set M has an upper bound (in this case the ordered set is said to be inductive), then there exists a maximal element in M.*

$$(x, y) \preceq (x', y') :\Longleftrightarrow x \le x', \ y \le y'$$

Illustration:



1° If $X_0 = X$, there is no to prove. Let $X_0 \neq X$, and let $h \in X \setminus X_0$. Put

$$Y := \text{lin}(X_0 \cup \{h\}) = X_0 + \mathbb{R}h = \{x + th \,|\, x \in X_0,\, t \in \mathbb{R}\}.$$

We will to extend x_0' from X_0 onto Y so that the extended linear functional $y' \in Y'$ still satisfy the majorization property: $y' \leq p\big|_Y$. It is obviously sufficient to choose a value γ of y' at the point h:

$$\gamma := \langle y', h \rangle.$$

Then

$$\langle y', x + th \rangle = \langle y', x \rangle + t\langle y', h \rangle = \langle x_0', x \rangle + t\gamma \quad (x \in X_0, t \in \mathbb{R}).$$

We want to have

$$\forall x \in X_0 \;\forall t \in \mathbb{R} \,\vdots\, \langle x_0', x \rangle + t\gamma \leq p(x + th). \tag{1}$$

For $t = 0$ this inequality is fulfilled by our assumption.
For $t > 0$ Equation (1) means if we put $t = \frac{1}{\alpha}$ $(\alpha > 0)$ that

$$\gamma \leq \alpha p\left(x + \frac{1}{\alpha}h \right) - \alpha\langle x_0', x \rangle \quad \forall \alpha > 0 \;\forall x \in X_0. \tag{2}$$

For $t < 0$ Eguation (1) means if we put $t = -\frac{1}{\alpha}$ $(\alpha > 0)$ that

$$\gamma \geq -\alpha p\left(x - \frac{1}{\alpha}h \right) + \alpha\langle x_0', x \rangle \quad \forall \alpha > 0 \;\forall x \in X_0. \tag{3}$$

It is clear that a real number γ satisfying both (2) and (3) exists iff

$$\forall \alpha, \beta > 0 \;\forall x, y \in X_0 \,\vdots\, \alpha p(x + \frac{1}{\alpha}h) - \alpha\langle x_0', x \rangle \geq -\beta p\left(y - \frac{1}{\beta}h \right) + \beta\langle x_0', y \rangle. \tag{4}$$

But 4 is indeed true:

$$\left(\alpha p\left(x + \frac{1}{\alpha}h \right) + \beta p\left(y - \frac{1}{\beta}h \right) \right) - (\,\alpha\langle x_0', x \rangle + \beta\langle x_0', y \rangle\,)$$

p is positively homog.

$$= (\,p(\alpha x + h) + p(\beta y - h)\,) - \langle x_0', \alpha x + \beta y \rangle$$

p is subadditive

$$\geq p(\alpha x + \beta y) - \langle x_0', \alpha x + \beta y \rangle \overset{p|_{X_0} \geq x_0'}{\geq} 0.$$

Hence, a desirable y' exists.
Thus, we have extended x_0' "by one dimension".
2° Now we apply the Zorn lemma. Put

$$M := \left\{ (Y, y') \,\middle|\, X_0 \Subset Y \Subset X,\, y' \in Y',\, y'\big|_{X_0} = x_0', y' \leq p\big|_Y \right\},$$

$$(Y, y') \prec (Z, z') :\Longleftrightarrow Y \Subset Z,\, y' = z'\big|_Y$$

(M consists from "partial" desirable extensions, and one such extension is "greater" than another one if the first one is an extension of the second one). This ordered set $(M, \prec)$ is *inductive*. Indeed, let $\{(Y_\alpha, y_\alpha')\}_{\alpha \in A}$ be a chain in $(M, \prec)$. Put

$$Y := \bigcup_{\alpha \in A} Y_\alpha$$

and define $y' \in Y'$ so: if $y \in Y$ then $y \in Y_\alpha$ for some α, and we put

$$\langle y', y\rangle := \langle y'_\alpha, y\rangle.$$

Since $\{(Y_\alpha, y'_\alpha)\}$ is a chain, it is clear that this definition does not depend on the choice of α, and it is evident that (Y, y') is an upper bound for the chain.

3⊳ Thus, by Zorn lemma, there exists a *maximal* element in $(M, \prec)$, say (Z, z'). If $Z \neq X$, then we can, as in 1⊳, construct an extension $(Y, y') \in M$, such that $(Z, z') \prec (Y, y')$ and $Z \neq Y$. But this contradicts to maximality of (Z, z'). Hence, $Z = X$, and z' is the desirable extension. ⊳

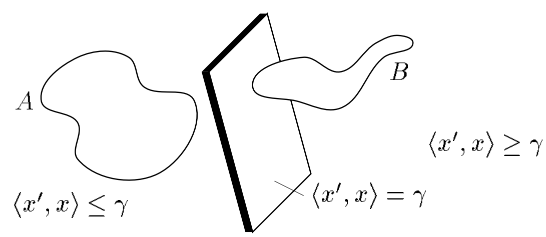
1.2.3 Separation Theorem

Let X be a vector space, let $A, B \subset X$, and let $x' \in X' \setminus 0$. We say that a *hyperplane* $\{x \in X \,|\, \langle x', x\rangle = \gamma\}$ *separates* A and B if

$$\begin{aligned} A &\subset \{x \in X \,|\, \langle x', x\rangle \geq \gamma\}, \\ B &\subset \{x \in X \,|\, \langle x', x\rangle \leq \gamma\}, \end{aligned}$$

that is, if

$$\langle x', A\rangle \leq \langle x', B\rangle. \tag{1}$$



Remark 1. Non-symmetry of A and B here is seeming: if we take $-x'$ instead of x' then A and B change their roles.

Recall that

$$\langle x', A\rangle := \{\langle x', x\rangle \,|\, x \in A\},$$

and for $P, Q \subset \mathbb{R}$

$$P \leq Q :\Longleftrightarrow \forall p \in P \ \forall q \in Q \,\vdots\, p \leq q.$$

Separation Theorem. *Let X be a TVS, and let A, B be non-empty convex subsets of X without common points, one of them, sayA, having a non-empty interior:*

$$A, B \in Conv,\ intA \neq 0,\ B \neq 0,\ A \cap B = 0.$$

Then there exists a non-zero continuous linear functional x^* on X ($x^* \in X^*$) that separates A and B.

Remark 2. If l is a *non-continuous* linear functional on a TVS X ($l \in X' \setminus X^*$), so that $l \neq 0$, then the half-spaces

$$\langle x \,|\, \langle l, x\rangle > 0\rangle \text{ and } \langle x \,|\, \langle l, x\rangle < 0\rangle$$

(obviously, convex and non-empty) can be separated by *no* $x^* \in X^*$. Prove this as an *exercise*.

◁ 0⊳

1) Hahn-Banach Theorem;
2) properties of neighbourhoods of 0 in TVS's;
3) elementary properties of convex sets and the definition and properties of Minkowski function.

1° At first we consider a special case, namely $0 \in \mathrm{int}A$, $B = \{b\}$. Consider the Minkowski function μA. We have by the definition of μA

$$\mu A \leq 1 \text{ on } A, \tag{2}$$
$$\mu A(b) \geq 1. \tag{3}$$

Put

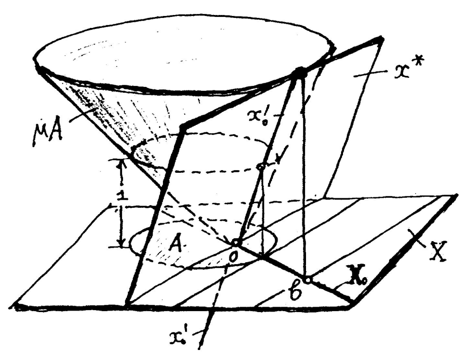
$$X_0 = \mathbb{R}b$$

and define a linear functional x_0' on X_0 by putting

$$\langle x_0', b \rangle := \gamma, \tag{4}$$

where γ is any real number between 1 and $\mu A(b)$ (see (3)):

$$1 \leq \gamma \leq \mu A(b). \tag{5}$$



(On the picture $\gamma = \mu A(b)$.) Since μA is nonnegative, we have

$$x_0' \leq \mu A \text{ on } X_0.$$

Now, μA is sublinear (see p. 20), so, by $0°1$), there exists a linear functional $x^* \in X'$ such that

$$x^*\big|_{X_0} = x_0' \text{ (that is } \langle x^*, b \rangle = \gamma(\neq 0!)\text{)} \tag{6}$$

and

$$x^* \leq \mu A. \tag{7}$$

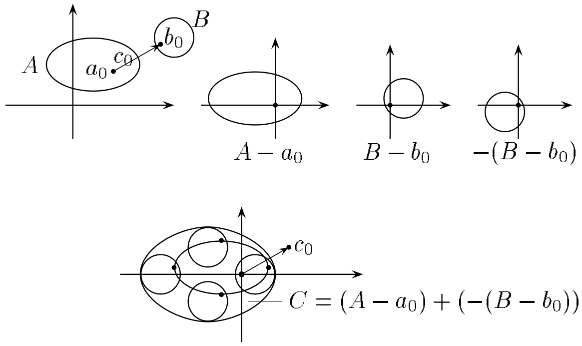
We claim that this functional x^* is continuous. Indeed, it follows from (7) and (2) that x^* is bounded from above (by 1) on A, but A is a neighbourhood of 0 by our supposition. So x^* is continuous by the lemma on p. 12.

Further, since $x^* \leq 1$ on A, we have

$$\langle x^*, A \rangle \leq 1 \leq \gamma = \langle x^*, b \rangle.$$

2° Now consider the general case. Let $a_0 \in \mathrm{int}A$, $b_0 \in B$. Put

$$C := (A - a_0) - (B - b_0),\ c_0 := b_0 - a_0.$$



Since translations, homothetic images, and sums of convex sets are again convex (by $0^\circ 3$)), the so defined C is convex. Now, since $a_0 \in \mathrm{int} A$, we have $0 \in \mathrm{int}(A - a_0)$ (by $0^\circ 2$)), and hence $0 \in C$ (since $A - a_0 \subset C$). Further, $c_0 \notin C$, otherwise we would have $c_0 = b_0 - a_0 = (a - a_0) - (b - b_0)$ for some $a \in A, b \in B$, which would imply $a = b$. But $A \cap B = 0$. Thus we have case 1° for C and c_0, so $\exists x^* \in X^*$ such that

$$\langle x^*, c_0\rangle \geq \langle x^*, C\rangle \overset{\substack{\text{def. of} \\ C \text{ and } c_0}}{=} \langle x^*, A - B + c_0\rangle \overset{\substack{\text{linearity} \\ \text{of } x^*}}{=} \langle x^*, A\rangle - \langle x^*, B\rangle + \langle x^*, c_0\rangle,$$

whence it follows that

$$\langle x^*, A\rangle \leq \langle x^*, B\rangle. \rhd$$

Remark. The condition that A has at least one interior point, was used only by proving the fact that x^* is continuous. Since in finite-dimensional case (that is for $X = \mathbb{R}^n$ with the usual metric topology) *every* linear functional is continuous (verify!), this condition may be omitted in this case.

1.2.4 Locally convex spaces

The most important class of TVS's is the class of logally convex (topological vector) spaces (LCS's).

Definition. A TVS (and its topology) is called *locally convex*, if in X there exists a base of *convex* neighbourhoods of 0.

In all the above examples of TVS's we have LCS's in fact.

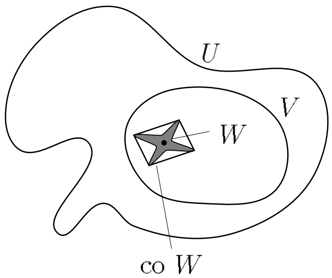
The typical examples of TVS's, which are not LCS's, are:

1. the space of all measurable functions on $[0, 1]$ with the topology of convergence in measure (see [2, p.107]);
2. the space l_p por $0 < p < 1$ (see [4, p. 247–248]).

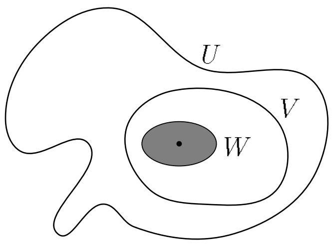
Lemma on neighbourhoods of 0 in LCS's. *In any LCS there exists a base of convex balanced open (resp., closed) neighbourhoods of 0.*

$\lhd$ 0°

1. In any TVS every neighbourhood of 0 contains a balanced open (resp., closed) neighbourhood of 0;
2. in any TVS convex hull of an open set is open (the lemma on conservation of openness);
3. in any vector space the convex hull of a balanced set is balanced (the proof is given below);
4. in any TVS the closure of a convex balanced set is a convex balanced set (the lemma on conservation of algebraical properties).



1° Let U be an arbitrary neighbourhood of 0. By the definition of an LCS, U contains some *convex* neighbourhood of 0 V. By 0°1), V contains some *balanced open* neighbourhood of 0 W. Then coW is a *convex balanced open* neighbourhood of 0 that is contained in U. Indeed, coW is balanced by 0°3), is open by 0°2), and is contained in V, since $W \subset V$ and since V is convex.



2° Let again U be an arbitrary neighbourhood of 0. By 0°1), U contains some *closed* neighbourhood of 0 V. By 1°, V contains some *convex balanced* neighbourhood of 0 W. Then the closure $\bar{W}$ is a *convex balanced closed* neighbourhood of 0 that is contained in U. Indeed, $\bar{W}$ is convex and closed by 0°4), and is contained in V, since $W \subset V$ and V is closed. $\triangleright$

1.2.5 Semi-norms and LCS's

By study of LCS's it is very useful the notions of norm and semi-norm:

Definition. A *semi-norm* in (or on) a vector space X is a function $p : X \to \mathbb{R}^{\bullet}$, that is *proper* and possesses the following properties:

1) $p(x) \geq 0 \ \forall x \in X$ (*non-negativity*);

2) $p(tx) = |t|\, p(x) \ \forall t \in \mathbb{R} \ \forall x \in X$ if the right-hand side has a sense (*homogenity*);

3) $p(x + y) \leq p(x) + p(y) \ \forall x, y \in X$ (*subadditivity*).

A *norm* is a semi-norm, that is equal to 0 at *no* $x \neq 0$ and is equal to $+\infty$ nowhere.

Remarks.

1. Thus, semi-norms are none other than *non-negative symmetric* (with respect to 0) sublinear functions.
2. In fact the property) *follows* from) and).
3. For semi-norms it holds not only the "triangle inequality") (as for all sublinear functions), but also the inequality
$$p(x - y) \geq |p(x) - p(y)|$$
("the length of a side of a triangle is greater than or equal to the (modulus of) difference of length of two other ones").

$\triangleleft$ $p(y + (x - y)) \leq p(y) + p(x - y)$ yields $p(x - y) \geq p(x) - p(y)$; switching the roles of x and y yields $p(y - x) \geq p(y) - p(x)$; but by) we have $p(y - x) = p(x - y)$. $\triangleright$

The "level sets"

$$\begin{aligned} B_p(r) &:= \{x \,|\, p(x) \leq r\} \quad r > 0 \\ \mathring{B}_p(r) &:= \{x \,|\, p(x) < r\} \quad r > 0 \end{aligned}$$

are called respectively the (*closed*) *ball* and the *open ball* of the radius r (with the center at 0), associated with a semi-norm p. For *unit* balls we write simply B_p and $\mathring{B}_p$:

$$B_p := B_p(1), \quad \mathring{B}_p := \mathring{B}_p(1).$$

Now for every nonempty *convex balanced* set B we denote by p_B the Minkowski function of B:

$$p_B(x) := \mu B(x) := \inf\{t > 0 \big| x \in tB\} \quad (x \in X)\,(\inf 0 := +\infty)$$

It appears that the correspondances

$$p \longmapsto B_p \text{ and } B \longmapsto p_B$$

are almost inverse one to another ("almost", since different sets B may have one and the same Minkowski function p_B):

Theorem on the correspondence between semi-norms and balanced convex sets. *Let X be a vector space.*

a) If p is a semi-norm in X, then its unit balls B_p and $\mathring{B}_p$ are balanced convex subsets of X, and

$$p_{B_p} = p_{\mathring{B}_p} = p.$$

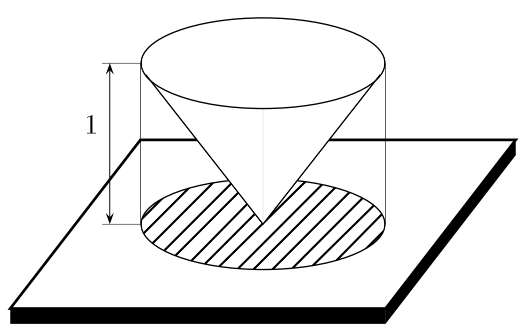
Moreover, if p is finite, then B_p and $\mathring{B}_p$ are absorbing; if p vanishes only at 0, then B_p and $\mathring{B}_p$ contain no (straight) line, passing through 0; hence if p is a norm, then B_p and $\mathring{B}_p$ are absorbing balanced convex sets, which contain no line, passing throuh 0.

b) If B is a balanced convex set, then its Minkowski function p_B is a semi-norm, and

$$\mathring{B}_{p_B} \subset B \subset B_{p_B}.$$

Moreover, if B is absorbing, then p_B is finite; if B contain no line, passing through 0, then p_B vanishes only at 0; hence if B is an absorbing balanced convex set, which contains no line, passing through 0, then p_B is a norm.

$\lhd$ All this follows easy from the definitions and the fact, that for any *convex* set A its Minkowski function μA is sublinear (see also Exercise 3 on p. 17). $\rhd$



Now bring in topology and discuss continuity of semi-norms.

Theorem on continuous semi-norms. *Let X be a TVS.*

a) If a semi-norm on X is continuous at 0 then it is continuous (everywhere).

b) A semi-norm p on X is continuous if its unit ball B_p is a neighbourhood of 0 and only if its open unit ball $\mathring{B}_p$ is a neighbourhood of 0. In this case the ball B_p is a closed set and the ball $\mathring{B}_p$ is an open set (which justifies the names "closed ball" and "open ball").

c) A balanced convex set B in X is a neighbourhood of 0 iff the associated semi-norm p_B is continuous. In this case it holds

$$\mathring{B}_{p_B} = intB, \quad B_{p_B} = \bar{B}.$$

◁ 0° The above theorem on the correspondance. [a)] This assertion follows at once from the inequality for difference of two sides of a triangle

$$|p(x+h) - p(x)| \leq p(h)$$

(see Remark 3 on p. 26).

[b)] Let p be a semi-norm on X. If p is continuous, then $\mathring{B}_p$ is open in X and B_p is closed in X as the pre-images of the open set $(-\infty, 1)$ and of the closed set $(-\infty, 1]$ respectivelly. Hence $\mathring{B}_p$ is in this case a neighbourhood of 0.

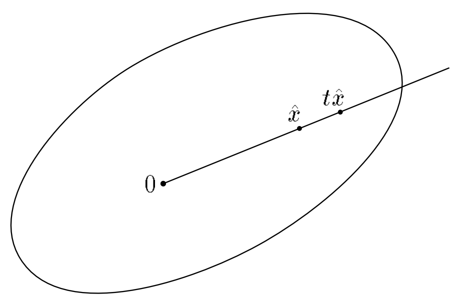
Vice versa, let B_p is a neighbourhood of 0. Let us prove that p is continuous. By a), it is sufficient to verify, that p is continuous at 0. Let it be given $\varepsilon > 0$. Then εB_p is a neighbourhood of 0 together with B_p, and it holds

$$p(\varepsilon B_p) = \varepsilon \underbrace{p(B_p)}_{\subset I_1} \subset \varepsilon I_1 = I_\varepsilon,$$

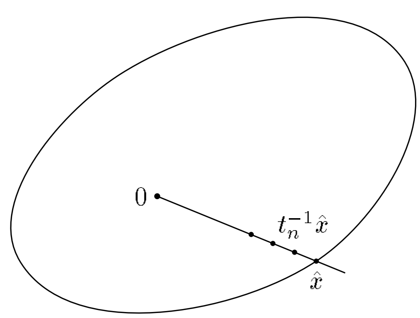
so p is continuous at 0.

[c)] Let B a balanced convex set in X. If B is a neighbourhood of 0, then B_{p_B} is also a neighbourhood of 0, since $B \subset B_{p_B}$ by 0°. Hence by b) p_B is continuous. Vice versa, if p_B is continuous, then by b) $\mathring{B}_{p_B}$ is open in X, and therefore B is a neighbourhood of 0, since $B \supset \mathring{B}_{p_B}$ by 0°.

Now let us prove that if B is a balanced convex neighbourhood of 0, then $\mathring{B}_{p_B} = \text{int}B$ and $B_{p_B} = \bar{B}$. Since p_B is in this case continuous, it follows that $\mathring{B}_{p_B}$ is open and B_{p_B} is closed (by b)). So it is sufficient to show that $\text{int}B \subset \mathring{B}_{p_B}$ and $B_{p_B} \subset \bar{B}$.



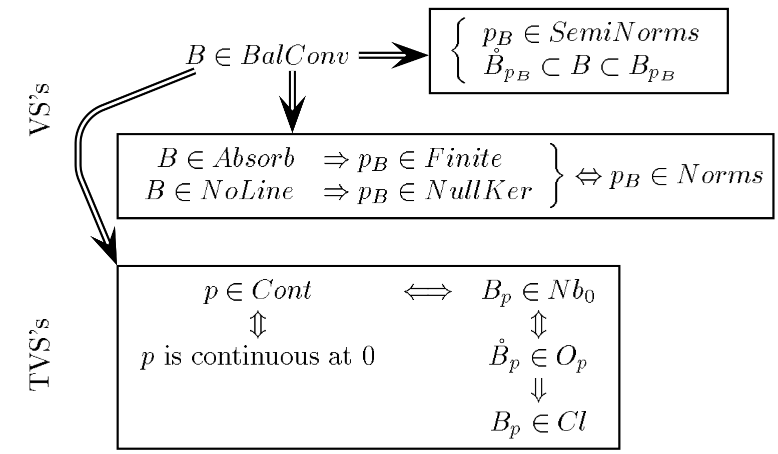
Let $\hat{x} \in \text{int}B$. Then B is a neighbourhood of $\hat{x}$. Since the mapping $(t, x) \mapsto tx$ is continuous at $(1, \hat{x})$, there exists $t > 1$ such that $t\hat{x} \in B$. Hence, by the definition of Minkovski function, $p_B(\hat{x}) < 1$, that is, $\hat{x} \in \mathring{B}_{p_B}$.

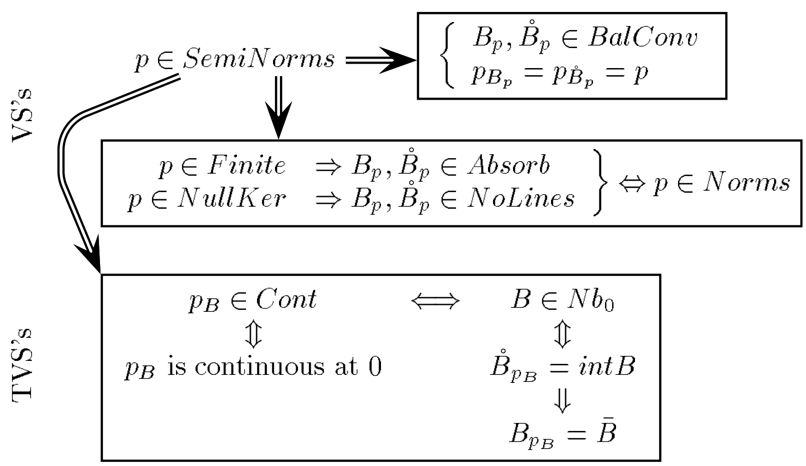


Let $\hat{x} \in B_{p_B}$, that is, $p_B(\hat{x}) \leq 1$. If $p_B(\hat{x}) < 1$, then $\hat{x} \in \mathring{B}_{p_B} \overset{0^\circ}{\subset} B \subset \bar{B}$. If $p_B(\hat{x}) = 1$ then by the definition of Minkowski function there exists a sequence $t_n \to 1$ $(t_n \geq 1)$, such that $t_n^{-1}\hat{x} \in B$. Since, again, the multiplication by scalar is continuous at $(a, \hat{x})$ and $t_n^{-1} \to 1$, we conclude that $\underbrace{t_n^{-1}\hat{x}}_{\in B} \to \hat{x}$ and hence $\hat{x} \in \bar{B}$. ▷

Remark 1. The continuity at 0 implies continuity (everywhere) for *any sublinear* function (prove!).

Summarize the two above theorems in the following diagram:





(the notations must be clear from the above exposition.)

Remark 2. We said above about continuity of functions $p : X \to R^\bullet$, but we did not define what it is! Of course, it should be understood so: we say that p is continuous at x if p is *finite* at some neighbourhood U of x and the restriction $p\big|_U$ is continuous at x in the usual sense. In particular, a continuous (that is, continuous at all points) function p is everywhere finite.

* * *

The notion of semi-norm allows to give the two following examples of LCS's:

Example 1. *Semi-normed and normed spaces.* Let p be a *finite* semi-norm (resp., norm) in a vector space X. We take as a base of neighbourhoods of 0 the system of all open balls $\mathring{B}_p(r)$, $r > 0$. It is easy to verify that all the conditions of the theorem on base of neighbourhoods of 0 are fulfilled:

a) each ball $\mathring{B}_p(r)$ is absorbing, since p is finite;

b) $t\mathring{B}_p(r) = \mathring{B}_p(tr)$;

c) $\mathring{B}_p\left(\frac{r}{2}\right) + \mathring{B}_p\left(\frac{r}{2}\right) \subset \mathring{B}_p(r)$ by convexity of p;

d) $\mathring{B}_p(r_1) \cap \mathring{B}_p(r_2) = \mathring{B}_p\ (\min(r_1, r_2))$.

Thus, the system $\mathring{B}_p(r)$, $r > 0$, defines a structure of TVS in X. Further, each our open ball is convex, by the above theorem on p. 27. So this TVS is an LCS. Such spaces are called *semi-normed* (resp., *normed*). *Normed* spaces are exactly *Hausdorff* semi-normed ones (verify!).

Example 2. *LCS, generated by a given system of semi-norms.* More generally, let $(p_\alpha)_{\alpha \in A}$ be any family of *finite* semi-norms in a vector space X. Take as a base of neighbourhoods of 0 the balls $\mathring{B}_{p_\alpha}(r)$, $\alpha \in A$, $r > 0$, *and* all their finite intersections. Just as in the above example, one verifies that this makes from X an LCS (called the LCS, *generated* by the system p_2).

It appears that the last example is the most general example of LCS. Viz., it holds
Theorem on generating by semi-norms. *In any LCS X the topology may be generated by a family of semi-norms. One may take as such a family the family of all continuous semi-norms on X.*

◁ 0°

1. Every neighbourhood of 0 in LCS contains a convex balanced open neighbourhood of 0 (the lemma on neighbourhoods of 0 in LCS);
2. a semi-norm p is continuous iff the ball $\mathring{B}_p$ is an open set (the theorem on semi-norms in TVS);
3. if B is an open balanced convex set in a TVS, then p_B is a continuous semi-norm in this TVS and $\mathring{B}_{p_B} = B$ (the same theorem).

1° Consider the family $\mathcal{P}$ of all continuous semi-norms on X. By 0°2) the balls $\mathring{B}_p$, $p \in \mathcal{P}$, are open, hence all the balls $\mathring{B}_p(r) = r\mathring{B}_p$, $r > 0$, $p \in \mathcal{P}$, and their finite intersections are open and hence are neighbourhoods of 0. Thus, the local convex topology, generated by the family $\mathcal{P}$, is *coarser* than the original topology.
2° On the orther hand, let U be a neighbourhood of 0 in X. By 0°1) U contains an open balanced convex neighbourhood of 0 B; by 0°3), the associated semi-norm p_B is continuous, that is, $p_B \in \mathcal{P}$, and $B = \mathring{B}_{p_B}$. Hence every neighbourhood of 0 in X contains a neighbourhood of 0 in the topology, generated by $\mathcal{P}$. Thus, the latter topology is *finer* than the original one. ▷

Remarks.

1. A family of semi-norms, generating the topology of LCS, is not uniquely defined. We may add to or exclude off the family, say, the semi-norm $2p$ (if p belongs to the family.)
2. The locally convex topology, generated by a system $\mathcal{P}$ of finite semi-norms on a vector space X, is the *weakest* locally convex topology on X, such that each semi-norm from $\mathcal{P}$ is continuous. (Verify!)

1.2.6 Strict separation theorem

Let X be a vector space, let $A, B \subset X$, and let $x' \in X'$. We say that x' *strictly separated* A and B, if there exists a non-empty open interval $I \subset \mathbb{R}$ such that

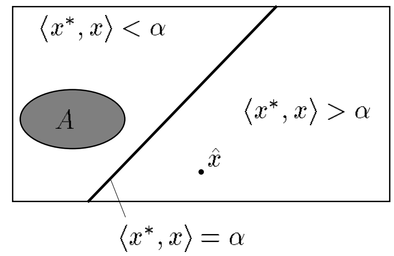
$$\langle x', A\rangle \leq I \leq \langle x', B\rangle,$$

or, which is equivalent, if

$$\sup \langle x', A\rangle < \inf\langle x', B\rangle.$$

Strict Separation Theorem.
Let X be an LCS, let A be a non-empty closed convex set in X, and let $\hat{x} \in X \setminus A$. Then $\exists x^* \in X^*$ that strictly separates A and $\hat{x}$, that is,

$$\underbrace{\sup \langle x^*, A\rangle}_{\equiv \alpha} < \langle x^*, \hat{x}\rangle.$$



It is clear that x^* satisfying this inequality, must be *non-zero*.

$\lhd$ 0°

1. Separation theorem for TVS;

2. the fact that every non-zero linear functional is open.

1° Since A is closed and $\hat{x} \notin A$, the complement $X \setminus A$ is an open neighbourhood of $\hat{x}$. Since X is an LCS, $\exists$ a *convex* neighbourhood U of $\hat{x}$ such that $U \subset X \setminus A$, that is, $U \cap A = 0$. By $0^\circ 1$), applied to A and U, $\exists x^* \in X^* \setminus 0$, such that

$$\langle x^*, A\rangle \leq \langle x^*, U\rangle,$$

or, equivalently,

$$\sup \langle x^*, A\rangle \leq \inf\langle x^*, U\rangle.$$

2° The latter infimum is *strictly* less than $\langle x^*, \hat{x}\rangle$, since the set $\langle x^*, U\rangle$ is, by $0^\circ 2$), open and, obviously, contains the point $\langle x^*, \hat{x}\rangle$. Hence it follows the desired inequality. $\rhd$

Here is a simple corollary of Strict Separation Theorem:
Lemma on non-triviality of annihilator. *Let X be an LCS, and let X_0 be a closed vector subspace of X, $X_0 \neq X$. Then $\exists$ a non-zero $x^* \in X^*$ such that*

$$\langle x^*, X_0\rangle = 0.$$

(For short, we write here and below 0 instead of more correct $\{0\}$.)
For any set $A \subset X$ the set

$$A^\perp := \{x^* \in X^* \,|\, \langle x^*, A\rangle = 0\}$$

is called the *annihilator* of A. So the lemma asserts that $X_0^\perp$ is not trivial, whence the name.

$\lhd$ 0°

1) Strict Separation Theorem;

2) the obvious fact, that a linear functional on a vector space, which is bounded from above or from below (on the *whole* space), is identically equal to 0.

1° By condition $\exists \hat{x} \in X \setminus X_0$. Hence by $0^\circ 1$) $\exists$ a (*nonzero!*) $x^* \in X^*$ such that

$$\sup \langle x^*, X_0\rangle < \langle x^*, \hat{x}\rangle.$$

2° Thus, x^* is bounded from above on X_0. So by $0^\circ 2$) $x^*\big|_{X_0} = 0$. $\rhd$

1.2.7 Totality of the dual space

On every Hausdorff LCS there exists "sufficiently many" continuous linear functionals (in this sense the space X^* is "total"):

Theorem on totality of the dual space. *Let X be a Hausdorff LCS, and let $\hat{x} \in X$. If $\langle X^*, \hat{x}\rangle = 0$ then $\hat{x} = 0$.*

$\lhd$ 0° Strict Separation Theorem.

1° Suppose that

$$\langle X^*, \hat{x}\rangle = 0, \tag{1}$$

but $\hat{x} \neq 0$. Then $\hat{x} \notin \{0\}$, and we can apply 0° to $\{0\}$ and $\hat{x}$ (since in a Hausdorff topological space every point is a closed set). We conclude that $\exists x^* \in X^*$:

$$0 \overset{\text{Obv.}}{=} \langle x^*, 0\rangle \overset{0^\circ}{<} \langle x^*, \hat{x}\rangle \overset{(1)}{=} 0. \rhd$$

Remark. For non-LCS's X it may be $X^* = \{0\}$ [2, p. 107].

1.3 Openness principle

This is the second "whale".

We shall prove this principle for so called F-spaces.

1.3.1 Definition and examples of F-spaces

Definition. A TVS X is called an *F-space* if:

1) its topology is generated by some metrics ϱ, which is *invariant relative to translations* in the sense that
$$\varrho(x, y) = \varrho(x - y, 0) \quad \forall x, y \in X;$$

2) the metric space (X, ϱ) is *complete.*

[It may be shown that the property to be complete doesn't depend on the choice of such a metric.]

We shall oft write simply "invariant" instead of "invariant relative to translations".

Usually one deals with *locally-convex* F-spaces; such spaces are called *Fréchet spaces* (whence the letter "F" comes).

Example. Every *Banach space*, that is, *complete* normed space, is a Fréchet space.

It a TVS satisfies the condition 1) above, we say, that this TVS is *metrizable.*

Thus, each F-space is metrizable. Here is an important general example of metrizable LCS's:

Example. Every LCS, generated by a *countable* systems of semi-norms (such LCS's are called *countably normed*), is metrizable, and the corresponding metrics may be chosen to be invariant. Viz., one may take the metric

$$\varrho(x, y) = \sum_{n=1}^{\infty} 2^{-n} \frac{p_n(x - y)}{1 + p_n(x - y)},$$

where $\{p_n\}$ is the generating system of semi-norms.

1.3.2 Opennes principle: Banach theorem on open mapping.

There is a number of versions of the openness principle. One of them is the Banach theorem on open mapping:

Banach theorem on open mapping *Let X and Y be F-spaces, and let $A : X \to Y$ be a continuous linear operator onto Y (that is, a surjective operator). Then A is an open mapping (that is, the image by A of every open set is an open set).*

$\lhd$ 0°

1) The *Baire theorem*: No complete metric space can be represented as the union of a countable family of nowhere-dense sets. (see [1, p. 69]);

2) in TVS $\overline{A+B} \supset \bar{A} + \bar{B}$ (the closure of a sum contains the sum of the closures) [prove this as an exercise and give an example, where the inclusion is *strict*];

3) in TVS the sum of two open sets is an open set (the lemma on conservation of openness);

4) the absorbtion property of neighbourhoods of 0 in TVS.

1° At first we shall prove that for any neighbourhood U of 0 in X the *closure* $\overline{AU}$ of its image contains some neighbourhood of 0 in Y.

Since the mapping $(x_1, x_2) \mapsto x_1 - x_2$ is continuous at $(0,0)$, $\exists$ a neighbourhood U' of 0 in X such that $U' - U' \subset U$. By $0^\circ 4)$, we have

$$X = \bigcup_{n=1}^{\infty} nU',$$

and therefore

$$Y = AX = \bigcup_{n=1}^{\infty} nAU'.$$

Hence by $0^\circ 1)$ (applied to Y which "is" a *complete* metric space) one of the sets $\overline{nAU'}$ contains a nonempty open set. Since the homothetic transform with the coefficient n is a homeomorphism of Y, we conclude that $\overline{AU'}$ contains some nonempty open set V. It follows that

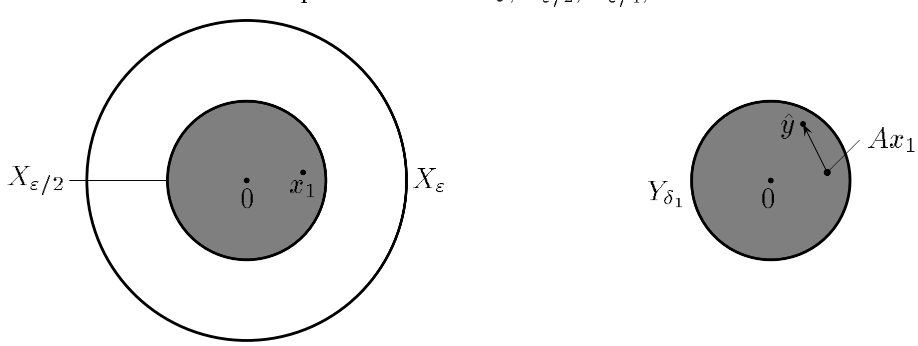
$$\overline{AU} \supset \overline{A(U'-U')} = \overline{AU' - AU'} \overset{0^\circ 2)}{\supset} \overline{AU'} - \overline{AU'} \supset V - V.$$

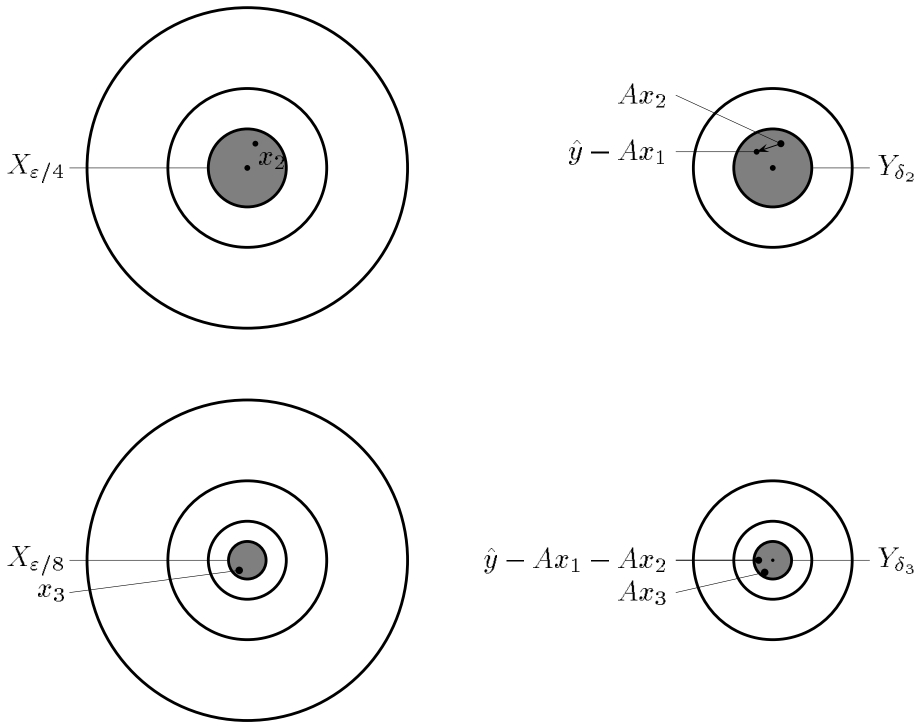
The set $V - V$ is open by $0^\circ 3)$ and obviously contains 0. Thus, $\overline{AU}$ contains a neighbourhood of 0.

2° Now let us prove that for any neighbourhood U of 0 in X already its image AU *itself* contains some neighbourhood of 0 in Y.

For each $\varepsilon > 0$ we denote by X_ε and Y_ε the closed (!) balls of radius ε with the center at 0 in X and Y respectively relative to fixed translation invariant metrics in X and Y, generating their topologies.

Now let it be given an arbitrary $\varepsilon > 0$. It is sufficient to verify that the image of X_ε contains Y_δ for some $\delta > 0$. Consider the sequence of balls $X_\varepsilon, X_{\varepsilon/2}, X_{\varepsilon/4}, \ldots$.





By 1° the image of $X_{\varepsilon/2}$ is *dense* in some Y_{δ_1}, the image of $X_{\varepsilon/4}$ ids dense in some Y_{δ_2}, etc. Without loss of generality we may assume that $\delta_n \downarrow 0$. Let us shown that the image of X_ε *contains* Y_{δ_1}.

Let $\hat{y} \in Y_{\delta_1}$. Since the image of $X_{\varepsilon/2}$ is dense in Y_{δ_1}, $\exists x_1 \in X_{\varepsilon/2}$ such that $\hat{y} - Ax_1 \in Y_{\delta_2}$. Since the image of $X_{\varepsilon/4}$ is dense in Y_{δ_2}, $\exists x_2 \in X_{\varepsilon/4}$ such that $(\hat{y} - Ax_1) - Ax_2 \in Y_{\delta_3}$, etc.

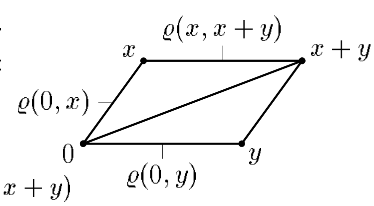
We claim that the series $\sum_{n=1}^{\infty} x_n$ converges in X to some element $\hat{x} \in X_\varepsilon$ and that $A\hat{x} = \hat{y}$. Indeed, if we denote by $|x|$ the distance (in our fixed invariant metric ϱ in X) between x and 0 and put $\hat{x}_n := x_1 + \ldots + x_n$, then

$$|\hat{x}_n| = |x_1 + \ldots + x_n| \overset{!}{\leq} |X_1| + \ldots + |x_n| \leq \frac{\varepsilon}{2} + \frac{\varepsilon}{2^2} + \ldots + \frac{\varepsilon}{2^n} < \varepsilon \qquad (1)$$

and for $m > n$

$$|\hat{x}_m - \hat{x}_n| = |x_{n+1} + \ldots + x_m| \overset{!}{\leq} |x_{n+1}| + \ldots + |x_m| \leq \frac{\varepsilon}{2^{n+1}} + \ldots + \frac{\varepsilon}{2^m} < \frac{\varepsilon}{2}.$$

The market inequality are true by the triangle inequality for metric and by the fact, that our metric is *translation invariant*: say,



$$\begin{aligned} |x+y| &= \varrho(0, x+y) \leq \varrho(0,x) + \varrho(x, x+y) \\ &= \varrho(0,x) + \varrho(0,y) = |x| + |y|. \end{aligned}$$

Hence the sequence $\{\hat{x}_n\}$ is a Cauchy sequence in X_ε, and therefore (by completeness of X) converges to some point $\hat{x}$, and we have

$$|\hat{x}| = \lim |\hat{x}_n| \overset{(1)}{\leq} \varepsilon$$

(since any metric is continuous function relative to generated by this metric topology), so that $\hat{x} \in X_\varepsilon$. By construction we have

$$|\hat{y} - A\hat{x}_n| = |\hat{y} - Ax_1 - Ax_2 - \ldots - Ax_n| < \delta_{n+1}, \quad n = 1, 2, \ldots,$$

so it holds

$$|\hat{y} - A\hat{x}| = \lim |\hat{y} - A\hat{x}_n| = 0,$$

that is, $\hat{y} = A\hat{x}$.

Thus, we have proved that the image of X_ε contains Y_{δ_1}.

3° In the case where A is one-to-one, it follows at once from 2°, that A^{-1} (which is, obviously, a *linear* mapping) is continuous at 0 and hence everywhere, so that A is a homeomorphism, and therefore A transforms open sets into open sets. In general case we argue as follows.

Let G be an arbitrary nonempty open set in X, $x \in G$ and U be a neighbourhood of 0 in X such that $x + U \subset G$. By 2° the image of U contains some neighbourhood V of 0 in Y. Then we have

$$AG \supset A(x + U) = Ax + AU \supset Ax + V.$$

Thus, AG contains some neighbourhood of each its point, that is AG is open. ▷

1.3.3 Banach theorem on inverse mapping

Here we derive a simple corollary of Banach theorem on open mapping (note, that this corollary was in fact proved in the process of proof of the mentioned theorem!):

Banach theorem on inverse mapping. *Any continuous linear bijection of F-spaces is a homeomorphism (that is the inverse mapping is (linear and) continuous).*

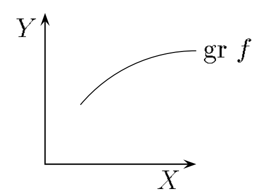
◁ By the Banach theorem on open mapping our mapping sends open sets into open sets. But this just means that pre-images of open sets by the inverse mapping are open sets. ▷

1.3.4 Theorem on closed graph

In conclusion of our discussion of the *openness* principle we derive from this principle the so-called theorem on *closed* graph.

Definition.

A mapping f from a topological space X into a topological space Y is said to be *closed*, if its *graph*



$$\mathrm{gr} f := \{(x, f(x)) \,|\, x \in X\}$$

is a closed subset of $X \times Y$.

Example. Every *continuous* function is closed (verify!).

Theorem on closed graph. *Every closed linear mapping from one F-space onto another such space is continuous.*

◁ 0°

1) Banach theorem on inverse mapping;

2) any closed linear subspace of an F-space is an F-space, and the product of any two F-spaces is an F-space (prove as an exercise!).

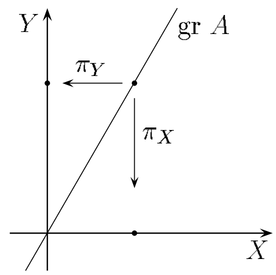
1° Let $A : X \to Y$ be a closed linear mapping of F-spaces. Ist graph $\mathrm{gr}A$ is a closed linear subspace in $X \times Y$ and hence, by 0°2), is an F-space. Denote by π_X and π_Y the projections of $X \times Y$ onto X and Y respectively. The mapping

$$\pi_X : (x, Ax) \longmapsto x, \ \mathrm{gr}A \longrightarrow X$$

is a continuous linear bijection, hence by 0°1) its inverse mapping π_X^{-1} is continuous. Therefore the mapping

$$A = \pi_Y \circ \pi_X^{-1}$$

is continuous.



▷

1.4 Boundedness principle

This is the third "whale". At first we discuss bounded sets and bounded operators in TVS.

1.4.1 Bounded sets

Bounded sets in TVS are the sets, which can be "absorbed" by any neighbourhood of 0:

Definition. A set A in a TVS is called *bounded* (the record: $A \in \mathrm{Bd}$) if for every neighbourhood of 0 U $\exists \delta > 0$ such that $\delta A \subset U$.

Examples.

1. Each point in a TVS is a bounded set. Indeed, every neighbourhood of 0 is absorbing.
2. A set A in $\mathbb{R}^\infty$ is bounded iff if is *coordinate -wise bounded*, that is, if for every $n \in \mathbb{N}$, the set $\{x_n \,|\, (x_1, \ldots, x_n, \ldots) \in A\}$ of all values of the n-th coordinate of its points is bounded in $\mathbb{R}$.

Elementary properties of bounded sets. Let X, Y be TVS, and let $A, B \subset X$. Then

a) $A \in \mathrm{Bd},\ B \subset A \Longrightarrow B \in \mathrm{Bd}$;

b) $A, B \in \mathrm{Bd}, \Longrightarrow A \cup B \in \mathrm{Bd}$;

c) $A, B \in \mathrm{Bd}, \Longrightarrow A + B \in \mathrm{Bd}$;

d) $A \in \mathrm{Bd},\ t \in \mathbb{R} \Longrightarrow tA \in \mathrm{Bd}$;

e) $A \in \mathrm{Bd} \Longrightarrow I_1 A \in \mathrm{Bd}$;

f) $A \in \mathrm{Bd} \Longrightarrow \bar{A} \in \mathrm{Bd}$;

g) $A \in \mathrm{Bd},\ f \in \mathcal{L}(X, Y) \Longrightarrow f(A) \in \mathrm{Bd}$;

h) if X is an LCS, then $A \in \mathrm{Bd} \Longrightarrow \mathrm{co}A \in \mathrm{Bd}$.

In orther words the property to be bounded is conserved by taking subsets, finite unions, linear combinations, closure, balanced hull and (in LCS) convex hull.

◁ This follows at once from the fact that each TVS has a base B, of neighbourhoods of 0, consisting from *closed balanced* sets such that $\forall U \in \mathrm{B}\, \exists V \in \mathrm{B} :\ V + V \subset U$, and from the fact that in each LCS we may choose the mentioned base B to consist from *convex* sets. ▷

According to intuition, *compact* sets are bounded:

Theorem on boundedness of compact sets. *In a TVS every compact set is bounded.*

$\lhd$ Let A be a compact set in a TVS X, and let U be a balanced neighbourhood of 0 in X. We need to show that there exists $\delta > 0$ such that $\delta A \subset U$. Since multiplication by scalar is continuous and since $0x = 0\ \forall x \in X$, for every $x \in A$, $\exists V_x \in Nb_x(X)\, \exists \delta_x > 0\ :\ I_{\delta_x} V_x \subset U$. These neighbourhoods V_x form a covering of A. By the definition of compactness we can choose a *finite* covering, say $V_{x_1}, \ldots, V_{x_n}$. Then $\delta := \min(\delta_{x_1}, \ldots, \delta_{x_n})$ will be our desired δ. Indeed,

$$\delta A \subset I_\delta A \subset I_\delta(V_{x_1} \cup \ldots \cup V_{x_n}) = I_\delta V_{x_1} \cup \ldots \cup I_\delta V_{x_n} \subset I_{\delta_{x_1}} V_{x_1} \cup \ldots \cup I_{\delta_{x_n}} V_{x_n} \subset U. \rhd$$

Remark. In finite-dimensional spaces it holds, roughly speaking, also the inverse assertion; namely, every bounded set is *relatively compact*, that is, has a compact closure:

$$A \in \text{RelComp} \ :\Longleftrightarrow \bar{A} \in \text{Comp}.$$

In infinite-dimensional case it is not so. E. g., in every infinite-dimensional normed space the closed unit ball is *not* compact (see p. 87).

(For l_2 it may be seen *directly* from the fact that $\|e_m - e_n\| = \sqrt{2}\ \forall m, n$.)

In LCS's boundedness of a set may be cheeked with the aid of *semi-norms*:

Characterization of boundedness in LCS's. Let X be an LCS and $\{p_\alpha\}$ be any generating system of semi-norms for X. Then a set A in X is bounded iff each *number* set $p_\alpha(A)$ is bounded (in $\mathbb{R}$).

$\lhd$ This follows at once from the relations between neighbourhoods of 0 and generating semi-norms in LCS. $\rhd$

Remark. Boundedness is in essence a sequential notion: it is sufficient to deal with bounded *sequences*, see the sequential characterization of boundedness below.

There is a close relation between bounded sequences and converging ones:

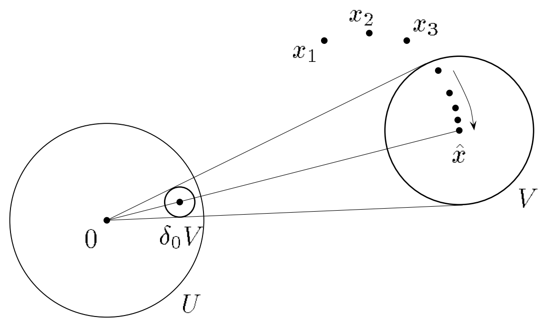
Lemma on bounded sequences and converging ones. *Let X be a TVS.*

a) Every converging sequence in X is bounded (that is the set of its points is bounded).

b) If a sequence $\{x_n\}$ in X is bounded, then $t_n x_n \to 0$ for each sequence $t_n \to 0$ in $\mathbb{R}$.

$\lhd$ a) This follows at once from the theorem on boundedness of compact sets. Here is a direct proof. Let $x_n \to \hat{x}$ and let U be a given *balanced* neighbourhood of 0 in X. Since $(t, x) \mapsto tx$ is continuous at $(0, \hat{x})$, there axist $\delta_0 > 0$ and a neighbourhood V of $\hat{x}$ such that $I_{\delta_0} V \subset U$. A fortiori we have $\delta_0 V \subset U$. Since $x_n \to \hat{x}$, $\exists n_0$ such that $x_n \in V$ for all $n \geq n_0$.

Further, since U is absorbing (as each neighbourhood of 0 is), $\exists \delta_i > 0$ such that $\delta_i x_i \in U$, $i = 1, \ldots, n_0 - 1$. Hence for $\delta = \min(\delta_0, \delta_1, \ldots, \delta_{n-1})$ we have $\delta x_i \in U$ for all i. Thus, $\{x_n\}$ is bounded.



b) Let $\{x_n\}$ be bounded and let $t_n \to 0$. Let us prove that $t_n x_n \to 0$. Let U be a given balanced neighbourhood of 0. Since $\{x_n\}$ is bounded, $\exists \delta > 0$ such that $\delta x_n \in U$ and hence $I_\delta x_n \subset U$ (by balancedness of U), $n = 1, 2, \ldots$. Since $t_n \to 0$, we have $t_n \in I_\delta$ for all sufficiently large n. Thus, for all such n it holds

$$t_n x_n \in I_\delta x_n \subset U,$$

that is $t_n x_n \to 0$. $\triangleright$

Remark. Assertion b) admits the inversion, as it follows from the following **Sequential characterization of boundedness.** A set in a TVS is bounded iff for every seguence $\{x_n\}$ of its points and every sequence $\{t_n\}$ of real numbers such that $t_n \to 0$ we have $t_n x_n \to 0$.

$\triangleleft$ An exercise for you. $\triangleright$

Corollary. In a (semi-) normed space a set A is bounded iff it is *bounded in the (semi-) norm*, that is, iff for some $c \geq 0$ we have $\|x\| \leq c \ \forall x \in A$.

Remark. In F-space boundedness *in the metric doesn't* imply boundedness! Indeed, balls (with the center at 0) in the metric, which generates the topology. are neighbourhoods of 0, and "as a rule" neighbourhoods of 0 are not bounded: if in a Hausdorff TVS there exists a bounded convex neighbourhood of 0 then this space is *normable* (see Kolmogorov's characterization of normable spaces in Chapter 3).

1.4.2 Bounded operators

For linear mapping "boundedness" means, by the definition, boundedness on bounded sets:
Definition. A linear mapping $A : X \to Y$, where X and Y are TVS's, is called *bounded*, if the image of every bounded set is a bounded set.

Example. If X, Y are normed spaces, then for a linear mapping $A : X \to Y$ the following conditions are equivalent:

a) A is continuous;

b) A is bounded;

c) A is bounded on the unit ball, that is, $\|x\| \leq 1 \Rightarrow \|Ax\| \leq c$ for some $c \geq 0$.

Theorem on boundedness and continuity *Let X, Y be TVS's and $A : X \to Y$ be a linear mapping.*

a) A is continuous $\Rightarrow$ A is bounded.

b) If X is a metrizable TVS, then A is bounded $\Rightarrow$ A is continuous.

◁ [a)] Let A is continuous, and let B be a bounded set in X. Let us prove that AB is bounded in Y. Let V be a given neighbourhood of 0 in Y. By continuity of A exists a neighbourhood U of 0 in X such that $AU \subset V$. By boundedness of B $\exists \delta > 0$ such that $\delta B \subset U$. Then

$$\delta(AB) = A(\delta B) \subset AU \subset V,$$

that is, AB is bounded.

[b)] Let X be a metrizable TVS, and let A be bounded. Suppose that A is not continuous. Then there exist $V \in \mathrm{Nb}_0(Y)$ such that for every $U \in \mathrm{Nb}_0(x)$ we have

$$AU \not\subset V.$$

If we take $U = \frac{1}{n} X_{\frac{1}{n}}$, we obtain:

$$\forall n \in \mathbb{N} \vdots A \frac{1}{n} X_{\frac{1}{n}} \not\subset V$$

(where X_ε is defined as on p. 33),

$$\forall n \in \mathbb{N} \exists x_n \in \frac{1}{n} X_{\frac{1}{n}} : Ax_n \notin V. \tag{1}$$

It is clear that $nx_n \to 0$ in X, and hence, the sequence $\{nx_n\}$ is bounded, by assertion a) of Lemma on p.37. Since A is bounded, it follows that the sequence $\{Anx_n\}$ is also bounded. So, by assertion b) of the some Lemma, we must have

$$\frac{1}{n}(Anx_n) \longrightarrow 0. \tag{2}$$

But $\frac{1}{n}(Anx_n) = Ax_n$ belongs to V for *no* n (by (1)), so (2) is impossible. The obtained contradiction shows that A is continuous. ▷

Remarks. Assertion b) remains true for so-called *bornological* LCS's X (and any LCS's Y). (See [Edw, p.652].)

That not every bounded linear mapping of TVS's is continuous, is seen from the following **Example.** Let, on a given vector space, there exist two different linear topologies with the same bounded sets (in the next chapter we shall encounter such spaces). Then the identity mapping from one of this TVS's onto another one will be bounded, but noncontinuous. (Thus, the "first" TVS cannot be metrizable.)

1.4.3 Equicontinuity and equiboundedness

In order to formulate the boundedness principle we need some notions.
Definition. Let X, Y be TVS's. A family of continuous linear mapping $A_\alpha : X \to Y$, $\alpha \in \mathcal{A}$, is called:

- *equicontinuous*, if for every neighbourhood V of 0 in Y exists a neighbourhood U of 0 in X such that $A_\alpha U \subset V$ $\forall \alpha \in \mathcal{A}$;

- *equibounded* (or *uniformly bounded*), if for every bounded set B in X exists a bounded set C in Y such that $A_\alpha B \subset C$ $\forall \alpha \in \mathcal{A}$;

- *pointwise bounded*, if for every point $x \in X$ the set $\{A_\alpha x \big| \alpha \in A\}$ is bounded in Y.

Remarks.

1. Thus we deal, in the above definition, with aquicontinuity *at* 0, but for linear mappings continuity at 0 and continuity are equivalent.

2. It holds

equi-continuity	$\Longrightarrow$	equi-boundedness	$\Longrightarrow$	pointwise boundedness

The secont implication is evident, and the first one may be proved just as one proves that continuity $\Rightarrow$ boundedness (for linear mappings).

Example. In the case of normed spaces X and Y the following conditions on a family $\{A_\alpha\}_{\alpha\in\mathcal{A}}$, $A_\alpha \in \mathcal{L}(X,Y)$, are equivalent (see Chapter 3):

a) $\{A_\alpha\}$ is equicontinuous;

a) $\{A_\alpha\}$ is equibounded;

a) norms of A_α, $\alpha\in\mathcal{A}$, are bounded from above: $\sup\limits_{\alpha\in\mathcal{A}} \|A_\alpha\| < \infty$.

1.4.4 Boundedness principle: Banach-Steinhaus theorem

There is a number of versions of the boundedness principle. One of them is the following
Banach-Steinhaus theorem. *Let X and Y be F-spaces. If a family of continuous linear mappings from X into Y is pointwise bounded, then it is equicontinuous.*

Sometimes results of this type are called also "*equicontinuity principles*".

$\lhd$ 0° Baire theorem.

1° Again we shall use notation $|a| : \varrho(0,a)$, where ϱ is a generating invariant metric of a given F-space. Emphasize, that $|\cdot|$ is *not* homogeneous ($|ta| \neq |t|\,|a|$!), but *is symmetric* ($|-x| = |x|$), since $\varrho(0,-x) = \varrho(x,0) = \varrho(0,x)$.

2° Let a family $\{A_\alpha\}_{\alpha\in\mathcal{A}}$, $A_\alpha \in \mathcal{L}(X,Y)$, is pointwise bounded. We need to show that for an arbitrary given $\varepsilon > 0$ exists a neighbourhood U of 0 in X such that

$$\forall\alpha\in\mathcal{A} \;\vdots\; A_\alpha U \subset Y_\varepsilon. \tag{1}$$

3° Put for $k = 1, 2, \ldots$

$$C_k := \left\{ x \in X \,\middle|\, \forall\alpha\in\mathcal{A} \;\vdots\; \left|\frac{1}{k}A_\alpha x\right| \le \frac{\varepsilon}{2} \right\} \overset{\text{Obv.}}{=} \bigcap_{\alpha\in\mathcal{A}} \left(\frac{1}{k}A_\alpha\right)^{-1}\left(Y_{\varepsilon/2}\right).$$

This set is *closed*, since $Y_{\varepsilon/2}$ is, and since all operators $\frac{1}{k}A_\alpha$ are continuous.

4° Each $x \in X$ lies in some C_k. Indeed, by the condition, the set $\{A_\alpha x \,|\, \alpha\in\mathcal{A}\}$ is bounded, so for some (sufficiently big) k it holds

$$\frac{1}{k}\{A_\alpha x \,|\, \alpha\in\mathcal{A}\} \subset Y_{\varepsilon/2},$$

which means

$$\forall\alpha\in\mathcal{A} \;\vdots\; \left|\frac{1}{k}A_\alpha x\right| \le \frac{\varepsilon}{2},$$

that is, $x \in C_k$. So

$$\bigcup_{k=1}^{\infty} C_k = X.$$

5° By 0°, at least one C_k is dense in some ball, hence (since C_k is closed) C_k contains some ball. Thus,

$$\exists k \in \mathbb{N}\ \exists V \in \mathrm{Nb}_0\ \exists x \in X \,\vdots\, x + V \subset C_k. \tag{2}$$

In particular, $x \in C_k$. It follows that

$$-x \in C_k. \tag{3}$$

Indeed,

$$\left|\frac{1}{k}A_\alpha(-x)\right| = \left|-\frac{1}{k}A_\alpha x\right| \overset{1^\circ}{=} \left|\frac{1}{k}A_\alpha x\right| \overset{x\in C_k}{\leq} \frac{\varepsilon}{2}.$$

6° We claim that for $U := \frac{1}{k}V$ eq. (1) is true. Indeed, for any $h \in V$ we have

$$\left|A_\alpha\left(\frac{1}{k}h\right)\right| \overset{\substack{\text{linearity}\\ \text{of } A_\alpha}}{=} \left|\frac{1}{k}A_\alpha(x+h) + \frac{1}{k}A_\alpha(-x)\right| \overset{\substack{\text{linearity}\\ \text{inequality}}}{\leq} \underbrace{\left|\frac{1}{k}A_\alpha(x+h)\right|}_{\overset{(2)}{\leq}\frac{\varepsilon}{2}} + \underbrace{\left|\frac{1}{k}A_\alpha(-x)\right|}_{\overset{(3)}{\leq}\frac{\varepsilon}{2}} \leq \varepsilon. \ \triangleright$$

Remarks.

1. The completeness of Y *was not used* in the proof and *may be omitted* in the assumptions of the theorem; but the completeness of X is an *essential* condition (see the next subsection).

2. The assertion of the theorem remains true, if

   1) X is "ultra-barrel" and Y is an arbitrary TVS, or if

   2) X is "barrel" and Y is an LCS (see [4, p. 636]).

3. For Banach spaces the boundedness principle takes the form of the so-called "principle of fixation of singularities" (see Chapter 3).

1.4.5 A counter-example

Let us give an example, showing that completeness of the "first" space is an essential condition for valadity of the boundedness principle.

Let k denote the space of all *finitary* real sequences, that is, of sequences $x = (x_1, x_2, \ldots)$ with only finite number of nonzero members (this number may be different for different sequences):

$$x = (\,x_1, x_2, \ldots, x_r, 0, 0, \ldots).$$

Equip k by the supremum-norm

$$\|x\| := \sup_n |x_n|$$

(this supremum is always finite, since we have only finite number of nonzero x_n). For $n = 1, 2, \ldots$ define a functional f_n on k by the formula

$$f_n(x) = nx_n.$$

It is clear that for each n this functional f_n is a continuous linear functional on k with the norm (see Chapter 3)

$$\|f_n\| = n.$$

So the sequence $\{f_n\}$ is *not* equicontinuous (see Example on p. 40). But this sequence *is* pointwise bounded: for every fixed point $x = (x_1, \dots, x_r, 0, 0, \dots)$ the set

$$\{f_n(x) \big| n = 1, 2, \dots\} = \{x_1, 2x_2, \dots, rx_r, 0\} \quad (\subset \mathbb{R})$$

is a *finite* set and therefore is a bounded set in $\mathbb{R}$.

1.4.6 "Pointwise completeness" of $\mathcal{L}(X,Y)$

As an application of the boundedness principle we prowe here that a *pointwise* limit of a *sequence* of continuous linear operators, acting from one F-space into another, is again a continuous linear operator.

Theorem on "pointwise completeness". *Let X, Y be F-spaces, and let $A_n \in \mathcal{L}(X,Y)$, $n = 1, 2, \dots$. Let for every $x \in X$ the sequence $A_n x$ converges in Y. Denote the limit by Ax*

$$Ax := \lim A_n x.$$

The the so defined operator A is linear and continuous:

$$A \in \mathcal{L}(X,Y).$$

$\lhd$ 0°

1. Boundedness principle;
2. continuity of arithmetic operations in TVS;
3. boundedness of convergent sequences.

1° *Linearity.* For all $t \in \mathbb{R}$, $x, x_1', x_2'' \in X$ we have

$$A(tx) = \lim A_n(tx) = \lim tA_n x \overset{!}{=} t \lim A_n x = tAx,$$

$$A(x' + x'') = \lim A_n(x' + x'') = \lim(A_n x' + A_n x'') \overset{!}{=} \lim A_n x' + \lim A_n x'' = Ax' + Ax'',$$

the marked equalities being true by continuity of multiplication by a scalar and of addition, resp. (in Y).

2° *Continuity.* For every $x \in X$ the sequence $\{A_n x\}$ is bounded by 0°3). This means that the sequence $\{A_n\}$ is pointwise bounded. Then by 0°1) it is equi-continuous, that is,

$$\forall \varepsilon > 0 \ \exists \delta > 0 \ : \ |x| \le \delta \implies |A_n x| \le \varepsilon \quad \forall n.$$

(Here $|\cdot|$ denotes the distance from 0 in the generating invariant metric.) Since $A_n x \to Ax$ we conclude by continuity of the function $|\cdot|$ that $|Ax| \le \varepsilon$ if $|x| \le \delta$. Thus A is continuous. $\rhd$

Remark. The completeness of Y is again a superfluous condition (see Remark 1 on p. 41), but the completeness of X is an essential condition as the following example shows.

Example. Let k be the (noncompete!) normed space of the *finite* sequences from Subsection 1.4.5. Put for $x = (x_1, x_2, x_3, \dots) \in k$

$$\langle x_n^*, x \rangle := x_1 + 2x_2 + 3x_3 + \dots + nx_n.$$

Clearly $x_n^* \in \mathcal{L}(k, \mathbb{R}) = k^*$ for all $n = 1, 2, \dots$, and $\langle x_n^*, x \rangle \to x_1 + 2x_2 + 3x_3 + \dots$ for every $x \in k$. The limit (linear) functional $(x_1, x_2, \dots) \mapsto x_1 + 2x_2 + \dots$ is *not* continuous, since it is not bounded on the unit ball in k.

2 Duality theory

In this chapter we discuss the notions of the topological dual space and of a topology, compatible with a given duality. In this connection we introduce so-called weak and weakened topology. At last we prove some results of convex analysis in LCS.

2.1 Topological dual space

Here we prove that any Hausdorff LCS and its topological dual space form a dual pair, and introduce a notion of topology, compatible with a given duality.

2.1.1 Definition and examples

We have already dealed with the space X^*. Now we give to it a name

Definition. Let X be a TVS. The set

$$X^* := \mathcal{L}(X, \mathbb{R})$$

of all continuous linear functionals on X is called the *topological dual* of X. [The set of *all* linear functionals on X is called the *algebraical dual* of X and denoted by X'.]

[In some books one used, vice vesra, the symbol X' for the topological dual space and the symbol X^* for the algebraical one.]

It is immediately verified that for any TVS's X and Y the space $\mathcal{L}(X, Y)$ of all continuous linear mappings from X into Y is a *vector space* with respect to natural operations of addition and multiplication by a scalar

$$\begin{aligned} (f_1 + f_2)(x) &:= f_1(x) + f_2(x), \\ (tf)(x) &:= t(f(x)), \end{aligned}$$

the *null* element of this vector space being the null functional.

We shall usually denote elements of X^* by x^* and write the value of a functional x^* on an element x in the *symmetric* form $\langle x^*, x \rangle$:

$$\langle x^*, x \rangle := x^*(x).$$

It is clear that $\langle \cdot, \cdot \rangle$ is a *bilinear form* on $X^* \times X$:

$$\langle \cdot, \cdot \rangle : X^* \times X \longrightarrow \mathbb{R}, \ (x^*, x) \longmapsto \langle x^*, x \rangle.$$

This form is called the *canonical pairing* of X and X^*.

Examples.

1. For $\mathbb{R}^n$ with the usual topology we have

$$(\mathbb{R}^n)^* = (\mathbb{R}^n)' = \mathbb{R}^n,$$

$$\langle x^*, x \rangle = \sum_{i=1}^{n} x_i^* x_i \quad (x = (x_1, \dots, x_n), x^* = (x_1^*, \dots, x_n^*)).$$

The canonical pairing is here just the inner product.

2. The *kernel-convex topology* τ_{kc} in a vector space X is the linear topology, for which a base of neighbourhoods of 0 consists of all absorbing balanced convex sets. This topology τ_{kc} is Hausdorff, and is the strongest (=finest) locally convex topology on X (verify!); each linear functional on X is continuous relative τ_{kc}, that is,

$$(X, \tau_{kc})^* = X'$$

(Verify!).

3. For the space $l_2$we have

$$l_2^* \approx l_2$$

in the sense, which will be explained in Chapter 4.

2.1.2 Duality between an LCS and its topological dual space

For an arbitrary TVS's the topological dual may be trivial ($X^* = \{0\}$). But for *Hausdorff LCS's* the topological duals are ever sufficiently rich: X^* and X form a dual pair relative to the canonical pairing, in the following sense:

Definition. Vector spaces X and Y are said to form a *dual pair* relative to a given bilinear mapping $\langle \cdot, \cdot \rangle : X \times Y \to \mathbb{R}$ (which is called a *pairing*), if the following conditions of "totality" are fulfilled:

1) *totality of* X:

$$\langle x, y \rangle = 0 \ \forall x \in X \Longrightarrow y = 0, \text{ or, equivalently, } y \neq 0 \Longrightarrow \exists x \in X : \langle x, y \rangle \neq 0;$$

item[2)] *totality of* Y:

$$\langle x, y \rangle = 0 \ \forall y \in Y \Longrightarrow x = 0, \text{ or, equivalently, } x \neq 0 \Longrightarrow \exists y \in Y : \langle x, y \rangle \neq 0.$$

We write in this case

$$X \overset{\langle\ ,\ \rangle}{\longleftrightarrow} Y.$$

Each element $y \in Y$ defines an element $\bar{y} \in X'$ by the formula

$$\bar{y}(x) := \langle x, y \rangle,$$

and by totality of X the mapping $y \mapsto \bar{y}$ is injective. So we identify y and $\bar{y}$ and assume that $Y \subset X'$, and analogously, $X \subset Y'$.

Theorem on duality of an LCS and its topological dual space. *Let X be a Hausdorff LCS. Then X and X^* form a dual pair relarive to the canonical pairing.*

$\lhd$ The totality of X holds just by definition of null functional. The totality of X^* is the content of the theorem on totality of the dual space. (p. 32). $\rhd$

In view if this theorem we consider *everywhere below* only Hausdorff locally convex topologies, if there is no special remark.

2.1.3 Topologies compatible with a given duality

In general on a given LCS there are many other locally convex topologies besides its original one, for which the topological dual space is the same. This justifies the following

Definition. Let X, Y be a dual pair. We say that a (linear) topology τ on X is compatible with the duality between X and Y, if

$$(X, \tau)^* = Y.$$

(Recall that we may assume $Y \subset X'$.)

Examples.

1. If X is a Hausdorff LCS then its topology is of course compatible with the duality $X \overset{\langle , \rangle}{\longleftrightarrow} X^*$.

2. For every vector space X the kernel-convex topology τ_{kc} in X is compatible with the duality $X \overset{\langle\,,\,\rangle}{\longleftrightarrow} X'$.

2.2 Weak and weakened topologies

In the last example τ_{kc} is the *strongest* locally convex topology compatible with the duality in question. It appeares that for a *each* dual pair X, Y there exist both the strongest and the weakest locally convex topologies on X among all locally convex topologies on X, which are compatible with the duality.

2.2.1 Definitions

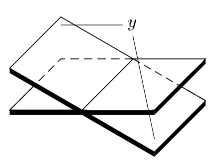
The existance of the former topology (which is called the *Mackey topology*) is a rather fine fact and we shall not discuss it; the existance of the latter (the weakest) topology we shall prove; it is the so-called "weak" topology:

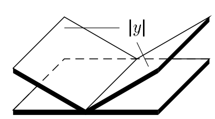
Definition. Let X, Y be a dual pair. The locally convex topology on X, generated by the system of all semi-norms of the norm $|y|\,,\ y \in Y$:

$$|y|\,(x) := |y(x)| = |\langle x, y\rangle|$$

is called the *weak topology* in X, defined by Y, and is denoted by

$$\sigma(X, Y).$$





(Notice that the weak topology is *Hausdorff*, since for every $x \neq 0$ exists by totality of Y an element y such that $\langle x, y\rangle \neq 0$, so that for this y we have $|y|\,(x) \neq 0$ and hence $x \notin \varepsilon \mathring{B}B_{|y|}$ for some $\varepsilon > 0$.)

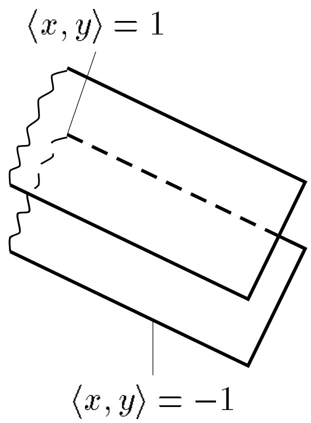
Exercises.

1. Show that the sets of the form $\left\{x \,\middle|\, \langle x, y\rangle < \alpha\right\}$, $y \in Y$, $\alpha \in \mathbb{R}$ ("open half-spaces") are open in $\sigma(X, Y)$, and the sets of the form $\left\{x \,\middle|\, \langle x, y\rangle \leq \alpha\right\}$, $y \in Y$, $\alpha \in \mathbb{R}$ ("closed half-spaces") are closed in $\sigma(X, Y)$.

2. Show that the "unit layers"

$$\left\{x \,\middle|\, \langle x, y\rangle \leq 1\right\},\ y \in Y,$$

and their finite intersections form a basic of balanced convex closed neighbourhoods of 0 for $\sigma(X, Y)$ (and it holds the analogous assertion with "closed" replaced by "open").



3. Show that $\sigma(X,Y)$ is the *weakest* linear topology in X such that all functionals $y \in Y$ ($\subset X'$) are continuous.
4. Show that $\sigma(\mathbb{R}^n, \mathbb{R}^n)$ is the usual topology in $\mathbb{R}^n$.
5. Show that $x_n \to x$ in $\sigma(X,Y) \Leftrightarrow \forall y \in Y \dot{:} \langle x_n, y\rangle \to \langle x, y\rangle$. (If all "test devices" say that a sequence converges, then this sequence really converges.)
6. Show that the bounded sets in $\sigma(X,Y)$ are exactly the sets A such that for every $y \in Y$ the set $\langle A, y\rangle$ is bounded in $\mathbb{R}$:

$$\forall y \in Y \dot{:} \sup |\langle A, y\rangle| < +\infty$$

(that is, the image $y(A)$ of A by y is bounded in $\mathbb{R}$). [*Hint*: use the characterization of boundedness in LCS's given on p. 25.]

Now return to our dual pair $X \overset{\langle\,,\,\rangle}{\longleftrightarrow} X^*$.

Definition. If X is an LCS then the topology $\sigma(X, X^*)$ is called the *weakened* topology in X; $\sigma(X, X^*)$ - open sets are called *weakly open* sets, and the word *weakly* is used in analogous manner for another properties related to the weakened topology. The topology $\sigma(X^*, X)$ is often called the weak* topology in X^* (one pronounces "weak star topology").

The name "weakened" is justified by the fact that the topology $\sigma(X, X^*)$ is always weaker than the original topology of X (verify!).

2.2.2 Compatibility of the weak topology with the duality

Here is the mentioned above result on the weak topology:

Theorem on minimum property of the weak topology. *Let $X \to Y$. Then $\sigma(X,Y)$ is the weakest:*

1) among all locally convex topologies τ on X such that each functional $\langle \cdot, y\rangle$, $y \in Y$, is continuous;

2) among all locally convex topologies τ on X, that are compatible with the duality.

In other words,

$$Y \subset (X, \tau)^* \Longrightarrow \tau \supset \sigma(X,Y),$$

$$(\, X, \sigma(X,Y)\,)^* = Y.$$

◁ 0Þ

1) The theorem on generating by semi-norms;
2) the fact that a linear functional bounded on the whole space is equal to 0;

3) **Lemma on linear dependence of functionals.** *Let X be a vector space and $f, f_1, \ldots, f_n$ be linear functionals on X. Then*

$$f \in lin\{f_1, \ldots, f_n\} \Longleftrightarrow ker f \supset \bigcap_{i=1}^{n} ker f_i. \tag{1}$$

Recall that $\operatorname{lin} A$ denote the linear hull of A, that is the set of all linear combinations of elements of A, and that

$$\ker f := \{x \,|\, f(x) = 0\}.$$

Let us prove this result of linear algebra:

$\lhd\lhd$ "$\Longrightarrow$" Let $f = \sum_{i=1}^{n} \lambda_i f_i$ and let $x \in \bigcap_{i=1}^{n} \ker f_i$, that is $f_i(x) = 0$ for $i = 1, \ldots, n$. Then $f(x) = \sum \lambda_i f_i(x) = 0$, that is $x \in \ker f$.

"$\Longleftarrow$" Let

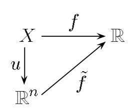
$$\ker f \supset \bigcap \ker f_i. \tag{2}$$

Consider the linear mapping

$$U : X \longrightarrow \mathbb{R}^n, \quad x \longmapsto (\, f_1(x), \ldots, f_n(x)\,)$$

and define on its image $\operatorname{im} u$ a linear functional $\tilde{f}$ by the formula

$$\tilde{f}(\, u(x)\,) := f(x).$$



This definition is correct by (2). [Indeed, if $u(x_1) = u(x_2)$, then $u(x_1 - x_2) = 0$, that is, $x_1 - x_2 \in \bigcap \ker f_i$, and hence, by (2), $x_1 - x_2 \in her f$, that is, $f(x_1) = f(x_2)$.] Now we extend $\tilde{f}$ *anyhow* to obtain a linear functional on the whole $\mathbb{R}^n$; as each linear functional on $\mathbb{R}^n$ this extension has the form

$$\tilde{f} : (\, \xi_1, \ldots, \xi_n\,) \longmapsto \sum_{i=1}^{n} \lambda_i \xi_i.$$

Hence, for every $x \in X$ we have

$$f(x) = \tilde{f}(\, u(x)\,) = \tilde{f}(\, f_1(x), \ldots, f_n(x)\,) = \sum_{i=1}^{n} \lambda_i f_i(x),$$

that is,

$$f = \sum_{i=1}^{n} \lambda_i f_i. \;\rhd\rhd$$

1° $(X, \tau)^* \supset Y \Rightarrow \tau \supset \sigma(X, Y)$ (this implication is even stronger than the right hand one in 1).
$\lhd\lhd$ If a locally convex topology τ in X is such that $(X; \tau)^* \supset Y$, then each functional y in Y (recall that we may assume $Y \subset X'$!) is continuous; hence each semi-norm $|y|$, $y \in Y$, is (as the composition of two continuous functions) continuous. This means, that $\sigma(X, Y)$ is generated by a *part* of the family of all continuous semi-norms on (X, τ). Since by 0°1) τ is generated just by the whole family, we conclude that $\tau \geq \sigma(X, Y)$. $\rhd\rhd$

2° $(X, \sigma(X,Y))^* \supset Y$. ◁◁ By the very definition of $\sigma(X,Y)$ each element $y \in Y$ is a *continuous* linear functional on $(X, \sigma(X,Y))$. ▷▷
3° $(X, \sigma(X,Y))^* \subset Y$. ◁◁ Let $l \in (X, \sigma(X,Y))^*$, that is, l is a linear functional on X, which is continuous with respect to $\sigma(X,Y)$. Then l is bounded on some neighbourhood of 0 in $\sigma(X,Y)$, that is, $\exists U \in \mathrm{Nb}_0(\sigma(X,Y))$:

$$l \text{ is bounded on } U. \tag{3}$$

Since each neighbourhood of 0 in $\sigma(X,Y)$ contains some finite intersection of "unit layers", $\exists y_1, \dots, y_n \in Y$ such that

$$U \supset \bigcap_{i=1}^{n} \{x \mid |\langle x, y_i \rangle| \le 1\}. \tag{4}$$

We claim that

$$\ker l \supset \bigcap_{i=1}^{n} \ker y_i. \tag{5}$$

Indeed,

$$\bigcap_{i=1}^{n} \ker y_i \overset{\text{obv.}}{\subset} \bigcap_{i=1}^{n} \{x \mid |\langle x, y_i \rangle| \le 1\} \overset{(4)}{\subset} U,$$

hence, by (3), l is bounded on $\bigcap_{i=1}^{n} \ker y_i$. By 0°2), we conclude that l is equal to 0 on this intersection, that is, (5) is true. It follows, by 0°3), that l is a linear combination of $y_1, \dots, y_n$, and therefore is an element of Y. ▷▷ ▷

2.2.3 Some common properties of all topologies which are compatible with a given duality

At first we shall prove a result on common properties of the weakened topology in an LCS and of the original one:

Theorem on the weakened topology. *Let X be an LCS. Then X with its original topology and X equipped with the weakened topology $\sigma(X, X^*)$, have the same:*

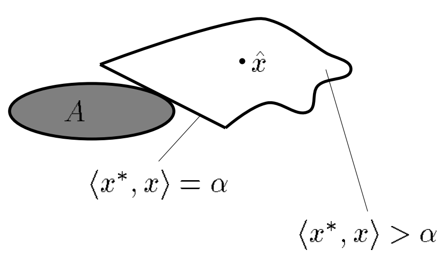
a) convex closed sets;

b) bounded sets.

◁ 0° Strict separation theorem.
1° a) Let A be a convex closed set in X. We have to show that A is weakly closed. Let $\hat{x} \notin A$. Then, by 0°, $\exists x^* \in X^*$ such that

$$\alpha := \sup \langle x^*, A \rangle < \langle x^*, \hat{x} \rangle.$$

So the half-space $\{x \mid \langle x^*, x \rangle > \alpha\}$ is a $\sigma(X, X^*)$-neighbourhood of $\hat{x}$ which contains no point of A. Thus A is $\sigma(X, X^*)$-closed.



b) We have to prove that every weakly bounded set is bounded in the original topology. We shall show this fact *later* (see Chapter 3) and *only* for (the most important) case of normed spaces. The proof for the general case may be found in [Edw, p.687]. $\rhd$

Corollary. Let X, Y be a dual pair. Then all locally convex topologies in X, which are compatible with this duality, have the same

a) convex closed sets;

b) bounded sets.

$\lhd$ By the theorem these sets are the same ones as for $\sigma(X, Y)$. $\rhd$

Remark. The weaker is a linear topology on a vector space the *more* it has of bounded sets and the *less* it has of closed sets. So a priory in the weakened topology there are more of bounded sets and there are less of closed sets than in the original one. And in fact, the original topology has *more* of closed sets! Only the closed *convex* sets are the same.

Exercise. Show that in l_2 the unit *sphere*

$$S_1 := \left\{x \,\middle|\, \|x\| = 1\right\}$$

is closed, but is *not* closed in the weakened topology (which is usually called simply "the *weak* topology in l_2"). Show, further, that weak closure of S_1 (that is, the closure in the weakened topology) is the unit *ball* B_1:

$$\overline{S_1}^{\text{Weak}} = B_1.$$

2.3 Elements of convex analysis in LCS's

The convex analysis deals with operations on convex functions and convex sets. A very fruitful idea is to consider, side by side with objects in a given space, "dual" objects in a space that is in duality with the original one.

2.3.1 Operators of convex analysis

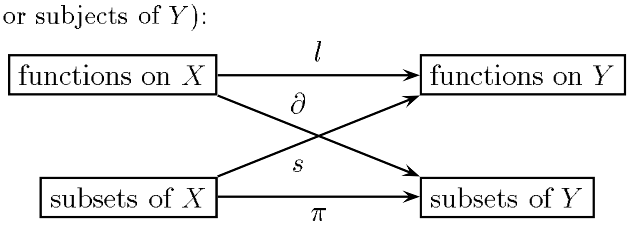
Let $X \overset{\langle,\rangle}{\longleftrightarrow} Y$ be a dual pair. Each elemant X in (say) X has a dual nature: on the one side, it is an *element* of X, and on the other side, it may be treated as a *function* $\langle x, \cdot\rangle$ on Y (recall we may assume that $X \subset Y'$).

So with each point in X we can bring into correspondence a "dual" object, viz. a (linear) function on Y. V. v., with each $\sigma(X, Y)$-continuous linear function on X (that is, continuous with respect to $\sigma(X, Y)$), we can bring into correspondence a point in Y $\big($since $\big(X, \sigma(X, Y)\big)^* = Y\big)$.

This transition to dual objects may be largely extended. The operators

$$l,\ \partial,\ s, \pi,$$

which are introduced below, bring object of X (functions on X or subjects of X) into "dual" object of Y (functions on Y or subjects of Y):



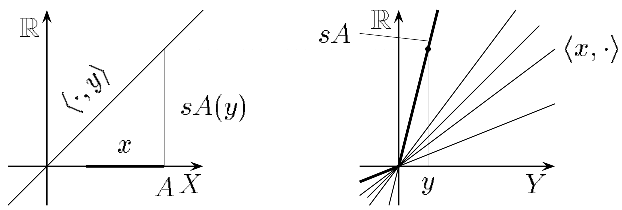
Operator s

This operator is a natural extension of the mapping

$$x \longmapsto \langle x, \cdot\rangle$$

which sends a *point* of X into a (linear) *function* on Y. Viz., for any *set* $A \subset X$ we define its *support function* $sA : Y \to \mathbb{R}$ as the supremum of all linear functions corresponding to the points of A:

$$sA := \bigvee_{x \in A} \langle x, \cdot \rangle. \tag{1}$$



In other words

$$sA(y) = \sup \langle A, y \rangle = \sup_A \langle \cdot, y \rangle. \tag{2}$$

Emphasize that in (1) we take the supremum of a *set* of (linear) functions $\langle x, \cdot \rangle$ *on* Y (indeed by $x \in A$), and in (2), dually, we take the supremum of a *fixed* (linear) function $\langle \cdot, y \rangle$ *on* X (over A).

Remarks.

1. The name "support" is somewhat misleading.

2. By the definition,

$$\sup 0 := -\infty \tag{3}$$

Notice that $\inf 0 := +\infty$, so that $\sup 0 < \inf 0$! The point is that for a "natural" extension of "$\le$" from the reals to the sets of reals we have $0 < \mathbb{R} < 0$, so that $\sup 0 = \inf \mathbb{R} = -\infty$, and $\inf 0 = \sup \mathbb{R} = +\infty$. This just means that such a natural extension is *not* an order relation.

Examples.

1. $\forall x \in X \,\vdots\, s\{x\} = \langle x, \cdot \rangle$.

2. $sX = \delta\{0\}$.

Exercises.

1. For the canonical duality $\mathbb{R}^n \leftrightarrow \mathbb{R}^n$

$$sB_1 = \|\cdot\|, \tag{4}$$

where $\|x\|^2 := x_1^2 + \ldots + x_n^2$ ($x = (x_1, \ldots, x_n)$), and B_1 is the closed unit ball.

2. For any non-empty $A_1, A_2 \subset X$

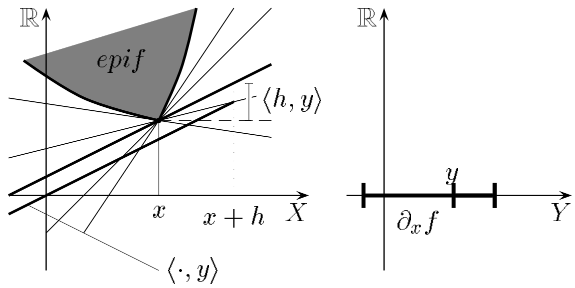
$$s(A_1 + A_2) = s(A_1) + s(A_2). \tag{5}$$

Operator ∂

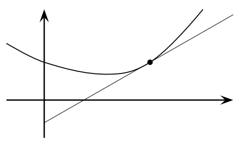
If $f = sA$, then, obviously, for any $x \in A$ it holds $f \ge \langle x, \cdot \rangle$. So if we want to reconstruct A, starting from sA, we have to consider all linear functions $\langle x, \cdot \rangle$ on Y whose graphs lie below the graph of sA, and A is just the set of the corresponding points x. This idea is the basic of the following definition.

The *subdifferential* $\partial_x f$ (the "∂" comes from "differential") of a function $f : X \to \bar{\mathbb{R}}$ at a point $x \in X$, where $f(x) \in \mathbb{R}$, is a subset of Y given by the formula

$$\partial_x f := \left\{ y \in Y \,\middle|\, \forall h \in X \,\vdots\, f(x+h) \ge f(x) + \langle h, y \rangle \right\}. \tag{6}$$



Thus, the subdifferential of f at a point x is the set of "slopes" of all "subtangent lines" to the (epi)graph of f at the point $(x, f(x))$. If such a subtangent line is unique, that is, if there exists the *tangent* line to the graph, we obtain the classical "differential" (see the pictuse to the left).



For *sublinear* function it is, of course, the subdifferential *at 0* that plays key role. This subdifferential is denoted simply by ∂:

$$\partial := \partial_0.$$

In explicit form, for $p \in \mathrm{Sublin}(X)$

$$\partial p := \{y \in Y \,|\, p \geq \langle \cdot, y \rangle\}. \tag{7}$$

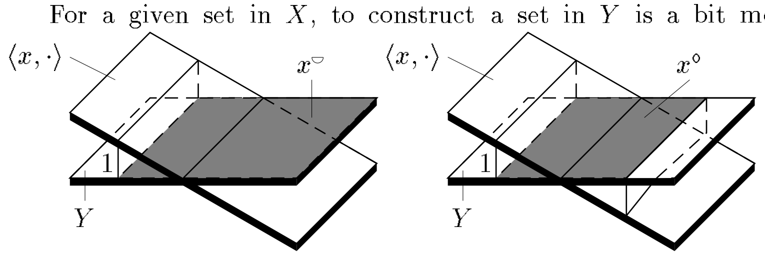
(The corresponding picture is just as on p. 50.)

Exercises.

1. $\forall y \in Y \vdots\, \partial\langle \cdot, y\rangle = \{y\}$ (the "derivative" of a linear function at each point coincides with this function itself).
2. For $\mathbb{R}^n \leftrightarrow \mathbb{R}^n$
$$\partial \|\cdot\| = B_1 \tag{8}$$
(notations as in Exercise 1 on p. 50).

Operator π

For a given set in X, to construct a set in Y is a bit more complicated. And there are here (at least) two different variants, viz. "balanced" and "non-balanced".



First of all, for a given *point* $x \in X$ it is natural to take as dual sets the following two sets in Y:

$$x^{\triangledown} := \{y \,|\, \langle x, y\rangle \leq 1\} \equiv \{x \leq 1\}, \tag{9}$$

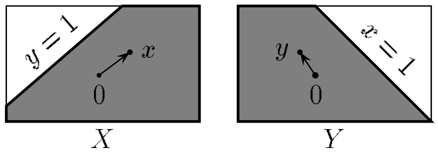
$$x^{o} := \{y \,|\, |\langle x, y\rangle| \leq 1\} \equiv \{|x| \leq 1\}. \tag{10}$$

Thus, if $x \neq 0$, then $x^{\triangledown}$ is a half-space in Y, and x^{o} is a "layer". For $x = 0$ we have

$$0^{\triangledown} = 0^{o} = Y. \tag{11}$$

Extend this duality to arbitrary sets. Consider at first the $^{\triangledown}$-case.

We say that a point $x \in X$ is a *friend* with a point $y \in Y$ if



$$\langle x, y\rangle \le 1. \tag{12}$$

Eq. (12) means that

1) $x \in y^{\triangledown}$, that is, x lies in the half-space $\{y \le 1\}$;
2) $y \in x^{\triangledown}$, that is, y lies in the half-space $\{x \le 1\}$.

This friendship is, obviously, a *symmetric* relation.

Exercises.

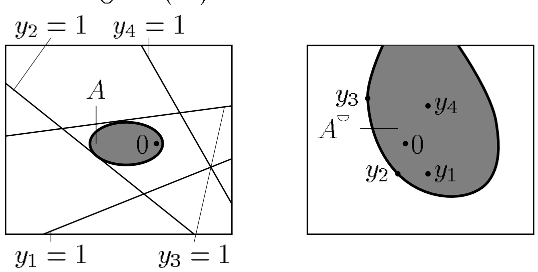
1. x and y are friends $\Rightarrow$ each points of $[0, x]\left(= I_1^+ x\right)$ is a friend with each point of $[0, y]$.
2. y' and y'' are friends with $x \Rightarrow$ each point of the straight line segment $[y', y'']$ is a friend with x.

Definition. For any set $A \subset X$, the set of all points $y \in Y$ that are friends with each points of A, is called the *one-sides polar* of A and is denoted by $A^{\triangledown}$ or πA ("π" comes from "polar"):

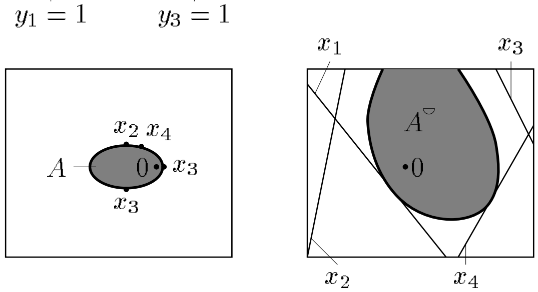
$$\pi A \equiv A^{\triangledown} := \left\{ y \,\middle|\, \forall x \in A \colon \langle x, y\rangle \le 1 \right\} \overset{(9)}{=} \bigcap_{x \in A} x^{\triangledown}. \tag{13}$$

It is clear that always $0 \in A^{\triangledown}$.

Geometrically, this definition admits two (dual one to another) interpretations, according with two meanings of (12):



1) $A^{\triangledown}$ *consists* from all points $y \in Y$ such that the half-space $\{y \le 1\}$ contains A;



2) $A^{\triangledown}$ is the *intersection* of all half-spaces $\{x \le 1\}$, $x \in A$.

Thus $x^{\triangledown} = \{x\}^{\triangledown}$.

Quite analogously, the $\circ$-case leads to balanced polars:

Definition. For $A \subset X$, the (*balanced*) *polar* A° of A is defined as

$$A^{\circ} := \left\{ y \,\middle|\, \forall x \in A \colon |\langle x, y\rangle| \le 1 \right\} = \bigcap_{x \in A} x^{\circ}. \tag{14}$$

It is clear that A° is indeed balanced.

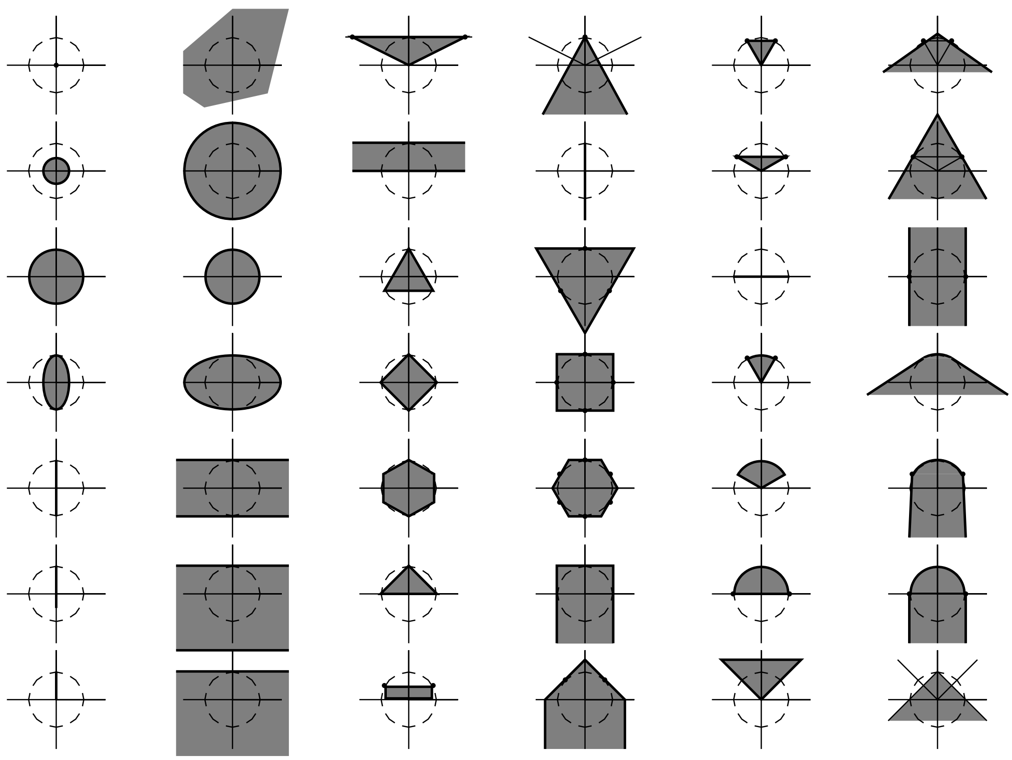
Two interpretations are now:

1) A° consists of all y such that the layer $\{|y| \le 1\}$ contains A;

1) A° is the intersection of all layers $\{|x| \le 1\}$, $x \in A$.

Again, of course, $x^{\circ} = \{x\}^{\circ}$.
Remark. $0^{\triangledown} = 0^{\circ} = Y$.

Some examples of one-sided polars are given below. In each pair of sets the left one is the one-sided polar of the right one, and v. v., the right one is the one-sided polar of the left one. Pay attention to the fact, that straight line segments of the boundary of A correspond to corner points of the boundary of $A^{\triangledown}$, and v. v., corner points of boundary A corresponds to straight line segments of boundary $A^{\triangledown}$. (Dotted line represents the unit circle.)



Examples of one-sided polars for $\mathbb{R}^2 \leftrightarrow \mathbb{R}^2$.

Exercises. [Hint to 1.-6.: use an appropriate from our two interpretations for polars.]

1. $A \subset B \Rightarrow A^{\triangledown} \supset B^{\triangledown},\ A^{\circ} \supset B^{\circ}$.
2. $(A \cup B)^{\triangledown} = A^{\triangledown} \cap B^{\triangledown},\ (A \cup B)^{\circ} = A^{\circ} \cap B^{\circ}$.
3. $(A \cap B)^{\triangledown} \supset A^{\triangledown} \cup B^{\triangledown},\ (A \cap B)^{\circ} \supset A^{\circ} \cup B^{\circ}$. (Here "$\supset$" cannot be replaced by "$=$"; give a counter-example!)
4. $\left(I_1^+ A\right)^{\triangledown} = A^{\triangledown},\ (I_1 A)^{\circ} = A^{\circ}\ \left(I_1^+ := I_1 \cap \mathbb{R}^+ = [0,1]\right)$.
5. $(\mathrm{co} A)^{\triangledown} = A,\ (\mathrm{co}\, A)^{\circ} = A^{\circ}$.
6. $(A \cup \{0\})^{\triangledown} = A^{\triangledown},\ (A \cup \{0\})^{\circ} = 0^{\circ}$.
7. $(-A)^{\triangledown} = -(A^{\triangledown})$. (So we can write without brackets: $-A^{\triangledown}$.)
8. $A^{\circ} = A^{\triangledown} \cap (-A^{\triangledown})$ (see 7.).
9. $A \in \mathrm{Cone} \Rightarrow A^{\triangledown} \in \mathrm{Cone}$. (The cone $-A^{\triangledown}$ (with minus!) is called the *dual cone* to A.)
10. $A \Subset X \Rightarrow A^{\triangledown} = A^{\circ} = A^{\perp}$. $\Big($Recall that $A \Subset X$ means that $A \in \mathrm{Lin}(X)$, and $A^{\perp}$ denotes the *annihilator* of A, defined by the formula $A^{\perp} := \left\{y \,\middle|\, \forall x \in A \ \vdots\ \langle x, y\rangle = 0\right\}\Big)$.
11. $\forall t \in \mathbb{R} \setminus 0 \ \vdots\ (tA)^{\triangledown} = t^{-1} A^{\triangledown},\ (tA)^{\circ} = t^{-1} A^{\circ}$. [Hint: $\langle tx, t^{-1} y\rangle = \langle x, y\rangle$.]

12. For the cannonical duality $\mathbb{R}^n \leftrightarrow \mathbb{R}^n$ it holds

$$B_r^{\circ} = B_{1/r},$$

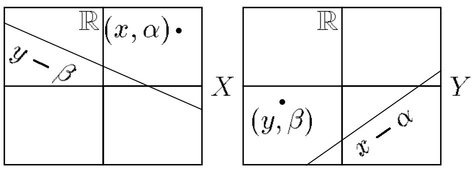
where B_r denotes the closed ball with the radius r and the center at 0.

13. $A \subset A^{\smile\smile} \subset A^{\circ\circ}$. [Hint: By the above two interpretations of polars, $A^{\smile\smile}$ is the intersection of all half-spaces $\{y \leq 1\}$, containing A, and $A^{\circ\circ}$ is the intersection of all layers $\{|y| \leq 1\}$, containing A.]

Operator l

In order to construct a (convex) function on Y starting from a (convex) function on X, we consider the *epigraph* of the function on X (a subset of $X \times \mathbb{R}$) and construct an *epigraph* of a function on Y (a subset of $Y \times \mathbb{R}$) in a manner quite similar to how we defined the one-sided polar.

We say that a point $(x, \alpha) \in X \times \mathbb{R}$ is a *friend* with a point $(y, \beta) \in Y \times \mathbb{R}$, and we



write

$$(x, \alpha) \sim (y, \beta),$$

if

$$\langle x, y \rangle \leq \alpha + \beta. \tag{15}$$

It is convenient to denote an affine function $x \mapsto \langle x, y \rangle - \beta$ by $y - \beta$, just as we write y instead of $\langle \cdot, y \rangle$:

$$(y - \beta)(x) := \langle x, y \rangle - \beta.$$

Then Eq. (15) is equivalent to each of the following two conditions:

1. the point (x, α) lies above the graph of $y - \beta$, i. e.,

$$(x, \alpha) \in \mathrm{epi}(y - \beta); \tag{16}$$

2. the point (y, β) lies above the graph of $x - \alpha$, i. e.,

$$(y, \beta) \in \mathrm{epi}(x - \alpha). \tag{17}$$

$\lhd$ $(x, \alpha) \in \mathrm{epi}(y - \beta) \overset{\text{def.}}{\Longleftrightarrow} (y - \beta)(x) \leq \alpha \Leftrightarrow \langle x, y \rangle - \beta \leq \alpha \Longleftrightarrow (15)$. $\rhd$

It is clear that this friendship relation is *symmetric*.

Thus, for a given point (x, α) in $X \times \mathbb{R}$, the set of all its friends is the half-space $\mathrm{epi}(x - \alpha)$ in $Y \times \mathbb{R}$.

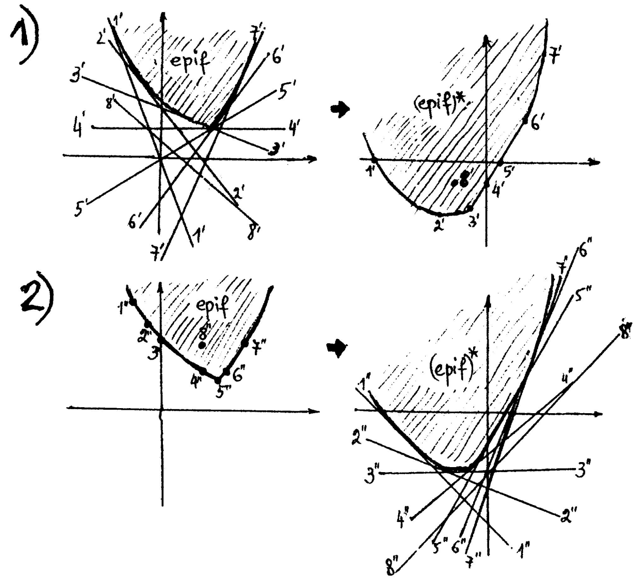
Exercises.

1. $(x, \alpha) \sim (y, \beta)$, $\alpha' > \alpha \Rightarrow (x, \alpha') \sim (y, \beta)$.
2. If two points in $X \times \mathbb{R}$ are friends with some point in $Y \times \mathbb{R}$, then each point of the straight line segment jointing them is also a friend with this point.

Now we construct a function on Y starting from a function $f : X \to \bar{\mathbb{R}}$. Consider the set $(\mathrm{epi} f)^*$ of all points in $Y \times \mathbb{R}$ that are friends with each point of $\mathrm{epi} f$:

$$(\mathrm{epi} f)^* := \left\{ (y, \beta) \middle| \forall (x, \alpha) \in \mathrm{epi} f \,\vdots\, (x, \alpha) \sim (y, \beta) \right\}.$$

This set can be described in two ways, according to (16) and (17), resp. (of two descriptions of polars):



$(\operatorname{epi} f)^*$ is the set of all (y, β) such that

$$\operatorname{epi} f \subset \operatorname{epi}(y - \beta);$$

$(\operatorname{epi} f)^*$ is the intersection of all $\operatorname{epi}(x - \alpha)$, such that

$$(x, \alpha) \in \operatorname{epi} f.$$

Since the last intersection is just the epigraph of the supremum of the functions $x - \alpha$, $(x, \alpha) \in \operatorname{epi} f$ (see p. 18), we see that $(\operatorname{epi} f)^*$ is the epigraph of the function $\bigvee\limits_{(x,\alpha)\in \operatorname{epi} f} (x - \alpha)$. This function is called the *Legendre-Young-Fenchel transform* of (or the *conjugate function* to) f and is denoted by lf (l comes from Legendre) or f^*:

$$lf \equiv f^* := \bigvee_{(x,\alpha)\in \operatorname{epi} f} (x - \alpha). \tag{18}$$

Thus, $(\operatorname{epi} f)^* = \operatorname{epi} f^*$.

Important Remark. If f has the value $-\infty$ if even at one point $\hat{x}$, then $f^* \equiv +\infty$.

$\lhd$ $\forall n \in \mathbb{N} \,\vdots\, (\hat{x}, -n) \in \operatorname{epi} f$, hence $f^* \overset{(18)}{\geq} \bigvee\limits_{n\in\mathbb{N}} (\hat{x} + n) \equiv +\infty$. $\rhd$

Exercise. Verify that $-\infty \in f(X) \Rightarrow f^* \equiv +\infty$, using the 1st description of $\operatorname{epi} f^*$.

The two descriptions of $\operatorname{epi} f^*$ can be reformulated so:

1) $\operatorname{epi} f^*$ consists of all points (y, β) such that

$$f \geq y - \beta;$$

2) if $-\infty \notin f(X)$ then $\operatorname{epi} f^*$ is the intersection of all $\operatorname{epi}(x - f(x))$, such that

$$x \in \operatorname{dom} f.$$

$\lhd$ 1) This follows from the fact that $\operatorname{epi} \varphi \subset \operatorname{epi} \psi \Leftrightarrow \varphi \geq \psi$.

2) First of all, a function $f : \mathbb{R} \to R^{\bullet}$ is *finite* on $\operatorname{dom} f$, so we *can* consider the affine function $x - f(x)$. Our assertion follows from the fact that for any $(x, \alpha) \in \operatorname{epi} f$ it holds $\alpha \geq f(x)$ (so that $x \in \operatorname{dom} f$) and hence $\operatorname{epi}(x - \alpha) \supset \operatorname{epi}(x - f(x))$. $\rhd$

Other representations of f^*. For any $f : X \to R^{\bullet}$

$$lf \equiv f^* = \bigvee_{x\in \operatorname{dom} f} x - f(x), \tag{19}$$

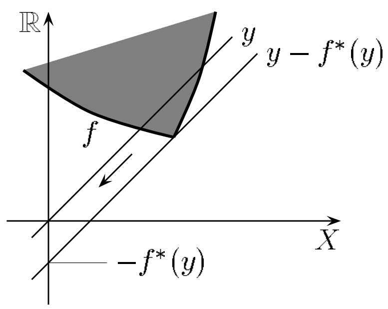
$$\forall y \in Y \,\vdots\, lf(y) \equiv f^*(y) = \sup_{x\in \mathrm{dom} f} (\,\langle x, y\rangle - f(x)\,) \tag{20}$$

$$= \sup_{x\in X} (\,\langle x, y\rangle - f(x)\,) \equiv \sup(y - f). \tag{21}$$

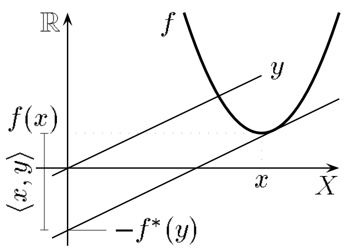
◁ Eq. (19) follows from 2). The first equality in (21) is an immediate consequence of (19), and the second one follows from the fact that $f(x) = +\infty$ for $x \notin \mathrm{dom}\, f$. ▷

Emphasize that in (19) we take the supremum of a *set of functions on* Y; and in (21), dually, we take the supremum of a *fixed function on* X.

The geometrical sense of (21) is elucidated by the first picture below:



$-f^*(y)$ is "the lowest point" of the projection of epi f on $\mathbb{R}$ "along" y.



For a smooth f (see the second picture) $-f^*(y)$ is the $\mathbb{R}$-coordinate of the point of intersection with the $\mathbb{R}$-axis of the "*tangent* line" to gr f taken for a point x where $f'(x) = y$. So for $X = \mathbb{R}$ f^* is defined by the equations

$$\begin{cases} f(x) + f^*(y) = xy \\ y = f'(x) \end{cases} \tag{22}$$

The transition from f to f^*, defined by (22), is called the *Legendre transform* in classical analysis.

Young inequality. For any $f : X \to \bar{\mathbb{R}}$

$$\forall \alpha \in X \;\forall y \in Y \,\vdots\, \langle x, y\rangle \le f(x) + f^*(y) \tag{23}$$

◁ Each point of epi$f$ is a friend with each point of epi f^*, hence

$$(\,x, f(x)\,) \sim (\,y, f^*(y)\,) \;▷$$

Since X and Y occur in our dual pair quite symmetrically, we can apply the same operator l to lf. The function

$$f^{**} \equiv l^2 f := l(lf) : X \longrightarrow \bar{\mathbb{R}}$$

is called the *second conjugate* function to f.

Description of f^{**}. For any $f : X \to \bar{\mathbb{R}}$ the function f^{**} is the supremum of all affine functions of the form $y - \beta$, $(y, \beta) \in Y \times \mathbb{R}$, that are less than f:

$$f^{**} = \bigvee_{\substack{(y,\beta)\in Y\times\mathbb{R} \\ y-\beta\leq f}} (y - \beta). \tag{24}$$

$\triangleleft$ By the definition of l, $f^{**} = l(f^*) = \bigvee_{(y,\beta)\in \operatorname{epi} f^*} (y - \beta)$, but, by), $\operatorname{epi} f^*$ consists just of all (y, β), such that $y - \beta \leq f$. $\triangleright$

Corollary. For any $f : X \to \bar{\mathbb{R}}$

$$f^{**} \leq f. \tag{25}$$

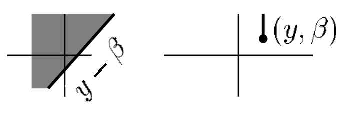
Examples.

1.

$$f \equiv +\infty \Longrightarrow f^* \equiv -\infty; \quad f \equiv -\infty \Longrightarrow f^* \equiv +\infty. \tag{26}$$

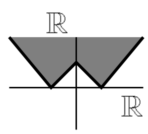
2.

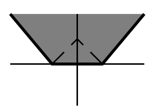
$$(\, y - \beta\,)^* = \delta\{y\} + \beta; \quad (\, \delta\{y\} + \beta\,)^* = y - \beta \tag{27}$$



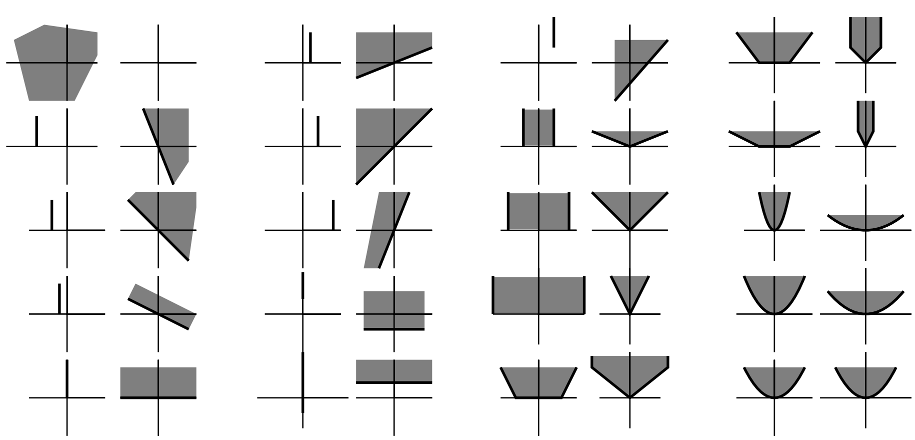
In particular ($\beta = 0$):

$$l\langle\cdot, y\rangle = \delta\{y\}; \quad l\delta\{y\} = \langle\cdot, y\rangle. \tag{28}$$

3. $f =$ 

$\Rightarrow f^{**} =$  .

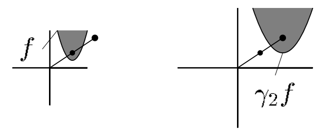
Some other examples of conjugate functions, for $\mathbb{R} \leftrightarrow \mathbb{R}$, are given below (the functions are represented by their epigraphs). In each pair of functions the left one is the conjugate of the right one, and v. v. Pay attention that again, as for polars, straight line segments of the graph of f correspond to corner points of $\operatorname{gr} f^*$, and v. v.



Examples of conjugate functions for $\mathbb{R} \leftrightarrow \mathbb{R}$.

Exercises.

1. $f \le g \Rightarrow f^* \ge g^*$.
2. $(f \wedge g)^* = f^* \vee g^*$.
3. $(f \vee g)^* \le f^* \wedge g^*$. Here "$\le$" cannot be replaced by "$=$". E. g., $\forall y_1, y_2 \in Y : y_1 \ne y_2 \;\vdots\; (y_1 \vee y_2)^* < y_1^* \wedge y_2^*$ (drawn the picture!).
4. $\forall \alpha \in \mathbb{R} \;\vdots\; (f + \alpha)^* = f^* - \alpha$.
5. For any $t > 0$ define an operator $\gamma_t : \mathcal{F}(X, \bar{\mathbb{R}}) \hookleftarrow$ (where $\mathcal{F}(X, \mathbb{R})$ is the space of all functions $f : X \to \bar{\mathbb{R}}$) by the formula



$$\operatorname{epi}(\gamma_t f) = t \operatorname{epi} f.$$

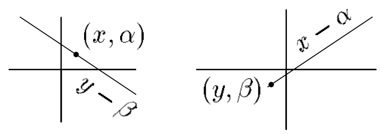
$$\text{a) } (tf)^* = \gamma_t(f^*); \quad \text{b) } (\gamma_t f)^* = tf^*.$$

6. For $\mathbb{R} \leftrightarrow \mathbb{R}$, if $f(x) = \frac{1}{2}ax^2$, then $f^*(y) = \frac{1}{2a}y^2$. (Thus, $\frac{1}{2}x^2$ goes into "itself". It is the unique function on $\mathbb{R}$ with this property.)
7. Put

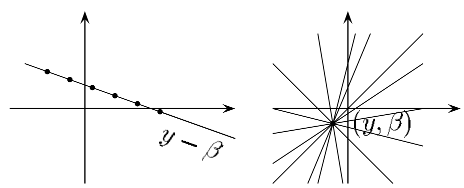
$$(x, \alpha) \dot{\sim} (y, \beta) :\Longleftrightarrow \langle x, y \rangle = \alpha + \beta \tag{29}$$

((x, α) and (y, β) are "lovers"). Obviously, it is a symmetric relation.

a) $(x, \alpha) \dot{\sim} (y, \beta) \Leftrightarrow (x, \alpha) \in \operatorname{gr}(y - \beta) \Leftrightarrow (y, \beta) \in \operatorname{gr}(x - \alpha)$.



b) If points (x_i, α_i), $i \in I$, lie on the graph of some affine function $y - \beta$, then all graphs $\operatorname{gr}(x_i - \alpha_i)$ go through the point (β). (That is why straight line segments of $\operatorname{gr} f$ give corner points of $\operatorname{gr} f^*$, and v. v., corner points give straight line segments. See the picture on p. 55.)



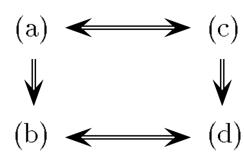
8. For smooth $f : \mathbb{R} \to \mathbb{R}$ it holds

$$(x, f'(x)) \dot{\sim} (f'(x), f^*(x)).$$

(See (22).)

9. Consider the following conditions:
   (a) $\langle x, y \rangle = f(x) + f^*(y)$,
   (b) $\langle x, y \rangle = f^{**}(x) + f^*(y)$,
   (c) $y \in \partial_x f$,
   (d) $x \in \partial_y f^*$.

Show that



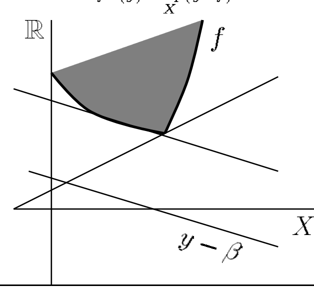
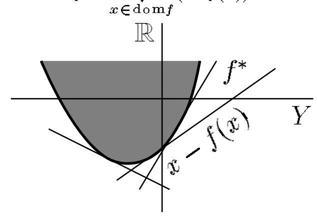
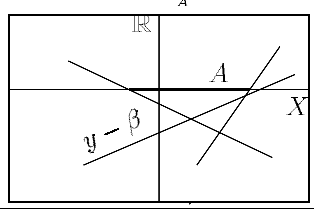
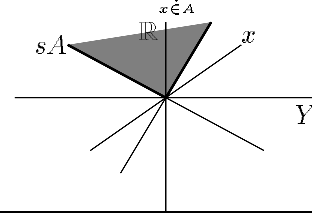
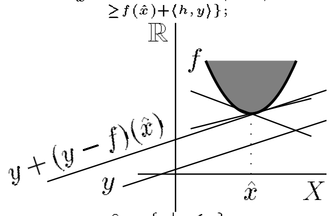
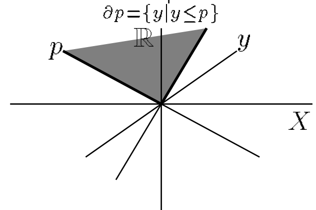
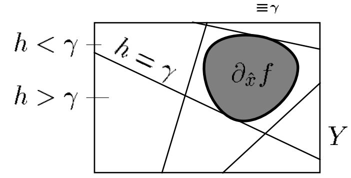
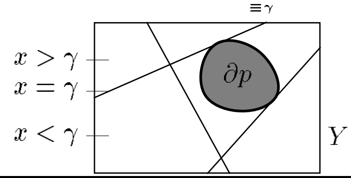
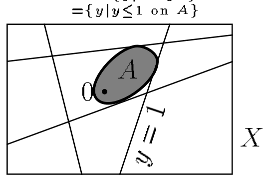
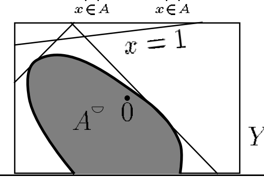
[Use the Young inequality.]

10. $\partial_x f \ne 0 \Rightarrow f^{**}(x) = f(x)$. [Use 9.]

Summarize all this.

Three descriptions of the operators of convex Analysis

$$\text{(below } f: X \to R^{\bullet},\ p \in \mathrm{Sublin}(X),\ A \subset X)$$

	friendship-representation	$\leq$−representation (in terms of subsets of X and functions on X)	$\bigcap$−representation (in terms of subsets of Y and functions on Y)
l	$\mathrm{epi} f^* = \{(y,\beta) \mid \forall (x,\alpha) \in \mathrm{epi} f \vdots (x,\alpha) \sim (y,\beta)\}$ $= \{(y,\beta) \mid f(x) \leq \alpha \Rightarrow \langle x,y\rangle \leq \alpha + \beta\}$	$\mathrm{epi} f^* = \{(y,\beta) \mid y - \beta \leq f\}$ $f^*(y) = \sup_X (y - f)$ 	$\mathrm{epi} f^* = \bigcap_{x \in \mathrm{dom} f} \mathrm{epi}(x - f(x))$ $f^* = \bigvee_{x \in \mathrm{dom} f} (x - f(x))$ 
s	$\mathrm{epi}\, s A = \{(y,\beta) \mid \forall x \in A \vdots (x,0) \sim (y,\beta)\}$ $= \{(y,\beta) \mid \forall x \in A \vdots \langle x,y\rangle \leq \beta\}$	$\mathrm{epi}\, s A = \{(y,\beta) \mid y - \beta \leq 0 \text{ on } A\}$ $s A(y) = \sup_A y$ 	$\mathrm{epi}\, s A = \bigcap_{x \in A} \mathrm{epi}\, x$ $s A = \bigvee_{x \in A} x$ 
∂	$f = \{y \mid \forall (x,\alpha) \in \mathrm{epi} f \vdots (x,\alpha) \sim (y, (y - f)(\hat{x}))$ $= \{y \mid f(x) \leq \alpha \Rightarrow \langle x,y\rangle \leq \alpha + (y - f)(\hat{x})$ $\partial p = \{y \mid \forall (x,\alpha) \in \mathrm{epi} f \vdots (x,\alpha) \sim (y,0)\}$ $= \{y \mid f(x) \leq \alpha \Rightarrow \langle x,y\rangle \leq \alpha\}$	$\partial_{\hat{x}} f = \{y \mid \forall h \in X \vdots f(\hat{x} + h) \geq f(\hat{x}) + \langle h,y\rangle\};$  $\partial p = \{y \mid y \leq p\}$ 	$\partial_{\hat{x}} f = \bigcap_{h \in X} \{y \mid \langle h,y\rangle \leq \underbrace{f(\hat{x} + h) - f(\hat{x})}_{\equiv \gamma}\}$  $\partial p = \bigcap_{x \in X} \{y \mid \langle x,y\rangle \leq \underbrace{p(x)}_{\equiv \gamma}\}$ 
π	$A^{\triangledown} = \{y \mid \forall x \in A \vdots x \sim y\}$ $= \{y \mid \forall x \in A \vdots \langle x,y\rangle \leq 1\}$	$A^{\triangledown} = \{y \mid A \subset y^{\triangledown}\}$ $= \{y \mid y \leq 1 \text{ on } A\}$ 	$A^{\triangledown} = \bigcap_{x \in A} x^{\triangledown} = \bigcap_{x \in A} \{x \leq 1\}$ 

Orded properties of l, s, ∂, π. The operators of convex analysis have the following properties (below I is an arbitrary index set):

monotony	$\bigcap$ *−property*
$f_1 \le f_2 \Rightarrow \begin{cases} f_1^* \ge f_2^* \\ \partial_x f_1 \subset \partial_x f_2 \text{ if } f_1(x) = f_2(x) \end{cases}$	$\Big(\bigwedge\limits_{\iota\in I} f_\iota\Big)^* = \bigwedge\limits_{\iota\in I} f_\iota^*$ $\partial_x\Big(\bigwedge\limits_{\iota\in I} f_\iota\Big) = \bigcap\limits_{\iota\in I} \partial_x f_\iota$ if all f_ι have at x a common value
$A_1 \subset A_2 \Rightarrow \begin{cases} sA_1 \le sA_2 \\ A_1^{\triangledown} \supset A_2^{\triangledown}, A_1^\circ \supset A_2^\circ \end{cases}$	$s\Big(\bigcup\limits_{\iota\in I} A_\iota\Big) = \bigvee\limits_{\iota\in I} sA_\iota$ $(\bigcup A_\iota)^{\triangledown} = \bigcap A_\iota^{\triangledown},\ (\bigcup A_\iota)^\circ = \bigcap A_\iota^\circ$

$\lhd$ 0°

1) $\operatorname{epi}(\bigvee f_\iota) = \bigcap \operatorname{epi} f_\iota$,

2) $\operatorname{epi}(\bigwedge f_\iota) = \overline{\bigcup \operatorname{epi} f_\iota}^{\,\mathrm{epi}}$, where the *epi-closure* $\overline{C}^{\,\mathrm{epi}}$ of a set $C \subset X \times \mathbb{R}$ is defined as

$$\overline{C}^{\,\mathrm{epi}} := \big\{(x,\alpha)\big|\, \alpha \ge \inf C_x\big\}, \quad C_x := \big\{\alpha \in \mathbb{R}\big|\, (x,\alpha) \in C\big\}.$$

3) $\forall (x,\alpha) \in C \vdots (x,\alpha) \sim (y,\beta) \Leftrightarrow \forall (x,\alpha) \in \overline{C}^{\,\mathrm{epi}} \vdots (x,\alpha) \sim (y,\beta)$.

1° The monotony properties follows from the $\bigcap$-properties. $\lhd\lhd$ E. g., $f_1 \le f_2 \overset{\text{obv.}}{\Longrightarrow} f_1 = f_1 \wedge f_2 \overset{\bigcap\text{-prop.}}{\Longrightarrow} f_1^* = f_1^* \vee f_2^* \overset{\text{obv.}}{\Longrightarrow} f_1^* \ge f_2^*$. $\rhd\rhd$

2° $\bigcap$-properties follow from the friend representations (f. r.). $\lhd\lhd$ E. g., let us prove that $(\bigwedge f_i)^* = \bigvee f_\iota^*$. We need show that $\operatorname{epi}(\bigwedge f_\iota)^* = \operatorname{epi}(\bigvee f_\iota^*)$. But, indeed,

$$\begin{aligned}
(y,\beta) \in \operatorname{epi}\Big(\bigwedge f_\iota\Big)^* &\overset{\text{f. r.}}{\Longleftrightarrow} \forall (x,\alpha) \in \operatorname{epi}\Big(\bigwedge f_\iota\Big) \vdots (x,\alpha) \sim (y,\beta) \\
&\overset{0^\circ 2)}{\Longleftrightarrow} \forall (x,\alpha) \in \overline{\bigcup \operatorname{epi} f_\iota}^{\,\mathrm{epi}} \vdots (x,\alpha) \sim (y,\beta) \\
&\overset{0^\circ 3)}{\Longleftrightarrow} \forall (x,\alpha) \in \bigcup \operatorname{epi} f_\iota \vdots (x,\alpha) \sim (y,\beta) \\
&\overset{\text{obv.}}{\Longleftrightarrow} \forall \iota\, \forall (x,\alpha) \in \operatorname{epi} f_\iota \vdots (x,\alpha) \sim (y,\beta) \\
&\overset{\text{f. r.}}{\Longleftrightarrow} \forall \iota \vdots (y,\beta) \in \operatorname{epi} f_\iota^* \\
&\overset{\text{obv.}}{\Longleftrightarrow} (y,\beta) \in \bigcap \operatorname{epi} f_\iota^* \\
&\overset{0^\circ 1)}{\Longleftrightarrow} (y,\beta) \in \operatorname{epi} \bigvee f_\iota^*. \ \rhd \rhd \rhd
\end{aligned}$$

Exercise. Prove $0^\circ 2)$ and $0^\circ 3)$.

Topological properties of operators l, ∂, s, π.

All this was so far a pure algebra (with the exception of $\bigcap$-properties). Now topology goes into the play.

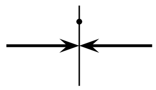
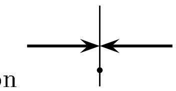
Definition 1. Let X be a topological space. A function $f : X \to \bar{\mathbb{R}}$ is called *closed* (it would more appropriate to say "closed from *below*", but it is long) if its *epi*graph is closed in $X \times \mathbb{R}$:

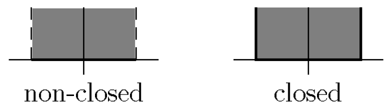
$$f \in \mathrm{Cl} :\Longleftrightarrow \operatorname{epi} f \in \mathrm{Cl}(X \times \mathbb{R}).$$

Remark. Another name of such functions is *lower semi-continuous.*

Examples.

1. The function $f \equiv +\infty$ (resp., $-\infty$) is closed.

2. On $\mathbb{R}$, the function  is *not* closed, the function  *is*.
3. $\delta A \in \mathrm{Cl} \Leftrightarrow A \in \mathrm{Cl}(X)$.



4. Each *continuous* function $X \to \mathbb{R}$ is closed. $\lhd$ The epigraph is the pre-image of the closed set $[0, +\infty)$ by the continuous mapping $(x, \alpha) \mapsto \alpha - f(x)$. $\rhd$

   (Notice that continuous functions are closed also "from above" and have a closed *graph*.)

Definition 2. Let $X \leftrightarrow Y$. We say that a function $X \to \bar{\mathbb{R}}$ is *weakly closed* if f is closed X being equipped with $\sigma(X, Y)$:

$$f \in \mathrm{wCl} :\Longleftrightarrow \mathrm{epi}\, f \in \mathrm{Cl}(\,(X, \sigma(X, Y)\,) \times \mathbb{R}\,).$$

Example. Let$X \leftrightarrow Y$. Then each affine function $y - \beta$, $(y, \beta) \in Y \times \mathbb{R}$, is weakly closed. $\lhd$ $y - \beta = \langle \cdot, y \rangle - \beta$ is continuous on $(\,X, \sigma(X, Y)\,)$. (Recall that $(\,X, \sigma(X, Y)\,)^* = Y$.) $\rhd$

Lemma on wClConv. *Let $X \leftrightarrow Y$.*

a) $\forall f : X \to \bar{\mathbb{R}} \;\vdots\; f^* \in wClConv$; if $f(x) \in \mathbb{R}$, then $\partial_x f \in wClConv(Y) := ClConv(\,Y, \sigma(X, Y)\,)$.

b) $\forall A \subset X \;\vdots\; A^{\triangledown}, A^{\circ} \in wClConv(Y)$; if $A \neq 0$, then $sA \in wClSublin$.

$\lhd$ All this follows at once from $\bigcap$-representations, from the fact that the intersection of any family of closed convex sets is also a closed convex one, and from the fact that each half-space $\{x \leq \gamma\}$ and each epigraph $\mathrm{epi}(x - \alpha)$ are closed by equipping of Y with $\sigma(X, Y)$. $\rhd$

Thus, the operators of convex analysis turn every thing to a weakly closed convex one.

2.3.2 Duality of properties to be bounded and to be absorbing

Here we prowe that by passing to polars the property to be weakly bounded goes to the property to be absorbing.

Lemma. *Let $X \leftrightarrow Y$. Then the polars of all finite sets in Y form a base of neighbourhoods of* 0 *for $\sigma(X, Y)$.*

$\lhd$ It follows at once from the relation

$$\{y_1, \dots, y_n\}^{\circ} = \bigcap_{i=1}^{n} \{|y_i| \leq 1\}, \tag{1}$$

which is none more than the $\bigcap$-representation of the polar of the set $\{y_1, \dots, y_n\}$ (see 14 on p. 52). $\rhd$

Theorem on Bddand Abs. *Let $X \leftrightarrow Y$, and let $A \subset X$. Then*

$$A \in Bdd(\,\sigma(X,Y)\,) \iff A^\circ \in Abs(Y). \tag{2}$$

$\lhd$ 0°

1) The lemma above.

2) $A \subset A^{\circ\circ}$. (See Exercise 13 on p. 54.)

3) $A \subset B \Rightarrow A^\circ \supset B^\circ$ (monotony of $^\circ$).

4) $\forall t \in \mathbb{R} \setminus 0 \,\vdots\, (tA)^\circ = t^{-1}A^\circ$. (See Exercise 11 on p. 53.)

1° "$\Rightarrow$". Let $A \in \mathrm{Bdd}(\,\sigma(X,Y)\,)$, and let y be an arbitrary point of Y. We need show that A° absorbes y. By $0^\circ 1)$, $y^\circ \in \mathrm{Nb}_0(\,\sigma(X,Y)\,)$. So $\exists t > 0 \,:\, tA \subset y^\circ$. Hence,

$$y \overset{0^\circ 2)}{\in} \{y\}^{\circ\circ} \overset{0^\circ 3)}{\subset} (\,tA\,)^\circ \overset{0^\circ 4)}{=} t^{-1}A^\circ,$$

that is, y is absorbed by A°.

2° "$\Leftarrow$". Let A° be absorbing. We need show that A is $\sigma(X,Y)$-bounded. Let U be an arbitrary neighbourhood of 0 in $\sigma(X,Y)$. By $0^\circ 1)$,

$$U \supset \{y_1, \ldots, y_n\}^\circ \tag{3}$$

for some $y_1, \ldots, y_n \in Y$. Since A° is absorbing and *balanced*, it holds, for sufficiently small $t > 0$, $ty_i \in A^\circ$, $i = 1, \ldots, n$, that is,

$$A^\circ \supset t\{y_1, \ldots, y_n\}. \tag{4}$$

Hence,

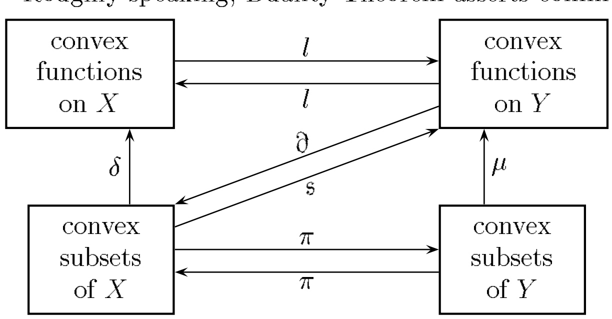
$$A \overset{0^\circ 2)}{\subset} A^{\circ\circ} \overset{(4),0^\circ 3)}{\subset} (\,t\{y_1, \ldots, y_n\}\,)^\circ \overset{0^\circ 4)}{=} t^{-1}\{y_1, \ldots, y_n\}^\circ \overset{(3)}{\subset} t^{-1}U.$$

Thus, A is absorbed by U. $\rhd$

2.3.3 Duality theorem

This is one of the central results of convex analysis.

Roughly speaking, Duality Theorem asserts commutativity of the following diagram:



	"hard" part		"easy" part
a)	$l^2 = id$	e)	$l\delta = s$
b)	$\pi^2 = id$	f)	$l\mu = \delta\pi$
c)	$s\partial = id$	g)	$\mu\pi = s$
d)	$\partial s = id$	h)	$\delta\partial = l$

More precisely:

Duality theorem. *Let $X \leftrightarrow Y$, $f : X \to R^\bullet$, $0 \neq A \subset X$, $p \in Sublin(X)$. Then:*

I (The "hard" part, with the weak topology, without δ and μ.) There hold equivalences:

a) $x^{**} = f \iff f \in wClConv$ *(Fenchel-Moreau theorem);*

b) $\left.\begin{array}{l} A^{\triangledown\triangledown} = A \iff A \in wClConv \text{ and } 0 \in A \\ A^{\circ\circ} = A \iff A \in \\ mboxwClConvBal \end{array}\right\}$ *(theorem on bipolar);*

c) $s\partial p = p \iff p \in wCl$ *(Minkovski theorem);*

d) $\partial s A = A \iff A \in wClConv$ *(Hörmander theorem).*

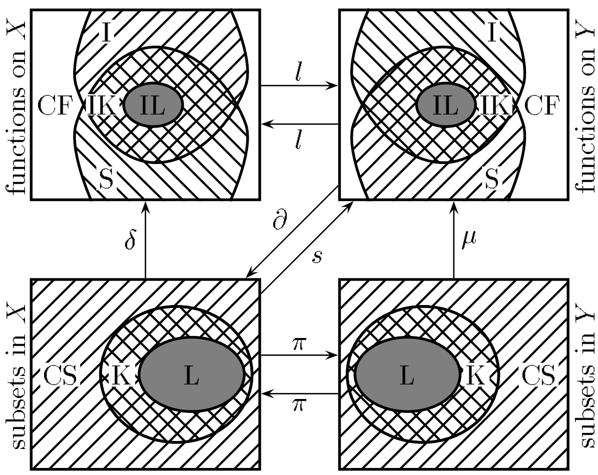
II (The "easy" part, without the weak topology, with δ and μ.) There hold equations:

e) $(\delta A)^* = sA;$

f) $(\mu A)^* = \delta A^{\triangledown};$

g) $\mu A^{\triangledown} = sA$ *if* $0 \in A;$

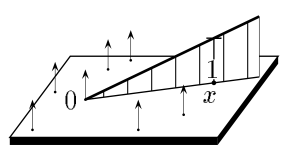
h) $\delta\partial p = p^*.$



Remark. Some more detailed, the commutative diagram looks as shown to the left. The equally shaded regions go one into another!
CF ... convex functions
CS ... convex sets
I ... indicator functions
IK ... ind. functions of cones (IK = I ∩ S)
IL ... ind. functions of lin. subspaces
K ... cones ($\mu = \delta$ on K)
K ... lin. subspaces ($\pi = {}^{\perp}$ on L)
S ... sublinear functions

◁ We begin from the "easy" part.

II. 0°
1) $\bigcap$-property of l;
2) $\bigcap$-representation of l and π;
3) $\mu A = \bigwedge_{x \in A} e_x$, where e_x is $\delta\{0\}$ for $x = 0$ and is



for $x \neq 0$ (see p. 19);
4) $l\delta\{x\} = \langle x, \cdot \rangle$ (see (27) on p. 57).

1° $\forall x \in X \,\vdots\, le_x = \delta x^{\triangledown}$. ◁◁ If $x = 0$ all is trivial: $e_0 = \delta\{0\}$, $l\delta\{0\} \overset{0^\circ 4)}{=} 0$, $0^{\triangledown} = Y$, $\delta Y = 0$. If $x \neq 0$ then

$$le_x \overset{0^\circ 2)}{=} \bigvee_{\xi \in \mathrm{dom}\, e_x} (\xi - e_x(\xi)) \overset{\text{def. of } e_x}{=} \bigvee_{t>0} (tx - t) \overset{\text{obv.}}{=} \bigvee_{t>0} t(x-1),$$

so

$$le_x(y) = \sup_{t>0} t(\langle x,y\rangle - 1) = \begin{cases} 0 & \text{if } \langle x,y\rangle \le 1, \text{ that is, if } y \in x^{\triangledown}, \\ +\infty & \text{if } \langle x,y\rangle > 1, \text{ that is, if } y \notin x^{\triangledown}. \end{cases}$$

But this just means that le_x is $\delta x^{\triangledown}$. ▷▷

2° e) ◁◁ $l\delta A \overset{\text{obv.}}{=} l \bigwedge_{x\in A} \delta\{x\} \overset{0^\circ 1)}{=} \bigvee_{x\in A} l\delta\{x\} \overset{0^\circ 4)}{=} \bigvee_{x\in A} \langle x,\cdot\rangle \overset{\text{def. of } s}{=} sA.$ ▷▷

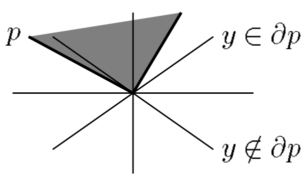
3° f) ◁◁ $l\mu A \overset{0^\circ 3)}{=} l \bigwedge_{x\in A} e_x \overset{0^\circ 1)}{=} \bigvee_{x\in A} le_x \overset{1^\circ}{=} \bigvee_{x\in A} \delta x^{\triangledown} \overset{\text{obv.}}{=} \delta \bigcap_{x\in A} x^{\triangledown} \overset{0^\circ 2)}{=} \delta A^{\triangledown}.$ ▷▷

4° g) ◁◁ Let $0 \in A$. Then $\forall y \in Y$:

$$\begin{aligned} \mu\pi A(y) &\overset{\text{def. of } \mu}{=} \inf\{\alpha > 0 \,|\, \alpha^{-1}y \in \pi A\} \overset{\text{def. of } \pi}{=} \inf\{\alpha > 0 \,|\, \langle A, \alpha^{-1}y\rangle \le 1\} \\ &\overset{\text{obv.}}{=} \inf\{\alpha > 0 \,|\, \langle A, y\rangle \le \alpha\} \overset{!}{=} \inf\{\alpha \in \mathbb{R} \,|\, \langle A, y\rangle \le \alpha\} \overset{\text{def. of sup}}{=} \sup\langle A, y\rangle \\ &\overset{\text{def. of } s}{=} sA(y). \end{aligned}$$

The equality marked by "!" is true since the set $\langle A, y\rangle\,(\subset \mathbb{R})$ contains 0 ($0 \in A!$), and hence cannot have negative majorants. ▷▷

5° h) ◁◁ It follows immediatelly from the definition of ∂p that



$$p^*(y) = \sup(y - p) = \begin{cases} 0 & \text{if } y \in \partial p, \\ +\infty & \text{if } y \notin \partial p. \end{cases}$$

But this just means that $p^* = \delta\partial p$. ▷▷

6° Thus, the easy part is over. Go to the hard one.

I 0°

1) Strict Separation theorem;
2) $f^{**} = \bigvee_{y-\beta\le f} (y - \beta)$ (see (24) on p. 57);
3) $f^{**} \le f$ (see (25) on p. 57);
4) $(X, \sigma(X,Y))^* = Y$;
5) Lemma on wClConv (see p. 61);
6) $\langle x,y\rangle \le \alpha + \beta \Leftrightarrow (x,\alpha) \in \operatorname{epi}(y - \beta)$ (see p. 54);
7) if $\varphi \equiv +\infty$ and $\psi \equiv -\infty$, then $\varphi^* = \psi$ and $\psi^* = \varphi$ (see Example 1 on p. 57);
8) $\operatorname{epi}\bigvee = \bigcap \operatorname{epi}$.

1° Each continuous linear functional l on $(X, \sigma(X,Y)) \times \mathbb{R}$ can be uniquely represented in the form

$$(x,\alpha) \longmapsto \langle x,y\rangle + \alpha\beta$$

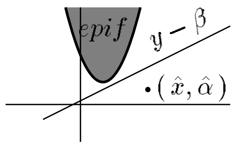
for some $(y,\beta) \in Y \times \mathbb{R}$. (Compare with the general form of a linear functional on $\mathbb{R}^2 = \mathbb{R}\times\mathbb{R}$.)
◁◁ The restriction of l on $X \times 0$ is a continuous linear functional on $(X, \sigma(X,Y))$ and hence can be uniquely represented in the form $\langle\cdot, y\rangle$ for some $y \in Y$, by $0^\circ 4)$. The restriction of l on $0 \times \mathbb{R}$ can be, obviously, represented in the form $\alpha \mapsto \alpha\beta$ for some $\beta \in \mathbb{R}$. Our assertion follows now from linearity of l. ▷▷

2° All implications "⇒" in a)–d) follows at once from $0^\circ 5)$. Let us prove "⇐".

3° a) ◁◁ Let $f \in$ wClConv. We have to show that $f^{**} = f$. If $f \equiv +\infty$, then this is true by $0^\circ 7)$. So we can assume that

$$\operatorname{dom} f \ne 0. \tag{1}$$

By $0^\circ 3)$, we need verify that $f \le f^{**}$, or, equinalently, that

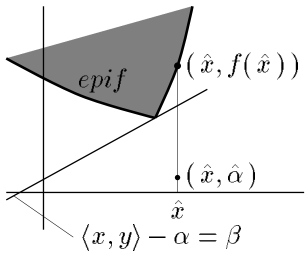


$$\operatorname{epi} f \supset \operatorname{epi} f^{**} \overset{0^\circ 2)}{=} \operatorname{epi} \bigvee_{y-\beta \le f} (y-\beta) \overset{0^\circ 8)}{=} \bigcap_{y-\beta\le f} (y-\beta).$$

So it suffices to prove that

$$\forall (\,\hat{x},\hat{\alpha}\,) \notin \operatorname{epi} f \; \exists (y,\beta) \in Y \times \mathbb{R} :\; y-\beta \le f,\; (\,\hat{x},\hat{\alpha}\,) \notin \operatorname{epi}(y-\beta). \tag{2}$$

Let $(\hat{x},\hat{\alpha}) \notin \operatorname{epi} f$. By the supposition, $\operatorname{epi} f$ is convex and closed in $(\,X,\sigma(X,Y)\,)\times\mathbb{R}$. Hence, by $0^\circ 1)$ and 1°,



$$\exists (z,\gamma) \in Y \times \mathbb{R} : \sup_{(x,\alpha)\in \operatorname{epi} f} \underbrace{(\,\langle x,z\rangle + \alpha\gamma\,)}_{(3')} < \underbrace{\langle \hat{x},z\rangle + \hat{\alpha}\gamma}_{(3'')}. \tag{3}$$

Now, γ cannot be > 0, since then the supremum in (3) would be equal to $+\infty$ (since $\operatorname{epi} f \neq 0$, by (1)). Further, consider two possible cases.

4° *Case 1:* $\hat{x} \in \operatorname{dom} f$. In this case $f(\hat{x}) \in \mathbb{R}$ and $(\,\hat{x}, f(\hat{x}) \in \operatorname{epi} f\,)$. So γ cannot $= 0$, since then for $(x,\alpha) = (\,\hat{x}, f(\hat{x})\,)$ we would have $(3') = (3'')$, which contradicts (3). Hence $\gamma < 0$. Deviding (3) by $-\gamma$ (which is positive!) and putting $y := \frac{z}{-\gamma}$, we obtain

$$\beta := \sup_{(x,\alpha)\in\operatorname{epi} f} (\,\langle x,y\rangle - \alpha\,) < \langle \hat{x},y\rangle - \hat{\alpha}. \tag{4}$$

It follows from (4) that

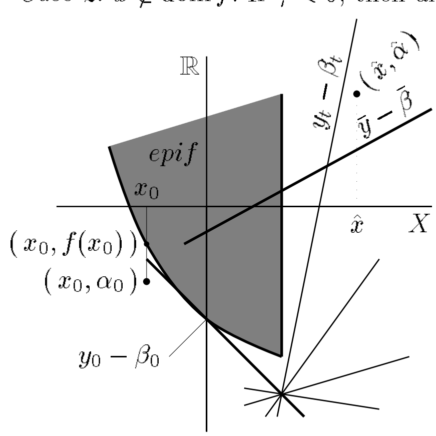
$$\forall (x,\alpha) \in \operatorname{epi} f \;\vdots\; \langle x,y\rangle \le \alpha+\beta \overset{0^\circ 6)}{\Longrightarrow} \operatorname{epi} f \subset \operatorname{epi}(y-\beta) \overset{\text{obv.}}{\Longrightarrow} y-\beta \le f$$

and

$$\langle \hat{x},y\rangle > \alpha+\beta \overset{0^\circ 6)}{\Longrightarrow} (\,\hat{x},\hat{\alpha}\,) \notin \operatorname{epi}(y-\beta).$$

So (2) is fulfilled. (Notice that the graph of $y-\beta$ is just the "β-level line" of our separating functional $(x,\alpha) \mapsto \langle x,y\rangle - \alpha$.)

5° *Case 2:* $\hat{x} \notin \operatorname{dom} f$. If $\gamma < 0$, then all is O. K., just as in Case 1.



If $\gamma = 0$, then (3) turns to (we will write now $\bar{y}$ instead of z)

$$\bar{\beta} := \sup_{x\in\operatorname{dom} f} \langle x,\bar{y}\rangle < \langle \hat{x},\bar{y}.\rangle \tag{5}$$

It follows from (5) that

$$\bar{y} - \bar{\beta} \le 0 \text{ on } \operatorname{dom} f, \tag{6}$$

$$(\,\bar{y}-\bar{\beta}\,)(\,\hat{x}\,) > 0. \tag{7}$$

Now, take any $x_0 \in \operatorname{dom} f$ (it is possible by (1)). Since f has values in $\mathbb{R}^{\bullet}$, $f(x_0) \in \mathbb{R}$. Take any finite $\alpha_0 < f(x_0)$. Then $(x_0,\alpha_0) \notin \operatorname{epi} f$. Just as in Case 1 we find $(y_0,\beta_0) \in Y \times \mathbb{R}$, such that

$$y_0 - \beta_0 \le f. \tag{8}$$

(That $(x_0, \alpha_0) \notin \mathrm{epi}(y_0 - \beta_0)$, is now unimportant for us: we need just *any* affine minorante for f.) Put for each $t \in \mathbb{R}$

$$y_t - \beta_t := (y_0 - \beta_0) + t(\bar{y} - \bar{\beta}).$$

By (6) and (8),

$$\forall t \geq 0 \,\vdots\, y_t - \beta_t \leq f.$$

By (7), $(y_t - \beta_t)(\hat{x}) \to +\infty$ as $t \to +\infty$. So for sufficiently big t it holds $(y_t - \beta_t)(\hat{x}) > \hat{\alpha}$, that is,

$$(\hat{x}, \hat{\alpha}) \notin \mathrm{epi}(y_t - \beta_t).$$

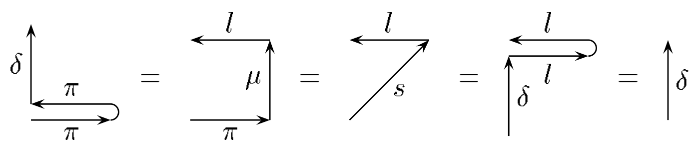
Thus, (2) is fulfilled for this (y_t, β_t). ▷▷

6° The rest is a play with diagrams: all other commutativity equations follows from the already proved ones.

b) ◁◁ Let $A \in$ wClConv, and let $0 \in A$. We need show that $\pi^2 A = A$, or, which is equivalent, that $\delta\pi^2 A = \delta A$. But, indeed,

$$\delta\pi^2 A \overset{f)}{=} l\mu\pi A \overset{g)}{=} lsA \overset{e)}{=} l^2\delta A \overset{a)}{=} \delta A$$

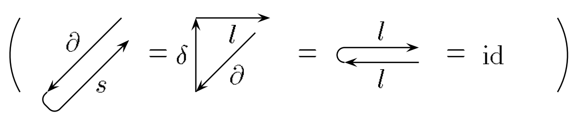
(we can apply a) to δA, since δA is convex and weakly closed together with A). This chain of equalities becomes more clear when represented graphically (see the "full" diagram before the formulation of the theorem):



The second implication (for balanced polars) follows at once from the first one and from the evident fact that $A^{\circ} = A^{\triangledown}$ for balanced sets A. ▷▷

7° c) ◁◁ Let $p \in$ wClSublin. Then we have

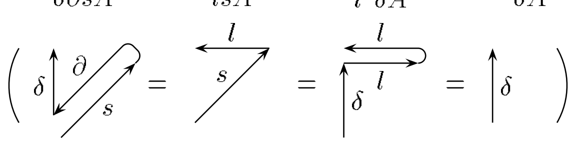
$$s\partial p \overset{e)}{=} l\delta\partial p \overset{h)}{=} l^2 p \overset{a)}{=} p$$



▷▷

8° d) ◁◁ Let $A \in$ wClConv. Then $\partial s A = A$, since the indicator functions of both sides are the

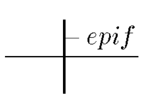
$$\delta\partial s A \overset{h)}{=} lsA \overset{e)}{=} l^2\delta A \overset{a)}{=} \delta A$$

same:


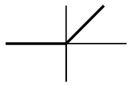
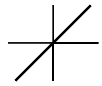
▷▷ ▷

Remarks.

1. The condition $f(x) \subset \mathbb{R}^{\bullet}$ in the theorem is essential. E. g., if $f(x) = \begin{cases} -\infty & \text{for } x = 0 \\ +\infty & \text{for } x \neq 0 \end{cases}$



then $f^* \equiv +\infty$ (see Important Remark on p. 55), so $f^{**} \equiv -\infty$, and $f \neq f^{**}$.

2. The condition $0 \in A$ in g) is essential. E. g., for $\mathbb{R} \leftrightarrow \mathbb{R}$, then $\mu\{1\}^{\triangledown} = \mu(-\infty, +1] =$ , but $s\{1\} =$ .

3. If $X \overset{\langle\, ,\, \rangle}{\longleftrightarrow} Y$, then $X \times \mathbb{R} \leftrightarrow Y \times \mathbb{R}$ with respect to the pairing

$$(\,(x, \alpha), (y, \beta)\,) \longmapsto \langle x, y \rangle + \alpha\beta$$

(that occurs in 1°), and the corresponding weak topology in $X \times \mathbb{R}$ coincides with the product topology of $(\,X, \sigma(X, Y)\,) \times \mathbb{R}$ (see Exercise below).

Exercise. (Product of dual pairs.) If $X_1 \overset{\langle\, ,\, \rangle_1}{\longleftrightarrow} Y_1$ and $X_2 \overset{\langle\, ,\, \rangle_2}{\longleftrightarrow} Y_2$, then

$$X_1 \times X_2 \overset{\langle\!\langle\, ,\, \rangle\!\rangle}{\longleftrightarrow} Y_1 \times Y_2,$$

where

$$\langle\!\langle (x_1, x_2), (y_1, y_2) \rangle\!\rangle := \langle x_1, y_1 \rangle_1 + \langle x_2, y_2 \rangle_2$$

Further

$$\sigma(\,X_1 \times X_2, Y_1 \times Y_2\,) = \sigma(X_1, Y_1) \times \sigma(X_2, Y_2).$$

2.3.4 Weak compactness of polars and subdifferentials

Here we apply the results of Sections 2.3.2 and 2.3.3 to the most interesting case of the dual pair $X \leftrightarrow X^*$, where X is a Hausdorff LCS.

Theorem on duality of boundedness and absorbingness for LCS's. *Let $X \in HLCS$, and let $A \subset X$, Then*

$$A \in Bdd(X) \iff A^{\circ} \in Absorb(X^*)$$

(where the polar is taken with respect to the dual pair $X \leftrightarrow X^*$).

◁ This follows from the theorem of Section 2.3.2 and the coincidence of the bounded sets in the original and in the weakened topology (theorem on p. 48, Part b)). ▷

Duality Theorem for LCS's *Let $X \in HLCS$. Then the theorem on p. 62 is true with Y replaced by X^* and wCl replaced by Cl.*

Here $f \in \mathrm{Cl} :\Leftrightarrow \mathrm{epi}\, f \in \mathrm{Cl}(X, \mathbb{R})$.

◁ This follows from Duality Theorem and the coincidence of the convex closed sets (and, hence, of the convex closed functions) for the original topology and for the weakened one (the theorem on p. 48, Part a)). ▷

Exercise. Prove, using this theorem, that $A^{\circ\circ} = \overline{\mathrm{co}I_1 A}$ for any $A \subset X$.

Now give one example of less direct application of Duality Theorem.

Theorem on compactness of polars and subdifferentials. *Let $X \in HLCS$.*

a) (Alaoglu-Bourbaki theorem.) Both U° and $U^{\triangledown}$ for every neighbourhood U of 0 in X are $\sigma(X^*, X)$-compact (that is, compact in the topology $\sigma(X^*, X)$.

b) The subdifferential ∂p of every continuous sublinear function p on X is a non-empty convex $\sigma(X^*, X)$-compact set.

◁ 0°

1) Duality Theorem (for LCS's);

2) The fact that for every balanced neighbourhood of 0 its Minkovski function (that is, the associated semi-norm) is continuous (see p. 27);

3) *Tichonov theorem*: the product of (any family of) compact spaces os a compact space.

1° At first we prove that b) $\Rightarrow$ a). Suppose that b) is already proved, and let U is a neighbourhood of 0 in X. Since U° is $\sigma(X^*, X)$-closed (see Section 2.3.1) and is contained in $U^\triangledown$, it is sufficient to show that $U^\triangledown$ is $\sigma(X^*, X)$-compact.

Since X is an LCS, there exists an open balanced convex neighbourhood V of 0 such that $V \in U$. The Minkovski function μV of this neighbourhood is continuous by $0^\circ 2$), and hence by b) its subdifferential $\partial \mu V$ is $\sigma(X^*, X)$-compact. But by $0^\circ 1$)

$$\partial \mu V = \pi V = V^\triangledown.$$

Thereby it is proved that $V^\triangledown$ is $\sigma(X^*, X)$-compact. It remains to notice that $U^\triangledown$ is contained in $V^\triangledown$ (because $U \supset V$) and is $\sigma(X^*, X)$-closed.

2° Now let us prove b). Let p be a continuous sublinear function on X, so that in particular p has only finite values. Then ∂p is convex (as every subdifferential) and is nonempty: if we had $\partial p = 0$ then we would have $\delta \partial p \equiv +\infty$, that is, by $0^\circ 1$), $lp \equiv +\infty$, and therefore $l^2 p = l(lp) \equiv -\infty$, but $l^2 p \overset{0^\circ 1)}{=} p \not\equiv -\infty$ ($0^\circ 1$) is applicable, since p is continuous and hence $p \in \mathrm{Cl}$). Thus, ∂p is a nonempty convex set.

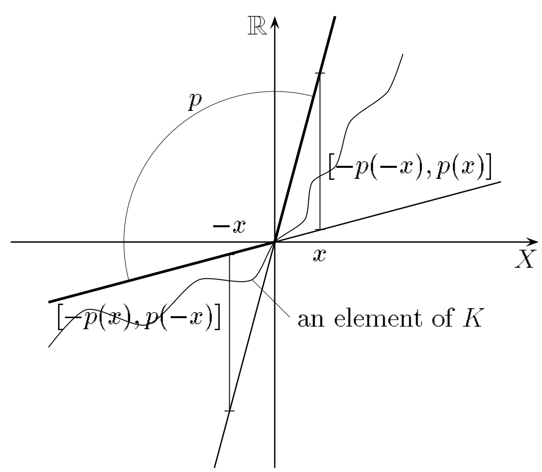
3° Put

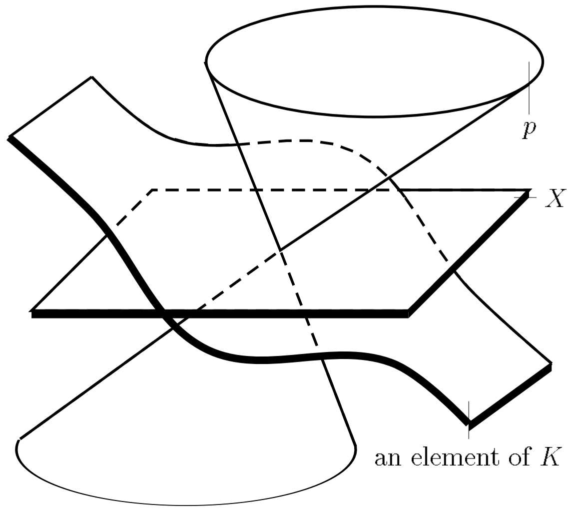
$$K := \prod_{x \in X} [-p(-x), p(x)]$$

(the product of segments $[-p(-x), p(x)]$ over all $x \in X$, with the product topology). By $0^\circ 3$) K is a compact space. (Notice that $\forall x \,\vdots\, -p(-x) \leq p(x)$. Why?)
We may consider K as a subspace of the space

$$\mathbb{R}^X := \mathcal{F}(X, \mathbb{R})$$

of all real-valued functions on X with the topology of simple (= pointwise) convergence (see p. 3). (Indeed a point of K is a family $\{t_x\}_{x \in X}$, $-p(-x) \leq t_x \leq p(x)$, and we can consider such a family as a function $X \to \mathbb{R}$, $x \mapsto t_x$.)





The elements of K are the functions on X such that their graphs lie between the graph of p and the symmetric cone.

4° The space X^* is a subset of $\mathcal{F}(X,\mathbb{R})$, consisting of all *continuous linear* functions on X, and the topology $\sigma(X^*,X)$ is evidently coincides with the topology, induced on X^* from $\mathcal{F}(X,\mathbb{R})$ (verify!). Further it is clear by the definition of the subdifferential that

$$\partial p = X^* \cap K$$

(the intersection being taken in $\mathcal{F}(X,\mathbb{R})$).

5° Now we claim that $X^* \cap K$ is *closed* in $\mathcal{F}(X,\mathbb{R})$. [NB X^* itself may be nonclosed in $\mathcal{F}(X,\mathbb{R})$! See Example on p. 42.] Indeed, let a function $f \in \mathcal{F}(X,\mathbb{R})$ belongs to the closure of $X^* \cap K$ in $\mathcal{F}(X,\mathbb{R})$. This means that $f = \lim f_\alpha$ in $\mathcal{F}(X,\mathbb{R})$, where $\{f_\alpha\}$ is a *net* of elements $f_\alpha \in X^* \cap K$. We need to show that $f \in X^* \cap K$, that is that f is a continuous linear functional on X, satisfying the condition

$$-p(-x) \le f(x) \le p(x). \tag{1}$$

By the definition of the topology of pointwise convergence we have

$$f(x) = \lim f_\alpha(x) \quad \forall x \in X.$$

So the fact, that f is linear, follows from linearity of f_α by passing to the limits:

$$\begin{aligned} f(x_1+x_2) &= \lim f_\alpha(x_1+x_2) = \lim\left(f_\alpha(x_1)+f_\alpha(x_2)\right) = \lim f_\alpha(x_1) + \lim f_\alpha(x_2) \\ &= f(x_1)+f(x_2), \\ f(tx) &= \lim f_\alpha(tx) = \lim\left(tf_\alpha(x)\right) = f \lim f_\alpha(x) = tf(x). \end{aligned}$$

Quite analogously the fact that f satisfies (1) follows at once from the fact that every f_α satisfies (1). At last continuity of f follows by (1) from continuity of p.

6° Thus our subdifferential $\partial p = X^* \cap K$ is a closed subset of the compact set K and hence is compact in the pointwise convergence topology. Since $\partial p \subset X^*$ and $\sigma(X^*,X)$ coincides with the topology, induced on X^* by the pointwise convergence topology, we conclude that ∂p is $\sigma(X^*,X)$-compact. ▷

Example. Let $X \leftrightarrow Y$ be a dual pair, and let $U \subset X$ be a neighbourhood of 0 in $\sigma(X,Y)$. Then $U \supset \{y_1,\dots,y_n\}^\circ$ for some $y_1,\dots,y_n \in Y$, and hence

$$U^\circ \subset \{y_1,\dots,y_n\}^{\circ\circ} \overset{\text{Exc. on p. 67}}{=} \overline{\mathrm{co} I_1\{y_1,\dots,y_n\}}$$

(where the closure is taken in $\sigma(Y,X)$). But the set $\{y_1,\ldots,y_n\}$ lies in a *finite*-dimensional subspace of Y, and the topology, induced on this subspace by $\sigma(Y,X)$, coincides with its usual linear topology (which is the *unique* Hausdorff linear topology in a finite-dimensional vector space). So $\sigma(Y,X)$-compactness of U° is evident here: it is simply compactness of a closed bounded set in finite-dimensional space.

3 Normed spaces

Normed spaces (NS) are a very special and a very important case of Hausdorff LCS's.

3.1 Preliminaries

We have already dealt with norms (see p. 26). Recall that a norm $\| \ \|$ is a *finite* semi-norm such that $\|x\| = 0 \Rightarrow x = 0$. For convinience we repeat the definition in an explicit form.

3.1.1 Definitions

A *norm* on a vector space X is a functional

$$\| \ \| : X \longrightarrow \mathbb{R},$$

that satisfies the following three conditions (”axioms of norm”):

a) $\|x\| \geq 0 \ \forall x \in X$, and $\|x\| = 0 \Rightarrow x = 0$ (positivity and nondegenerary);

b) $\|tx\| = |t| \, \|x\| \ \forall t \in \mathbb{R} \, \forall x \in X$; in particular, $\|-x\| = \|x\|$ (positive homogeneity and symmetry);

c) $\|x + y\| \leq \|x\| + \|y\| \ \forall x, y \in X$ (subadditivity = sublinearity = convexness).

A *normed space* is a vector space, equipped with a norm. It one want to emphasize that a norm is defined just on X, one writes $\| \ \|_X$ or $\| \ \|^X$.

It is easy seen, that the formula

$$\varrho(x, y) := \|x - y\|$$

defines a *metric* on X, which is invariant (relative to translations). The topology generated by this metric coincides evidently with the (*locally convex Hausdorff*) topology generated by the norm. This topology has the *closed balls*

$$B_r := \{x \,|\, \|x\| \leq r\}$$

as a base of closed balanced convex neighbourhoods of 0, and has the *open balls*

$$\mathring{B}B_r := \{x \,|\, \|x\| < r\}$$

as a base of open balanced convex neighbourhoods of 0.

The balls B_1 and $\mathring{B}B_1$ are called, naturaly, the (*closed unit ball* and the *open unit ball*, resp.

It is clear that for any net (generalized sequence) $\{x_\alpha\}$ in a NS X we have

$$x_\alpha \longrightarrow x \Longleftrightarrow \|x_\alpha - x\| \longrightarrow 0.$$

The most important class of NS's are the Banach spaces. A *Banach space* (BS) is a NS, that is complete with respect to the metric generated by the norm.

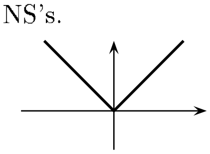
Clearly, a sequence $\{x_n\}$ in a NS is a *Cauchy sequence* iff

$$\forall \varepsilon > 0 \ \exists N \in \mathbb{N} : n \geq N, \, m \geq N \Longrightarrow \|x_n - x_m\| \leq \varepsilon.$$

3.1.2 Examples

Here are some basic examples of finite-dimensional and infinite-dimensional NS's.

1. On the real line $\mathbb{R}$ there exists the unique (up to a scalar factor) norm, viz. $|x|$. The corresponding topology is the usual topology of $\mathbb{R}$.



2. On $\mathbb{R}^n$ there are, among many other, such norms:

$$\|x\|_1 := \sum_{i=1}^{n} |x_i| \, ,$$

and, more generally,

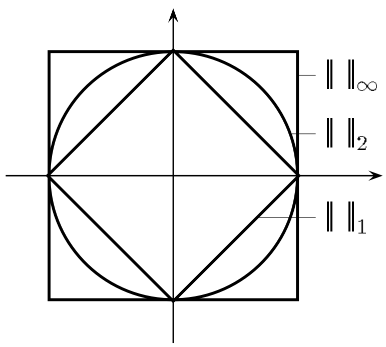
$$\|x\|_p := \left(\sum |x_i|^p\right)^{\frac{1}{p}} \quad (1 \le p < \infty),$$

$$\|x\|_\infty := \max_{1\le i\le n} |x_i|\,.$$

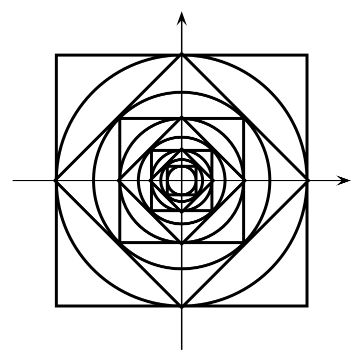
In the most important case $p = 2$ we obtain the well-familiar *Euclidean norm*

$$\|x\|_2^2 = \sum |x_i|^2\,.$$

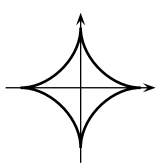
The unit balls in these norms are of the following form (in, say, $\mathbb{R}^2$):



We see that a ball may be *nonsmooth*! Notice, that all these norms generate one and the same topology in $\mathbb{R}^n$ (that is, are *equivalent*, see Section 3.4.1):

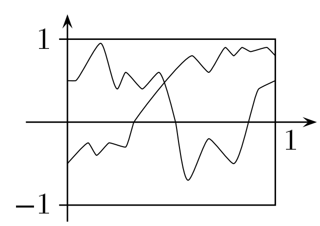


Exercise. Verify that, for $0 < p < 1$, $n \ge 2$, $\| \ \|_p$ is *not* a norm (though the corresponding "balls" form a base of neighbourhoods of 0 of the usual topology in $\mathbb{R}^n$.



3. $C([0,1])$. The space of all continuous functions $x : [0,1] \to \mathbb{R}$, with the norm

$$\|x\|_0 := \max_{t\in[0,1]} |x(t)|\,.$$



4. $C^r([0,1])$. The space of all r times continuously differentiable functions $[0,1] \to \mathbb{R}$, with the norm

$$\|x\|_r := \max\left(\|x\|_0, \|x'\|_0, \ldots, \left\|x^{(r)}\right\|_0\right).$$

5. l_2. The space of all sequences $x = (x_1, x_2, \ldots)$, $x_i \in \mathbb{R}$, such that $\sum x_i^2 < \infty$, which the norm

$$\|x\|_2^2 := \sum x_i^2.$$

It is a classical example of so-called *Hilbert space.*

6. l_p. More generally, for any $1 \le p < \infty$, the space of all real sequences such that $\sum |x_i|^p < \infty$, with the norm

$$\|x\|_p := \left(\sum |x_i|^p\right)^{\frac{1}{p}}.$$

7. l_∞. The space of all *bounded* real sequences, with the norm

$$\|x\|_\infty := \sup_{i\in\mathbb{N}} |x_i| .$$

8. c_0. The space of all *hull* sequences (that is, convergent to 0), with the same norm, as in .

Exercise. Prove that in all these examples we have BS's.

3.1.3 Subspaces and products of normed spaces

If X is a NS, and $Y \Subset X$, then we can equip Y by the *induced norm*

$$\| \ \|_Y := \| \ \|_X \big|_Y .$$

If X and Y are NS's, we can equip $X \times Y$ by the norm

$$\|(x,y)\|_p := \|(\ \|x\|, \|y\|\ )\|_p^{\mathbb{R}^2} ,$$

for any $1 \le p \le \infty$.

It is easy seen, that, in both cases, all the axioms of norm are fulfilled, and that, in the product case, all the norms generate one and the same topology, viz. the product topology.

Exercises.

1. The *closed* vector subspace of a BS, with the induced norm, is again a BS.
2. The product of two BS's, with any of the norms, introduced above, is again a BS.

3.1.4 Continuous mapping in NS's

The definition of continuous mapping for TVS's, "translated" into the language of norms, looks as follows: a mappping $f : Y \to Y$ (X, Y being NS's) is *continuous* at $\hat{x} \in X$ if

$$\forall \varepsilon > 0 \ \exists \delta > 0 \ : \ \|x - \hat{x}\| \le \delta \Longrightarrow \|f(x) - f(\hat{x})\| \le \varepsilon.$$

Examples of continuous mappings.

1. The norm $\| \ \| : X \to \mathbb{R}$ itself, for any NS X.
2. As for every TVS, the arithmetical operations in a given NS.
3. *Every linear* mapping $\mathbb{R}^n \to \mathbb{R}^m$ (with respect to any norm $\| \ \|_p$, $1 \le p \le \infty$, in $\mathbb{R}^n$ and any norm $\| \ \|_q$, $1 \le p \le \infty$, in $\mathbb{R}^m$).
4. The linear functional
$$\delta_\tau \ : \ C(\,[0,1]\,) \longrightarrow \mathbb{R}, \ x \longmapsto x(\tau),$$
$\tau \in [0,1]$ being fixed. This functional is called *Dirac δ-function* concentrated at a point τ. It is an example of so-called *generalized functions* (or *distributions*).
5. Linear functionals of the form
$$l : C(\,[0,1]\,) \longrightarrow \mathbb{R}, \quad x \longmapsto \int_0^1 x(t)y(t)dt,$$
$y \in C(\,[0,1]\,)$ being fixed.
6. Linear operators of the form
$$A : C(\,[0,1]\,) \longrightarrow C(\,[0,1]\,), \quad x \longmapsto Ax,$$
$$Ax(t) := \int_0^1 K(s,t)x(s)ds,$$
$K \in C(\,[0,1] \times [0,1]\,)$ being fixed (*integral operators* with the *kernel* K).

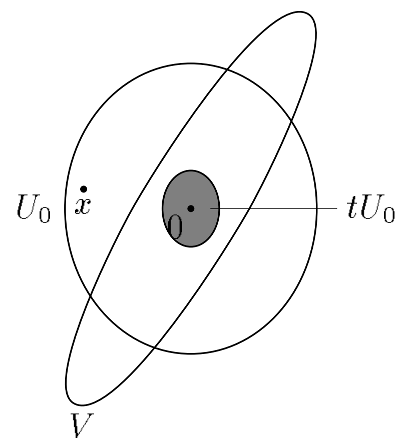
3.1.5 The place of NS's among TVS's

A TVS is called *normable* if its topology is generated by a norm. Since each NS is (as a TVS) Hausdorff and locally convex, it is clear that normable spaces are to be looked for among Hausdorff LCS's.

Kolmogorov's theorem. *Let X be a Hausdorff LCS. Then X is normable iff there exists a bounded neighbourhood of* 0.

$\triangleleft$ 0°
1) Theorem on continuous semi-norms (p. 27);
2) the fact that in any LCS there exists a base of convex balanced closed neighbourhoods of 0.

1° The part "only if" is trivial: the unit ball in the corresponding NS is such a neighbourhood of 0.



2° The part "if": Let U be a bounded neighbourhood of 0. Then, by $0^\circ 2$), $\exists U_0 : U_0 \subset U,\ U_0 \in \mathrm{BalConvCl}Nb_0$. (This U_0 is *bounded* together with U.) By $0^\circ 1$), the corresponding semi-norm p_{U_0} is continuous. We claim that p_{U_0} is a norm. Indeed, p_{U_0} is finite as any continuous function. Let us show that $x \neq 0 \Rightarrow p_{U_0}(x) \neq 0$:

$$x \neq 0 \overset{X \text{ is Hausd.}}{\Longrightarrow} \exists V \in Nb_0 : x \notin V \overset{U_0 \in \mathrm{Bdd}}{\Longrightarrow} \exists t > 0 : tU_0 \subset V$$

$$\Longrightarrow x \notin tU_0 \overset{\text{def. of } \mu}{\Longrightarrow} \underbrace{\mu U_0}_{p_{U_0}}(x) \geq t > 0 \Longrightarrow p_{U_0}(x) \neq 0.$$

Thus, p_{U_0} is a norm.

3° Now we claim that the topology $\tau(p_{U_0})$ generated by p_{U_0} coincides with the original topology τ of X. Indeed, by $0^\circ 1$),

$$B_{p_{U_0}} = U_0,$$

hence, the balls $\varepsilon U_0,\ \varepsilon > 0$, form a base of neighbourhoods of 0 for $\tau(p_{U_0})$. But they form also a base of neighbourhoods 0 for τ: $V \in Nb_0(\tau) \overset{U_0 \in \mathrm{Bdd}}{\Longrightarrow} \exists \varepsilon > 0 : \varepsilon U_0 \subset V$. Thus, $\tau = \tau(p_{U_0})$. $\triangleright$

So in a *non*normable Hausdorff LCS any neighbourhood of 0 is *unbounded*.

3.2 The normed space $\mathcal{L}(X, Y)$

Here we show that for NS's X and Y the space $\mathcal{L}(X,Y)$ has a natural structure of a NS, and we study this structure.

3.2.1 Characterization of bounded sets and bounded operators in NS's

Since NS's are *metrizable*, we know from the theorem on boundedness and continuity (p. 38), that a linear mapping between NS's is continuous iff it is bounded. More precisely, it holds

Lemma. *Let $X, Y \in NS,\ B \subset X,\ A \in L(X,Y)$. Then:*

a) $B \in Bdd(X) \Leftrightarrow \|B\| \in Bdd(\mathbb{R}) \Leftrightarrow \exists c > 0\ \forall x \in B : \|x\| \leq c$; *that is, a set in a NS is bounded (with respect to the generated topology) iff it is bounded in norm.*

b)

$$A \in \mathcal{L}(X,Y) \Longleftrightarrow a \in Bdd(X,Y) \Longleftrightarrow A(B_1(X)) \in Bdd(Y)$$
$$\Longleftrightarrow \exists c > 0 : \|x\| \leq 1 \Longrightarrow \|Ax\| \leq c.$$

(See Corollary and Example on p. 38.)

3.2.2 Norm of an operator

By the assertion b) of the lemma above, we can put for any $A \in \mathcal{L}(X,Y)$ $(X, Y \in \mathrm{NS})$

$$\|A\| := \sup_{\|x\| \leq 1} \|Ax\|. \tag{1}$$

Exercises.

1. Verify that the defined function $\| \ \| : \mathcal{L}(X,Y) \to \mathbb{R}$ is a *norm* (called the *operator norm*).
2. Verify that "$\|x\| \leq 1$" in (1) may be replaced by "$\|x\| = 1$", that is
$$\|A\| = \sup_{\|x\|=1} \|Ax\|. \tag{2}$$
3. Prove that for all $A \in \mathcal{L}(X,Y)$ and all $x \in X$
$$\|Ax\| \leq \|A\|\|x\|. \tag{3}$$
This inequality and analogous one (see below) will be used very often. We shall relate to all of them as to *basic inequalities for norms.*

Thus, for every normed spaces X and Y we obtain a *normed* space $\mathcal{L}(X,Y)$. If the contrary is not specified, we *ever* assume that $\mathcal{L}(X,Y)$ is equipped with the norm (2).
Basic inequality for norms of operators. Let X, Y, Z be NS's, and let $A \in \mathcal{L}(X,Y)$, $B \in \mathcal{L}(Y,Z)$:

$$X \xrightarrow{A} Y \xrightarrow{B} Z.$$

Then the norm of the (evidently continuous) composition $BA := B \circ A$ satisfies the estimate

$$\|BA\| \leq \|B\|\|A\|.$$

$\lhd$ $\forall x \in X$,

$$\|(BA)x\| = \|B(Ax)\| \overset{(3)}{\leq} \|B\|\|Ax\| \overset{(3)}{\leq} \|B\|\|A\|\|x\|,$$

hence

$$\|BA\| = \sup_{\|x\|\leq 1} \|(BA)x\| \leq \|B\|\|A\|. \ \rhd$$

3.2.3 Completeness of $\mathcal{L}(X,Y)$

We know that $\mathcal{L}(X,Y)$ equipped with the "topology of pointwise convergence" may be "non-complete", if X is not complete (see p. 42). As to the space $\mathcal{L}(X,Y)$ equipped with *norm* topology, it is the completeness of Y that plays the key role, that is, defines the completeness of $\mathcal{L}(X,Y)$:
Theorem on completeness of $\mathcal{L}(X,Y)$. *Let X be a NS, and let Y be a BS. Then the space $\mathcal{L}(X,Y)$ (with the operator norm) is complete (that is, is a BS).*

$\lhd$ 0▷
1) Basic inequality for norms;
2) continuity of arithmetical operations.

1▷ Let $\{A_n\}$ be a Cauchy sequence in $\mathcal{L}(X,Y)$, that is,

$$\forall \varepsilon > 0 \ \exists n_\varepsilon \ : n \geq n_\varepsilon, \ m \geq n_\varepsilon \Longrightarrow \|A_n - A_m\| \leq \varepsilon. \tag{1}$$

We need show that there exists $A \in \mathcal{L}(X,Y)$ such that

$$\|A_n - A\| \overset{n\to\infty}{\longrightarrow} 0. \tag{2}$$

2° For every fixed $x \in X$ the sequence $\{A_n x\}$ is a Cauchy sequence in Y. In view of (1), this follows from the estimate

$$\|A_n x - A_m x\| = \|(A_n - A_m)x\| \overset{0^\circ 1)}{\leq} \|A_n - A_m\|\|x\|. \tag{3}$$

Since Y is complete, $\exists \lim A_n x =: Ax$. We claim that so defined A is what we need.
3° $A \in L(X, Y)$. Indeed, for all $x, y \in X$ and all $\alpha, \beta \in \mathbb{R}$

$$\begin{aligned} A(\alpha x + \beta y) &= \lim (\, A_n(\alpha x + \beta y)\,) = \lim (\, \alpha A_n x + \beta A_n y\,) \overset{0^\circ 2)}{=} \alpha \lim A_n x + \beta \lim A_n y \\ &= \alpha A x + \beta A y. \end{aligned}$$

4°

$$\forall m \geq n_\varepsilon \,\vdots\, A - A_m \in \mathcal{L}(X, Y) \text{ and } \|A - A_m\| \leq \varepsilon. \tag{4}$$

Indeed, the mapping $y \mapsto \|y - A_m x\|$, $Y \to \mathbb{R}$ is continuous, for any fixed n and x, as the composition of translation by $-A_m x$ and $\|\ \|$, which are both continuous mappings. Since $A_n x \to Ax$ in Y (for fixed x), we conclude, that

$$\|A_n x - A_m x\| \longrightarrow \|Ax - A_m x\| = \|(\, A - A_m\,)x\|.$$

Hence, for $x \in B_1(X)$ and $m \geq n_\varepsilon$, it holds

$$\|(\, A - A_m\,)x\| = \lim_{n\to\infty} \underbrace{\|A_n x - A_m x\|}_{\overset{(3)}{\leq} \|A_n - A_m\| \underbrace{\|x\|}_{\leq 1}} \overset{(1)}{\leq} \varepsilon,$$

that is $A - A_m$ is bounded on $B_1(X)$ and, therefore, is continuous, and

$$\|A - A_m\| = \sup_{\|x\|\leq 1} \|(\, A - A_m\,)x\| \leq \varepsilon.$$

5° Since $A = A_m + (\, A - A_m\,)$, it follows from the first assertion (4) that $A \in \mathcal{L}(X, Y)$. Then the second assertion (4) means that (2) is true. $\triangleright$

3.2.4 Dual normed space

Applying the results of Sections 3.2.3 and 3.2.4 to the case $Y = \mathbb{R}$, we obtain that the topological dual space

$$X^* = \mathcal{L}(X, \mathbb{R})$$

to every normed space X may be equipped with the norm

$$\|x^*\| := \sup_{\|x\|\leq 1} |\langle x^*, x\rangle| \overset{!}{=} \sup_{\|x\|\leq 1} \langle x^*, x\rangle = \sup \langle x^*, B_1\rangle = sB_1(x^*), \tag{1}$$

and is *Banach* space with respect to this norm (since $\mathbb{R}$ is complete!).

If the contrary is not specified, we assume ever that X^* is the Banach space with the norm (1).

Exercise. Prove the equation, marked with "!" in (1). [Use the fact that B_1 is balanced.]

The basic inequality for norm takes in our case the form

$$|\langle x^*, x\rangle| \leq \|x^*\|\|x\|. \tag{2}$$

Examples.

1. For $1 \le p \le \infty$, it holds $\left(\mathbb{R}^n, \|\ \|_p\right)^* = \left(\mathbb{R}^n, \|\ \|_q\right)$, where q is defined by the equation

$$\frac{1}{p} + \frac{1}{q} = 1. \tag{3}$$

   In particular, for $p = 2$ the normed dual space *coincides* with the original one.

2. For $1 \le p \le \infty$, it holds $l_p^* := (l_p)^* = l_q$, where q is again defined by (3). As to $p = \infty$, the normed dual space to l_∞ is the space ba of "finitely additive measures" on $\mathbb{N}$ (see [3, p. 261]); note that l_1 may be considered as the space of (countably additive) measures on $\mathbb{N}$.

3. $C([0,1])^*$ may be identified with the space of all *measures* μ on $[0,1]$:

$$\langle \mu, x \rangle := \int_0^1 x d\mu \quad \text{(the integral of } x \text{ with respect to the measure } \mu),$$

$$\|\mu\| = \operatorname{Var} \mu \quad \text{(the total variation of } \mu).$$

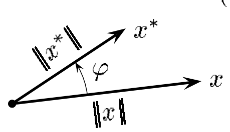
   In particular, Dirac delta-function δ_τ "is" the measure of the unit mass ($\|\delta_\tau\| = 1$), "concentrated" at τ.

3.3 Applications of Hahn-Banach theorem in NS's

Here we derive some special "normed space" corollaries of the "first whole" of functional analysis.

3.3.1 Banach lemma on zero angle

In the case $X = (\mathbb{R}^n, \|\ \|_2)$ the basic inequality for norms (2) on p. 76 has a simple geometrical interpretation: the modulus of the inner product (scalar product) of two vectors is less or equal to the product of their lengths ($|\cos\varphi| \le 1$, see the picture).



$\langle x^*, x \rangle = \|x^*\|\,\|x\| cos\varphi$

For arbitrary NS's Equation (2) may be used for *defining* cosine of the "angle" φ between a vector x^* in X^* and a vector x in X (let these vectors lie in different spaces!):

$$\cos\varphi(x^*, x) := \frac{\langle x^*, x \rangle}{\|x^*\|\|x\|} \quad (x^* \ne 0, x \ne 0).$$

In particular, we say that x^* and x are *orthogonal* if $\cos\varphi(x^*, x) = 0$:

$$x^* \perp x :\Longleftrightarrow \langle x^*, x \rangle = 0,$$

and we say that x^* and x *have the same direction* if $\cos\varphi(x^*, x) = 1$:

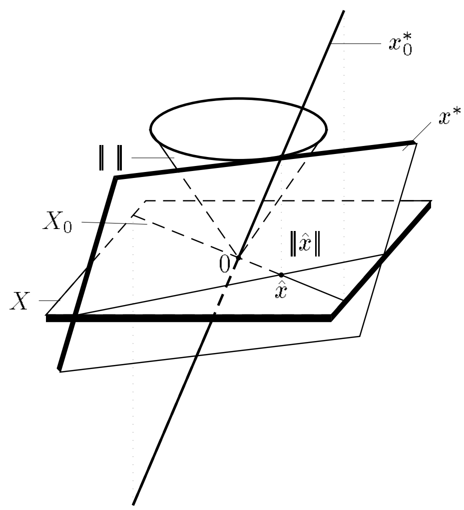
$$x^* \uparrow\uparrow x :\Longleftrightarrow \langle x^*, x \rangle = \|x^*\|\|x\|.$$

In the finite-dimensional case ($X = \mathbb{R}^n$) for every vector $x \in \mathbb{R}^n$ there exist ever a vector $x^* \in (\mathbb{R}^n)^* = \mathbb{R}^n$ that has the same direction as x; we can just take $x^* = x$. In the case of infinite-dimensional NS's, this fact remains valid (although x and x^* lie now in *different* spaces!), but it is by no means trivial.

Banach lemma on zero angle. *Let X be a NS. Then for every $\hat{x} \in X$ there exists a vector $x^* \in X^* \setminus 0$ that has the same direction as x. An equivalent statement is: for every $\hat{x} \in X$ there exists $x^* \in X^*$ such that*

$$\|x^*\| = 1 \text{ and } \langle x^*, \hat{x}\rangle = \|\hat{x}\|.$$

$\triangleleft$ 0° Hahn-Banach theorem.
1° First of all, the equivalence of our two statement follows from the fact that if $x^* \neq 0$ and $x^* \uparrow\uparrow x$ then $\frac{x^*}{\|x^*\|}$ satisfies:



2° If $\hat{x} = 0$ then any $x^* \neq 0$ fits. Let $\hat{x} \neq 0$. Put

$$X_0 := \operatorname{lin}\{\hat{x}\} = \{t\hat{x} \big| t \in \mathbb{R}\},$$

and define a linear functional x_0^* on X_0 by putting

$$\langle x_0^*, \hat{x}\rangle := \|\hat{x}\| \tag{1}$$

(it is clear that x_0^* is uniquely defined by its value at $\hat{x}$). We have, evidently,

$$x_0^* \leq := \| \ \| \big|_{X_0}.$$

So, by 0°, $\exists x^* \in X'$ such that

$$x^*\big|_{X_0} = x_0^*, \tag{2}$$

$$x^* \leq \| \ \|. \tag{3}$$

We claim that this x^* is what we need.
3° We have

$$\sup_{\|x\|\leq 1} \langle x^*, x\rangle \stackrel{(3)}{\leq} \sup_{\|x\|\leq 1} \|x\| = 1,$$

so x^* is bounded by 1 on the unit ball and hence is continuous, and $\|x^*\| \leq 1$. But at the point $\frac{\hat{x}}{\|\hat{x}\|} \in B_1$ our x^* has the value 1:

$$\left\langle x^*, \frac{\hat{x}}{\|\hat{x}\|} \right\rangle \stackrel{(2)}{=} \left\langle x_0^*, \frac{\hat{x}}{\|\hat{x}\|} \right\rangle \stackrel{(1)}{=} 1, \tag{4}$$

so

$$\|x^*\| = \sup_{x \in B_1} \langle x^*, x \rangle = 1.$$

At last,

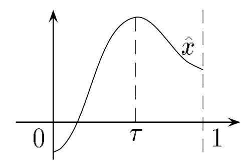
$$\langle x^*, \hat{x} \rangle \overset{(4)}{=} \|\hat{x}\|.$$

$\triangleright$

Exercise. Prove by analogous argument the following *theorem on extension with presentation of norm*:
Let X be a NS, let $X_0 \Subset X$, and let x_0^* be a continuous linear functional on X_0 (equipped with the induced norm). Then there exists $x^* \in X^*$ such that

$$x^*\Big|_{X_0} = x_0^* \quad \text{and} \quad \|x^*\| = \|x_0^*\|.$$

Example. Let $X = C(\,[0,1]\,)$. For any $\hat{x}$ from this space we have



$$\hat{x} \uparrow\uparrow \delta_\tau \quad \text{or} \quad \hat{x} \uparrow\uparrow (-\delta_\tau),$$

where τ is any of the points where $|\hat{x}|$ attains its maximal value.
Indeed,

$$\langle \delta_\tau, \hat{x} \rangle = \hat{x}(\tau) = \pm\|x\|$$

(recall that $\|\delta_\tau\| = 1$). We see that for, say, $\hat{x} \equiv 1$ there exists a *whole continuum* $\{\delta_\tau\}_{\tau \in [0,1]}$ of (linearly independent) vectors with the same direction!

3.3.2 Canonical imbedding $X \hookrightarrow x^{**}$

Let X be a NS. Since $X \leftrightarrow X^*$, we know that

$$X \Subset (\,X^*\,)' \tag{1}$$

(see p. 2.1.2). But we have more in this case:
Theorem on the canonical imbedding. *Let $X \in NS$. Denote the imbedding (1) by i:*

$$i : X \longrightarrow (\,X^*\,)', \ \langle ix, x^* \rangle = \langle x, x^* \rangle \tag{2}$$

(the first pairing corresponding to $(\,X^*\,)' \leftrightarrow X^*$, the second one corresponding to $X \leftrightarrow X^*$). Then
a) $iX \subset X^{**} := (\,X^*\,)^*$,

b) $i : X \to X^{**}$ is an isometry, that is,

$$\|ix\|_{X^{**}} = \|x\|_X \quad \forall x \in X. \tag{3}$$

Recall that X^* denotes the *normed* dual space to X, and X^{**} is the *normed* dual space to this X^*.
$\triangleleft$ 0°
1) Basic inequality for norms;
2) Banach lemma on zero angle.
1° a): Let $x \in X$. We have for any $x^* \in B_1(X^*)$

$$|\langle ix, x^* \rangle| \overset{(2)}{=} |\langle x, x^* \rangle| \overset{0°1)}{\leq} \|x\| \underbrace{\|x^*\|}_{\leq 1} \leq \|x\|, \tag{4}$$

that is, ix is bounded (by $\|x\|$) on $B_1(X^*)$, and hence $ix \in X^{**}$.
2° b): By $0^\circ 2$), applied to our x, $\exists \hat{x}^* \in X^* \, : \, \|\hat{x}^*\| = 1$ and

$$\langle x, \hat{x}^* \rangle = \|x\|. \tag{5}$$

From (4) and (5) follows that

$$\sup_{\|x^*\| \leq 1} |\langle ix, x^* \rangle| = \|x\|,$$

that is $\|ix\|_{X^{**}} = \|x\|$. ▷

3.3.3 Reflexive NS's

By the theorem on the canonical imbedding we can consider any NS X as a *subspace* of the NS X^{**}. A NS X is called *reflexive* if $X^{**} = X$ (that is, if each continuous linear functional on, X^* is "generated" by some element of X):

$$N \in \text{Refl}\, NS :\Longleftrightarrow X^{**} = X.$$

Remarks.

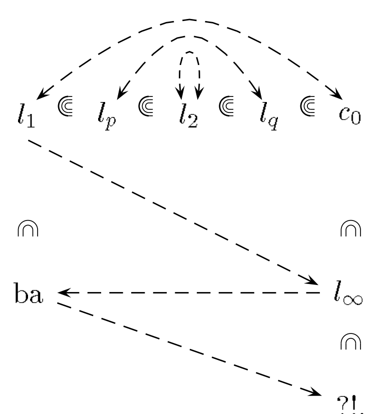
1. Every reflexive NS is a *Banach* space. Indeed X^{**} is a BS as every normed dual space (see p. 87).
2. In the case of general TVS's one distinguishes reflexive and so called "semireflexive" spaces. For NS's these notions coincide. (See [4, p. 693].)
3. For a reflexive NS X the *weak* topology in X^*, associated with the duality $X \leftrightarrow X^*$, coincides with the *weakened* topology of X^*:

   $$\sigma(X^*, X) = \sigma(X^*, X^{**}$$

   (just since $X = X^{**}$).
4. If X is a reflexive NS then X^* is also a reflexive NS (indeed, $(\, X^* \,)^{**} = (\, X^{**} \,)^* = X^*$).

Examples.

1. The spaces $l_p, \; 1 < p < \infty$, are reflexive (see p. 88).
2. The spaces c_0 and l_∞ are *not* reflexive (see the picture where dotted lines denote the passage to the normed dual space). (See [3, p. 408], [6, p. 203].)



$\Subset$ denotes dense linear imbedding

$(1 < p < 2)$, $(2 < p < \infty)$

Exercises.

1. A NS X is reflexive iff the topology of X^* (as a NS) is compatible with the duality $X \leftrightarrow X^*$.
2. Prove that
$$X \Subset_{\to} Y \Longrightarrow Y^* \Subset_{\to} X^*$$

[Here $X, Y \in NS$, X being not assumed to be equipped with the norm induced from Y, $\Subset_{\to}$ denotes *continuous* dense linear imbedding, and $\subset_{\to}$ denotes *continuous* linear imbedding. (Recall that a *dense imbedding* is an imbedding with a dense image.) For example, $p \le q \Rightarrow l_p \Subset_{\to} l_q$.]

3.3.4 Dual operators

The notion of duality may be extended onto *operators*. Let $X_1 \overset{\langle\,,\,\rangle_1}{\longleftrightarrow} Y_1$ and $X_2 \overset{\langle\,,\,\rangle_2}{\longleftrightarrow} Y_2$. We say that a linear operator $A : X_1 \to X_2$ *agrees* with the dualities $\langle\ ,\ \rangle_1$ and $\langle\ ,\ \rangle_2$, if $\forall y_2 \in Y_2$ the functional

$$x_1 \longmapsto \langle y_2, AX_1\rangle, X_1 \longrightarrow \mathbb{R}$$

(which is evidently linear) is given by some element of Y_1 (recall that $Y_1 \Subset X_1'$); such an element, if exists, must be *unique* by the totality properties of dual pairs (verify!). So the formula

$$\langle A^* y_2, x_1\rangle_1 := \langle y_2, Ax_1\rangle_2 \quad (x_1 \in X_1, y_2 \in Y_2)$$

defines in this case an operator from Y_2 into Y_1, which is obviously linear. This operator A^* is called *dual operator* to A. Emphasize that A^* acts in the opposite direction as compared with A.

Exercises.

1. Every $A \in L(X,Y)$ (X and Y being arbitrary vector spaces) agrees with $X \leftrightarrow X'$ and $Y \leftrightarrow Y'$, so we obtain
$$A^* : Y' \longrightarrow X'$$
(the *algebraically dual* operator).
2. Every $A \in \mathcal{L}(X,Y)$ (X and Y being arbitrary Hausdorff LCS's) agrees with $X \leftrightarrow X^*$ and $Y \leftrightarrow Y^*$, so we obtain
$$A^* : Y^* \longrightarrow X^*$$
(the *topologically dual* operator).

Example. If $A \in L(\mathbb{R}^n, \mathbb{R}^m) = \mathcal{L}(\mathbb{R}^n, \mathbb{R}^m)$ is an operator defined by a matrix, then $A^* \in L(\mathbb{R}^m, \mathbb{R}^n) = \mathcal{L}(\mathbb{R}^m, \mathbb{R}^n)$ is the operator defined by the *transpose* of this matrix.

Remark. It is clear that operation $A \mapsto A^*$ is *linear* (when defined): $(\alpha A + \beta B)^* = \alpha A^* + \beta B^*$. It is also easy to see that this operation changes order of composed mappings: $(A \circ B)^* = B^* \circ A^*$.

Theorem. *Let X and Y be NS's, and let $A \in \mathcal{L}(X,Y)$. Then*

$$A^* \in \mathcal{L}(Y^*, X^*) \quad \textit{and} \quad \|A^*\| = \|A\|.$$

$\lhd$ 0°
1) The basic inequality for norms;
2) Banach lemma on zero angle.

1° For any $y^* \in B_1(Y^*)$ and $x \in B_1(X)$ we have

$$|\langle A^* y^*, x\rangle| \overset{\text{def.}}{=} |\langle y^*, Ax\rangle| \overset{0^\circ 1)}{\le} \underbrace{\|y^*\|}_{\le 1} \|Ax\| \overset{0^\circ 1)}{\le} \|A\| \underbrace{\|x\|}_{\le 1} \le \|A\|.$$

It follows that $\forall y^* \in B_1(Y^*)$ $\|A^*y^*\| \leq \|A\|$, that is A^* is bounded on $B_1(Y^*)$ and hence is continuous, and, moreover, that

$$\|A^*\| = \sup_{\|y^*\|\leq 1} \|A^*y^*\| \leq \|A\|.$$

2° Show that, v. v., $\|A\| \leq \|A^*\|$. We need verify that $\forall x \in B_1(X)$ $\|Ax\| \leq \|A^*\|$. By 0°2) *applied to* Ax, $\exists y^* \in Y^*$ such that

$$\|y^*\| = 1 \quad \text{and} \quad \langle y^*, Ax\rangle = \|Ax\|.$$

Hence, indeed,

$$\|Ax\| = \langle y^*, Ax\rangle \overset{\text{def.}}{=} \langle A^*y^*, x\rangle \overset{0^\circ 1)}{\leq} \|A^*\| \underbrace{\|y^*\|}_{=1} \underbrace{\|x\|}_{\leq 1} \leq \|A^*\|.$$

▷

3.4 Applications of Openness Principle in NS's

Since each Banach space is a Fréchet space, Banach theorems on open mapping and on inverse mapping and the theorem on closed graph are true for BS's.

Here we give two specifically "normed space" corollaries of the mentioned Banach theorems.

3.4.1 Equivalent norms

Let we have two norms $\| \ \|_1$ and $\| \ \|_2$ on a vector space. We say that $\| \ \|_1$ is *stronger* than $\| \ \|_2$ and write

$$\| \ \|_1 \succeq \| \ \|_2$$

if the topology generated by $\| \ \|_1$ is stronger than the one generated by $\| \ \|_2$. We say that $\| \ \|_1$ and $\| \ \|_2$ are *equivalent* and write

$$\| \ \|_1 \sim \| \ \|_2$$

if they generate the same topology.

Exercises.

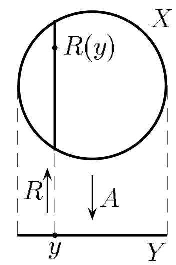
1. $\| \ \|_1 \succeq \| \ \|_2 \Leftrightarrow \exists \alpha > 0 : \| \ \|_1 \geq \alpha\| \ \|_2 \Leftrightarrow B_1(\| \ \|_1) \in \mathrm{Bdd}(\| \ \|_2)$.
2. $\| \ \|_1 \sim \| \ \|_2 \Leftrightarrow \exists \alpha, \beta, \gamma, \delta > 0 : \alpha\| \ \|_2 \leq \| \ \|_1 \leq \beta\| \ \|_2, \gamma\| \ \|_1 \leq \| \ \|_2 \leq \delta\| \ \|_1$.
3. In $\mathbb{R}^n$ all $\| \ \|_p$, $1 \leq p \leq \infty$, are equivalent.
4. In k (finite sequences) $p > q \Rightarrow \| \ \|_p \overset{\not\succeq}{\preceq} \| \ \|_q \Leftrightarrow: \| \ \|_p \prec \| \ \|_q$.

Theorem on comparable norms. *Let* $\| \ \|_1$ *and* $\| \ \|_2$ *be two norms on a vector space* X*. If* $\| \ \|_1 \succeq \| \ \|_2$ *and* X *is Banach space with respect to each of these two norms, then* $\| \ \|_1 \sim \| \ \|_2$.

◁ 0° Banach theorem on inverse mapping.
1° The identity mapping id: $(X, \| \ \|_1) \to (X, \| \ \|_2)$ is bijective (evidently) and continuous (by the definition of a stronger norm). Hence, by 0°, this mapping is a homeomorphism. But this just means that topologies, generated by $\| \ \|_1$ and $\| \ \|_2$, coincide. ▷

This theorem is used usually in the following form (see exercises 1 and 2):
Let $\| \ \|_1$ *and* $\| \ \|_2$ *be two norms on* X*, such that* X *is BS with respect to each of them. If* $\exists \alpha > 0$ *such that* $\| \ \|_1 \geq \alpha\| \ \|_2$*, then* $\exists \beta > 0$ *such that* $\| \ \|_2 \geq \beta\| \ \|_1$.

3.4.2 Lemma on a right inverse operator



It is obvious that for every *surjective* mapping $A : X \to Y$ (where X and Y are arbitrary sets) there exists some *right inverse mapping* R, that is, a mapping $R : Y \to X$ such that $A \circ R = \mathrm{id}$ ($\Leftrightarrow A(\,R(y)\,) = y\ \forall y \in Y$).

The following result affirms, that for *Banach* spaces X and Y and a *continuous linear* mapping A such a mapping R can be chosen to be "bounbed" in a sense.

Lemma on a right inverse operator. *Let $X, Y \in BS$, $A \in \mathcal{L}(X, Y)$, and let A be surjective ($AX = Y$). Then there exist a mapping $R : Y \to X$ and a real number $\alpha > 0$ such that*

$$A \circ R = id \tag{1}$$

and

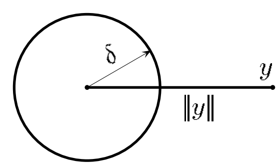
$$\|R(y)\| \le \alpha \|y\| \quad \forall y \in Y. \tag{2}$$

It should be emphasized that such an operator R is in general *neither linear nor even continuous* (but is, of course, continuous at 0). Condition (2) expresses the mentioned "boundedness in a sense".

$\lhd$ 0° Banach theorem on open mapping.

1° By 0°, A is open. Hence the image $A\mathring{B}B_1^x$ is open in Y. Since, evidently, 0 belongs to this image, we conclude that $\exists \delta > 0 \,:\, A\mathring{B}B_1^x \supset B_\delta^Y$, that is,

$$\forall y \in Y \,:\, \|y\| \le \delta,\ \exists x_y \in X \,:\, \underbrace{\|x_y\| < 1}_{(a)}, \underbrace{Ax_y = y}_{(b)}.$$



2° Fix any x_y for each y *with* $\|y\| = \delta$ and define R as the extension by linearity of this mapping on the rays, going out of 0:

$$R(y) := \begin{cases} 0 & \text{if } y = 0, \\ \frac{\|y\|}{\delta} x_{\delta \frac{y}{\|y\|}} & \text{if } y \ne 0. \end{cases}$$

3° This R is what we need. Indeed, $\forall y \in Y \setminus 0$

$$AR(y) = A\Big(\frac{\|y\|}{\delta} x_{\delta \frac{y}{\|y\|}} \Big) = \frac{\|y\|}{\delta} \underbrace{Ax_{\delta \frac{y}{\|y\|}}}_{\overset{(b)}{=} \delta \frac{y}{\|y\|}} = y,$$

$$\|R(y)\| = \Big\| \frac{\|y\|}{\delta} x_{\delta \frac{y}{\|y\|}} \Big\| = \frac{\|y\|}{\delta} \underbrace{\Big\| x_{\delta \frac{y}{\|y\|}} \Big\|}_{\overset{(a)}{<} 1} \le \underbrace{\frac{1}{\delta}}_{=:\alpha} \|y\|,$$

and for $y = 0$ all is also O.K. $\rhd$

3.5 Applications of Boundedness Principle in NS's

Here we at first reformulate Boundedness Principle as applied to NS's, and then we apply this principle for study of the weak(ened) topology in NS's.

3.5.1 Boundedness Principle for NS's

First of all recall onde more characterizations for NS's of all related notions (see p. 38, 39, and 74).

Characterizations of boundedness, equi-boundedness and equi-continuity in NS's. Let X and Y be NS's.

a) For a set $B \subset X$ the following conditions are equivalent:

1. B is bounded (in X considered as a TVS),
2. B is *bounded in norm*, i. e.

$$\|B\| \in \mathrm{Bdd}(\mathbb{R}) \Longleftrightarrow \exists c > 0 \,:\, x \in B \Longrightarrow \|x\| \leq c.$$

b) For an operator $A \in L(X, Y)$ the following conditions are equivalent:

1. A is continuous (that is, $A \in \mathcal{L}(X, Y)$),
2. A is bounded (as a mapping between two TVS's),
3. A is bounded on the unit ball, i. e.

$$AB_1(X) \in \mathrm{Bdd}(Y) \Longleftrightarrow \exists c > 0 \,:\, \|x\| \leq 1 \Longrightarrow \|Ax\| \leq c.$$

c) For a family of continuous linear operator $\{A_\alpha\}_{\alpha\in\mathcal{A}}$, $A_\alpha \in \mathcal{L}(X, Y)$ the following conditions are equivalent:

1. $\{A_\alpha\}$ is equi-continuous,
2. $\{A_\alpha\}$ is equi-bounded,
3. $\{A_\alpha\}$ is *bounded in operator norm*, that is $\exists c > 0 \,:\, \|A_\alpha\| \leq c \ \forall \alpha \in \mathcal{A}$.

Taking into account this characterization we can reformulate the boundedness principle for the case of NS so:

Boundedness principle for NS. Let X be a BS and Y be a NS, and let $\{A_\alpha\}_{\alpha\in\mathcal{A}}$ be a family of continuous linear operator from X into Y. Then the following conditions are equivalent:

a) $\sup\limits_{\alpha\in\mathcal{A}} \|A_\alpha\| < \infty$ ($\Leftrightarrow \exists c > 0 \,:\, \|A_\alpha\| \leq c \ \forall \alpha \in \mathcal{A}$);

b) $\sup\limits_{\alpha\in\mathcal{A}} \|A_\alpha x\| < \infty \ \forall x \in X$ ($\Leftrightarrow \forall x \in X \, \exists c > 0 \,:\, \|A_\alpha x\| \leq c \ \forall \alpha \in \mathcal{A}$).

The nontrivial part of this equivalence is the implication b) $\Rightarrow$ a). The latter may be rewritten in the equivalent "negative form":

$$\sup_\alpha \|A_\alpha\| = \infty \Longrightarrow \exists x \in X \,:\, \sup_\alpha \|A_\alpha x\| = \infty.$$

This implication is called the *principle of fixation of singularities*. It means that if a family of operators is "bad" in the sense that the values of the operators on the unit ball are unbounded (although the values of each individual operator are bounded), then we can find a "fixed" point (say, on the unit sphere), where things are "bad".

3.5.2 Criterium of the weak convergence of sequences

Here we give a criterium of convergence of *sequences* in a NS with respect to the *weakened* topology and in a dual NS with respect to the *weak* topology. In order to formulate the result, we need a notion of fundamental set:

Definition. A set F in a TVS X is called *fundamental* if its linear hull is dense in X, that is

$$\overline{\mathrm{lin}\, F} = X.$$

Theorem (criterium of weak convergence of sequences). *Let X be a NB.*

a) A sequence $\{x_n\}$ in X converges to an element $\hat{x} \in X$ in the topology $\sigma(X, X^*)$ iff

1) $\{x_n\}$ is bounded in the norm, that is $\sup_n \|x_n\| < \infty$;

2) there exists a fundamental set G in X^* (equipped with the norm topology) such that for every $x^* \in G$

$$\langle x^*, x_n \rangle \longrightarrow \langle x^*, \hat{x} \rangle. \tag{1}$$

b) If X is a Banach space, then a sequence $\{x_n^*\}$ in X^* converges to an element $\hat{x}^* \in X^*$ in the topology $\sigma(X^*, X)$ iff

1) $\{x_n^*\}$ is bounded in the norm, that is $\sup_n \|x_n^*\| < \infty$;

2) there exists a fundamental set F in X such that for every $x \in F$

$$\langle x_n^*, x \rangle \longrightarrow \langle \hat{x}^*, x \rangle. \tag{2}$$

If (2) is true we say that the sequence $\hat{x}_n^*$ *converges to* $\hat{x}^*$ *at the point (element)* x.
Thus we have the following scheme:

$$\begin{array}{rcl} x_n \overset{\sigma(X,X^*)}{\longrightarrow} \hat{x} & \overset{\text{def.}}{\Longleftrightarrow} & \forall x^* \in X^* \,:\, \langle x^*, x_n \rangle \longrightarrow \langle x^*, \hat{x} \rangle \\ & \Longleftrightarrow & \left\{ \begin{array}{l} 1) \; \exists c > 0 \, \forall n \in \mathbb{N} \,:\, \|x_n\| \le c \\ 2) \; \exists G \in \, \mathrm{Fund}(X^*) \,:\, \forall x^* \in G \,:\, \langle x^*, x_n \rangle \longrightarrow \langle x^*, \hat{x} \rangle \end{array} \right. \end{array} \tag{3}$$

$$\begin{array}{rcl} x_n^* \overset{\sigma(X,X^*)}{\longrightarrow} \hat{x}^* & \overset{\text{def.}}{\Longleftrightarrow} & \forall x \in X \,:\, \langle x_n^*, x \rangle \longrightarrow \langle \hat{x}^*, x \rangle \\ & \overset{X \in \text{BS}}{\Longleftrightarrow} & \left\{ \begin{array}{l} 1) \; \exists c > 0 \, \forall n \in \mathbb{N} \,:\, \|x_n^*\| \le c \\ 2) \; \exists F \in \, \mathrm{Fund}(X) \,:\, \forall x \in F \,:\, \langle \hat{x}^*, x \rangle \longrightarrow \langle \hat{x}^*, x \rangle \end{array} \right. \end{array} \tag{4}$$

$\lhd$ 0°
1) Theorem on completeness of the dual NS;
2) theorem on the canonical imbedding X into X^{**};
3) boundedness principle.

1° The equivalences marked by "def." are immediate corollaries of the definition of $\sigma(X, Y)$ (and were the content of Exercise 4 on p. 46).

2° The assertion a) follows from b), since by 0°1) the space X^* is Banach space and since the canonical imbedding X into X^{**} is an isometry by 0°2). So it remains to prove the right equivalence in (4). Let us prove it.

3° "$\Longrightarrow$" Let for every $x \in X$ it holds $\langle x_n^*, x \rangle \to \langle \hat{x}^*, x \rangle$. Then 2) in (4) is fulfilled with $F = X$. Further for each $x \in X$ the sequence $\langle \hat{x}^*, x \rangle$, being convergent, is bounded (in $\mathbb{R}$). Thus, the sequence $\hat{x}^*$ is pointwise bounded, and by 0°3) we conclude that the norms $\|\hat{x}^*\|$ are bounded, so that 1) in (4) is also fulfilled.

4° "$\Longleftarrow$" Let 1) and 2) in (4) be fulfilled, and let $\hat{x} \in X$. We have to show that $\langle x_n^*, \hat{x} \rangle \to \langle \hat{x}^*, \hat{x} \rangle$. First of all it follows from 2) that the sequence $\langle x_n^*, x \rangle$ converges to $\langle \hat{x}^*, x \rangle$ also for every element x of the linear hull of F. Indeed, if $x = t_1 x_1 + \ldots + t_k x_k$, $t_i \in \mathbb{R}$, $x_i \in F$, then

$$\begin{aligned} \langle x_n^*, x \rangle &= \langle x_n^*, t_1 x_1 + \ldots + t_n x_n \rangle \\ &= t_1 \underbrace{\langle x_n^*, x_1 \rangle}_{\to \langle \hat{x}^*, x_1 \rangle} + \ldots + t_n \underbrace{\langle x_n^*, x_k \rangle}_{\to \langle \hat{x}^*, x_k \rangle} \longrightarrow t_1 \langle \hat{x}^*, x_1 \rangle + \ldots + t_n \langle \hat{x}^*, x_k \rangle = \langle \hat{x}^*, x \rangle. \end{aligned}$$

5° Further, it follows from 1) that it holds also $\|\hat{x}^*\| \le c$. Indeed for all $x \in \operatorname{lin} F$ we have by 4°

$$\langle \hat{x}^*, x \rangle = \lim_{n\to\infty} \underbrace{\langle x_n^*, x \rangle}_{\le \underbrace{\|x_n^*\|}_{\le c} \|x\|} \le c\|x\|. \tag{5}$$

Since $\operatorname{lin} F$ is dense in X and since $\hat{x}^*$ and $\| \, \|$ are continuous functions we conclude that

$$\langle \hat{x}^*, x \rangle \le c\|x\| \quad \forall x \in X,$$

so that

$$\|\hat{x}^*\| = \sup_{\|x\|=1} \langle \hat{x}^*, x \rangle \le c.$$

6° At last let us show that $\langle x_n^*, \hat{x} \rangle \to \langle \hat{x}^*, \hat{x} \rangle$. Let be given $\varepsilon > 0$. Choose $\tilde{x} \in \operatorname{lin} F$ so that

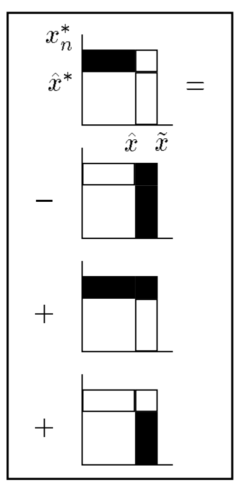
$$\|\hat{x} - \tilde{x}\| \le \frac{\varepsilon}{2c+1} \tag{6}$$

(it is possible since $\operatorname{lin} F$ is dense in X). Now by 4° it holds $\langle x_n^*, \tilde{x} \rangle \to \langle \hat{x}^*, \tilde{x} \rangle$, so for all sufficiently great n we have

$$|\langle x_n^*, \tilde{x} \rangle - \langle \hat{x}^*, \tilde{x} \rangle| \le \frac{\varepsilon}{2c+1} \tag{7}$$

and hence

$$\begin{aligned}
|\langle x_n^*, \hat{x} \rangle - \langle \hat{x}^*, \hat{x} \rangle| &= |\langle x_n^*, \hat{x} \rangle - \langle x_n^*, \tilde{x} \rangle + \langle x_n^*, \tilde{x} \rangle - \langle \hat{x}^*, \tilde{x} \rangle + \langle \hat{x}^*, \tilde{x} \rangle - \langle \hat{x}^*, \hat{x} \rangle| \\
&\le |\langle x_n^*, \hat{x} \rangle - \langle x_n^*, \tilde{x} \rangle| + |\langle x_n^*, \tilde{x} \rangle - \langle \hat{x}^*, \tilde{x} \rangle| + |\langle \hat{x}^*, \tilde{x} \rangle - \langle \hat{x}^*, \hat{x} \rangle| \\
&\overset{(7)}{\le} |\langle x_n^*, \tilde{x} - \hat{x} \rangle| + \frac{\varepsilon}{2c+1} + |\langle \hat{x}^*, \tilde{x} - \hat{x} \rangle| \\
&\le \underbrace{\|x_n^*\|}_{\overset{1)}{\le} c} \underbrace{\|\tilde{x} - \hat{x}\|}_{\overset{(6)}{\le} \frac{1}{2c+1}} + \frac{\varepsilon}{2c+1} + \underbrace{\|\hat{x}^*\|}_{\overset{\circ}{\le} c} \underbrace{\|\tilde{x} - \hat{x}\|}_{\overset{(6)}{\le} \frac{1}{2c+1}} \\
&\le \frac{\varepsilon}{2c+1}(c+1+c) = \varepsilon.
\end{aligned}$$



▷

Remarks.

1. The condition 1) is essential (both in a) and in b)), see Example 1 below.

2. The condition of completeness of X in) is essential, see Example 2 below.

Examples.

1. Consider in the space l_2 the sequence

$$\begin{aligned} x_1 &= (1,0,0,0,\ldots), \\ x_2 &= (0,2,0,0,\ldots), \\ x_3 &= (0,0,3,0,\ldots), \\ &\vdots \end{aligned}$$

and in the dual space l_2^* $(= l_2)$ the set G of all elements of the form

$$(0,\ldots,0,1,0,0,\ldots).$$

The set G is fundamental in l_2^*, and $\langle x^*, x_n \rangle \to 0$ for every $x^* \in G$, but x_n does not converge to 0 in $\sigma(l_2, l_2^*)$. Indeed, for the elements, say,

$$a^* = \left(1, \frac{1}{2}, \frac{1}{3}, \ldots\right), \quad \left(\text{or } b^* = \left(1,0,0, \underbrace{\frac{1}{2}}_{4^{th}\text{ place}}, 0,\ldots,0, \underbrace{\frac{1}{3}}_{9^{th}\text{ place}}, 0,0,\ldots\right)\right)$$

it holds

$$\langle a^*, x_n \rangle \equiv 1 \not\to 0 \; (\text{or even } \langle b^*, x_n \rangle = n \longrightarrow \infty).$$

The point is that $\|x_n\| \to \infty$.

2. Consider the space k of *finite* sequences with the max-norm, defined on p. 41, and the same sequence of continuous linear functionals on k as there, only now we will denote them by x_n^*:

$$\langle x_n^*, x \rangle := n x_n \quad (x = (x_1, x_2, \ldots)).$$

This sequence x_n^* converges to 0 at every element $x \in k$, that is it converges to 0 in $\sigma(k^*, k)$, but the norms $\|x_n^*\|$ $(= n)$ are not bounded. The point is that k is not complete.

3.5.3 Equivalence of boundedness, weak boundedness and relative weak compactness

Here we study interrelations between the properties of boundedness and compactness in norm and weakened topologies.

It is well known that in finite-dimensional spaces every closed bounded set is compact. In infinite-dimensional case this is not so:

Theorem on noncompactness. *In no infinite-dimensional NS the closed unit ball is compact.*

$\lhd$ Let X be an infinite-dimensional NS. Assume that the closed unit ball B in X is compact. All the balls of the radius $\frac{1}{2}$ form a covering of B. By compactness of B there exists a finite subcovering, say $B_1, \ldots, B_N$. For every $n \geq N$ there exists an n-dimensional linear subspace $\tilde{X}$ in X, containing the centers of these N balls. Denote by $\tilde{B}, \tilde{B}_1, \ldots, \tilde{B}_N$ the intersections of $B, B_1, \ldots, B_n$ with $\tilde{X}$. It is clear that $\tilde{B}$ is a ball in $\tilde{X}$ of the radius 1 and $\tilde{B}_1, \ldots, \tilde{B}_N$ are the balls in $\tilde{X}$ of the radius $\frac{1}{2}$.

If we denote the n-dimensional volume of $\tilde{B}$ by V, then it is clear that the n-dimensional volume of each ball $\tilde{B}_i$ is equal to $\left(\frac{1}{2}\right)^n V$. Since $\tilde{B} \subset \tilde{B}_1 \cup \ldots \cup \tilde{B}_N$ we must have $V \leq N\left(\frac{1}{2}\right)^n V$. But for sufficiently great n this is impossible. $\triangleright$

The weakened topology behaves "better" (is more like finite-dimensional ones): weakly closed weakly bounded sets in NS are weakly compact. This is a half of the next theorem. The second half is the equivalence of weak boundedness and boundedness (in norm).

Recall that a set A in a topological space X is called *relatively compact*, if its closure is compact:

$$A \in \mathrm{RelComp}(X) :\Longleftrightarrow \bar{A} \in \mathrm{Comp}(X).$$

Below we write for short $\sigma(X, X^*)$ instead of $(X, \sigma(X, X^*))$. It is supposed as ever that X^* is equipped with its natural norm.

Theorem on equivalence of three properties. *Let X be a NS and $A \subset X$. Then the following conditions are equivalent:*

a) A is dounded, that is bounded in norm;

b) A is weakly bounded, that is bounded in $\sigma(X, X^*)$;

c) A is relatively weakly compact, that is relatively compact in $\sigma(X, X^*)$.

Thus for every NS X

$$\mathrm{Bd}(X) = \mathrm{Bd}(\sigma(X, X^*)) = \mathrm{RelComp}(\sigma(X, X^*)).$$

$\triangleleft$ 0°

1) Theorem on the canonical imbedding;
2) characterization of $\sigma(X, X^*)$-bounded sets (p. 46, Exercise 5);
3) theorem on completeness of the dual NS;
4) boundedness principle for NS;
5) theorem on boundedness of compact sets in TVS;
6) elementary properties of bounded sets in TVS;
7) theorem on compactness (of bipolars and subdifferentials).

1° a) $\Rightarrow$ b). It is evident, since the weakened topology is weaker than the original one.

2° b) $\Rightarrow$ a). Let $A \subset X$ is bounded in $\sigma(X, X^*)$. By 0°1) we can consider A as a family of continuous linear functionals on X^*. By 0°2) the set $x^*(A) = \{\langle x^*, x\rangle \,|\, x \in A\}$ is bounded (in $\mathbb{R}$) $\forall x^* \in X^*$. This means that A when considered as a subset of X^{**} is pointwise-bounded. Now X^* is a Banach space by 0°3), so we can apply 0°4) and conclude that A is bounded in the norm of X^{**}. But the imbedding of X into X^{**} is isometric by 0°1), so A is bounded in the norm of X.

3° c) $\Rightarrow$ b). This follows at once from 0°5).

4° b) $\Rightarrow$ c). We shall prove this implication *only for reflexive* X. Let A be bounded in $\sigma(X, X^*)$. Then its $\sigma(X, X^*)$-closure $\overline{A}^{\sigma(X,X^*)}$ will be also $\sigma(X, X^*)$-bounded by 0°6) and hence will be bounded in the norm by 0°2). So

$$\overline{A}^{\sigma(X,X^*)} \subset B_R^X$$

for some $R > 0$, where B_R^X denotes the closed ball in X of the radius R with the center at 0. The polar $(B_R^X)^\circ (= (B_R^X)^\bullet)$ with respect to the canonical duality $X \leftrightarrow X^*$ is the ball $B_{1/R}^{X^*}$ and hence is a neighbourhood of 0 in X^*. So by 0°7) the polar of $B_{1/R}^{X^*}$ with respect to the duality $X^* \leftrightarrow X^{**}$ is $\sigma(X^{**}, X^*)$-compact. The latter polar is the ball $B_R^{X^{**}}$. But $X^{**} = X$ by our supposition on reflexivity of X, and we may conclude that the ball B_R^X is $\sigma(X, X^*)$-compact. It follows that $\overline{A}^{\sigma(X,X^*)}$ is also $\sigma(X, X^*)$-compact as a $\sigma(X, X^*)$-closed subset of B_R^X. Thus A is relatively $\sigma(X, X^*)$-compact. $\triangleright$

Test question. Why the implication b) $\Rightarrow$ a) does not contradict Example 2 on p. 87 (where the sequence x_n^* is $\sigma(k^*, k)$-bounded, but the norms $\|x_n^*\|$ are not bounded!)?

Corollary on noncoincidence. In no infinite-dimensional NS the weakened topology coincides with the original norm topology.

(In finite-dimensional spaces these topologies coincide.)
$\lhd$ If they coincided, then the closed unit ball would be compact, at conflict with the above theorem on noncompactness. $\rhd$

Remark. This corollary shows that for every infinite-dimensional NS its norm topology and its weakened topology may serve an example of two different locally convex topologies (on one and the same vector space) with the same bounded sets.

3.6 Some additional information

Here we discuss some topologies and convergencies in the space $\mathcal{L}(X, Y)$ and give one result concerning an important notion of the same spectrum of an operator.

3.6.1 Topologies in $\mathcal{L}(X, Y)$

There are three basic topologies in the space $\mathcal{L}(X, Y)$ of all continuous linear mappings from a NS X to a NS Y. We define them for the general case of LCS X, Y by prescribing a base of neighbourhoods of 0 according to the following table (where $A, A_\alpha \in \mathcal{L}(X, Y)$):

Name	Symbol	A base of neighbourhoods of 0	A generating system of semi-norms	Convergence
weak operator topology	$\mathrm{w}(X, Y)$	$U_{x,y^*,\varepsilon} := \{A \,\vert\, \vert\langle y^*, Ax\rangle\vert \le \varepsilon\}$ $x \in X,\ y^* \in Y^*,\ \varepsilon > 0$	$P_{x,y^*}(A) := \vert\langle y^*, Ax\rangle\vert$ $x \in X,\ y^* \in Y^*$	$A_\alpha \overset{w(X,Y)}{\longrightarrow} A \iff$ $\forall x \in X : A_\alpha x \overset{\sigma(Y,Y^*)}{\longrightarrow} Ax$ $\iff \forall x \in X\ \forall y^* \in Y^* :$ $\langle y^*, A_\alpha x\rangle \to \langle y^*, Ax\rangle$
strong operator topology	$\mathrm{s}(X, Y)$	$U_{x,V} := \{A \,\vert\, Ax \in V\}$ $x \in X,\ V \in \mathrm{Nb}_0(Y)$	$P_{x,\pi}(A) := \pi(Ax)$ $x \in X,\ \pi \in \Pi$, where Π is a generating system of semi-norms in Y	$A_\alpha \overset{s(X,Y)}{\longrightarrow} A \iff$ $\forall x \in X : A_\alpha x \to Ax$
topology of uniform convergence on the bounded set	$\mathrm{b}(X, Y)$	$U_{B,V} := \{A \,\vert\, A(B) \subset V\}$ $B \in \mathrm{Bd}(X),\ V \in \mathrm{Nb}_0(Y)$	$P_{B,\pi}(A) := \sup_{x \in B} \pi(Ax)$ $B \in \mathrm{Bd}(X),\ \pi \in \Pi$, where Π is a generating system of semi-norms in Y	$A_\alpha \overset{b(X,Y)}{\longrightarrow} A \iff$ $\forall B \in \mathrm{Bd}(X)$: $A_\alpha x \to Ax$ uniformly in $x \in B$

It is clear that

$$w \le s \le b$$

("$\le$" means "is weaker").

Remarks.

1. If $X = \mathbb{R}^n$, $Y = \mathbb{R}^m$, then $w = s = b$ is the usual topology in the $n \times m$-dimensional space $\mathcal{L}(\mathbb{R}^n, \mathbb{R}^m)$ (of all $n \times m$-matrices).

2. In the case $Y = \mathbb{R}$ we have

$$w(X, \mathbb{R}) = s(X, \mathbb{R}) = \sigma(X^*, X).$$

3. In the case of NS X, Y the topology $b(X, Y)$ coincides with the norm topology in $\mathcal{L}(X, Y)$ and is called also the *uniform operator topology*; in the special case $Y = \mathbb{R}$ this topology wears the name "*strong topology* in X^*".

4. Thus there is some confusion in the terminology: for a NS X

$$\text{"uniform topology in } \mathcal{L}(X,\mathbb{R})\text{"} = \text{"strong topology in } X^*\text{"},$$
$$\text{"strong topology in } \mathcal{L}(X,\mathbb{R})\text{"} = \text{"weak topology in } x^*\text{"}.$$

The "operator terminology" does not agree with the "duality one", but both are traditional!

3.6.2 The spectrum of an operator

Here we consider an important notion of the spectrum of an operator. By study of questions concerning spectra, it is essential to deal with vector space *over* $\mathbb{C}$.

Recall that a complex number λ is called an *eigenvalue* of a linear operator $A : X \to X$, where X is a vector space over $\mathbb{C}$, if there exists a *nonzero* vector $x \in X$ (called the *eigenvector* of A associated with λ) such that

$$Ax = \lambda x.$$

(Emphasize that A "acts" in one and the same space.) In this case the operator $A - \lambda 1$, where 1 denotes the identity operator:

$$1 \equiv 1_X := \mathrm{id}_X,$$

has no inverse operator, since $(A - \lambda 1)x = 0$ for $x \neq 0$.

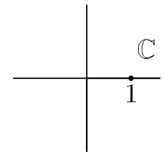
For finite-dimensional X the non-existence of $(A - \lambda 1)^{-1}$ is equivalent to the fact that λ is an eigenvalue of A. But in infinite-dimensional case this is not already so, and things are some more complex (and interesting).

Definition. Let X be a vector space over $\mathbb{C}$, and let $A \in L(X, X)$. The set of all $\lambda \in \mathbb{C}$ such that the operator $A - \lambda 1$ *is not bijective*, is called the *spactrum* of A and denoted by $\sigma(A)$.

The eigenvalues of A form a part of $\sigma(A)$, characterized by the fact that $A - \lambda 1$ *is not injective*. This part of the spectrum is called the *point spectrum*. For orther $\lambda \in \sigma(A)$ the operator $A - \lambda 1$ *is injective*, but *is not surjective* (in the finite-dimensional case such a thing is impossible).

Examples.

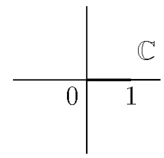
1° For the identity operator 1 the spectrum consists from the single point 1, which is its eigenvalue.



2° For the operator of "multiplication by t" in the space $C([0,1])$, defined by the formula

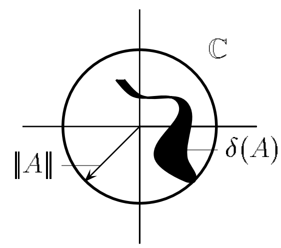
$$(Ax)(t) := tx(t),$$

the spectrum is the segment $[0,1]$ of the real axis, and the point spectrum is *empty*.



Theorem on the spectrum. *Let X be a Banach space over $\mathbb{C}$. Then the spectrum of every operator $A \in \mathcal{L}(X,X)$ is nonempty and closed and iscontained in the disc of radius $\|A\|$ with the center at 0.*

0°
1) The inequality for norms;
2) theorem on convergence of operator series (see below).
1° We shall prove only the last assertion:



◁
$$\lambda \in \sigma(A) \Longrightarrow |\lambda| \le \|A\|.$$

Let $|\lambda| > \|A\|$. Let us show that in this case the operator $(A-\lambda 1)^{-1}$ exists and is given by the following (convergent in norm) series:

$$(A-\lambda 1)^{-1} = -\frac{1}{\lambda}\Big(1 + \frac{1}{\lambda}A + \frac{1}{\lambda^2}A^2 + \ldots\Big). \tag{1}$$

(Notice that *for numbers* the latter equality is just the formula for the sum of a geometrical progression.)
2° Since $\left\|\frac{1}{\lambda}A\right\| = \frac{1}{|\lambda|}\|A\| < 1$, the number series $\sum\limits_0^\infty \left\|\frac{1}{\lambda}A\right\|^k$ converges. Hence by 0°2) the operator series $\sum\limits_0^\infty \left(\frac{1}{\lambda}A\right)^k$ converges and its sum is a continuous linear operator in X. Thus the right-hand side of (1) is a continuous linear operator in X. Let us verify that this operator is inverse to $A-\lambda 1$.
3° For every $n \in \mathbb{N}$ we have

$$\begin{aligned}(A-\lambda 1)\Big(-\frac{1}{\lambda}\Big(1+\frac{1}{\lambda}A+\ldots+\frac{1}{\lambda^n}A^n\Big)\Big) &= \Big(-\frac{1}{\lambda}\Big(1+\frac{1}{\lambda}A+\ldots+\frac{1}{\lambda^n}A^n\Big)\Big)(A-\lambda 1)\\ &= -\frac{1}{\lambda}\Big(\cancel{A}+\cancel{\frac{1}{\lambda}A^2}+\ldots+\frac{1}{\lambda^n}A^{n+1}-\lambda 1-\cancel{A}-\cancel{\frac{1}{\lambda}A^2}-\ldots-\cancel{\frac{1}{\lambda^{n-1}}A^n}\Big)\\ &= 1-\frac{1}{\lambda^{n+1}}A^{n+1}.\end{aligned} \tag{2}$$

But

$$\left\|\frac{1}{\lambda^{n+1}}A^{n+1}\right\| \overset{0°1)}{\le} \underbrace{\left\|\frac{1}{\lambda}A\right\|}_{<1}{}^{n+1} \longrightarrow 0 \text{ as } n \longrightarrow \infty,$$

so in the limit we obtain from (2)

$$(A-\lambda 1)\Big(-\frac{1}{\lambda}\Big(1+\frac{1}{\lambda}A+\ldots\Big)\Big) = \Big(-\frac{1}{\lambda}\Big(1+\frac{1}{\lambda}A+\ldots\Big)\Big)(A-\lambda 1) = 1,$$

just what we had to verify. $\triangleright$

Theorem on convergence of operator series on NS's. *Let X be a NS, and let Y be a Banacg space. Let $\{A_k\} \subset \mathcal{L}(X, Y)$ be a sequence of operators such that*

$$\sum_{k=1} \infty \|A_k\| < \infty.$$

Then the series $\sum A_k$ converges in $b(X, Y)$ (that is, in $\mathcal{L}(X, Y)$ considered as a NS) to some operator $A \in \mathcal{L}(X, Y)$, and

$$\|A\| \leq \sum \|A_k\|.$$

$\triangleleft$ The proof is quite analogous to the corresponding proof in the "usual analysis ($X = Y = \mathbb{R}$), only one writes $\| \quad \|$ instead of $| \quad |$. $\triangleright$

4 Hilbert spaces

Now we turn to the most special spaces, the theory of which is in some aspects the most rich. Notice that this theory is a working tool of quantum mechanics. (We consider as ever only the spaces over $\mathbb{R}$, although the complex spaces are also of great importance.)

4.1 Basic definitions and examples

We introduce here basic definitions, give some example and discuss the place of Hilbert spaces among other normed spaces.

4.1.1 Scalar product

The notion of a Hilbert space (HS) is a generalization of the notion of an Euclidean (finite-dimensional) space:

Definition. A vector space X is called *pre-Hilbert space* (pre-HS) if it is given a *scalar product* (one says also *inner product*) on X, i. e. a bilinear function

$$(\ |\ ) : X \times X \longrightarrow \mathbb{R},$$

which is

1) *symmetric*: $(x|y) = (y|x)\ \forall x, y \in X$;

2) *nonnegatively definite*: $(x|x) \geq 0\ \forall x \in X$.
A pre-HS is called a *Hilbert space* if we have additionaly

3) $(x|x) = 0 \Rightarrow x = 0$ ("Hausdorff");

4) X is complete in the norm, defined by the formula

$$\|x\|^2 := (x|x).$$

That the last formula defines really a norm, will be proved later (see Corollary on the next page).

Examples.

1. The finite-dimensional space $\mathbb{R}^n$ with the usual scalar product

$$(x|y) := \sum_{i=1}^{n} x_i y_i \quad (x = (x_1, \ldots, x_n), y = (y_1, \ldots, y_n))$$

is a HS.

2. The space l_2 of all real sequences $x = (x_1, x_2, \ldots)$ such that $\sum x_i^2 < \infty$ with the scalar product

$$(x|y) := \sum_{i=1}^{\infty} x_i y_i$$

is a HS.

3. For any set $\mathcal{A}$ the space $l_2(\mathcal{A})$ of all families $\{x_\alpha\}_{\alpha\in\mathcal{A}}$ of real number x_α such that

$$\sum_{\alpha\in\mathcal{A}} x_\alpha^2 < \infty \tag{1}$$

with the scalar product

$$(x|y) := \sum_{\alpha\in\mathcal{A}} x_\alpha y_\alpha \tag{2}$$

is a HS. (The space l_2 is a special case of $l_2(\mathcal{A})$ with $\mathcal{A} = \mathbb{N}$.) The sum in (1) is meant as $\sup_{\mathcal{A}_0} \sum_{\alpha\in\mathcal{A}_0} x_\alpha^2$, where $\mathcal{A}_0$ runs over all *finite* subsets of $\mathcal{A}$. It is easy to prove that (1) may be fulfilled, only if *at most countable* number of x_α are nonzero; so the sum in (2) may be meant as a usual series. The space $l_2(\mathcal{A})$ is *separable* (that is possesses a countable dense subset) if $\mathcal{A}$ is countable, and nonseparable othervise.

4. The space $C([0,1])$ with the scalar product

$$(x|y) := \int_0^1 x(t)y(t)dt$$

is a Hausdorff separable pre-HS, but is not a HS. We shall denote it by $C_2([0,1])$.

5. The space $C_2(\mathbb{R})$ of all continuous real-valued functions on $\mathbb{R}$ such that

$$\int_{-\infty}^{\infty} x(t)^2 dt < \infty$$

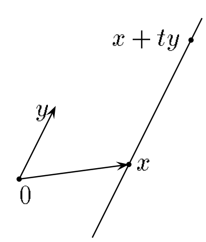
with the scalar product

$$(x|y) := \int_{-\infty}^{\infty} x(t)y(t)dt$$

is a nonseparable and noncomplete pre-HS.

One of the key properties of a scalar product is
Schwartz (-Cauchy-Bunjakowski) inequality Let $(\ |\ )$ be a scalar product and $\|\ \|$ be the associated norm. Then

$$\big|(x|y)\big| \le \|x\|\,\|y\|. \tag{3}$$



$\lhd$ For all $t \in \mathbb{R}$ we have by the property) of a scalar product

$$0 \le (x+ty|x+ty) = \underbrace{(x|x)}_{\|x\|^2} + 2t(x|y) + t^2\underbrace{(y|y)}_{\|y\|^2}.$$

Hence the quadratic trinomial in the right-hand side must have a nonpositive discriminant:

$$(x|y)^2 - \|x\|^2\|y\|^2 \le 0. \rhd$$

Remarks.

1. The Schwartz inequality is very like the basic inequality for norms in normed spaces, and as we shall see it is not by accident.

2. The inequality (3) means that $(\ |\ )$ is continuous with respect to the associated semi-norm.

Corollary. The formula $\|x\|^2 := (x|x)$ defines a finite semi-norm, which is a norm iff $(\ |\ )$ satisfies condition 3) of the definition above.

$\lhd$ All the properties of a semi-norm are evident to be fulfilled, besides the triangle inequality. The latter follows from the Schwartz inequality:

$$\|x+y\|^2 = (x+y|x+y) = \underbrace{(x|x)}_{\|x\|^2} + 2\underbrace{(x|y)}_{\leq \|x\|\,\|y\|} + \underbrace{(y|y)}_{\|y\|^2} \leq (\|x\| + \|y\|)^2. \ \rhd$$

Thus every pre-HS is a semi-normed space, every Hausdorff pre-Hilbert space is a normed space, and every HS is a Banach space.

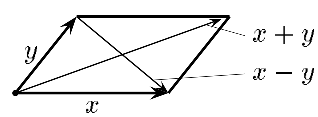
4.1.2 Parallelogram identity

What normed spaces are "Hilbertizable"? The answer is bound up with a well-familiar property of parallelograms: the sum of the squares of the lengths of the diagonals is equal to the sum of the squares of the lengths of the sides.

Parallelogram identity. For any scalar product in a vector space X

$$\|x+y\|^2 + \|x-y\|^2 = 2\|x\|^2 + 2\|y\|^2 \quad \forall x, y \in X.$$

$$\begin{aligned} \lhd (x+y|x+y) + (x-y|x-y) &= (x|x) + 2(x|y) + (y|y) \\ &+ (x|x) - 2(x|y) + (y|y) \\ &= 2(x|x) + 2(y|y) \ \rhd \end{aligned}$$



Theorem on Hilbertizability. *Let X be a normed space. Then a scalar product $(\ |\ )$ on X such that $\|x\|^2 = (x|x)$ exists iff $(\ |\ )$ satisfies the parallelogram identity.*

$\lhd$ The part "only if" was just proved, and the proof of the part "if" may be found in [KF, pp. 160–162] or [KG, nv. 547]. $\rhd$

Examples. 1. The space $\mathbb{R}^n$ with the norm

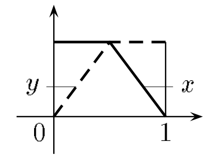
$$\|x\|_p := \left(\sum_{i=1}^{n} |x_i|^p \right)^{\frac{1}{p}} \quad (x = (x_1, \ldots, x_n)),$$

$p \geq 1$, is "Hilbertian" only for $p = 2$. Indeed the vectors

$$x = (1, 1, 0, \ldots, 0) \quad y = (1, -1, 0, \ldots, 0)$$

satisfy the parallelogram identity only for $p = 2$.

2. The space $C([0,1])$ (with the max-norm) is not Hilbertizable, since the pair of functions on the picture does not satisfy the parallelogram identity.



4.1.3 Orthogonality

Since a scalar product satisfies the basic inequality for norms, we can define the *angle* φ between any two vectors x, y in a pre-HS by the formula

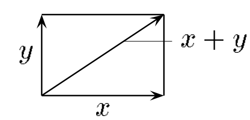
$$\left(x \middle| y\right) = \|x\| \, \|y\| \cos\varphi.$$

In particular two vectors x and y are *orthogonal*, if this angle is equal to $90°$:

$$x \perp y :\Longleftrightarrow \left(x \middle| y\right) = 0.$$

Pythagorean theorem. *In a pre-HS if $x \perp y$ then $\|x+y\|^2 = \|x\|^2 + \|y\|^2$.*

$$\lhd \ \left(x+y \middle| x+y\right) = \underbrace{\left(x \middle| x\right)}_{\|x\|^2} + 2\underbrace{\left(x \middle| y\right)}_{0} + \underbrace{\left(y \middle| y\right)}_{\|y\|^2}. \ \rhd$$



For any set A in a pre-HS the set of all vectors, orthogonal to each vector of A, is called the *orthogonal complement* of A and denoted by $A^\perp$:

$$A^\perp := \left\{ y \in X \,\middle|\, \left(x \middle| y\right) = 0 \ \forall x \in A \right\}.$$

Remark. It is easy to see that $A^\perp$ is a closed subspace in X:

$$A^\perp \in \mathrm{ClSub}(X).$$

4.2 Geometry of a Hilbert space

Here we prove the main fact of the geometry of a Hilbert space, the existence of the orthogonal projection, and derive some important corollaries.

4.2.1 Orthogonal projection

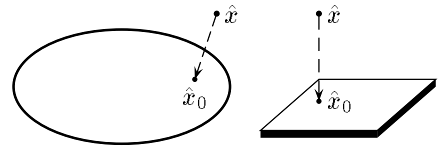
We shall need later only the projections on vector subspaces, but will prove here a more general result.

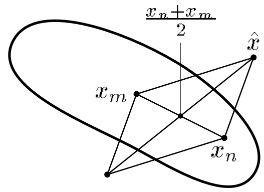
Theorem on orthogonal projection. *Let X be a HS, A be a nonempty closed convex set in X, and $\hat{x}$ be an arbitrary point of X. Then there exists a point $\hat{x}_0 \in A$, which is the closest point to $\hat{x}$ among the points of A, that is which minimizes the distance from $\hat{x}$ to the points of A:*

$$\|\hat{x} - \hat{x}_0\| = \min_{x \in A} \|\hat{x} - x\|.$$

This point is unique and is called the orthogonal projection $pr_A \hat{x}$ of $\hat{x}$ into A.

Remark. In the case of *affine subspace* A the projecting vector is orthogonal to A (see the next subsection), which justifies the name "orthogonal projection".





◁ 0° Parallelogram identity.
1° Put

$$d := \inf_{x\in A}\|\hat{x}-x\|,$$

and let x_n be a sequence in A, such that

$$d = \lim_{n\to\infty}\|\hat{x}-x_n\|.$$

Let us show that $\{x_n\}$ is a Cauchy sequence. By 0° applied to the parallelogram on the picture we have

$$\underbrace{\|(\hat{x}-x_n)+(\hat{x}-\hat{x_m})\|^2}_{2\underbrace{\left\|\hat{x}-\frac{x_n+x_m}{2}\right\|}_{\leq d,\text{since } \frac{x_n+n_m}{2}\in A}} + \underbrace{\|(\hat{x}-x_n)-(\hat{x}-x_m)\|^2}_{\|x_n-x_m\|} = 2\underbrace{\|\hat{x}-x_n\|}_{\to d}^2 + 2\underbrace{\|\hat{x}-x_m\|}_{\to d}^2.$$

If we choose n_0 so that $\|\hat{x}-x_n\|^2 \leq d^2+\varepsilon$ for $n \geq n_0$ then it holds for $n, m \geq n_0$

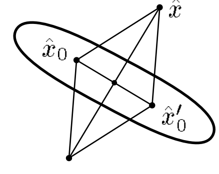
$$\|x_n-x_m\|^2 \leq 2(d^2+\varepsilon)+2(d^2+\varepsilon)-(2d)^2 = 4\varepsilon.$$

Thus $\{x_n\}$ is a Cauchy sequence.
2° Put $\hat{x}_0 := \lim x_n$. Then by continuity of the norm we have

$$\|\hat{x}-\hat{x}_0\| = \lim\|\hat{x}-x_n\| = d.$$

The existence of the closest point is proved.
3° The uniqueness follows again from 0°: if $\|\hat{x}-\hat{x}_0\| = \|\hat{x}-\hat{x}_0'\| = d$ for $\hat{x}_0, \hat{x}_0' \in A$, then it holds analogously

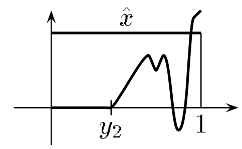
$$\Bigg(2\underbrace{\left\|\hat{x}\frac{\hat{x}_0+\hat{x}_0'}{2}\right\|}_{\geq d}\Bigg)^2 + \|\hat{x}_0-\hat{x}_0'\|^2 = 2\underbrace{\|\hat{x}-\hat{x}_0\|}_{d}^2 + 2\underbrace{\|\hat{x}-\hat{x}_0'\|}_{d}^2,$$



whence it follows that $\|\hat{x}_0-\hat{x}_0'\| \leq 0$, that is $\|\hat{x}_0-\hat{x}_0'\| = 0$ ▷.

Remarks.

1. The condition of completeness of X is essential. For example in the space $C([0,1])$ with the scalar product $\int xy dt$ there exists in the closed subspace

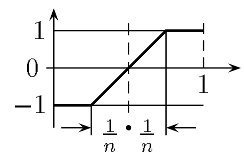


$$A := \left\{ x \,\middle|\, x(t) = 0 \text{ if } t \in \left[0, \frac{1}{2}\right] \right\}$$

no point closest to the point $\hat{x} \equiv 1$ (verify!).

2. Using the above theorem one may establish the following fact: in a HS any sequence

$$A_1 \supset A_2 \supset A_3 \supset \dots$$



of nonempty *closed convex bounded* sets has a nonempty intersection. Notice that in arbitrary normed spaces this is not so (in $C([0,1])$) with the usual max-norm it may be given such a counter-example, based on the sequence $\{x_n\}$ of the functions x_n, represented on the picture.

3. The theorem remains true if we omit the condition of completeness of X (that is for Hausdorff pre-HSs), but add the condition that A lies in a *finite-dimensional* subspace of X. (The point is that every finite-dimensional normed space is complete.)

4.2.2 Canonical decomposition into a direct sum

For any closed subspace (by subspaces we mean, as ever, *linear* subspaces) of a HS we can decompose the HS into the direct sum of this subspace and its orthogonal complement:

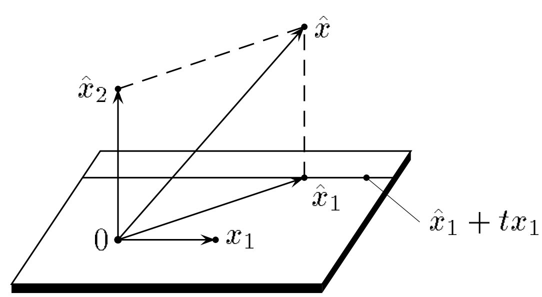
Theorem on the canonical decomposition. *Let X be a HS and X_1 be its closed subspace. Then X is the direct sum of X_1 and $X_2 := X_1^{\perp}$:*

$$X = X_1 \oplus X_2.$$

The assertion remains true if we omit the condition of completeness of X (that is for Hausdorff pre-Hilbert spaces) but add the condition that X_1 is finite-dimensional.

$\lhd$ 0° Theorem on orthogonal projection with Remark to this theorem.

1° Let $\hat{x}$ be an arbitrary point in X. By 0° there exist the orthogonal projection of $\hat{x}$ into X_1. But



$$\hat{x}_1 := \mathrm{pr}_{X_1} \hat{x}, \quad \hat{x}_2 := \hat{x} - \hat{x}_1.$$

We claim that $\hat{x}_2 \in X_1^{\perp}$. Indeed, let $x_1 \in X_1$. Consider the straight line in X_1, passing through the point $\hat{x}_1$ in the direction of the vector x_1. Since $\hat{x}_1$ is the closest to $\hat{x}$ point of X_1, the function

$$\begin{aligned} t \longmapsto \|\hat{x} - (\hat{x}_1 + tx_1)\|^2 &= (\hat{x} - \hat{x}_1 - tx_1 | \hat{x} - \hat{x}_1 - tx_1) \\ &= (\hat{x} - \hat{x}_1 | \hat{x} - \hat{x}_1) - 2t(\hat{x} - \hat{x}_1 | x_1) + t^2(x_1 | x_1) \end{aligned}$$

takes its minimal value at $t = 0$. Hence the derivative in t must be equal to 0 at $t = 0$. But this derivative is equal to $-2(\hat{x} - \hat{x}_1 | x_1)$. It follows that $(\hat{x} - \hat{x}_1 | x_1) = 0$. Thus, $\hat{x}_2 \in X_1^{\perp}$.

2° We have proved that X is the sum of X_1 and X_2. That this sum is direct, follows from the fact that $X_2 = X_1^{\perp}$. Indeed, if $x \in X_1 \cap X_2$, then $x \perp x$, that is $(x | x) = 0$ and hence $x = 0$. $\rhd$

Corollary on the orthogonal complement. Let X be a HS and Y be a closed subspace in X. Then
$$Y^{\perp\perp} := \left(Y^{\perp}\right)^{\perp} = Y.$$

Thus, every closed subspace in a HS is the orthogonal complement of its orthogonal complement.
$\lhd$ It is evident that $Y \subset \left(Y^{\perp}\right)^{\perp}$. Let us prove that $Y \supset Y^{\perp\perp}$. Let $x \in Y^{\perp\perp}$. Write its decomposition
$$x = x_1 + x_2, \quad x_1 \in Y,\ x_2 \in Y^{\perp}.$$
We claim that $x_2 = 0$. Indeed, $x \perp x_2$, hence
$$0 = \left(x \middle| x_2\right) = \underbrace{\left(x_1 \middle| x_2\right)}_{0} + \underbrace{\left(x_2 \middle| x_2\right)}_{\|x_2\|^2},$$
whence it follows $\|x\|_2 = 0$. Thus, $x = x_1 \in Y$. $\rhd$

Remark. The conditon of completeness of X in the theorem on the canonical decomposition is essential. Indeed, consider again the pre-Hilbert space $C([0,1])$ with the scalar product $\int_0^1 xydt$ and the closed subspace $Y = \{x \mid x(t) = 0 \text{ if } t \in [0, \frac{1}{2}]\}$. We have $Y^{\perp} = \{x \mid x(t) = 0 \text{ if } t \in [\frac{1}{2}, 1]\}$. But
$$Y + Y^{\perp} = \left\{x \middle| x\left(\frac{1}{2}\right) = 0\right\} \neq X.$$

4.2.3 Self-duality

A key property of HSs is that we can identify a HS with its dual space.
First of all the condition $\left(x \middle| x\right) = 0 \Rightarrow x = 0$ means that for any HS we have
$$X \overset{(\,|\,)}{\longleftrightarrow} X,$$
that is X forms a dual pair with itself with respect to the own scalar product as the pairing. But we have more than that:

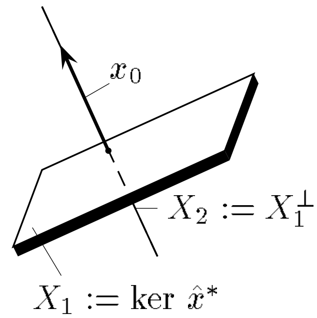
Theorem on the dual space to a HS. *Let X be a HS.*

a) Every element $\hat{x}^* \in X^*$ may be represented by the unique way in the form
$$\langle \hat{x}^*, x\rangle = \left(\hat{x} \middle| x\right) \quad (x \in X),$$
where $\hat{x} \in X$; moreover it holds
$$\|\hat{x}\|_X = \|\hat{x}^*\|_{X^*}.$$
b) V. v. for every $\hat{x} \in X$ the functional $\hat{x}^* := \left(\hat{x} \middle| \cdot\right)$ belongs to X^*, and
$$\|\hat{x}^*\|_{X^*} = \|\hat{x}\|_X.$$

This means just that 1) the norm topology in a HS is *compatible* with the duality $X \overset{(\,|\,)}{\longleftrightarrow} X$, and 2) the (bijective) correspondence $\hat{x}^* \leftrightarrow \hat{x}$ is an *isometry*.
$\lhd$ 1° Assertion b) follows easily from the Schwartz inequality and the fact that $\left(x \middle| x\right) = \|x\|^2$ (give details!).

2° Let us prove a). The uniqueness of desired $\hat{x}$ follows at once from the mentioned property $\left(x \middle| x \right) = 0 \Rightarrow x = 0$, and the equality of the norms follows from the equality in b). It remains to find the desired $\hat{x}$.

Put
$$X_1 := \ker \hat{x}^* \stackrel{\text{def.}}{=} \left\{ x \middle| \langle \hat{x}^*, x \rangle = 0 \right\}.$$



If $X_1 = X$, then $\hat{x}^* = 0$, and we can take $\hat{x} = 0$. Let $X_1 \neq X$. Put $X_2 := X_1^{\perp}$. We claim that X_2 is one-dimensional. Indeed choose any nonzero $x_0 \in X_2$ (it is possible since $X_1 \oplus X_2 = X$) and show that $\{x_0\}$ is a basis in X_2. We can assume that $\|x_0\| = 1$.

At first we must have $\langle x^*, x_0 \rangle \neq 0$, since otherwise we would have $x_0 \in \ker \hat{x}^* = X_1$ and hence $x_0 = 0$. Then each vector $x \in X_2$ may be written as
$$x = \frac{\langle x^*, x \rangle}{\langle x^*, x_0 \rangle} x_0.$$
Indeed the difference
$$x - \frac{\langle x^*, x \rangle}{\langle x^*, x_0 \rangle} x_0$$
belongs both X_2 (evidently) and X_1 (since $\hat{x}^*$ vanishes at this difference), hence is equal to 0. Thus $\{x_0\}$ is a basis in X_2.

3° Now let us show that we can take
$$\hat{x} := \langle \hat{x}^*, x_0 \rangle x_0.$$
Indeed, both our original functional $\hat{x}^*$ and the functional $\left(\hat{x} \middle| \cdot \right)$ are equal to 0 on X_1 and have one and the same value $\langle \hat{x}^*, x_0 \rangle$ at x_0, that is are equal on X_2. Hence they coincide (since $X = X_1 \oplus X_2$). $\triangleright$

Thus we can identify X^* with X (as normed spaces) by means of he rule
$$\langle x, y \rangle = \left(x \middle| y \right).$$

Rorollary on reflexivity. Every HS is reflexive.

$\triangleleft$ $X^{**} = \big(\underbrace{X^*}_{=X}\big)^* = X^* = X.$ $\triangleright$

Remark. A HS is more than reflexive; it "coincides" not merely with its *second* "reflection", but already with the *first* one:
$$X^* = X.$$

4.3 Hilbert basic

Here we show, how the notion of a basic for finite-dimensional Euclidean spaces may be generalized to HSs.

4.3.1 Orthogonal systems

A family $\{e_\alpha\}_{\alpha \in \mathcal{A}}$ of vectors in a pre-HS X is called an *orthogonal system* (ONS) if
$$\left(e_\alpha \middle| e_\beta \right) = \begin{cases} 1 & \text{for } \alpha = \beta, \\ 0 & \text{for } \alpha \neq \beta. \end{cases}$$

Remark. Every ONS is a linearly independent family. (Verify!)

Lemma on finite ONS *Let X be a Hausdorff pre-HS, and $\{e_1, \ldots, e_n\}$ be an ONS in X. Then the subspace $E := lin\{e_1, \ldots, e_n\}$, generated by $e_1, \ldots, e_n$, is closed, and for every point $x \in X$ its orthogonal projection into E is given by the formula*

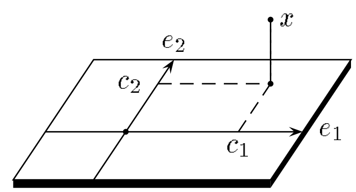
$$pr_E x = \sum_{k=1}^{n} c_k e_k,$$

where

$$c_k := \left(x \middle| e_k \right).$$

◁ 0° Theorem on the canonical decomposition.

1° Let a sequence $x_i = c_{1i}e_1 + \ldots + c_{ni}e_n$ in E converges in X to some element $x \in X$. Then by continuity of scalar product we have for each $k = 1, \ldots, n$



$$\left(x_i \middle| e_k \right) = c_{ki} \longrightarrow \left(x \middle| e_k \right) = c_k.$$

On the base of continuity of arithmetic operations in a TVS we conclude that

$$c_{1i}e_1 + \ldots + c_{ni}e_n \longrightarrow c_1 e_1 + \ldots c_n e_n.$$

Hence by uniqueness of the limit in a Hausdorff TVS it holds

$$x = c_1 e_1 + \ldots + c_n e_n \in E.$$

Thus, E is closed.

2° By 0° we have for every $x \in X$

$$x = x' + x'', \text{ where } x' = \mathrm{pr}_E x \in E,\ x'' \in E^{\perp}.$$

Hence

$$x = \overbrace{\alpha_1 e_1 + \ldots + \alpha_n e_n}^{x'} + x''$$

for some real α_i. Taking the scalar product with e_k gives

$$c_k = \left(x \middle| e_k \right) = \alpha_k,$$

whence it follows that

$$x' = c_1 e_1 + \ldots + c_k e_k. \ \triangleright$$

It turns out that for every point x in a pre-HS its *coordinates* $c_\alpha := \left(x \middle| e_\alpha \right)$ with respect to any ONS $\{e_\alpha\}_{\alpha \in \mathcal{A}}$ from an element of the HS $l_2(\mathcal{A})$:

Bessel inequality. Let X be a Hausdorff pre-HS, and $\{e_\alpha\}_{\alpha \in \mathcal{A}}$ be an ONS in X. Then

$$\sum_{\alpha \in \mathcal{A}} \left(x \middle| e_\alpha \right)^2 \leq \|x\|^2 \quad \forall x \in X. \tag{1}$$

(The sum is meant in the sence explained on p. 93 just after eq. (2).)

◁ 0°

1) Lemma on finite ONS;

2) Pythagorean theorem.

1° Let $x \in X$. By the definition of the sum in (1) it is sufficient to verify (1) for a *finite* $\mathcal{A}' \subset \mathcal{A}$. In this case by 0°1) we have

$$x = \sum_{\alpha \in \mathcal{A}'} c_\alpha e_\alpha + y, \text{ where } c_\alpha := (\, x \,|\, e_\alpha \,) \text{ and } y \perp e_\alpha \ \forall \alpha \in \mathcal{A}'.$$

Hence by 0°2)

$$\|x\|^2 = \sum_{\alpha \in \mathcal{A}'} \|c_\alpha e_\alpha\|^2 + \|y\|^2 = \sum_{\alpha \in \mathcal{A}'} c_\alpha^2 + \underbrace{\|y\|^2}_{\geq 0},$$

whence it follows (1). ▷

4.3.2 Hilbert basises and the theorem on isomorphism

An ONS in a pre-HS is called *maximal* (or *total*, *complete*), if there is no strictly greater ONS in X, that is, if its orthogonal complement consists only from 0.
Remark. In every pre-HS there exist a maximal ONS. It is quite easy to verify (!), using the Zorn lemma.

In a finite-dimensional space any maximal ONS will be of course a *basis* in algebraical sense. It turns out that in *Hilbert* spaces every maximal ONS is a *Hilbert* basis in the sense of the following
Definition. An ONS $\{e_\alpha\}_{\alpha \in \mathcal{A}}$ in a pre-HS X is called a *Hilbert basis* of X, if for every vector $x \in X$ it holds

$$x = \sum_{\alpha \in \mathcal{A}} c_\alpha e_\alpha, \tag{1}$$

where

$$c_\alpha := (\, x \,|\, e_\alpha \,). \tag{2}$$

The inequality (1) means that 1) only *countable* number of the coefficients c_α are nonzero, that is there exist a countable subset $\mathcal{A}' \subset \mathcal{A}$ (depending on x) such that $c_\alpha = 0 \ \forall \alpha \in \mathcal{A} \setminus \mathcal{A}'$; 2) if we number *anyhow* $\alpha \in \mathcal{A}'$ into a sequence $\alpha_1, \alpha_2, \ldots$, then the partial sums

$$S_n := \sum_{k=1}^{n} c_{\alpha_k} e_{\alpha_k} \tag{3}$$

converge (in norm) to x:

$$\|S_n - x\| \longrightarrow 0 \quad \text{as } n \longrightarrow \infty. \tag{4}$$

Remark. The representation (1) is *unique*, that is if we have a representation (1) with *some* c_α, thene these c_α are given by (2). Indeed, by continuity of scalar product we have for each $\alpha \in \mathcal{A}$

$$(\, x \,|\, e_\alpha \,) = \lim_{n \to \infty} (\, S_n \,|\, e_\alpha \,),$$

but

$$(\, S_n \,|\, e_\alpha \,) = \left(\sum_{l=1}^{n} c_{\alpha_l} e_{\alpha_l} \,\Big|\, e_\alpha \right) = \begin{cases} c_{\alpha_k} = c_\alpha & \text{if } \alpha = \alpha_k \text{ and } n \geq k, \\ 0 = c_\alpha & \text{if } \alpha \notin \mathcal{A}' \end{cases}$$

Example. In the HS $l_2(\mathcal{A})$ the family $\{\delta_\alpha\}_{\alpha \in \mathcal{A}}$, where

$$\delta_\alpha := \{\delta_{\alpha\beta}\}_{\beta \in \mathcal{A}}, \quad \delta_{\alpha\beta} := \begin{cases} 1 & \text{if } \alpha = \beta, \\ 0 & \text{if } \alpha \neq \beta, \end{cases}$$

is a Hilbert basis. In particular, in l_2 the vectors

$$(1,0,0,0,\dots),(0,1,0,0,\dots),\dots$$

form a Hilbert basis.

Characterizations of Hilbert basises. Let X be a HS, and $\{e_\alpha\}_{\alpha\in\mathcal{A}}$ be an ONS in X. Then the following conditions are equivalent:

a) $\{e_\alpha\}$ is a Hilbert basis;

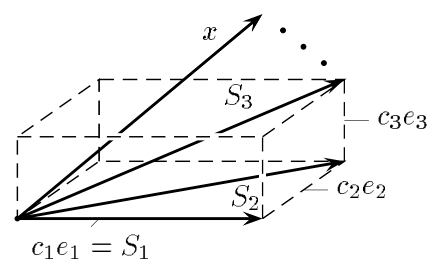
b) for every $x \in X$ it holds the *Parseval equality*

$$\|x\|^2 = \sum_{\alpha\in\mathcal{A}} \big(x \big| e_\alpha\big)^2;$$

c) $\{e_\alpha\}$ is maximal (that is $\{e_\alpha\}^\perp = \{0\}$);

d) $\{e_\alpha\}$ is fundamental (that is $\overline{lin\{e_\alpha\}} = X$).

The parseval equality is just the infinite-dimensional formula for the square of the length of the diagonal in a rectangular parallelepiped.



◁ 0°
1) Definition of a Hilbert basis;
2) Bessel inequality;
3) Pythagorian theorem;
4) theorem on the canonical decomposition;
5) lemma on finite ONS;
6) theorem on the orthogonal projection.

1° a) $\Rightarrow$ b). Let $\{e_\alpha\}$ be a Hilbert basis, and let x be an arbitrary element of X. By 0°1) we have

$$x = \lim_{n\to\infty} S_n, \quad S_n := \sum_{k=1}^{n} c_k e_k, \quad c_k := \big(x \big| e_k\big)$$

(for short we write e_k instead of e_{α_k}). By continuity of norm we conclude that

$$\|x\|^2 = \lim \|S_n\|^2 \overset{0°3)}{=} \lim \sum_{k=1}^{n} c_k^2 = \sum_{k=1}^{\infty} c_k^2.$$

2° b) $\Rightarrow$ c). Let $\{e_\alpha\}$ be such that for each $x \in X$ the Parseval equality is fulfilled, and let $x \in \{e_\alpha\}^\perp$. Then $\big(x \big| e_\alpha\big) = 0\ \forall\alpha$, and hence by the Parseval equality for x we have $\|x\|^2 = 0$, that is $x = 0$.

3° c) $\Rightarrow$ d). Let $\{e_\alpha\}$ be maximal, that is $\{e_\alpha\}^\perp = \{0\}$. Put $X_1 := \overline{\mathrm{lin}\{e_\alpha\}}$, $X_2 := X_1^\perp$. It is clear that $X_2 = \{0\}$. By 0°4) it holds $X = X_1 \oplus X_2 = X_1$, that is $\{e_\alpha\}$ is fundamental.

4° d) $\Rightarrow$ a). Let $\{e_\alpha\}$ be fundamental, that is $X = \overline{\text{lin}\{e_\alpha\}}$, and let x be an arbitrary element of X. Put $c_\alpha := (x|e_\alpha)$. By 0°2)

$$\sum_{\alpha\in\mathcal{A}} c_\alpha^2 \le \|x\|^2 < \infty,$$

hence, as it was noticed on p. 94, only a countable number of c_α may be nonzero, say, c_α with $\alpha \in \mathcal{A}'$, $\mathcal{A}'$ being countable; so $c_\alpha = 0$ if $\alpha \notin \mathcal{A}'$. Number $\mathcal{A}'$ *anyhow* into a sequence and shall for short use notations $e_1, e_2, \dots$ and $c_1, c_2, \dots$ for corresponding e_α and c_α. We need to show that

$$\|S_n - x\| \longrightarrow 0, \quad s_n := \sum_{k=1}^{n} c_k e_k.$$

By 0°5)

$$S_n = \text{pr}_{E_n} x, \quad E_n := \text{lin}\{e_1, \dots, e_n\}.$$

Since $E_1 \subset E_2 \subset \dots$, we have evidently

$$\left\|x - \text{pr}_{E_1} x\right\| \ge \left\|x - \text{pr}_{E_2} x\right\| \ge \dots,$$

that is

$$\|x - S_1\| \ge \|x - S_2\| \ge \dots.$$

Thus it is sufficient to prove that

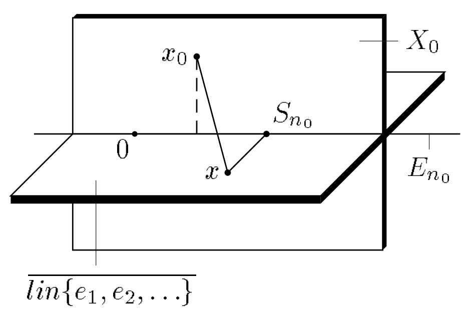
$$\forall \varepsilon > 0\, \exists n_0 \in \mathbb{N} \,:\, \|x - S_{n_0}\| \le \varepsilon.$$

5° Let be given $\varepsilon > 0$. Since $x \in \overline{\text{lin}\{e_\alpha\}}$, we can find a finite number of e_α, say $e'_1, \dots, e'_N$, such that

$$\|x - x_0\| \le \varepsilon \quad \text{for some } x_0 \in \text{lin}\{e'_1, \dots, e'_N\}.$$

Among $e'_1, \dots, e'_N$ there are some "our" elements from the sequence $e_1, e_2, \dots$ and some "foreign" ones. Choose n_0 to be greater than the numbers of all "our" e_k occuring among $e'_1, \dots, e'_N$, and put

$$X_0 := \text{lin}\Big\{e_1, \dots e_{n_0}, \underbrace{f_1, \dots f_{m_0}}_{\text{"foreign" elements among } e'_1,\dots,e'_N}\Big\}.$$



It is clear that $x_0 \in X_0$. We claim that

$$\text{pr}_{X_0} x = S_{n_0} \,\Big(= \text{pr}_{E_{n_0}} x\Big).$$

Indeed by 0°5)

$$\text{pr}_{X_0} x = \underbrace{(x|e_1)e_1 + \dots + (x|e_{n_0})e_{n_0}}_{S_{n_0}} + \underbrace{(x|f_1)}_{0} f_1 + \dots + \underbrace{(x|f_{m_0})}_{0} f_{m_0} = S_{n_0}$$

$\Big((x|f_i) = 0$ since $c_\alpha = 0$ for $\alpha \notin \mathcal{A}'\Big)$. But by 0°6) the orthogonal projection of x into X_0 minimized the distance from x to the points of X_0, so

$$\|x - S_{n_0}\| \le \|x - x_0\| \le \varepsilon. \;\triangleright$$

Remark. The condition of completeness of X is essential for the implication c) $\Rightarrow$ d) to be true. For example, consider in l_2 the subspace

$$\operatorname{lin}\{x_1, e_2, e_3, \ldots\},$$

where

$$x_1 = \left(1, \frac{1}{2}, \frac{1}{3}, \ldots\right),\ e_2 = (0,1,0,0,\ldots),\ e_3 = (0,0,1,0,\ldots), \ldots,$$

and let X be this subspace with the induced from l_2 scalar product. Then the ONS $\{e_2, e_3, \ldots\}$ is maximal (verify!), but evidently is not fundamental.

Theorem on existence and equi-cardinality of Hilbert basises. *Let X be a HS. Then X possesses a Hilbert basis, and all the Hilbert basises in X have one and the same cardinality.*

This cardinality is called the *Hilbert dimension* of X.

$\lhd$ 0°

1) Existence of a maximal ONS in every pre-HS (see the Remark on p. 102);
2) Characterizations of Hilbert basises.

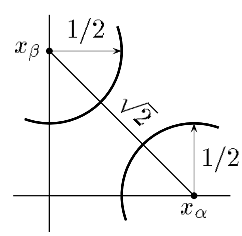
1° By 0°1) in X there exists a maximal ONS. By 0°2) this ONS is a Hilbert basis.

2° In finite-dimensional X equicardinality of all Hilbert basises follows from the analogous algebraical fact, since in this case the notions of a (algebraical) basis and of a Hilbert basis coincide.

3° Now let X possesses a countable Hilbert basis $\{e_n\}_{n\in\mathbb{N}}$. Then X is infinite-dimensional (since $\{e_n\}$ is, as every ONS, a linearly independent family) and is separable (since $\{e_n\}$ is by 0°2)) fundamental, that is line$\{e_n\}$ is dense in X, whence it follows, that the countable set of all linear combination of $\{e_n\}$ with *rational* coefficients is dense in X). It implies that each other Hilbert basis $\{x_\alpha\}_{\alpha\in A}$ contains an infinite number of elements. If $\mathcal{A}$ was uncountable,then there would exist in X an uncountable number of disjoint balls, say of the radius $\frac{1}{2}$ with the centers at x_α, which contradicts the separability of X.

4° The case of uncountable dimensions requires an appealing the set theory, and we shall omit it.

$\rhd$



Example. The Hilbert dimension of $l_2(T)$ is equal to the cardinality of T. In particular l_2 is countable-dimensional in the sense of Hilbert dimension.

Two pre-HS X and Y are called *isomorphic* if there exists a bijective mapping $i : X \to Y$, which preserves both the linear structure (i. e. is linear) and the scalar product, i. e.

$$(ix_1 | ix_2)_Y = (x_1 | x_2)_X \quad \forall x_1, x_2 \in X.$$

Theorem on isomorphism of HSs. *Two HSs are isomorphic iff they have one and the same Hilbert dimension.*

$\lhd$ 0°

1) Example on p. 102;
2) characterizations of Hilbert basises.

1° The necessity is obvious, since by an isomorphism a Hilbert basis is transformed into a Hilbert basises.
2° The sufficiency is an immediate corollary of the following two facts:

1. a HS with a Hilbert basis $\{e_\alpha\}_{\alpha\in\mathcal{A}}$ is isomorphic to $l_2(\mathcal{A})$;
2. if $\mathcal{A}$ and $\mathcal{B}$ have one and the same cardinality then $l_2(\mathcal{A})$ and $l_2(\mathcal{B})$ are isomorphic.

Let us prove the first fact (the second one is obvious):
3° Let HS X has a Hilbert basis $\{e_\alpha\}_{\alpha\in\mathcal{A}}$. By (viz. by the Parseval equality) the mapping

$$i : x \longmapsto \{c_\alpha\}_{\alpha\in\mathcal{A}}, \quad c_\alpha := \left(x \middle| e_\alpha\right),$$

is an *isometry* of X into $l_2(\mathcal{A})$: $\|ix\|^2 = \sum\limits_{\alpha\in\mathcal{A}} c_\alpha^2 = \|x\|^2$. Hence i is injective and its image $i(X)$ is a closed subspace in $l_2(\mathcal{A})$. But $i(X)$ contains all images

$$i(e_\alpha) = \delta_\alpha, \quad \alpha \in \mathcal{A},$$

which form a Hilbert basis of $l_2(\mathcal{A})$ by 0°1), so that by 0°2) (part d)) a closed subspace, containing them, must coincide with the whole $l_2(\mathcal{A})$. Thus i is also surjective.
4° At last the fact that i preserves scalar product, follows from the fact that i is isometry, since a scalar product may be reconstructed on the basis of the associated norm $\left(2\left(x \middle| y\right) = \|x+y\|^2 - \|x\|^2 - \|y\|^2\right)$. ▷

Remark. Thus $l_2(\mathcal{A})$ may be regarded as an "arithmetic realization" of all "abstract" HS which have the Hilbert dimension equal card$\mathcal{A}$.

4.3.3 The separable case; the orthogonalization procedure

Here we consider an important special case of separable (pre-)HS.
Theorem on countable Hilbert dimension. *A HS is separable iff it has the countable Hilbert dimension.*

◁ This was in fact proved in Part 3° of the proof of the theorem of existence and equicardinality of Hilbert basises. ▷

Corollary. All the separable HS are isomorphic (to l_2).

In another words, l_2 is the "arithmetic realization" of all separable HS.
The separable case is distinguished by the fact, that the existence of a Hilbert basis may be proved without the assumption on completeness:
Theorem on orthogonalization. *Let X be a separable Hausdorff pre-HS. Then there exists a countable Hilbert basis in X. And what is more, there exists a countable ONS in X which satisfies all the conditions a)–d) from the theorem on characterizations of Hilbert basises.*
This ONS may be constructed by means of the orthogonalization procedure, *described in the proof below.*

Remark. The condition that a pre-HS is Hausdorff, is equivalent to condition) from the definition of a pre-HS.

◁ 0°
1) The following fact of the theory of metric spaces: If a set is dense in a metric space X, then this set is dense also in the *completion* of X;

2) lemma on finite ONS;
3) characterizations of Hilbert basises in HS.

1° *Orthogonalization procedure.* Let X be a Hausdorff pre-HS and $\{x_n\}$ be any countable linearly independent family in X. Put

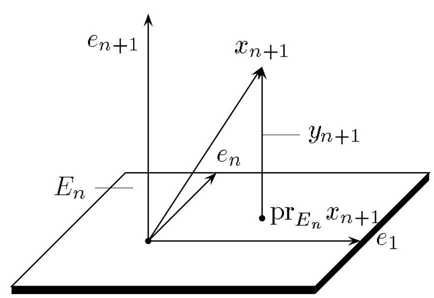
$$\begin{aligned}
&y_1 := x_1, && e_1 := \tfrac{y_1}{\|y_1\|}, \quad \ldots\\
&y_2 := x_2 - \mathrm{pr}_{E_1} x_2 \overset{0°2)}{=} x_2 - \big(x_2 \big| e_1\big) e_1, && e_2 := \tfrac{y_2}{\|y_2\|},\\
&\vdots && \vdots\\
&y_{n+1} := x_{n+1} - \mathrm{pr}_{E_n} x_{n+1} \overset{0°2)}{=} x_{n+1} - \sum_{k=1}^{n} \big(x_{n+1} \big| e_k\big) e_k, && e_{n+1} := \tfrac{y_{n+1}}{\|y_{n+1}\|}, \quad \ldots\\
&\vdots && \vdots
\end{aligned}$$

$$\begin{aligned}
\ldots \quad & E_1 := \mathrm{lin}\{e_1\} = \mathrm{lin}\{x_1\},\\
& E_2 := \mathrm{lin}\{e_1, e_2\} = \mathrm{lin}\{x_1, x_2\},\\
& \vdots\\
\ldots \quad & E_{n+1} := \mathrm{lin}\{e_1, \ldots, e_{n+1}\} = \mathrm{lin}\{x_1, \ldots, x_{n+1}\}\\
& \vdots
\end{aligned}$$

This algorithm is called the *orthogonal procedure.*

By the construction $y_{n+1} \in E_n^{\perp}$ $\forall n$, so $\{e_n\}$ is an ONS in X. It is clear that

$$\mathrm{lin}\{e_1, e_2, \ldots\} = \mathrm{lin}\{x_1, x_2, \ldots\}.$$



2° Let $\{x_n\}$ be a countable dense family in X. Delete successively from $\{x_n\}$ each element which is a linear combination of the preceding ones. It is clear that the obtained family $\{y_n\}$ will be linearly independent and *fundamental*. Apply to $\{y_n\}$ the orthogonalization procedure. We claim that the resulting ONS $\{e_n\}$ is our desired one.

Indeed, since $\mathrm{lin}\{e_1, e_2, \ldots\} = \mathrm{lin}\{y_1, y_2, \ldots\}$, the family $\{e_n\}$ will be also fundamental, so the condition d) is fulfilled.

3° In order to prove that the rest conditions are fulfilled, we consider the *completion* $\bar{X}$ of X in the metric generated by the norm (recall that X supposed to be Hausdorff). It is clear that both linear structure and the scalar product may be extended onto $\bar{X}$ in a natural way ("by continuity") and that the resulting pre-HS is a HS.

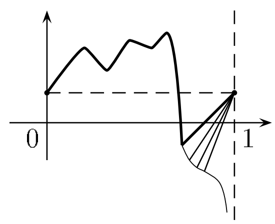
4° By 0°1) the ONS $\{e_n\}$ remains fundamental in $\bar{X}$. Therefore by 0°3) it satisfies the three rest conditions a)–c) in $\bar{X}$ and hence, evidently, in X. ▷

Example. In the pre-HS $C_2([0,1])$ (with the scalar product $\int_0^1 xy dt$) the functions

$$1, \sqrt{2}cos 2\pi t, \sqrt{2}sin 2\pi t, \sqrt{2}cos 4\pi t, \sqrt{2}\sin 4\pi t, \ldots$$

form a Hilbert basis.

This follows (verify!) from 1) the theorem on uniform approximation of continuous functions on $[0, 1]$, satisfying the condition $x(0) = x(1)$, by the trigonometric polynomials, and 2) the fact that any continuous function on $[0, 1]$ may be approximated in $C_2([0, 1])$ by functions satisfying the mentioned condition:



References

[1] A. N. Kolmogorov, S. V. Fomin: *Základy teorie funkcí a funkcionální analýzy,* Praha, SNTL, 1975.

[2] č. . äÚÓÓŤ, É. Ź. ÄÍŤŽÓ: ůÍŤÓŽť ŽŤÁ, ĆÓŮÍ, ćůÍ, 1977.

[3] ć. Óš, ś. í. : čŤŚtŠ ÓÔŚÚÓ. íÓž 1, ĆÓŮÍ, č, 1962.

[4] ź. ĚššŮ: ůÍŤÓŽť ŽŤÁ, ĆÓŮÍ, ĆŤ, 1969.

[5] ä. ÓŮŤš: ůÍŤÓŽť ŽŤÁ, ĆÓŮÍ, ĆŤ, 1967.

[6] Ä. Ä. äŤŤŽŽÓ, Ä. . ÉŤÝŤŤ: íŚÓŚž Ť ÁšŤ ůÍŤÓŽÓÓ ŽŤÁ, ĆÓŮÍ, ćůÍ, 1979.

[7] L. Mišík: *Funkcionální analýza,* Alfa, Bratislava, 1989.

[8] K. Najzar: *Funkcionální analýza,* SPN, Praha, 1988.

[9] . ÚŮÓ, ś. ŹŤž: ŹŤŽÓśŚŤ ůÍŤÓŽÓÓ ŽŤÁ Ť íŚÓŤŤ ÓÔŚÚÓÓ, ĆÓŮÍ, ĆŤ, 1983.

[10] ůÍŤÓŽť ŽŤÁ, ĎÔÓ ĆÚŚžÚŤŚŮÍ ŤŠŽŤÓÚŚÍ, 2-Ś ŤÁš., ĆÓŮÍ, ćůÍ, 1972.