Reading the Follwing Chapter on Norms

Our present strategy in Analysis II is to develop a theory of differentiation *simultaneously* for functions of one real or complex variables and for vector valued functions of several variables. For this purpose we need all the linear algebra you have learned in Linear Algebra I; but we need a perhaps deeper understanding of norms than you have considered so far. The following pages should help you to refresh your memory on some vector space theory and to have before you the relevant facts on norms.

These notes include some references to vector spaces of functions and among these vector spaces of functions of Riemann integrable functions. We shall discuss Riemann integration after our discussion of the theory of differentiability. Accordingly, in your reading of the material now you may defer such references as e.g. in 1.2(iv)–(vi) and 1.6(iii); if you wish, you may also skip for now the subsection entitled "Complete metric spaces and Banach spaces" from 1.8 to 1.11. The subsections on "Hilbert spaces" and "the geometry of real inner product spaces" comprising 1.15–1.23 is a recall of what you saw in Linear Algebra and is perhaps a good application of what you learned in Analysis I about angles in the context of the complex exponential function.

Important reading, however, is the subsection on "Finite dimensional normed vector spaces", on "Linear maps", on "Linear forms", and on the "Continuity of linear maps" from 1.24 to 1.31, and do read 1.12 and 1.13. Also the subsection on "The operator norm" contains material we need if we want to define the exponential function on spaces of matrices (1.32–1.35). You will see that the material on finite dimensional real or complex vector spaces links nicely with what we have said on compact metric spaces and the Bolzano-Weierstraß Theorem.

All section numbers occuring in the text which do not refer to number in this Chapter refer to the orange book (Analysis I).

Chapter 1
Normed Vector Spaces

1. Normed Vector Spaces

We need background information on vector spaces such as it is provided in the courses on linear algebra. Therefore we have to borrow information from that line of mathematical investigation. In particular, we need here the theory of *normed* vector spaces; we will recall the most salient features of norms, because it is the norms that provide us with the metrics we need. In the standard introductory courses on linear algebra, not all the facts on normed vector spaces are introduced ot the extend they are needed in the calculus of several variables. We shall emphasize the connections between linear algebra and analysis by discussion relevant applications right away. A restriction to finite dimensional vector spaces would be counterproductive; however, many applications do pertain to spaces like $\mathbb{R}^n$ and $\mathbb{C}^n$. We rely on the fact that the theory of vector spaces is developed from axioms, with a start out of set theory, in the same spirit as we developed analysis.

Norms on Vector Spaces

The theory of norms is the link between the theories of vector space and the theory of metric spaces. Inevitably, we have to discuss normed vector spaces it at this point.

For the sake of completeness we recall the definition of a vector space over a field $\mathbb{K}$ which in our situation will **always** be either the field $\mathbb{R}$ of real numbers or the field $\mathbb{C}$ of complex numbers.

Definition 1.1. A *vector space* V over a field $\mathbb{K}$ is a set endowed with an *addition* $(x, y) \mapsto x + y: V \times V \to V$ and a *scalar multiplication* $(r, x) \mapsto r{\cdot}x: \mathbb{K} \times V \to V$ such that the following axioms are satisfied:

ADD

$(V, +)$ is a commutative group, that is, the axioms ADD of 1.37 are satisfied:

$(\mathcal{C})$	$(\forall x, y \in K)$	$x + y = y + x$	(Commutativity)
$(\mathcal{A})$	$(\forall x, y, z \in K)$	$(x + y) + z = x + (y + z)$	(Associativity)
$(\mathcal{N})$	$(\exists 0 \in K)(\forall x \in K)$	$0 + x = x + 0 = x$	(Neutral Element)
$(\mathcal{I})$	$(\forall x \in K)(\exists y \in K)$	$x + y = y + x = 0$	(Inverse elements)

SCAL

Addition and scalar multiplication are linked through the following axioms:

($\mathcal{S}$i)	$(\forall x \in V)$	$1{\cdot}x = x$	(Action of identity)
($\mathcal{S}$ii)	$(\forall r, s \in \mathbb{K},\ x \in V)$	$r{\cdot}(s{\cdot}x) = (rs){\cdot}x$	(Associativity)
($\mathcal{S}$iii)	$(\forall r \in \mathbb{K},\ x, y \in V)$	$r{\cdot}(x + y) = r{\cdot}x + s{\cdot}x$	(Distributivity 1)
($\mathcal{S}$iv)	$(\forall r, s \in \mathbb{K}, x \in V)$	$(r + s){\cdot}x = r{\cdot}x + s{\cdot}x$	(Distributivity 2)

The elements of a vector space are called *vectors* , in the context of a vector space, the field elements tend to be called *scalars* . (The German words are „Vektor" and „Skalar", the French expressions are ≪vecteur≫ and ≪scalaire≫.)

A subset $W \subseteq V$ of a vector space is called a *vector subspace* of V if $W + W \subseteq W$ and $\mathbb{K}{\cdot}W \subseteq W$, that is, if W is *closed under addition and scalar multiplication.* □

We note expressly that we never say what a vector is but rather how we manipulate vectors. This is in complete agreement with our axiomatic introduction of real numbers in Chapter 1.

Notice at once that a vector subspace W of a a vector space V is a vector space in its own right with respect to the addition and scalar multiplications that is induced on on W by the addition and scalar multiplication of V.

The German words for vector subspace are „Untervektorraum" or „Teilvektorraum." Therefore German students (of all ages) tend to come up with the "translation" *subvectorspace.*
While this "translation" may be understood, one should accept the fact that the English language does not allow the concatenation of nouns resulting in new ones; as a consequence, the proper function of prefixes emerges in a different fashion. There is no such thing as a "subvector." The nonassociativity of the German conglomeration of nouns is evidenced by this example: „Unter(vektorraum)" is a legitimate handling of the prefix „Unter-", while „(Untervektor)raum" would by an absurd association. [1]

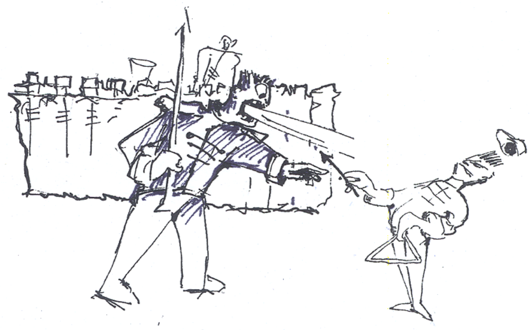
We encountered, in passing, numerous examples of vector spaces:

---

1 One of my teachers used to offer an even better example for the nonassociativity of German conglomeration of nouns; his example was „Mädchenhandelsschule".

Recently, a colleague of mine at the University of Tübingen cites a newspaper report: *Am 7/8. Mai 2000 fand in Tübingen ein „Hallenflohmarkt mit Kindertauschbörse" statt.* See also Mark Twain: The awful German Language, in: A Tramp abroad, Penguin Books 1997, pp 390ff. "Some German words are so long they have a perspective. Observe these examples: *Freundschaftsbezeigungen. Dilettantenaufdringlichkeiten. Stadtverordnetenversammlungen. Untervektorraumkonstruktionen.* These this are not words, they are alphabetical processions. And they are not rare; one can open a German math book any time and see them marching majestically across the page... They impart a martial thrill to the meekest subject." [Exercise. Check Mark Twain and find out where the quotation of the original text has been (slightly) modified.]

vector



subvector

Examples 1.2. **(i)** All n-tuple-spaces $\mathbb{K}^n$, $n = 1, 2, \ldots$, are vector spaces under componentwise addition and scalar multiplication:

$$\begin{aligned}(x_1, \ldots, x_n) + (y_1, \ldots, y_n) =& (x_1 + y_1, \ldots, x_n + y_n), \\ r{\cdot}(x_1, \ldots, x_n) =& (rx_1, \ldots, rx_n).\end{aligned}$$

These examples have central significance for the foundation of the entire theory.

(ii) Let X be any nonempty set and $\mathbb{K}^X$ the set of all functions $f\colon X \to \mathbb{K}$ with pointwise addition and scalar multiplication

$$(f, g) \mapsto f + g, \text{ respectively, } (r, f) \mapsto r{\cdot}f,$$

$$\begin{aligned}(f + g)(x) =& f(x) + g(x) \\ (r{\cdot}f)(x) =& rf(x).\end{aligned}$$

This makes $\mathbb{K}^X$ a vector space. Since analysis is the theory of functions, obviously this is a crucial example because many sets of functions occurcing in analysis emerge as vector subspaces of $\mathbb{K}^X$.

In reality, **(i)** is a special case of **(ii)**, because an n-tuple $(x, \ldots, x_n)$ is none other than a function $j \mapsto x_j\colon \{1, \ldots, n\} \to \mathbb{K}$. (Cf. the topic space of sequences $\mathbb{K}^{\mathbb{N}}$ in Remark 2.39ff., or spaces of functions, implicitly in 3.23, in E5.2).

(iii) Let $B(X) \leq \mathbb{K}^X$ denote the set of all *bounded* function. Then $B(X)$ is a vector subspace (cf. E5.2).

(iv) If X is a real interval, then the set $C^n(X) \subseteq \mathbb{R}^X$ of all n-times continuously differentiable functions and the set $C^\infty(X)$ of all smooth functions on I are vector subspaces of $\mathbb{R}^X$. (For $n = 0$ this statements includes the case of the vector space $C(X)$ of continuous functions.)

(v) The set $I([a, b])$ of all Riemann integrable functions on $[a, b]$ is a vector subspace of $B([a, b])$. The set $S([a, b])$ of all step functions on $[a, b]$ and the set $C([a, b])$ are vector subspaces of $I([a, b])$ (cf. 5.18).

(vi) Given a continuous function $f: I \to \mathbb{R}$, the set of all solutions $u: I \to \mathbb{R}$ of the linear differential equation $u'(t) = f(t)u(t)$ is a vector subspace of $C^1(I)$.

(vii) The set of all increasing real valued functions on an interval I is *not* a vector subspace of $\mathbb{R}^I$, nor is the set of all monotone functions a vector subspace.□

Exercise E1.1. Take stock of all assertions in 1.2. In particular, verify (vi) and (vii).

In dealing with the space of real numbers and the space of complex numbers, we were able to proceed with analysis only after we had introduced metrics. The distance of two numbers x and y was defined to be $d(x, y) = |y-x|$. This made all sets of numbers into a metric space. For this purpose we used the *absolute value* or *norm* $|\cdot|$. The idea of a *norm* can be extended to the vector spaces $\mathbb{K}^n$ and other vector spaces such as we saw in the case of $B(X)$ in 5.3 (in the case of $X = [a, b]$). Let us now systematically deal with this concept.

We now collect the defining properties of a norm $\|\cdot\|$, being guided by the properties of the absolute value (cf. 1.73.).

Definition 1.3. A *norm* on a vector space V over $\mathbb{K} = \mathbb{R}$ or $\mathbb{K} = \mathbb{C}$ is a function $\|\cdot\|: V \to \mathbb{R}$ if the following conditions are satisfied:
(i) $(\forall x \in V)\, \|x\| \geq 0$ and $(\|x\| = 0 \Leftrightarrow x = 0)$,
(ii) $(\forall x, y \in V)\, \|x + y\| \leq \|x\| + \|y\|$, (*triangle inequality*),
(iii) $(\forall r \in \mathbb{K}, x \in V)\, \|r\cdot x\| = |r|\cdot\|x\|$. □

The triangle inequality is equivalent to the following condition:
(iv) $(\forall x, y \in V)\, \big|\|x\| - \|y\|\big| \leq \|x - y\|$. □

Exercise E1.2. Prove the equivalence of (iii) and (iv) in 1.3.

On the n-tuple vector spaces, there are special norms which have particular significance.

Proposition 1.4. *For $x = (x_1, \ldots, x_n)$ in $\mathbb{K}^n$ we define*

(i) $\|x\|_1 = |x_1| + \cdots + |x_n|$.
(ii) $\|x\|_2 = \sqrt{|x_1|^2 + |x_2|^2 + \cdots + |x_n|^2}$.
(iii) $\|x\|_\infty = \max\{|x_1|, \ldots, |x_n|\}$.
Then all of these functions define norms on $\mathbb{K}^n$.

Proof . Exercise. □

Exercise E1.3. Prove 1.4
[Hint. The verification of 1.3(i, iii) is easy. The triangle inequality is easy for $\|\cdot\|_p$ for $p = 1$ and $p = \infty$. In the case $p = 2$, it is a consequence of the Cauchy-Schwarz inequality. Indeed, since the squaring function is strictly antitone, we first notice that the triangle inequality is equivalent to

(ii*) $(\forall x, y \in V)\ \|x + y\|^2 \leq \|x\|^2 + 2\|x\|{\cdot}\|y\| + \|y\|^2$.
If we now set $(x|y) = \sum_{j=1}^{n} x_j\overline{y_j}$, then $\|x\|_2^2 = (x|x)$, and $\|x+y\|_2^2 = (x+y \mid x+y) = (x|x)+(x|y)+(y|x)+(y|y) = \|x\|^2+(x|y)+\overline{(x|y)}+\|y\|^2 = \|x\|^2+2\operatorname{Re}(x|y)+\|y\|^2$. Thus since $\operatorname{Re}(x|y) \leq |(x|y)|$, for the 2-norm, (ii*) is implied by
(ii**) $(\forall x,\ y \in V)\ |(x|y)| \leq \|x\|_2{\cdot}\|y\|_2$,
and this is the Cauchy Schwarz inequality (in the complex version of 5.9 for $\mathbb{K} = \mathbb{C}$).]

Examples i) and iii) are special cases of the more general Definition
(iv) $\|x\|_p = \sqrt[p]{|x_1|^p + \cdots + |x_n|^p}$.

It can be shown that $\|{\cdot}\|_p$ is a norm on $\mathbb{K}^n$ for all $p = 1, 2, \ldots, \infty$, but the proof is harder.

Definition 1.5. A vector space $(V, \|{\cdot}\|)$ with a norm $\|{\cdot}\|: V \to \mathbb{R}$, is called a *normed vector space*. We shall often refer to V itself as normed vector space. The norm $\|{\cdot}\|_2$ on $\mathbb{K}^n$ is called a *euclidian norm*. □

Automatically, every vector subspace of a normed vector space is a normed vector space.

Example 1.6. (i) The vector space $B(X)$ of all bounded functions $X \to \mathbb{K}$ on a set X is a normed vector space with respect to the norm defined by

$$\|f\|_\infty = \sup\{|f(x)| : x \in X\}.$$

This norm is often referred to as the *sup-norm*. All vector subspaces of $B(X)$ are normed spaces with respect to the sup-norm.

(ii) If X is a compact metric space, then the space $C(X)$ of all continuous functions $f: X \to \mathbb{K}$ is a vector subspace of $B(X)$ and is, therefore, a normed space (cf. 3.52).

(iii) The vector space $C([a, b])$ of all continuous and hence integrable function (see 5.18(i)) on $[a, b]$ is a normed space with respect to each of the norms

$$\|f\|_1 = \int |f|, \quad \|f\|_2 = \sqrt{\int |f|^2}, \quad \text{and } \|f\|_\infty = \sup\{|f(x)| : a \leq x \leq b\}.$$

On the vector space $I([a, b])$, the functions $\|{\cdot}\|_1$ and $\|{\cdot}\|_2$ fail to be norms. Indeed, the characteristic function f of the singleton set is nonzero, but $\|f\|_1 = \|f\|_2 = 0$.
The conditions (ii), (iii), and (iv) of 1.4 are satisfied and, in addition, the condition
(i′) $(\forall x \in V)\quad \|x\| \geq 0$. □

A function $\|{\cdot}\|$ which satisfies (i′), (ii), (iii) is called a *seminorm*, and a space with a seminorm is called a *seminormed space*.

Proposition 1.7. *Let $(V, \|{\cdot}\|)$ be a normed vector space and $X \subseteq V$ an arbitrary subset. If we set $d(x, y) = \|x - y\|$, then (X, d) is a metric space.*

Proof . Exercise. □

Exercise E1.4. Prove Proposition 1.7.

As in every metric space we may consider spherical neighborhoods. The closed unit balls $B_1(0) = \{x \in V : \|x\| \le 1\}$ in the normed vector space $V = \mathbb{R}^2$, with respect to the norms $\|\cdot\|_p$, $p = 1, 2, \infty$ look as follows

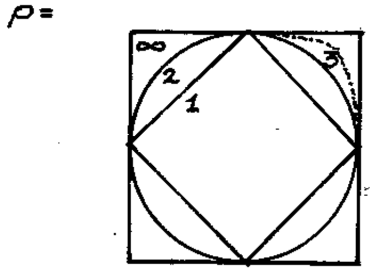


Figure 1.1

Complete metric spaces and Banach spaces

By 1.7 above, the concept of a Cauchy sequence (see 2.42) is meaningful in any normed vector space. In 2.43 we proved that every Cauchy sequence in $\mathbb{R}$ and $\mathbb{C}$ converges. In the paragraph following 2.42 we observe that in $\mathbb{Q}$ with the natural metric given by $d(x, y) = |y - x|$ there are divergent Cauchy sequences. We shall now give a particular name to those metric spaces in which this does not occur.

Definition 1.8. (i) A metric space (X, d) is called *complete*, if every Cauchy sequence converges.

(ii) A complete normed vector space over $\mathbb{R}$ or $\mathbb{C}$ is called a *Banach space.* □

From Theorem 2.43 we know that $\mathbb{R}$ and $\mathbb{C}$ are Banach spaces (of dimension 1) in their own right.

Proposition 1.9. *Assume that X is a subspace of a complete metric space Y. Then the following statements are equivalent.*
(i) *X is closed in Y.*
(ii) *X is complete.*

Proof . (i)⇒(ii): Let $(x_n)_{n\in\mathbb{N}}$ be a Cauchy sequence in X. But then this sequence is a Cauchy sequence in Y. By hypothesis, Y is complete, hence $y = \lim_{n\to\infty} x_n$ in Y. Now X is closed by (i), and thus $y \in X$ (cf. E3.4(iv)). Somit hat die Folge einen Limes in X.

(ii)⇒(i): Let y be an accumulation point of X in Y; we have to show $y \in X$. For every natural number n there is an $x_n \in X \cap U_{1/n}(y)$, that is, $d(y, x_n) < \frac{1}{n}$.

Clearly, $y = \lim x_n$. Claim: $(x_n)_{n\in\mathbb{N}}$ is a Cauchy sequence in X. For a proof let $\varepsilon > 0$. Let us pick N so, that $n > N$ implies $d(y, x_n) < \frac{\varepsilon}{2}$. If now $m, n > N$, then $d(x_m, x_n) \le d(x_m, y) + d(y, x_n) < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon$. Thus the claim is proved. By (ii) the space X is complete. Thus the Cauchy sequence $(x_n)_n$ has a limit x in X. Thus $y = \lim x_n = x \in X$, and this is what we had to show. □

Proposition 1.10. (i) *Let X be an arbitrary nonempty set. Then $B(X)$ is a Banach space.*

(ii) *If X is a metric space, then $C(X) \cap B(X)$ is a Banach space.*

(iii) *If X is a compact metric space then $C(X)$ is a Banach space.*

Proof. (i) Let $(f_k)_{k\in\mathbb{N}}$ be a Cauchy sequence in $B(X)$ and let $x \in X$ be arbitrary. Then $(f_k(x))_{k\in\mathbb{N}}$ is a Cauchy sequence in $\mathbb{K}$ and thus has a limit $f(x) \in \mathbb{K}$ by Theorem 2.43. Now let $\varepsilon > 0$ be given. Since (f_k) is a Cauchy sequence there is a natural number N so that $\|f_k - f_p\| < \varepsilon/2$ holds for all $k, p > N$. Thus $\|f_k(x) - f_p(x)\| < \varepsilon/2$ for all $x \in X$. The function $r \mapsto |r - f_p(x)|: \mathbb{K} \to \mathbb{R}$ is continuous (cf. E3.12(iii)). Therefore $|f(x) - f_p(x)| = \lim_k |f_k(x) - f_p(x)| \le \varepsilon/2$ for all $x \in X$. Consequently, $\|f - f_p\| = \sup\{|f(x) - f_p(x)| : x \in X\} \le \varepsilon/2 < \varepsilon$. This shows $f = \lim f_p$, and because of $|f(x) - f_p(x)| \le \varepsilon/2$ for all x we note $|f(x)| \le |f_p(x)| + \varepsilon \le \|f_p\| + \varepsilon$ for any fixed p. This shows that f is bounded and thus $f \in B(X)$. (We observed, by the way that $\|f\| \le \|f_p\| + \varepsilon$.)

(ii) By (i) and the preceding Proposition 1.9 we have to show that $C(X) \cap B(X)$ is closed in $B(X)$. Thus let $f = \lim f_n$ with $f_n \in C(X)$. We have to show that f is continuous in any $x \in X$. Thus let $\varepsilon > 0$ be given. First we find an N so, that $d(f, f_n) < \varepsilon/3$ for $n > N$. Now let $n > N$. Then $f_n: X \to \mathbb{K}$ is continuous; hence we find a $\delta > 0$ such that $d\big(f_n(u), f_n(x)\big) < \varepsilon/3$ for $d(u, x) < \delta$. Thus for all of these u we have

$$d\big(f(x), f(u)\big) \le d\big(f(x), f_n(x)\big) + d\big(f_n(x), f_n(u)\big) + d\big(f_n(u), f(u)\big) < \frac{\varepsilon}{3} + \frac{\varepsilon}{3} + \frac{\varepsilon}{3} = \varepsilon.$$

This shows the of f in x.

(iii) After Theorem 3.52, every continuous real or complex valued function on a compact space is bounded. Therefore we now have $C(X) \subseteq B(X)$, that is $C(X) \cap B(X) = C(X)$, and the assertion now follows from (ii). □

We noted before that $\mathbb{K}^n$ is a special case of $\mathbb{K}^X$ with $X = \{1, \ldots, n\}$. If X is finite, then $\mathbb{K}^X = B(X)$. Thus we obtain at once from 1.10(i):

Corollary 1.11. *With respect to the norm $\|\cdot\|_\infty$, the vector space $\mathbb{K}^n$ is a Banach space.* □

It is now rather useful that we can *compare* the various norms on $\mathbb{K}^n$.

Proposition 1.12. *For all $x \in \mathbb{K}^n$ we have*

(i) $\|x\|_2 \le \sqrt{n}\|x\|_\infty$, *and*

(ii) $\|x\|_\infty \le \|x\|_2$.

Proof. (i) Let $m = \|x\|_\infty = \max\{\, |x_j| \mid j = 1, \dots, n\,\}$. Then $\|x\|_2^2 = \sum_{j=1}^n x_j\overline{x_j} \le \sum_{j=1}^n m^2 = nm^2$. Extracting square roots on both sides gives (i).

(ii) If $m = |x_k|$, then $m^2 = x_k\overline{x_k} \le \sum_{j=1}^n x_j\overline{x_j} = \|x\|_2^2$. □

Therefore a set $U \in \mathbb{K}^n$ is open with respect to $\|\cdot\|_2$ if an only if it is open with respect to $\|\cdot\|_\infty$; the two norms and their associated metrics define the same topology on $\mathbb{K}^n$. Thus topological concepts like the convergence of sequences in $\mathbb{K}^n$ of the continuity of functions between subsets of $\mathbb{K}^m$ and $\mathbb{K}^n$ does not depend on our choice of one of these two norms. Likewise a sequence is a $\|\cdot\|_2$-Cauchy sequence if and only if it is a $\|\cdot\|_\infty$-Cauchy sequence. Thus we may conclude

Corollary 1.13. *In $\mathbb{R}^n$ and in $\mathbb{C}^n$, any Cauchy sequence with respect to the euclidean norm $\|\cdot\|_2$ converges. In particular, $\mathbb{K}^n$ is a Banach space with respect to the euclidian norm $\|\cdot\|_2$.* □

In 1.27 we shall prove an even better result.

Not every normed vector space is complete as is illustrated by the following example.

We consider the Banach space $B(\mathbb{N})$ of all bounded sequences $(x_1, x_2, \dots)$ in $\mathbb{R}$ the vector subspace of all sequences $(x_1, x_2, \dots)$, which have only finitely many nonzero terms x_k, that is, we consider $V = \{(x_n)_{n\in\mathbb{N}} \colon (\exists m)(\forall n) n \ge m \Rightarrow x_n = 0\}$. By Proposition 1.9, V is complete iff it is closed in the Banach space $B(\mathbb{N})$. We now show that this is not the case: Define $f_1 = (1, 0, 0, 0, \dots)$, $f_2 = (1, 1/2, 0, 0, 0, \dots)$, $\dots$, $f_n = (1, 1/2, \dots, 1/n, 0, 0, 0, \dots)$. Then $(f_k)_{k\in\mathbb{N}}$ is a sequence in V. Its limit in $B(\mathbb{N})$ is the sequence $(1, 1/2, \dots, 1/n, \dots)$, all of whose members are nonzero and which therefore does not belong to V.

Uniform convergence versus pointwise convergence

If a sequence of functions $f_n \colon X \to \mathbb{K}$ converges in Banach space $B(X)$ with respect to $\|\cdot\|_\infty$ towards f, then we say that it *converges uniformly* to f. We have to distinguish this type of convergence meticulously from that by which the sequence of numbers $f_n(x)$ converges to $f(x)$ for each $x \in X$. If this is the case we say that the sequence of the functions f_n *converges pointwise* to f.

Every sequence which converges uniformly to a function f converges pointwise but not vice versa, as the following example shows.

Indeed, consider the sequence $f_n \colon [0, 1] \to \mathbb{R}$ defined as follows:

$$f_n(x) = \begin{cases} 2nx & \text{for } x \in [0, 1/2n[, \\ 2 - 2nx & \text{for } x \in [\frac{1}{2n}, \frac{1}{n}[, \\ 0 & \text{for } x \in [1/n, 1]. \end{cases}$$

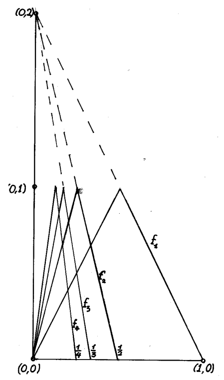


Figure 1.2

This sequence converges pointwise, but not uniformly to the zero function. However, if a sequence of functions converges pointwise to a function f and if, in addition, it converges uniformly on, then it converges uniformly to f.

Exercise E1.5. Let X be an arbitrary set. Prove the following assertions on a sequence $(f_n)_{n\in\mathbb{N}}$ and an element f in $B(X)$.

(i) $(f_n)_n$ converges uniformly to f iff

(U) $$(\forall\varepsilon > 0)(\exists N \in \mathbb{N})(\forall n > N)(\forall x \in X)\, |f(x) - f_n(x)| < \varepsilon.$$

(ii) $(f_n)_n$ converges pointwise to f iff

(P) $$(\forall x \in X)(\forall\varepsilon > 0)(\exists N \in \mathbb{N})(\forall n > N) \qquad |f(x) - f_n(x)| < \varepsilon.$$

[Hint. In case (ii) we reproduce the definition. In case (i), show that (U) and $\lim_n \|f - f_n\| = 0$ are equivalent.]

By Proposition 1.10, any uniform limit of a sequence of continuous functions is continuous. In particular, for any compact metric space X such as for instance $X = [a, b]$, the vector space $C(X)$ of all continuous functions $f\colon X \to \mathbb{K}$ is a Banach space with respect to the sup-norm. In Part (iii) of Proposition 1.10 compactness was used only in order to secure that the limit function is bounded.

We have observed just now that $C([a, b])$ is a Banach space with respect to the sup-norm. A similar theorem also holds for Riemann integrable functions as we shall prove now.

Theorem 1.14. *Assume that a sequence $f_n\colon [a, b] \to \mathbb{R}$ of real valued functions on the compact interval $[a, b]$ converges uniformly to a function $f\colon [a, b] \to \mathbb{R}$, and assume further that all terms f_n of the sequence are integrable, then the limit function f is integrable and $\int f = \lim_n \int f_n$.*

In particular, the space $I([a, b])$ of all Riemann integrable functions on $[a, b]$ is a Banach space with respect to the sup-norm and $\int\colon I([a, b]) \to \mathbb{R}$ is a continuous linear form.

Proof . Let $\varepsilon_0 > 0$. Set $\varepsilon = \varepsilon_0/(b-a+1)$. Then there exists an n such that $\|f - f_n\|_\infty < \varepsilon/3$. That is,

$$f_n - \varepsilon/3 < f < f_n + \varepsilon/3. \tag{$*$}$$

By the Riemann Criterion 5.12 there are step functions $s, t \in S[a,b]$ such that $s < f_n < t$ and $\int(t-s) < \varepsilon/3$. Now $s - \varepsilon/3 < f_n - \varepsilon/3 < f < f_n + \varepsilon/3 < t + \varepsilon/3$ and $\int\big((t+\varepsilon/3) - (s-\varepsilon/3)\big) < \varepsilon/3 + 2(b-a)\varepsilon/3 < \varepsilon_0/3 + 2\varepsilon_0/3 = \varepsilon_0$. The Riemann Criterion now shows that f is integrable. Furthermore we know from the uniformity of the convergence of f_n to f that for all sufficiently large natural numbers m we have $\|f - f_m\| < \frac{\varepsilon}{3}$. For these m we have $|\int f - \int f_m| = |\int(f - f_m)| \le \int |f - f_m| \le \int \varepsilon/3 = (b-a)\varepsilon/3 \le \varepsilon_0/3$. Therefore $\lim_{m\to\infty} \int f_m = \int f$. □

We point out specifically that the statement $\int(\lim f_n) = \lim(\int f_n)$ in 1.14 means, that in the case of uniform convergence one may exchange integration and passage to the limit.

Hilbert spaces

Returning for a moment to the spaces $\mathbb{K}^n$ we recall that the euclidian norm was defined by $\|x\|_2^2 = |x_1|^2 + \ldots + |x_n|^2$.

In proving the triangle inequality for the norm $\|\cdot\|$ we used a two argument function from which this norm arises; let us summarize what we did in that proof. For two vectors $x = (x_1, \ldots, x_n)$ and $y = (y_1, \ldots, y_n)$ in $\mathbb{K}^n$ we define the following number in $\mathbb{K}$:

$$(x \mid y) = x_1\overline{y_1} + \cdots + x_n\overline{y_n} \in \mathbb{K}, \tag{4}$$

also written as $\langle x, y\rangle$ or $x{\cdot}y$. It is called *scalar product* or *inner product* of x and y.

□

In terms of the inner product we have $\|x\|_2^2 = (x \mid x)$. We record the properties of this inner product.

Definition 1.15. Let V be a real or complex vector space. A function $(\cdot|\cdot) : V \times V \to \mathbb{K}$ is called an *inner product* if it satifies the following conditions.

(i) $(x + y \mid z) = (x \mid z) + (y \mid z)$ and $(x \mid y + z) = (x \mid y) + (x \mid z)$ for all $x, y, z \in V$.
(ii) $(rx \mid y) = r(x \mid y)$ and $(x \mid ry) = \overline{r}(x \mid y)$ for all $x, y \in V$ and $r \in \mathbb{K}$.
(iii) $(x \mid y) = \overline{(y \mid x)}$ for all $x, y \in V$.
(iv) $(x \mid x) \ge 0$ for all $x \in V$, and
(v) $(x \mid x) = 0$ implies $x = 0$.

In the case of the inner product on $\mathbb{K} = \mathbb{R}$, one may omit the overlines.

Sometimes an inner product is also called a positive definite *sesquilinear form* in reference to the conjugate linearity in the second argument, as the Latin prefix “sesqui” means “one-and-a-half times.” Another expression that is frequently applied to an inner product is "positive definite *hermitian form.*" If all conditions with the exception of (v) are satisfied, one speaks of a *positive semidefinite hermitean form.* □

Proposition 1.16. *For $K = \mathbb{R}$, $\mathbb{C}$, on the vector space $\mathbb{K}^n$ the function defined by* (4) *above is an inner product.*

Proof . Exercise. □

Exercise E1.6. Prove Proposition 1.16.

Proposition 1.17. (Cauchy-Schwarz Inequality) *An inner product on a complex vector space satisfies*

$$|(f \mid g)|^2 \leq (f \mid f)(g \mid g). \tag{5}$$

Proof . Exercise. □

Exercise E1.7. (i) Prove 1.17(5).

(ii) Prove the following assertion:

For any inner product on a real or complex vector space V, the function $\|\cdot\|: V \to \mathbb{R}$ defined by $\|x\| = \sqrt{(x|x)}$ is a norm.

[Hint. (i) Invoke 5.9.

(ii) Use (i) and the arguments of Exercise E1.2.]

Example 1.18. On the space $I = I([a, b])$ of all real valued functions on $[a, b]$ that are Riemann integrable, the function $(\cdot \mid \cdot): I \times I \to \mathbb{R}$ $(f \mid g) = \int fg$ is a positive semidefinite function which satisfies (1), (2) and (3), but fails to satisfy (4). Such a function is called a positive semidefinite hermitian form.

Observe that $\|f\|_2 = \sqrt{(f \mid f)}$. In Example 1.6(b) we noted that $\|\cdot\|_2$ is not a norm on $I([a, b])$, but only a seminorm.

Let $\{ r_n : n \in \mathbb{N} \}$ be an enumeration of the set of rational numbers in $[0, 1]$ and let $f_n \in I[0, 1]$ denote the characteristic function of $\{r_1, \ldots, r_n\}$. Then $f_n \leq f_{n+1}$, $n = 1, 2, \ldots$, and $\int f_n = 0$ as well as $\int |f_n - f_m|^2 = 0$. Thus $(f_n)_{n\in\mathbb{N}}$ is a Cauchy sequence in $I([0, 1])$ with respect to the seminorm $\|\cdot\|_2$. It converges pointwise (from below) to the characteristic function f of $\mathbb{Q} \cap [0, 1]$, a function which is not Riemann integrable (see discussion following Definition 5.11(ii)). All of the functions f_n have the norm $\|f_n\|_2 = 0$, but neither of them is 0. The convergence of f_n to f is far from uniform since $\|f_n - f_m\|_\infty = 1$ for $m \neq n$ and $\|f - f_n\|_\infty = 1$ for all n. □

Definition 1.19. A real or complex vector space with an inner product $(\cdot|\cdot)$, that is, a positive definite hermitian form, is called an *inner product space.* A Banach space whose norm arises from a positive definite hermitean form is called a *Hilbert space.* □

Sometimes an inner product space is called a *pre-Hilbert-space.*

Examples 1.20. (i) The vector spaces $\mathbb{R}^n$ (and $\mathbb{C}^n$) are Hilbert spaces with respect to the inner product.

(ii) Let V be the set of all sequences $(x_k)_{k\in\mathbb{N}}$, $x_k \in \mathbb{K}$, such that $\sum_{k=1}^{\infty} |x_k|^2$ converges. Thus set is a vector subspace of $\mathbb{K}^{\mathbb{N}}$. For each pair of elelements $x = (x_k)_{k\in\mathbb{N}}$ and $y = (y_k)_{k\in\mathbb{N}}$. Then the infinite series $\sum_{k=1}^{\infty} x_k\overline{y_k}$ converges absolutely, and its sum yields a number $(x \mid y)$ in such a fashion that $(\cdot|\cdot)$ is a positive definite hermitean form on the vector space V making it into a Hilbert space. □

The inner product space V of 1.20(ii) is denoted ℓ^2.

Exercise E1.8. Prove, that ℓ^2 *is a Hilbert space.* □

The geometry of real inner product spaces

We are interested in the geometry of inner product spaces, notably in the special case of $\mathbb{R}^n$. We shall utilize our knowldge of complex numbers and the exponential function. The Cauchy-Schwarz Inequality says $|(x \mid y)| \le \|x\|{\cdot}\|y\|$. If x as well as y are nonzero, we can form the vectors $u = \|x\|^{-1}{\cdot}x$ and $v = \|y\|^{-1}{\cdot}y$, both of which are *unit vectors*, that is vectors of length 1. In that case the inner product $(u \mid v)$ is a number in $[-1, 1]$. What is the significance of this number?

Let us assume momentarily that $u, v \in \mathbb{R}^2$. In that case we identify $\mathbb{R}^2$ with $\mathbb{C}$ under the bijection $(x, y) \mapsto x + iy$ and observe that

$$\begin{aligned}\operatorname{Re}(u\overline{v}) =& \operatorname{Re}\big((u_1 + u_2 i)(v_1 - v_2 i)\big) = \operatorname{Re}\big(u_1v_1 + u_2v_2 + i(-u_1v_2 + u_2v_1)\big)\\ =& u_1v_1 + u_2v_2 = (u|v).\end{aligned}$$

Now let $|u| = |v| = 1$, say, $u = e^{is}$ and $v = e^{it}$ with unique numbers $s, t \in [0, 2\pi[$ by Proposition 3.38(i). Note that for $r \in \mathbb{R}$ we have $\cos r = \operatorname{Re} e^{ir} = \operatorname{Re} e^{-ir} = \operatorname{Re} e^{i(2\pi - r)} = \cos(2\pi - r)$, since exp has period $2\pi i$ (cf. 3.36). Then $(u \mid v) = \operatorname{Re}(u\overline{v}) = \operatorname{Re}(e^{is}e^{-it}) = \operatorname{Re} e^{i(s-t)} = \cos(s - t) = \cos|s - t| = \cos(2\pi - |s - t|)$

Definition 1.21. Let (u_1, u_2), (v_1, v_2) be two nonzero vectors in $\mathbb{R}^2$; set $u = u_1 + u_2 i$ and $v = v_1 + v_2 i$ in $\mathbb{C}$ and write $\frac{u}{|u|} = e^{is}$ and $\frac{v}{|v|} = e^{it}$ with unique real numbers s, $t \in [0, 2\pi[$. Then the number

$$\alpha(u, v) = \min\{|s - t|, 2\pi - |s - t|\} \in [0, \pi]$$

is the called the *nonoriented angle* between u and v. (Cf. Definition 3.39ff.) □

The number $|s-t|$ depends on the set $\{u, v\}$ and not on the ordered pair (u, v), that is, it does not change if we exchange the roles of u and v; this is why we call it the "nonoriented" angle between u and v.

The following figure shows that for equal values of $\cos s$ and $\cos t$ two different angles s and t are conceivable which are still distinguished by the numbers $\sin s$ and $\sin t$. (Cf. Proposition 4.50(∗).)

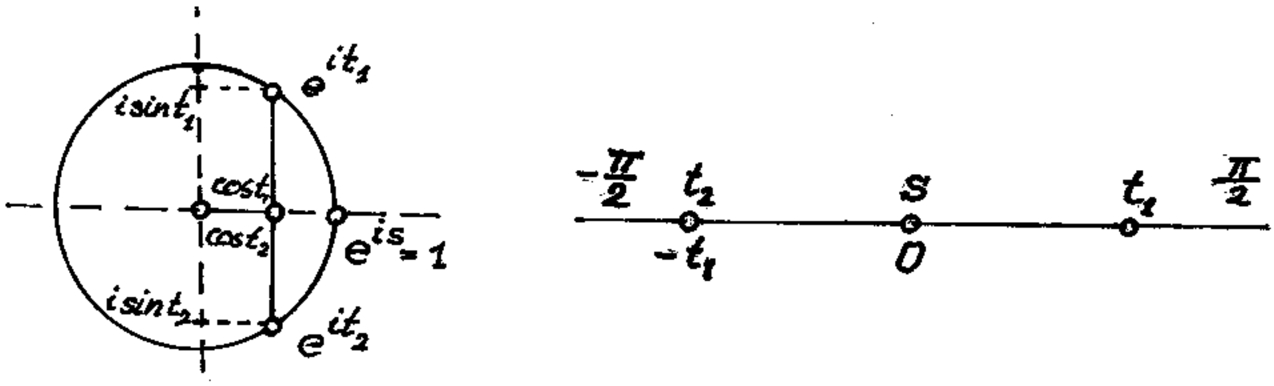


Figure 1.3

Now we have to get away from $\mathbb{R}^2$. We ask the question whether in an arbitrary real vector space V with an inner product $(\cdot|\cdot)$ we can interpret $(f|g)$ for two unit vectors f and g as the cosine of an angle between f and g. For this purpose we attempt to transport the geometry of the plane into V in an appropriate fashion

Lemma 1.22. *Let V be an arbitrary real vector space with an inner product $(\cdot|\cdot)$. Let f, g be two linearly independent vectors in V. We give $\mathbb{R}^2$ the inner product given by $\big((u_1, u_2)|(v_1, v_2)\big) = u_1v_1 + u_2v_2$. Then there is a linear map $L: \mathbb{R}^2 \to V$ such that $(Lx|Ly) = (x|y)$ for x, $y \in \mathbb{R}^2$ whose range is the span of f and g.*

Proof . Now let f and g be unit vectors in V. First we consider the vector $h = g - (f \mid g)f$ in V.

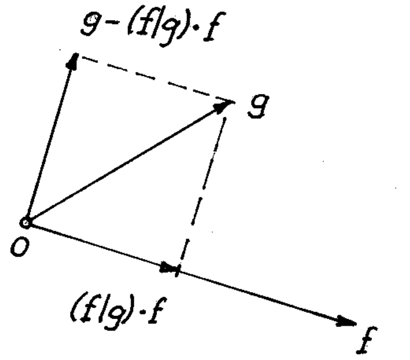


Figure 1.4

If this vector is 0, then $g = (f \mid g)f$. Since f and g are unit vectors, this means $(f \mid g) = \pm 1$. In both of these cases, $(f \mid g)$ is indeed the cosine of the angle between f and g, which is 0 for $+1$ and π for -1. Now assume the case $h \neq 0$ and set $e = \|h\|^{-1}{\cdot}h$. Then e is a unit vector such that

$$(f \mid e) = \big(f \mid \|h\|^{-1}{\cdot}(g - (f \mid g){\cdot}f)\big) = \|h\|^{-1}\big((f \mid g) - (f \mid g)(f \mid f)\big) = 0.$$

This motivates us to choose the wanted map L so that $L(1,0) = f$ and $L(0,1) = e$. In other words we define

$$L(x, y) = x \cdot f + y \cdot e.$$

Then L satisfies the following conditions
(i) $L(a + b) = L(a) + L(b)$ for all $a, b \in \mathbb{R}^2$, and
(ii) $L(r \cdot a) = r \cdot L(a)$ for all $a \in \mathbb{R}^2$, $r \in \mathbb{R}$.

If we take $a = (x, y)$ and $b = (x', y')$ in $\mathbb{R}^2$, then on the one hand we have $(a \mid b) = xx' + yy'$. On the other hand we observe $(f \mid f) = (e \mid e) = 1$ and $(e \mid f) = 0$, from which we conclude $(L(a) \mid L(b)) = (x \cdot f + y \cdot e \mid x' \cdot f + y' \cdot e) = xx' + yy'$. Thus we have
(iii) $(L(a) \mid L(b)) = (a \mid b)$ for all $a, b \in \mathbb{R}^2$.

In particular, (3) implies $\|L(a)\| = \|a\|$. This is why L is called an *isometry* from $\mathbb{R}^2$ into V. If $L(a) = L(b)$, then (1) and (2) imply $L(b - a) = 0$ and thus $\|b - a\| = \|L(b - a)\| = 0$, that is $b - a = 0$. An isometry therefore is always injective. The conditions (1) and (2) of *linearity* of L ensure that the vector space properties of $\mathbb{R}^2$ are faithfully translated into vector space properties of V. The vector subspace of all $x \cdot f + y \cdot e$, $(x, y) \in \mathbb{R}^2$ is in every respect a faithful image of $\mathbb{R}^2$. One says that it is *isometrically isomorphic* to $\mathbb{R}^2$ via the isometry L. Measuring of lengths and angles yield the exact same results in this vector subspace of V as in $\mathbb{R}^2$. In particular, $(f \mid g)$ is the cosine of the angle between the unit vectors f and g. Also, by our definition of the angle between f and g is a number $t \in [0, \pi]$ Thus we have the following result which sharpens the Cauchy-Schwarz inequality (in the real case).

Theorem 1.23. *If f and g are two vectors in a vector space with an inner product $(\cdot \mid \cdot)$, then*

$$(f \mid g) = \|f\| \cdot \|g\| \cos t, \tag{6}$$

where $t \in [0, \pi]$ is the nonoriented angle between f and g. □

We notice that the angle is undefined if one of the two vectors f and g vanishes. But then (6) remains valid in the sense that it both sides of (6) are zero. If both f and g are nonzero, We may interpret $\|g\| \cos t$ as the length of the projection of g on the straight line spanned by f, and likewise $\|f\| \cos t$ as the length of the projection of f on the straight line spanned by g.

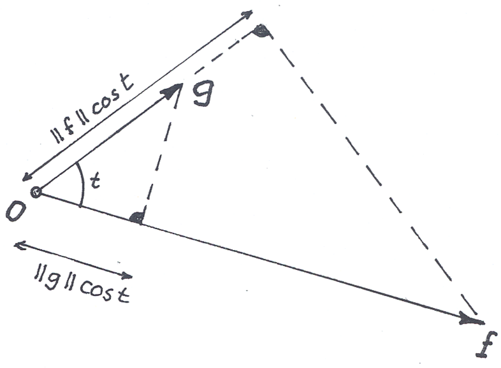


Figure 1.5

In particular we say that two vectors are *perpendicular* or *orthogonal* to each other if $(f \mid g) = 0$ holds.

Everything that was said for vector spaces with an inner product holds, in particular for $\mathbb{R}^n$.

Finite dimensional normed vector spaces

If V is a finite dimensional normed vector space, then for every basis $\{e_1, \ldots, e_n\}$ we find an *isomorphism of vector spaces* $\varphi\colon \mathbb{K}^n \to V$ by setting $\varphi(x_1, \ldots, x_n) = x_1{\cdot}e_1 + \cdots + x_n{\cdot}e_n$. If we set $\|(x_1, \ldots, x_n)\| \stackrel{\text{def}}{=} \|\varphi(x_1, \ldots, x_n)\|_{\mathbb{V}}$, then we define on $\mathbb{K}^n$ a norm $\|\cdot\|$ such that $\|\varphi^{-1}(x)\| = \|x\|_{\mathbb{V}}$. Therefore φ preserves norms; it is an *isometry*. In considering norms on finite dimensional vector spaces we may restrict our attention to the n-tuple spaces $\mathbb{K}^n$.

Lemma 1.24. *Let* $\|\cdot\|$ *be* any *norm on* $\mathbb{K}^n$, $n \geq 1$. *Then the function* $x \mapsto \|x\| : (\mathbb{K}^n, \|\cdot\|_\infty) \to \mathbb{R}$ *is continuous with respect to the metric* d *on* $\mathbb{K}^n$ *given by* $d(u, v) = \|v - u\|_\infty$ *with the max-norm.*

Proof . We set $e_1 = (1, 0, \ldots, 0)$, $e_2 = (0, 1, 0, \ldots, 0)$ etc. If we set $C = \|e_1\| + \cdots \|e_n\|$, then $C \geq \|e_1\| > 0$, and

$$\begin{aligned}
\big|\|x\| - \|y\|\big| &\leq \|x - y\| \\
= \|(x_1, \ldots, x_n) - (y_1, \ldots, y_n)\| &= \|(x_1 - y_1, \ldots, x_n - y_n)\| \\
= \|(x_1 - y_1){\cdot}e_1 + \cdots + (x_n - y_n){\cdot}e_n\| &\leq |x_1 - y_1|{\cdot}\|e_1\| + \cdots + |x_n - y_n|{\cdot}\|e_n\| \\
&\leq \|x - y\|_\infty{\cdot}(\|e_1\| + \cdots \|e_n\|) \\
&= C\|x - y\|_\infty.
\end{aligned}$$

Now for any $\varepsilon > 0$, the relation $\|x - y\|_\infty < \frac{\varepsilon}{C}$ implies $\big|\|x\| - \|y\|\big| < \varepsilon$. □

Let $\mathbb{S}$ denote the surface $\{x \in K^n : \|x\|_2 = 1\}$ of the euclidean unit ball, called the *unit sphere.* If $\mathbb{K} = \mathbb{R}$ then $\mathbb{S}$ is an "$n - 1$-dimensional surface" in the n-dimensional space $\mathbb{R}^n$; therefore one also writes $\mathbb{S} = \mathbb{S}^{n-1}$ and calls $\mathbb{S}$ the

$n-1$-*sphere*. The boundary of the unit ball in $\mathbb{C}^n$ is

$$\{(x_1 + y_1 i, \ldots, x_n + y_n i) : x_1^2 + y_1^2 + \cdots + x_n^2 + y_n^2 = 1\},$$

and if we identify $\mathbb{R}^{2n}$ with $\mathbb{C}^n$ via $(x_1, y_1, \ldots, x_n, y_n) \mapsto (x_1 + y_1 i, \ldots, x_n + y_n i)$, then this is the $2n-1$–sphere $\mathbb{S}^{2n-1}$.)

If D denotes the unit disk or unit ball $\{z \in \mathbb{K} : |z| \leq 1\}$ in $\mathbb{K}$, then the $\|\cdot\|_\infty$-unit ball B_∞ in $\mathbb{K}^n$ is exactly $\underbrace{D \times \cdots \times D}_{n\,\text{times}} \subseteq \mathbb{K}^n$. This set contains the $n-1$–sphere $\mathbb{S}^{n-1} = \{(x_1, \ldots, x_n) : |x_1|^2 + \cdots + |x_n|^2 = 1\}$. By 1.23, the function $\nu \colon (\mathbb{K}^n, \|\cdot\|) \to \mathbb{R}$, $\nu(x) = \|x\|_2$ is continuous and $\mathbb{S} = \nu^{-1}(\{1\})$, the subset $\mathbb{S}$ is closed in B_∞, since for continuous functions, inverse images of closed sets are closed.

We could have proved the following lemma long ago, but now it is needed. Indeed the essence was proved in 3.49(iii).

Lemma 1.25. (i) *If X_j, $j = 1, \ldots, n$ are compact metric spaces then the product space $X_1 \times \cdots \times X_n$ is compact with respect to the metric given by*

$$D((x_1, \ldots, x_n), (y_1, \ldots, y_n)) = \max\{d_1(x_1, y_1), \ldots, d_n(x_n, y_n)\}.$$

(ii) *For each $r \geq 0$ the ball $B_r(0) = \{x \in \mathbb{K}^r : \|x\|_\infty \leq r\} = \{(x_1, \ldots, x_n) : |x_1|, \ldots, |x_n| \leq r\}$ is compact.*

Proof. (i) For $n = 1$ the Lemma is trivial. For $n = 2$ it was proved in 3.49(iii). It is now an easy exercise to apply induction to prove (i) in full generality.

(ii) Let $D_r \subseteq \mathbb{K}$ denote the closed disc of radius r. Then D_r is compact by the Bolzano-Weierstraß-Theorem for $\mathbb{K}$. (See 3.46 and 3.50.) Now we notice $B_r(0) = D_r^n \subseteq \mathbb{K}^n$ and apply (i) above. □

Exercise E1.9. Write down the details of the proof by induction of Lemma 1.25(i), using 3.49(iii).

By 1.25(ii) the $\|\cdot\|_\infty$-unit ball $B_1(0)$ in $\mathbb{K}^n$ is compact and $\mathbb{S} \subseteq B_1(0)$. By 3.49(i) a closed subspace of a compact metric space is compact. Therefore,
(comp) *the $n-1$–sphere $\mathbb{S}^{n-1}$ in $\mathbb{R}^n$ is compact.*
In Proposition 1.12 we observed that each of the norms $\|\cdot\|_2$ and $\|\cdot\|_\infty$ is dominated by a multiple of the other. This is an important fact and we now show that this remains true for any pair of norms on $\mathbb{K}^n$.

Definition 1.26. Two norms $\|\cdot\|$ and $\|\cdot\|_*$ on any $\mathbb{K}$-vector space V are said to be *equivalent norms*, if there are positive numbers c and C such that

$$(\forall x \in V) \quad c\|x\|_* \leq \|x\| \leq C\|x\|_*. \tag{7}$$ □

Now we prove a surprising and important result.

Theorem 1.27. *Two arbitrary norms on a finite dimensional vector space over $\mathbb{R}$ or $\mathbb{C}$ are equivalent.*

Proof. First we consider an arbitrary norm $\|\cdot\|$ on $\mathbb{K}^n$ and show that it is equivalent to the euclidean norm $\|\cdot\|_2$. The function $s \mapsto \|s\|: \mathbb{S} \to \mathbb{R}$ is continuous by 1.24 when the $\|\cdot\|_\infty$-metric is considered on $\mathbb{S}$ relative to which $\mathbb{S}$ is compact as we observed in (comp) above. Thus the image $\|\mathbb{S}\|$ is a compact set of nonnegative real numbers. By the Theorem of the Maximum 3.52 the numbers $a = \min \|\mathbb{S}\|$ and $A = \max \|\mathbb{S}\|$ are well defined. Thus there is a $v \in \mathbb{S}$ with $a = \|v\|$. If we had $a = 0$ then $v = 0$ by 1.3(i), and then $0 = \|v\|_2 = 1$, and obvious contradiction. Therefore $0 < a$. Now let $0 \neq x \in \mathbb{K}^n$. Then $s \overset{\text{def}}{=} \|x\|_2^{-1} \cdot x \in \mathbb{S}$. Hence $a \leq \|s\| = \|x\|_2^{-1} \cdot \|x\| \leq A$. This is equivalent to $a\|x\|_2 \leq \|x\| \leq A\|x\|_2$. These inequalities also hold for $x = 0$.

Next, if a second norm $\|\cdot\|_*$ is given on $\mathbb{K}^n$, then by the first part of the proof there are positive numbers $b < B$ such that $b\|x\|_2 \leq \|x\|_* \leq B\|x\|_2$ for all $x \in \mathbb{K}^n$. But then $\frac{a}{B}\|x\|_* \leq \|x\| \leq \frac{A}{b}\|x\|_*$ for all $x \in \mathbb{K}^n$. We set $c = \frac{a}{B}$ and $C = \frac{A}{b}$; then the theorem is proved for $\mathbb{K}^n$. Since every n-dimensional $\mathbb{K}$-vector space is isomorphic to $\mathbb{K}^n$ the theorem holds for all n-dimensional $\mathbb{K}$ vector spaces. □

The most important conclusion is that, *in a finite dimensional* $\mathbb{K}$*-vector space, the properties of a subset to be* open, closed, bounded, compact, or connected *do not depend on the choice of a particular norm.* Furthermore *the properties of a sequence to be* a Cauchy sequence or to converge *also do not depend on the choice of a norm.*

Exercise E1.10. Prove the following proposition.
Let V be a finite dimensional vector space over $\mathbb{K}$ and let A be a subset. Moreover let $(x_n)_{n\in\mathbb{N}}$ be a sequence in V and x an element of V. Each norm $\|\cdot\|$ on V turns V into a metric space via $d(x, y) = \|x - y\|$. The following statements therefore are all meaningful:
(a) *A is open [respectively, closed] in V.*
(b) *A is bounded in V.*
(c) *$x = \lim_{n\to\infty} x_n$.*
(d) *x is an accumulation point of A.*
(e) *x is a cluster point of $(x_n)_{n\in\mathbb{N}}$.*
(f) *A ist compact.*
(g) *A ist connected.*
(h) *A is pathwise connected.*
The truth of each of these assertions is independent of the choice of $\|\cdot\|$. □

This permits us to prove a generalisation of the theorem of Bolzano and Weierstrass which we proved for one-dimensional vector spaces in Chapter 3.

Theorem 1.28. (Bolzano-Weierstrass) *For a subset X of a finite dimensional real or complex vector space the following statements are equivalent*
(i) *X is compact.*
(ii) *X is closed and bounded.*

Proof. (i)⇒(ii): Every compact subset of a metric space is closed and bounded by 3.45.

(ii)⇒(i): It is no loss of generality to consider $V = \mathbb{K}^n$. By Theorem 1.27 and Exercise E1.8 above, (ii)⇒(i) is proved for any norm if (ii)⇒(i) is proved for the max-norm given by $\|(x_1, \ldots, x_n)\| = \max\{|x_1|, \ldots, |x_n|\}$. We establish the implication for this one.

By (ii) there is an $r \geq 0$ such that $\|x\|_\infty \leq r$ for all $x \in X$. Thus X is a closed subset of the $\|\cdot\|_\infty$-ball $B_r(0)$ which is compact by 1.25(ii). Since X is assumed to be closed in (ii), and since by 3.49(i) a closed subspace of a compact space is compact, X is compact as asserted. □

We recall from a warning following Theorem 3.46 that that compactness (and closedness) are topological properties while boundedness in a metric space is not a topological property. This applies again in the situation of the more general Bolzano-Weierstraß Theorem 1.28. Notice, however, that a set which is *norm*-bounded with respect to one norm in $\mathbb{K}^n$ is norm bounded with respect to all norms. Metric boundedness does not have this agreeable property of norm boundedness as we saw after 3.46.

Corollary 1.29. *All closed balls in a finite dimensional real or complex vector space are compact.* □

Exercise E1.11. Consider the Banach space $B(\mathbb{N}, \mathbb{K})$ of all real bounded sequences. Show that the unit ball $B_1(0) = \{(x_n)_{n\in\mathbb{N}} : |x_n| \leq 1\}$ is *not* compact with respect to the sup-norm.
[Hint. Consider the sequence $(e_n)_n \in \mathbb{N}$ in $B(\mathbb{N})$ given by $e_n = (\delta_{nm})_{m\in\mathbb{N}}$ with

$$\delta_{nm} = \begin{cases} 1 & \text{if } m = n \\ 0 & \text{otherwise.} \end{cases}$$

(This function $(n, m) \mapsto \delta_{nm}$ is sometimes called the *Kronecker delta*.) Observe that $p \neq q$ implies $\|e_p - e_q\|_\infty = 1$. Suppose that x were a cluster point of $(e_p)_{p\in\mathbb{N}}$. Then the open ball $U_{1/2}(x)$ of radius $\frac{1}{2}$ would contain infinitely many of the different elements e_p (see 3.41). Let e_p and e_q be a pair of these with $p \neq q$. Then $1 = \|e_p - e_q\| \leq \|e_p - x\| + \|x - e_q\| < \frac{1}{2} + \frac{1}{2} = 1$, a contradiction.]

One can show that the only Banach spaces in which the closed balls are compact are the finite dimensional ones.

Linear maps

Recall from linear algebra:

Definition 1.30. A function or map *linear* if it satisfies the following conditions
(i) $(\forall u, v \in V)\ L(u + v) = L(u) + L(v)$,
(ii) $(\forall v \in V, r \in \mathbb{K})\ L(r\cdot v) = r\cdot L(v)$. □

The set $\mathrm{Hom}(V,W) \subseteq W^V$ of linear maps $L: V \to W$ or vector space *homomorphisms* is a vector subspace of W^V. If V and W are finite dimensional we choose bases $e_1,\ldots,e_n$ of V and $f_1,\ldots,f_m$ of W, then there are, firstly, m linear forms ω_j, $j = 1,\ldots,m$ such that $y = \sum_{j=1}^m \omega_j(y)\cdot f_j$. The numbers $a_{jk} = \omega_j\big(L(e_k)\big)$ form the *matrix* $(a_{jk})_{\substack{j=1,\ldots,m\\k=1,\ldots,n}}$ of the linear map L with respect to the bases e_k and f_j. We note

$$\begin{aligned} L(x) &= \sum_{j=1}^m \omega_j\big(L(x)\big)\cdot f_j, \quad \text{and}\\ L(x) &= L(\sum_{k=1}^n x_k\cdot e_k) = \sum_{k=1}^n x_k\cdot L(e_k) = \sum_{k=1}^n\sum_{j=1}^m x_k\omega_j\big(L(e_k)\big)\cdot f_j\\ &= \sum_{\substack{j=1,\ldots,m\\k=1,\ldots,n}} a_{jk}x_k\cdot f_j, \quad \text{also}\\ \omega_j\big(L(x)\big) &= \sum_{k=1}^n a_{jk}x_k. \end{aligned}$$

The choice of a basis of V is tantamount with the choice of an isomorphism of $\mathbb{K}$-vector spaces $V \to \mathbb{K}^n$. If, instead of $V \cong \mathbb{K}^n$, we even have equality $V = \mathbb{K}^n$, then we have the so-called *standard basis* $e_k = (0,\ldots,1,\ldots,0)$ with 1 in the k-th position. Then in case of $V = \mathbb{K}^n$ and $W = \mathbb{K}^m$ there is a natural (linear) bijection between the vector space $\mathrm{Hom}(V,W)$ of all linear maps $L: V \to W$ and the vector space $M_{mn}(\mathbb{K})$ of all $m\times n$-matrices $(a_{jk})_{\substack{j=1,\ldots,m\\k=1,\ldots,n}}$. Since the latter is clearly an mn-dimensional vector space we have

$$\dim \mathrm{Hom}(V,W) = (\dim V)(\dim W). \tag{8}$$

Linear forms

A particularly simple type are the linear maps $f: \mathbb{K}^n \to \mathbb{K}$. By the preceding discussion they are characterized by a row matrix $(a_1,\ldots,a_n)$ and $f(x_1,\ldots,x_n) = a_1x_1+\cdots+a_nx_n \overset{\text{def}}{=} \langle a,x\rangle$ where $x = (x_1,\ldots,x_n)$, and $a = (a_1,\ldots,a_n)$. Therefore there is a bijective map $\mathbb{K}^n \to \mathrm{Hom}(\mathbb{K}^n,\mathbb{K})$ which associates with a vector $a = (a_1,\ldots,a_n) \in \mathbb{K}^n$ the linear form $x \mapsto \langle a,x\rangle : \mathbb{K}^n \to \mathbb{K}$.

We notice that in $\mathbb{R}^2$ the set $\{x = (x_1,x_2) : \langle a,x\rangle = (x \mid a) = 0\}$ is exactly the set of all vectors which are perpendicular to a by 1.23.

If $\mathbb{K} = \mathbb{C}$ then $\langle a,x\rangle = a_1x_1+\cdots+a_nx_n = (x|\overline{a})$ according to the definition of the inner product in the complex case by (4) preceding 1.15. Thus in $\mathbb{K}^2$, then set $\{x : \langle a,x\rangle = 0\}$ is the set of vectors which are perpendicular to $\overline{a}$. We shall from now on deal with the real case. But we point out, that in principle, an extension to the case of $\mathbb{C}^n$ is easily possible if this remark is observed. In the real case we have $\langle a,x\rangle = (x \mid a) = (a \mid x)$.

If $a \neq 0$, then this set is a straight line through the origin. In $\mathbb{R}^n$ the corresponding set of all x satisfying $\langle a,x\rangle = (x \mid a) = 0$ is a plane. Following this pattern we say that for $a \neq 0$ the set $\{x \in \mathbb{R}^n : \langle a,x\rangle = (a \mid x)\}$ is a *hyperplane* through the origin. If x_0 is an arbitrary point in $\mathbb{R}^n$ and $a_0 \stackrel{\text{def}}{=} (a \mid x_0)$, then the vector $x - x_0$ is perpendicular to a iff $(a \mid x - x_0) = 0$, that is, if $(a \mid x) = a_0$. This is the case exactly when x is contained in that hyperplane which contains the point x_0 and is perpendicular to a. Notably, the vector $\check{a} = \frac{a_0}{\|a\|^2}{\cdot}a$ satisfies $(a \mid \check{a}) = a_0$, where $\|{\cdot}\|$ is the euclidean norm. We note $\|\check{a}\| = \frac{|a_0|}{\|a\|}$. Thus we have:

Remark. *For a nonzero vector a in an arbitrary real vector space V the set $\{x \in V\colon (x \mid a) = (a \mid x) = a_0\}$ is a hyperplane which is perpendicular to a and has the distance $\|a\|^{-1}{\cdot}\|a_0\|$ from the origin.* □

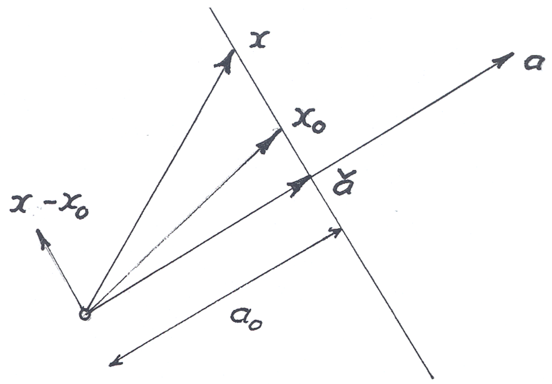


Figure 1.6

Continuity of linear maps

Theorem 1.31. *A linear map between finite dimensional vector spaces is continuous.*

Proof. We consider a linear map $L\colon \mathbb{K}^n \to \mathbb{K}^m$; we know that the special form of the domain and codomain is no restriction of generality. Now L is described by a matrix $(a_{jk})_{\substack{j=1,\dots,m\\k=1,\dots,n}}$ through the formula

$$(9) \qquad L(x) = (\sum_{k=1}^{n} a_{1k}x_k, \dots, \sum_{k=1}^{n} a_{mk}x_k), \quad x = (x_1, \dots, x_n).$$

After Theorem 1.27 it is immaterial which norms we consider on $\mathbb{K}^n$ and $\mathbb{K}^m$. We choose the max-norms and compute

$$\begin{aligned}
\|Lx - Ly\|_\infty &= \left\| \left(\sum_{k=1}^n a_{1k}(x_k - y_k), \dots, \sum_{k=1}^n a_{mk}(x_k - y_k) \right) \right\|_\infty \\
&= \max \left\{ \left| \sum_{k=1}^n a_{1k}(x_k - y_k) \right|, \dots, \left| \sum_{k=1}^n a_{mk}(x_k - y_k) \right| \right\} \\
&\le \max \left\{ \sum_{k=1}^n |a_{1k}| \cdot |x_k - y_k|, \dots, \sum_{k=1}^n |a_{mk}||x_k - y_k| \right\} \\
&\le n \cdot \max\{|a_{jk}| : j = 1, \dots, m;\ k = 1, \dots, n\} \cdot \max\{|x_k - y_k| : k = 1, \dots, n\} \\
&\le C \cdot \|x - y\|_\infty
\end{aligned}$$

where $C = n \cdot \max\{|a_{jk}| : j = 1, \dots, m;\ k = 1, \dots, n\}$. Of course, this proves the desired continuity. □

The automatic continuity of linear maps between finite dimensional vector spaces breaks down if the domain is infinite dimensional. For instance if we consider on $C([0,1], \mathbb{R})$ the norm given by $\|f\|_2 = \sqrt{\int f^2}$, then the sequence f_n defined by $f_n(x) = x^n$ converges to the zero function $f = 0$ for this norm. (Proof?) The function $L\colon C([0,1]) \to \mathbb{R}$, $L(\varphi) = \varphi(1)$, is a linear form. But $L(f_n) = 1^n = 1$ but $L(f) = f(1) = 0$. Thus L fails to be continuous even though the range is one-dimensional.

The operator norm

Consider two finite dimensional normed vector spaces V and W. We claim that the finite dimensional vector space $\operatorname{Hom}(V, W)$ (see (8)) can be given a norm which depends in a natural way on the norms of V and W. We first have to assign to a linear map $L\colon V \to W$ a number $\|L\| \in \mathbb{R}$. The norm $\|\cdot\|_V$ in V gives us a unit ball $B = \{x \in V : \|x\|_V \le 1\}$. Since V is finite dimensional, B is compact by 1.29. By Theorem 1.31, the function $L\colon V \to W$ is continuous. The function $y \mapsto \|y\|_W : W \to \mathbb{R}$ is continuous by 1.24 (cf. also 1.27). Then we have a continuous function $x \mapsto \|L(x)\|_W\colon B \to \mathbb{R}$ which is bounded on the compact space by Theorem 3.52 and even attains a maximum. We set

$$(10_1) \qquad \|L\| = \max\{\|L(x)\| : x \in V,\ \|x\|_V \le 1\}.$$

We remark

$$(10_2) \qquad \|L\| = \sup\{\|L(x)\| : x \in V,\ \|x\|_V \le 1\}.$$

If $L\colon V \to W$ is a linear map between not necessarily finite dimensional normed vector spaces, then the continuity of L is equivalent with the boundedness of the set $\{L(x) : \|x\| \le 1\}$. In this case, $(10)_2$ is still a viable definition while $(10)_1$ does not work in general.

Theorem 1.32. *Assume that V and W are finite dimensional normed $\mathbb{K}$-vector spaces. The function $L \mapsto \|L\|\colon \mathrm{Hom}(V,W) \to \mathbb{R}$ defined in* (10_1)*, equivalently,* (10_2) *above is a norm on the vector space* $\mathrm{Hom}(V,W)$ *satisfying*

$$(11) \qquad (\forall L \in \mathrm{Hom}(V,W),\ v \in V)\ \|Lv\|_W \le \|L\|{\cdot}\|v\|_V.$$

If U, V, and W are finite dimensional vector spaces, then

$$(12) \qquad \big(\forall T \in \mathrm{Hom}(U,V),\ S \in \mathrm{Hom}(V,W)\big) \quad \|ST\| \le \|S\|{\cdot}\|T\|.$$

Proof. Exercise. □

Definition 1.33. The norm on $\mathrm{Hom}(V,W)$ according to Theorem 1.32 is called *operator norm.* □

Exercise E1.13. (i) Prove 1.32.

(ii) Prove the following proposition:

Proposition. *Let $V = \mathbb{K}^n$ and $W = \mathbb{K}^m$ and consider the max-norms on both of these. Let $L\colon V \to W$ have the matrix $A = (a_{jk})_{\substack{j=1,\dots,m\\ k=1,\dots,n}}$. Then the operator norm of L is computed as*

$$\|L\| = \max_{j=1,\dots,m} \sum_{k=1}^{n} |a_{jk}|.$$

If $\mathbb{K} = \mathbb{R}$ and $m = n = 2$ and if L has the matrix $\begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$, then the operator norm (for the max-norm on $\mathbb{R}^2$) is 1.

If one has enough linear algebra background to know the concept of an eigenvalue, then one notices that the characteristic polynomial of L is λ^2 and thus the only eigen-value of L is 0.

[Hint. (i) Verify all properties of a norm 1.3(i),(ii),(iii). For a proof of (11) take an arbitrary $v \in V$. If $\|v\|_V \le 1$, then (11) is immediate from (10). Now let $v \ne 0$; then $w \stackrel{\text{def}}{=} \|v\|_V^{-1}{\cdot}v$ has norm 1; thus $\|Lw\| \le \|L\|$ by (10), and then (11) follows upon multiplying with $\|v\|_V$. For a proof of (12) assume $\|u\|_U \le 1$; then $\|STu\|_W \le \|S\|{\cdot}\|Tu\|_V \le \|S\|{\cdot}\|T\|$ by (11) and (10); now form the sup over all $\|u\|_U \le 1$ on the left side.

(ii) Write $v = \begin{pmatrix} v_1 \\ \vdots \\ v_n \end{pmatrix}$ and $Lv = \begin{pmatrix} w_1 \\ \vdots \\ w_m \end{pmatrix}$. Then $Lv = Av$, whence $w_j = a_{j1}v_1 + \cdots + a_{jn}v_n$, $j = 1, \dots, m$. Use this to show $|w_j| \le \|v\|_\infty \sum_{k=1}^n |a_{jk}|$. Then $\|Lv\|_\infty \le \max_{j=1,\dots,m} \sum_{k=1}^n |a_{jk}|$. Conversely, assume that $\sum_{k=1}^n |a_{j_0k}|$ is the maximal one among the sums on the right hand side. Choose v by setting $v_k = \operatorname{sgn} a_{j_0k}$. Then $\|v\|_\infty = 1$ and $\|Lv\| = \sum_{k=1}^n a_{j_0k}v_k = \sum_{k=1}^n |a_{j_0k}|$.]

The explicit calculation of an operator norm in terms of matrix coefficients like in E1.10(ii) is a rare event; one does not expect that this is easily possible for other norms on V and W.

Given enough information on Hilbert spaces and their endomorphisms, one can show that the operator norm of a selfadjoint endomorphism is the maximum of the absolute values of the eigenvalues.

The vector space $A \stackrel{\text{def}}{=} \text{Hom}(V,W)$ has dimension $(\dim V)(\dim W)$ by (8) above and is therefore finite dimensional. Hence it is a Banach space by 1.11 and 1.27. The vector space $\text{Hom}(V,V)$ is closed under multiplication (composition of linear self-maps of V) and is an excellent example for what one calls a *Banach algebra.*

Definition 1.34. A *Banach algebra* A (over $\mathbb{K}$) is a Banach space (over $\mathbb{K}$) with an associative multiplication $(x,y) \mapsto xy : A \times A \to A$ such that $s{\cdot}xy = (s{\cdot}x)y = x(s{\cdot}y)$ for all $s \in \mathbb{K}$ and $x, y \in A$, and that $\|xy\| \le \|x\|{\cdot}\|y\|$ for all x and y in A. We say that A is *unital* if it has a multiplicative identity element. □

Apart from the example of $\text{Hom}(V,V)$, which is isomorphic to the space $M_n(\mathbb{K})$ of $n \times n$ matrices over $\mathbb{K}$ with the operator norm, the fields $\mathbb{R}$ and $\mathbb{C}$, equipped with their absolute values, are Banach algebras over $\mathbb{R}$, respectively $\mathbb{C}$.

Since convergence of sequences is defined in all metric spaces (see Definition 2.13) and every Banach space is in particular a metric space with $d(x,y) = \|x-y\|$, the definition of convergence of an infinite series of numbers can be generalized from $\mathbb{R}$ and $\mathbb{C}$ without any problems to infinite series $\sum_{n=0}^{\infty} x_n$, to any Banach space E, where $x_n \in E$, and it should be clear what convergence and absolute convergence should be. Indeed we shall say that $\sum_{n=0}^{\infty} x_n$ has a sum x in E if $x = \lim_{n\to\infty}(x_0 + \cdots + x_n)$, and we shall call $\sum_{n=0}^{\infty} x_n$ absolutely convergent if the infinite series of nonnegative numbers $\sum_{n=0}^{\infty} \|x_n\|$ converges. In E all Cauchy sequences converge by definition; the proof of 2.50 applies with $\|\cdot\|$ in place of $|\cdot|$ and shows that absolutely convergent series converge.

Now let us consider a unital Banach algebra A (such as $\text{Hom}(V,V)$) and $a \in A$. Then induction shows $\|a\|^n \le \|a\|^n$. Now let $\sum_{n=0}^{\infty} a_n z^n$ be a power series in $\mathbb{C}$ with radius of convergence $\rho > 0$ (see Definitions 2.57 and 2.58 and Theorem 2.59). If $\|a\| < \rho$, then $\sum_{n=0}^{\infty} |a_n|{\cdot}\|a\|^n$ converges. Since $\|a_n{\cdot}a^n\| \le |a_n|{\cdot}\|a\|^n$ the power series $\sum_{n=0}^{\infty} a_n{\cdot}a^n$ converges absolutely. In particular, it converges.

This applies, in particular to the power series $\sum_{n=0}^{\infty} \frac{1}{n!} z^n$. Thus for all elements $u \in A$ the infinite series $1 + u + \frac{1}{2!}u^2 + \frac{1}{3!}u^3 + \cdots$ converges absolutely and thus defines a function

$$\exp\colon A \to A, \quad \exp u = \sum_{n=0}^{\infty} \frac{1}{n!} u^n.$$

The proof of Theorem 2.66 of the Convergence of the Convolution (Cauchy product) of two abolutely convergent infinite series generalizes at once to absolutely convergent infinite series in a Banach algebra; all we need to do is again to replace $|\cdot|$ by $\|\cdot\|$. This allows us to raise the question whether for the exponential function $\exp\colon A \to A$ on a Banach algebra A the functional equation of Theorem 2.68 remains valid. An inspection of the proof of 2.68 leads us up to the binomial formula (47). Unfortunately in a Banach algebra, in general multiplication is not commutative. In $M_2(\mathbb{R})$ let $x = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$ and $y = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}$ we have $xy = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} \neq \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} = yx$. Then $(x+y)^2 = x^2 + xy + yx + y^2 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \neq \begin{pmatrix} 2 & 0 \\ 0 & 0 \end{pmatrix} = x^2 + 2xy + y^2$. However, if we *assume* that x and y are two elements in A which commute, i.e. satisfy $xy = yx$, then the binomial formula holds (cf. 2.11), and then the proof of Theorem 2.68 shows that $\exp(x+y) = (\exp x)(\exp y)$ holds. In particular, if $s, t \in \mathbb{K}$ and $a \in$, the elements $x = s{\cdot}a$ and $t{\cdot}a$ commute. Therefore $\exp(s+t){\cdot}a = (\exp s{\cdot}a)(\exp t{\cdot}y)$. Let us record this in the following

Theorem 1.35. (The Exponential Function on a Banach Algebra). *Let A be a unital Banach algebra over $\mathbb{K} \in \{\mathbb{R}, \mathbb{C}\}$. Then the infinite series $1 + x + \frac{1}{2!}x^2 + \frac{1}{3!}x^3 + \cdots$ converges absolutely for all $x \in A$ and its sum defines a function $\exp\colon A \to A$. If x and y are commuting elements of A, then $\exp(x+y) = (\exp x)(\exp y)$. In particular,*

$$(*) \qquad (\forall a \in A,\ s, t \in \mathbb{K})\ \exp(s+t){\cdot}a = (\exp s{\cdot}a)(\exp t{\cdot}a). \qquad \square$$

Upon taking $s = 1$ and $t = -1$ in $(*)$, and observing $\exp 0 = 1$ we note that $\exp a$ is invertible for all $a \in A$ $(\exp a)^{-1} = \exp -a$.

Everything we said on the exponential function in a unital Banach algebra applies, in particular, to the Banach Algebra $\mathrm{Hom}(V, V)$ of all self-maps of a finite dimensional normed $\mathbb{K}$-vector space V, where $\mathrm{Hom}(V, V)$ is equipped with the operator norm. So, in particular, everything applies to the algebra $M_n(\mathbb{K})$ of $n \times n$-matrices over the field $\mathbb{K}$, equipped with the operator norm when $M_n(\mathbb{K})$ and $\mathrm{Hom}(\mathbb{K}^n, \mathbb{K}^n)$ are identified.

Affine maps

We call that a function $A\colon V \to W$ between vector spaces is called *affine*, if there is a linear function $L\colon V \to W$ and a vector $v \in V$ such that $A(x) = L(x) + v$ for all $x \in V$. □

We note that $v = A(0)$; thus a function A is affine iff $A - A(0)$ is linear. An affine function fixes the origin if and only if it is linear. Affine functions map straight lines into straight lines. (Indeed this property characterizes affine maps.)

Postscript

Once we have a norm on a vector space V, it becomes at once a metric space via a metric d defined by $d(x, y) = \|y - x\|$ (see 1.7). And as soon as we have a metric space before us, everything that was said in Analysis I becomes instantly available for V and all of its subsets to which we assign the metric induced from that of V. Among the very first examples of normed vector spaces we have observed spaces of bounded functions on a set X (Example 1.6), in which we find many spaces of functions treated in Analysis I when $X = [a, b]$. We note at once that most of these spaces are infinite dimensional (unless X is finite, in which case we retrieve the familiar vector spaces $\mathbb{R}^n$ or $\mathbb{C}^n$). Therefore, strong emphasis is placed on possibly infinite dimensional normed vector spaces.

The appropriate concept is that of *completeness* of a metric space which was introduced now in 1.8(i) and which specializes at once to normed vector spaces giving us readily the idea of a *Banach space* (1.8(ii)). In this environment it is now very natural to introduce the concept of *uniform convergence* of a sequence of functions $f_n: X \to \mathbb{K}$ defined on a metrix space X (such as e.g. $X = [a, b]$) because it is nothing else but convergence in some metric space which we practiced in Analysis I. There arises for students the problem of distinguishing between uniform convergence and *pointwise convergence* (cf. exercise E1.5 and the paragraphs preceding it); we address this issue although for the topics discussed in this book, pointwise convergence plays a relatively subordinate role. We recall at this point that in Chapter 5 of Analysis I we dealt with integration without invoking the concept of uniform convergence of functions. Now that we understand this idea, we quickly prove that with respect to the sup-norm $\|\cdot\|_\infty$, the vector space $I[a, b]$ of Riemann integrable real valued functions on the compact interval $[a, b]$ is a Banach space and that the Riemann integral $\int: I(a, b) \to \mathbb{R}$ is a *continuous* linear functional (see 1.14); the proof is remarkably casual with the aid of the Riemann Criterion 5.12.

Hilbert spaces are special Banach spaces (1.15) and we know them from elementary linear algebra in the form of the euclidian space $\mathbb{R}^n$ with the inner product $(x \mid y) = x_1y_2 + \ldots + x_ny_n$. Even though we shall not deal with infinite dimensional Hilbert spaces very much in this book we wanted to make sure that at least the basic facts of finite dimensional Hilbert space theory is available; if we restrict our attention to the real field, which we shall do increasingly in this book, we could just as well speak of the geometry of euclidean space. In this context, we resume the topic of the concept of an *angle* which we initiated in Analysis I in 3.39. We emphasize now that angles between nonzero vectors pertain to Hilbert spaces (or inner product spaces) and *not* to normed spaces. The concept remains delicate in the present context as well, and we settle it in a rigorous fashion by showing, *using the given inner product*, that the real span of to linearly independent vectors in any inner product space is "isometric" to the complex plane (as two dimensional real vector space), and in it we know how to handle angles. Students tend to overlook the fact that the "familiar" geometry of two- or three-dimensional euclidian space depends crucially on the chosen inner product. We shall return to the concept of an angle shortly when we have the concept of an arc length (Exercise E1.7 following Theorem 1.11 in Chapter 7 below)

For the most part in Analysis II we have to handle finite dimensional normed spaces such as $\mathbb{R}^n$ or $\mathbb{C}^n$. They deserve a special discussion and not everything that is said here is routinely part of a Linear Algebra course. The first important basic fact is that on a finite dimensional normed space (that we know to be a Banach space), *all norms are equivalent* (1.27)—false for infinite dimensional vector spaces.

Thus all topological concepts on finite dimensional normed spaces are independent of the choice of a norm (E1.28). Another principal result characteristic for finite dimensional normed vector spaces is the characterisation of compact subsets (1.28).

In the foundations of the theory of differential calculus for functions of several variables we need the idea of linear maps (or "operators"); of course the concept itself is a central theme of any linear algebra course; but we need some aspects that may not be considered standard fare in such courses, for instance, the continuity of linear maps between finite dimensional normed spaces (1.31)—not true in the infinite dimensional situation!—and the fact that the *operator norm* (1.32, 1.33) makes the space $\mathrm{Hom}(V, W)$ of all linear maps $V \to W$ between finite dimensional normed vector spaces is itself a (finite dimensional) normed vector space and thus a Banach space. (This remains intact for infinite dimensional Banach spaces V and W, and for the space $\mathrm{Hom}(V, W)$ of all *continuous* linear operators $V \to W$.)

We have stressed the fact that for finite dimensional vector spaces V and W, any linear map $L: V \to W$ has a matrix *after one has selected a basis in eache of* V *and* W. If, however, $V = \mathbb{R}^n$ and $W = \mathbb{R}^m$, the selection has been already made for us in the form of the standard bases; therefore $\mathrm{Hom}(\mathbb{R}^n, \mathbb{R}^m)$ is indeed naturally isomorphic to the space $M_{mn}(\mathbb{K})$ of $m \times n$ matrices over $\mathbb{K}$. One could– and often does identify a linear map $\mathbb{R}^n \to \mathbb{R}^m$ with a matrix. (Cf. 1.30 and the remarks which follow.) There is always the little trickery that if we want to have $m \times n$-matrices with the alphabetic order "first m then n" we have to consider $\mathrm{Hom}(\mathbb{R}^n, \mathbb{R}^m)$, i.e. "first n then m; this is, in the final evaluation due to the fact that analysts always write functions on the *left* of the argument. A "reverse Polish notation", preferred by some algebraists, would obliterate this quirk.

The Banach spaces $\mathrm{Hom}(V, V)$ and M_{nn} with respect to the operator norm have an added feature. They have a multiplication satisfying $\|ab\| \leq \|a\| \cdot \|b\|$. They are our prime examples of a *Banach algebra.* We don't have to dwell at great length on this concept here; however, it is the right setting for an important generalisation of the exponential function. This is important enough for us to discuss in 1.35.